

UNIFIED STOCHASTIC PDF DYNAMICS OPERATOR (USPDO): NEURAL OPERATOR-BASED MODELING OF NONLINEAR STOCHASTIC SYSTEMS UNDER DIVERSE NOISE EXCITATIONS

Pengtai Liao

College of Information Engineering, Tarim University, Alar, Xinjiang 843300, China

Abstract - In order to solve the problem that the Probability Density Function (PDF) evolution of nonlinear stochastic systems is difficult to describe uniformly under the action of complex noise, this study proposes a unified stochastic PDF dynamics operator (USPDO) model; this can realize the unified modeling of PDF dynamics under different random excitation conditions. Based on the correspondence between the stochastic differential equation and the PDF evolution equation, the model constructs a unified operator framework including a drift-diffusion operator, a fractional Lévy operator and a transition operator; meanwhile, it learns the mapping relationship between random parameters and probability density evolution via a neural operator structure. In the training process, the physical constraint residual loss is introduced; this guarantees that the predicted PDF meets the non-negative and probability normalization constraints. Aiming at the scenarios of Gaussian White Noise, Ornstein-Uhlenbeck colored noise, Lévy heavy-tailed noise and Poisson impulse noise, this study builds a verification sample using the relevant time series in the Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH) Noise Stress Test Database and the University of California, Riverside (UCR) Time Series Classification Archive. Meanwhile, it constructs a reference PDF based on the kernel density estimation results of the verification sample; it also analyzes the relationship between the data series and the corresponding random noise hypothesis through statistical feature tests. The experiment is evaluated by Kullback-Leibler (KL) divergence, Wasserstein distance, Negative Log-Likelihood (NLL), Rare Event Probability Error (REPE) and other indicators. The results show that under the tested multiple types of random excitation conditions, the USPDO model's KL divergence is about 0.02–0.04, the Wasserstein distance is about 0.04–0.07, the NLL is about 0.18–0.31, and the REPE is about 0.02–0.04. This suggests that the model can effectively maintain the non-negative and normalized constraints of the probability density function, thus realizing the PDF dynamic reconstruction under different noise conditions. The proposed method provides a unified modeling framework for probabilistic dynamic analysis in complex stochastic systems.

Keywords: Nonlinear stochastic system; PDF dynamics analysis; USPDO model; Random noise excitation; Rare event.

1. Introduction

In recent years, the application of nonlinear stochastic systems has been continuously expanded in aerospace, intelligent transportation, and energy equipment. Their modeling and analysis have become important directions in the research of control science and complex systems [1,2]. A prominent feature of such systems is that their dynamical processes are restricted by nonlinear structures and continuously affected by random disturbances; this leads to strong uncertainty and

complexity in state evolution [3]. Under the coupling effect of multiple types of random excitation (such as Gaussian, heavy-tailed, and pulse), the system may produce distribution broadening, tail thickening or multimodal structure; these phenomena are directly related to the stability of the system and the risk assessment of extreme events. Traditional methods mostly rely on low-order statistical features such as mean and variance; these methods reflect overall trends but fail to capture the fine-grained evolution laws of distribution patterns. Thus, obvious shortcomings exist in dealing with heavy-tailed

distributions, impulsive noise, or non-stationary disturbances.

Probability Density Function (PDF) dynamics analysis provides a new approach to solving this problem. It can reveal the evolution of peak height, tail distribution, and multi-modal structure by studying the distribution's changes over time; it can also quantitatively characterize the impact of noise on escape time, chaos degree, and the probability of rare events occurring. Existing studies have attempted to construct analytical models based on the Fokker-Planck equation (FPE) or stochastic averaging methods. However, when facing multiple types of complex noise, they are often limited by the complex form of operators and high numerical calculation costs [4]. Especially in high-dimensional or non-Gaussian scenarios, traditional approaches struggle to balance physical constraints and computational efficiency; this challenge limits their application and promotion in engineering systems.

Grounded upon this, the study proposes the Unified Stochastic PDF Dynamics Operator (USPDO) model. This model couples the stochastic differential equation (SDE) and probability density evolution equations under a unified operator framework. Meanwhile, a learnable approximation of complex operators is realized through a neural operator structure, thus achieving systematic modeling of probability distribution dynamics under the excitation of multiple types of noise. The research objects encompass nonlinear stochastic systems driven by Lévy-type heavy-tailed noise, Gaussian White Noise (GWN), Poisson impulsive noise, and Ornstein-Uhlenbeck (O-U) colored noise. The verification data are obtained from the Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH) Noise Database and the University of California, Riverside (UCR) Time Series Archive. The objective is to simultaneously achieve accurate reconstruction of distribution pattern evolution and preservation of physical constraints and rare event characteristics, offering support for system stability analysis and extreme risk assessment. The innovations of this study are mainly reflected in the following three aspects:

(1) The study constructs a unified operator framework, integrating drift-diffusion, correlation, fractional, and jump operators into the same modeling system; it achieves a unified description under multiple noise conditions.

(2) A neural operator structure is introduced, ensuring the non-negativity, normalization of the PDF, and its compliance with dynamic laws through a conditional input mechanism and physics-constrained residual learning.

(3) The study conducts systematic verification using public multi-source datasets, covering

Gaussian, non-Gaussian, colored, and impulse noise scenarios; it demonstrates the model's adaptability and generalization ability in complex environments.

2. Materials and Methods

2.1 Related Work

In the research on nonlinear stochastic systems, accurately characterizing the PDF dynamics analysis of systems in complex noise environments has always been an important issue in control science and stochastic dynamics. Bai et al. introduced a multi-scale representation strategy for stochastic nonlinear systems with time delay based on the Path Integral method. They converted the delay term into an equivalent Markov process, solving the transient and steady-state PDF through the Path Integral. This successfully revealed the system's stochastic bifurcation phenomenon based on different parameters, and reduced the computational burden of high-dimensional problems by using sparse matrix storage **Error! Reference source not found.** Er et al. expanded the evolutionary exponential polynomial closure method. Through variable transformation and the solution of the non-stationary Fokker-Planck-Kolmogorov Equation (FPKE), they realized the PDF dynamics' approximation of nonlinear oscillators with potential barrier constraints under multi-modulated GWN. Their results showed that the probability distribution shape was extremely sensitive to noise correlation and potential barrier structure [6]. Spanos and Pomaro focused on nonlinear dynamic systems with fractional derivatives. They derived the expected values of equivalent damping and stiffness using the conditional spectrum and harmonic balance method. They verified the accuracy of the steady-state power spectrum and PDF solution through Monte Carlo simulation, which provided a new idea for analyzing complicated systems considering memory effects [7].

In addition, with the diversification of stochastic noise types, researchers began to explore the modeling of nonlinear systems under Gaussian, non-Gaussian, and impulse excitations. Tang et al. analyzed the fractional-damping Morse oscillator driven by GWN based on a stochastic averaging method. They derived the analytical expressions of drift and diffusion coefficients, obtained the steady-state PDF's solution, and successfully revealed the law of stochastic P-bifurcation [8]. Koverda and Skokov proposed to describe self-similar stochastic processes using nonlinear stochastic equations from the perspective of non-equilibrium statistical mechanics. The obtained solutions showed power-law characteristics in spectral density and PDF, and they studied the stability of stochastic processes through entropy analysis [9]. Chen et al. presented a dimension-reduced probability density evolution

equation for high-dimensional nonlinear systems excited by multiplicative white noise. They projected the high-dimensional system into a one-dimensional or two-dimensional FPKE to solve the transient PDF.

This method effectively captured the tail characteristics and markedly reduced the computational complexity [10]. Peng et al. constructed a one-step Transition Probability Density Function (TPDF) matrix for systems excited by Poisson white noise, which could handle impulse interference under high arrival rate conditions. The results showed that the proposed Path Integral method maintained high accuracy even under intricate noise driving [11]. Landgraf et al. further reviewed the probability prediction methods in nonlinear systems. They pointed out that accurate PDF prediction was crucial to system performance in model predictive control applications. Besides, various approximation methods were classified and numerically evaluated, furnishing a systematic reference for stochastic control problems [12].

In recent years, learning-based and data-driven modeling methods have been widely used in the analysis of complex systems. The related research mostly focuses on feature selection, ensemble learning, and prediction performance optimization. For example, Gul et al. constructed a k-Nearest Neighbors (kNN) integrated model through feature weighting and a sub-sampling mechanism to improve the prediction stability under complex data conditions [13]. Ali et al., based on the extended neighborhood rule and random projection strategy, improved the classification accuracy of kNN

integrated on multiple datasets [14]. Ali et al. further proposed a weighted Fisher scoring method to alleviate the influence of high-dimensional feature redundancy on model performance [15]. Khan et al. constructed a two-stage integration framework to deal with the forecasting problem under the condition of unbalanced data [16]. This kind of method shows strong adaptability in data-driven prediction and pattern recognition tasks. However, its modeling goal mainly focuses on the statistical prediction or classification performance of output variables; it usually does not involve the operator-level correspondence between SDE and probability density evolution, nor does it uniformly model PDF dynamic structures under different random excitation mechanisms. Faced with the reliability and extreme event prediction needs of complex engineering systems under strong noise conditions, it is necessary to change the idea of a single analytical derivation. Therefore, this study proposes a unified operator framework of the USPDO model, optimizing PDF dynamics modeling through neural operator learning and physics-constrained residual minimization. Meanwhile, it ensures non-negativity and probability conservation, providing an efficient and robust statistical modeling tool for engineering systems. In order to further summarize the characteristics and limitations of existing research methods in the random PDF dynamics modeling, this study organizes the relevant research work, as shown in Table 1:

Table 1. A comparative summary of methods related to random PDF dynamics

Author	Method	Target system	Types of noise that can be handled	Computational complexity	Limitation
Bai et al.	Path Integral	Time-delay nonlinear stochastic system	Gaussian noise	High	High-dimensional problems have high computational costs
Er et al.	FPKE & Exponential polynomial closure	Nonlinear oscillatory system	Multi-modulated Gaussian noise	Medium	Mainly targets Gaussian noise
Spanos & Pomaro	Harmonic balance & Monte Carlo	Fractional-order nonlinear system	Gaussian random excitation	High	Numerical computation costs are relatively high
Chen et al.	DR-PDEE	High-dimensional stochastic dynamical system	Multiplicative white noise	Medium	Limited adaptability to complex non-Gaussian noise
Peng et al.	Path Integral & TPDF	Poisson noise system	Poisson impulse noise	High	Difficult to extend to multi-noise situations

