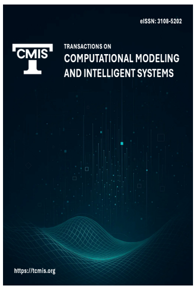

Transactions on
Computational Modeling
and Intelligent Systems

Volume 2, January 2026, Article 10028
https://doi.org/10.65112.tcmis.10028

eISSN: 3108-5202
www.tcmis.org

REVIEW ARTICLE

A selective review of modern stochastic modeling: SDE/SPDE numerics, data-driven identification, and generative methods with applications in biomathematics

Yassine Sabbar[1] and Kottakkaran Sooppy Nisar[2,*]

[1] *IMIA Laboratory, T-IDMS, Department of Mathematics, FST Errachidia, Moulay Ismail University of Meknes, P.O. Box 509, 52000, Boutalamine, Morocco.*
[2] *Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia.*
[*] *Corresponding Author. Email: n.sooppy@psau.edu.sa (K. S. Nisar)*

Article Information

Received: 8 December 2025
Accepted: 18 December 2025
Published 7 January 2026

AMS Classification:
60H10, 60H30, 60J74, 92D30

Abstract

This review maps 2020-2025 developments in stochastic modeling, highlighting non-standard approaches and their applications to biology and epidemiology. It brings together four strands: (1) core models for systems that evolve with randomness; (2) learning key parts of those models directly from data; (3) methods that can generate realistic synthetic data in continuous time; and (4) numerical techniques that keep simulations stable, accurate, and faithful over long runs. The objective is practical: help researchers quickly see what is new, how the pieces fit together, and where important gaps remain. We summarize tools for estimating changing infection or reaction rates under noisy and incomplete observations, modeling spatial spread, accounting for sudden jumps and heavy tails, and reporting uncertainty in a way that is useful for decisions. We also highlight open problems that deserve near-term attention: separating true dynamics from noise when data are irregular; learning spatial dynamics under random influences with guarantees of stability; aligning training with the numerical method used in applications; preserving positivity and conservation in all simulations; reducing cost while controlling error for large studies; estimating rare but important events; and adopting clear, comparable reporting standards. By organizing the field around these aims, the review offers a concise guide to current methods, their practical use, and the most promising directions for future work in biology and epidemiology.

Keywords: Stochastic modeling, stochastic differential equations, stochastic partial differential equations, neural stochastic differential equations, operator learning, numerical methods

Nomenclatures

Abbrev.	Meaning	Abbrev.	Meaning
ABC	Approximate Bayesian Computation	MLE	Maximum Likelihood Estimation
CI	Confidence Interval	MLMC	Multilevel Monte Carlo
CRPS	Continuous Ranked Probability Score	NPE	Neural Posterior Estimation
DSM	Denoising Score Matching	ODE	Ordinary Differential Equation
EnKF	Ensemble Kalman Filter	PDE	Partial Differential Equation
ESS	Effective Sample Size	PINO	Physics-Informed Neural Operator
FD	Finite Difference	PINN	Physics-Informed Neural Network
IMEX	Implicit-Explicit	PSD	Positive Semidefinite
MCMC	Markov Chain Monte Carlo	SDE	Stochastic Differential Equation
SG-MCMC	Stochastic-Gradient Markov Chain Monte Carlo	SPDE	Stochastic Partial Differential Equation
SIR	Susceptible-Infected-Recovered	SMC	Sequential Monte Carlo
SIS	Susceptible-Infected-Susceptible	SRK	Stochastic Runge-Kutta
UQ	Uncertainty Quantification		

1. Introduction

Stochastic modeling offers a mathematically rigorous and physically interpretable framework for analyzing dynamical systems whose evolution is influenced by intrinsic randomness [1], unresolved multiscale interactions [2], or incomplete specification of the governing mechanisms [3]. In contrast to deterministic formulations (where the system trajectory is uniquely determined by a prescribed set of ordinary or partial differential equations [4]) stochastic models incorporate random variables or stochastic processes directly into the governing equations [5]. This probabilistic formulation enables explicit quantification of epistemic and aleatory uncertainty, reproduces variability across independent realizations [6], and captures non-deterministic phenomena, such as intermittency, heavy-tailed fluctuations, or regime switching, that lie beyond the representational capacity of purely deterministic models [7].

For finite-dimensional systems, the canonical representation is the stochastic differential equation (SDE):

$$\mathrm{d}X_t = f(X_t, t)\,\mathrm{d}t + G(X_t, t)\,\mathrm{d}W_t, \tag{1}$$

where $X_t \in \mathbb{R}^d$ denotes the state vector, $f : \mathbb{R}^d \times [0, \infty) \to \mathbb{R}^d$ is the drift function describing the deterministic dynamics, $G : \mathbb{R}^d \times [0, \infty) \to \mathbb{R}^{d\times m}$ is the diffusion coefficient matrix, and W_t is an m-dimensional standard Wiener process modeling temporally continuous Gaussian perturbations [8]. The drift term f typically encodes system-specific physical laws or conservation principles, whereas the diffusion term G parameterizes the intensity, anisotropy, and possible state-dependence of the noise [9]. When the state space is infinite-dimensional, one arrives at stochastic partial differential equations (SPDEs), where the solution $X(t, \xi)$ depends on both temporal (t) and spatial (ξ) variables, and the stochastic forcing may act locally or globally in space [10]. SPDEs arise naturally in the modeling of spatially distributed systems subject to random forcing, such as turbulent flows [11], population dispersal [12], or reaction-diffusion processes with environmental fluctuations [13].

The classical Itô calculus framework has been generalized to accommodate diverse noise structures [14], including: (i) *Jump-diffusion processes* [15], where compensated Poisson random measures introduce finite-activity discontinuities representing abrupt or rare events [16]; (ii) *Lévy-driven systems* [17], which allow for heavy-tailed increments and infinite-activity jump processes [18], capturing burst-like dynamics and extreme variability [19]; and (iii) *Fractional Brownian motion* and *Volterra-type kernels* [20, 21], which encode long-range temporal dependence, self-similarity, and memory effects [22]. In many applications, the diffusion term $G(X_t, t)$ is itself state-dependent, creating a multiplicative noise structure and introducing bidirectional coupling between the system dynamics and the stochastic perturbations [23].

From a statistical perspective, stochastic models serve both as *generative mechanisms* and as *inferential frameworks* [24, 25]. As generative models, they produce synthetic sample paths that preserve empirical distributional properties [26], temporal correlation structures [27], and extreme-event statistics observed in experimental or observational datasets [28]. As inferential frameworks, they provide a principled foundation for recovering latent parameters, hidden states [29], and entire posterior distributions from noisy, incomplete, or irregularly sampled measurements [30].

Parameter estimation for SDEs and SPDEs is supported by a broad methodological arsenal, including maximum likelihood estimation [31], the generalized method of moments [32], and Bayesian inference frameworks [33] that yield full posterior distributions and enable comprehensive uncertainty quantification [34]. Computational advances now permit the application of Markov chain Monte Carlo (MCMC) methods (such as Hamiltonian Monte Carlo and particle MCMC [35]) to efficiently explore high-dimensional posterior spaces [36]. For online estimation in systems with streaming data, sequential Monte Carlo methods, such as the particle filter and the ensemble Kalman filter, have proven indispensable [37]. In scenarios where the likelihood function is analytically intractable, likelihood-free inference techniques, including approximate Bayesian computation and synthetic likelihood methods [38], offer practical alternatives by relying on forward simulation and carefully chosen summary statistics in place of explicit likelihood evaluation [39].

Recent theoretical advances in high-frequency sampling asymptotics have yielded consistent and asymptotically efficient estimators for both drift and diffusion coefficients under substantially weaker smoothness and regularity conditions than those previously required [40]. These results hold even under model misspecification [41], endogenous observation noise [42], or irregular sampling schedules, thereby extending their applicability to empirical datasets where idealized sampling assumptions are violated [43]. In the context of SPDEs, spectral estimation methods (combined with finite-dimensional Galerkin projections [44]) enable accurate recovery of dynamical modes from partial, noisy, and spatially sparse measurements [45]. Parallel developments in computational statistics have given rise to hybrid methodologies that integrate stochastic modeling with machine learning [46]. In particular, Neural SDE frameworks [47], physics-informed neural operators (PINO) [48], and operator-learning architectures can now recover the functional structure of the drift f and diffusion G in (1) directly from data, while rigorously enforcing physical and statistical constraints [49]. This fusion of stochastic analysis, numerical approximation, and data-driven inference represents a decisive step toward predictive modeling of complex, high-dimensional, and uncertainty-dominated systems [50].

Over the past decade, several research frontiers have expanded both the theoretical foundations and the computational reach of stochastic modeling [51, 52]. One prominent direction addresses the treatment of high-dimensional and multiscale systems [53], where the number of interacting stochastic degrees of freedom can be in the hundreds or thousands [54], as encountered in turbulent fluid dynamics [55], molecular simulations [56], and eddy-resolving climate models [57]. In such regimes, direct time integration of the full stochastic system is often computationally prohibitive [58]. To overcome this barrier, dimensionality-reduction techniques (such as dynamic mode decomposition, stochastic principal component analysis, and manifold learning [59]) combined with reduced-order and homogenized models [60], have enabled the derivation of effective lower-dimensional stochastic dynamics that preserve essential statistical, spectral, and dynamical properties of the original system while significantly reducing computational cost [61].

Another active research area concerns non-Lipschitz and stiff stochastic dynamics [62], which arise in systems with super-linear drifts, polynomial nonlinearities, or singular coefficients [63]. Classical explicit discretizations, including the Euler-Maruyama scheme [64], may fail in such settings due to numerical instability or divergence. To address these challenges, tamed and balanced numerical integrators have been developed to control the growth of nonlinear terms without sacrificing strong or weak convergence properties [65]. In addition, implicit-explicit (IMEX) schemes [66], exponential integrators [67], and stochastic Rosenbrock-type methods [68] have been successfully adapted to handle severe stiffness and disparate temporal scales, ensuring stability while maintaining high-order accuracy [69].

The modeling of heavy-tailed fluctuations and discontinuous trajectories has likewise progressed significantly [70]. Notable developments include SDEs driven by α-stable Lévy processes [71], which generate infinite-variance increments and capture heavy-tailed statistics [72]; compound Poisson processes [73], which model finite-activity but potentially large-amplitude jumps [74]; and state-dependent jump intensities [75], allowing the jump rate to vary adaptively with the evolving system state [76]. Such models have become indispensable in quantitative finance [77], where they describe market shocks and volatility clustering; in reliability engineering [78], where they model catastrophic system failures; and in epidemiology, where they capture abrupt changes in transmission due to superspreading events or public-health interventions [79].

Memory-dependent stochastic processes constitute another rapidly advancing frontier [80]. Fractional SDEs and SPDEs [81], incorporating Caputo or Riemann-Liouville derivatives [82], accurately model anomalous diffusion [83], viscoelastic response [84], and long-range temporal correlations observed in diverse physical, biological [85], and socioeconomic systems [86]. The inherently nonlocal structure of fractional operators poses severe computational challenges [87], which have been mitigated by fast convolution algorithms [88], kernel compression strategies [89], and spectral approximations [90], thereby reducing memory and time complexity from $O(N^2)$ to nearly $O(N \log N)$ without significant loss of accuracy [91].

The integration of stochastic modeling with modern machine learning has given rise to a new class of data-driven inference methods [92]. Neural SDEs parameterize the drift and diffusion terms using deep neural networks [93], enabling direct estimation from raw, noisy, or irregularly sampled data while preserving stochastic well-posedness [94]. Operator-learning paradigms [95], including physics-informed neural operators, generalize this framework to SPDEs, learning mappings between infinite-dimensional function spaces under physical and statistical constraints [96]. Furthermore, diffusion-based generative models reinterpret high-dimensional data synthesis as the numerical solution of forward-reverse SDEs, providing a probabilistic bridge between stochastic analysis and state-of-the-art generative learning [97, 98].

Advances in statistical inference have paralleled these methodological innovations [99]. Scalable Bayesian inference for SDE and SPDE models now leverages surrogate modeling [100], Gaussian process emulators [101], and multi-fidelity Monte Carlo schemes to accelerate likelihood evaluations [102]. This has enabled full posterior sampling in previously intractable high-dimensional settings, thereby facilitating rigorous uncertainty quantification in complex stochastic systems [103]. When exact likelihood evaluation is impossible, likelihood-free approaches such as approximate Bayesian computation (ABC) [104] and synthetic likelihood estimation have emerged as powerful alternatives, relying on forward simulation combined with carefully chosen summary statistics [105].

The impact of these developments is evident across multiple disciplines. In epidemiology, stochastic compartmental models augmented with demographic noise [106], Lévy-driven transmission rates [107], and adaptive behavioral feedback mechanisms have achieved superior realism in reproducing epidemic variability [108], persistence probabilities, and extinction events [109]. In quantitative finance, regime-switching jump-diffusion models accurately capture sudden volatility spikes and heavy-tailed return distributions [110, 111]. In climate science, SPDEs driven by spatially correlated colored noise provide a principled framework for representing structured uncertainty in coupled ocean-atmosphere systems [112]. In robotics and control, stochastic optimal control problems (often formulated via backward SDEs) enable robust decision-making under uncertainty, supporting safe trajectory planning and real-time risk-aware operations in dynamic environments [113].

This review delivers a synthesis of recent advances in stochastic modeling with a deliberate emphasis on the *coupled* evolution of (i) stochastic theory (SDEs/SPDEs, including jumps and memory effects), (ii) statistical inference and identifiability under partial/noisy observation, and (iii) algorithmic design principles that determine long-time reliability and practical usability in scientific workflows [114]. While the mathematical foundations of SDEs and SPDEs are well established, the 2020-2025 period has seen a marked consolidation and cross-fertilization of three research streams that are often surveyed *in isolation*: (a) data-driven system identification (e.g., neural SDEs and learned drift-diffusion models), (b) operator learning and physics-informed

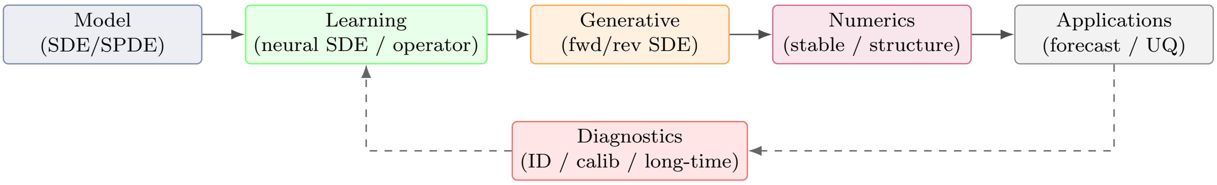

Figure 1. Methodological pipeline.

surrogates for stochastic dynamics (including PINO-type approaches), and (c) generative modeling via forward-reverse SDEs and probability-flow formulations, alongside renewed progress in structure-preserving stochastic numerics. The novel contribution of the present review is therefore not a standalone account of any single strand, but a *unifying taxonomy* that organizes these developments as an end-to-end methodological pipeline from model specification (including heavy tails, jumps, and fractional effects), to learnable representations and inference under incomplete data, to simulation schemes calibrated for stiffness, non-Lipschitz nonlinearities, and fidelity to invariant measures. In addition to summarizing results, we provide critical analysis of where and why methods fail (e.g., drift-diffusion confounding, solver-induced bias, and degradation under sparse/biased biological sampling) and we translate the methodological landscape into application-level guidance by mapping tools to concrete biological and epidemiological tasks such as estimating time-varying transmission or reaction rates, forecasting under reporting artifacts, modeling spatial spread via SPDEs, and quantifying uncertainty for decision support. By consolidating and analyzing works from 2020 to 2025, this review aims to clarify what is now mature, what remains fragile, and which research directions are most promising for building theoretically grounded, computationally scalable, and practically impactful stochastic models in biology and epidemiology [115].

Related surveys typically focus on *one* methodological axis: neural SDE identification and continuous-time latent-variable learning, operator learning/PINOs for PDE-SPDE surrogates, diffusion-modeling and reverse-time SDE sampling, or stochastic numerics for stability and long-time accuracy. These works provide deep coverage within their respective lanes, but they rarely make the *cross-cutting* connections that determine end-to-end performance in scientific applications, for example, how identifiability under partial observation interacts with solver bias, or how generative reverse-time formulations depend on discretization fidelity and data quality. Our review complements these surveys by integrating the four strands into a single pipeline-oriented taxonomy, emphasizing shared failure modes, maturity levels, and minimal evaluation/reporting practices. This perspective is tailored to biomathematics, where practical deployment requires simultaneously credible inference, robust simulation, and interpretable translation to decision-relevant quantities.

