A NOVEL GPR-BASED PREDICTION MODEL FOR CYCLIC BACKBONE CURVES OF REINFORCED CONCRETE SHEAR WALLS

Zeynep Tuna Deger, Ph.D. and Gulsen Taskin Kaya, Ph.D.
Istanbul Technical University
{zeynep.tuna@itu.edu.tr, gulsen.taskin@itu.edu.tr}

Abstract
Backbone curves are used to characterize nonlinear responses of structural elements by simplifying the cyclic force – deformation relationships. Accurate modeling of cyclic behavior can be achieved with a reliable backbone curve model. In this paper, a novel machine learning-based model is proposed to predict the backbone curve of reinforced concrete shear (structural) walls based on key wall design properties. Reported experimental responses of a detailed test database consisting of 384 reinforced concrete shear walls under cyclic loading were utilized to predict seven critical points to define the backbone curves, namely: shear at cracking point ($V_{cr}$); shear and displacement at yielding point ($V_y$ and $\delta_y$); and peak shear force and corresponding displacement ($V_{max}$ and $\delta_{max}$); and ultimate displacement and corresponding shear ($V_u$ and $\delta_u$). The predictive models were developed based on the Gaussian Process Regression method (GPR), which adopts a non-parametric Bayesian approach. The ability of the proposed GPR-based model to make accurate and robust estimations for the backbone curves was validated based on unseen data using a hundred random sampling procedure. The prediction accuracies (i.e., ratio of predicted/actual values) are close to 1.0, whereas the coefficient of determination ($R^2$) values range between 0.90 – 0.97 for all backbone points. The proposed GPR-based backbone models are shown to reflect cyclic behavior more accurately than the traditional methods, therefore, they would serve the earthquake engineering community for better evaluation of the seismic performance of existing buildings.

Keywords: nonlinear modeling, reinforced concrete shear walls, cyclic behavior, backbone curves, machine learning, Gaussian Process Regression

1 Introduction

1.1 Background on Shear Wall Modeling

Reinforced concrete shear walls are commonly used to resist earthquake loading to provide substantial lateral strength and stiffness. As they are designed to undergo inelastic deformations under high-intensity ground shaking, it is essential to understand the nonlinear behavior and reflect the hysteretic responses in analytical models as accurately as possible. There are various well-accepted modeling approaches in the literature (e.g. summarized in NIST GCR 17-917-45 [1]), that can be classified into two main approaches, namely: finite
element and member modeling. The finite element modeling approach includes a detailed establishment of the problem by typically describing constitutive relations, and geometric nonlinearity is at the stress-strain level. Despite its advantages, such as being able to yield comprehensive results and capture local responses, the finite element modeling approach is typically cumbersome and computationally expensive for large-scale modeling and post-processing. Member modeling approaches, i.e., macro-models, are considered to be more convenient from this point of view, particularly for practitioner engineers, due to their ease-of-use and computational efficiency.

Numerous macro-models are available in the literature that can make reasonably accurate predictions for the structural response of reinforced concrete shear walls. Macro-models are also referred to as discrete finite element models in the literature [2]. Among macro-models, distributed and lumped plasticity models are those that have been adopted by various seismic design standards (e.g., ASCE 41-17 [3], Eurocode 8 [4]) and implemented into various analysis software [5, 6, 7] for modeling of reinforced concrete shear walls. The distributed plasticity model (also referred to as the fiber model) discretizes wall cross-section using nonlinear uniaxial fiber elements, assuming that plane sections remain plane after loading, whereas the lumped plasticity approach adopts the use of elastic elements with effective stiffness and zero-length plastic hinges at regions where material nonlinearity is expected to occur (that is, typically member ends). When compared to the fiber model, the lumped plasticity model has several disadvantages, such as neglecting (i) deformations at section neutral axis during loading-unloading, (ii) contribution of horizontal elements (e.g., beams, slabs) that are connected to the walls, (iii) influence of the change in axial load on wall shear strength and stiffness. Nevertheless, the method stands out with its ease-of-use and computational efficiency in practice. It should also be noted that both models have a common disadvantage, that is, shear-flexure interaction is not taken into account as coupled axial-flexural (i.e., P-M) behavior is represented by stress-strain relations of concrete and reinforcing steel along with generally elastic or elastoplastic horizontal spring to reflect the shear behavior. Other modeling techniques in literature account for shear-flexure interaction based on strut and tie method [8, 9, 10, 11] and multiple vertical line element model [10, 12, 13]; however, they are out of the scope of this paper.

Zero-length plastic hinges in the lumped plasticity model can be employed in two forms: (i) rotational springs, (ii) fiber sections. This paper focuses on the use of rotational springs. Nonlinear static and dynamic analyses are conducted based on such plastic hinge models characterized by force-deformation relationships, also known as "backbone curves", to ensure load and deformation paths follow but not exceed them. Backbone curves are constructed by connecting peak points at the first cycle of each increment [3], whereas they can be approximated by a series of linear segments broken at critical points for the sake of practicality (Fig. 1).