In Table 1, existing methods mostly focus on the research of a single noise mechanism or a specific system. They still have certain limitations when dealing with multi-source random excitation, non-local operators, and complex noise coupling problems. At the same time, they also face challenges in terms of computational efficiency and scalability. In addition, in recent years, methods of the Physics-Informed Neural Network (PINN) and neural operator have been gradually applied to the modeling tasks of complex dynamic systems and

partial differential equations. Raissi et al. put forward a PINN framework and realized the joint learning of data constraints and physical laws by integrating physical equation residuals into neural network optimization processes [17]; Yang et al. further proposed a Bayesian Physics-Informed Neural Network (B-PINN); they used a probability inference method to deal with the uncertainty in noisy PDE problems [18]. In the direction of operator learning, Lu et al. proposed a new network with small generalization error-Deep Operator

Network (DeepONet); their study learned the nonlinear mapping relationship between input function spaces and output function spaces via branch networks and backbone network structure [19]. Kovachki et al. systematically summarized the theoretical framework of neural operators; their work proved that this framework could be used as a function space mapping model in the process of solving complex partial differential equations [20]. The above methods expand the application scope of neural networks in physical system modeling. However, PINN mainly faces the solution approximation under the condition of given equations; neural operators mainly learn the general function space mapping. None of them models the unified operator relationship between SDE and PDF dynamic evolution under the condition of multiple random excitations. In contrast, the USPDO does not regress or classify the PDF as a pure data object. Instead, it takes the infinitesimal generator of stochastic dynamics as the core of modeling, uniformly characterizing the influence mechanisms of Gaussian diffusion, Lévy heavy-tailed perturbations, and Poisson jumps on PDF dynamics analysis at the operator level. While introducing a learning structure, this framework explicitly retains physical constraints such as probability conservation and non-negativity. Thus, it differs from existing learning-based methods oriented toward feature engineering or prediction performance in PDF dynamics modeling under multi-noise excitation conditions; this forms a unified operator-driven modeling approach for stochastic dynamics problems. PINN usually approaches a specific equation solution through a physical residual constraint network for a single partial differential equation with given parameters. Unlike PINN, USPDO faces the problem of PDF dynamic evolution under multi-random excitation; it learns the operator mapping relationship between stochastic system parameters and probability density evolution by introducing noise parameters as a conditional input. Compared with the neural operator method, USPDO learns the mapping relationship between input and output functions; it also embeds random dynamic operators (e.g., drift diffusion, fractional Lévy and Poisson jump) into the network training process, so that the model can keep the consistency of PDF evolution laws under different noise conditions.

2.2 Design Process of the USPDO Model

• Concept of PDF dynamics analysis

The essence of PDF dynamics analysis is to study the shape characteristics of the PDF dynamics analysis over time for nonlinear stochastic systems across diverse random excitations. Let the system

state variable be x_t , and its probability distribution be described by the PDF $p(x, t)$. When the system is driven by noise, the shape of the state distribution does not remain fixed. Instead, it manifests as changes in peak height, stretching or contraction of the distribution tail, emergence or disappearance of multi-modal structures, and formation of heavy-tailed characteristics [21]. When considering the basic SDE, the system dynamics can be written as:

$$dx_t = f(x_t)dt + g(x_t)\xi(t)dt \quad (1)$$

$f(x_t)$ denotes the drift term, $g(x_t)$ represents the diffusion coefficient, and $\xi(t)$ refers to the random excitation. In the basic derivation of Equation (1), it is assumed that $\xi(t)$ is zero-mean GWN, with statistical properties satisfying $\langle \xi(t) \rangle = 0$ and $\langle \xi(t)\xi(s) \rangle = \delta(t-s)$. Under this assumption, the system state process is Markovian. The PDF corresponding to this stochastic process satisfies the Fokker-Planck equation:

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x}(f(x)p(x,t)) + \frac{1}{2} \frac{\partial^\alpha}{\partial |x|^\alpha}(D(x)p(x,t)) \quad (2)$$

$D(x) = g^2(x)$ represents the diffusion intensity, which degrades to the form of the classic Fokker-Planck equation when $\alpha = 2$. When $\alpha \in (0,2)$, α denotes the order of the fractional derivative, describing the non-local diffusion and jump effects caused by non-Gaussian heavy-tailed random excitation. The order parameter $\alpha \in (0,2)$ is directly derived from the characteristic exponent of the Lévy stable process. According to the Lévy-Khintchine representation theorem, the infinitesimal generator of a symmetric α -stable Lévy random process corresponds to a spatial fractional operator at the probability density evolution level; its order is consistent with the characteristic exponent α of the Lévy stable distribution. Here, the fractional

operator $\frac{\partial^\alpha}{\partial |x|^\alpha}$ in Equation (2) adopts the definition of the Riesz fractional derivative. This operator corresponds to the infinitesimal generator of the symmetric α -stable Lévy random process, and its expression in the Fourier space is $\mathcal{F} \left[\frac{\partial^\alpha p(x,t)}{\partial |x|^\alpha} \right] = -|k|^\alpha \hat{p}(k,t)$; $\hat{p}(k,t)$ represents the PDF's Fourier transform $p(x,t)$. This generalized form is consistent with the traditional equation when the system satisfies the Markov property, and the noise is GWN. When the noise exhibits memory effects or heavy-tailed characteristics, the fractional derivative term can reflect long-range dependence or jump behaviors, thereby achieving a unified characterization of white

noise, colored noise, and Lévy-type random excitation. Correlated noise refers to a random process with finite correlation time, whose statistical properties do not satisfy the δ -correlation assumption of white noise.

In this study, correlated noise is specifically modeled using the O-U process, whose stochastic differential form can be expressed as: $d\eta(t) = -\frac{1}{\tau}\eta(t)dt + \sigma dW(t)$. Here, τ is the correlation time, σ is the noise intensity, and $W(t)$ is the standard Wiener process. By introducing the auxiliary variable $\eta(t)$ and expanding the state space, a joint PDF $p(x, \eta, t)$ is established, and its evolution satisfies:

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}(f(x)p) - \frac{\partial}{\partial \eta}\left(-\frac{1}{\tau}\eta p\right) + \sigma^2 \frac{\partial^2 p}{\partial \eta^2} \quad (3)$$

η denotes the correlated noise term driving the system, and its evolution follows a first-order O-U process, which characterizes the autocorrelation of the noise. Equation (3) corresponds to a special case of O-U colored noise, used to describe a first-order linear autoregressive correlation structure. For general cases of nonlinear or multi-scale correlated noise, further generalization can be achieved by introducing memory kernel functions or generalized Langevin equations. When considering jump-type random excitation, the evolution of the PDF includes a non-local term:

$$\frac{\partial p(x,t)}{\partial t} = \int_{\mathbb{R}} [p(x-\zeta, t) - p(x, t)]\phi(\zeta)d\zeta \quad (4)$$

$\phi(\zeta)$ refers to the distribution of jump amplitudes. The impact of noise on the system's long-term stability, instantaneous escape time, degree of chaos, and the probability of rare event occurrences can be quantified by observing these shape features' evolution. The PDF's global behavior can often be characterized by quantile trajectories, defined as:

$$Q_\alpha(t) = \inf\{x: \int_{-\infty}^x p(y, t)dy \geq \alpha\} \quad (5)$$

$Q_\alpha(t)$ stands for the distribution quantile at time t . Compared to only studying the mean and variance, the PDF dynamics analysis can reveal global statistical laws and non-Gaussian behavior, thus offering a more complete description for characterizing complicated stochastic systems [22].

• The implementation principle of the USPDO model

The USPDO model's implementation principle is based on the strict correspondence between SDE and PDF dynamics analysis equations. It integrates dynamical processes under various noise excitations through a unified operator framework and uses

neural operators to realize a learnable approximation of complex operators.

Under this framework, the core issue of this study is whether neural operators can accurately reconstruct the dynamic evolution of PDF under various random noises in the unified operator structure. Meanwhile, it preserves the non-negativity and normalization constraints of the PDF as well as the probabilistic characteristics of rare events. It should be noted that although the model training process adopts a learning approach based on physics-constrained residuals, its goal is not to solve a single partial differential equation under fixed parameter conditions. Instead, it aims to learn the probability density evolution operator parameterized by system parameters and noise characteristics, thereby realizing a unified mapping of PDF dynamics under different random excitation conditions. On this basis, this study further introduces a stochastic differential representation containing multiple types of noise terms; it aims to describe the extended forms of the basic SDE in Equation (1) under different random excitation conditions within a unified framework. This representation is used to characterize the state evolution process of the system under the combined action of Gaussian diffusion, heavy-tailed random perturbations, and impulsive excitation. It is important to clarify that the Lévy noise term in this study is used to characterize the continuous heavy-tailed diffusion effect caused by stable Lévy processes. In contrast, the Poisson jump term describes independent impulsive perturbations with a finite occurrence rate and finite amplitude. The two are distinguished as random excitations from different sources in this study, and the unified representation of the general Lévy-Itô decomposition is not adopted. Lévy noise modifies the decay rate of the PDF's tail primarily through fractional operators, thereby forming a heavy-tailed distribution structure. Poisson jumps, by contrast, alter the distribution position of probability mass in the state space via finite-amplitude transitions at random moments; this in turn generates local peaks or abrupt increases in tail probability within the PDF. During the time evolution process, the state variables of nonlinear stochastic systems are simultaneously affected by both deterministic drift terms and multiple types of noise terms [23], and their basic form can be expressed as:

$$dx_t = f(x_t, t)dt + g(x_t, t)dW_t + h(x_t, t)dL_t + r(x_t, t)dN_t, \quad (6)$$

$f(x_t, t)$ describes the deterministic drift of the system; the term $g(x_t, t)dW_t$ represents GWN driven by Brownian motion; $h(x_t, t)dL_t$ corresponds to the heavy-tailed noise of the stable

Lévy process; $r(x_t, t)$ reflects the impulsive jumps introduced by the Poisson process.

