Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen

doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P2538-cd

Data-driven Stochastic Model Updating with Diffusion Models

Tairan Wang

Department of Aeronautics and Astronautics Engineering, University of Southampton, United Kingdom. E-mail: tairan.wang@soton.ac.uk

Sifeng Bi

Department of Aeronautics and Astronautics Engineering, University of Southampton, United Kingdom. E-mail: sifeng.bi@soton.ac.uk

Stochastic model updating is a vital technique in engineering that can calibrate the input parameters of the computational model to reflect the real-world physical system while accounting for the existence of uncertainties. However, traditional methods such as the Bayesian approach always struggle with high-dimensional and nonlinear problems. Thus, there is a trend to adopt data-driven approaches to solve stochastic model updating problems since of their remarkable capability to process high-dimensionality and nonlinearity. Apart from utilising neural networks as a surrogate for forward models, the conditional invertible neural networks (cINNs), a type of flow-based deep generative model, can serve as an inverse surrogate to address stochastic model updating problems alternatively. Recently, another group of deep generative models called Diffusion Models has become very popular in generation tasks because of their better ability to handle complex distributions, flexibility in network architecture and stability in training. In this work, the feasibility of leveraging diffusion models to resolve stochastic model updating problems is investigated. Diffusion models transform a simple latent distribution (e.g., Gaussian noise) into a complex distribution that aligns with observation data through a gradual iterative process. In contrast to cINNs, diffusion models build up complexity in the learned distribution progressively through a series of Markov chains, allowing for more accurate modelling of complex systems with high uncertainty. A 3 DOF spring-mass system was adopted as an example. The training dataset is formed by input parameters generated from the prior distribution and synthetic observation data obtained from the forward numerical model. This work presents diffusion models as a potential alternative to conventional Bayesian approaches for stochastic model updating, with advantages in accuracy, uncertainty, and flexibility for complicated, real-world applications.
Keywords: stochastic model updating, uncertainty analysis, deep generative models, machine learning

1. Background

All the engineering models are not entirely correct since the inevitable discrepancies between the experimental observations and the numerical simulations caused by the existence of aleatory and epistemic uncertainties (Bi et al. 2023). The model updating technique was developed in the deterministic domain by adopting the sensitivity-based approach (John E. Mottershead, Link, and Friswell 2011) to calibrate the input parameters of the numerical models. The parameters of interest then become those affected by the uncertainties, so the stochastic model updating was then proposed (J.E. Mottershead et al. 2006). A variety of approaches demonstrated their capability for stochastic model updating problems, such as the optimisation-based approach and Bayesian model updating framework (Beck and Katafygiotis 1998). The optimisation-based approaches define an objective function that can quantify the differences between the simulation and observation. The minimised objective function will result in optimised input parameters. The Bayesian model updating framework, on the other hand, was established based on the Bayesian inference, which can obtain the posterior distribution of uncertain parameters from prior knowledge and the likelihood function. The likelihood computation can be achieved through various methods, such as Euclidean and Bhattacharyya distance-based likelihood functions, which are widely used as an approximation to enhance computational efficiency. Some sampling algorithms like Markov chain Monte Carlo (MCMC) and Transitional Markov chain Monte Carlo are also utilised to obtain the posterior distribution. The above structure makes the Bayesian model updating framework powerful while at the same time making it inefficient in some instances. It leads to the tendency to leverage data-driven techniques to facilitate model updating problems.

In the present studies, a lot of work has been done by developing novel neural network-based surrogate models to replace the computationally expensive forward numerical (e.g., finite element models (FEMs)). However, the drawbacks of traditional model updating methods, like high computational demand and being overwhelmed by high-dimensional data, remain. The deep generative models, e.g., Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs), and Conditional Invertible Neural Networks (cINNs), which are widely used in the image generation field and other generation tasks, become potential tools to tackle stochastic model updating problems. They share similar objectives with model updating, which is to generate unknown data given some conditions after training by acknowledged data. Recently, the VAEs (Lee et al. 2024), GANs (Yuan et al. 2023) and cINNs (Zeng, Todd, and Hu 2023) have all been successfully implemented in the stochastic model updating area. Diffusion models (Ho, Jain, and Abbeel 2020) are a different class of deep generative models that have recently become more popular in generation tasks because of their superior handling of complex distributions, network architecture flexibility, and training stability, which opens the possibility of employing them for addressing stochastic model updating problems.

