This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Feng, De-Cheng; Chen, Shi-Zhi; Azadi Kakavand, Mohammad Reza; Taciroglu, Ertugrul
Probabilistic Model Based on Bayesian Model Averaging for Predicting the Plastic Hinge Lengths of Reinforced Concrete Columns

Published in:
Journal of Engineering Mechanics

DOI:
[10.1061/\(ASCE\)EM.1943-7889.0001976](https://doi.org/10.1061/(ASCE)EM.1943-7889.0001976)

Published: 01/10/2021

Document Version
Peer reviewed version

Please cite the original version:

Feng, D-C., Chen, S-Z., Azadi Kakavand, M. R., & Taciroglu, E. (2021). Probabilistic Model Based on Bayesian Model Averaging for Predicting the Plastic Hinge Lengths of Reinforced Concrete Columns. *Journal of Engineering Mechanics*, 147(10), [04021066]. [https://doi.org/10.1061/\(ASCE\)EM.1943-7889.0001976](https://doi.org/10.1061/(ASCE)EM.1943-7889.0001976)

Probabilistic Model Based on Bayesian Model Averaging for Predicting the Plastic Hinge Lengths of Reinforced Concrete Columns

De-Cheng Feng¹, Shi-Zhi Chen², Mohammad Reza Azadi Kakavand³, and Ertugrul Taciroglu⁴

¹Associate Professor, Key Laboratory of Concrete and Prestressed Concrete Structures of the Ministry of Education, Southeast University, Nanjing 211189, China.

²Lecturer, School of Highways, Chang'an University, Xi'an 710064, China.

³Postdoc Researcher, Department of Civil Engineering, School of Engineering, Aalto University, Espoo 02150, Finland

⁴Professor, Department of Civil & Environmental Engineering, University of California, Los Angeles, CA 90095, USA (corresponding author, Email: etacir@ucla.edu).

ABSTRACT

A probabilistic model is devised for predicting the plastic hinge lengths (PHLs) of reinforced concrete (RC) columns. Seven existing parametric models are evaluated first using a comprehensive database comprising PHL measurements from 133 RC column tests. It is observed that the performances of these seven models are fair (as opposed to strong), and their predictions bear significant uncertainties. A novel technique is devised to combine them into a weighted-average super model wherein the weights are determined via Bayesian inference. This approach naturally produces the weights' statistical moments, and thus, the resulting model is a probabilistic one that is amenable for performance-based seismic design and assessment analyses. Prediction comparisons indicate that the proposed super-model has a higher performance than all prior models. The new model is easily expandable should more test data become available.

Keywords: Plastic hinge length, reinforced concrete column, probabilistic model, model uncertainty, model averaging, Bayesian inference, performance-based seismic engineering

INTRODUCTION

Plastic hinge length (PHL) of reinforced concrete (RC) columns is a key engineering parameter used in seismic design and analysis of RC structures, representing the localized region of inelastic column deformations. With PHL, a column's displacement ductility can be easily obtained by integrating the corresponding inelastic curvature profile (Ning and Li, 2015). PHL is also required in finite element analyses, which require a correct representation of this localized region to capture load redistribution patterns and thus an objective overall structural response (Scott and Hamutçuoğlu, 2008; Feng et al., 2016; Feng and Ren, 2017; Feng et al., 2019).

The determination of the PHL can be achieved by experimental investigations and/or high-fidelity numerical simulations. The experimental studies are restricted by cost and time, and they can only provide a case-by-case prediction of the PHL. On the other hand, the high-fidelity numerical simulations (Kakavand and Allahvirdizadeh, 2019) are time-consuming and will introduce numerous modeling uncertainties—namely, finite element type, mesh sizes, and other untested model parameters. When a PHL is needed for calculating the structural ductility in the design stage, it will not be possible/feasible to conduct *ad hoc* experiments or high-fidelity simulations.

What is needed, therefore, is a rapid prediction of the PHL value for structural design and/or performance assessment. In this regard, numerous mechanics-driven empirical models have been derived for predicting the PHL, for example, by [Park et al. \(1982\)](#); [Priestley and Park \(1987\)](#); [Paulay and Priestley \(1992\)](#); [Sheikh and Khoury \(1993\)](#); [Sheikh et al. \(1994\)](#); [Panagiotakos and Fardis \(2001\)](#); [Lu et al. \(2005\)](#); [Berry et al. \(2008\)](#); [Bae and Bayrak \(2008\)](#); [Elmenschawi et al. \(2012\)](#); [Ning and Li \(2015\)](#). These models are typically based on fitting functional forms to data and comprise a variety of physical parameters that are deemed to have strong correlations with observed PHL values. However, as comprehensively studied by [Ning and Li \(2015\)](#), PHL mechanisms embody multiple phenomena (e.g., strain penetration, tension shift, moment gradient) and are affected by a large number of parameters (e.g., column dimensions, material properties, and reinforcing details). It is inherently difficult to fully consider them in a single model. Not surprisingly, it can be observed from the equations of existing PHL models that most emphasize only a few of the mechanisms at a time, and thus, their accuracies are limited, and their predictions have large variations.

The large scatter of the PHLs predicted by different models is a crucial model uncertainty problem. To resolve this issue, [Ning and Li \(2015\)](#) proposed a new and probabilistic model. They carefully analyzed the PHL mechanisms and selected four key parameters, which all have a coefficient representing uncertainties. By using the generalized likelihood uncertainty estimation method ([Ning and Li, 2017, 2018](#)) and a test database of PHL measurements, they obtained the statistical distributions of the coefficients and established a probabilistic model. While the performance of this model is greatly enhanced compared to previous ones, there is still space for improvement since the model only consists of four parameters, and the random coefficients are constrained to pre-defined ranges.