This stochastic process is transformed into an integro-partial differential equation at the PDF level, with the unified form expressed as:

$$\frac{\partial p}{\partial t} = \mathcal{L}_{\text{drift-diff}}[p] + \mathcal{L}_{\text{levy}}[p] + \mathcal{L}_{\text{jump}}[p] \quad (7)$$

The first part is the drift-diffusion operator, the second part is the non-local fractional-order operator corresponding to Lévy noise, and the third part is the jump operator corresponding to impulsive noise [24]. The linear superposition form of the operator in Equation (7) is based on the assumption that different random excitations are mutually independent or uncorrelated in the sense of increments. Under this assumption, the infinitesimal generator of the joint probability density evolution operator can be expressed as the linear sum of the generators of each random mechanism. If there are correlations, state-dependent couplings, or nonlinear interactions between noise sources, cross terms or nonlinear operator corrections may appear in the joint generator. The above linear superposition form no longer holds strictly. Such cases are not within the scope of this study. The local drift-diffusion operator directly originates from the broadening effect of Gaussian random disturbances on the probability distribution [25-27], and its expression is:

$$\mathcal{L}_{\text{drift-diff}}[p] = -\frac{\partial}{\partial x}(f(x, t)p) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(D(x, t)p) \quad (8)$$

The diffusion coefficient $D(x, t) = g^2(x, t)$ reflects the coupling relationship between noise intensity and state [28]. When the system is driven by Lévy noise, the PDF's evolution includes a heavy-tailed effect, and its operator form can be represented as:

$$\mathcal{L}_{\text{levy}}[p] = \kappa_\alpha \int_{\mathbb{R}} \frac{p(x+\zeta, t) - p(x, t)}{|\zeta|^{1+\alpha}} d\zeta \quad (9)$$

The adopted fractional operator is consistent with the definition of the Riesz fractional derivative in Equation (2) to ensure the consistency of probability density evolution in both physical and statistical senses. The parameter α determines the distribution tail's decay rate, and κ_α controls the jump intensity. When the system is driven by the Poisson process, the state may have finite-amplitude jumps at random times, and the corresponding operator reads:

$$\mathcal{L}_{\text{jump}}[p] = \lambda \int_{\mathbb{R}} [p(x - \zeta, t) - p(x, t)] \phi(\zeta) d\zeta \quad (10)$$

λ stands for the impulsive occurrence frequency, and $\phi(\zeta)$ denotes the jump amplitude distribution.

USPDO unifies the above different operators into the neural operator framework.

It embeds the system parameters and noise features into the network by using the conditional input mechanism; meanwhile, it switches the dynamic operators corresponding to four types of random excitation only by adjusting the noise feature parameters of the input layer, without rebuilding the network structure. Thus, the operator mapping can maintain flexible and adjustable expression ability under different noises. In this framework, the Gaussian diffusion operator mainly governs the overall broadening speed of the PDF; the fractional and jump operators affect the distribution's tail decay characteristics and the redistribution of probability mass in the state space, respectively. Hence, statistical features such as distribution broadening and tail thickening can be characterized under a unified operator structure. This unified operator framework is directly derived from the structural form of the infinitesimal generator of the SDE. Based on this, when the system dynamics contain fractional diffusion terms or non-integer order derivatives, the extension can be achieved simply by introducing the corresponding fractional operators into the operator set. Thus, the modeling framework is structurally adaptable to stochastic systems with fractional dynamic characteristics. At the network output end, the unnormalized function $u(x, t)$ is obtained first, and then the PDF prediction value is generated through exponential mapping and integral normalization:

$$\hat{p}(x, t) = \frac{\exp(u(x, t))}{\int_{\mathbb{R}} \exp(u(y, t)) dy} \quad (11)$$

Equation (11) is based on the idea of energy normalization, that is, ensuring non-negativity through exponential mapping of the network output, and then maintaining probability conservation through integral normalization. This form is equivalent to the extension of Softmax normalization in the continuous space. This structure ensures that the PDF remains non-negative at all times and satisfies the normalization condition [29]. Equation (11) is not an assumption of the general analytical form of PDF, but rather a representation method for numerical modeling and network parameterization. Its function is to introduce non-negativity and normalization constraints without presetting the specific function form of the solution. In specific implementation, the USPDO uses a multi-layer feedforward neural network structure to implement operator mapping. The network consists of an input layer, several hidden layers, and an output layer; the hidden layers adopt a structure of alternating stacked linear mappings and GELU activation functions; the output layer is a linear mapping to generate an unnormalized function.

The network's input vector is composed of state variables and noise parameters jointly. In other words, spatial variables, temporal variables, and random excitation features are concatenated to form a unified input vector; this describes the PDF dynamic states under different random conditions. The network outputs function values; then, it undergoes exponential mapping and integral normalization according to Equation (11) to yield the final prediction results of the PDF. During training, USPDO does not rely on directly fitting the sample distribution, but constructs a loss function through weak-form residuals. It substitutes the predicted distribution into the control equation and calculates the residuals at spatiotemporal sampling points, thereby constraining the network output to follow dynamic laws. Its loss is defined as:

$$\mathcal{L}_{\text{res}} = \int_{\Omega} \left| \frac{\partial \hat{p}}{\partial t} - \mathcal{L}[\hat{p}] \right|^2 d\Omega \quad (12)$$

Ω represents the spatio-temporal domain sampling set. At the implementation level, random excitation parameters (e.g., diffusion intensity, Lévy stability index, Poisson pulse rate, and correlation time) are embedded into the network as conditional inputs together with state variables; thus, a unified expression of operator mapping can be achieved under different noise mechanisms. This conditional input mechanism is accomplished via vector concatenation; it combines random excitation parameters with spatiotemporal variables into a unified input vector. This allows the network to learn the PDF dynamic relationships corresponding to different random excitation conditions within the same structure. Since various operators differ distinctly in numerical scales, USPDO introduces adaptive weight balancing mechanisms and incorporates all residual terms into the total loss function in a unified manner. During training, the total loss function is composed of various operator residuals and a regularization term, and its overall form can be expressed as follows:

$$\mathcal{L}_{\text{total}} = \sum_i \omega_i \mathcal{L}_i + \lambda_{KL} \mathbf{L}_{KL} \quad (13)$$

$\mathcal{L}_i$ denotes the residual loss term corresponding to different operators; ω_i is the weight coefficient of the corresponding residual term; $\mathbf{L}_{KL}$ refers to the Kullback-Leibler (KL) divergence regularization term; it constrains the information difference between the predicted probability density function and the reference distribution; λ_{KL} represents the regularization coefficient. The weights are dynamically adjusted during the training process. Equation (12) presents the definition of the overall residual term. In actual calculation, the residuals

corresponding to various operators are recorded separately, such as the drift-diffusion residual, Lévy operator residual, and jump operator residual. Equation (13) represents the weighted combination of these sub-item residuals to form the total loss; the weights are dynamically adjusted through an adaptive balancing mechanism to ensure a balanced training contribution of operators at different scales. Specifically, the weight of each residual term is normalized and updated according to its average residual amplitude in the current training batch. Thus, the residual error of an operator is prevented from dominating the overall optimization process on the numerical scale; meanwhile, the contributions of different physical operators are kept relatively balanced in training. The computational complexity of the non-local operator part is usually high. Hence, the model adopts an approximate form based on the Monte Carlo method, expressing the fractional operator or jump operator as a summation approximation of finite samples:

$$\mathcal{J}[p](x) \approx \frac{1}{M} \sum_{j=1}^M \frac{p(x+\zeta_j, t) - p(x, t)}{|\zeta_j|^{1+\alpha}} \quad (14)$$

Equation (14) is not obtained through the analytical derivation of Equation (9), but is approximately expressed by the numerical value formed through Monte Carlo sampling of the non-local integration operator in Equation (9). ζ_j is a jump variable randomly sampled from the noise distribution $\phi(\zeta)$; this variable calculates the integral operator in equation (9) in the form of a Monte Carlo approximation, thereby reducing the numerical complexity of the non-local operator and maintaining differentiability. The integral term of the non-local operator is approximated by summation with a finite number of random samples; the number of samples is consistent with the experimental settings. Each operator calculation adopts 128 random samples for estimation to strike a balance between computational efficiency and numerical accuracy. $\mathcal{J}[p(x)]$ represents the integral operator for non-local fractional or jump terms, which is used to approximate the long-range effects caused by Lévy noise and Poisson noise; meanwhile, its numerical implementation adopts Monte Carlo sampling approximation. During training, random samples are regenerated from the corresponding noise distribution for each iteration batch. The sampling results are averaged to obtain the operator approximation value, thereby reducing the variance fluctuation caused by random sampling.

This processing method enables the non-local operator to be embedded in the neural network in a differentiable form and participate in training together with the local drift-diffusion operator. Since the sampling scale remains fixed during training and the operator approximation value is calculated by batch averaging, a relatively stable numerical performance can be maintained in training iterations. After the model training is completed, the network can output the corresponding PDF function $\hat{p}(x, t)$ at any time; this result is a direct numerical approximation of the solution to the Fokker-Planck equation, thus realizing the explicit solution of PDF dynamics under different noise conditions. In the specific calculation process, random samples are first generated from the corresponding noise distribution according to the current type of random excitation. Then, the approximation value of the non-local operator is calculated. Subsequently, combined with the local drift-diffusion operator, it forms the residual term and participates in the update of network parameters. Through the combination of this unified operator framework and neural operators, the dynamic evolution of PDF under different noise excitations can be fully described in an end-to-end structure. As a result, a unified modeling process is formed for the statistical characteristics of nonlinear stochastic systems.

From the perspective of the implementation process of USPDO, the model's final objective is to achieve two key tasks. First, this model outputs the nonlinear stochastic system's PDF at any time and under any noise condition, to fully characterize the system state distribution's real-time evolution. Second, it ensures that the PDF satisfies the physical constraints of non-negativity and normalization, and can accurately reflect the impact of different noise mechanisms on the distribution shape.

3. Results and Discussion

3.1 Verification of the PDF Dynamics Modeling Effect of the USPDO Model under Multiple Types of Noise Excitation

• Verification data collection and preprocessing process

This study selects two public and authoritative datasets to verify the USPDO model's modeling capability for PDF dynamics of nonlinear stochastic systems under multiple random excitations. The first data type comes from the MIT-BIH Noise Stress Test Database on PhysioNet [30] (<http://www.physionet.org/physiobank/database/nstdb/>). This database includes baseline wander noise, myoelectric interference noise, and electrode motion noise, with a sampling rate of 360 Hz and each segment lasting approximately 30 minutes. This database can be used to construct inputs for GWN and non-Gaussian disturbances. The second type of data comes from subsets of the UCR Time Series Classification Archive of hexagon-ml (https://www.cs.ucr.edu/~eamonn/time_series_data_2018/), namely Nonlinear Gas Furnace, Electrocardiogram (ECG)5000, and Earthquakes.