Similar to other latent space-based methods, diffusion models can also build connections between a complex distribution and the latent distribution (normally standard Gaussian distribution) by adding Gaussian noise on the initial distribution over a series of time steps gradually until a pure Gaussian noise distribution is obtained, which is called the forward diffusion process. In the reverse diffusion process, the pure Gaussian noise can denoise progressively to generate a sample. In the training phase, the noise added at each time step of the forward process is predicted by minimising the difference between the true noise in the forward process and the predicted noise in the reverse process.

In this contribution, a novel stochastic model updating method based on diffusion models is proposed. The diffusion model is set up and trained by the training dataset obtained by sampling input parameters from the pre-defined prior distribution and collecting the corresponding output data of the forward model. The well-trained diffusion model can then generate the posterior distribution by taking the observation data as the condition. A classic three-degree-of-freedom spring-mass system is adopted as an example. The training dataset is formed by stiffness parameters and the natural frequencies, and the synthetic observation data are output data derived by taking samples drawn from the true distribution as input parameters.

The remainder of the paper is organised as follows: Section 2 introduces the basic theory of diffusion models. In Section 3, the framework for diffusion models-based stochastic model updating is presented. This framework is implemented in a case study involving a three-degree-of-freedom spring-mass system to demonstrate the capabilities of the novel framework in Section 4. Section 5 presents the conclusions and perspectives.

2. Diffusion Models

Diffusion models are generative models that take influence from thermodynamic diffusion processes (Sohl-Dickstein et al. 2015). These models have proven to be remarkably effective in many fields such as computer vision, audio generation, and text generation (Chen et al. 2024), etc. The diffusion model consists of a forward process and a backward process. The fundamental idea is to represent the process of reverse generation as the opposite of a forward diffusion process that turns data into noise, as illustrated in Fig 1.

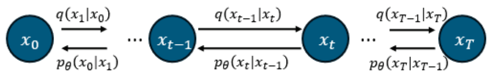


Fig. 1. Schematic of forward and backward process in diffusion models.

2.1.1 *Forward Diffusion Process*

In the forward diffusion process, a known distribution x_0 is converted into a pure noise distribution x_T by adding Gaussian noise incrementally over T Markov chains (i.e., time steps). Thus, the forward diffusion process can be expressed mathematically as follows Eq. (1):

$$q(x_t|x_{t-1}) = \mathcal{N}\left(x_t; \sqrt{1-\beta_{t-1}} \cdot x_{t-1}, \beta_t \mathbf{I}\right) \quad (1)$$

where β_t are the variance schedule parameters of the noise added at each step bounded by $\beta_t \in [0,1]$, which normally starts from a relatively small value and increases gradually; t refers to each time step; and $x_0 \sim q(x_0)$ is a sample from the known data distribution initially. The term $\mathcal{N}$ and $\mathbf{I}$ denote normal distribution and identity matrix, respectively.