Figure 1: (a) Reversed cyclic test setup; (b) loading protocol; (c) load – deformation response [14].
The history of the lumped plasticity approach dates back to the 1960s [15, 16, 17, 18], whereas the earliest introduction of the use of backbone curves in a seismic performance-based evaluation guideline was in 1997 [19, 20]. In later years, cyclic backbone curves for shear walls were recommended in FEMA 356 (2000) [21] and ASCE 41 (2006) [22], which updated the modeling parameters based on the relevant research studies (Wallace et al. for squat walls [23]). Definition of backbone curves and associated modeling parameters for reinforced concrete shear walls in the most recent version of ASCE 41 [3] are presented in Fig.2. The multilinear backbone curve model recommended for shear-controlled walls includes five points at cracking (F), yielding (B), peak strength (C), ultimate displacement (D), and residual strength (E). The backbone curve model for flexure-controlled shear walls (Fig.2(b)) is less refined (excludes the cracking point) as the deformations related to shear are negligible compared to the deformations associated with flexure, whereas the -x axis represents rotation (versus drift) for such walls. The normalized shear force (Q/Q_y) and drift ratio (Δ/h) values at such points are defined according to wall axial load level (as well as shear stress level for flexure-controlled walls), where normalized shear force at cracking is constant and equal to (V_cr/V_y = 0.6) where V_y equals to V_n, i.e., the nominal shear strength calculated based on ACI 318 [24]. It is worth noting that the definition of backbone curves in ASCE 41 was very similar for reinforced concrete columns up until ASCE 41-13 [25], where column modeling parameters were used to obtain based on a combination of column axial load level, reinforcement ratio, and shear stress ratio depending on the column failure mode. However, a major revision has been made in ASCE 41-17, which now provide parameters in the form of equations rather than the past table form. This revision suggests that predictive equations ensure better accuracy and easier use, whereas a similar study is needed for reinforced concrete shear walls, in parallel with this study. Nonlinear modeling of reinforced concrete shear walls based on lumped plasticity approach has been investigated by various researchers for walls with different failure types. The lumped plasticity approach was addressed in terms of plastic hinge length [26, 27, 28] and line-element models [29] for flexure-controlled walls, whereas semi-empirical equations were developed for backbone curves of low-rise shear-controlled walls by Carrillo and Alcocer [30]. More recently, Epackachi et al. addressed the ASCE 41 cyclic backbone curve definitions by recommending equations for peak shear strength for shear-controlled walls [31, 1].

As summarized above, traditional approaches are available in the literature to define backbone curves of reinforced concrete shear walls; however, they either require interpolation and generally provide conservative lower-bound estimates with unknown dispersion (ASCE 41-17) or do not provide the entire behavior characteristics (e.g., equations for lateral strength only [1]). Not to mention, traditional approaches do not provide recommendations on modeling parameters for walls with shear-flexure interaction (according to ASCE 41-17 [3], 1.5 ≤ h_w/h_w ≤ 3). This study aims to fill this gap by approaching the problem from a machine-learning (ML) point of view, which has not been explored before, and to propose predictive models to obtain cyclic backbone curves of RC shear walls based on key wall design properties (concrete compressive strength, horizontal boundary reinforcing ratio, etc.). The prediction of the backbone curves is essentially a multi-output regression problem. Although the use of machine learning methods in the field of structural analysis and material modeling date back to the late 1980s [32, 33], such methods have received a lot of attention in the field of structural earthquake engineering [34], in particular, for modeling the physical relations of a structural system as an input-output problem via clustering, classification, or regression [35, 36]. More recently, the machine learning methods have been utilized to predict earthquake-induced building damage [37], to define backbone curve models for reinforced concrete columns [38].

2 Experimental database

To develop predictive models for shear wall backbone curves based on the machine learning approach, the key wall properties (Fig.3) and the critical backbone variables are designated as input variables (features) and output variables, respectively. To achieve this, more than 500 shear wall specimens from various experimental studies conducted worldwide are collected based on a review of available research [39, 40, 41]. Repaired or strengthened walls, specimens tested under monotonic loading, as well as specimens constructed with diagonal web reinforcement, confinement devices, composite materials (e.g., I-section confinement), or hollow sections are excluded, resulted in 384 specimens useful for this study.
The database includes specimens that have different cross-sections (rectangular [255/384] or barbell-shaped or flanged [129/384]), different end conditions (cantilever [341/384] or rotations fixed on top [43/384]), and different failure types (130 shear-controlled specimens, 126 specimens with shear-flexure interaction, and 128 flexure-controlled specimens). Shear-controlled specimens are those reported to fail in diagonal tension failure or web crushing by reaching their shear strength before flexural yielding occurs, whereas flexure-controlled specimens are those that yield in flexure before reaching their shear strength, thus show damages such as concrete spalling and crushing and/or reinforcing bar buckling at the boundary elements. Specimens that contain both damage types are classified as walls controlled by shear-flexure interaction. Fig. 3 presents the distribution of key design parameters for each failure mode. Further details on the database can be obtained elsewhere [42].

2.1 Definition of backbone curves

Backbone curve of each specimen is obtained using cyclic base shear-top lateral displacement loops by taking the average of responses in positive and negative regions, as shown in Fig. 4(a) for a representative specimen.
Deger ZT, Kaya GT. A Novel GPR-based Prediction Model For Cyclic Backbone Curves of RC Shear Walls, Preprint.

The deformation components are not separated; that is, the total deformation included contributions from flexure, shear, and slip of the reinforcement.

Backbone curves include shear force and displacement at four critical stages, namely: cracking, yielding, reaching shear strength, and reaching ultimate displacement. The shear ($V_{cr}$) and displacement ($\delta_{cr}$) at the cracking point are related through the elastic wall properties; therefore, the following seven backbone variables are investigated in this study: shear at cracking point ($V_{cr}$), shear and displacement at yielding point ($V_y$ and $\delta_y$), shear and displacement at shear capacity ($V_{max}$ and $\delta_{max}$), as well as shear and displacement at displacement capacity ($V_u$ and $\delta_u$).