After a comprehensive evaluation of the existing PHL models, it can be observed that a few consistently over- and under-estimate the PHL values. Therefore, a new model that averages the results by these models can be contemplated—here, in a Bayesian sense. In fact, such an approach can be viewed as belonging to an existing branch of model-averaging methods in the field of statistics ([Hoeting et al., 1999](#); [Wasserman, 2000](#)), which has been widely applied in numerous fields such as ecology and economics (see, e.g., [Dormann et al., 2018](#); [Steel, 2017](#)). The basic idea is to combine several regression models into a single one through weighted averaging to reduce the model uncertainty. When incorporated with Bayesian inferencing, the probabilistic distributions of the quantity of interest can also be captured ([Raftery et al., 2005](#)). With this proposed PHL model, the prediction accuracy will be enhanced as the model uncertainty is reduced in the averaging process. Thus, it will be more accurate when used to calculate the column ductility in design or determine the localized region scale in analysis. Besides, as the model is a probabilistic one, it will provide not only the point estimation but also a prediction band, which can be further used in a wider field like uncertainty quantification. The latter feature may open a new perspective for structural design or analysis (e.g., reliability-based design), as one can directly produce the structural response with a confidence interval, rather than a deterministic value ([Feng et al., 2020c](#)).

The present paper aims at devising a probabilistic PHL model using a Bayesian-inference-based model averaging method. First, seven well-known parametric PHL models are introduced and evaluated using a database of 133 PHL tests. Then, the theoretical background of model averaging is presented. This is followed by the introduction of the new probabilistic PHL model, which is subsequently validated using the test database. A summary and conclusions are presented in the end.

EVALUATION OF SEVEN EXISTING PARAMETRIC PHL MODELS

Model Introduction

A total of seven state-of-the-art empirical PHL models proposed by different researchers are selected in this study. The detailed expressions of these parametric models are summarized in Table 1, and a brief and chronological introduction for each model is presented in what follows.

Priestley and Park (1987) proposed an empirical PHL model for RC bridge columns. The model was developed based on the test results of 20 bridge columns and only considers the influences of column height L and longitudinal bar diameter d_b on PHL. The model is simple and easy to use; however, the parameters considered are presently deemed insufficient.

Paulay and Priestley (1992) improved the model by Priestley and Park by adding a more detailed description of the strain penetration effect by including the yield strength f_y of the longitudinal bars in the second term.

Panagiotakos and Fardis (2001) developed a new model using an extensive database of 1012 tests. This more elaborate model accounts for the shear span-to-depth ratio $\frac{L}{H}$. Moreover, it refines the strain penetration effect by introducing a coefficient α_{sl} that represents rebar pull-out phenomena.

Lu et al. (2005) updated the empirical parameters of the original Priestley and Park model using a larger database of 53 tests. This naturally led to a more reliable model. Nevertheless, the too-simple functional form remained the same.

Similarly, Berry et al. (2008) extended Paulay and Priestley's original model by updating the empirical parameters with a larger database (37 tests), and introduced a new influence factor—namely, the bond strength between concrete and longitudinal bars, which is represented by the term $\sqrt{f'_c}$.

Bae and Bayrak (2008) performed an analytical derivation based on four large-scale RC column tests and identified three key factors—namely, the ratio of applied axial force to the nominal axial load capacity P/P_0 , the reinforcing ratio A_s/A_g , and the shear span-to-depth ratio L/H . They then established a new parametric model featuring these parameters.

Ning and Li (2015) examined the main PHL mechanisms and devised up a model by considering the strain penetration, tension shift, and moment gradient effects. They also used four random parameters in the model to reflect model uncertainty. Data from 133 PHL tests were used to identify and approximate the distributions of these parameters.

Performance Evaluation

The database compiled by Ning and Li (2015) is directly adopted here to evaluate the performance of the seven existing PHL models, as they have rigorously analyzed the required sample size for the development of their own PHL model. The database contains 133 PHL measurements from column tests, and each sample has 12 characteristic parameters—namely, the column height L , the section height H and width B , the concrete compressive strength f'_c , the longitudinal bar diameter d_b and yield strength f_y , the longitudinal reinforcing ratio ρ_l , the stirrup spacing s and strength f_{yv} , the stirrup reinforcing ratio ρ_v , the axial load ratio P/P_0 , and the loading type (monotonic as 1 and cyclic as 2). Details of this database is provided in Appendix I.

Fig. 1 displays the comparisons of measured and predicted PHL ratios (L_p/H). As seen, all models exhibit significant discrepancies with distinct regions of bias against the test data. For example, the model by Paulay and Priestley (1992) overestimates the PHL for $L_p/H < \sim 1.5$

TABLE 1. Summary of the existing parametric PHL models

Literatures	Equations
Priestley and Park (1987)	$L_p = 0.08L + 6d_b$
Paulay and Priestley (1992)	$L_p = 0.08L + 0.022f_y d_b$
Panagiotakos and Fardis (2001)	$L_p = 0.12\frac{L}{H} + 0.014\alpha_{sl}f_y d_b$
Lu et al. (2005)	$L_p = 0.077L + 8.16d_b$
Berry et al. (2008)	$L_p = 0.0375L + 0.12f_y d_b / \sqrt{f'_c}$
Bae and Bayrak (2008)	$L_p = \left(0.3\frac{P}{P_0} + 3\frac{A_s}{A_g} - 0.1\right)L + 0.25H \geq 0.25H$
Ning and Li (2015)	$L_p = \left(0.042 + 0.072\frac{P}{P_0}\right)L + 0.298H + 6.407d_b$

Note: L_p is the PHL; L is the column height; H is the section height; d_b , f_y and A_s are the diameter, yielding stress and area of the longitudinal bar, respectively; f'_c is the concrete compressive strength; α_{sl} is the coefficient for slip which equals 1 if considering bar pull-out effect or 0 if not; P and P_0 are the applied axial load and the axial load capacity, respectively; A_g is the gross area of the section.

while models by Panagiotakos and Fardis (2001); Berry et al. (2008); Bae and Bayrak (2008) underestimate the PHL.

For a more detailed examination, the ratios between tested and predicted PHLs for the 133 samples are also presented in Fig. 2. The mean value, standard derivation (St.D.), maximum, minimum, as well as coefficient of variation (C.O.V.) of the same ratios are summarized in Table 2. From this figure and table, it can be observed that the models by Paulay and Priestley (1992); Lu et al. (2005); Ning and Li (2015) achieve relatively better performances in terms of mean prediction accuracy. The results by Panagiotakos and Fardis (2001); Bae and Bayrak (2008) are relatively poor, on the other hand, whose mean test-to-prediction ratios are ~ 2.0 . Moreover, the C.O.V.s of measured-to-predicted PHL ratios for all seven models are all above 30%. The models by Ning and Li (2015) and Bae and Bayrak (2008) displays the best and the worst performances, whose C.O.V.s are 30.9% and 58.1%, respectively.