Since the forward diffusion process consists of a series of Markov chains, the joint distribution that represents the sequence of transformation from x_0 to x_T can be written as Eq. (2):

$$q(x_T|x_0) = \prod_{t=1}^{T} \mathcal{N}\left(x_t; \sqrt{1-\beta_{t-1}} \cdot x_{t-1}, \beta_t \mathbf{I}\right) \quad (2)$$

The above equation is then reparametrised by a parameter α_t as shown in Eq. (3),

$$\alpha_t = 1 - \beta_t \quad (3)$$

and the cumulative product is then calculated as,

$$\overline{\alpha_t} = \prod_{i=1}^{t} \alpha_i \quad (4)$$

Therefore, Eq. (2) can be rewritten as Eq. (5) by substituting Eq. (3) and Eq. (4) into it.

$$q(x_T|x_0) = \mathcal{N}\left(x_t; \sqrt{\overline{\alpha_t}} \cdot x_0, (1-\overline{\alpha_t})\mathbf{I}\right) \quad (5)$$

Let a random noise $z_t \sim \mathcal{N}(0, I)$, the transitional x_t at any time step can be computed as shown in Eq. (6).

$$x_t = \sqrt{\overline{\alpha_t}} \cdot x_0 + \sqrt{1-\overline{\alpha_t}} \cdot z_t \quad (6)$$

Hence, the initial data is corrupted by adding a series of Gaussian noise incrementally and each step is conditional on the previous step.

2.1.2 *Reverse Diffusion Process*

The reverse diffusion process is envisioned as an iterative denoising sequence that recovers the initial distribution x_0 once it is diffused from the pure noise x_T. The reverse is also a Markov chain process that estimates the previous latent variable x_{t-1} from the current noisy state x_t. Each step of the reverse process can be expressed mathematically as Eq. (7).

$$p_\theta(x_{t-1}|x_t) = \mathcal{N}\left(x_{t-1}; \mu_\theta(x_t, t), \Sigma_\theta(x_t, t)\right) \quad (7)$$

where θ represents the parameters of the diffusion model; $\mu_\theta(\cdot)$ and $\Sigma_\theta(\cdot)$ denote the mean and covariance matrix of the transitional data at each time step to be estimated during training.

The reverse process can also be expressed as the inverse of the forward process as shown in Eq. (8), which makes each step computable.

$$q(x_{t-1}|x_t, x_0) = \mathcal{N}\left(x_{t-1}; \tilde{\mu}_\theta(x_t, x_0), \tilde{\beta}_t \mathbf{I}\right) \quad (8)$$

where $\tilde{\mu}_\theta(x_t, x_0) = \frac{\sqrt{\overline{\alpha_{t-1}}} \cdot \beta_t}{1-\overline{\alpha_t}} \cdot x_0 + \frac{\sqrt{\alpha_t}(1-\overline{\alpha_{t-1}})}{1-\overline{\alpha_t}} x_t$, and $\tilde{\beta}_t = \frac{1-\overline{\alpha_{t-1}}}{1-\overline{\alpha_t}} \beta_t$.

2.1.3 *Training Process*

Similar to the Variational Autoencoders (VAEs), the diffusion models are trained by optimising the variational lower bound (VLB) on the negative log-likelihood function.

$$L = \mathbb{E}[-\log p_\theta(x_0)] \leq \mathbb{E}_q\left[-\log \frac{p_\theta(x_{0:T})}{q(x_{1:T}|x_0)}\right]$$
$$= \mathbb{E}_q\left[-\log p(x_T) - \sum_{t \geq 1} \log \frac{p_\theta(x_{t-1}|x_t)}{q(x_t|x_{t-1})}\right] \quad (9)$$

The loss function in Eq. (9) can be further simplified into Eq. (10) by using the Kullback–Leibler divergences between the transitional distributions at each time step of the diffusion process:

$$L = D_{KL}\left(q(x_T|x_0) || p(x_T)\right) +$$
$$D_{KL}\left(q(x_{t-1}|x_t, x_0) || p(x_{t-1}|x_t)\right) - \log p_\theta(x_0|x_1)$$

In practice, only the noise added and removed at each time step is adopted to construct the loss function. Thus, the objective function becomes to calculate the MSE between true noise ϵ and model-predicted noise ϵ_θ.

$$L_{simple} = \mathbb{E}_{x_0, t, \epsilon}[\|\epsilon - \epsilon_\theta(x_t, t)\|^2] \quad (11)$$

The optimal values for model parameters θ can be obtained through backpropagation and stochastic gradient descent. After training, the target data are generated by generating a sample from the noise distribution $x_T \sim \mathcal{N}(0, \mathrm{I})$ and performing the learned reverse diffusion progress.