As summarized in Fig. 1, our review is organized as a pipeline linking stochastic model specification (SDE/SPDE) to data-driven learning (neural SDEs and operator learning), generative SDE formulations (forward-reverse/score-based diffusion), and structure-preserving numerics for stable simulation and uncertainty quantification in biological and epidemiological applications [93, 96, 97, 161, 162].

The remainder of this article is structured to provide a coherent progression from foundational concepts to advanced applications. Section 2 examines data-driven identification strategies for stochastic systems, with particular attention to neural SDE frameworks and operator-learning methodologies that enable the direct inference of drift and diffusion structures from data. Section 3 explores the emerging role of SDE-based formulations in generative modeling, highlighting their connection to forward-reverse processes and diffusion-based learning. Section 4 offers a detailed survey of numerical integration techniques for SDEs and SPDEs, emphasizing stability, convergence, and long-time accuracy in challenging regimes. Representative applications in biology and epidemiology are presented in Section 5, and the review concludes in Section 6 with a synthesis of open research challenges and promising future directions.

2. Learning stochastic dynamics from data

The growing availability of high-resolution, large-scale datasets has shifted the emphasis in stochastic modeling from prescribing dynamics solely on the basis of first-principles derivations to inferring them directly from empirical observations. Within this data-centric paradigm, two complementary methodologies have emerged. The first addresses finite-dimensional systems through neural SDEs [116], in which both drift and diffusion structures are represented by flexible, trainable mappings estimated from data. The second focuses on infinite-dimensional systems via operator-learning approaches for SPDEs [117], which aim to approximate stochastic evolution operators in a mesh-independent fashion. Both frameworks are designed to assimilate sparse or noisy measurements, generalize robustly across dynamical regimes, and preserve the analytical properties imposed by the underlying stochastic calculus.

2.1. Neural SDEs and drift-diffusion identification

In the neural SDE setting, the drift and diffusion terms are replaced by parameterized functions learned from data, while the governing Itô form is retained [93]. A general formulation on $\mathbb{R}^d$ is

$$\mathrm{d}X_t = f_\theta(X_t,t)\,\mathrm{d}t + G_\theta(X_t,t)\,\mathrm{d}W_t, \qquad X_0 \sim \mu_0, \tag{2}$$

where $X_t \in \mathbb{R}^d$ is the state vector at time t, $f_\theta : \mathbb{R}^d \times [0,T] \to \mathbb{R}^d$ is the drift, $G_\theta : \mathbb{R}^d \times [0,T] \to \mathbb{R}^{d\times m}$ is the diffusion coefficient matrix, θ denotes all trainable parameters, W_t is an m-dimensional standard Wiener process, and μ_0 is the distribution of the initial state. When discrete samples $\{x_{t_k}\}_{k=0}^n$ with time steps $\Delta t_k = t_{k+1} - t_k$ are available, one common identification approach is maximum likelihood under a discretized transition model [118]. Using the Euler-Maruyama approximation [119], the conditional distribution is

$$X_{t_{k+1}} \mid X_{t_k} \approx \mathcal{N}\Big(x_{t_k} + f_\theta(x_{t_k},t_k)\,\Delta t_k,\ \Sigma_\theta(x_{t_k},t_k)\,\Delta t_k\Big),$$

where $\Sigma_\theta := G_\theta G_\theta^\top$ is the diffusion covariance. Maximization of this Gaussian likelihood yields an estimate of θ [120]. For state-dependent diffusion, higher-order schemes such as Milstein or Itô-Taylor expansions improve accuracy [121]. Positive semidefiniteness of Σ_θ is enforced via Cholesky factorization, and drift growth is regulated through tamed or monotonicity-preserving architectures to ensure well-posedness and strong solution uniqueness [122].

An alternative approach trains in continuous time by differentiating through the SDE solver. Fixing noise seeds $\xi_k \sim \mathcal{N}(0, I_m)$, a single solver step is denoted by

$$X_{t_{k+1}} = \Phi_\theta(X_{t_k}, t_k;\, \xi_k),$$

where Φ_θ encodes the numerical update rule [123]. Gradients $\nabla_\theta \mathbb{E}[\mathcal{L}(X_{0:n})]$ are obtained via direct backpropagation or through the adjoint SDE, the latter reducing memory from $\mathcal{O}(n)$ to $\mathcal{O}(1)$ in the number of steps [124]. For partially observed systems $Y_k = HX_{t_k} + \varepsilon_k$, where H is the observation matrix and ε_k is measurement noise, the neural SDE can be coupled with a differentiable filter (such as an ensemble Kalman update or amortized particle weighting) and trained by maximizing a filtering log-likelihood or a variational lower bound [125].

When likelihood-based methods are unreliable, for example under irregular sampling or unknown observation noise, moment-based estimation offers a robust alternative [126]. For any smooth function $g : \mathbb{R}^d \to \mathbb{R}$, Itô's formula gives the local martingale:

$$M_t^g = g(X_t) - g(X_0) - \int_0^t \Big(\nabla g^\top f_\theta + \tfrac{1}{2}\,\mathrm{Tr}[\Sigma_\theta \nabla^2 g]\Big)(X_s, s)\,\mathrm{d}s.$$

Enforcing that empirical averages of $M_{t_k}^g$ vanish yields unbiased estimating equations even with noisy or irregularly spaced data [127]. The diffusion can also be estimated from the quadratic variation:

$$\frac{1}{\Delta t_k}(X_{t_{k+1}} - X_{t_k})(X_{t_{k+1}} - X_{t_k})^\top \approx \Sigma_\theta(X_{t_k}, t_k),$$

with bias corrections applied for finite Δt_k [128]. In models with jumps, drift inference may employ Girsanov's theorem [129], while the jump component can be estimated from its Lévy-Khintchine representation [130], using, for example, compound Poisson thinning or variational approximations for infinite-activity cases [131].

Careful treatment of identifiability is essential: without constraints, drift effects can be spuriously absorbed into the diffusion term [132]. Common remedies include imposing structural restrictions on G_θ (e.g., diagonal or low-rank form), penalizing deviations $\|\widehat{\Sigma}-\Sigma_\theta\|$ over short time increments, or leveraging stationary distributions [133]. If π is the invariant measure of (2), f_θ and Σ_θ satisfy the stationary Fokker-Planck equation

$$\nabla\cdot\big(f_\theta \pi\big) - \tfrac{1}{2}\sum_{i,j}\partial_i\partial_j\big(\Sigma_{\theta,ij}\pi\big) = 0,$$

which can serve as a weak constraint when long-run data are available [134]. In practice, training objectives often combine the transition likelihood, quadratic-variation penalties, weak Fokker-Planck constraints, and stability regularizers that enforce linear growth bounds or one-sided Lipschitz conditions to promote ergodicity [135].

Uncertainty quantification may be addressed via Bayesian neural SDEs [136], where priors are placed on θ and posteriors are explored using stochastic-gradient MCMC or ensemble variational inference [137], or by nonparametric bootstrap resampling over both trajectories and noise seeds [138]. Robustness to out-of-distribution states can be enhanced by spectral normalization in the drift and by imposing lower bounds on diffusion coefficients [139]. In controlled systems with exogenous inputs u_t, the drift $f_\theta(x,t,u)$ can be parameterized in a control-affine form, enabling integration with stochastic optimal control through Hamilton-Jacobi-Bellman surrogates or pathwise policy-gradient methods [140].

2.2. Operator learning for SPDEs (PINO, Neural Operators)

Stochastic partial differential equations on spatial domains $\Omega \subset \mathbb{R}^p$ can be expressed in abstract form as

$$\mathrm{d}u(t) = \mathcal{A}_\phi\big(u(t)\big)\,\mathrm{d}t + \mathcal{B}_\phi\big(u(t)\big)\,\mathrm{d}W_t, \qquad u(0)=u_0, \quad t\in[0,T], \tag{3}$$

where $u(t):\Omega\to\mathbb{R}^q$ denotes the state variable at time t, $\mathcal{A}_\phi$ is a (possibly nonlinear) differential operator parameterized by ϕ, $\mathcal{B}_\phi$ is a noise-multiplication operator, and W_t is a cylindrical Wiener process on an appropriate Hilbert space. The initial condition is $u_0\in\mathcal{H}$, where $\mathcal{H}$ is the state space [141]. The learning objective is to approximate the solution operator

$$\mathcal{S} : (u_0, f, \xi) \mapsto u(\cdot),$$

which maps initial data u_0, deterministic forcing terms $f:\Omega\times[0,T]\to\mathbb{R}^q$, and stochastic realizations ξ to full trajectories u [142]. Neural operator architectures approximate $\mathcal{S}$ directly in function space, enabling mesh-independent generalization and evaluation across discretizations [96]. Two principal designs have been particularly effective.

The first class, Fourier or kernel neural operators, alternates local linear mappings with global integral transforms of the form

$$v_{k+1} = \sigma\big(Wv_k + \mathcal{K}_\psi v_k\big),$$

where v_k denotes the feature representation at layer k, W is a local (pointwise) linear transformation, σ is a nonlinear activation, and $\mathcal{K}_\psi$ is a global integral operator with kernel $K_\psi(x,y)$ parameterized either in physical space or via truncated Fourier multipliers [143]. Stacking such layers yields a mapping $\mathcal{G}_\Psi:\mathcal{X}\to\mathcal{Y}$ that approximates $\mathcal{S}$. In the stochastic setting, randomness enters through sampled forcings f or noise realizations ξ, and the training objective minimizes trajectory- or distribution-based losses over ensembles of noise samples [144]. Stability is promoted by constraining spectral multipliers (e.g., bounding spectral radius) and incorporating skip connections designed to mimic the dissipative structure of the underlying PDE [145].

The second class, PINO, augments data-driven learning with weak enforcement of the governing equations in (3) [146]. Given a set of test functions $\{\varphi_j\}_{j=1}^J \subset \mathcal{H}$, the variational residual is defined by

$$\mathcal{R}_\phi(u) := \sum_{j=1}^{J} \Big\langle \partial_t u - \mathcal{A}_\phi(u),\ \varphi_j \Big\rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the $L^2(\Omega)$ inner product. Stochastic forcing is incorporated either by sampling noise paths ξ and enforcing Itô weak identities, or by matching statistical quantities (e.g., covariances, energy spectra) derived from theory [147]. The composite training objective takes the form

$$\min_{\Psi,\phi} \ \mathbb{E}_\xi \big[\|\mathcal{G}_\Psi(u_0, f, \xi) - u\|_{\mathcal{Y}}^2 \big] + \alpha\, \mathbb{E}_\xi \big[\|\mathcal{R}_\phi(\mathcal{G}_\Psi(u_0, f, \xi))\|^2 \big] + \beta\, \mathcal{J}_{\mathrm{BC/IC}},$$

where $\|\cdot\|_{\mathcal{Y}}$ denotes the L^2 norm over $\Omega \times [0, T]$, $\mathcal{J}_{\mathrm{BC/IC}}$ enforces boundary and initial conditions, and $\alpha, \beta > 0$ are weighting parameters [148]. This formulation reduces data requirements and mitigates nonphysical extrapolation in regimes where stochastic forcing excites unresolved spatial or temporal scales.

Accurate noise representation is critical. A common approach is to expand $\xi(t, x)$ in a Karhunen-Loève basis $\{\phi_i(x)\}_{i=1}^r$ [149],

$$\xi(t, x) = \sum_{i=1}^{r} a_i(t)\, \phi_i(x),$$

where $a_i(t)$ are temporally stochastic coefficients and the ϕ_i are spatial modes, possibly learned jointly with the operator [150]. For multiplicative noise, maintaining Itô-Stratonovich consistency requires either explicit correction terms in the residual or learned commutator terms that approximate the conversion [151]. If an invariant measure π is known or estimable, an auxiliary ergodic loss

$$\mathbb{E}_\pi[\mathcal{Q}(u)] \approx \frac{1}{T} \int_0^T \mathcal{Q}\big(\mathcal{G}_\Psi(u_0, f, \xi)(t)\big)\, \mathrm{d}t$$

is used to align surrogate long-time statistics with those of the true system, where $\mathcal{Q}$ denotes a chosen diagnostic such as energy, enstrophy, or a spectral functional [152].

Generalization across meshes and domains is supported by spectral parameter tying to ensure resolution scalability, coordinate-free message passing for unstructured meshes, and discrete conservation constraints such as divergence-free enforcement for incompressible flows or positivity for density fields [153]. For stiff or multiscale SPDEs, multi-rate operator architectures couple coarse global updates with local fine-scale correctors, analogous to heterogeneous multiscale methods [154]. Uncertainty quantification can be incorporated via Bayesian priors on operator weights, ensembles over noise realizations [155], or linearization-based covariance approximations using Gauss-Newton methods [156].

A hybrid approach couples finite-dimensional latent SDEs with learned spatial bases, representing the solution as $u(t, x) \approx \sum_{i=1}^r c_i(t)\, \phi_i(x)$, where $c(t) \in \mathbb{R}^r$ evolves under a neural SDE and $\{\phi_i\}$ are trainable spatial modes [157]. This yields interpretable low-rank stochastic dynamics with mesh-independent inference and efficient sampling, applicable to tasks such as stochastic control, data assimilation, and probabilistic forecasting [158].

Table 1. Critical view of learning stochastic dynamics from data (Sec. 2). Compact summary of advantages and limits to guide method selection and reporting.

Method	**What is learned**	**Key advantages**	**Key limits / failure modes**	**When to use**
Neural SDEs and drift-diffusion identification				
MLE (EM) [93, 118–120]	Learn (f_θ, G_θ) via Gaussian transition likelihood (Euler-Maruyama).	Simple, scalable; transparent objective; easy PSD/regularization.	Discretization bias for coarse Δt; brittle under irregular sampling/noise; confounding possible.	Dense/moderate sampling; benchmarking, prototyping.