The Nonlinear Gas Furnace subset sequences reflect the gas furnace system's nonlinear changing response under external inputs; the ECG5000 subset contains a large number of electrocardiographic signal sequences that exhibit obvious non-Gaussian fluctuation characteristics; the Earthquakes subset provides earthquake time series, where event occurrences show characteristics similar to a Poisson process, making it suitable for simulating impulse noise effects. The two datasets' descriptions are summarized in Table 2:

Table 2. Information on two public datasets

Dataset/noise type	Description	Sampling rate (Hz)	Duration/Length	Frequency range (Hz)	Characteristic statistics	Noise parameter range
MIT-BIH baseline wander noise	Breathing and low-frequency interference from electrode drift	360	30min	0.1-0.5	The peak value is approximately 0.05–0.3 mV, with low-frequency fluctuations.	$D \in [10^{-4}, 10^{-3}]$
MIT-BIH myoelectric interference noise	High-frequency muscle electrical activity superimposed signal	360	30min	20-100	The variance is relatively large, and the kurtosis is around 5-8.	$\sigma \in [0.1, 0.5]$
MIT-BIH electrode motion noise	Transient pulse disturbance of electrode displacement	360	30min	0-10 (broadband)	High peakness, with remarkable non-Gaussianity	$\lambda \in [0.05, 0.2]$
UCR Nonlinear Gas Furnace colored noise	Nonlinear input-output sequence of the gas stove	0.5	296 points	Dynamic correlation, primarily low frequency	Significant autocorrelation, $\tau \approx 2 - 10s$	$\sigma \in [0.05, 0.2]$

UCR ECG5000 Lévy-type heavy-tailed noise	Electrocardiogram signal time series	140–250	5000 points	Low frequency - mid frequency	Kurtosis: 5–9, presence of non-Gaussian fluctuations	$D \in [0.01, 0.1]$
UCR Earthquakes Poisson impulse noise	Earthquake event time series	Discrete events	500 points	Irregular pulses	Power-law tail distribution, with evident rare events	$\lambda \in [0.1, 0.5]$

In the preprocessing stage, all data are unified into equidistant time series, and the time step is set to $\Delta t = 0.01$ s to adapt to the operator framework of USPDO. The three types of noise from MIT-BIH first undergo band-pass filtering and detrending processing, followed by zero-mean normalization and standard deviation normalization to $\sigma^2 = 1$. This indicates that they can be used as inputs for GWN and non-Gaussian disturbances. This normalization performs only linear scaling on signal amplitudes to eliminate dimensional differences among datasets. It improves numerical stability during model training while preserving the relative shape and tail structure of the original probability distribution. For the Nonlinear Gas Furnace sequences, the autocorrelation function is calculated to obtain the correlation time constant $\tau \in [2, 10]$ s, and the diffusion coefficient $D \approx 0.05 - 0.2$ is obtained based on spectrum estimation. For the ECG5000 data, the non-stationary noise components are extracted through a difference method, and kurtosis and skewness are calculated to determine non-Gaussianity, with a typical kurtosis $\approx 5 - 8$. The Earthquakes subset, as an event-driven signal, uses a window statistical method to estimate the impulse intensity $\lambda \approx 0.1 - 0.5$ events/day on the scale of $\Delta t = 1$ day; this subset constructs the impulse amplitude distribution $\phi(\zeta)$ according to the magnitude distribution. The parameter settings mentioned above are determined mainly based on the statistical properties of each dataset. For instance, the correlation time characterizes the temporal dependence structure of colored noise, and the pulse statistics window estimates the occurrence rate of Poisson events. All processing follows the principle of retaining the statistical characteristics of stochastic processes. This ensures that different datasets still reflect their original stochastic dynamic properties under a unified operator framework. In order to further verify the corresponding relationship between different data sources and random noise models, this study tests the statistical characteristics of the processed time series. For the O-U colored noise scenario, the temporal correlation of the sequence is analyzed and verified by the autocorrelation function; the relevant time parameters are determined according to the correlation attenuation characteristics. For the Lévy-type heavy-tailed noise scenario, in order to verify the heavy-tailed structure, its non-Gaussian property

is analyzed by skewness, kurtosis and tail probability distribution characteristics. For the Poisson impulse noise scenario, the pulse occurrence pattern is statistically analyzed by the event interval, and the Poisson process parameters are estimated via the event frequency. The above statistical test is used to confirm the consistency between the random disturbance features extracted from different datasets and the corresponding noise hypothesis, rather than changing the random dynamic properties of the original data.

In order to obtain the reference PDF needed for model evaluation, this study uses the empirical distribution estimation method based on the processed time series samples to construct the reference PDF. Specifically, in each noise scenario, statistical sampling is carried out according to the corresponding time series state samples; the continuous probability density representation is obtained by kernel density estimation methods, and it is used as the reference distribution in the calculation of KL divergence, Wasserstein distance and other probability consistency indicators. Among them, the reference PDF is only estimated independently by the real time series samples in the verification set; the output results of the USPDO model are not used to participate in the construction. Kernel density estimation uses the Gaussian kernel function, and bandwidth parameters are adaptively determined according to the number of samples and data scale, so as to avoid the influence of manual setting on the reference distribution. The reference PDF is employed to uniformly evaluate the consistency between the model's predicted distribution and the actual statistical distribution under different noise conditions. This method does not presuppose the function form of PDF and can maintain the distribution characteristics of data itself under different noise conditions. GWN corresponds to the distribution broadening characteristics; colored noise reflects the influence of temporal correlation; heavy-tailed noise and impulse noise retain the tail probability and extreme event structure. All processed datasets are divided into 80% for training, 10% for parameter tuning, and 10% for validation. This division supports the verification of PDF dynamics for nonlinear stochastic systems by the USPDO model under various noise excitations. Through the unified time-scale processing and amplitude normalization strategy, data from different sources can be trained within the

same operator learning framework; this reduces the impact of data-scale differences on the model learning process and enhances the model's adaptability across diverse random excitation scenarios.

• The design scheme

In the research design, the USPDO model is verified based on multi-type noise excitation scenarios constructed from the MIT-BIH and UCR subsets. The experiment has two core objectives. First, it tests whether the model can accurately

reconstruct the time-dependent evolution law of PDF under the action of GWN, O-U colored noise, Lévy-type heavy-tailed noise, and Poisson impulse noise; second, it evaluates whether USPDO can reasonably reflect the statistical characteristics of the system's long-term stability, transient escape time, and extreme event occurrence probability under the condition of ensuring the probability distribution's physical constraints. The experimental environment and model hyperparameters are exhibited in Tables 3 and 4 to ensure the reproducibility of the experiment:

Table 3. Experimental environment configuration

Environment type	Equipment	Configuration instruction
Hardware	Operating system	Ubuntu 20.04 LTS
	Central Processing Unit (CPU)	Intel Xeon Silver 4310 @ 2.10GHz×32 cores
	Random Access Memory (RAM)	128 GB DDR4
	Graphics Processing Unit (GPU)	NVIDIA A100 40GB×2
Software	CUDA / cuDNN	CUDA 11.8 / cuDNN 8.4
	Deep learning framework	PyTorch 2.0.1
	Python	3.10.12
	Primary dependency library	NumPy 1.24, SciPy 1.10, Matplotlib 3.7

Table 4. The main hyperparameters of the USPDO model

Hyperparameter categories	Value or range
Learning rate	1×10^{-3} , cosine annealing decreases to 1×10^{-5}
Batch size	64
Number of epochs	500
Optimizer	AdamW, the momentum parameter is set to 0.9, and the squared gradient accumulation decay factor is set to 0.999.
Activation function	GELU
Residual weight balancing parameter	[0.1, 1.0] adaptive adjustment
Number of sampling points (spatial-temporal domain)	2,000 points per curve
Non-local integral sampling	128
Loss function	Residual loss + KL divergence regularization term
Learning rate scheduling	CosineAnnealingLR

Note: In the loss function, during training, a combination of weighted residual loss and KL divergence regularization term is adopted. The total loss is composed of the weighted sum of the residuals of each operator and the superposition of the KL regularization term.

For Objective 1 (PDF dynamics reconstruction), KL divergence and Wasserstein Distance (WD) are used as comparison indicators with the true distribution. This ensures that the probability distribution output by the model is consistent with the reference in terms of global shape.

For Objective 2 (physical consistency and rare event

capture), Negative Log-Likelihood (NLL) and Quantile Error (QE) are adopted to characterize the tail behavior of the distribution. Besides, Rare Event Probability Error (REPE) is employed to measure the model's fitting accuracy for the occurrence probability of extreme events. Table 5 summarizes the entire verification indicator system:

Table 5. Verification indicator system based on the USPDO model

Objective category	Verification indicators	Explanation	Acceptable error
Objective 1: Accuracy of PDF dynamics reconstruction	KL	It measures the difference between the predicted probability distribution and the true distribution in terms of overall shape, reflecting the model's information retention ability when reconstructing the PDF; the smaller the value, the higher the distribution consistency.	Under the conditions of this experiment, when the KL value is less than about 0.1, it is considered that the predicted distribution has good overall consistency with the reference distribution. (KL<0.1)
	WD	It evaluates the overall deviation of the model output distribution from the reference	In the numerical experiments of this study, when WD is less than about

		distribution in probability space, which can intuitively reflect the temporal accuracy of the PDF dynamics analysis trajectory.	0.15, the probability distance between the predicted distribution and the reference distribution is within an acceptable range. (WD<0.15)
Objective 2: Physical consistency and rare event capture	NLL	It is used to assess the physical consistency and likelihood fitting accuracy of the model output PDF; the smaller the value, the better the predicted distribution satisfies non-negativity and normalization constraints.	At the experimental scale of this study, when NLL is less than about 0.4, the model is generally considered to have good probabilistic fitting ability for the observed samples. (NLL<0.4)
	QE	It reflects the error in quantile locations between the predicted distribution and the true distribution, used to characterize the accuracy of the distribution's tail shape and non-Gaussian features.	When QE is less than about 0.1, it indicates that the predicted distribution matches the true distribution well within the main quantile intervals. (QE<0.1)
	REPE	It characterizes the model's prediction error for extreme or rare event probabilities, illustrating the model's reliability in capturing low-probability, high-impact events.	In this study experiment, it is considered that the model can accurately depict the probability of rare events when REPE is less than about 0.05. (REPE<0.05)

