HOKKAIDO UNIVERSITY

Title	Neural Network with Variational Monte Carlo for Bayesian Inference and Response Estimation of Nonlinear Systems
Author(s)	Lin, Huian; Kitahara, Takeshi
Relation	Proceedings of the 10th International Symposium on Reliability Engineering and Risk Management (ISRERM 2026), June 28 - July 1, 2026, Sapporo, Japan
Citation	Proceedings of the 10th International Symposium on Reliability Engineering and Risk Management (ISRERM 2026)
Issue Date	2026-07-01
Doc URL	https://hdl.handle.net/2115/100426
Type	conference paper
File Information	C000016_1.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

10th International Symposium on Reliability Engineering and Risk Management (ISRERM 2026),
28 June-01 July 2026, Sapporo, Japan
Kitahara, T., Miyamori, Y. & Beer, M. (editors)

Neural Network with Variational Monte Carlo for Bayesian Inference and Response Estimation of Nonlinear Systems

Huian Lin[1], and Takeshi Kitahara[2*, 3]

[1]*Graduate School of Engineering, Kanto Gakuin University, Yokohama, Japan.*
E-mail: linhuian.lha@gmail.com
[2]*Department of Civil Engineering, Kanto Gakuin University, Yokohama, Japan.*
*E-mail: kitaharar@kanto-gakuin.ac.jp, *Corresponding Author*
[3]*Graduate School of Global Environmental Studies, Kyoto University, Kyoto, Japan.*
E-mail: kitahara.takeshi.8t@kyoto-u.ac.jp

In the current environment, where health diagnosis, seismic response assessment, and control design of bridges, high-rise buildings, and smart structures are becoming critical issues, methods for parameter identification and dynamic response estimation are needed for systems where highly nonlinear structural systems are subjected to highly nonstationary disturbances, such as earthquakes and wind. This study proposes an innovative computational framework for estimating structural parameters of seismically isolated bridges using neural network variational Monte Carlo (NN-VMC). Conventional parameter identification methods, such as transition Markov chain Monte Carlo, often suffer from low computational efficiency and convergence problems in high-dimensional parameter spaces. To address these issues, we developed a new NN-VMC approach that combines neural network approximation with the probabilistic inference of variational Monte Carlo. We applied the proposed framework to the estimation problem of key structural parameters of a two-degree-of-freedom seismically isolated bridge model, using seismic response data as observational input. Using numerical example, we validated the parameter-estimation accuracy and computational efficiency of the NN-VMC method. The proposed NN-VMC method is expected to contribute to real-time structural integrity monitoring, damage identification, and seismic risk assessment of critical infrastructure.

Keywords: Bayesian updating, Posterior distribution, Neural network, Variational inference.

1 Introduction

In today's complex society, solving engineering problems often involves numerous uncertainties. Quantifying these uncertainties requires high computational efficiency and accuracy. To address this problem, this study aims to use seismic response acceleration time history data to infer key bridge structural parameters using Bayesian theory and to further evaluate the seismic performance of bridges. Variational Inference is a probability density approximation method in machine learning that approximates difficult-to-calculate probability densities via an optimization process and is particularly used in Bayesian inference to approximate posterior distributions. In Bayesian inference, inferring unknown quantities requires calculation via posterior densities, Bayesian inference requires calculations based on posterior densities to estimate unknown quantities, but explicitly determining the posterior density of complex models is often difficult. Traditional methods like Markov Chain Monte Carlo (MCMC) approximate the posterior through sampling, but they incur high computational costs and struggle to handle large-scale data. In contrast, Variational Inference approximates the target density by solving an optimization problem, featuring faster computation speed and better scalability, making it suitable for large-scale data and complex models. Variational Inference and Variational Bayes are highly similar in their core ideas and mathematical principles, sharing the same optimization framework (e.g., minimizing the Kullback-Leibler divergence and maximizing the Evidence Lower Bound).
Therefore, this study aims to advance statistical research on Variational Inference for an application to time-series models. Additionally, it seeks to develop a method to address the problem of excessive computational load, enabling it to demonstrate excellent performance in both high-dimensional and large-scale data processing. The proposed NN-VMC method combines deep neural networks with Variational Inference: neural networks serve as the theoretical basis for parameterizing variational distributions and provide a complete algorithm implementation framework. This method cites and improves the reparameterization trick from Auto-Encoding Variational Bayes and integrates the principles of Stochastic Variational Inference. Based on time-series seismic observation acceleration data, it can approximate the posterior density that is difficult to calculate directly through an optimization process, thereby providing a new solution for parameter identification and uncertainty quantification of large-scale civil engineering structures.
This work builds on prior studies. Kitahara et al. (2022) applied Transitional Markov Chain Monte Carlo to estimate parameters of seismic-isolated bridges using a two-degree-of-freedom model. Blei et al. (2017) provided a comprehensive overview of Variational Inference, while Kingma and Welling (2013) introduced efficient inference via the reparameterization trick. Carleo et al. (2017) demonstrated the capability of neural networks to represent complex distributions. Based on these developments, this study advances variational inference for time-series structural