Continued on next page

Method	What is learned	Key advantages	Key limits / failure modes	When to use
Higher-order [121]	Milstein / Itô-Taylor for state-dependent diffusion.	Less bias; more efficient; better for multiplicative noise.	More assumptions/complexity; still sensitive to partial observation.	When EM underfits transitions; multiplicative noise.
Pathwise / adjoint [123, 124]	Differentiate through solver (or adjoint) for task losses.	Flexible beyond likelihood; memory savings; integrates control/forecasting.	Noisy gradients; solver-induced bias; stiff/non-Lipschitz instability.	Long-horizon forecasting/control; non-likelihood learning.
Partial obs + filter [125]	Learn latent SDE from $Y_k = HX_{t_k} + \varepsilon_k$ via differentiable filtering.	Realistic for biology; principled link to observation model.	Filter degeneracy/linearization bias; identifiability worsens.	Hidden states; sensor noise; irregular sampling.
Martingale moments [126, 127]	Estimate via Itô-martingale identities / estimating equations.	Robust to irregular grids; interpretable constraints; complements MLE.	Test-function choice matters; can be less efficient.	Irregular sampling; uncertain noise models.
Quadratic variation [128]	Estimate Σ_θ from short-time increments (with corrections).	Direct diffusion signal; useful diagnostic vs confounding.	Sensitive to measurement noise; finite-Δt bias.	High-frequency data; auxiliary loss/diagnostic.
Jumps (Lévy) [129–131]	Infer drift+jumps (Girsanov / Lévy-Khintchine / variational thinning).	Captures bursts/heavy tails; relevant for abrupt interventions.	Hard identifiability; masked by noise/discretization; high cost.	Shock-driven/event data; heavy-tailed regimes.
Identifiability + stationarity [132–135]	Structure on G_θ + weak Fokker-Planck / ergodic constraints.	Targets drift-diffusion confounding; improves scientific credibility.	Stationarity may fail; PDE constraints stiff; risk underfitting transients.	Long time series; quasi-stationary regimes.
UQ (Bayes/ens.) [136–139]	Posterior/ensemble/bootstraps over θ and noise seeds.	Calibrated uncertainty; supports decisions; detects overconfidence.	Expensive; calibration degrades OOD; needs reporting standards.	Forecasting/policy; when uncertainty is critical.
Operator learning for SPDEs (Neural Operators, PINO)				
Neural operators [96, 142–145]	Learn solution operator $\mathcal{S} : (u_0, f, \xi) \mapsto u(\cdot)$.	Mesh-independent surrogates; fast sampling/UQ after training.	OOD failures (regimes/geometries); stability not automatic; nonphysical outputs.	Multi-query emulation; UQ/optimization loops.
PINO [146–148]	Data loss + weak SPDE residual (variational constraint).	Less data; reduces nonphysical extrapolation; residual diagnostics.	Residual weighting delicate; rare-event/tails hard; Itô/Strat subtle.	Known physics + limited labels; extrapolation-sensitive tasks.
Noise basis (KL) [149, 150]	Low-rank noise expansions; learn modes jointly.	Tractable randomness; sample-efficient; interpretable modes.	Rank truncation bias; non-identifiability; tail distortion.	Dominant-mode stochasticity; moderate dimensions.
Itô/Strat fixes [151]	Consistency corrections for multiplicative noise.	Reduces structural bias; improves transfer.	Implementation burden; often neglected (silent error).	Multiplicative-noise SPDEs; sensitive physics.
Ergodic loss [152]	Match invariant-measure diagnostics (energy/moments/spectra).	Improves distributional credibility; supports risk endpoints.	Needs long runs/robust estimators; may trade short-term accuracy.	Long-horizon stats; tail-risk applications.
Mesh/domain gen. [153, 154]	Coordinate-free design + constraints (positivity, divergence-free); multirate.	Portability; fewer artifacts; helps stiff/multiscale regimes.	Engineering effort; constraints reduce flexibility; tuning issues.	Irregular meshes; conservation structure; stiffness.
UQ for operators [155, 156]	Bayes/ensembles or covariance via linearization (GN).	Credible uncertainty propagation; supports design decisions.	OOD calibration hard; Gaussian approx misses tails; cost.	High-stakes prediction; limited SPDE truth.
Hybrid latent SDE [157, 158]	Low-rank $u(t, x) \approx \sum c_i(t)\phi_i(x)$ with neural SDE for $c(t)$.	Interpretable; efficient sampling; mesh-independent inference.	Rank choice critical; misses localized/rare extremes.	Coherent-structure systems; fast probabilistic prediction.

Table 1 highlights a maturity gradient across the learning pipeline. For neural SDEs, likelihood- and solver-based training are now well established, but *identifiability* (drift-diffusion confounding) and *solver-induced bias* remain the main blockers for reliable scientific inference under partial observation [124, 132, 135]. For operator learning, mesh-independent surrogates have matured rapidly, yet the dominant open challenges are *out-of-distribution generalization* and *long-time statistical fidelity*, especially under multiplicative noise and rare-event regimes [145, 151, 152]. These observations motivate the reporting standards proposed later (distributional scores, coverage diagnostics, and invariant-measure checks) as minimal safeguards for biological deployment.

3. Generative Modeling as SDEs

Generative modeling aims to construct stochastic transformations that map a tractable reference distribution onto a complex target distribution [24]. Recent advances have established that score-based generative models and denoising diffusion models can be formulated entirely within the stochastic differential equation (SDE) framework introduced in (1) [97]. In this setting, a *forward* process progressively perturbs samples from the data distribution p_{data} into a simple reference law p_{ref}, typically a standard Gaussian, over a fixed time horizon [159]. A *reverse-time* stochastic process then reconstructs new samples by inverting this diffusion [160], as illustrated in Fig. 2.

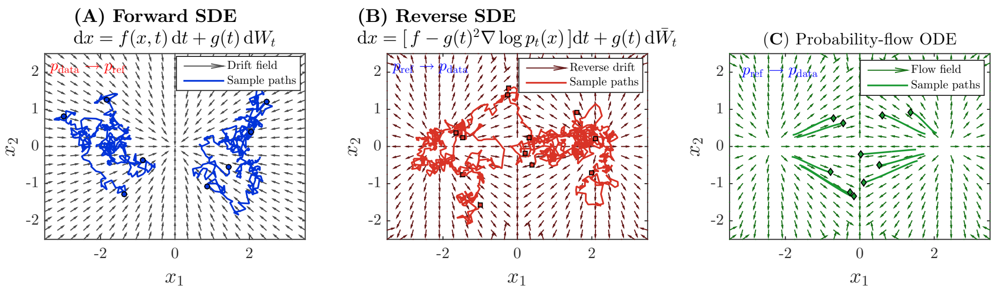

Figure 2. Forward-reverse SDE framework and its deterministic probability-flow ODE counterpart. Panel (A) shows the forward SDE transporting p_{data} to a reference Gaussian p_{ref}. Panel (B) shows the reverse SDE mapping p_{ref} back to p_{data} via the learned score function $\nabla \log p_t(x)$. Panel (C) depicts the equivalent probability-flow ODE sharing the same marginals as the forward and reverse SDEs but evolving deterministically. Vector fields represent drift components; colored trajectories indicate sample paths.

Let p_t denote the marginal distribution of the forward process at time t. Under regularity conditions, Anderson's time-reversal theorem [161] yields the reverse-time SDE:

$$\mathrm{d}x(t) = \Big[f(x(t), t) - g(t)^2\, \nabla_x \log p_t\big(x(t)\big) \Big]\, \mathrm{d}t + g(t)\, \mathrm{d}\bar{w}(t), \tag{4}$$

where $\bar{w}(t)$ is a standard Wiener process evolving backward in time, f and g are the drift and diffusion coefficients of the forward dynamics, and $\nabla_x \log p_t$ is the score function of p_t. Since p_t is not known in closed form, the score function is approximated by a trainable surrogate $\hat{s}_\varphi(x, t)$, estimated through score-matching or denoising-score-matching objectives, ensuring asymptotic consistency in the limit of infinite data and model capacity [162].

Numerical generation is performed by discretizing (4) with schemes such as Euler-Maruyama [163], stochastic Runge-Kutta [164], or predictor-corrector integrators [165]. The choice of scheme directly influences sampling bias and variance, with high-order solvers yielding greater fidelity for a given computational budget [166]. An important alternative is the *probability-flow ordinary differential equation* associated with the forward process, which evolves deterministically yet preserves the same marginal distributions as the reverse SDE, often permitting larger stable time steps [167].

In scientific applications, the forward dynamics can be tailored to preserve statistical invariants, enforce spectral constraints, or emulate dissipative mechanisms from coarse-grained physics [168]. The learned reverse process then reconstructs fine-scale structure and corrects higher-order moments, enabling physically consistent synthesis in high-dimensional contexts such as turbulent flows, molecular dynamics, and climate modeling [169]. Computational efficiency can be further enhanced through adaptive time stepping, variance-reduction strategies (e.g., control variates), and reduced-order latent representations [170].

A natural generalization introduces exogenous control inputs u_t into the drift, yielding a controlled reverse-time process of the form

$$f(x,t,u_t) - g(t)^2\,\nabla_x \log p_t(x,t),$$

which connects generative modeling with stochastic optimal control [171]. In this interpretation, the reverse dynamics act as an optimal feedback policy minimizing a cost functional equivalent to the negative log-likelihood, enabling targeted sampling subject to trajectory or terminal constraints [172]. This formulation has direct implications for data assimilation, scenario generation, and constraint-aware simulation [173].

By embedding generative modeling entirely within the SDE formalism and incorporating domain-specific constraints, one obtains a unifying mathematical structure capable of preserving statistical integrity while extending these methods from static data generation to the synthesis and control of complex stochastic dynamical systems.

Table 2. Critical view of generative modeling as SDEs (Sec. 3). Compact summary of the forward-reverse SDE pipeline, the main solver choices, and the dominant failure modes (score error compounding, discretization bias, tail miscalibration, and constraint-induced bias) relevant for scientific applications.

Component / method	**Core idea / what is learned**	**Key advantages (critical view)**	**Key limits / failure modes**	**When to use**
Forward diffusion and reverse-time sampling				
Forward SDE (data $\to$ ref) [24, 97, 159]	Defines a noising process transporting p_{data} to a tractable p_{ref} (often Gaussian) over $t \in [0,T]$.	Mathematically clean bridge between diffusion models and SDEs; controllable noise schedule; enables likelihood/ELBO-style training.	Poor schedule choice induces ill-conditioned learning (vanishing/exploding SNR); may destroy relevant invariants; can amplify dataset bias.	Generic high-dimensional generation; as a baseline pipeline before adding constraints.
Reverse-time SDE [160, 161]	Samples by integrating the time-reversed SDE with drift correction via the score $\nabla_x \log p_t(x)$.	Flexible sampling; supports conditional generation; can incorporate domain constraints via drift modifications.	Sensitive to score error (compounds backward); discretization bias; numerical instability for stiff dynamics.	When exact sampling is not required but high fidelity is; conditional/scenario generation.
Score estimation (DSM / SM) [162]	Learns $\hat{s}_\varphi(x,t) \approx \nabla_x \log p_t(x)$ via score matching or denoising score matching.	Avoids explicit density evaluation; scalable to high dimension; asymptotically consistent (idealized limit).	Finite-data/model bias; poorly calibrated tails/rare events; sensitive to measurement noise and covariate shift.	Large datasets; high-dimensional fields; when density evaluation is infeasible.
Numerical solvers and sampling accelerations				
Euler-Maruyama [163]	First-order discretization of reverse SDE; simplest sampler.	Fast, simple; robust baseline; easy to implement for benchmarking.	Large step-size bias; many steps needed for quality; unstable if drift/score is stiff.	Quick experiments; low-budget runs; baseline comparisons.
Higher-order SDE solvers (SRK) [164, 166]	Stochastic Runge-Kutta / higher-order discretizations to reduce bias/variance per step.	Better fidelity per compute; fewer steps for same quality; improved stability in moderate stiffness.	More complex; requires additional evaluations; gains depend on regularity assumptions.	When EM is too biased/slow; medium-to-high accuracy sampling.

Continued on next page

Component / method	**Core idea / what is learned**	**Key advantages (critical view)**	**Key limits / failure modes**	**When to use**
Predictor-corrector [165, 166]	Alternates prediction with score-based correction (e.g., Langevin-type refinements).	Often improves sample quality at fixed steps; can reduce discretization artifacts; practical robustness.	Heuristic tuning; can increase compute; may over-smooth or mishandle sharp features.	Image/field synthesis; practical high-quality sampling under limited steps.
Probability-flow ODE [167]	Deterministic ODE sharing the same marginals as the SDE; enables ODE solvers/adaptive stepping.	Deterministic trajectories; often larger stable steps; easier likelihood computation in some settings.	ODE discretization error still matters; can lose stochastic diversity if not handled carefully; may be sensitive to solver tolerances.	Fast sampling with adaptive solvers; when determinism and reproducibility matter.
Adaptive stepping + variance reduction [170]	Controls local error and reduces variance via control variates / adaptive step schedules / reduced-order latents.	Efficiency gains for stiff or multiscale systems; better compute-quality tradeoff.	Complexity and tuning; may break guarantees if heuristics dominate; risk of biased estimates.	Scientific simulation surrogates; multiscale/stiff dynamics; compute-constrained workflows.
Scientific constraints, control, and reliability				
Physics-consistent reconstruction [168, 169]	Reverse process restores fine scales and higher-order statistics under constraints (moments, spectra, conservation).	Better long-time/statistical fidelity; improves interpretability for scientific use.	Hard for rare events and tails; constraint enforcement may be approximate; evaluation must be distributional.	Scientific field synthesis; uncertainty propagation; scenario generation with plausibility constraints.
Controlled generative SDEs / control link [171–173]	Introduces control inputs u_t in the drift; interprets reverse dynamics as feedback policy for targeted sampling.	Connects diffusion models to stochastic control and data assimilation; enables constrained/goal-conditioned generation.	Control objectives can conflict with fidelity; sensitive to model bias; stability under constraints can be nontrivial.	Data assimilation; constraint-aware scenarios; goal-directed sampling under terminal/path constraints.

Table 2 indicates that the dominant bottleneck is typically *not* the forward SDE itself but the interaction between (i) score-estimation error (especially in tails/rare events), and (ii) discretization error that compounds during reverse-time integration [162, 166]. In sparse or biased biological datasets, these effects can manifest as miscalibrated uncertainty and overconfident extrapolation, motivating distributional diagnostics (e.g., coverage, CRPS/PIT) and sensitivity checks to solver tolerances as minimal safeguards for deployment [167, 170].

4. Numerical methods for SDEs and SPDEs

The accurate and efficient numerical integration of stochastic differential equations and their infinite-dimensional counterparts is central to both predictive modeling and statistical inference. The numerical discretization must preserve the fundamental properties of the underlying stochastic dynamics, including convergence order, stability, invariant measures, and qualitative behavior over long time horizons. This section surveys recent methodological developments across four interconnected areas.

4.1. Strong and Weak Convergence, and Stability

In the numerical treatment of stochastic differential equations, strong convergence quantifies the mean-square error between the numerical approximation $X_T^{\Delta t}$ and the exact solution X_T at a fixed terminal time T [174],

$$\left(\mathbb{E}\,\|X_T - X_T^{\Delta t}\|^2\right)^{1/2} = \mathcal{O}\big((\Delta t)^{\gamma}\big),$$

while weak convergence measures the error in expectations of sufficiently smooth observables φ,

$$\big|\mathbb{E}\,\varphi(X_T) - \mathbb{E}\,\varphi(X_T^{\Delta t})\big| = \mathcal{O}\big((\Delta t)^{\beta}\big).$$

For drift-diffusion systems with non-globally Lipschitz drift, polynomial growth, or multiplicative noise, modern analyses derive orders γ and β under one-sided Lipschitz and coercivity conditions, often with tamed, balanced, or implicit updates to ensure well-posedness of the discrete dynamics [62,175,176]. In infinite dimensions, the same notions apply to semidiscrete SPDEs in Hilbert spaces, where temporal error interacts with spatial discretization error through the smoothing properties of the linear generator and the regularity of the noise [177]. Stability considerations extend beyond mean-square boundedness to include almost-sure exponential stability [178], boundedness in probability, and preservation of qualitative invariants over long horizons [179]. Empirical slopes for representative strong and weak errors, and their asymptotic rates, are illustrated in Fig. 3.

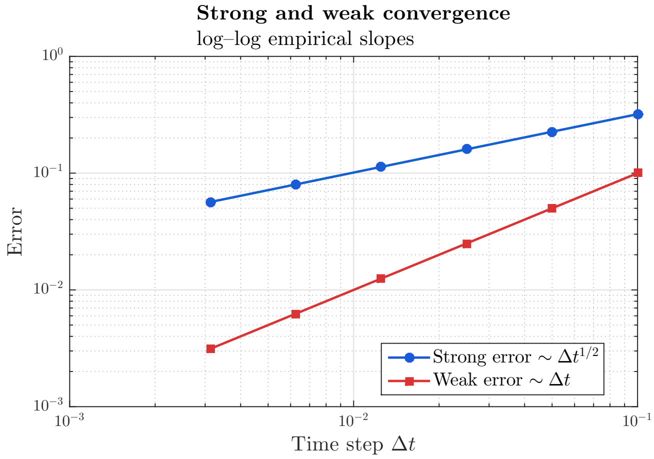

Figure 3. Strong and weak convergence on a log-log scale for representative schemes. The strong error exhibits the expected order one half for an Euler-Maruyama discretization with multiplicative noise, whereas the weak error attains order one for smooth test functionals. Dashed guides indicate the asymptotic slopes used to assess rates. The same methodology extends to semidiscrete SPDEs, where temporal and spatial errors interact through the smoothing properties of the linear part and the regularity of the stochastic forcing.

4.2. Tamed, implicit, and balanced schemes

Classical explicit schemes may diverge when applied to SDEs with superlinear drift growth. Tamed schemes modify the drift term to control growth without sacrificing explicitness [180]. For instance, a tamed Euler method replaces $f(x)$ by $f(x)/(1 + \alpha\Delta t\|f(x)\|)$, ensuring boundedness of increments while preserving strong convergence of order one under appropriate conditions [181]. Implicit schemes, such as backward Euler-Maruyama, provide improved stability properties, particularly for stiff stochastic systems and SPDEs with dissipative operators [182]. Balanced methods interpolate between explicit and implicit updates, adjusting drift and diffusion contributions to simultaneously control stability and accuracy [183]. These approaches have been extended to higher-order stochastic Runge-Kutta methods and to adaptive time-stepping algorithms that refine resolution in regions of high stochastic activity [184].