Note: Let the true probability density function be $p(x, t)$ and the model-predicted PDF be $\hat{p}(x, t)$. The KL divergence measures the information difference between two probability distributions, and its mathematical form is $KL(p \parallel \hat{p}) = \int p(x, t) \ln \frac{p(x, t)}{\hat{p}(x, t)} dx$. The Wasserstein distance describes the optimal transport distance of probability mass in the state space, expressed as $(p, \hat{p}) = \inf_{\gamma \in \Pi(p, \hat{p})} \int |x - y| d\gamma(x, y)$. $\Pi(p, \hat{p})$ denotes the set of joint distributions with marginal distributions p and $\hat{p}$. The negative log-likelihood is defined as $NLL = -\frac{1}{N} \sum_{i=1}^N \log \hat{p}(x_i)$; x_i represents an observed sample. The quantile error is represented by the difference between the quantile functions of the predicted and reference distributions at a given probability level α , namely $QE = |Q_{\hat{p}}(\alpha) - Q_p(\alpha)|$. The rare-event probability error compares the probability differences of the two distributions above the threshold τ , and the expression is $REPE = |P_{\hat{p}}(X > \tau) - P_p(X > \tau)|$. These indicators evaluate the PDF's prediction quality from five perspectives: information divergence, probability transport distance, sample likelihood consistency, quantile position deviation, and extreme-event probability estimation.

3.2 Analysis of Parameter Sensitivity and Uncertainty Propagation

To evaluate the impact of modeling assumptions and key parameters in the USPDO framework on PDF reconstruction results, this study first conducts parameter sensitivity analysis and uncertainty propagation research. The analysis focuses on key parameters in random excitation characteristics, including the characteristic exponent α of stable

Lévy noise, the occurrence rate λ of the Poisson impulse process, and the correlation time τ of O-U colored noise. While keeping the network structure, training strategy, and sampling scale unchanged, a single parameter is perturbed at a time, and the remaining parameters are fixed at baseline values. By comparing the overall shape, tail distribution, and rare event probability of the obtained PDF under different parameter settings, the model's response characteristics and stability to parameter changes are qualitatively evaluated. The final sensitivity comparison results are shown in Table 6:

Table 6. Sensitivity comparison of PDF reconstruction results under different parameter values

Parameter type	Value range	PDF main body shift	Tail decay change	Rare event probability change
α	1.2/1.5/1.8	Slight	Significant	Significant
λ	0.2/0.5/0.8	Moderate	Moderate	Significant
τ	0.1/0.5/1.0	Significant	Weak	Weak

Note: The qualitative descriptions "slight", "moderate", "significant", and "weak" are used to characterize the magnitude of the impact of parameter perturbation on the statistical characteristics of the PDF relative to the baseline parameter conditions. "PDF main body shift" is measured by the relative change in the mean value and main peak position of the PDF; "tail decay change" is evaluated by the change in the slope of the probability density in the high quantile interval under a logarithmic coordinate; "rare event probability change" is compared by the relative change in the probability of events above a preset threshold. Among them, "slight" corresponds to a relative change of less than approximately 10%,

"moderate" corresponds to approximately 10%–30%, "significant" corresponds to more than approximately 30%; "weak" indicates that the impact of the parameter on the corresponding statistic is at a low level. The value range of each parameter is selected from the effective interval under the constraints of model numerical stability and physical consistency.

The results show that different parameters have distinct impacts on the dynamic characteristics of the PDF. Among them, the Lévy exponent α has the most significant influence on the distribution tail shape and rare event probability. As α decreases, the heavy-tailed characteristic of the distribution strengthens, and the probability of high - amplitude events increases significantly; changes in the Poisson impulse rate λ mainly affect the multi-peak structure

of the PDF and the frequency of extreme jump events; the O-U noise correlation time τ mainly affects the diffusion speed and peak position of the PDF main body, with a relatively weak impact on tail behavior. Overall, within a reasonable range of parameter perturbations, the USPDO model can stably capture the PDF dynamics analysis trend and key statistical characteristics; the proposed framework has good robustness and interpretability under parameter uncertainty.

3.3 Verification and analysis based on the USPDO Model

• Reconstruction analysis of PDF dynamics

The reconstruction analysis results of PDF dynamics for the two datasets are illustrated in Table 7 and Figure 1, respectively:

Table 7. Reconstruction and analysis results of PDF dynamics of the MIT-BIH dataset

Time (min)	KL result			WD result		
	Baseline wander noise	Myoelectric interference noise	Electrode motion noise	Baseline wander noise	Myoelectric interference noise	Electrode motion noise
1	0.013	0.015	0.018	0.024	0.025	0.029
2	0.014	0.012	0.019	0.029	0.035	0.039
3	0.011	0.018	0.018	0.026	0.026	0.039
4	0.015	0.018	0.024	0.032	0.036	0.042
5	0.014	0.019	0.023	0.032	0.038	0.045
6	0.017	0.017	0.026	0.03	0.037	0.041
7	0.015	0.016	0.026	0.037	0.033	0.047
8	0.022	0.022	0.029	0.035	0.039	0.052
9	0.017	0.025	0.026	0.038	0.046	0.045
10	0.024	0.028	0.033	0.046	0.047	0.05
11	0.022	0.029	0.032	0.050	0.046	0.063
12	0.026	0.023	0.034	0.045	0.055	0.057
13	0.027	0.029	0.033	0.047	0.047	0.058
14	0.027	0.029	0.034	0.049	0.052	0.063
15	0.027	0.029	0.036	0.049	0.051	0.065
16	0.030	0.032	0.039	0.048	0.059	0.064
17	0.027	0.031	0.040	0.058	0.061	0.067
18	0.026	0.034	0.045	0.057	0.063	0.072
19	0.032	0.034	0.045	0.059	0.064	0.071
20	0.034	0.038	0.044	0.064	0.062	0.075
21	0.037	0.037	0.042	0.063	0.068	0.076
22	0.035	0.037	0.047	0.065	0.071	0.078
23	0.035	0.041	0.045	0.064	0.068	0.08
24	0.034	0.042	0.046	0.072	0.077	0.076
25	0.039	0.046	0.053	0.067	0.076	0.086
26	0.042	0.043	0.050	0.073	0.077	0.084
27	0.042	0.046	0.053	0.074	0.079	0.087
28	0.043	0.046	0.056	0.075	0.083	0.089
29	0.042	0.050	0.054	0.075	0.087	0.096
30	0.041	0.051	0.055	0.079	0.089	0.094

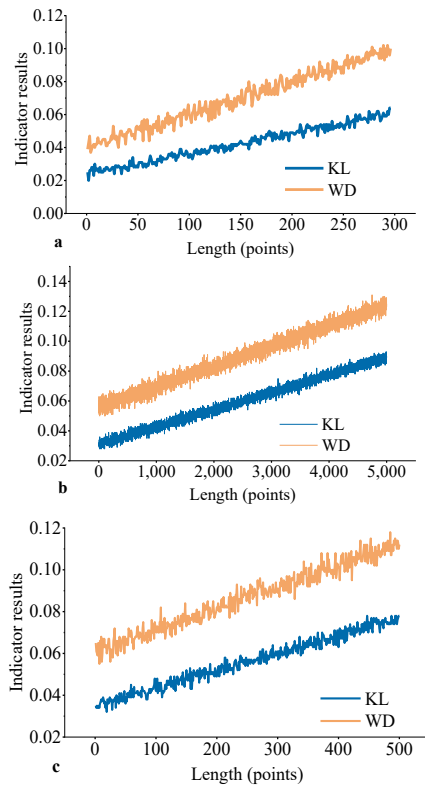


Figure 1: Reconstruction and analysis results of PDF dynamics of the UCR dataset ((a): Nonlinear Gas Furnace; (b): ECG5000; (c): Earthquakes)

Note: In Figure 1, the horizontal axis represents the time series sampling points Length (points), and the vertical axis denotes the Error metric value; the blue curve is KL divergence, and the orange curve refers to WD.

In Table 7, the USPDO model exhibits good PDF reconstruction capability under the three types of noise conditions in the MIT-BIH dataset. For baseline wander noise, KL increases slowly from 0.013 at the 1st minute to 0.041 at the 30th minute, and WD rises from 0.024 to 0.079. The overall growth is stable, indicating that the model can effectively track the distribution changes under low-frequency disturbances. In the myoelectric interference scenario, KL reaches 0.028 at the 10th minute, and WD is close to 0.047, reflecting the challenge of high-frequency components to distribution reconstruction; however, the error is still within an acceptable range. Under electrode motion noise, KL is 0.055, and WD is approximately 0.094 at the 30th minute, substantially exceeding those of baseline wander; the USPDO model has some deviation when handling non-Gaussian pulses. In general, USPDO can still maintain the main evolving characteristics of the distribution shape across different noise excitations, and its stability and accuracy are more prominent, especially under Gaussian and colored noise conditions.

In the UCR dataset shown in Figure 1, the O-U colored noise of Nonlinear Gas Furnace presents a gentle error growth: KL increases from 0.024 to

0.061, and WD rises from 0.041 to 0.099. The USPDO model can well capture the temporal dependence of the relevant stochastic process. In the ECG5000 data, Lévy-type heavy-tailed noise leads to a relatively high error level: KL reaches 0.089 at 5000 points, and WD reaches 0.124; this reflects the model's difficulty in dealing with extreme fluctuations. The Earthquakes data exhibit Poisson impulse characteristics: KL and WD rise from 0.035 to 0.077 and from 0.061 to 0.112. Although the error increases gradually, it remains within a reasonable range overall. Comprehensively, the USPDO model performs best under GWN and O-U colored noise; its error increases under non-Gaussian pulse and heavy-tailed noise conditions, but it still maintains effective reconstruction of the distribution shape. This reveals that the model has strong adaptability to different noise types.