models and provides a scalable framework for parameter estimation and uncertainty quantification. The remainder of the paper introduces the methodology, presents experimental validation, and concludes with future research directions.

2 Related Work

2.1 Analysis Model

The analysis focused on an existing LRB seismic-isolated bridge, which was modeled as a 2DOF spring-mass system with fixed pier foundations, as shown in **Figure 1**. The structural specifications of the model are presented in **Table 1**, where both the design values and current values after aging are recorded. In the Bayesian update process, the ratio of the current values to the design values (listed in the rightmost column of the table) was used as the true value. The mass was treated as a fixed value and not included in the parameters to be updated. For the hysteretic restoring force characteristics, the seismic-isolation bearings were assigned to a perfectly elastoplastic bilinear model with zero secondary stiffness, and the RC (Reinforced Concrete) piers were modeled using a deteriorated Takeda model. Additionally, damping constants were set as Rayleigh damping: 0% for the seismic-isolation bearings and 2% for the RC piers. The generation of pseudo-seismic response observation data and the calculation of likelihoods for model parameter updates required the implementation of nonlinear time-history response analysis. For this analysis, the Newmark β method ($\gamma = 1/2$ and $\beta = 1/4$) was adopted, with a time step of 0.001 seconds. In this study, the parameters to be identified for the 2DOF system included, for the seismic-isolation bearings: initial stiffness (Kb1), post-yield stiffness (Kb2), and yield load (Qb), for the RC piers: initial stiffness (Kp) and yield load (Qp). These five parameters were identified as the key targets in this research.

Table 1. Structural specifications.

Structural Parameters		Design Value [A]	Current Value [B]	B/A
Superstructure	Mass Ms (ton)	604	604	—
Isolation rubber bearing	yield load Qb (MN)	1.12	0.56	0.5
	initial stiffness Kb1 (MN/m)	40.0	48.0	1.2
	post-yield stiffness Kb2 (MN/m)	6.00	7.20	1.2
RC pier	Mass Mp (ton)	346	346	—
	yield load Qp (MN)	3.37	2.02	0.6
	initial stiffness Kp (MN/m)	110.1	88.1	0.8

2.2 Pseudo-Seismic Observation Data

Since the seismic response measurements of the superstructure of the target bridge were used as observation data, a nonlinear time-history response analysis of the reference model (design value model) was conducted to simulate the observation data, and the acceleration time history of the superstructure mass point under the input ground motion was calculated. For Bayesian updating, the observation data used in this study were the pseudo-observations derived from the superstructure's response acceleration data during the NN-VMC and TMCMC updating processes by KobeNakayamate_NS, Takatori.NS ground motion was input to the ground surface of the analytical model. In this case, considering that actual seismic response observations are affected by factors such as measurement errors, white Gaussian noise with a standard deviation of 5% relative to the RMS (Root Mean Square) of the absolute response acceleration was added to the full-time data in this study to generate the pseudo-observation data. The maximum acceleration of each input ground motion is KobeNakayamate_NS (Max: -818.016(cm/s/s)), Takatori.NS (Max: -642.0(cm/s/s)) and Sendai.NS (Max: 1519.8 (cm/s/s)). Additionally, the seismic wave data KobeNakayamate_NS is completely new and untrained by the neural network, For Takatori_NS, the data were used for training the neural network. The Sendai.NS ground motion, which was used for seismic response estimation, was employed to examine the seismic response.