4.3. Long-time accuracy and invariant measures

For ergodic SDEs, a central requirement on a numerical method is the faithful approximation of the invariant law over long horizons [185]. If π denotes the stationary distribution of the exact process, a method with step size Δt induces a Markov chain with stationary law $\pi^{\Delta t}$ [186]. Long-time accuracy entails weak convergence $\pi^{\Delta t} \Rightarrow \pi$ as $\Delta t \to 0$ and a bias in ergodic averages that decays at an optimal rate [187]. The asymptotic bias of an observable $\mathcal{Q}$ admits Poisson-equation representations that expose its dependence on the generator and the local truncation error [188]; this viewpoint underpins modified-equation corrections and Talay-Tubaro-type expansions for invariant-measure error [189]. In practice, drift and diffusion adjustments, symmetrization, and schemes that preserve reversibility or detailed balance reduce long-time bias while retaining acceptable finite-time accuracy [190]. In infinite dimensions, structure-preserving time integrators for SPDEs target invariants and statistical diagnostics dictated by the physics, such as energy spectra [191], enstrophy [192], or mass [193], and are analyzed using spectral gaps or Wasserstein-contractivity of the associated Markov semigroups [194]. Figure 4 illustrates the approach of time averages toward the invariant expectation and highlights the reduced residual bias of an implicit method relative to explicit Euler-Maruyama at a fixed step size.

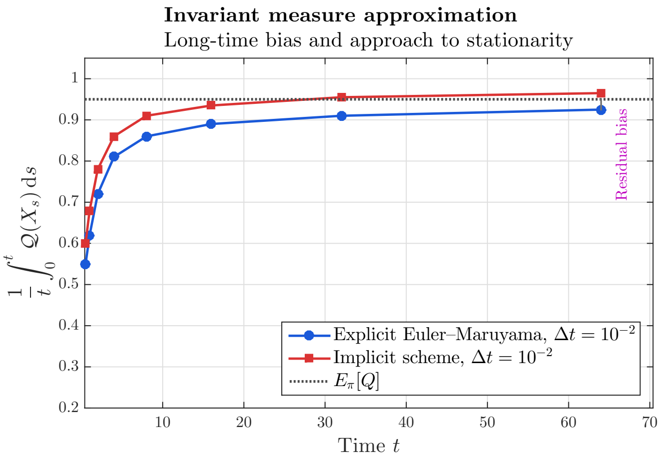

Figure 4. Convergence of time averages to the invariant expectation for an ergodic observable $\mathcal{Q}$. Curves show $\frac{1}{t}\int_0^t \mathcal{Q}(X_s)\,\mathrm{d}s$ for an explicit Euler-Maruyama discretization and an implicit method at the same step size. The dotted line marks $\mathbb{E}_\pi[\mathcal{Q}]$, and the vertical segment at the final time indicates the residual bias. The implicit scheme approaches the stationary value more rapidly and with smaller long-time bias, consistent with theoretical predictions based on generator perturbations and Poisson-equation error analysis.

4.4. Multilevel Monte Carlo and Rare-Event simulation

Multilevel Monte Carlo reduces work by coupling approximations across a hierarchy in time and, when relevant, space [195]. Let φ_ℓ denote an approximation of $\varphi(X_T)$ at level ℓ with step Δt_ℓ (and mesh width h_ℓ if spatial discretization is present). The telescoping identity:

$$\mathbb{E}[\varphi_L] = \mathbb{E}[\varphi_0] + \sum_{\ell=1}^{L} \mathbb{E}\big[\varphi_\ell - \varphi_{\ell-1}\big]$$

is estimated by Monte Carlo averages with N_ℓ *coupled* samples [196]. Three rates govern complexity: the weak bias $\mathcal{O}(\Delta t_L^\alpha)$, the variance decay $\mathrm{Var}(\varphi_\ell - \varphi_{\ell-1}) = \mathcal{O}(\Delta t_\ell^\beta)$, and the per-sample cost $\mathcal{O}(\Delta t_\ell^{-\gamma})$. With the optimal allocation $N_\ell \propto \sqrt{\mathrm{Var}(\varphi_\ell - \varphi_{\ell-1})/C_\ell}$,

$$\mathcal{C}_{\mathrm{MLMC}} = \begin{cases} \mathcal{O}(\varepsilon^{-2}), & \beta > \gamma, \\ \mathcal{O}(\varepsilon^{-2}\log^2\varepsilon), & \beta = \gamma, \\ \mathcal{O}\big(\varepsilon^{-2-(\gamma-\beta)/\alpha}\big), & \beta < \gamma, \end{cases}$$

which improves on single-level sampling whenever $\beta \geq \gamma$. For Lipschitz SDEs with Euler-Maruyama couplings one typically has $\alpha = 1$, $\beta \approx 1$, and $\gamma \approx 1$, leading to $\mathcal{O}(\varepsilon^{-2}\log^2\varepsilon)$ [197]. Antithetic constructions or Milstein-type ideas without Lévy areas often restore $\beta > \gamma$ and the canonical $\mathcal{O}(\varepsilon^{-2})$ [198]. In SPDEs, multilevel couplings combine temporal integrators with spectral or finite-element hierarchies; multi-index Monte Carlo extends the telescoping structure across time-space refinement directions [199].

Rare-event probabilities demand variance control beyond stratification by level [200]. Importance sampling with optimally tilted measures derived from large-deviation asymptotics, splitting or subset simulation with adaptively placed intermediate thresholds, and adaptive multilevel splitting or interacting particle systems yield estimators that remain efficient as the probability decreases [201]. These mechanisms integrate naturally with the multilevel decomposition by biasing each difference $\varphi_\ell - \varphi_{\ell-1}$, by coupling biased path pairs, or by embedding control-based surrogates of the optimal change of measure [202]. The pipeline in Fig. 5 summarizes this workflow from hierarchy construction, through telescoping and estimator assembly, to variance-reduction layers for the tail regime.

Overall performance reflects the balance between bias order, coupling variance, long-time statistical fidelity, and variance-reduction design. In practice, structure-preserving or higher-order discretizations are paired with MLMC (or multi-index MLMC) for bulk uncertainty quantification, while importance sampling, splitting, or adaptive multilevel splitting is activated selectively for tail-sensitive observables.

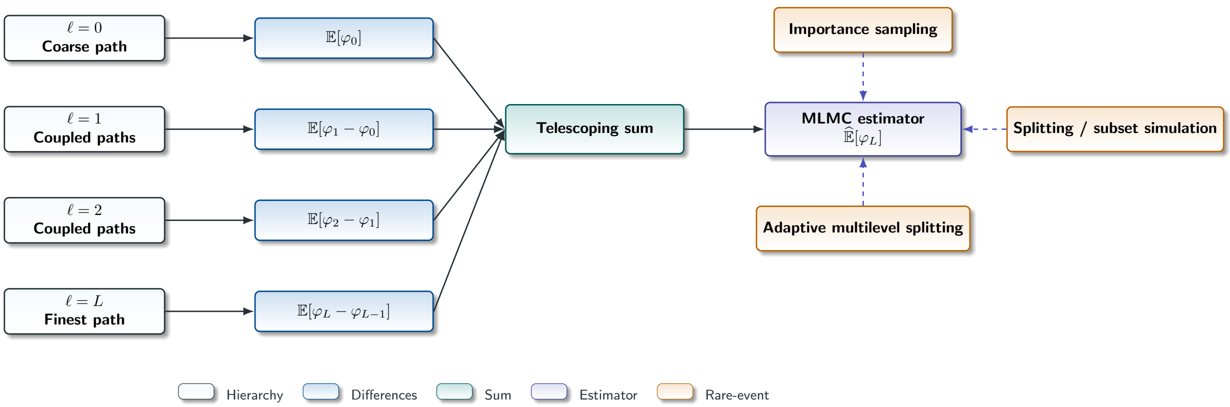

Figure 5. Schematic representation of the MLMC pipeline augmented with rare-event variance-reduction strategies.

5. Applications in biological and epidemiological modeling

To strengthen practical relevance, we synthesize the pipeline from Sections 2-3 into a reproducible workflow for biological and epidemiological studies, where data are typically sparse, biased, and partially observed [203–209]. Model choice is scale-driven: finite-dimensional compartmental / reaction-network systems support neural-SDE identification for time-varying rates, whereas spatial spread and dispersal motivate SPDE formulations and operator-learning surrogates (Section 2.2). Because under-reporting, reporting delays, and heterogeneous testing form an explicit noisy measurement channel, latent-state learning should be coupled with differentiable filtering/assimilation (Section 2.1) rather than trained directly on raw case counts. Reliable inference then combines transition-likelihood or pathwise objectives with quadratic-variation and martingale moment constraints, supplemented (when long time series permit) by weak stationarity or invariant-measure checks to mitigate drift-diffusion confounding and solver-induced bias. Forecasting and uncertainty quantification follow via ensemble simulation, Bayesian neural SDEs, or calibrated bootstrap procedures, evaluated using both pointwise and distributional criteria (e.g., CI coverage and distributional scores) to avoid overconfident extrapolation. Finally, scenario generation and counterfactual analysis can be posed within the generative SDE framework (Section 3), where controlled reverse-time dynamics enable constraint-aware synthesis while preserving target marginals, and operator-learning surrogates provide computationally efficient sampling for spatial forecasts.

Table 3 maps the method families in Sections 2-4 to biological/epidemiological use-cases (2020-2025): (A) neural SDE identification, (B) operator learning and physics-informed PDE/SPDE surrogates, (C) reverse-time generative SDEs, (D) structure-preserving stochastic numerics, and (E) inference/surrogates/phenomena. An entry ✓ denotes at least one relevant published work or public preprint; ✗ denotes a gap (no result found as of August 2025). Across domains, the central objectives are consistent: estimation of time-varying transmission/reaction rates, partial-observation inference, preservation of positivity/invariants, and uncertainty propagation for policy-oriented forecasting.

On the SDE side, neural SDEs provide flexible yet mechanistically grounded parameterizations for noisy biological dynamics and epidemic time series [210], with discretized likelihood calibration via Euler-Maruyama remaining a practical default for discretely sampled outbreaks [221]. When stiffness and signal-to-noise demand higher fidelity, Milstein/Itô-Taylor/SRK schemes improve simulation-and-fit efficiency [217]. Positive-semidefinite diffusion parameterizations (e.g., Cholesky factors) stabilize multivariate learning [222], while tamed or monotonicity-preserving architectures mitigate superlinear drift pathologies arising from saturating incidence or autocatalytic kinetics [223]. Continuous-time training through differentiable SDE solvers has been explored in biophysical and biopharmaceutical identification tasks [224]. In practice, partial and biased observation is the rule, motivating differentiable filters (EnKF/particle) for joint state-parameter learning [225] and martingale/moment estimators as complementary, interpretable checks [226]. Bayesian neural SDEs (e.g., SG-MCMC and variational ensembles) deliver posterior trajectories and calibrated intervals for epidemic indicators [227].

When spatial coupling, diffusion, and nonlocal effects dominate, operator learning extends the toolkit beyond low-dimensional SDEs. Neural operators (Fourier/kernel variants) and PINN/PINO-style training have been adapted to SPDE reaction-diffusion and metapopulation settings, improving extrapolation while retaining mechanistic constraints [228, 229]. Long-time alignment through ergodic/statistical losses helps match stationary behavior in biochemical and ecological networks [230], multirate/multiscale architectures address separated time scales [231], and hybrid low-rank constructions (learned spatial bases with latent temporal SDEs) yield compact surrogates for spatially aggregated epidemics [232]. By contrast, rigorous mesh/domain generalization with explicit conservation/positivity guarantees and operator-level Bayesian UQ remains underreported (✗).

Generative SDEs provide complementary capability: score-based diffusion models enable strong priors and data augmentation for biological trajectories and epidemic curves [233], and predictor-corrector samplers support accurate path synthesis and parameter exploration in low-data regimes [234]. Open opportunities remain in controlled reverse-time formulations linked to stochastic optimal control, variance reduction, and reduced-order sampling under mechanistic constraints (✗).

Credible forecasting depends on discretizations that respect biological structure-positivity of compartments, invariance of feasible sets, and stability under stiffness. Recent studies establish strong/weak convergence and stability for SDE/SPDE schemes used in epidemiology, including linearly implicit and explicit positivity-preserving integrators [216, 218, 220]. For stochastic reaction-diffusion SPDEs, unconditionally stable backward-Euler/implicit finite-difference schemes support practical step-size rules [212], and explicit positivity-preserving methods can provide extinction guarantees in subcritical regimes [220]. Still underexplored in applied pipelines are tamed/balanced schemes for superlinear drifts, adaptive time stepping, and multilevel Monte Carlo for forward UQ (✗), despite clear relevance to bursty transmission and regime shifts.

Finally, modern inference and surrogates complete the application pipeline: sequential Monte Carlo supports online calibration under delays and regime changes [215]; likelihood-free neural posterior estimation amortizes inference across mechanistic stochastic epidemic simulators for rapid scenario analysis [211]; neural surrogates for stochastic SIS-type models reduce computational cost in design/control loops [214]; interacting jump-diffusions and common-noise effects clarify synchronization and variability in metapopulations and environmentally forced systems [213]; and stochastic Runge-Kutta methods remain practical tools for vector-borne dynamics and multiscale uncertainty propagation [217]. Across single-cell dynamics, pharmacological response, vector-borne and waterborne outbreaks, and respiratory epidemics, four recurring themes emerge: mechanistically informed learning, partial-observation inference, structure-preserving numerics, and scalable UQ/surrogates for decision support-with the most stable pipelines combining identifiable drift parameterizations, PSD-constrained diffusions, filter-based training, and positivity-preserving or implicit time integration.

Table 3: Recent works (2020-2025) applying the surveyed methods to biology and epidemiology. ✓ = at least one study found; ✗ = no results found to date.

Method	Area	Status	Refs	Notes
A. Neural SDE identification and learning				
Neural SDEs (learned drift/diffusion)	Neural SDE	✓	[210]	scDiffEq from single-cell time series.
MLE (Euler-Maruyama transitions)	SDE ID	✓	[221]	Cholera data; discretized likelihood.
Higher-order integrators (Milstein/Itô-Taylor/SRK)	SDE numerics	✓	[217]	Dengue SDE simulations (SRK).
PSD diffusion (Cholesky)	SDE ID	✓	[222]	PSD covariance parameterization.
Tamed / monotone architectures	Robust learning	✓	[223]	Stable learning for nonlinear epidemic dynamics.
Continuous-time training via SDE solvers	System ID	✓	[224]	Pharmacological dynamics (CT-ESN).
Partial obs. + differentiable filters	State-param. inf.	✓	[225]	Particle/EnKF-style inference (monkeypox).
Moment / martingale estimation	Stochastic proc.	✓	[226]	CTMC moments / MFPT bounds.
Bayesian neural SDEs	Bayesian inf.	✓	[227]	COVID-19 posterior dynamics (Germany).
Bootstrap UQ (seeds/paths)	UQ	✗	-	No results found.
B. Operator learning and PINNs for SPDE / PDE models				
Neural operators (Fourier/kernel)	Operator (SPDE)	✓	[228]	Physics-informed fractional epidemic models.
PINO / PINNs	Operator learn.	✓	[229]	PINN-style training for epidemic dynamics.
Ergodic/statistical losses	Long-time align.	✓	[230]	Ergodicity-guided inference.
Mesh/domain gen. + conservation	Operator (SPDE)	✗	-	No results found.
Multirate / multiscale operators	Operator learn.	✓	[231]	Multilevel modeling for comp. epidemiology.
Operator-level UQ	Operator (SPDE)	✗	-	No results found.
Hybrid latent SDE + spatial bases	Hybrid	✓	[232]	Low-rank hybrid for COVID-19.
C. Generative SDEs and reverse-time modeling				
Reverse-time SDE / prob.-flow ODE	Generative SDE	✓	[233]	Score-based reverse-time view.