In summary, it can be concluded from this performance evaluation that the existing models' results bear significant prediction errors with large scatter and apparent bias. The reason may be that the plastic hinge mechanism is very complicated, including several influential effects, e.g., moment gradient, tension shift, and strain penetration, etc. However, all these mechanisms can hardly be included in one unified model that is easy and applicable in practical engineering. Thus, great simplification and assumptions are introduced when deriving these empirical models. Also, these models generally select a limited number of variables (around five), but there is no consensus that which kind of variables will govern the plastic hinge formation. All of the above factors lead to a large discrepancy when using the empirical models predicting the PHL, which can also be denoted as the model uncertainty.

MODEL AVERAGING METHOD BASED ON BAYESIAN INFERENCE

Fig. 1. Scatter diagrams of tested and predicted PHL ratios

TABLE 2. Summary of the results by existing parametric PHL models

Reference	Measured-to-predicted PHL ratios				
	Mean	St.D.	Maximum	Minimum	C.O.V. (%)
Priestley and Park (1987)	1.28	0.44	2.54	0.36	34.2
Paulay and Priestley (1992)	0.95	0.32	1.99	0.27	33.5
Panagiotakos and Fardis (2001)	2.63	0.87	5.75	0.77	33.1
Lu et al. (2005)	1.13	0.37	2.25	0.33	32.8
Berry et al. (2008)	1.64	0.67	3.94	0.41	40.8
Bae and Bayrak (2008)	1.79	1.04	4.80	0.51	58.1
Ning and Li (2015)	0.98	0.30	2.02	0.30	30.9

Model Averaging

Given the observations of bias and scatter in existing models, it appears possible to develop a new and improved model for predicting PHL through Bayesian model-averaging. The requisite mathematical background is briefly introduced herein.

For a regression problem with inputs $\mathbf{x}$ and outputs $\mathbf{y}$, various prediction models $\mathcal{M}$ with different performance can be combined through numerous techniques. These models may be of the same type with different parameter values (e.g., subsets in linear regression) or different models constructed different approaches (e.g., combinations of linear regression, neural networks, etc.). In a standard procedure, a “best” single model that bears the lowest training error among all competing models would be selected for future predictions. However, other competing models might provide more plausible estimations than the overall best single model for different ranges of input parameters (Feng et al., 2020b, 2021). The typical approach for making predictions using a single model may also lead to bias.

The concept of model averaging that will be pursued here will be to combine the individual prediction models into a super model, $\mathbb{M}$, through weighted averaging. Through this approach, the prediction result is not merely counting on a single “best” model, but on an ensemble of all

(a) Priestley and Park (1987)

(b) Paulay and Priestley (1992)

(c) Panagiotakos and Fardis (2001)

(d) Lu et al. (2005)

(e) Berry et al. (2008)

(f) Bae and Bayrak (2008)

prediction models (Friedman et al., 2001; Feng et al., 2020a). The process of model averaging can be summarized as

$$\mathbb{M}(\mathbf{x}) = \sum_{i=1}^n w_i \cdot \mathcal{M}_i(\mathbf{x}) \quad (1)$$

where $\mathcal{M}_i$ are the available single models with $i = \{1, 2, \dots, n\}$, and the corresponding weights are w_i , which satisfy

$$\sum_{i=1}^n w_i = 1. \quad (2)$$

Denoting the model and weight vectors respectively as

$$\begin{aligned} \mathbf{M} &= [\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_n] \\ \mathbf{w} &= [w_1, w_2, \dots, w_n], \end{aligned} \quad (3)$$

the final model can be expressed as

$$\mathbb{M}(\mathbf{w}, \mathbf{x}) = \mathbf{w}^T \cdot \mathbf{M}(\mathbf{x}). \quad (4)$$

The weights can then be determined per the following equation,

$$\mathbf{w} = \arg \min_{\mathbf{w}} \left[\mathbf{y} - \mathbf{w}^T \mathbf{M}(\mathbf{x}) \right]^2. \quad (5)$$

Model averaging can be treated as a procedure for the models to complement each other, which is advantageous as component models may have their own merits and deficiencies relative to each other. Moreover, it can be easily shown that the averaged model will produce a lower-bound value for the objective function in Eq. (5), as any single model $\mathcal{M}_{i=j}$ is merely a special case of model averaging with $w_{i=j} = 1$, $w_{i \neq j} = 0$. To wit,

$$\left[\mathbf{y} - \mathbf{w}^T \mathbf{M}(\mathbf{x}) \right]^2 \leq \left[\mathbf{y} - w_{i=j} \mathcal{M}_{i=j}(\mathbf{x}) \right]^2 = \left[\mathbf{y} - \mathcal{M}_j(\mathbf{x}) \right]^2. \quad (6)$$

As the core part of the model averaging method, several conventional deterministic approaches can be adapted to determine the optimal weights $\mathbf{w}$, such as the Gauss-Newton algorithm, genetic algorithm, etc. However, these deterministic approaches cannot provide a reasonable reflection of the model uncertainty, as they attempt to draw a unique solution while omitting other potential solutions with equal or nearly-the-same importance. In fact, all of these possibilities should be properly considered to reduce the model uncertainty and to improve performance (Raftery et al., 2005). As such, Bayesian inference, which is a probabilistic approach, is adopted herein to estimate the weights $\mathbf{w}$.