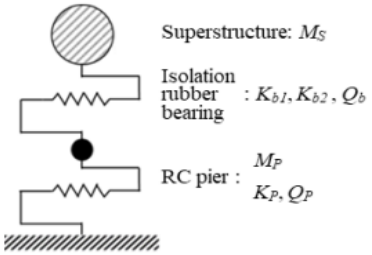


Figure1. Analytical model.

3 NN-VMC Framework

3.1 Feature Engineering

In the Neural Network Variational Monte Carlo (NN-VMC) framework for seismic isolated bridge parameter identification, feature engineering plays a pivotal role in transforming raw seismic acceleration time-history data into multi-dimensional feature vectors suitable for neural network training. The significance of this step is manifested at three fundamental levels:

First, from an information theory perspective, raw seismic acceleration time histories typically contain thousands to tens of thousands of sampling points. Directly using such high-dimensional data as neural network input leads to the notorious "curse of dimensionality." A high-dimensional input space not only significantly increases network parameters but also causes convergence difficulties, elevated overfitting risks, and excessive computational resource consumption. Through feature extraction, the input dimensionality can be reduced to 28 while maintaining, or even enhancing, information density.

Second, from an earthquake engineering perspective, feature extraction endows each input dimension with clear physical significance. The 28-dimensional feature vector constructed in this study encompasses amplitude parameters (PGA, PGV, PGD, etc., 6 items), spectral parameters (predominant frequency f_p, mean period T_m, etc., 4 items), response spectrum parameters (spectral accelerations S_a at 10 critical periods plus spectrum intensity SI and ASI, 12 items), energy and duration parameters (Arias intensity I_a, cumulative absolute velocity CAV, significant duration D_{5-95}, etc., 4 items), and statistical parameters (skewness γ_1, kurtosis γ_2, etc., 2 items). All these features are derived from theoretical foundations in earthquake engineering rather than arbitrary signal processing operations.

Third, from the perspective of parameter identification, feature extraction establishes a physical bridge between ground motion characteristics and structural response. In the 2-degree-of-freedom (2-DOF) isolated bridge model used in this project, the nonlinear structural response depends strongly on the spectral energy distribution of the ground motion near the structure's natural frequencies. The response spectrum feature $S_a(T, \zeta)$ precisely encodes this critical information — the spectral acceleration value at the first modal period T_1 determines the first-mode response amplitude. In contrast, the energy features I_a and CAV quantify the total input energy of the ground motion and directly correlate with the cumulative plastic damage of the structure. The statistics for all 28-dimensional features are presented in **Table 2**.

Within the NN-VMC project framework, the Feature Engineering module serves as an indispensable bridge, integrating "Offline Surrogation" with "Online Instantaneous Inference." Data generation phase: guided by the generate_surrogate model, the 2-DOF mechanical surrogate simulator outputs dynamic responses (acceleration sequences $a(t)$) spanning tens of seconds. Through the feature extraction function, the system immediately distills 28 critical mechanical parameters, covering peak indicators, energy metrics, frequency components, and structural response spectra specific to target frequency bands (S_a, SI, ASI clusters). Instantaneous inference phase: real-time observed data orchestrated by the vmc_bridge model undergo the exact same purification module. They are subsequently concatenated with the similarly extracted seismic waveform features to forge a 56-dimensional panoramic vector ($28_{eq} + 28_{resp}$), which is then fed into the pre-trained VAE encoder. This End-to-End consistent feature stripping paradigm constitutes the shortest mapping manifold from the discrete physical domain into the smooth neural network latent space (ϕ).

Table 2. 28-dimensional features.