Continued on next page

Table 3: Recent works (2020-2025) applying the surveyed methods to biology and epidemiology. ✓ = at least one study found; ✗ = no results found to date. (Continued)

Method	**Area**	**Status**	**Refs**	**Notes**
Predictor-corrector samplers	Generative SDE	✓	[234]	Parameter estimation strategies.
Adaptive/VR/reduced-order sampling	Generative SDE	✗	-	No results found.
Controlled reverse-time (control link)	Generative SDE	✗	-	No results found.
D. Numerical analysis for SDEs / SPDEs				
Strong/weak conv. + stability	Numerics	✓	[216, 218, 220]	Baseline + SIS implicit + positivity schemes.
Implicit SPDE (BE + implicit FD)	SPDE num.	✓	[212]	Unconditionally stable reaction-diffusion.
Linearly implicit SIS schemes	SDE num.	✓	[218]	Stability/accuracy for stiff incidence.
Tamed schemes (superlinear drift)	SDE num.	✗	-	No results found.
Balanced schemes	SDE num.	✗	-	No results found.
Adaptive time stepping	SDE/SPDE num.	✗	-	No results found.
Invariant-measure fidelity	Long-time num.	✓	[216]	Long-horizon bias / invariant measures.
Positivity / invariance preserving	Structure	✓	[220]	Positivity + extinction guarantees.
MLMC / multi-index MLMC	UQ	✗	-	No results found.
E. Inference, surrogates, and phenomena				
Sequential Monte Carlo	Inference	✓	[215]	Online state + time-varying parameter tracking.
Likelihood-free NPE	Inference	✓	[211]	Amortized posteriors for epidemic simulators.
ANN surrogates for stochastic SIS	Surrogates	✓	[214]	LM-trained emulator for SIS.
Jumps / common-noise phenomena	Phenomena	✓	[213]	Jump-diffusion + common-noise variability.
Stochastic Runge-Kutta	SDE num.	✓	[217]	Dengue uncertainty propagation.

6. Outlook and open problems

The convergence of data-driven identification, continuous-time generative modeling, and structure-preserving numerics is reshaping stochastic modeling for biological and epidemiological systems. Table 4 consolidates the most actionable research directions by pairing each open problem with a precise objective, plausible methodological routes, and reportable evaluation criteria.

For neural SDEs under partial observation (especially with state-dependent diffusion) priority must be given to verifiable identifiability and finite-sample guarantees that prevent drift-diffusion confounding and remain robust to irregular sampling. For operator learning, the central requirement is to extend mesh-independent approximation to stochastic forcing while providing stability and well-posedness guarantees (strong solutions, pathwise stability, geometric ergodicity, and, in Bayesian settings, posterior contraction). These guarantees are essential to support principled recovery of dynamics and uncertainty rather than ad hoc curve-fitting (Table 4, "Identifiability" and "Consistency & well-posedness").

Training should also be *discretization-aware*: objectives ought to be optimized under the same integrator used at inference to eliminate train-test discretization mismatch. This motivates memory-stable adjoints and AD-compatible schemes with strong stability/structure properties (e.g., tamed or balanced Milstein variants, linearly implicit IMEX updates, and positivity-preserving projections), together with in-loop adaptivity that separates discretization error from statistical error during learning (Table 4, "Solver-aware training" and "Adaptive error control").

Finally, evaluation should move beyond pointwise intervals toward calibrated, shift-robust sequential forecasting: distinguish aleatoric from epistemic uncertainty, incorporate dependence-aware conformal calibration, and report distributional diagnostics (CRPS/energy scores, PIT) alongside long-horizon bias in invariant statistics. Rare-event and tail-risk quantification remains comparatively underdeveloped; coupling rare-event biasing with multi-index MLMC is a promising route to predictable variance-cost behavior for exceedance probabilities (Table 4, "Tail risk & rare events" and "Benchmarking & reporting").

Table 4. Open problems distilled into goals, candidate methods, and reportable metrics for learned stochastic dynamics and structure-aware numerics.

Theme	**Goal**	**Candidate methods**	**Metrics**
Identifiability (neural SDE)	Disentangle drift vs. diffusion under partial observation and irregular sampling.	Structured diffusion (factorized/low-rank), short-increment penalties; stationary-moment constraints; uniform empirical-process bounds; confounding tests.	Drift/diffusion MSE; misattribution rate; CI coverage.
Consistency & well-posedness	Guarantee strong solutions, stability, ergodicity, and (Bayesian) posterior contraction.	One-sided Lipschitz & linear-growth priors; Lyapunov drift/minorization; Wasserstein contractivity; sieve priors and small-ball conditions.	Stability constants; mixing rate; ergodic bias; contraction rate.
Solver-aware training	Remove train-inference discretization mismatch.	Discretization-aligned losses; reversible/checkpointed adjoints; step-size-aware objectives; gradient-stability control for stiffness.	Generalization gap (train vs. inference); gradient growth; solver-induced bias.
Structure-preserving SPDE integrators	Maintain invariants (mass/positivity) and spectra over long horizons.	Linearly implicit IMEX; positivity-preserving projections; energy/enstrophy-aware updates; invariant-region enforcement.	Invariant error; spectral discrepancy; long-time bias.
Adaptive error control (in-loop)	Separate discretization vs. statistical error and adapt time/mesh during training.	Dual-weighted residuals; MLMC/MIMC gradient estimators; online mesh/time refinement.	Work-RMSE tradeoff; estimator variance; speedup at fixed error.
Tail risk & rare events	Estimate exceedance probabilities with predictable variance-cost.	Schrödinger-bridge or Girsanov tilting; adaptive splitting/AMS; multi-index MLMC couplings.	Relative error at fixed cost; variance-decay exponent; ESS.
Benchmarking & reporting	Make results comparable and auditable across methods/costs.	Fixed data splits; structural baselines; standardized metric suite; compute/energy reporting.	Reproducible leaderboard; compute-normalized scores; energy/carbon budget.

7. Conclusion

This review set out to give researchers a compact, navigable map of recent (2020-2025) advances in stochastic modeling for biology and epidemiology, with two goals: (i) to locate and clarify *novelty* across theory, inference, learning, and numerics; and (ii) to surface *open problems* and underused methods with clear relevance to biological and epidemiological practice. We organized the field along a pipeline from Itô SDE/SPDE formulations to neural SDE identification and operator learning, reverse-time generative SDEs, and structure-preserving integrators, so that readers can see where each contribution fits and how choices in representation, solver, and inference jointly determine scientific reliability.

Concretely, we (a) distilled an identification toolkit for drift-diffusion learning under partial observation (with PSD-constrained diffusion, monotonicity/taming, and differentiable filtering), (b) summarized operator-learning surrogates for spatial dynamics with weak-physics and ergodic/statistical alignment, and (c) reviewed solver-aware computation that maintains feasibility, stability, and long-time statistics while enabling scalable uncertainty quantification. We also charted gaps that invite immediate work in biology/epidemiology: identifiability under irregular sampling and state-dependent noise; operator learning under random forcing with well-posedness/ergodicity guarantees; discretization-aware training with memory-stable adjoints; positivity- and invariant-preserving schemes in routine pipelines; multilevel/multi-index Monte Carlo for tail-sensitive policy metrics; controlled reverse-time SDEs for assimilation and counterfactuals; and operator-level Bayesian UQ with mesh/domain generalization.

By mapping what is new, what is missing, and where methods align with biological and epidemiological objectives (time-varying transmission, spatial spread, regime shifts, and decision-oriented UQ), this review aims to shorten the path from method selection to credible results, reduce reinvention, and focus effort on the most impactful open directions.

Acknowledgments

None.

Funding

None.

Conflict of interest

There is no conflict of interest to disclose.

Author contributions

Yassine Sabbar: Software, Writing - Original draft, Writing- Review & Editing. **Kottakkaran Sooppy Nisar:** Validation, Writing- Review & Editing.

Declaration of using AI tools

The authors declare that they have not used any type of generative artificial intelligence for the writing of this manuscript, nor for the creation of images, graphics, tables, or their corresponding captions.

References

[1] S. Goldstein, B. Misra, and M. Courbage, "On intrinsic randomness of dynamical systems," J. Stat. Phys., vol. 25, no. 1, pp. 111–126, 1981. https://doi.org/10.1007/BF01008481

[2] M. Jebahi, F. Dau, J. L. Charles, and I. Iordanoff, "Multiscale modeling of complex dynamic problems: An overview and recent developments," Arch. Comput. Methods Eng., vol. 23, no. 1, pp. 101–138, 2016. https://doi.org/10.1007/s11831-014-9136-6

[3] N. Demopoulos, Foundations for a Unified Stochastic Framework: Bridging Physics. Biology, and Socio-Economic Systems, 2025.

[4] M. W. Hirsch, "The dynamical systems approach to differential equations," Bull. Am. Math. Soc., vol. 11, no. 1, pp. 1–64, 1984. https://doi.org/10.1090/S0273-0979-1984-15236-4

[5] Q. Lü and X. Zhang, Mathematical control theory for stochastic partial differential equations, vol. 101. Cham: Springer, 2021. https://doi.org/10.1007/978-3-030-82331-3

[6] M. S. Eldred, L. P. Swiler, and G. Tang, "Mixed aleatory-epistemic uncertainty quantification with stochastic expansions and optimization-based interval estimation," Reliab. Eng. Syst. Saf., vol. 96, no. 9, pp. 1092–1113, 2011. https://doi.org/10.1016/j.ress.2010.11.010

[7] R. C. Smith, Uncertainty quantification: theory, implementation, and applications. Society for Industrial and Applied Mathematics, 2024. https://doi.org/10.1137/1.9781611977844

[8] H. Schurz, "A brief review on stability investigations of numerical methods for systems of stochastic differential equations," Netw. Heterog. Media, vol. 19, no. 1, pp. 355–383, 2024. https://doi.org/10.3934/nhm.2024016

[9] N. G. Van Kampen, "Stochastic differential equations," Phys. Rep., vol. 24, no. 3, pp. 171–228, 1976. https://doi.org/10.1016/0370-1573(76)90029-6

[10] C. Salvi, M. Lemercier, and A. Gerasimovics, "Neural stochastic pdes: Resolution-invariant learning of continuous spatiotemporal dynamics," Adv. Neural Inf. Process. Syst., vol. 35, pp. 1333–1344, 2022.

[11] F. Flandoli and E. Luongo, Stochastic partial differential equations in fluid mechanics, vol. 2330. Singapore: Springer, 2023. https://doi.org/10.1007/978-981-99-0385-6

[12] E. E. Holmes, M. A. Lewis, J. E. Banks, and R. R. Veit, "Partial differential equations in ecology: Spatial interactions and population dynamics," Ecology, vol. 75, no. 1, pp. 17–29, 1994. https://doi.org/10.2307/1939378

[13] F. Arceci, M. Maurelli, D. Morale, and S. Ugolini, (2021, September). A reaction diffusion model with a stochastic boundary condition. In INdAM Workshop: Mathematical modeling and Analysis of degradation and restoration in Cultural Heritage (pp. 1-16). Singapore: Springer Nature Singapore.

[14] L. C. Evans, An introduction to stochastic differential equations, vol. 82. American Mathematical Soc., 2012.

[15] C. M. Ahn and H. E. Thompson, "Jump-diffusion processes and the term structure of interest rates," J. Finance, vol. 43, no. 1, pp. 155–174, 1988.

[16] Hewitt, K. (1970). Probabilistic approaches to discrete natural events: a review and theoretical discussion. Economic Geography, 46(sup1), 332-349. https://doi.org/10.2307/143148

[17] I. Eliazar and J. Klafter, "Lévy-driven Langevin systems: Targeted stochasticity," J. Stat. Phys., vol. 111, no. 3, pp. 739–768, 2003. https://doi.org/10.1023/A:1022894030773

[18] T. Bollerslev and V. Todorov, "Estimation of jump tails," Econometrica, vol. 79, no. 6, pp. 1727–1783, 2011. https://doi.org/10.3982/ECTA9240

[19] C. Bayer and S. Breneis, "Markovian approximations of stochastic Volterra equations with the fractional kernel," Quant. Finance, vol. 23, no. 1, pp. 53–70, 2023. https://doi.org/10.1080/14697688.2022.2139193

[20] A. Bondi and F. Flandoli, (2023). On the Kolmogorov equation associated with Volterra equations and Fractional Brownian Motion. arXiv preprint arXiv:.

[21] V. Pipiras and M. S. Taqqu, Long-range dependence and self-similarity, vol. 45. Cambridge university press, 2017. https://doi.org/10.1017/CBO9781139600347

[22] W. B. Gong, Y. Liu, V. Misra, and D. Towsley, "Self-similarity and long range dependence on the internet: A second look at the evidence, origins and implications," Comput. Netw., vol. 48, no. 3, pp. 377–399, 2005. https://doi.org/10.1016/j.comnet.2004.11.026

[23] P. C. Young, P. McKenna, and J. Bruun, "Identification of non-linear stochastic systems by state dependent parameter estimation," Int. J. Control, vol. 74, no. 18, pp. 1837–1857, 2001. https://doi.org/10.1080/00207170110089824

[24] K. Sankaran and S. P. Holmes, "Generative models: An interdisciplinary perspective," Annu. Rev. Stat. Appl., vol. 10, no. 1, pp. 325–352, 2023. https://doi.org/10.1146/annurev-statistics-033121-110134

[25] L. Pacchiardi, (2022). Statistical inference in generative models using scoring rules (Doctoral dissertation, University of Oxford).

[26] R. Cid Fernandes, L. Sodré, H. R. Schmitt, and J. R. Leao, "A probabilistic formulation for empirical population synthesis: Sampling methods and tests," Mon. Not. R. Astron. Soc., vol. 325, no. 1, pp. 60–76, 2001. https://doi.org/10.1046/j.1365-8711.2001.04366.x

[27] D. J. Albers and G. Hripcsak, "Using time-delayed mutual information to discover and interpret temporal correlation structure in complex populations," Chaos, vol. 22, no. 1, p. 013111, Mar. 2012. https://doi.org/10.1063/1.3675621

[28] P. Naveau, A. Hannart, and A. Ribes, "Statistical methods for extreme event attribution in climate science," Annu. Rev. Stat. Appl., vol. 7, no. 1, pp. 89–110, 2020. https://doi.org/10.1146/annurev-statistics-031219-041314

[29] R. J. Zhu, T. Peng, T. Cheng, X. Qu, J. Huang, D. Zhu, et al., (2025). A survey on latent reasoning. arXiv preprint arXiv:.

[30] S. Khatibisepehr, B. Huang, and S. Khare, "Design of inferential sensors in the process industry: A review of Bayesian methods," J. Process Contr., vol. 23, no. 10, pp. 1575–1596, 2013. https://doi.org/10.1016/j.jprocont.2013.05.007

[31] F. S. Richards, "A method of maximum-likelihood estimation," J. R. Stat. Soc. Series B Stat. Methodol., vol. 23, no. 2, pp. 469–475, 1961. https://doi.org/10.1111/j.2517-6161.1961.tb00430.x

[32] J. M. Wooldridge, "Applications of generalized method of moments estimation," J. Econ. Perspect., vol. 15, no. 4, pp. 87–100, 2001. https://doi.org/10.1257/jep.15.4.87

[33] G. E. Box and G. C. Tiao, Bayesian inference in statistical analysis. John Wiley & Sons, 2011.

[34] J. N. Nielsen, H. Madsen, and P. C. Young, "Parameter estimation in stochastic differential equations: An overview," Annu. Rev. Contr., vol. 24, no. 1, pp. 83–94, 2000. https://doi.org/10.1016/S1367-5788(00)00005-5

[35] R. M. Neal, (2011). MCMC using Hamiltonian dynamics. Handbook of markov chain monte carlo, 2(11), 2. 10.1201/b10905-6

[36] Septier, F., & Peters, G. W. (2015). An overview of recent advances in Monte-Carlo methods for Bayesian filtering in high-dimensional spaces. Theoretical aspects of spatial-temporal modeling, 31-61. https://doi.org/10.1007/978-4-431-55336-6_2

[37] N. Kantas, A. Doucet, S. S. Singh, and J. M. Maciejowski, (2009). An overview of sequential Monte Carlo methods for parameter estimation in general state-space models. IFAC Proceedings Volumes, 42(10), 774-785. https://doi.org/10.3182/20090706-3-FR-2004.00129

[38] D. T. Frazier and C. Drovandi, "Robust approximate Bayesian inference with synthetic likelihood," J. Comput. Graph. Stat., vol. 30, no. 4, pp. 958–976, 2021. https://doi.org/10.1080/10618600.2021.1875839

[39] J. M. Lueckmann, G. Bassetto, T. Karaletsos, and J. H. Macke, (2019, January). Likelihood-free inference with emulator networks. In Symposium on Advances in Approximate Bayesian Inference (pp. 32-53). PMLR.