• Analysis of physical consistency and rare events

The analysis results of physical consistency and rare events for the two datasets are depicted in Figures 2 and 3:

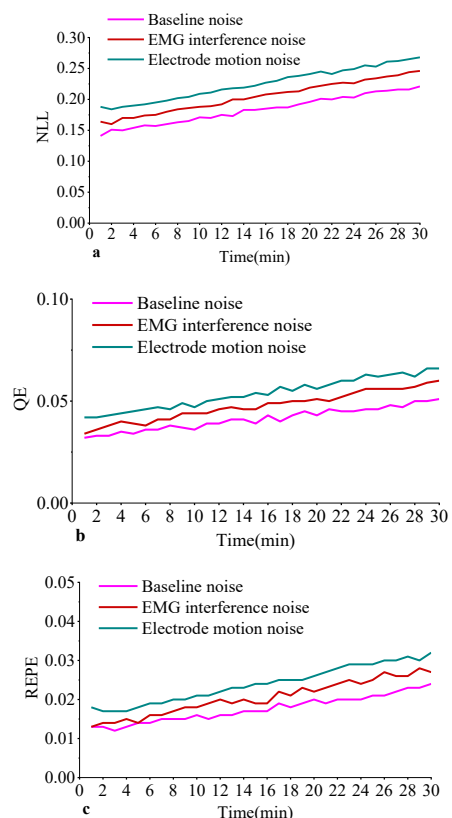


Figure 2: Physical consistency and rare event analysis results of the MIT-BIH dataset ((a): NLL; (b): QE; (c): REPE)

Note: In Figure 2, the horizontal axis refers to the evolution duration Time (min); the vertical axis represents the NLL, QE, and REPE in sequence. The three curves correspond respectively to the working conditions of Baseline noise, EMG interference noise and Electrode motion noise.

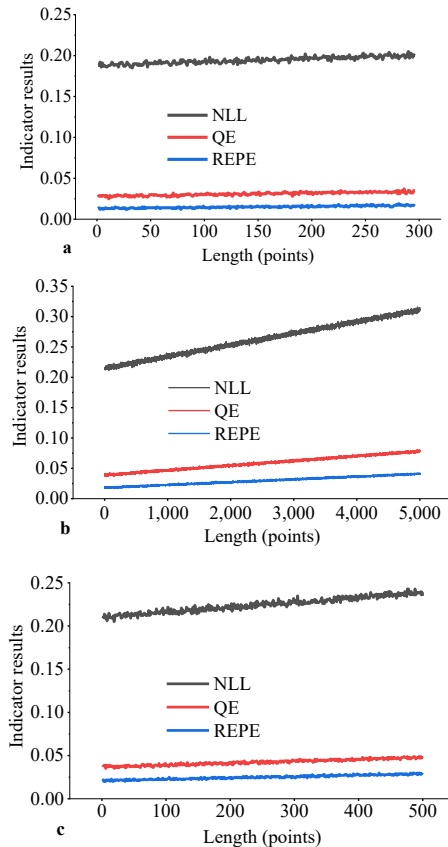


Figure 3: Analysis results of physical consistency and rare events of the UCR dataset ((a): Nonlinear Gas Furnace; (b): ECG5000; (c): Earthquakes)

Note: In Figure 3, the horizontal axis denotes the time series sampling points Length (points), and the vertical axis is the Error metric value; the black curve represents NLL, the red curve is QE, and the blue curve indicates REPE.

The USPDO model can effectively maintain the physical rationality of the probability distribution across different noise conditions and characterize the occurrence characteristics of extreme events to a certain extent. Taking the MIT-BIH data in Figure 2 as an example, under baseline wander noise, the NLL value only increases gradually from 0.145 to 0.221 in the 1–30-minute interval, and its magnitude remains stable; the QE is always lower than 0.05. This reveals that the model can grasp the PDF constraints under stable Gaussian disturbances. Under the condition of myoelectric noise, due to the enhanced high-frequency components, the error indicators increase accordingly. At the 30th minute, the NLL reaches 0.247, and the QE is approximately 0.059; this means the difficulty of tail fitting increases slightly. The electrode motion noise shows stronger non-Gaussian pulse characteristics, and the REPE increases gradually from 0.016 to 0.032. Although USPDO has deviations in estimating the probability of sudden events, the overall trend can reflect the rare event occurrence pattern of the actual distribution.

In Figure 3, the UCR dataset further demonstrates the USPDO model's generalization ability in

nonlinear stochastic systems. In the Gas Furnace sequence, the NLL ranges from approximately 0.188 to 0.201 with an extremely small range, which shows that USPDO has good stability under O-U colored noise. In the ECG5000 data, the NLL reaches 0.311, and the QE is approximately 0.078 at 5000 points, suggesting that the Lévy-type heavy-tailed distribution poses a great challenge to the model. Still, the REPE is controlled within 0.041, so the model can still capture the probability of extreme fluctuations. In the Earthquakes sequence, the REPE remains stable between 0.021 and 0.029, indicating that the model can effectively characterize the statistical characteristics of rare events under Poisson impulse noise. Overall, the USPDO model maintains high physical consistency under Gaussian and correlated noise. Although the error increases under complex non-Gaussian and pulse noise, USPDO still has reliable prediction capabilities for tail distributions and extreme events.

• Supplementary verification analysis of diversified datasets

To further verify the USPDO model's generalization and robustness under different random excitation environments, the study selects six public datasets from the Kaggle platform. These datasets cover six typical scenarios: GWN, O-U colored noise, mixed noise, non-stationary colored noise, Poisson impulse noise, and Lévy-type heavy-tailed noise. Specific data sources include: 5G SINR Estimation Dataset (GWN, <https://www.kaggle.com/datasets/programmer3/5g-sinr-estimation-dataset>); Aging Bridge SHM Time-Series Dataset (O-U colored noise, <https://www.kaggle.com/datasets/programmer3/aging-bridge-shm-time-series-dataset>); Mechanical Gear Vibration Dataset (Gaussian-Poisson mixed noise, <https://www.kaggle.com/datasets/hieudaotrung/gear-vibration>); Electric Motor Vibrations Dataset (non-stationary colored noise, <https://www.kaggle.com/datasets/amirberenji/electric-motor-vibrations-dataset>); Seismic-Benchmark Dataset (Poisson impulse noise, <https://www.kaggle.com/datasets/elemento/earthquakes-timeseries-analysis-nefs-sac>); All the Earthquakes Dataset (1990-2023) (Lévy-type heavy-tailed noise, <https://www.kaggle.com/datasets/alessandrobellio/the-ultimate-earthquake-dataset-from-1990-2023>). All data are uniformly sampled ($\Delta t=0.01s$), detrending, and normalized; they are divided into an 80% training set, a 10% parameter tuning set, and a 20% validation set. The model evaluates PDF dynamic performance through five indicators: KL, WD, NLL, QE, and REPE. The results are demonstrated in Table 8:

Table 8. The verification results of the USPDO model on multi-source public noise datasets

Dataset	Noise type	KL	WD	NLL	QE	REPE
5G SINR Estimation	GWN	0.022	0.041	0.179	0.043	0.018
Aging Bridge SHM	O-U colored noise	0.024	0.042	0.183	0.046	0.019
Mechanical Gear Vibration	Gaussian-Poisson mixed noise	0.033	0.055	0.201	0.05	0.027
Electric Motor Vibrations	Non-stationary colored noise	0.029	0.049	0.193	0.052	0.024
Seismic-Benchmark	Poisson impulsive noise	0.038	0.065	0.215	0.059	0.032
All the Earthquakes	Lévy-type heavy-tailed noise	0.042	0.073	0.228	0.063	0.037

Note: The results of all indicators in Table 8 are the mean of five independent repeated experiments (for details of the standard deviation, please refer to the appendix).

In Table 8, the KL and WD of the USPDO model remain at low levels across different noise conditions. Among them, the results under GWN and O-U color noise are the best, indicating that the model can accurately reconstruct the PDF dynamics analysis of stationary diffusion systems. For mixed and non-stationary noise scenarios such as mechanical and motor vibrations, the error increases slightly, but the overall form is consistent; this reflects the good adaptability of USPDO to non-stationary and correlated random processes. In seismic data, Poisson impulse and Lévy heavy-tailed noise lead to increased tail probability deviation, but NLL and REPE remain in stable ranges; USPDO can reliably characterize the probability of rare events and extreme fluctuations. In summary, the performance of USPDO on multi-source real data verifies its unified modeling logic from noise excitation to probability response; meanwhile, it can be widely applied to PDF dynamic research of complex stochastic systems such as structural health monitoring, signal recognition, and seismic risk analysis.

• Comparative analysis with classic and learning-based PDF solvers

In response to the lack of quantitative comparison between USPDO and existing PDF dynamics analysis methods, this study introduces representative classical numerical methods and learning-based methods as benchmarks without changing the aforementioned dataset and noise settings. Meanwhile, a comparative analysis is conducted on the performance of PDF dynamics analysis. Specifically, three types of methods are selected: (i) the finite difference method based on the FPE (Crank-Nicolson, CN), as a classical deterministic PDF solver; (ii) Monte Carlo SDE trajectory simulation, which obtains the reference PDF through a large number of sample statistics; (iii) the Physics-Informed Neural Network (PINN), as a typical learning-based PDE solving method. All methods are evaluated for the same low-dimensional nonlinear stochastic system and noise configuration as in Section 4.4; the PDF is calculated under a unified time window and spatial discretization scale.