Feature ID	Symbol	Feature Name
F1	PGA	Peak Ground Acceleration
F2	PGV	Peak Ground Velocity
F3	PGD	Peak Ground Displacement
F4	R_av	PGA-to-PGV Ratio
F5	a_rms	Root-Mean-Square Acceleration
F6	SMA	Sustained Maximum Acceleration
F7	f_p	Predominant Frequency
F8	T_m	Mean Period
F9	epsilon	Bandwidth Parameter
F10	N0+	Expected Zero-Crossing Rate
F11-F20	Sa(T)	Spectral Acceleration
F21	SI	Housner Spectrum Intensity
F22	ASI	Acceleration Spectrum Intensity
F23	I_a	Arias Intensity
F24	CAV	Cumulative Absolute Velocity
F25	D_5-95	Significant Duration D_5-95
F26	SED	Specific Energy Density
F27	gamma_1	Skewness
F28	gamma_2	Kurtosis

3.2 Variational Inference (VI)

Due to the difficulty of calculating the marginal likelihood in the Bayesian inference framework, this high-dimensional integral is usually not analytically tractable, which is precisely the motivation for variational inference (VI). The core idea of the VI principle is to approximate the posterior distribution $p(\theta|\mathcal{D})$ using a simple parameterized distribution $q_\phi(\theta)$. $\mathcal{D}$ denotes the time-series model matrix of the input seismic wave acceleration, referred to herein as the observed data. The formula for approximating the posterior distribution using the variational distribution $q_\phi(\theta|\mathcal{D})$ is as follows:

$$q_\phi^*(\theta|\mathcal{D}) = \underset{q\in Q}{\operatorname{argmin}} D_{KL}(q_\phi(\theta|\mathcal{D}) \parallel p(\theta|\mathcal{D})) \tag{2}$$

The loss function consists of a reconstruction term and a regularization term, as follows:

$$\mathcal{L} = \underbrace{\mathbb{E}_{q_\phi}[\log p(\mathcal{D}|\theta)]}_{Reconstruction\ Term} - \underbrace{D_{KL}(q_\phi(\theta|\mathcal{D}) \parallel p(\theta))}_{Regularization\ Term} \tag{3}$$

3.3 Neural Network Architecture

The core innovation of the proposed Neural Network Variational Monte Carlo (NN-VMC) framework lies in the synergetic integration of surrogate modeling and amortized variational inference. Rather than adopting the standard generative Variational Autoencoder (VAE) paradigm, which lacks physical interpretability in the latent space, this framework strictly constrains the latent manifold to represent explicit physical properties of a structural system (e.g., stiffness degradation, yield strength, damping capacities). To achieve this, the architecture bifurcates into two distinct, deeply coupled Multi-Layer Perceptron (MLP) networks: The Surrogate Physics Model (acting as the definitive deterministic decoder), and the Variational Inference Encoder (acting as the probabilistic posterior approximator). By sequentially training and subsequently coupling these networks, the framework effectively bypasses the intractable $\mathcal{O}(N)$ computational bottleneck inherent in traditional Markov Chain Monte Carlo (MCMC) techniques, reducing Bayesian inference from hours to sub-seconds while maintaining rigorous physical constraints. As shown in **Figure 2**. Architecture of the Variational Autoencoder (VAE) within the NN-VMC framework. The encoder network $q_\varphi(z|x)$ maps 56-dimensional input features (28 earthquake + 28 structural response features) to a 10-dimensional output comprising 5 variational means μ_φ and 5 log-variances $\log \sigma^2_\varphi$. Through the reparameterization trick and a per-parameter sigmoid bijector, the latent variables are mapped to the physical parameter space $\Theta = [\theta_{lo}, \theta_{hi}]$. The decoder is a frozen surrogate physics network that predicts normalized response features from the sampled parameters. Training maximizes the evidence lower bound (ELBO) via β-annealed KL divergence, with gradients flowing analytically through the entire pipeline back to the encoder weights.