[40] J. S. Chorowski, Statistics for diffusion processes with low and high-frequency observations. Germany: Humboldt Universitaet zu Berlin, 2016.

[41] M. Sorensen, (2024). Efficient estimation for ergodic diffusion processes sampled at high frequency. arXiv preprint arXiv:.

[42] Y. Aït-Sahalia and J. Jacod, "Analyzing the spectrum of asset returns: Jump and volatility components in high frequency data," J. Econ. Lit., vol. 50, no. 4, pp. 1007–1050, 2012. https://doi.org/10.1257/jel.50.4.1007

[43] H. G. Feichtinger and K. Gröchenig, "Theory and practice of irregular sampling," in Wavelets. CRC Press, 2021, pp. 305–363. https://doi.org/10.1201/9781003210450-10

[44] F. Auteri and N. Parolini, "A mixed-basis spectral projection method," J. Comput. Phys., vol. 175, no. 1, pp. 1–23, 2002. https://doi.org/10.1006/jcph.2001.6855

[45] C. Gross and M. Iwen, "Sparse spectral methods for solving high-dimensional and multiscale elliptic PDEs," Found. Comput. Math., vol. 25 , pp. 765–811, 2025.

[46] T. Bai and P. Tahmasebi, "Hybrid geological modeling: Combining machine learning and multiple-point statistics," Comput. Geosci., vol. 142, p. 104519, 2020. https://doi.org/10.1016/j.cageo.2020.104519

[47] X. Liu, T. Xiao, S. Si, Q. Cao, S. Kumar, and C. J. Hsieh, (2019). Neural sde: Stabilizing neural ode networks with stochastic noise. arXiv preprint arXiv:.

[48] Z. Li, H. Zheng, N. Kovachki, D. Jin, H. Chen, B. Liu, et al., "Physics-informed neural operator for learning partial differential equations. ACM/IMS," J. Data Sci., vol. 1, no. 3, pp. 1–27, 2024. https://doi.org/10.6339/24-JDS1158

[49] J. L. Wu, K. Kashinath, A. Albert, D. Chirila, and H. Xiao, "Enforcing statistical constraints in generative adversarial networks for modeling chaotic dynamical systems," J. Comput. Phys., vol. 406, p. 109209, 2020. https://doi.org/10.1016/j.jcp.2019.109209

[50] J. Kwon, S. Ahn, H. Park, and M. Kwon, "Stochastic approximate inference of latent information in epidemic model: A data-driven approach," Signal Processing, vol. 232, p. 109919, 2025. https://doi.org/10.1016/j.sigpro.2025.109919

[51] A. Ruszczyński and A. Shapiro, "Stochastic programming models," Handb. Oper. Res. Manag. Sci., vol. 10, pp. 1–64, 2003. https://doi.org/10.1016/S0927-0507(03)10001-1

[52] A. Shapiro, D. Dentcheva, and A. Ruszczynski, Lectures on stochastic programming: modeling and theory. Society for Industrial and Applied Mathematics, 2021. https://doi.org/10.1137/1.9781611976595

[53] A. Cohen, W. Dahmen, R. A. DeVore, and A. Kunoth, "Multiscale and High-Dimensional Problems," Oberwolfach Reports, vol. 14, no. 1, pp. 1001–1051, 2018. https://doi.org/10.4171/owr/2017/17

[54] Y. Zhang, G. Wang, F. Su, and Z. Qiao, "The theory of response analysis of fuzzy stochastic dynamical systems with a single degree of freedom," Earthquake Eng. Struct. Dynam., vol. 25, no. 3, pp. 235–251, 1996. https://doi.org/10.1002/(SICI)1096-9845(199603)25:3<235::AID-EQE552>3.0.CO;2-3

[55] K. R. Sreenivasan, "Fluid turbulence," Rev. Mod. Phys., vol. 71, no. 2, pp. S383–S395, 1999. https://doi.org/10.1103/RevModPhys.71.S383

[56] F. Noé, A. Tkatchenko, K. R. Müller, and C. Clementi, "Machine learning for molecular simulation," Annu. Rev. Phys. Chem., vol. 71, no. 1, pp. 361–390, Apr. 20 2020. https://doi.org/10.1146/annurev-physchem-042018-052331

[57] R. C. Malone, R. D. Smith, M. E. Maltrud, and M. W. Hecht, "Eddy-resolving ocean modeling," Los Alamos Sci., vol. 28, pp. 223–231, 2003.

[58] M. Aoki, Optimization of stochastic systems: topics in discrete-time systems. Elsevier, 2016.

[59] Y. Huang, X. F. Zha, J. Lee, and C. Liu, "Discriminant diffusion maps analysis: A robust manifold learner for dimensionality reduction and its applications in machine condition monitoring and fault diagnosis," Mech. Syst. Signal Process., vol. 34, no. 1-2, pp. 277–297, 2013. https://doi.org/10.1016/j.ymssp.2012.04.021

[60] Abdulle, A., & Bai, Y. (2014). Reduced-order modelling numerical homogenization. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 372(2021), 20130388. https://doi.org/10.1098/rsta.2013.0388

[61] I. Mezić, "Spectral properties of dynamical systems, model reduction and decompositions," Nonlinear Dyn., vol. 41, no. 1, pp. 309–325, 2005. https://doi.org/10.1007/s11071-005-2824-x

[62] C. Kelly and G. J. Lord, "Adaptive Euler methods for stochastic systems with non-globally Lipschitz coefficients," Numer. Algorithms, vol. 89, no. 2, pp. 721–747, 2022. https://doi.org/10.1007/s11075-021-01131-8

[63] X. Chen, G. Dos Reis, and W. Stockinger, "Wellposedness, exponential ergodicity and numerical approximation of fully super-linear McKean-Vlasov SDEs and associated particle systems," Electron. J. Probab., vol. 30, none, pp. 1–50, 2025. https://doi.org/10.1214/25-EJP1278

[64] Y. Hu, "Semi-implicit Euler-Maruyama scheme for stiff stochastic equations," in Stochastic Analysis and Related Topics V: The Silivri Workshop, 1994. Boston, MA: Birkhäuser Boston, 1996, pp. 183–202. https://doi.org/10.1007/978-1-4612-2450-1_9

[65] T. Sands, "Treatise on analytic nonlinear optimal guidance and control amplification of strictly analytic (non-numerical) methods," Front. Robot. AI, vol. 9, p. 884669, Oct. 4 2022. https://doi.org/10.3389/frobt.2022.884669

[66] U. M. Ascher, S. J. Ruuth, and B. T. Wetton, "Implicit-explicit methods for time-dependent partial differential equations," SIAM J. Numer. Anal., vol. 32, no. 3, pp. 797–823, 1995. https://doi.org/10.1137/0732037

[67] M. Hochbruck and A. Ostermann, "Exponential integrators," Acta Numer., vol. 19, pp. 209–286, 2010. https://doi.org/10.1017/S0962492910000048

[68] T. Y. A. Averina and K. A. Rybakov, "Rosenbrock-type methods for solving stochastic differential equations," Numer. Anal. Appl., vol. 17, no. 2, pp. 99–115, 2024. https://doi.org/10.1134/S1995423924020010

[69] S. Blanes, A. Iserles, and S. Macnamara, "Positivity-preserving methods for ordinary differential equations," Model. Math. Anal. Numer., vol. 56, no. 6, pp. 1843–1870, 2022. https://doi.org/10.1051/m2an/2022042

[70] Y. Sabbar and S. F. Aldosary, "A Novel Stochastic SVIR Model Capturing Transmission Variability Through Mean-Reverting Processes and Stationary Reproduction Thresholds," Mathematics, vol. 13, no. 13, p. 2097, 2025. https://doi.org/10.3390/math13132097

[71] A. Khan, Y. Sabbar, and A. Din, "Stochastic modeling of the Monkeypox 2022 epidemic with cross-infection hypothesis in a highly disturbed environment," Math. Biosci. Eng., vol. 19, no. 12, pp. 13560–13581, Sep. 15 2022. https://doi.org/10.3934/mbe.2022633

[72] N. Privault and L. Wang, "Stochastic SIR Lévy jump model with heavy-tailed increments," J. Nonlinear Sci., vol. 31, no. 1, p. 15, 2021. https://doi.org/10.1007/s00332-020-09670-5

[73] Y. Sabbar, (2021). Mathematical Analysis of Some Stochastic Infectious Disease Models with White Noises and Lévy Jumps (Doctoral dissertation, Université Sidi Mohamed Ben Abdellah de Fès (Maroc)).

[74] Y. Sabbar, "Asymptotic extinction and persistence of a perturbed epidemic model with different intervention measures and standard lévy jumps," Bulletin of Biomathematics, vol. 1, no. 1, pp. 58–77, 2023. https://doi.org/10.59292/bulletinbiomath.2023004

[75] A. Alkhazzan, Y. Guoqiang, W. Zhou, S. M. A. Shah, and Y. Sabbar, "Stochastic modeling of worm propagation in PLC-based industrial systems with general incidence and Lévy noise," Chaos Solitons Fractals, vol. 202, p. 117563, 2026. https://doi.org/10.1016/j.chaos.2025.117563

[76] Y. Sabbar, "Refining extinction criteria in a complex multi-stage epidemic system with non-Gaussian Lévy noise," Commun. Nonlinear Sci. Numer. Simul., vol. 149, p. 108911, 2025. https://doi.org/10.1016/j.cnsns.2025.108911

[77] M. C. Mariani and I. Florescu, Quantitative finance. John Wiley and Sons, 2019. https://doi.org/10.1002/9781118630006

[78] K. C. Kapur and M. G. Pecht, Reliability engineering. John Wiley and Sons, 2014. https://doi.org/10.1002/9781118841716

[79] B. M. Althouse, E. A. Wenger, J. C. Miller, S. V. Scarpino, A. Allard, L. Hébert-Dufresne, et al., "Superspreading events in the transmission dynamics of SARS-CoV-2: Opportunities for interventions and control," PLoS Biol., vol. 18, no. 11, p. e3000897, Nov. 12 2020. https://doi.org/10.1371/journal.pbio.3000897

[80] B. Pei, L. Feng, Y. Li, and Y. Xu, "Non-markovian dynamics: The memory-dependent probability density evolution equations," Nonlinear Dyn., vol. 113, no. 11, pp. 1–19, 2025. https://doi.org/10.1007/s11071-025-11078-3

[81] B. P. Moghaddam, A. Babaei, A. Dabiri, and A. Galhano, "Fractional stochastic partial differential equations: Numerical advances and practical applications—A state of the art review," Symmetry (Basel), vol. 16, no. 5, p. 563, 2024. https://doi.org/10.3390/sym16050563

[82] C. Li, D. Qian, and Y. Chen, "On Riemann-Liouville and caputo derivatives," Discrete Dyn. Nat. Soc., vol. 2011, no. 1, p. 562494, 2011. https://doi.org/10.1155/2011/562494

[83] R. Metzler, W. G. Glöckle, and T. F. Nonnenmacher, "Fractional model equation for anomalous diffusion," Physica A, vol. 211, no. 1, pp. 13–24, 1994. https://doi.org/10.1016/0378-4371(94)90064-7

[84] K. Adolfsson, M. Enelund, and P. Olsson, "On the fractional order model of viscoelasticity," Mech. Time-Depend. Mater., vol. 9, no. 1, pp. 15–34, 2005. https://doi.org/10.1007/s11043-005-3442-1

[85] C. H. Lee, C. H. Juan, H. H. Chen, J. P. Hong, T. W. Liao, I. French, et al., "Long-Range Temporal Correlations in Electroencephalography for Parkinson's Disease Progression," Mov. Disord., vol. 40, no. 2, pp. 266–275, Feb. 2025. https://doi.org/10.1002/mds.30074

[86] Voronenko, V. I., Kovalov, B. L., Kharchenko, M. O., Hrytsenko, P. V., & Omelianenko, V. A. (2024). The Development of the digital transformation of socio-economic and ecological systems.

[87] L. Beghin, F. Mainardi, and R. Garrappa, Eds. Nonlocal and fractional operators. Cham, Switzerland: Springer, 2021. https://doi.org/10.1007/978-3-030-69236-0

[88] J. Ren and Y. Wang, "Overview of object detection algorithms using convolutional neural networks," Journal of Computer and Communications, vol. 10, no. 1, pp. 115–132, 2022.

[89] Z. Lyu, T. Yu, F. Pan, Y. Zhang, J. Luo, D. Zhang, et al., "A survey of model compression strategies for object detection," Multimedia Tools Appl., vol. 83, no. 16, pp. 48165–48236, 2024. https://doi.org/10.1007/s11042-023-17192-x

[90] R. Motorga, V. Mureșan, M. L. Ungureșan, M. Abrudean, H. Vălean, and I. Clitan, "Artificial intelligence in fractional-order systems approximation with high performances: Application in modelling of an isotopic separation process," Mathematics, vol. 10, no. 9, p. 1459, 2022. https://doi.org/10.3390/math10091459

[91] M. Hosseini, M. Horton, H. Paneliya, U. Kallakuri, H. Homayoun, and T. Mohsenin, (2019, June). On the complexity reduction of dense layers from o (n2) to o (nlogn) with cyclic sparsely connected layers. In Proceedings of the 56th Annual Design Automation Conference 2019 (pp. 1-6). https://doi.org/10.1145/3316781.3317873

[92] I. Pan, L. R. Mason, and O. K. Matar, "Data-centric Engineering: Integrating simulation, machine learning and statistics. Challenges and opportunities," Chem. Eng. Sci., vol. 249, p. 117271, 2022. https://doi.org/10.1016/j.ces.2021.117271

[93] M. Shen and C. Cheng, (2025). Neural SDEs as a Unified Approach to Continuous-Domain Sequence Modeling. arXiv preprint arXiv:.

[94] A. Arafa, K. Banawan, K. G. Seddik, and H. V. Poor, "Sample, quantize, and encode: Timely estimation over noisy channels," IEEE Trans. Commun., vol. 69, no. 10, pp. 6485–6499, 2021. https://doi.org/10.1109/TCOMM.2021.3092413

[95] L. Yang, S. Liu, T. Meng, and S. J. Osher, "In-context operator learning with data prompts for differential equation problems," Proc. Natl. Acad. Sci. USA, vol. 120, no. 39, p. e2310142120, Sep. 26 2023. https://doi.org/10.1073/pnas.2310142120

[96] N. Kovachki, Z. Li, B. Liu, K. Azizzadenesheli, K. Bhattacharya, A. Stuart, et al., "Neural operator: Learning maps between function spaces with applications to pdes," J. Mach. Learn. Res., vol. 24, no. 89, pp. 1–97, 2023.

[97] H. Cao, C. Tan, Z. Gao, Y. Xu, G. Chen, P. A. Heng, et al., "A survey on generative diffusion models," IEEE Trans. Knowl. Data Eng., vol. 36, no. 7, pp. 2814–2830, 2024. https://doi.org/10.1109/TKDE.2024.3361474

[98] P. Kumar, "Diffusion Models and Generative Artificial Intelligence: Frameworks, Applications and Challenges," Arch. Comput. Methods Eng., vol. 32, no. 7, pp. 1–44, 2025. https://doi.org/10.1007/s11831-025-10266-z

[99] P. Johri, C. F. Aquadro, M. Beaumont, B. Charlesworth, L. Excoffier, A. Eyre-Walker, et al., "Recommendations for improving statistical inference in population genomics," PLoS Biol., vol. 20, no. 5, p. e3001669, May 31 2022. https://doi.org/10.1371/journal.pbio.3001669

[100] Z. Hu, H. Wang, and Q. Zhou, "A MCMC method based on surrogate model and Gaussian process parameterization for infinite Bayesian PDE inversion," J. Comput. Phys., vol. 507, p. 112970, 2024. https://doi.org/10.1016/j.jcp.2024.112970

[101] J. Donnelly, S. Abolfathi, J. Pearson, O. Chatrabgoun, and A. Daneshkhah, "Gaussian process emulation of spatio-temporal outputs of a 2D inland flood model," Water Res., vol. 225, p. 119100, Oct. 15 2022. https://doi.org/10.1016/j.watres.2022.119100

[102] D. V. Patel, J. Lee, M. W. Farthing, P. K. Kitanidis, and E. F. Darve, (2024). Multi-fidelity Hamiltonian Monte Carlo. arXiv preprint arXiv:.