Bayesian Inference

Bayesian inference can be used to establish a model with unknown parameters (here $\mathbf{w}$) from data—a process that produces the statistical distributions of those parameters, instead of merely their optimal deterministic values (Beck, 2010). According to Bayes' theory, the posterior probability distribution function (PDF) of $\mathbf{w}$ can be derived as (Jang and Smyth, 2017; Shirzad et al., 2018):

$$P(\mathbf{w}|\mathbf{y}) = c^{-1} \cdot P(\mathbf{y}|\mathbf{w}) \cdot P(\mathbf{w}) = \frac{P(\mathbf{y}|\mathbf{w}) \cdot P(\mathbf{w})}{\int P(\mathbf{y}|\mathbf{w}) \cdot P(\mathbf{w}) \, d\mathbf{w}} \quad (7)$$

where $P(w)$ is the prior PDF, which provides the prior knowledge of the unknown parameters w ; $P(y|w)$ is the likelihood function reflecting the error level of the estimation results based on w ; y is the given dataset (outputs); $c = \int P(y|w) \cdot P(w) dw$ is the Bayesian evidence, which is a constant and can be calculated through a multi-dimensional integration; $P(w|y)$ is the posterior PDF of w representing the final results after fully consideration of the initial prior PDF and training dataset (Soyoz et al., 2010; Li et al., 2013).

Bayesian inference combines prior knowledge of target parameters with a training dataset to infer the posterior PDFs. However, a closed-form solution of the posterior PDF is typically not possible due to the difficulty in performing the multi-dimensional integration of the Bayesian evidence c in Eq. (7). Considering that the Bayesian evidence c is a constant, Eq. (7) can be reformulated as

$$P(w|y) \propto P(y|w) \cdot P(w), \quad (8)$$

which indicates that the posterior PDF is positively proportional to the likelihood function, and thus the non-normalized posterior PDF can be easily obtained.

To this end, the Markov Chain Monte Carlo (MCMC) method, which is a stochastic simulation technique, has been widely utilized in Bayesian inference on the basis of Eq. (8) (Goller et al., 2012). MCMC uses a Markov Chain approach to numerically draw random samples that conform to a target posterior PDF based on a balance condition. The generated samples are then gathered together to estimate the target posterior PDF (Zhang et al., 2013). In this procedure, the most common sampling strategy is the Metropolis-Hastings (MH) algorithm (Hastings, 1970), and the crucial detailed balance condition of a Markov chain is guaranteed by an “acceptance probability” α , which is given by

$$\alpha = \min \left[1, \frac{P(y|w_c)}{P(y|w_i)} \right] \quad (9)$$

where w_c is a randomly generated candidate weight sample vector; w_i is the i -th accepted weight sample vector. Considering the relationship displayed in Eqs. (7) - (9), the acceptance probability can be treated as

$$\alpha = \min \left[1, \frac{\text{likelihood}(w_c)}{\text{likelihood}(w_i)} \right] = \min \left[1, \frac{P(w_c|y)}{P(w_i|y)} \right] = \min \left[1, \frac{\text{posterior}(w_c)}{\text{posterior}(w_i)} \right] \quad (10)$$

Through this transformation, it can be seen that the specific value of Bayesian evidence c is no longer required in MCMC. The sample sequence drawn in MCMC would eventually conform to the posterior PDF $P(w|y)$. The pseudocode for the MCMC method featuring the MH algorithm is presented in Algorithm 1.

In the MH algorithm, first, the initial samples w_0 are generated through a proposal PDF $q(w)$ defined by the user, which represents the prior PDF of target parameters and should be convenient for drawing samples (e.g., a uniform or a normal distribution). Next, for each iteration, a candidate sample w_c is drawn from $q(w_c|w_i)$, which means that the next potential sample is drawn only by considering the last sample set. This operation aligns with the definition of a Markov chain in that a Markov chain describes a sequence of possible variable states in a system where the probability of future states merely depends on the current state and the transient probability between the states is fixed. With this property, the system’s statistical properties become predictable and will conform to the target PDF, even though the system is stochastic. The acceptance or dismissal of the

Input: test data $\mathbf{y}$, proposal PDF $q(\mathbf{w})$, sample number M , likelihood function $P(\mathbf{y}|\mathbf{w})$

Output: Samples $[\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_M]$ that conform to the posterior PDF $P(\mathbf{w}|\mathbf{y})$

Initialization: $i = 1, \mathbf{w}_0$;

while $i \leq M$ **do**

 Sample $\mathbf{w}_c \sim q(\mathbf{w}_c|\mathbf{w}_i)$;

 Sample $u \sim U[0, 1]$;

$\alpha = \min \left[1, \frac{P(\mathbf{y}|\mathbf{w}_c)}{P(\mathbf{y}|\mathbf{w}_i)} \right]$;

if $\alpha > u$ **then**

$\mathbf{w}_i = \mathbf{w}_c$;

else

$\mathbf{w}_i = \mathbf{w}_{i-1}$;

end

$i = i + 1$;

end

Algorithm 1: MCMC method with the MH algorithm

candidate sample $\mathbf{w}_c$ is determined by a comparison between the acceptance probability α and a random number $u \in [0, 1]$. The sample with a higher likelihood—i.e., a sample that yields a lower prediction error—would more likely be accepted because its acceptance probability is larger. After several iterations, the sample sequence drawn from this Markov Chain would converge to a target PDF, which is the posterior PDF of $\mathbf{w}$ (Lam et al., 2018; Guan et al., 2011).

In MCMC, the likelihood of each candidate sample $P(\mathbf{y}|\mathbf{w}_c)$ is estimated using the bias between the predictions $\mathbb{M}(\mathbf{w}_c, \mathbf{x})$ and the actual values $\mathbf{y}(\mathbf{x})$, which can be expressed as

$$P(\mathbf{y}|\mathbf{w}_c) = C \cdot \exp \left[\frac{\mathbf{y}(\mathbf{x}) - \mathbb{M}(\mathbf{w}_c, \mathbf{x})}{2\sigma^2} \right] \quad (11)$$

where σ is the assumed St.D. of the likelihood function and $C = (2\pi\sigma)^{-0.5}$ is a constant normalization coefficient.