Figure 4: (a) Load – deformation response [14]; (b) backbone curve definition

Shear and flexure behaviors can be assessed together in Fig. 5(a) based on backbone curves plotted with identifying colors for different failure modes. Observations indicate that deformation capacity is influenced by the level of shear strength as also shown in literature [39, 43, 44, 45]. Wall shear strength degrades with increasing ductility, leading to shear failure (red backbone curves), whereas, higher displacements can be reached under low shear levels (blue backbone curves).

Fig. 5(b) demonstrates the distribution of the seven backbone variables in the wall database. The seven backbone variables are compared to ASCE 41-17 recommended modeling parameters (a, b, c, d, e, f, g values shown in Fig. 2) in Tables 1 and 2 for shear-controlled and flexure-controlled walls, respectively. Note that dominant wall types in Table 1 and Table 2 are identified based on reported failure modes of the specimens, as data for walls controlled by flexure based on ASCE 41-17 classification criteria (i.e., shear-controlled if $h_w/l_w<1.5$, flexure-controlled if $h_w/l_w>3$) is too scarce to make an assessment. Comparisons indicate that ASCE 41 modeling parameters achieve a fairly good agreement for walls subjected to low axial load with the following exceptions for the sake of conservatism: (i) ASCE 41 neglects the significant increase in the slope after yield point (between $V_y$ and $V_{max}$), (ii) ASCE 41 residual strength recommendations ($V_u$) are much lower for shear-controlled walls, (iii) ASCE 41 does not address the data dispersion, therefore, can be too conservative in some cases. For other wall types, higher discrepancies are observed except for well-predicted $f$, (i.e. $V_{cr}/V_y$) value for shear-controlled walls. It is noted that ASCE 41 does not provide modeling parameters for walls that have shear-flexure interaction.

3 Background on Gaussian Process Regression

Gaussian Process Regression (GPR) is obtained by generalizing linear regression. The GPR is a collection of random variables, with a joint Gaussian distribution, hence is a probabilistic multivariate regression method. Let us $D = \{(x_i, f_i), i = 1, \ldots, n\}$ refer to a given training set, where $f_i = f(x_i)$ is the value of the function at the sample $x_i$. The Gaussian Process (GP), a generalization of the Gaussian probability distribution, is represented as follows:

$$f(x) \sim GP(m(x), k(x, x_j))$$

where $x$ and $k(x, x_j)$ refer to mean and covariance function, respectively. The covariance matrix express the similarity between the pair of random variables and determined by means of squared exponential (SE) kernel.
Figure 5: (a) Backbone curves of 384 specimens in the database, (b) Distribution of seven backbone curve variables

Table 1: Comparisons of database backbone values with ASCE 41-17 recommendations for shear-controlled walls

| Parameter          | \( \frac{(A_s-A'_s)E+P_{\text{tw}}}{f_{\text{ct}}w} \) | ASCE 41-17   | Database   |
|--------------------|----------------------------------|-------------|------------|
| \( V_{\text{max}}/V_y \) | ≤ 0.5 | 1.0 | 1.43±0.37  |
|                     | >0.5 | 1.0 | 1.18±0.05  |
| \( V_u/V_{\text{max}} \) | ≤ 0.5 | 0.2 | 0.65±0.29  |
|                     | >0.5 | 0.0 | 0.75±0.13  |
| \( V_{\text{cr}}/V_y \) | ≤ 0.5 | 0.6 | 0.58±0.17  |
|                     | >0.5 | 0.6 | 0.61±0.23  |
| \( \delta_{\text{max}}/h \) | ≤ 0.5 | 1.0 | 0.97±0.48  |
|                     | >0.5 | 0.75 | 1.21±0.19  |
| \( \delta_u/h \) | ≤ 0.5 | 2.0 | 1.84±1.19  |
|                     | >0.5 | 1.0 | 2.05±0.46  |
| \( \delta_y/h \) | ≤ 0.5 | 0.4 | 0.42±0.3   |
|                     | >0.5 | 0.4 | 0.57±0.08  |

The function given as

\[
k(x_i, x_j) = exp(||x_i - x_j||^2/(2\sigma^2))
\]

where \( \sigma \) is the kernel width that needs to be optimized. The predictive function, \( f^* \), can be obtained for a given test sample \( X* \), as following,

\[
p(f_*|X*, X, f) = N(f_*|\mu_*, \sigma_*)
\]
Deger ZT, Kaya GT. A Novel GPR-based Prediction Model For Cyclic Backbone Curves of RC Shear Walls, Preprint.

Table 2: Comparisons of database backbone values with ASCE 41-17 recommendations for flexure-controlled walls

| Confined boundary | ASCE 41-17 | Database | ASCE 41-17 | Database |
|-------------------|------------|----------|------------|----------|
| \( \frac{(A_s-A'_s)k_E}{\tau w_t w_f E} \) | \( V_{max}/V_y \) | \( c = V_v/V_{max} \) |
| <=0.1 <=4 Yes | 1.0 | 1.29 | 0.75 | 0.75 |
| <=0.1 >=6 Yes | 1.0 | 1.38 | 0.4 | 0.70 |
| >0.25 <=4 Yes | 1.0 | 1.37 | 0.6 | 0.75 |
| >0.25 >=6 Yes | 1.0 | 1.22 | 0.3 | 0.73 |
| <=0.1 <=4 No | 1.0 | 1.27 | 0.6 | 0.70 |
| <=0.1 >=6 No | 1.0 | 1.0 | 0.3 | 0.75 |
| >0.25 <=4 No | 1.0 | 1.34 | 0.25 | 0.62 |
| >0.25 >=6 No | 1.0 | 1.22 | 0.3 | 0.73 |

where

\[
\mu_s = k_s^T[K(X, X) + \sigma_n^2 I]^{-1}y \quad \text{(4)}
\]

\[
\sigma^2_s = k(X_s, X_s) - k_s^T[K(X, X) + \sigma_n^2 I]^{-1}k_s \quad \text{(5)}
\]

\( k_s \) is a vector holding the covariance values between training samples \( X \) and the test samples \( X_s \). The \( \sigma^2_n \) and \( \sigma^2_s \) refer to the noise variance and the confidence measure associated with the model to output, respectively.