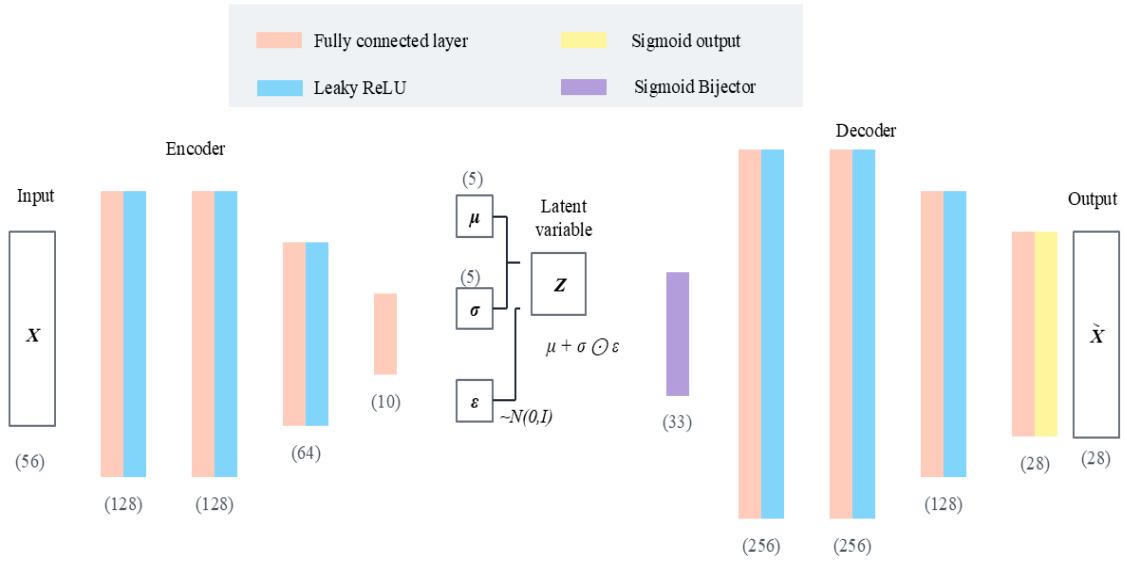


Figure 2. Encoder-Decoder architecture of VAE.

4 Numerical Example

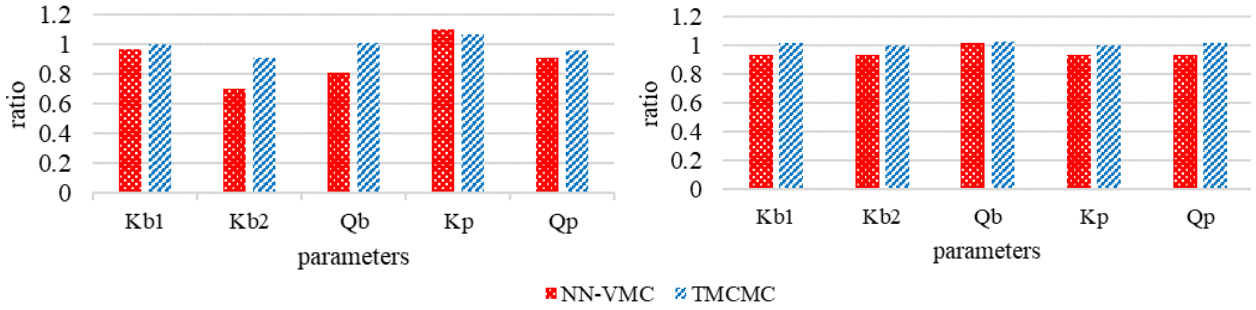
First, in **Figure 3**, an explanation begins with (1) KobeNakayamate_NS. The seismic wave data KobeNakayamate_NS is completely new to the neural network, meaning it is unfamiliar and inexperienced with the network. Instead of inputting the full-length seismic wave data, only the strong-motion portion was extracted, from $t_{\text{start}} = 15\text{s}$ to $t_{\text{end}} = 30\text{s}$. For accuracy analysis, it should be emphasized that 1 in the figure represents the baseline; the closer the value is to 1, the higher the accuracy. The Kb1 value from NN-VMC is close to 1, while that from TMCMC is even more