PROBABILISTIC PARAMETRIC PHL MODEL

In order to develop the probabilistic PHL model using the Bayesian inference, a database of PHL measurements is compiled, which can be denoted as $\mathcal{D} = \{(\mathbf{X}_1, Y_1), (\mathbf{X}_2, Y_2), \dots, (\mathbf{X}_N, Y_N)\}$, where $\mathbf{X}_i$ are the input variables, such as the column length, section dimensions, material properties, etc.; Y_i is the output variable (here, the PHL measurements); $N = 133$ is the total number of samples. Meanwhile, there are seven PHL prediction models with varying accuracies and variance. The predicted PHL values by each model can be compactly expressed with the vector $\mathcal{L}_p = \{L_p^1, L_p^2, \dots, L_p^n\}$, where $n = 7$ is the number of models considered in the present study (see, Table 1). The objective is determining the weights $\mathbf{w} = \{w_1, w_2, \dots, w_n\}$ for the seven models and obtaining a more precise prediction model L_p^{avg} , say,

$$L_p^{\text{avg}}(\mathbf{w}, \mathbf{X}) = \mathbf{w}^T \cdot \mathcal{L}_p(\mathbf{X}) = \sum_{j=1}^{n=7} w_j \cdot L_p^j(\mathbf{X}) \quad (12)$$

Moreover, the developed model L_p^{avg} is not just a deterministic model like all L_p^j considered here, but a probabilistic one. The MCMC method captures both the optimal value of each weight w_i , but also produces the PDF of each weight. Consequently, the model uncertainty is inherently represented in the proposed model, as shown in

$$\text{PDF}(L_p^{\text{avg}}) = \sum_{j=1}^{n=7} \text{PDF}(w_j) \cdot L_p^j(\mathbf{X}). \quad (13)$$

The entire procedure of training the proposed probabilistic parametric PHL model is illustrated in Fig. 3. The initial values of the seven weights w_j are all set equal to each other (1/7), which means that at the beginning, the prior knowledge of these models' contributions is that they are even. Then, a proposal PDF $q(\mathbf{w})$, which is assumed as a normal distribution, is used to select the candidate weight samples $\mathbf{w}_c$. Finally, a series of samples that are in accordance with the prior distributions of these weights $\mathbf{w}_{k=1,2,\dots,M}$ are drawn through the MCMC+MH algorithm. The likelihood of each sample point is calculated using the root-mean-square error with the training dataset $\mathcal{D}$, namely,

$$\begin{aligned} P[\mathcal{D}|L_p^{\text{avg}}(\mathbf{w}_c, \mathbf{X})] &= \prod_{i=1}^N P[Y_i|L_p^{\text{avg}}(\mathbf{w}_c, \mathbf{X}_i)] \\ &= \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma_i^2}} \cdot \exp\left[-\frac{Y_i - L_p^{\text{avg}}(\mathbf{w}_c, \mathbf{X}_i)}{2\sigma_i^2}\right] \\ &= C \cdot \exp\left[\sum_{i=1}^N \frac{Y_i - L_p^{\text{avg}}(\mathbf{w}_c, \mathbf{X}_i)}{2\sigma_i^2}\right] \end{aligned} \quad (14)$$

where $C = \prod_{i=1}^N (2\pi\sigma_i)^{-0.5}$ is a constant normalization coefficient.

The prior distribution of these weights represent the relative contribution of L_p^j in L_p^{avg} . The total MCMC sample size is set as $M = 5 \times 10^4$, and the beginning 5000 samples within the “burn-in period” are omitted. The samples after convergence, i.e., $\mathbf{w}_{k=1,2,\dots,M}$, are collected to calculate the distribution characteristics, including mean and St.D. of the weights $\mathbf{w}$. The sampling results and the obtained distribution of each weight are displayed in Figs. 4 and 5.

With the obtained distribution of model weights $\mathbf{w}$, a new prediction distribution of PHL result can be made through Eq. (12). The specific distribution characteristics of then seven weights are listed in Table. 3. It can be seen the mean values of all weight distributions are greatly varied, with some even having negative values. This represents the different deviation direction of each model L_p^j within the average model L_p^{avg} . The performance of this converged probabilistic parametric PHL model compared with current models L_p^j are be presented next.

MODEL PERFORMANCE

To evaluate the model performance, the collected PHL databased is also used here. Since the model is a probabilistic one, both the mean and variance values could be obtained.

Fig. 3. The training procedure for the probabilistic parametric PHL model

TABLE 3. Distribution characteristics of the model weights

Characteristic	Model weights w						
	w_1	w_2	w_3	w_4	w_5	w_6	w_7
Mean	-1.77	-0.39	1.02	1.77	-1.21	-0.02	1.60
St.D.	0.23	0.12	0.08	0.28	0.09	0.04	0.11

Mean Prediction Results and Comparisons

If the mean values of the seven weights in Table 3 are substituted to the model Eq. (1), one obtains the mean prediction of the PHL model. Table 4 summarizes the accuracy of prediction results by the seven existing models and this mean prediction by the probabilistic model. The root mean square error (RMSE)—a more general measure for prediction accuracy—is also used to evaluate and compare the performance of the models, which is given by

$$\text{RMSE} = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^N (Y_i - \hat{Y}_i)^2} \quad (15)$$

where Y_i represents the actual value for each sample i ; $\hat{Y}_i$ is the predicted value by each model.

It can be seen from Table 4 that the current model L_p^{avg} attains a better precision than all the other models, with its RMSE=0.28 being the lowest. Among the existing models, the most eminent

Fig. 4. Sampling results of the model weights

one would be that by Ning and Li (2015) with a corresponding RMSE=0.36. Using these two RMSE values, the current model's performance improvement on L_p^{avg} over the model by Ning and Li (2015) is thus 22.2%, indicating the significantly high accuracy of the new model.

The St.D. and C.O.V. values of the measured-to-predicted PHL ratios by each model are also displayed in Table. 4. Among all models, the proposed model L_p^{avg} is the best with Mean=1.0, St.D.=0.26 and C.O.V.=25.9%, which illustrates that this model's prediction results not only are the most precise but bear the smallest variance against data. Similarly, it displays a significant improvement over the model by Ning and Li (2015), which was the best prediction model. The improvement rates on mean, St.D. and C.O.V. are +2.0%, -13.3% and -16.2%, respectively. In Fig. 6, the prediction results by the best existing model and the proposed one are presented in detail to show the improvement more directly. Overall, the proposed model L_p^{avg} 's performance is

Fig. 5. PDFs of the model weights

statistically better than all current seven models.