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(x_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\epsilon_\theta(x_t, t)\right) + \sigma_t \epsilon \quad (12)$$

where $\epsilon \sim \mathcal{N}(0,1)$ is the predicted noise generated by the trained model and σ_t refers to the standard deviation value.

3. Stochastic Model Updating with Diffusion Models

Stochastic model updating is the technique for refining computational models of physical systems considering the existence of uncertainties from both measurements and model parameters and inherent variability. In the conventional Bayesian model updating approach, the posterior distribution is derived from evaluating the likelihood function and the prior distribution, which could be expensive. On the contrary, the evaluation of the likelihood function and elaborate sampling process can be omitted in the diffusion model-based stochastic model updating.

In the diffusion model-based model updating framework, the input data are corrupted into the latent space (pure noise) and then the denoising process is guided by the condition, which is summarised by the feature extractor from the observation data. A well-designed neural network is adopted in the denoising process to predict the noise to be removed. A schematic diagram of the diffusion model-based model updating framework is shown in Fig 2.

Thus, in the training phase, the training dataset includes the input data samples and the corresponding output data. The training process is carried out by adjusting the predicted noise in the denoising process in agreement with the added noise in the forward process. The well-trained model is capable of conditional input data generation by giving observation data to accomplish the model updating task.

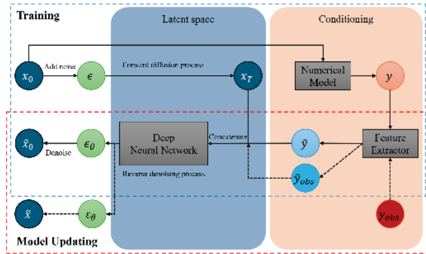


Fig. 2. Diffusion model-based model updating.

The posterior distribution of the parameter of interest is generated by a well-trained diffusion model given the true observation data as the condition. In the training process, both the input and output data are embedded as training data. Thus, the forward diffusion process introduced before turns into Eq. (13), where the parameter y refers to the condition (output data in this case).

$$q(x_t|x_{t-1}, y) = \mathcal{N}\left(x_t; \sqrt{1-\beta_{t-1}} \cdot x_{t-1}, \beta_t \mathbf{I}\right) \quad (13)$$

The forward process is not affected due to the involvement of the condition, but the conditional information is brought in (Dhariwal and Nichol 2021). For the reverse process, the joint distribution is written as Eq. (14).

$$q(x_T|x_0, y) = q(x_T)\prod_{t=1}^{T} p_\theta(x_{t-1}|x_t, y) \quad (14)$$

The objective function (Eq. (15)) is established to measure the differences between the predicted noise and the true noise with the condition of observation data.

$$L = \mathbb{E}[\|\epsilon - \epsilon_\theta(x_t, t, y)\|^2] \quad (15)$$

4. Case Study: 3-DOF Spring-Mass System

In this study, a classic 3-degree-of-freedom spring-mass system (shown in Fig. 3) is employed as a simulation-based case study for stochastic model updating by utilising the diffusion model-based model updating method introduced before. All the springs in the system are assumed to follow the linear elasticity assumption. The three stiffness parameters k_4, k_5, and k_6 and the three masses m_1, m_2 and m_3 are set to be constant variables and do not need update, the values are denoted in Table 1. The other three stiffness parameters k_1, k_2, and k_3 are random variables with uncertainty to be calibrated in this example. The three natural frequencies f_1 f_2, and f_3 are the quantities of interest whose uncertainties are controlled by the uncertain input parameters.