accurate. The accuracy of Kb2 from both NN-VMC and TMCMC is unsatisfactory. The accuracy of Qb from NN-VMC is relatively low, whereas TMCMC yields excellent accuracy for Qb. The Kp results from both NN-VMC and TMCMC exceed 1, but the difference is small, indicating comparable accuracy. Similarly, the Qp accuracies of the two methods are comparable, though TMCMC is slightly more precise. Overall, TMCMC is more accurate. NN-VMC shows low accuracy in Kb2 and Qb, but performs well on other parameters. Nevertheless, given that the neural network has been trained on only 41 seismic wave records, its performance is satisfactory, and there remains considerable room for improvement. For (2) Takatori_NS, the data were used for training the neural network. Therefore, the full-duration data from t∈[0s,120s] were input to the NN-VMC. For TMCMC, however, only the strong-motion part from $t \in [0s, 20s]$ was used. In terms of accuracy, NN-VMC and TMCMC are comparable, though TMCMC is marginally more accurate. According to the computational efficiency comparison between NN-VMC and TMCMC in **Table 3**, NN-VMC can complete inference within 0.1–0.4 seconds, whereas TMCMC requires a much longer time. If full-length seismic wave data (e.g., $t \in [0s, 120s]$) or multi-peak seismic data are used, TMCMC may take hours to finish inference.

Table 3 compares computational efficiency. NN-VMC completes inference within 0.1–0.4 seconds, whereas TMCMC requires significantly longer time and may take hours for long-duration or complex seismic inputs. It should be noted that NN-VMC and TMCMC are based on fundamentally different computational paradigms shown in **Table 4**. The reported NN-VMC time includes only the online inference phase—feature extraction, a forward pass through the trained VAE encoder, and Monte Carlo sampling—without any physical simulation. In contrast, TMCMC time includes the full Bayesian updating process, including prior sampling, repeated likelihood evaluations via physical simulations, tempering, and MCMC sampling. This difference stems from the amortized inference framework of NN-VMC, where computational cost is shifted to an offline phase (dataset generation, surrogate model training, and VAE training). As a result, online inference becomes extremely fast. Such comparisons are consistent with standard practices in amortized inference methods. From an application perspective, this comparison is particularly relevant for long-term and real-time structural health monitoring, where inference must be performed repeatedly. TMCMC performs all computations after each earthquake, whereas NN-VMC leverages prior offline training to enable immediate inference once new seismic data is available. Therefore, the offline cost is not a hidden advantage but a deliberate preparation for future events.

In summary, although nn-vmc_time excludes offline training cost and tmcmc_time includes full online computation, the comparison is methodologically justified. NN-VMC achieves comparable accuracy while enabling significantly faster inference, making it suitable for real-time and large-scale applications.

About computational time comparison between NN-VMC and TMCMC, it is necessary to clearly state that these two methods are based on completely different computational paradigms. For more detailed comparisons, please refer to



(1) KobeNakayamate_NS (no trained) (2) Takatori_NS (trained)

Figure 3. Comparison of inference results between NN-VMC and TMCMC.

Table 3. Comparison of the computational times of online inference speed (NN-VMC and TMCMC)

seismic wave	NN-VMC (sec)	TMCMC (sec)	How fast(times)
KobeNakayamate_NS	0.4081536	734.8235912	1800
Takatori_NS	0.1419747	1147.390515	8082

Table 4. Side-by-Side Comparison of Computation Contents

Item	NN-VMC (time)	TMCMC (time)
Measurement scope	Online inference phase only	Entire online inference phase
Physical simulation	Not included (replaced by trained surrogate)	Included
Likelihood evaluation	Not required	1,000 evaluations × all annealing steps
Model updating / learning	Not included (trained models are loaded)	Weights and covariance updated at every step

Figure 4 compares posterior parameter distributions using KobeNakayamate_NS (untrained data). A key observation is that the left Y-axis (TMCMC) and right Y-axis (NN-VMC) use different scales, reflecting differences in density magnitude despite similar distribution centers. For Kb1, both methods accurately identify the true value, with comparable distribution widths. TMCMC shows slight right skewness, whereas NN-VMC is more symmetric and sharper, indicating a more concentrated posterior and comparable uncertainty quantification. Kb2 exhibits the largest discrepancy. TMCMC produces a wide, right-skewed distribution with a flat peak, suggesting poor identifiability. In contrast, NN-VMC yields a narrow and concentrated posterior near the true value, demonstrating superior uncertainty quantification, although some deviation from the true value remains. For Qb, both methods produce narrow and symmetric distributions with low dispersion, indicating stable estimates. However, NN-VMC shows a slight deviation from the true value. For Kp and Qp, both methods correctly identify the true values with narrow and well-centered distributions. NN-VMC tends to produce sharper posteriors, while TMCMC shows slightly broader distributions. Overall, both methods demonstrate low uncertainty and reliable performance for these parameters.