As an important factor in general Bayesian inference problems, the prior assumption of the model weights should have an influence on the results. Consequently, a further test was carried out to analyze the influence of initial model weights on the final prediction results. Instead of setting uniform prior weights for 7 models (hereafter denoted as Case A), two different prior assumptions are studied. First, each model's prior weight is randomly generated, and 10 independent rounds were carried out for analysis (denoted as Case B). Besides that, one more test was conducted whose prior weight distributions follow each empirical model's accuracy performance to achieve informative priors (denoted as Case C). More specifically, the initial weights of 7 models w_i were

Fig. 6. Prediction results by the best existing model and proposed model

TABLE 4. Prediction results by the seven existing models and proposed probabilistic model

Measures	Existing models							L_p^{avg}	Improvement over best model
	L_p^1	L_p^2	L_p^3	L_p^4	L_p^5	L_p^6	L_p^7		
RMSE	0.44	0.43	0.76	0.40	0.57	0.52	0.36	0.28	−22.2%
Mean	1.28	0.95	2.63	1.13	1.64	1.79	0.98	1.00	+2.0%
St.D.	0.44	0.32	0.87	0.37	0.67	1.03	0.30	0.26	−13.3%
C.O.V. (%)	34.21	33.50	33.05	32.78	40.80	58.09	30.93	25.92	−16.2%

designed according to the RMSE of each model, say,

$$w_i = \frac{1/\text{RMSE}_i}{\sum_{i=1}^7 1/\text{RMSE}_i} \quad (16)$$

The prior weights for the three cases are listed in Appendix II. Based on these prior weights, corresponding probabilistic PHL models were developed through the procedure illustrated in Fig. 3. For brevity, the obtained distribution characteristics of the posterior model weights are also given in Appendix II, and the final prediction performances of these models on PHL are listed in Table 5. Note the result for Case B is the average of the 10 rounds. It can be seen that in these tests—although the initial model weights vary dramatically and so as some of their final weights and distributions—the prediction accuracies of all developed models are nearly identical with tiny fluctuation in the coefficient of variance. This result revealed that the performance of generated PHL prediction model would not be impacted by the prior assumption of the model weights and also proves the correctness of the Markov chain utilized in the training of this model.

Probabilistic Prediction Results

Unlike current PHL prediction models, the proposed model is probabilistic, as it is based on Bayes' theory, which produces a PDF of PHL rather than a single value. Fig. 7 displays the probabilistic results by the proposed model L_p^{avg} . In the figure, the red dots are the actual PHL value of the

TABLE 5. Prediction results by models with different prior model weights

Case	RMSE	Mean	St.D.	C.O.V(%)
A	0.28	1.00	0.26	25.92
B	0.28	1.00	0.26	26.01
C	0.28	1.00	0.26	26.08

samples, the blue line is the mean prediction for the samples by the proposed model, while the shaded area is the confidence interval in terms of $\text{Mean} \pm \text{St.D}$ by the model. From Fig. 7, it can be observed that more than 95% of the true PHL test values lie within the high confidence region of the proposed model, and even the few outliers are still located near the boundaries of the high confidence region.

For a more detailed appraisal of the proposed probabilistic predictions, four typical columns with various section heights are selected—namely, the specimens H60-10.5-C0-1-30, H90-7.5-C2-2-30 by Barrera et al. (2011), HSC-S1-3 by Yang et al. (2013), and Unit 1 by Tanaka (1990). The PHL prediction results for these four column tests obtained by each model are shown in Fig. 8. The results predicted by the proposed model L_p^{avg} can be approximately described by a normal function. The PHL results predicted by existing models can simply be treated as the quantiles of the PDF predicted by the proposed model L_p^{avg} with various probability densities. This confirms the findings in the mean prediction results presented earlier that the probabilistic prediction model provides more robust results.

Fig. 7. Probabilistic results by the proposed model

CONCLUSIONS

This paper presented a novel probabilistic model for predicting the PHL of RC columns devised using a Bayesian approach. Seven current existing PHL models are evaluated first through a database containing 133 PHL measurements. Their performances were found to range from fair to poor in terms of prediction accuracy and generally poor in terms of variance against test data. To attain more accurate and robust results, these models were combined through weighted

Fig. 8. Predicted probabilistic PHL results for four typical columns

averaging. The posterior PDFs of the weights were determined through Bayesian inference applied to the test data. The advantage of the resulting model was manifold: (i) it consistently provided more accurate results with least variance; (ii) it offered PDFs of weights, thus offering a smooth pathway to modeling uncertainty quantification; and finally, (iii) the method by which the model was constructed is amenable to assimilate more data and other additional models of different types (parametric or otherwise). It should also be noted that the model will have satisfactory performance when the input variables fall in the range of the database, as the model is trained using the 133 samples. In order to improve the generalization ability, more data outside the hypothesis of the 133 samples are required.

APPENDIX I. THE PHL DATABASE SUMMARY

The PHL database is collected by Ning and Li (2015), which includes 133 PHL samples from 14 published papers. The full data is provided in Ning and Li (2015). To save space, here only the statistical properties of the database are summarized in Table 6.

TABLE 6. Statistical properties of parameters included in PHL database

Parameter	Unit	Min.	Max.	Mean	St.D.	Type
L	mm	630.00	3050.00	1712.62	407.88	Input
B	mm	150.00	610.00	273.87	107.15	Input
H	mm	140.00	610.00	288.31	107.48	Input
f'_c	MPa	24.80	175.00	70.27	30.89	Input
d_b	mm	8.00	32.00	16.39	5.27	Input
f_y	MPa	339.00	572.00	498.07	44.65	Input
ρ_l	%	0.82	6.10	2.31	0.91	Input
s	mm	50.00	220.00	93.59	28.92	Input
f_{yv}	MPa	325.00	952.00	530.92	158.69	Input
ρ_v	%	0.32	6.72	2.11	1.26	Input
P/P_0	-	0.00	0.77	0.35	0.19	Input
Loading Type	-	1.00	2.00	1.64	0.48	Input
PHL ratio	-	0.35	2.43	1.11	0.39	Output

Note: Min.=Minimum; Max.=Maximum; St.D.=Standard derivation.