The comparison indicators are consistent, using KL, WD, REPE, and the computation time for a complete PDF dynamics analysis as measurement standards to ensure comparability between different methods. The results are shown in Table 9:

Table 9. Quantitative comparison results between USPDO and different PDF dynamics analysis methods

Method	KL	WD	REPE	Calculation time (relative)
Crank-Nicolson (FPE)	0.021	0.038	0.029	High
Monte Carlo SDE	0.018	0.035	0.026	Very high
PINN	0.031	0.057	0.041	Medium
USPDO	0.024	0.042	0.028	Low

Note: The KL, WD, and REPE indicators are the statistical averages of the results obtained by modeling the PDF dynamics of the same nonlinear stochastic system under various random excitation conditions considered in Section 4.4 (including GWN, O-U colored noise, Lévy-type heavy-tailed noise, and Poisson impulse noise). Specifically, each method first calculates the time average of the corresponding indicators within a unified time interval; it then takes the average of the results for different noise types to reflect the method's overall modeling performance under multi-noise conditions. The computation time is the relative time consumption to complete a full PDF dynamics analysis process, normalized by the Monte Carlo method. The computation time is defined as $T_{rel} = T_{method}/T_{MC}$; T_{MC} denotes the average computation time of the Monte Carlo SDE method under the same system, noise configuration, and time window. According to the numerical interval of the relative computation time, the computation cost is qualitatively classified: $T_{rel} \geq 1.0$ (very high), $0.5 \leq T_{rel} < 1.0$ (high), $0.1 \leq T_{rel} < 0.5$ (medium), $T_{rel} < 0.1$ (low). Moreover, the KL, WD, and REPE results in this table are the mean values of five independent repeated experiments, and the detailed standard deviations are provided in the appendix.

In Table 9, different methods exhibit significant differences in PDF dynamics modeling performance. The difference in the KL indicator between USPDO (0.024) and the Crank–Nicolson method (0.021) is approximately 0.003. The difference in the WD indicator between USPDO (0.042) and the same method (0.038) is about 0.004. The REPE indicator differs from the Monte Carlo method by roughly 0.002. The overall errors are of the same order of magnitude. In-depth analysis shows that the Crank–Nicolson-based finite difference method has high numerical accuracy under low-dimensional and regular grid conditions. However, its computation cost increases markedly with spatial dimension and time step, and its extensibility to non-Gaussian heavy-tailed noise and non-local operators is limited. Monte Carlo SDE simulation can theoretically approximate the true distribution. But it is highly sensitive to the number of samples, requiring a large number of trajectories for rare event probability estimation, resulting in significant computation overhead. The PINN method has certain flexibility in unifying PDE constraints and data-driven approaches. Nevertheless, in long-term evolution and tail regions, it is difficult to simultaneously ensure convergence stability and PDF normalization constraints. In contrast, USPDO keeps the accuracy of PDF reconstruction equal to that of classical methods; meanwhile, it avoids the need for step-by-step numerical iteration or large-scale sampling through operator layer learning; this shows more stable characteristics in terms of computational efficiency and adaptability to noisy conditions. Compared with PINN-type methods that directly solve partial differential equations using physical information residuals, USPDO does not solve a single partial differential equation. Instead, it learns the PDF dynamic operator mapping parameterized jointly by system parameters and random excitations. Thus, it does not require reconstruction of the solution process under varying noise conditions and enables unified prediction within the same operator structure. Specifically, traditional numerical methods for the Fokker–Planck equation require stepwise discrete iteration over the entire spatiotemporal grid. Monte Carlo methods rely on a large number of random trajectory samples to obtain stable statistical results. In contrast, by learning the mapping between random excitations and PDF dynamics analysis operators, USPDO yields the corresponding probability distribution with only one forward computation during inference; this remarkably reduces computational cost while maintaining accuracy. The comparison results indicate that USPDO provides a trade-off among accuracy, computation cost, and extensibility; this makes it suitable for PDF dynamics modeling scenarios under multi-noise excitation.

• A comparative analysis with neural operator learning methods

In order to further analyze the differences between USPDO and neural operator methods in PDF dynamic modeling tasks, this study selects representative DeepONet and Fourier Neural Operator (FNO) as comparative models. Different from the comparison between the classical PDF solver and the physical constraint solving method in Section 3.3.4, this section mainly focuses on the modeling ability of learning operator models to the evolution mapping relationship of PDF under the condition of random parameter changes. Among them, DeepONet learns the mapping relationship between the input and output functions through branch networks and backbone networks; FNO captures the global association in function spaces via frequency-domain feature transformation. All models adopt the same data division method, noise excitation conditions and training settings; they are tested on the multi-noise PDF dynamics task constructed by MIT-BIH and UCR datasets. The evaluation indicators include KL, WD and REPE; they respectively measure the prediction ability of the model for the overall distribution difference of PDF, probability transmission error and tail event probability. All the results are the Mean±standard deviation (SD) of five independent repeated experiments, and the experimental results are listed in Table 10:

Table 10. Performance comparison results of USPDO and different models in PDF dynamics modeling tasks

Method	KL↓	WD↓	REPE↓
DeepONet	0.046±0.0048	0.086±0.0075	0.044±0.0051
FNO	0.041±0.0042	0.078±0.0068	0.039±0.0047
USPDO	0.024±0.0028	0.042±0.0039	0.028±0.0030

It can be found from Table 10 that both DeepONet and FNO can learn the mapping relationship between the input conditions of random systems and the PDF output; nevertheless, there are still some errors under the multi-random excitation conditions. DeepONet relies on the branch and backbone networks to jointly represent the input and output functions, and it can effectively fit the mapping relationship of general function spaces. However, for the PDF evolution process with various noise mechanisms, its adaptability to random dynamic structural changes is limited. FNO extracts global features through Fourier transform; this further reduces the distribution error in PDF dynamic modeling. However, FNO mainly focuses on the mapping relationship between input and output functions and does not explicitly integrate the dynamic operator structures corresponding to different random noises.

In contrast, USPDO introduces random excitation parameters into PDF dynamic modeling processes through a conditional input mechanism; it also combines drift diffusion, fractional Lévy and Poisson jump operator constraints, so that this model can simultaneously learn the change patterns of data distribution and stochastic dynamic evolution mechanism. Under the same test conditions, USPDO obtains lower errors in KL, WD and REPE, which are about 0.017, 0.036 and 0.011 lower than those of FNO; this indicates that USPDO can learn the PDF mapping relationship and keep the consistency of the evolution law of random systems by using physical constraints. The experiment further shows that the main difference between USPDO and traditional neural operator methods is that it does not simply learn the function space mapping; instead, it realizes the PDF evolution modeling for multi-noise conditions under the constraint of stochastic dynamic operators.

3.4 Discussion

From the numerical results, it can be seen that the USPDO model can effectively describe the PDF dynamic characteristics of nonlinear systems under different noise excitations. The scenario error of heavy-tailed noise and impulse noise is caused by the sampling deviation caused by non-local fractional order and Monte Carlo approximation of the transition operator. Gaussian diffusion operator only contains local differential terms, and there is no sampling approximation error, so the fitting accuracy is higher. Figures 2 and 3 reveal that the stable changes of NLL and QE reflect the physical consistency of the model's output PDF under non-negativity and normalization constraints; the low REPE indicates that the USPDO has credibility in the statistical estimation of rare event probabilities. According to the quantitative results in Table 8, it can be further observed that USPDO achieves comparable performance with classical numerical methods on KL, WD, and REPE indicators; however, it exhibits noticeably lower computational cost than Monte Carlo trajectory simulation. The reason for this difference is that USPDO directly learns the mapping between random excitations and probability distribution evolution through operator-level modeling; thus, it avoids the dependence on massive random trajectory samples in the Monte Carlo method. Meanwhile, compared with PINN methods that directly approximate the solution of partial differential equations, USPDO explicitly represents different stochastic mechanisms in PDF dynamics analysis as operator structures; besides, it embeds noise parameters through a conditional input mechanism, so that the model can adapt to various random excitation conditions in a unified framework. Such operator-level modeling preserves

probability conservation and non-negativity constraints; at the same time, it improves the model's ability to represent PDF dynamics features under complex noise conditions. On the whole, the model accurately reproduces the spatiotemporal characteristics of probability flow under various noise drives. At the engineering application level, such PDF dynamic descriptions can be further used to evaluate the impact of stochastic perturbations on system operational reliability. For example, real-time estimation of the evolutionary trend of state probability distributions provides a statistical basis for risk assessment, trajectory planning, or control strategy design in complex systems.

Based on these results, the PDF analysis logic of the USPDO model can be understood as a mapping learning mechanism from noise excitation to probability response. In other words, it realizes the unified modeling of the system's steady-state and transient behaviors by learning the dynamic impact of different noise characteristics on the probability manifold. This logic has clear application significance in engineering and natural systems. In structural dynamics, it can be used to identify vibration distributions and energy diffusion paths under Gaussian and non-Gaussian disturbances; in signal and seismic sequence analysis, the system's transition from stationary to non-stationary states can be identified through the PDF broadening rate, thereby evaluating the probability of system instability or extreme events. Comprehensive analysis shows that USPDO can quantify the variation laws of system probability characteristics under different noise conditions; it is also a general tool for multi-noise dynamic modeling in nonlinear stochastic systems, providing an effective solution for system stability prediction and risk characterization in complex environments.

4. Conclusions

On the whole, the proposed USPDO model demonstrates effective modeling capabilities for the PDF dynamics of nonlinear stochastic systems under multiple types of random excitations. Through system verification in different noise scenarios, the results show that this model can maintain the overall stability of the probability distribution; it also has good characterization accuracy for tail structure and rare events. Compared with the traditional low-order statistical analysis methods, it can completely capture the statistical characteristics of distribution tails and rare events. In addition, it can provide a general modeling scheme for multi-noise coupled stochastic systems.

It should be noted that although USPDO generally performs stably under various noise conditions, its modeling accuracy still shows significant differences under different random excitation scenarios.

Combined with the numerical results in Section 4, it can be observed that under Lévy-type heavy-tailed noise and Poisson impulse perturbation conditions, as the time scale lengthens or the tail probability region expands, KL, WD, and REPE all show a relatively obvious increasing trend. The current USPDO framework still has certain limitations in characterizing PDFs of extremely heavy-tailed distributions, multi-peak jump structures, and high quantile intervals. Especially in the accurate modeling of low-probability high-impact events, its prediction error is more prominent compared to Gaussian and O-U colored noise scenarios. In addition, as the Lévy stability index decreases or the Poisson pulse rate increases, the PDF tail decay characteristics and rare-event probability variations become more pronounced; this increases the difficulty of accurate modeling in the distribution tail region.

In addition, due to the large differences in sampling rate, time scale, and non-stationary characteristics among different datasets, the model's parameter sharing and operator expression capabilities are still limited when generalizing across data sources and application scenarios. Meanwhile, this study verifies the effectiveness of USPDO in PDF dynamics reconstruction mainly from the perspective of numerical modeling. The stability analysis and convergence theory of the operator learning process still require further investigation. For instance, theoretical characterizations of operator approximation error and training convergence properties need to be established under a more general framework of stochastic dynamical systems. This cross-domain inconsistency is particularly obvious in complex engineering signals and natural event sequences; this may lead to estimation biases in PDF tail behavior and transient jump probabilities. In addition, since the model adopts Monte Carlo sampling approximation when solving non-local fractional operators and jump operators, its numerical accuracy is related to the sampling size; this may introduce additional approximation errors in cases of small tail probabilities or high jump frequencies.