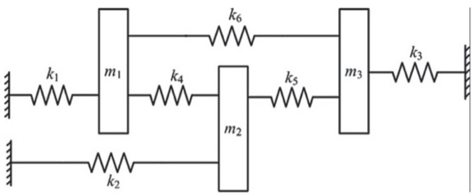


Fig. 3. 3-DOF spring-mass system.

The parameters of interest (k_1-k_3) are assumed to follow a uniform prior distribution, and the target distributions are determined as Gaussian distributions with various means and standard deviations, as illustrated in Table. 1.

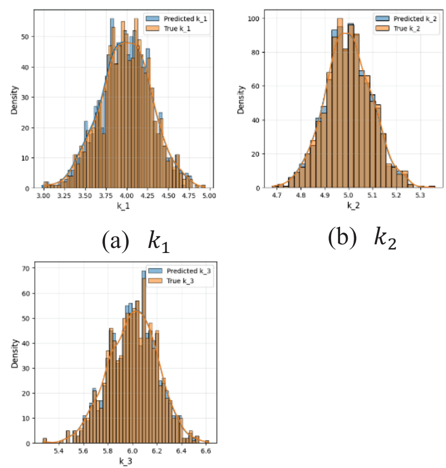
The training dataset is obtained by generating input samples from the prior distribution and collecting the corresponding model output. The synthetic observation data is derived from the output obtained by sampling input parameters from the target distribution. This model updating task is basically to validate if the distribution of input parameters generated by the diffusion model, taking the synthetic observation data as the condition, is in agreement with the target distribution.

Table 1. Uncertain characteristics of the 3-DOF system.

Parameter	Prior Distribution	Target Distribution
k_1	$U(3.0,7.0)$	$\mathcal{N}(\mu_1 = 4.0, \sigma_1 = 0.3)$
k_2	$U(3.0,7.0)$	$\mathcal{N}(\mu_2 = 5.0, \sigma_2 = 0.1)$
k_3	$U(3.0,7.0)$	$\mathcal{N}(\mu_3 = 6.0, \sigma_3 = 0.2)$
	Constant variables do not need update	
$k_4 - k_6$ $m_1 - m_3$	$k_4 = k_5 = k_6 = 5\ N/m$, $m_1 = 0.7\ kg, m_2 = 0.5\ kg, m_3 = 0.3\ kg$.	

The deep neural network for denoising process is designed as a fully connected neural network (FCNN). The training dataset consists of 10,000 sets of samples generated from the prior distribution and their corresponding natural frequencies. In the training process, the Adam optimiser is selected with a learning rate of 4e-4. The loss function is calculated as the MSE between the predicted noise and the generated random noise at each time step. The model is trained until the loss converges to a relatively small value and remains volatile within a reasonable interval.

In the reverse sampling process, a total of 1,000 sets of samples from the Gaussian noise, together with 1,000 sets of observation data, the natural frequencies, are generated by taking the input parameters sampled from the target distribution. Thus, a total of 1,000 sets of predicted data points were generated, constituting the posterior distribution of the model updating. The posterior distribution of the stochastic model updating using the diffusion model is shown in Fig. 4 below, with the target distribution for validation.



(a) k_1 (b) k_2

(c) k_3

Fig. 4. Posterior distribution for three stiffness parameters (a) k_1, (b) k_2, and (c) k_3.

According to the model updating results, all three parameters are well calibrated as the posterior distribution PDFs are all in good agreement with the target distribution PDFs. Not only the mean values but also the standard deviations of the calibrated parameters match the target values closely. The model updating problem is successfully solved by the diffusion model-based model updating framework. This result makes the diffusion model-based model updating a promising novel alternative to traditional Bayesian model updating framework.

5. Conclusion

In this study, the feasibility of the implementation of the diffusion model-based approach to the stochastic model updating problem is proven. Distinguishing from traditional model updating methods, the diffusion model-based approach can bypass the complexity of investigating the likelihood function and the computationally expensive sampling algorithms. Different from developing novel surrogate models using neural networks, the diffusion model-based model updating method can substitute for the conventional approach. The diffusion model-based stochastic model updating leverages the flexibility of deep learning models and the capability of high-dimensional or time-series data. Compared to other deep generative models, the generation process of the diffusion models is probabilistic, so the generation results might vary even with the same condition, but it can perform accurate predictions based on distributions.