APPENDIX II. PRIOR AND POSTERIOR MODEL WEIGHTS

The prior weights for the 7 models for Case A, B and C are listed in Table 7, and the distribution parameters for the posterior weights for Case A, B and C are listed in Table 8.

TABLE 7. Prior weights for Case A, B and C

Case	No.	w_1	w_2	w_3	w_4	w_5	w_6	w_7
A	-	0.1428	0.1428	0.1428	0.1428	0.1428	0.1428	0.1428
	1	0.1467	0.1940	0.2436	0.1195	0.2278	0.0614	0.0069
	2	0.1748	0.1020	0.1414	0.3183	0.0330	0.0743	0.1562
	3	0.0340	0.0085	0.0966	0.2872	0.1239	0.2818	0.1679
B	4	0.1628	0.1149	0.1973	0.1067	0.1943	0.0332	0.1907
	5	0.0231	0.2177	0.1697	0.0678	0.2860	0.0377	0.1980
	6	0.1881	0.0022	0.1932	0.2415	0.0898	0.2849	0.0002
	7	0.2319	0.1645	0.0741	0.1537	0.0848	0.2034	0.0876
	8	0.0596	0.0915	0.2796	0.1818	0.1781	0.0031	0.2062
	9	0.0151	0.1776	0.2677	0.2222	0.0340	0.0024	0.281
	10	0.2797	0.0208	0.2331	0.1413	0.1225	0.0897	0.1129
C	-	0.1528	0.1554	0.0883	0.1682	0.1176	0.1300	0.1877

DATA AVAILABILITY STATEMENT

All data, models, or code generated or used during the study are available from the corresponding author by request.

ACKNOWLEDGEMENTS

The authors would like to acknowledge financial support from the National Natural Science Foundation of China (Grant No. 52078119), which enabled the first and second authors to spend terms as Visiting Scholars at UCLA.

REFERENCES

- Bae, S. and Bayrak, O. (2008). "Plastic hinge length of reinforced concrete columns." *ACI Structural Journal*, 105(3), 290–300.
- Barrera, A. C., Bonet, J., Romero, M., and Miguel, P. (2011). "Experimental tests of slender reinforced concrete columns under combined axial load and lateral force." *Engineering Structures*, 33(12), 3676–3689.
- Beck, J. L. (2010). "Bayesian system identification based on probability logic." *Structural Control and Health Monitoring*, 17(7), 825–847.
- Berry, M. P., Lehman, D. E., and Lowes, L. N. (2008). "Lumped-plasticity models for performance simulation of bridge columns." *ACI Structural Journal*, 105(3), 270–279.

TABLE 8. Distribution parameters for posterior weights for Case A, B and C

Case	No.		w_1	w_2	w_3	w_4	w_5	w_6	w_7
A	-	Mean	-1.77	-0.39	1.02	1.77	-1.21	-0.02	1.60
		St.D.	0.23	0.12	0.08	0.28	0.09	0.04	0.11
B	1	Mean	-1.88	-0.42	1.02	1.96	-1.22	-0.01	1.54
		St.D.	0.36	0.18	0.09	0.54	0.07	0.04	0.13
	2	Mean	-1.83	-0.34	1.04	1.83	-1.23	-0.01	1.54
		St.D.	0.56	0.23	0.07	0.87	0.09	0.05	0.19
	3	Mean	-1.99	-0.40	1.03	2.05	-1.21	-0.01	1.52
		St.D.	0.62	0.26	0.08	0.97	0.09	0.05	0.19
	4	Mean	-1.93	-0.40	1.03	1.94	-1.20	-0.02	1.56
		St.D.	0.30	0.17	0.08	0.43	0.09	0.04	0.12
	5	Mean	-1.72	-0.35	1.02	1.70	-1.22	-0.02	1.59
		St.D.	0.43	0.18	0.08	0.58	0.09	0.04	0.14
	6	Mean	-1.82	-0.35	1.04	1.83	-1.23	-0.01	1.54
		St.D.	0.53	0.26	0.09	0.87	0.08	0.05	0.18
	7	Mean	-1.81	-0.35	1.03	1.82	-1.22	-0.01	1.54
		St.D.	0.56	0.26	0.10	0.84	0.09	0.05	0.16
	8	Mean	-2.13	-0.51	1.00	2.35	-1.19	0.00	1.48
		St.D.	0.44	0.19	0.08	0.63	0.10	0.04	0.15
	9	Mean	-1.98	-0.44	1.02	2.12	-1.22	0.00	1.50
		St.D.	0.39	0.16	0.08	0.57	0.08	0.04	0.15
	10	Mean	-1.93	-0.47	1.00	2.09	-1.20	0.00	1.53
		St.D.	0.24	0.15	0.08	0.35	0.08	0.04	0.11
C	-	Mean	-1.71	-0.38	1.01	1.72	-1.22	-0.01	1.58
		St.D.	0.34	0.18	0.08	0.52	0.09	0.04	0.14

Dormann, C. F., Calabrese, J. M., Guillerá-Arroita, G., Matechou, E., Bahn, V., Bartoń, K., Beale, C. M., Ciuti, S., Elith, J., Gerstner, K., et al. (2018). “Model averaging in ecology: a review of bayesian, information-theoretic, and tactical approaches for predictive inference.” *Ecological Monographs*, 88(4), 485–504.

Elmenshaw, A., Brown, T., and El-Metwally, S. (2012). “Plastic hinge length considering shear reversal in reinforced concrete elements.” *Journal of Earthquake Engineering*, 16(2), 188–210.

Feng, D.-C., Cetiner, B., Reza Azadi Kakavand, M., and Taciroglu, E. (2021). “Data-driven approach to predict the plastic hinge length of reinforced concrete columns and its application.” *Journal of Structural Engineering*, 147(2), 0402332.

Feng, D.-C., Liu, Z.-T., Wang, X.-D., Chen, Y., Chang, J.-Q., Wei, D.-F., and Jiang, Z.-M. (2020a). “Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach.” *Construction and Building Materials*, 230, 117000.