4 Implementation of Gaussian Process Regression and Results

4.1 Experimental Setup

The key wall properties (Fig.3) and the critical backbone variables are designated as input variables (features) and output variables for the regression analyses, respectively. Note that the problem is a multi-output regression since multiple outputs are needed to be estimated. Training and test datasets are randomly selected as disjoint sets from the database at a ratio of 90% and 10%, respectively, and this random splitting is repeated a hundred times. The purpose of a hundred trials is to assess the proposed method’s generalization capability and robustness. Because of the randomization, training/test pairings vary in each realization, resulting in a hundred different learning models. The GPR-based models are constructed on each training dataset, whereas the accuracy of the resulting models is evaluated based on associated test dataset. The validation results are then averaged over a hundred trials for each output. It should be noted that experiments with fewer training datasets (e.g. 80% for training and 20% for testing) are also considered; however, a large number of trials are overtrained, implying that the amount of the training set is not sufficient to generate a reliable learning model. Based on this experience, the size of the training set is increased to 90% of all data.

The kernel parameters of the GPR method are automatically tuned using gradient-based optimization routines to maximize the marginal likelihood.

The performance of models is evaluated based on three criteria: the mean coefficient of determination (R^2), relative root mean square error (RelRMSE), and the prediction accuracy, that is, the ratio of predicted to
Deger ZT, Kaya GT. A Novel GPR-based Prediction Model For Cyclic Backbone Curves of RC Shear Walls, Preprint.

actual values (P/A). The experiments are conducted based on a hundred trials to ensure the reliability and robustness of the proposed model, and the averaged results are reported. All the codes are developed in Matlab on a personal computer with Dual Core Intel Core i7 CPU (3.50GHz), 16GB RAM, and MacOS operating system.

Data pre-processing has been shown as a useful technique to improve the quality of machine learning algorithms by transforming the data to a more informative format. Among several data transformation techniques available, Box-cox and log transformation methods are applied to the outputs. The original form of Box-Cox transformation equation is as shown in Eq. 6.

$$y(\lambda) = \begin{cases} y^{\frac{1}{\lambda}} - 1, & \text{if } \lambda \neq 0 \\ \log(y), & \text{if } \lambda = 0 \end{cases}$$

(6)

where $y$ is the data value and $\lambda$ is a value between -5 and 5. The optimal $\lambda$ is the one that transforms the data set to the best approximation of normal distribution. The log transformation corresponds to $\lambda$ equals zero. It is observed that the use of log transformation does not improve the results, whereas Box-Cox (optimal $\lambda = -0.3$) leads to an increase in model accuracy for all outputs, particularly for $\delta_y$, $\delta_{max}$, and $\delta_u$ (Table 3).

Table 3: Effect of data transformation on GPR model results (Averaged values over 100 trials).

| Output | No Transformation | Box-Cox Transformation |
|--------|-------------------|------------------------|
| $V_y$  | 0.80 0.36         | 0.84 0.33              |
| $\delta_y$ | 0.72 0.55       | 0.82 0.50              |
| $V_{max}$ | 0.87 0.28       | 0.89 0.28              |
| $\delta_{max}$ | 0.84 0.43      | 0.91 0.33              |
| $V_u$  | 0.57 0.64         | 0.62 0.59              |
| $\delta_u$ | 0.77 0.40       | 0.89 0.29              |

4.2 Experimental Results and Discussion

Prediction accuracy and correlation values based on machine learning experiments are summarized in Table 4 for each output. The results reveal that the GPR method achieves better correlations ($R^2>0.8$) for walls that failed in shear (S) or shear-flexure interaction (SF), and the relatively low correlation coefficients (mean $R^2$ around 0.65) reduce the performance of models. It is noted that modeling parameters in the horizontal axis of backbone curves ($\delta_y$, $\delta_{max}$, $\delta_u$) were also studied in terms of chord rotation and curvature for flexure-controlled walls and drift ratio for the shear-controlled walls; however, correlation coefficients were much lower. Therefore, the results presented in Fig. 6 focus on walls that reached their shear strength, that is, those controlled by shear or shear-flexure interaction, and the prediction models are proposed for such walls for the seven backbone variables.