In conclusion, this work proposed a novel stochastic model updating method on the basis of the popular diffusion models that bridge the gap between deep generative models and stochastic model updating problems after the other models (VAEs, GANs, cINNs) have all been utilised in the model updating. The diffusion model-based stochastic model updating has the potential to be practical in high-dimensional and non-linear spaces.

Reference

Beck, J. L., and L. S. Katafygiotis. 1998. 'Updating Models and Their Uncertainties. I: Bayesian Statistical Framework'. *Journal of Engineering Mechanics* 124 (4): 455–61. https://doi.org/10.1061/(ASCE)0733-9399(1998)124:4(455).

Bi, Sifeng, Michael Beer, Scott Cogan, and John Mottershead. 2023. 'Stochastic Model Updating with Uncertainty Quantification: An Overview and Tutorial'. *Mechanical Systems and Signal Processing* 204 (December):110784. https://doi.org/10.1016/j.ymssp.2023.110784.

Chen, Minshuo, Song Mei, Jianqing Fan, and Mengdi Wang. 2024. 'An Overview of Diffusion Models: Applications, Guided Generation, Statistical Rates and Optimization'. arXiv. https://doi.org/10.48550/arXiv.2404.07771.

Dhariwal, Prafulla, and Alex Nichol. 2021. 'Diffusion Models Beat GANs on Image Synthesis'. arXiv. https://doi.org/10.48550/arXiv.2105.05233.

Ho, Jonathan, Ajay Jain, and Pieter Abbeel. 2020. 'Denoising Diffusion Probabilistic Models'. In *Advances in Neural Information Processing Systems*, 33:6840–51. Curran Associates, Inc. https://proceedings.neurips.cc/paper/2020/hash/4c5bcfec8584af0d967f1ab10179ca4b-Abstract.html.

Lee, Sangwon, Taro Yaoyama, Masaru Kitahara, and Tatsuya Itoi. 2024. 'Latent Space-Based Stochastic Model Updating'. arXiv. https://doi.org/10.48550/arXiv.2410.03150.

Mottershead, J.E., C. Mares, S. James, and M.I. Friswell. 2006. 'Stochastic Model Updating: Part 2—Application to a Set of Physical Structures'. *Mechanical Systems and Signal Processing* 20 (8): 2171–85. https://doi.org/10.1016/j.ymssp.2005.06.007.

Mottershead, John E., Michael Link, and Michael I. Friswell. 2011. 'The Sensitivity Method in Finite Element Model Updating: A Tutorial'. *Mechanical Systems and Signal Processing* 25 (7): 2275–96. https://doi.org/10.1016/j.ymssp.2010.10.012.

Sohl-Dickstein, Jascha, Eric A. Weiss, Niru Maheswaranathan, and Surya Ganguli. 2015. 'Deep Unsupervised Learning Using Nonequilibrium Thermodynamics'. arXiv. https://doi.org/10.48550/arXiv.1503.03585.

Yuan, Zi-Qing, Yu Xin, Zuo-Cai Wang, Ya-Jie Ding, Jun Wang, and Dong-Hui Wang. 2023. 'Structural Nonlinear Model

Updating Based on an Improved Generative Adversarial Network'. *Structural Control and Health Monitoring* 2023 (1): 9278389. https://doi.org/10.1155/2023/9278389.

Zeng, Jice, Michael D. Todd, and Zhen Hu. 2023. 'Probabilistic Damage Detection Using a New Likelihood-Free Bayesian Inference Method'. *Journal of Civil Structural Health Monitoring* 13 (2): 319–41. https://doi.org/10.1007/s13349-022-00638-5.