Feng, D.-C., Liu, Z.-T., Wang, X.-D., Jiang, Z.-M., and Liang, S.-X. (2020b). “Failure mode

- classification and bearing capacity prediction for reinforced concrete columns based on ensemble machine learning algorithm.” *Advanced Engineering Informatics*, 45, 101126.
- Feng, D.-C. and Ren, X.-D. (2017). “Enriched force-based frame element with evolutionary plastic hinge.” *Journal of Structural Engineering*, 143(10), 06017005.
- Feng, D. C., Ren, X. D., and Li, J. (2016). “Implicit gradient delocalization method for force-based frame element.” *Journal of Structural Engineering*, 142(2), 04015122.
- Feng, D.-C., Wu, G., and Ning, C.-L. (2019). “A regularized force-based timoshenko fiber element including flexure-shear interaction for cyclic analysis of rc structures.” *International Journal of Mechanical Sciences*, 160, 59–74.
- Feng, D.-C., Xie, S.-C., Xu, J., and Qian, K. (2020c). “Robustness quantification of reinforced concrete structures subjected to progressive collapse via the probability density evolution method.” *Engineering Structures*, 202, 109877.
- Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The elements of statistical learning*. Springer series in statistics, New York, NY, USA.
- Goller, B., Beck, J. L., and Schuëller, G. I. (2012). “Evidence-based identification of weighting factors in bayesian model updating using modal data.” *Journal of Engineering Mechanics*, 138(5), 430–440.
- Guan, X., Jha, R., and Liu, Y. (2011). “Model selection, updating, and averaging for probabilistic fatigue damage prognosis.” *Structural Safety*, 33(3), 242–249.
- Hastings, W. K. (1970). “Monte carlo sampling methods using markov chains and their applications.” *Biometrika*, 57(1), 97–109.
- Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (1999). “Bayesian model averaging: a tutorial.” *Statistical Science*, 382–401.
- Jang, J. and Smyth, A. (2017). “Bayesian model updating of a full-scale finite element model with sensitivity-based clustering.” *Structural Control and Health Monitoring*, 24(11), e2004.
- Kakavand, M. R. A. and Allahvirdizadeh, R. (2019). “Enhanced empirical models for predicting the drift capacity of less ductile rc columns with flexural, shear, or axial failure modes.” *Frontiers of Structural and Civil Engineering*, 13(5), 1251–1270.
- Lam, H.-F., Yang, J.-H., and Au, S.-K. (2018). “Markov chain monte carlo-based bayesian method for structural model updating and damage detection.” *Structural Control and Health Monitoring*, 25(4), e2140.
- Li, H., Mao, C. X., and Ou, J. P. (2013). “Identification of hysteretic dynamic systems by using hybrid extended kalman filter and wavelet multiresolution analysis with limited observation.” *Journal of Engineering Mechanics*, 139(5), 547–558.
- Lu, Y., Gu, X., and Guan, J. (2005). “Probabilistic drift limits and performance evaluation of reinforced concrete columns.” *Journal of Structural Engineering*, 131(6), 966–978.

- Ning, C.-L. and Li, B. (2015). “Probabilistic approach for estimating plastic hinge length of reinforced concrete columns.” *Journal of Structural Engineering*, 142(3), 04015164.
- Ning, C.-L. and Li, B. (2017). “Probabilistic development of shear strength model for reinforced concrete squat walls.” *Earthquake Engineering & Structural Dynamics*, 46(6), 877–897.
- Ning, C.-L. and Li, B. (2018). “Analytical probabilistic model for shear strength prediction of reinforced concrete beams without shear reinforcement.” *Advances in Structural Engineering*, 21(2), 171–184.
- Panagiotakos, T. B. and Fardis, M. N. (2001). “Deformations of reinforced concrete members at yielding and ultimate.” *ACI Structural Journal*, 98(2), 135–148.
- Park, R., Priestley, M., and Gill, W. D. (1982). “Ductility of square-confined concrete columns.” *Journal of the Structural Division*, 108(4), 929–950.
- Paulay, T. and Priestley, M. N. (1992). *Seismic design of reinforced concrete and masonry buildings*. Wiley New York.
- Priestley, M. and Park, R. (1987). “Strength and ductility of concrete bridge columns under seismic loading.” *ACI Structural Journal*, 84(1), 61–76.
- Raftery, A. E., Gneiting, T., Balabdaoui, F., and Polakowski, M. (2005). “Using bayesian model averaging to calibrate forecast ensembles.” *Monthly Weather Review*, 133(5), 1155–1174.
- Scott, M. and Hamutçuoğlu, O. (2008). “Numerically consistent regularization of force-based frame elements.” *International journal for numerical methods in engineering*, 76(10), 1612–1631.
- Sheikh, S. A. and Khoury, S. S. (1993). “Confined concrete columns with stubs.” *ACI Structural Journal*, 90(4), 414–414.
- Sheikh, S. A., Shah, D. V., and Khoury, S. S. (1994). “Confinement of high-strength concrete columns.” *ACI Structural Journal*, 91(1), 100–100.
- Shirzad, N., Mahsuli, M., Ghahari, F., and Taciroglu, E. (2018). “Bayesian identification of soil-foundation stiffness of building structures.” *Structural Control & Health Monitoring*, 25(3), e2090.
- Soyoz, S., Feng, M. Q., and Shinozuka, M. (2010). “Structural reliability estimation with vibration-based identified parameters.” *Journal of Engineering Mechanics*, 136(1), 100–106.
- Steel, M. F. (2017). “Model averaging and its use in economics.” *arXiv preprint arXiv:1709.08221*.
- Tanaka, H. (1990). “Effect of lateral confining reinforcement on the ductile behaviour of reinforced concrete columns.” Ph.D. thesis, University of Canterbury, Christchurch, New Zealand.
- Wasserman, L. (2000). “Bayesian model selection and model averaging.” *Journal of Mathematical Psychology*, 44(1), 92–107.

- 469 Yang, K., Shi, Q. X., and Zhao, J. H. (2013). “Plastic hinge length of high strength concrete
470 columns confined by high strength stirrups.” *Engineering Mechanics*, 30(2), 254–259.
- 471 Zhang, J., Wan, C., and Sato, T. (2013). “Advanced markov chain monte carlo approach for finite
472 element calibration under uncertainty.” *Computer-Aided Civil and Infrastructure Engineering*,
473 28(7), 522–530.