Due to overtraining, trials with poor generalization capability such as those with the highest correlation in training but very low correlation in testing, are omitted, resulting in the mean values shown with the red diamond in Fig. 6 and the median values tabulated in Table 4. The results indicate that coefficient of determination values are over 80% with low relative RMSE values (around 0.3) for all seven backbone variables; except for the shear force corresponding to ultimate deformation (i.e., $V_u$), the mean $R^2$ is around 60% with higher error. As seen in the boxplots, prediction accuracy is very close to 1.0 with low standard deviations for all seven backbone variables, except for the shear force at ultimate deformation. The lower performance for $V_u$ implies that input-output relation can not be established for this output, potentially because some specimens reached their ultimate deformation, whereas, for the others, ultimate deformation.
Table 4: Averaged performances of the GPR-based models over 100 trials

| Output | R² (median) | Mean RelRMSE | Mean P/A | Std. Dev. P/A |
|--------|-------------|---------------|-----------|--------------|
| Vᶜʳ    | S, SF       | 0.71          | 0.44      | 1.08         | 0.47         |
|        | All         | 0.67          | 0.54      | 1.07         | 0.46         |
| Vʸ     | S, SF       | 0.84          | 0.33      | 1.02         | 0.29         |
|        | All         | 0.79          | 0.52      | 1.06         | 0.41         |
| δʸ     | S, SF       | 0.82          | 0.5       | 1.06         | 0.46         |
|        | All         | 0.64          | 0.55      | 1.11         | 0.66         |
| Vᵐᵃˣ   | S, SF       | 0.89          | 0.28      | 1.04         | 0.26         |
|        | All         | 0.80          | 0.69      | 1.09         | 0.57         |
| δᵐᵃˣ   | S, SF       | 0.91          | 0.33      | 1.05         | 0.38         |
|        | All         | 0.79          | 0.42      | 1.03         | 0.57         |
| Vᵘ     | S, SF       | 0.62          | 0.50      | 1.3          | 0.99         |
|        | All         | 0.79          | 0.42      | 1.03         | 0.57         |
| δᵘ     | S, SF       | 0.89          | 0.29      | 1.03         | 0.29         |
|        | All         | 0.79          | 0.33      | 1.05         | 0.36         |

Figure 6: GPR model performances for the backbone variables over 100 trials.

is estimated at 80% of shear strength, as mentioned earlier. This may have changed the structure of the problem, resulting in relatively poor predictions.

The correlation between the actual and predicted values is assessed based on the proposed GPR-based predictive model that achieves the highest R², the P/A closest to 1.0, and the lowest error simultaneously (Fig. 7) for each backbone variable. The scatter plots indicate that data are closely distributed around the line of y = x, implying that the proposed models are able to make accurate predictions when compared to the actual values. Results demonstrate that the proposed GPR-based models achieve excellent performance in predicting shear strength V_max and corresponding displacement (δ_max) with highest coefficients of determination (R² equals to 0.96 and 0.97, respectively) and smallest relative RMSE. The proposed model for ultimate
deformation ($V_u$) also achieves good agreement with the actual data even though mean values over 100 trials were lower as this output is more data-dependent.

To evaluate the influence of each wall property (i.e. feature, Fig.3) on the model output, optimized lengthscale hyperparameters of the kernel function are used in the GPR model as each hyperparameter controls the spread of each input. The hyperparameter optimization is carried out by computing the partial derivatives of the negative log-marginal likelihood function with respect to the given parameter. As small values show greater relevance, their inverses are used for feature ranking. The most influential features are axial load ratio ($P/(A_g f_c')$) and aspect ratio ($h_w/l_w$) for force-related backbone variables ($V_{cr}$, $V_y$, $V_{max}$, and $V_u$), whereas shear span ratio ($M/Vl_w$) is more critical for the displacement variables ($\delta_y$, $\delta_{max}$, $\delta_u$).

The performance of the proposed models is also compared to ASCE 41 modeling recommendations based on 20 unseen test data in Table 5. Mean and standard deviation of the predicted/actual (P/A) values are...

Figure 7: Scatter plots for correlations of the proposed model for each backbone variable
assessed as performance indicators for each backbone variable, where ASCE 41 values are back-calculated from the normalized modeling parameters. Mean P/A values are close to 1.0 for all seven backbone variables for the GPR-based models, whereas higher errors are observed based on ASCE 41. The standard deviation of the P/A is also much smaller for the GPR-based models. Results indicate that the proposed models outperform the traditional ASCE 41 approach for a wider spectrum of walls (versus only low axial loads, as discussed in Table 1 with higher accuracy.

| Output | ASCE 41 | GPR |
|--------|---------|-----|
|        | P/A mean | P/A dispersion | P/A mean | P/A dispersion |
| $V_y$  | 2.17     | 2.41 | 0.97   | 0.16  |
| $\delta_y$ | 1.41     | 1.08 | 0.98   | 0.32  |
| $V_{max}$ | 1.28     | 0.52 | 0.99   | 0.18  |
| $\delta_{max}$ | 0.90     | 0.44 | 0.99   | 0.17  |
| $V_u$  | 0.40     | 0.28 | 1.04   | 0.39  |
| $\delta_u$ | 1.27     | 0.55 | 1.02   | 0.18  |

Machine learning experiments revealed that unusually thin ($l_w/t_w>20$) shear walls and slender walls ($h_w/l_w>3.0$) decrease the correlation and the prediction accuracy. The former is expected as typical $l_w/t_w$ for shear walls is reported as 13.3 [49] and reaches $l_w/t_w=19$ for around 350cm-long shear walls [50] which corresponds to the maximum wall length in the database. The latter factor decreasing correlations ($h_w/l_w$) can be associated with the typical wall behavior, that is, squat wall responses are governed by shear and slender wall responses are dominated by flexural behavior. However, experimental results showed that squat walls may fail in flexure if adequate reinforcement and detailing are provided [51], whereas slender walls may fail in shear (web crushing) [52]. Therefore, backbone relations of slender walls need to be analyzed separately even though they happened to fail in shear.