In view of the above problems, the follow-up research will be further expanded from the following aspects under the premise of maintaining the unified operator framework. First, for heavy-tailed and impulse noise conditions, this study introduces residual weighting or quantile constraint mechanisms more sensitive to tail regions to enhance the ability to characterize extreme event probabilities. Second, the study enhances the model's adaptability to non-stationary random excitation by combining a multi-scale or time-dependent operator structure. Third, it improves the model's generalization stability in different application fields through cross-dataset joint

training or parameter adaptive strategy. Besides, as the system dimension increases further, the spatiotemporal sampling size and non-local integral computation required by the PDF dynamics analysis equation may increase accordingly. Consequently, future work will combine dimensionality reduction strategies or more efficient operator approximation methods to improve the model's scalability.

Data Availability

This study selects two open and authoritative datasets, all of which can be directly downloaded via the provided links. The first dataset is the MIT-BIH Noise Stress Test Database from PhysioNet (DOI: <https://doi.org/10.13026/C2HS3T>). The second dataset is a subset of the UCR Time Series Classification Archive from hexagon-ml (https://www.cs.ucr.edu/~eamonn/time_series_data_2018/)—Nonlinear Gas Furnace, electrocardiogram (ECG)5000, Earthquakes.

Code availability

Details of the implementation code for the proposed USPDO model are provided in the appendix. The code fixes the random seed (SEED=42) and embeds a grid-search mechanism for hyperparameter optimization. The core hyperparameters correspond to the settings in Table 3: learning rate starting at 1×10^{-3} and annealed to 1×10^{-5} via CosineAnnealingLR; batch size: 64; training epochs: 500; AdamW optimizer parameters: $\beta_1=0.9/\beta_2=0.999$; GELU activation function; 2000 spatiotemporal sampling points; 128 sampling points for non-local integrals, etc. The loss function integrates residual loss and KL divergence regularization; the residual weights are adaptively adjusted from 0.1 to 1.0.

References

- [1] Wang, L., Li, X., Peng, J., et al. (2024). Response analysis of projectile system under Gaussian noise excitation using path integral method. *International Journal of Aerospace Engineering*, 2024(1), 7720612. <https://doi.org/10.1155/ijae/7720612>
- [2] Pires, C. A., Docquier, D., & Vannitsem, S. (2024). A general theory to estimate information transfer in nonlinear systems. *Physica D: Nonlinear Phenomena*, 458, 133988. <https://doi.org/10.1016/j.physd.2023.133988>
- [3] Angelani, L., De Gregorio, A., & Garra, R. (2025). Generalized time-fractional kinetic-type equations with multiple parameters. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 35(2), 17. <https://doi.org/10.1063/5.0243533>
- [4] Sykora, H. T., Kuske, R., & Yurchenko, D. (2024). Stochastic dynamics of mechanical systems with

- impacts via the step matrix multiplication based path integration method. *Nonlinear Dynamics*, 112(11), 9095–9116. <https://doi.org/10.1007/s11071-024-09513-y>
- [5] Bai, Y., Wang, L., Xu, W., et al. (2025). Stochastic dynamics analysis of nonlinear systems with time delay based on path integration method. *Journal of Sound and Vibration*, 615, 119169. <https://doi.org/10.1016/j.jsv.2025.119169>
- [6] Er, G. K., Luo, J., & Lu, V. P. (2025). Evolutionary stochastic characteristics of nonlinear oscillator with one side barrier due to multiple modulated Gaussian white noise. *Communications in Nonlinear Science and Numerical Simulation*, 145, 108696. <https://doi.org/10.1016/j.cnsns.2025.108696>
- [7] Spanos, P. D., & Pomaro, B. (2025). Stochastic response spectrum determination of nonlinear systems endowed with fractional derivative elements. *Nonlinear Dynamics*, 113, 14791–14805. <https://doi.org/10.1007/s11071-025-11028-z>
- [8] Tang, Y., Peng, Y., He, G., et al. (2023). P-bifurcation analysis for a fractional damping stochastic nonlinear equation with Gaussian white noise. *Fractal and Fractional*, 7(5), 408. <https://doi.org/10.3390/fractalfract7050408>
- [9] Koverda, V. P., & Skokov, V. N. (2025). An analytical approach for describing self-similar random processes with positive variables. *Physica A: Statistical Mechanics and its Applications*, 673, 130684. <https://doi.org/10.1016/j.physa.2025.130684>
- [10] Chen, J., Sun, T., Spanos, P. D., et al. (2025). Efficient stochastic response analysis of high-dimensional nonlinear systems subject to multiplicative noise via the DR-PDEE. *Journal of Computational Physics*, 531, 113929. <https://doi.org/10.1016/j.jcp.2025.113929>
- [11] Peng, J., Wang, L., Wang, B., et al. (2024). A new path integration method for the stochastic system under Poisson white noise excitation based on a probability mapping. *Journal of Sound and Vibration*, 571, 118037. <https://doi.org/10.1016/j.jsv.2023.118037>
- [12] Landgraf, D., Völz, A., Berkel, F., et al. (2023). Probabilistic prediction methods for nonlinear systems with application to stochastic model predictive control. *Annual Reviews in Control*, 56, 100905.
- [13] Gul, N., Mashwani, W. K., Aamir, M., et al. (2023). Optimal model selection for k-nearest neighbours ensemble via sub-bagging and sub-sampling with feature weighting. *Alexandria Engineering Journal*, 72, 157–168. <https://doi.org/10.1016/j.aej.2023.03.075>
- [14] Ali, A., Khan, Z., & Khan, D. M., et al. (2024). An optimal random projection k nearest neighbors ensemble via extended neighborhood rule for binary classification. *IEEE Access*, 12, 61401–61409. <https://doi.org/10.1109/ACCESS.2024.3392729>
- [15] Ali, A., Khan, Z., & Aldahmani, S. (2025). Optimized feature selection in high-dimensional gene expression data using weighted differential gene expression analysis. *Applied Soft Computing*, 180, 113329. <https://doi.org/10.1016/j.asoc.2025.113329>
- [16] Khan, Z., Ali, A., Aldahmani, S., et al. (2025). Customer lifetime value modelling via two stage selected trees ensembles. *IEEE Access*, 13, 115236–115247. <https://doi.org/10.1109/ACCESS.2025.3584709>
- [17] Raissi, M., Perdikaris, P., Karniadakis, G. E., et al. (2018). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378(C), 686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>
- [18] Yang, L., Meng, X., Karniadakis, G. E., et al. (2021). B-PINNs: Bayesian physics-informed neural networks for forward and inverse PDE problems with noisy data. *Journal of Computational Physics*, 425, 109913. <https://doi.org/10.1016/j.jcp.2020.109913>
- [19] Lu, L., Jin, P., Pang, G., et al. (2021). Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3(3), 218–229. <https://doi.org/10.1038/s42256-021-00302-5>
- [20] Kovachki, N., Li, Z., Liu, B., et al. (2023). Neural operator: Learning maps between function spaces with applications to PDEs. *Journal of Machine Learning Research*, 24(89), 1–97. <http://jmlr.org/papers/v24/21-1524.html>
- [21] Luo, Y., Dang, C., Broggi, M., et al. (2025). Stochastic dynamic response analysis via dimension-reduced probability density evolution equation (DR-PDEE) with enhanced tail-accuracy. *Probabilistic Engineering Mechanics*, 79, 103735. <https://doi.org/10.1016/j.pro bengmech.2025.103735>
- [22] Paul, R., Mauger, J., Bulsara, A. R., et al. (2023). Noise driven coupled nonlinear systems: A “dynamical multilayer perceptron” demonstrating XOR functionality. *IEEE Access*, 11, 144306–144324. <https://doi.org/10.1109/ACCESS.2023.3344161>
- [23] Hägglund, K., & Axell, E. (2021). Adaptive demodulation in impulse noise channels. *IEEE Transactions on Vehicular Technology*, 71(2), 1685–1698. <https://doi.org/10.1109/TVT.2021.3133652>
- [24] Chen, J. B., & Lyu, M. Z. (2024). Probabilistic response determination of high-dimensional nonlinear dynamical systems enforced by parametric multiple Poisson white noises.

- Nonlinear Dynamics, 112(13), 11283–11298. <https://doi.org/10.1007/s11071-024-09592-x>
- [25] Heyes, D. M., Dini, D., & Smith, E. R. (2020). Statistical analysis and molecular dynamics simulations of the thermal conductivity of Lennard–Jones solids including their pressure and temperature dependencies. *Physica Status Solidi (B)*, 257(10), 2000344. <https://doi.org/10.1002/pssb.202000344>
- [26] Cortés, J. C., Romero, J. V., Roselló, M. D., et al. (2024). Probabilistic analysis of the steady state of weakly perturbed linear oscillators subject to a class of Gaussian inputs. *Chaos, Solitons & Fractals*, 187, 115451. <https://doi.org/10.1016/j.chaos.2024.115451>
- [27] Chen, J., Yang, J., Shen, K., et al. (2022). Probability density analysis of nonlinear stochastic dynamics of horizontal axis wind turbine blades. *Ocean Engineering*, 261, 111806. <https://doi.org/10.1016/j.oceaneng.2022.111806>
- [28] Lyu, M. Z., Feng, D. C., Chen, J. B., et al. (2024). A decoupled approach for determination of the joint probability density function of a high-dimensional nonlinear stochastic dynamical system via the probability density evolution method. *Computer Methods in Applied Mechanics and Engineering*, 418, 116443. <https://doi.org/10.1016/j.cma.2023.116443>
- [29] Luong, D., Balaji, B., & Rajan, S. (2022). Structured covariance matrix estimation for noise-type radars. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–13. <https://doi.org/10.1109/TGRS.2022.3184597>
- [30] Goldberger, A. L., Amaral, L. A. N., Glass, L., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23), e215–e220. <https://doi.org/10.1161/01.CIR.101.23.e215>