[103] J. Zhang, "Modern Monte Carlo methods for efficient uncertainty quantification and propagation: A survey," Wiley Interdiscip. Rev. Comput. Stat., vol. 13, no. 5, p. e1539, 2021. https://doi.org/10.1002/wics.1539

[104] M. Nourali, "Comparison of likelihood-free inference approach and a formal Bayesian method in parameter uncertainty assessment: Case study with a single-event rainfall-runoff model," J. Hydrol. Eng., vol. 26, no. 3, p. 05020049, 2021. https://doi.org/10.1061/(ASCE)HE.1943-5584.0002048

[105] D. T. Frazier, D. J. Nott, C. Drovandi, and R. Kohn, "Bayesian inference using synthetic likelihood: Asymptotics and adjustments," J. Am. Stat. Assoc., vol. 118, no. 544, pp. 2821–2832, 2023. https://doi.org/10.1080/01621459.2022.2086132

[106] T. Granger, T. M. Michelitsch, M. Bestehorn, A. P. Riascos, and B. A. Collet, "Stochastic compartment model with mortality and its application to epidemic spreading in complex networks," Entropy (Basel), vol. 26, no. 5, p. 362, Apr. 25 2024. https://doi.org/10.3390/e26050362

[107] Y. Sabbar, "Exploring threshold dynamics of a behavioral epidemic model featuring two susceptible classes and second-order jump-diffusion," Chaos Solitons Fractals, vol. 186, p. 115216, 2024. https://doi.org/10.1016/j.chaos.2024.115216

[108] K. S. Nisar and Y. Sabbar, "Long-run analysis of a perturbed HIV/AIDS model with antiretroviral therapy and heavy-tailed increments performed by tempered stable Lévy jumps," Alex. Eng. J., vol. 78, pp. 498–516, 2023. https://doi.org/10.1016/j.aej.2023.07.053

[109] D. Kiouach and Y. Sabbar, (2019, April). Threshold analysis of the stochastic Hepatitis B epidemic model with successful vaccination and Levy jumps. In 2019 4th World Conference on Complex Systems (WCCS) (pp. 1-6). IEEE.

[110] D. Kiouach and Y. Sabbar, (2019, February). Modeling the impact of media intervention on controlling the diseases with stochastic perturbations. In AIP Conference Proceedings (Vol. 2074, No. 1, p. 020026). AIP Publishing LLC. https://doi.org/10.1063/1.5090643

[111] D. Wiwatanapataphee, Y. H. Wu, W. Sawangtong, P. Sawangtong, and W. Australia, "Modeling anomalous diffusion and volatility in the Australian national electricity market using a space-fractional Black-Scholes framework," AIMS Math., vol. 10, no. 5, pp. 12388–12420, 2025. https://doi.org/10.3934/math.2025560

[112] K. A. Newhall, "A Primer on Stochastic Partial Differential Equations with Spatially Correlated Noise," Annu. Rev. Condens. Matter Phys., vol. 16 , pp.195–208, 2025.

[113] E. M. B. Bouhadjar, Backward SDEs and Applications to Optimal Control Problems (Doctoral dissertation, Université Mohamed Khider (Biskra-Algérie)), 2024.

[114] U. Sirisha, S. P. Praveen, P. N. Srinivasu, P. Barsocchi, and A. K. Bhoi, "Statistical analysis of design aspects of various YOLO-based deep learning models for object detection," International Journal of Computational Intelligence Systems, vol. 16, no. 1, p. 126, 2023. https://doi.org/10.1007/s44196-023-00302-w

[115] M. Kaveh and M. S. Mesgari, "Application of meta-heuristic algorithms for training neural networks and deep learning architectures: A comprehensive review," Neural Process. Lett., vol. 55, no. 4, pp. 1–104, Oct. 31 2022.

[116] N. Gaby, X. Ye, and H. Zhou, "Neural control of parametric solutions for high-dimensional evolution pdes," SIAM J. Sci. Comput., vol. 46, no. 2, pp. C155–C185, 2024. https://doi.org/10.1137/23M1549870

[117] Z. Gao, L. Yan, and T. Zhou, "Adaptive operator learning for infinite-dimensional Bayesian inverse problems," SIAM/ASA J. Uncertainty Quantif., vol. 12, no. 4, pp. 1389–1423, 2024. https://doi.org/10.1137/24M1643815

[118] K. Suresh, C. Severn, and D. Ghosh, "Survival prediction models: An introduction to discrete-time modeling," BMC Med. Res. Methodol., vol. 22, no. 1, p. 207, Jul. 26 2022. https://doi.org/10.1186/s12874-022-01679-6

[119] A. Kaneko, "Multi-stage Euler-Maruyama methods for backward stochastic differential equations driven by continuous-time Markov chains," Jpn. J. Ind. Appl. Math., vol. 41, no. 2, pp. 1223–1276, 2024. https://doi.org/10.1007/s13160-024-00649-3

[120] J. Cai, R. J. B. Goudie, C. Starr, and B. D. M. Tom, "Dynamic factor analysis with dependent Gaussian processes for high-dimensional gene expression trajectories," Biometrics, vol. 80, no. 4,ujae131, Oct. 3 2024. https://doi.org/10.1093/biomtc/ujae131

[121] H. Hoel and S. Ragunathan, "Higher-order adaptive methods for exit times of Itô diffusions," IMA J. Numer. Anal., vol. 44, no. 5, pp. 2821–2863, 2024. https://doi.org/10.1093/imanum/drad077

[122] N. J. Higham, "Cholesky factorization," Wiley Interdiscip. Rev. Comput. Stat., vol. 1, no. 2, pp. 251–254, 2009. https://doi.org/10.1002/wics.18

[123] M. Gonzalez, N. Fernandez Pinto, T. Tran, H. Hajri, and N. Masmoudi, "Seeds: Exponential sde solvers for fast high-quality sampling from diffusion models," Adv. Neural Inf. Process. Syst., vol. 36, pp. 68061–68120, 2023.

[124] D. N. Tanyu, J. Ning, T. Freudenberg, N. Heilenkötter, A. Rademacher, U. Iben, et al., "Deep learning methods for partial differential equations and related parameter identification problems," Inverse Probl., vol. 39, no. 10, p. 103001, 2023. https://doi.org/10.1088/1361-6420/ace9d4

[125] W. Bao, S. Feng, and K. Zhang, (2024). Application of Kalman Filter in Stochastic Differential Equations. arXiv preprint arXiv:.

[126] G. Boaretto and M. P. Laurini, "DSGE Estimation Using Generalized Empirical Likelihood and Generalized Minimum Contrast," Entropy (Basel), vol. 27, no. 2, p. 141, Jan. 30 2025. https://doi.org/10.3390/e27020141

[127] W. Davis and B. Buffett, "Estimation of drift and diffusion functions from unevenly sampled time-series data," Phys. Rev. E, vol. 106, 1-1, p. 014140, Jul. 2022. https://doi.org/10.1103/PhysRevE.106.014140

[128] P. R. Hansen and G. Horel, (2009). Quadratic variation by Markov chains. Univ. of Aarhus Dept. of Economics Research Paper, (2009-13). https://doi.org/10.2139/ssrn.1367519

[129] G. H. Choe, "Girsanov's theorem," in Stochastic Analysis for Finance with Simulations. Cham: Springer International Publishing, 2016, pp. 137–145. https://doi.org/10.1007/978-3-319-25589-7_8

[130] A. L. Lewis, (2001). A simple option formula for general jump-diffusion and other exponential Lévy processes. Available at SSRN 282110. https://doi.org/10.2139/ssrn.282110

[131] L. J. Powers, J. Nešlehová, and D. A. Stephens, "Pricing American Options in an Infinite Activity Lévy Market: Monte Carlo and Deterministic Approaches Using a Diffusion Approximation," in Numerical Methods in Finance: Bordeaux, June 2010. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 291–321. https://doi.org/10.1007/978-3-642-25746-9_9

[132] V. Guan, J. Janssen, H. Rahmani, A. Warren, S. Zhang, E. Robeva, et al., (2024). Identifying drift, diffusion, and causal structure from temporal snapshots. arXiv preprint arXiv:.

[133] Y. Chen and M. J. Wainwright, (2015). Fast low-rank estimation by projected gradient descent: General statistical and algorithmic guarantees. arXiv preprint arXiv:.

[134] S. N. Sharma and H. G. Patel, "The Fokker-Planck equation," in Stochastic Control. Rijeka: IntechOpen, 2010, pp. 1–20.

[135] L. Borcea, T. Hohage, and B. Kaltenbacher, "Computational Inverse Problems for Partial Differential Equations," Oberwolfach Reports, vol. 14, no. 2, pp. 1463–1549, 2018. https://doi.org/10.4171/owr/2017/24

[136] Z. Zou, X. Meng, A. F. Psaros, and G. E. Karniadakis, "NeuralUQ: A comprehensive library for uncertainty quantification in neural differential equations and operators," SIAM Rev., vol. 66, no. 1, pp. 161–190, 2024. https://doi.org/10.1137/22M1518189

[137] P. Bera and S. Bhanja, (2025). Quantification of Uncertainties in Probabilistic Deep Neural Network by Implementing Boosting of Variational Inference. arXiv preprint arXiv:.

[138] X. Chen, (2024). Enhancing Estimation and Uncertainty Quantification in Stochastic Optimization: Importance Sampling and Bootstrap Resampling (Doctoral dissertation, University of California, Davis).

[139] X. Feng, H. Sha, Y. Zhang, Y. Su, S. Liu, Y. Jiang, et al., "Reliable deep learning in anomalous diffusion against out-of-distribution dynamics," Nat. Comput. Sci., vol. 4, no. 10, pp. 761–772, Oct. 2024. https://doi.org/10.1038/s43588-024-00703-7

[140] M. Fuhrman, F. Masiero, and G. Tessitore, "Stochastic equations with delay: Optimal control via BSDEs and regular solutions of Hamilton-Jacobi-Bellman equations," SIAM J. Contr. Optim., vol. 48, no. 7, pp. 4624–4651, 2010. https://doi.org/10.1137/080730354

[141] S. Lototsky and B. Rozovskii, "Stochastic differential equations: a Wiener chaos approach," in From stochastic calculus to mathematical finance: the Shiryaev festschrift. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 433–506. https://doi.org/10.1007/978-3-540-30788-4_23

[142] S. V. Lototsky and B. L. Rozovsky, Stochastic partial differential equations, vol. 11. Cham: Springer, 2017. https://doi.org/10.1007/978-3-319-58647-2

[143] M. Liu-Schiaffini, J. Berner, B. Bonev, T. Kurth, K. Azizzadenesheli, and A. Anandkumar, (2024). Neural operators with localized integral and differential kernels. arXiv preprint arXiv:.

[144] O. R. Dunbar, N. H. Nelsen, and M. Mutic, "Hyperparameter optimization for randomized algorithms: A case study on random features," Stat. Comput., vol. 35, no. 3, pp. 1–28, 2025. https://doi.org/10.1007/s11222-025-10587-w

[145] S. Li, S. U. Khan, K. Yu, E. A. Ismail, and F. A. Awwad, "Stability Analysis of Spectral Technique for Three-Connected Fractional Stochastic Integro-Delay Differential Systems," Int. J. Theor. Phys., vol. 64, no. 4, pp. 1–29, 2025. https://doi.org/10.1007/s10773-025-05969-9

[146] S. Chen, (2025). Knowledge-Guided Machine Learning Methods and Applications to Scientific Data Simulation and Upscaling (Doctoral dissertation, University of Pittsburgh).

[147] N. Chen, Stochastic methods for modeling and predicting complex dynamical systems. Berlin/Heidelberg, Germany: Springer International Publishing, 2023. https://doi.org/10.1007/978-3-031-22249-8

[148] G. Picci, "Stochastic Noises, Observation, Identification and Realization with," in Encyclopedia of Complexity and Systems Science. New York, NY: Springer, 2009, pp. 8749–8766. https://doi.org/10.1007/978-0-387-30440-3_525

[149] M. Grigoriu, "Evaluation of Karhunen-Loève, spectral, and sampling representations for stochastic processes," J. Eng. Mech., vol. 132, no. 2, pp. 179–189, 2006. https://doi.org/10.1061/(ASCE)0733-9399(2006)132:2(179)

[150] B. Haasdonk, K. Urban, and B. Wieland, "Reduced basis methods for parameterized partial differential equations with stochastic influences using the Karhunen-Loeeve expansion," SIAM/ASA J. Uncertainty Quantif., vol. 1, no. 1, pp. 79–105, 2013. https://doi.org/10.1137/120876745

[151] R. Kupferman, G. A. Pavliotis, and A. M. Stuart, "Itô versus Stratonovich white-noise limits for systems with inertia and colored multiplicative noise," Phys. Rev. E Stat. Nonlin. Soft Matter Phys., vol. 70, 3 Pt 2, p. 036120, Sep. 2004. https://doi.org/10.1103/PhysRevE.70.036120

[152] J. Aaronson, An introduction to infinite ergodic theory (No. 50). American Mathematical Soc., 1997. https://doi.org/10.1090/surv/050

[153] B. Després, M. Dumbser, J. R. Kamm, and M. Torrilhon, "High-Resolution Mathematical and Numerical Analysis of Involution-Constrained PDEs," Oberwolfach Reports, vol. 10, no. 3, pp. 2691–2747, 2014. https://doi.org/10.4171/owr/2013/47

[154] A. Abdulle and G. R. de Souza, "Explicit stabilized multirate method for stiff stochastic differential equations," SIAM J. Sci. Comput., vol. 44, no. 4, pp. A1859–A1883, 2022. https://doi.org/10.1137/21M1439018

[155] K. Koval and R. Nicholson, "Non-intrusive optimal experimental design for large-scale nonlinear Bayesian inverse problems using a Bayesian approximation error approach: K.

Koval, R. Nicholson," J. Sci. Comput., vol. 104, no. 3, p. 98, 2025. https://doi.org/10.1007/s10915-025-03008-7

[156] J. Watson and J. Peters, (2021, May). Advancing trajectory optimization with approximate inference: Exploration, covariance control and adaptive risk. In 2021 American Control Conference (ACC) (pp. 1231-1236). IEEE. https://doi.org/10.23919/ACC50511.2021.9482657

[157] H. U. Molla and J. Qiu, "Numerical approximations of coupled forward-backward SPDEs," Stochastic Anal. Appl., vol. 41, no. 2, pp. 291–326, 2023. https://doi.org/10.1080/07362994.2021.2011318

[158] Z. Hao, S. Liu, Y. Zhang, C. Ying, Y. Feng, H. Su, et al., (2022). Physics-informed machine learning: A survey on problems, methods and applications. arXiv preprint arXiv:.

[159] P. Glasserman, Gradient estimation via perturbation analysis, vol. 116. Springer Science & Business Media, 1991.

[160] A. Taghvaei, (2024, December). Time-Reversal of Stochastic Maximum Principle. In 2024 IEEE 63rd Conference on Decision and Control (CDC) (pp. 2597-2602). IEEE. https://doi.org/10.1109/CDC56724.2024.10886762

[161] H. T. Nguyen, J. B. Andersen, G. F. Pedersen, P. Kyritsi, and P. C. Eggers, "Time reversal in wireless communications: A measurement-based investigation," IEEE Trans. Wirel. Commun., vol. 5, no. 8, pp. 2242–2252, 2006. https://doi.org/10.1109/TWC.2006.1687740

[162] B. E. Anderson, R. A. Guyer, T. J. Ulrich, and P. A. Johnson, "Time reversal of continuous-wave, steady-state signals in elastic media," Appl. Phys. Lett., vol. 94, no. 11, p. 111908, 2009. https://doi.org/10.1063/1.3097811

[163] A. Li, W. Fang, H. Zhao, L. Lu, G. Yang, and M. Xu, (2025). MaRS: A Fast Sampler for Mean Reverting Diffusion based on ODE and SDE Solvers. arXiv preprint arXiv:.