As noted above, the proposed GPR-based backbone curve models are generalized and can be used regardless of failure modes. However, results demonstrate that the predictions are more reliable especially for shear-controlled and shear-flexure controlled walls. It is worth advising the potential users of the proposed GPR-based models in that identification of dominant wall behavior (failure mode) based on ASCE 41 may lead to conservative assessments. For example, some moderate aspect ratio walls ($1.5<h_w/l_w<3$) would be classified as shear-flexure controlled even though they would actually fail in flexure, creating a false impression that backbones of such walls can be predicted based on the proposed GPR-based models. Therefore, prior to the use of the proposed models for a new shear wall specimen, a more accurate failure mode prediction [53, 54] is recommended.

It is also noted that other machine-learning methods including k-nearest neighbour regression (KNN) [55], least absolute shrinkage and selection operator (LASSO) [56], neighbourhood component analysis (NCA) [57], random forest [58], regularized logistic regression (RLR) [59], and support vector regression (SVR) [60] were also utilized in this study in an attempt to obtain higher correlations and accuracies; however, none of them were as accurate as the GPR method.

## 5 Summary and Conclusions

Backbone curves are critical component modeling tools in element-based nonlinear modeling to represent the response under reversal drift cycles. It is essential to simulate the hysteretic response as accurately as possible to achieve realistic assessments as nonlinear static and dynamic analyses are conducted based on such curves.

This study aims to address this need by developing predictive models for backbone curves based on a comprehensive shear wall experimental database using a powerful machine learning method, that is Gaussian
Deger ZT, Kaya GT. A Novel GPR-based Prediction Model For Cyclic Backbone Curves of RC Shear Walls, *Preprint*.

Process Regression (GPR). Seven backbone curve variables associated with four limit states, namely: cracking, yielding, peak shear strength, and ultimate deformation capacity are predicted in terms of twelve wall design properties (e.g. concrete compressive strength, reinforcement details, wall geometry). To achieve this, the backbone curve variables and wall design properties are designated as outputs and inputs, respectively. The database is divided into train and test data sets, whereas the performances of the proposed models are evaluated based on statistical metrics (e.g. coefficient of determination, relative root mean square error, prediction accuracy) based on test data sets. The prediction performances of the proposed GPR models are also compared to traditional modeling approaches (ASCE 41-17) and shown to be superior.

For best performance, the proposed models are recommended for walls that satisfy the following conditions: (i) wall behavior is controlled by shear or shear-flexure interaction, (ii) \( l_w/t_w \leq 20 \), and (iii) \( h_w/l_w \leq 3 \).

The results reveal that one can predict the shear force - displacement relations, thus, the existing capacity and seismic performance of RC shear walls, close to accurate using the proposed GPR-based models. This study will contribute to the literature by providing tools that achieve dual objectives of high accuracy performance and computational efficiency to ensure more realistic performance evaluations.

6 Acknowledgements

The project has been supported by funds from the Scientific and Technological Research Council of Turkey (TUBITAK) under Project No: 218M535. Opinions, findings, and conclusions in this paper are those of the authors and do not necessarily represent those of the funding agency.
Deger ZT, Kaya GT. A Novel GPR-based Prediction Model For Cyclic Backbone Curves of RC Shear Walls, *Preprint*.

**Notation**

- \( A_g \) = gross cross section area of wall
- \( \delta_{cr} \) = cracking displacement
- \( \delta_y \) = yield displacement
- \( \delta_{max} \) = displacement when peak shear is reached
- \( \delta_u \) = ultimate displacement
- \( \varepsilon \) = the noise term
- \( E(X|Y) \) = expected value of random variable \( X \) given \( Y \)
- \( f'_{cc} \) = concrete compressive strength
- \( f_{yb} \) = longitudinal boundary reinforcement yield strength
- \( f_{yl} \) = longitudinal web reinforcement yield strength
- \( f_{ysh} \) = transverse boundary reinforcement yield strength
- \( f_{yt} \) = transverse web reinforcement yield strength
- \( h_w \) = wall height
- \( h_w/l_w \) = aspect ratio
- \( I \) = the identity matrix
- \( l_w \) = wall length
- \( \mathcal{L} \) = the loss function
- \( m \) = number of specimens
- \( M/Vl_w \) = shear-span ratio
- \( n \) = number of design parameters
- \( P/(A_g f'_{cc}) \) = axial load ratio
- \( P(A|B) \) = probability of event \( A \) given event \( B \) occured
- \( s \) = stirrups spacing
- \( t_w \) = wall thickness
- \( \rho_{bl} \) = longitudinal boundary reinforcement ratio
- \( \rho_{l} \) = longitudinal web reinforcement ratio
- \( \rho_{sh} \) = transverse boundary reinforcement ratio
- \( \rho_{t} \) = transverse web reinforcement ratio
- \( w \) = weighing vector of features
- \( X \) = \( m \times n \) dimensional matrix holding all the features
- \( x_i \) = \( i^{th} \) feature value of the ML model
- \( y \) = \( m \) dimensional output vector
- \( V_{cr} \) = shear force at cracking
- \( V_y \) = shear force at yielding
- \( V_{max} \) = peak shear force
- \( V_u \) = shear force when ultimate displacement is reached
References

[1] R Hamburger, GG Deierlein, DE Lehman, DG Lignos, LN Lowes, R Pekelnicky, PB Shing, P Somers, and JW Van de Lindt. Recommended modeling parameters and acceptance criteria for nonlinear analysis in support of seismic evaluation, retrofit, and design. *Rep. No. NIST GCR*, pages 17–917, 2017.

[2] Fabio Taucer, Enrico Spacone, and Filip C Filippou. *A fiber beam-column element for seismic response analysis of reinforced concrete structures*, volume 91. Earthquake Engineering Research Center, College of Engineering, University, 1991.

[3] ASCE-41. *ASCE Standard, ASCE/SEI, 41-17, Seismic Evaluation and Retrofit of Existing Buildings*. American Society of Civil Engineers, 2017.