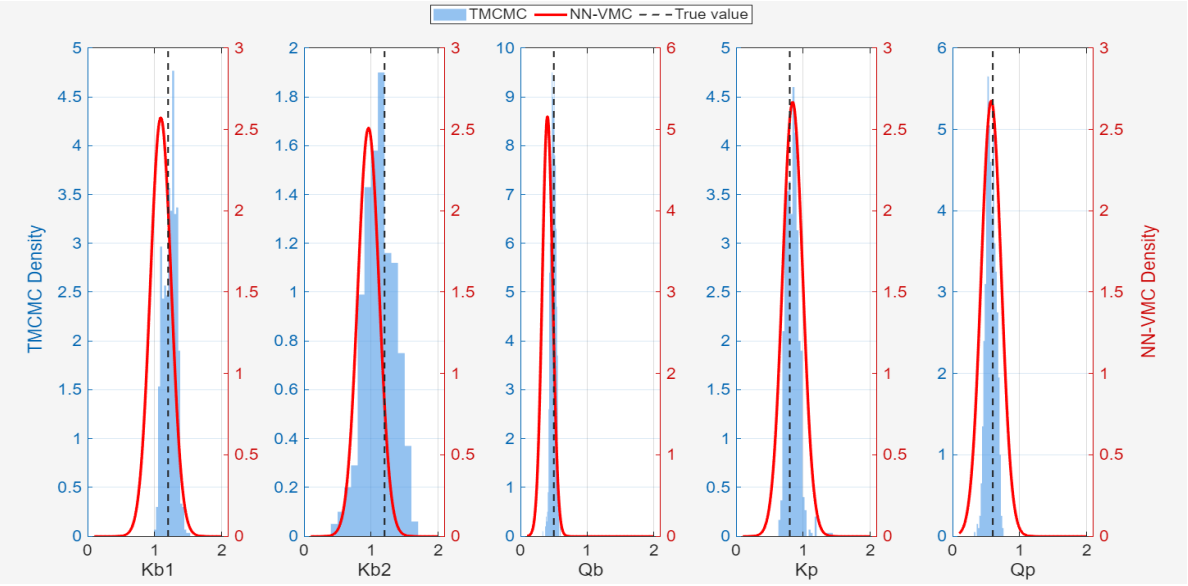


Figure 4. Parameter Distribution Comparison (KobeNakayamate_NS (no trained))

5 Conclusions and Future Research

This study proposes an innovative computational framework for inferring structural parameters of seismic-isolated bridges. Its core innovation lies in the development of a Physics-Informed Neural Network Variational Monte Carlo (NN-VMC) method based on surrogate-assisted amortized variational autoencoders, which instantaneously identifies five unobserved physical parameters of a two-degree-of-freedom base-isolated bridge from historical seismic records. According to the comparative analysis results, NN-VMC achieves accuracy comparable to TMCMC, validating its suitability as a complete alternative to TMCMC. In future research, it is critical to increase the variety and quantity of seismic wave data for model training, enabling the model to continuously accumulate experience in analyzing seismic data.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number JP24K00973.

References

Takeshi Kitahara, Masaru Kitahara, Michael BEER: Influence of Seismic Response Observation Data on Structural Parameter Updates Using Transitional MCMC, Proceedings of the 25th Symposium on Seismic Design of Bridges and Other Structures, pp. 125-130, 2022. (In Japanese)

Blei, D. M., Kucukelbir, A., & McAuliffe, J. D. (2017). Variational inference: A review for statisticians. Journal of the American Statistical Association, 112(518), 859-877.

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational Bayes. arXiv preprint arXiv:1312.6114.

Carleo, G., & Troyer, M. (2017). Solving the quantum many-body problem with artificial neural networks. Science, 10 Feb 2017 Vol 355, Issue 6325 pp. 602-606 DOI: 10.1126/science.aag2302v.