[4] British Standard. *Eurocode 8: Design of structures for earthquake resistance. Part 3:1998–3*, 2005.

[5] Perform3D. *Nonlinear Analysis and Performance Assessment of 3D Structures, PERFORM-3D*. Computers and Structures, Inc., Berkeley, California., 2016.

[6] ETABS. *Integrated Analysis, Design and Drafting of Building Systems*. Computers and Structures, Inc., Berkeley, California., 2017.

[7] Silvia Mazzoni, Frank McKenna, Michael H Scott, Gregory L Fenves, et al. Opensees command language manual. *Pacific Earthquake Engineering Research (PEER) Center*, 264, 2006.

[8] Alfonso Vulcano, Viltelmo V Bertero, and Vincenzo Colotti. Analytical modeling of rc structural walls. In *Proceedings, 9th world conference on earthquake engineering*, volume 6, pages 41–46, 1988.

[9] Marios Panagiotou, José I Restrepo, Matthew Schoettler, and Geonwoo Kim. Nonlinear cyclic truss model for reinforced concrete walls. *ACI Structural Journal*, 109(2):205, 2012.

[10] M Fischinger, K Rejec, and T Isakovic. Modeling inelastic shear response of rc walls. In *Proceedings, 15th World Conference on Earthquake Engineering*, volume 2120, 2012.

[11] Kristijan Kolozvari, Lauren Biscombe, Farhad Dashti, Rajesh P Dhakal, Aysegul Gogus, M Fethi Gullu, Richard S Henry, Leonardo M Massone, Kutay Orakcal, Fabian Rojas, et al. State-of-the-art in nonlinear finite element modeling of isolated planar reinforced concrete walls. *Engineering Structures*, 194:46–65, 2019.

[12] Kristijan Kolozvari, Kutay Orakcal, and John W Wallace. Modeling of cyclic shear-flexure interaction in reinforced concrete structural walls. i: Theory. *Journal of Structural Engineering*, 141(5):04014135, 2015.

[13] Kristijan Kolozvari, Kamiar Kalbasi, Kutay Orakcal, Leonardo M Massone, and John Wallace. Shear–flexure-interaction models for planar and flanged reinforced concrete walls. *Bulletin of Earthquake Engineering*, 17(12):6391–6417, 2019.

[14] Leonardo M Massone, Sebastián Díaz, Ignacio Manríquez, Fabián Rojas, and Ricardo Herrera. Experimental cyclic response of rc walls with setback discontinuities. *Engineering Structures*, 178:410–422, 2019.

[15] Ray W Clough. Effect of stiffness degradation on earthquake ductility requirements. In *Proceedings of Japan earthquake engineering symposium*, 1966.

[16] Ray W Clough and Kalman L Benuska. Nonlinear earthquake behavior of tall buildings. *Journal of the Engineering Mechanics Division*, 93(3):129–146, 1967.

[17] Melbourne Fernald Giberson. *The response of nonlinear multi-story structures subjected to earthquake excitation*. PhD thesis, California Institute of technology, 1967.

[18] Toshikazu Takeda, Mete A Sozen, and N Norby Nielsen. Reinforced concrete response to simulated earthquakes. *Journal of the structural division*, 96(12):2557–2573, 1970.

[19] Building Seismic Safety Council. *Nehrp guidelines for the seismic rehabilitation of buildings* (fema publication 273), 1997.
Deger ZT, Kaya GT. A Novel GPR-based Prediction Model For Cyclic Backbone Curves of RC Shear Walls, Preprint.

[20] Building Seismic Safety Council. Nehrp commentary on the guidelines for the seismic rehabilitation of buildings (fema publication 274), 1997.

[21] FEMA-356. Prestandard and commentary for the seismic rehabilitation of building. Rehabilitation, 221, 2000.

[22] ASCE-41. ASCE Standard, ASCE/SEI, 41-06, Seismic Evaluation and Retrofit of Existing Buildings. American Society of Civil Engineers, 2006.

[23] JW Wallace. Lightly reinforced wall segments. In New Information on the Seismic Performance of Existing Concrete Buildings Seminar Notes, pages 0–100. Earthquake Engineering Research Institute Oakland, CA, 2006.

[24] ACI Committee and International Organization for Standardization. Building code requirements for structural concrete (ACI 318-19) and commentary. American Concrete Institute, 2019.

[25] ASCE-41. ASCE Standard, ASCE/SEI, 41-13, Seismic Evaluation and Retrofit of Existing Buildings. ASCE.: Standard. American Society of Civil Engineers, 2013.

[26] Alfredo Bohl and Perry Adebar. Plastic hinge lengths in high-rise concrete shear walls. ACI Structural Journal, 108(2):148, 2011.

[27] Anna C Birely. Seismic performance of slender reinforced concrete structural walls. University of Washington, 2012.

[28] Ilker Kazaz and Polat Guklan. Süněklik düzeyi yüksek betonarme perdelerdeki hasar sınırları. Teknik Dergi, 23(114):6113–6140, 2012.

[29] Joshua S Pugh, Laura N Lowes, and Dawn E Lehman. Nonlinear line-element modeling of flexural reinforced concrete walls. Engineering Structures, 104:174–192, 2015.

[30] Julian Carrillo and Sergio M Alcocer. Backbone model for performance-based seismic design of rc walls for low-rise housing. Earthquake Spectra, 28(3):943–964, 2012.

[31] Siamak Epackachi, Nikhil Sharma, Andrew Whittaker, Ronald O Hamburger, and Ayse Hortacsu. A cyclic backbone curve for shear-critical reinforced concrete walls. Journal of Structural Engineering, 145(4):04019006, 2019.

[32] H Adeli and C Yeh. Perceptron learning in engineering design. Computer-Aided Civil and Infrastructure Engineering, 4(4):247–256, 1989.

[33] RD Vanluchene and Roufei Sun. Neural networks in structural engineering. Computer-Aided Civil and Infrastructure Engineering, 5(3):207–215, 1990.

[34] Hadi Salehi, Saptarshi Das, Shantanu Chakrabarty, Subir Biswas, and Rigoberto Burgueño. Damage identification in aircraft structures with self-powered sensing technology: A machine learning approach. Structural Control and Health Monitoring, 25(12):e2262, 2018.

[35] Bruno Sudret, Chu Mai, and Katerina Konakli. Assessment of the lognormality assumption of seismic fragility curves using non-parametric representations. Earthquake Engineering & Structural Dynamics, 43(May):2075–2095, 2014.

[36] Youbao Jiang, Jun Luo, Guoyu Liao, and Yulai Zhao. An efficient method for generation of uniform support vector and its application in structural failure function fitting. Structural Safety, 54:1–9, 2015.

[37] Yu Zhang, Henry V. Burton, Han Sun, and Mehrdad Shokrabadi. A machine learning framework for assessing post-earthquake structural safety. Structural Safety, 72:1–16, 2018.

[38] Huan Luo and Stephanie German Paal. Machine learning-based backbone curve model of reinforced concrete columns subjected to cyclic loading reversals. Journal of Computing in Civil Engineering, 32(5):04018042, 2018.

[39] Zeynep Tuna Deger and Cagri Basdogan. Empirical expressions for deformation capacity of reinforced concrete structural walls. ACI Structural Journal, 116(6), 2019.

[40] NEES. Nees: Shear wall database, 2017. Accessed: 2020-04-11.
Deger ZT, Kaya GT. A Novel GPR-based Prediction Model For Cyclic Backbone Curves of RC Shear Walls, Preprint.

[41] Sofia Grammatikou, Dionysis Biskinis, and Michael N Fardis. Strength, deformation capacity and failure modes of rc walls under cyclic loading. *Bulletin of earthquake engineering*, 13(11):3277–3300, 2015.

[42] Zeynep Tuna Deger, G Taskin Kaya, F Sutcu, B Topaloglu, and S Tahaei. Betonarme perdelerde enerji sönümleme davranısının meta-modelleme yöntemleriyle İncelenmesi (in Turkish). Technical report, The Scientific and Technological Research Council of Turkey, Ankara, Turkey, 2021.

[43] RG Oesterle, AE Fiorato, and William Gene Corley. *Reinforcement details for earthquake-resistant structural walls*. Portland Cement Association, 1981.

[44] MJ Nigel Priestley, Ravinda Verma, and Yan Xiao. Seismic shear strength of reinforced concrete columns. *Journal of structural engineering*, 120(8):2310–2329, 1994.

[45] PEER/ATC-72. *Modeling and Acceptance Criteria for Seismic Design and Analysis of Tall Buildings*. Applied Technology Council, Redwood City, CA, 2010.

[46] MATLAB. *version 7.10.0 (R2019a)*. The MathWorks Inc., Natick, Massachusetts, 201.

[47] Yue Jiang, Bojan Cukic, and Tim Menzies. Can data transformation help in the detection of fault-prone modules? In *Proceedings of the 2008 workshop on Defects in large software systems*, pages 16–20, 2008.

[48] George EP Box and David R Cox. An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(2):211–243, 1964.

[49] John W Wallace. Behavior, design, and modeling of structural walls and coupling beams—lessons from recent laboratory tests and earthquakes. *International Journal of Concrete Structures and Materials*, 6(1):3–18, 2012.

[50] CE Magna-Verdugo and SK Kunnath. Non-linear response analysis of reinforced concrete shear walls using multispring macro-models. In *Proc. of the 10th National Conf. in Earthq. Eng*, 2014.

[51] Thomas N Salonikios and Andreas J Kappos. Cyclic load behavior of low-slenderness reinforced. *ACI Structural Journal*, 96(4), 1999.

[52] Chadchart Sittipunt, Sharon L Wood, Panitan Lukkunaprasit, and Pichai Pattararatthanakul. Cyclic behavior of reinforced concrete structural walls with diagonal web reinforcement. *Structural Journal*, 98(4):554–562, 2001.

[53] Sujith Mangalathu, Hansol Jang, Seong-Hoon Hwang, and Jong-Su Jeon. Data-driven machine-learning-based seismic failure mode identification of reinforced concrete shear walls. *Engineering Structures*, 208:110331, 2020.

[54] Zeynep Tuna Deger and Gulsen Taskin Kaya. Glass-box model representation of seismic failure mode prediction for conventional reinforced concrete shear walls. Unpublished results.

[55] T. Cover and P. Hart. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theor.*, 13(1):21–27, September 2006.

[56] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society (Series B)*, 58:267–288, 1996.

[57] Jacob Goldberger, Geoffrey E Hinton, Sam Roweis, and Russ R Salakhutdinov. Neighbourhood components analysis. *Advances in neural information processing systems*, 17:513–520, 2004.

[58] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[59] David W Hosmer Jr, Stanley Lemeshow, and Rodney X Sturdivant. *Applied logistic regression*, volume 398. John Wiley & Sons, 2013.

[60] Vladimir N. Vapnik. *Statistical Learning Theory*. Wiley-Interscience, 1998.