[164] A. Solin, E. Tamir, and P. Verma, "Scalable inference in SDEs by direct matching of the Fokker-Planck-Kolmogorov equation," Adv. Neural Inf. Process. Syst., vol. 34, pp. 417–429, 2021.

[165] S. Xue, M. Yi, W. Luo, S. Zhang, J. Sun, Z. Li, et al., "Sa-solver: Stochastic adams solver for fast sampling of diffusion models," Adv. Neural Inf. Process. Syst., vol. 36, pp. 77632–77674, 2023.

[166] E. Dussaud, W. W. Symes, P. Williamson, L. Lemaistre, P. Singer, B. Denel, et al., "Computational strategies for reverse-time migration," SEG Tech. Prog. Exp. Abstr., vol. 2008, pp. 2267–2271, 2008. (Society of Exploration Geophysicists.)

[167] S. Chen, S. Chewi, H. Lee, Y. Li, J. Lu, and A. Salim, "The probability flow ode is provably fast," Adv. Neural Inf. Process. Syst., vol. 36, pp. 68552–68575, 2023.

[168] M. M. Chertkov, "Mixing artificial and natural intelligence: From statistical mechanics to AI and back to turbulence," J. Phys. A Math. Theor., vol. 57, no. 33, p. 333001, 2024. https://doi.org/10.1088/1751-8121/ad67bb

[169] G. Fabiani, (2024). Machine learning-based Modelling and Numerical Analysis of the Emergent Dynamics of Complex Systems (Doctoral dissertation, University of Naples Federico II).

[170] A. Agrawal and C. McComb, (2025). Adaptive Multi-Fidelity Reinforcement Learning for Variance Reduction in Engineering Design Optimization. arXiv preprint arXiv:.

[171] P. Sanchez and S. A. Tsaftaris, (2022). Diffusion causal models for counterfactual estimation. arXiv preprint arXiv:.

[172] H. Hassani, R. Razavi-Far, M. Saif, and L. Lin, (2024). Towards Sample-Efficiency and Generalization of Transfer and Inverse Reinforcement Learning: A Comprehensive Literature Review. arXiv preprint arXiv:.

[173] P. Behnoudfar and N. Chen, (2025). RL-DAUNCE: Reinforcement Learning-Driven Data Assimilation with Uncertainty-Aware Constrained Ensembles. arXiv preprint arXiv:. https://doi.org/10.2139/ssrn.5259995

[174] W. Rüemelin, "Numerical treatment of stochastic differential equations," SIAM J. Numer. Anal., vol. 19, no. 3, pp. 604–613, 1982. https://doi.org/10.1137/0719041

[175] W. Fang and M. B. Giles, (2016, August). Adaptive Euler-Maruyama method for SDEs with non-globally Lipschitz drift. In International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing (pp. 217-234). Cham: Springer International Publishing.

[176] M. Bossy and K. Martínez, (2024). Strong convergence of the exponential Euler scheme for SDEs with superlinear growth coefficients and one-sided Lipschitz drift. arXiv preprint arXiv:.

[177] A. Lang, "Almost sure convergence of a Galerkin approximation for SPDEs of Zakai type driven by square integrable martingales," J. Comput. Appl. Math., vol. 236, no. 7, pp. 1724–1732, 2012. https://doi.org/10.1016/j.cam.2011.10.003

[178] Z. Jia and C. Li, "Almost sure exponential stability of uncertain stochastic Hopfield neural networks based on subadditive measures," Mathematics, vol. 11, no. 14, p. 3110, 2023. https://doi.org/10.3390/math11143110

[179] M. Sharp, (2024). Quantitative Preservations of Ulam Stability-type Estimates. arXiv preprint arXiv:.

[180] C. E. Bréhier, "Approximation of the invariant distribution for a class of ergodic SDEs with one-sided Lipschitz continuous drift coefficient using an explicit tamed Euler scheme," ESAIM Probab. Stat., vol. 27, pp. 841–866, 2023. https://doi.org/10.1051/ps/2023017

[181] G. Rohith and N. K. Sinha, "Routes to chaos in the post-stall dynamics of higher-dimensional aircraft model," Nonlinear Dyn., vol. 100, no. 2, pp. 1705–1724, 2020. https://doi.org/10.1007/s11071-020-05604-8

[182] R. D'Ambrosio, "Numerical Methods for Stochastic Differential Equations," in Numerical Approximation of Ordinary Differential Problems: From Deterministic to Stochastic Numerical Methods. Cham: Springer Nature Switzerland, 2023, pp. 291–363. https://doi.org/10.1007/978-3-031-31343-1_9

[183] F. Djeumou, T. J. Lew, N. Ding, M. Thompson, M. Suminaka, M. Greiff, et al., (2024, November). One model to drift them all: Physics-informed conditional diffusion model for driving at the limits. In 8th Annual Conference on Robot Learning.

[184] C. K. Wai, L. K. Choon, and R. Idris, "Adaptive time-stepping for Runge-Kutta method for ordinary differential equations," Universiti Malaysia Terengganu Journal of Undergraduate Research, vol. 3, no. 1, pp. 25–36, 2021. https://doi.org/10.46754/umtjur.v3i1.189

[185] Z. Liu and Z. Liu, (2024). Unique Ergodicity of Implicit Methods for Monotone SDEs and SPDEs driven by Nondegenerate Multiplicative Noise. arXiv preprint arXiv:.

[186] J. Kuntz, P. Thomas, G. B. Stan, and M. Barahona, "Stationary distributions of continuous-time Markov chains: A review of theory and truncation-based approximations," SIAM Rev., vol. 63, no. 1, pp. 3–64, 2021. https://doi.org/10.1137/19M1289625

[187] A. Ni, "Ergodic and foliated kernel-differentiation method for linear responses of random systems," J. Nonlinear Sci., vol. 35, no. 5, p. 90, 2025. https://doi.org/10.1007/s00332-025-10186-z

[188] M. Perman, J. Pitman, and M. Yor, "Size-biased sampling of Poisson point processes and excursions," Probab. Theory Relat. Fields, vol. 92, no. 1, pp. 21–39, 1992. https://doi.org/10.1007/BF01205234

[189] D. Talay and L. Tubaro, "Expansion of the global error for numerical schemes solving stochastic differential equations," Stochastic Anal. Appl., vol. 8, no. 4, pp. 483–509, 1990. https://doi.org/10.1080/07362999008809220

[190] J. S. Lamb and J. A. Roberts, "Time-reversal symmetry in dynamical systems: A survey," Physica D, vol. 112, no. 1-2, pp. 1–39, 1998. https://doi.org/10.1016/S0167-2789(97)00199-1

[191] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," Nat. Rev. Phys., vol. 3, no. 6, pp. 422–440, 2021. https://doi.org/10.1038/s42254-021-00314-5

[192] R. Dascaliuc, C. Foias, and M. S. Jolly, "Estimates on enstrophy, palinstrophy, and invariant measures for 2-D turbulence," J. Differ. Equ., vol. 248, no. 4, pp. 792–819, 2010. https://doi.org/10.1016/j.jde.2009.11.020

[193] D. Goodair and D. Crisan, Stochastic Calculus in Infinite Dimensions and SPDEs. Springer Nature, 2024. https://doi.org/10.1007/978-3-031-69586-5

[194] F. Y. Wang, "Exponential contraction in Wasserstein distances for diffusion semigroups with negative curvature," Potential Anal., vol. 53, no. 3, pp. 1123–1144, 2020. https://doi.org/10.1007/s11118-019-09800-z

[195] M. B. Giles, "Multilevel monte carlo methods," Acta Numer., vol. 24, pp. 259–328, 2015. https://doi.org/10.1017/S096249291500001X

[196] F. Cornalba and J. Fischer, "Multilevel Monte Carlo methods for the Dean-Kawasaki equation from fluctuating hydrodynamics," SIAM J. Numer. Anal., vol. 63, no. 1, pp. 262–287, 2025. https://doi.org/10.1137/23M1617345

[197] W. Fang and M. B. Giles, "Adaptive Euler-Maruyama method for SDEs with nonglobally Lipschitz drift," Ann. Appl. Probab., vol. 30, no. 2, pp. 526–560, 2020. https://doi.org/10.1214/19-AAP1507

[198] C. Bayer, P. K. Friz, S. Riedel, and J. Schoenmakers, "From rough path estimates to multilevel Monte Carlo," SIAM J. Numer. Anal., vol. 54, no. 3, pp. 1449–1483, 2016. https://doi.org/10.1137/140995209

[199] P. Mycek, (2024). Hierarchical methods for deterministic and stochastic partial differential equations (Doctoral dissertation, Université de Bordeaux).

[200] Y. Bai, Z. Huang, H. Lam, and D. Zhao, "Overconservativeness of variance-based efficiency criteria and probabilistic efficiency in rare-event simulation," Manage. Sci., vol. 70, no. 10, pp. 6852–6873, 2024. https://doi.org/10.1287/mnsc.2023.4973

[201] C. Hartmann and L. Richter, "Nonasymptotic bounds for suboptimal importance sampling," SIAM/ASA J. Uncertainty Quantif., vol. 12, no. 2, pp. 309–346, 2024. https://doi.org/10.1137/21M1427760

[202] H. Singh, F. Xia, A. Subbaswamy, A. Gossmann, and J. Feng, "A hierarchical decomposition for explaining ML performance discrepancies," Adv. Neural Inf. Process. Syst., vol. 37,128516–128555, 2024.

[203] E. C. Duru, M. C. Anyanwu, and M. G. C. Ezike, "A mathematical model to investigate the effect of misdiagnosis and wrong treatment in the co-circulation and co-infection of Malaria and Zika virus disease," Bulletin of Biomathematics, vol. 3, no. 1, pp. 79–110, 2025. https://doi.org/10.59292/bulletinbiomath.1711811

[204] C. Nkeki and I. Mbarie, "On a mathematical model and the efficacy of control measures on the transmission dynamics of chickenpox," Bulletin of Biomathematics, vol. 3, no. 1, pp. 37–61, 2025. https://doi.org/10.59292/bulletinbiomath.1707079

[205] L. L. Iwa, A. Omame, and S. C. Inyama, "A fractional-order model of COVID-19 and Malaria co-infection," Bulletin of Biomathematics, vol. 2, no. 2, pp. 133–161, 2024. https://doi.org/10.59292/bulletinbiomath.2024006

[206] B. Shiri, E. D. Khiabani, and D. Baleanu, "Analysis and analytical solution of incommensurate fuzzy fractional nabla difference systems in neural networks," An International Journal of Optimization and Control: Theories & Applications, vol. 15, no. 4, pp. 610–624, 2025. https://doi.org/10.36922/IJOCTA025130067

[207] Kravitz, H., Duron, C., Nieves, B., & Brio, M. (2025). Data-driven Optimization and Parameter Estimation for a Metric Graph Epidemic Model with applications to COVID-19 spread in Poland: A real-world example of optimization for a challenging Rosenbrock-type objective function. An International Journal of Optimization and Control: Theories & Applications, vol. 15, no. 4, pp. 750-778, 2025. https://doi.org/10.36922/IJOCTA025220106

[208] Z. Hammouch, M. Yavuz, and N. Özdemir, "Numerical solutions and synchronization of a variable-order fractional chaotic system," Mathematical Modelling and Numerical Simulation with Applications, vol. 1, no. 1, pp. 11–23, 2021. https://doi.org/10.53391/mmnsa.2021.01.002

[209] F. Evirgen, E. Uçar, S. Uçar, and N. Özdemir, "Modelling influenza a disease dynamics under Caputo-Fabrizio fractional derivative with distinct contact rates," Mathematical Modelling and Numerical Simulation with Applications, vol. 3, no. 1, pp. 58–73, 2023. https://doi.org/10.53391/mmnsa.1274004

[210] M. E. Vinyard, A. W. Rasmussen, R. Li, A. M. Klein, G. Getz, and L. Pinello, (2023). scDiffEq: drift-diffusion modeling of single-cell dynamics with neural stochastic differential equations. bioRxiv, 2023-12.

[211] P. Chatha, F. Bu, J. Regier, E. Snitkin, and J. Zelner, (2024). Neural posterior estimation for stochastic epidemic modeling. arXiv preprint arXiv:.

[212] N. Ahmed, M. W. Yasin, S. M. Ali, A. Akguel, A. Raza, M. Rafiq, et al., "Computational aspects of an epidemic model involving stochastic partial differential equations," Int. J. Mod. Phys. C, vol. 34, no. 11, p. 2350146, 2023. https://doi.org/10.1142/S0129183123501462

[213] Y. V. Vuong, M. Hauray, and E. Pardoux, "Conditional propagation of chaos in a spatial stochastic epidemic model with common noise," Stoch. Partial Differ. Equ., vol. 10, no. 3, pp. 1180–1210, 2022. https://doi.org/10.1007/s40072-022-00268-4

[214] N. Anwar, K. Shahzadi, M. A. Zahoor Raja, I. Ahmad, M. Shoaib, and A. K. Kiani, "Intelligent adaptive networks for chaos analysis in stochastic SIS differential epidemic model with the impact of unpredictable environmental factors," Fluct. Noise Lett., vol. 24, no. 02, p. 2550006, 2025. https://doi.org/10.1142/S0219477525500063

[215] A. Fadikar, A. Stevens, N. Collier, K. B. Toh, O. Morozova, A. Hotton, et al., (2024, May). Towards Improved Uncertainty Quantification of Stochastic Epidemic Models Using Sequential Monte Carlo. In 2024 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW) (pp. 843-852). IEEE. https://doi.org/10.1109/IPDPSW63119.2024.00151

[216] Hoang, M. T., & Ehrhardt, M. (2025). Differential equation models for infectious diseases: Mathematical modeling, qualitative analysis, numerical methods and applications: MT Hoang, M. Ehrhardt. SeMA Journal, 1-36.

[217] N. Anwar, I. Ahmad, H. Javaid, A. K. Kiani, M. Shoaib, and M. A. Z. Raja, "Stochastic Runge-Kutta for numerical treatment of dengue epidemic model with Brownian uncertainty," Mod. Phys. Lett. B, vol. 39, no. 03, p. 2450408, 2025. https://doi.org/10.1142/S0217984924504086

[218] Z. Yang, H. Yang, and Y. Wang, "Numerical analysis of linearly implicit θ methods for a stochastic SIS epidemic model," Commun. Nonlinear Sci. Numer. Simul., vol. 152 , p. 109140, 2026.

[219] Anwar, N., Ahmad, I., Kiani, A. K., Shoaib, M., & Raja, M. A. Z. (2024). Novel neuro-stochastic adaptive supervised learning for numerical treatment of nonlinear epidemic delay differential system with impact of double diseases. International Journal of Modelling and Simulation, 1-23.

[220] H. Yang and J. Huang, "Strong convergence and extinction of positivity preserving explicit scheme for the stochastic SIS epidemic model," Numer. Algorithms, vol. 95, no. 4, pp. 1475–1502, 2024. https://doi.org/10.1007/s11075-023-01617-7

[221] Y. M. Marwa, I. S. Mbalawata, S. Mwalili, and W. M. Charles, "Stochastic dynamics of cholera epidemic model: Formulation, analysis and numerical simulation," Z. Angew. Math. Phys., vol. 7, no. 5, pp. 1097–1125, 2019.

[222] S. Rawat and S. Martiniani, "Element-wise and recursive solutions for the power spectral density of biological stochastic dynamical systems at fixed points," Phys. Rev. Res., vol. 6, no. 4, p. 043179, Sep. 5 2024. https://doi.org/10.1103/PhysRevResearch.6.043179

[223] B. Prasse and P. Van Mieghem, "Network reconstruction and prediction of epidemic outbreaks for general group-based compartmental epidemic models," IEEE Trans. Netw. Sci. Eng., vol. 7, no. 4, pp. 2755–2764, 2020. https://doi.org/10.1109/TNSE.2020.2987771

[224] R. Anantharaman, A. Abdelrehim, A. Jain, A. Pal, D. Sharp, A. Edelman, et al., "Stably accelerating stiff quantitative systems pharmacology models: Continuous-time echo state networks as implicit machine learning," IFAC-PapersOnLine, vol. 55, no. 23, pp. 1–6, 2022. https://doi.org/10.1016/j.ifacol.2023.01.004

[225] V. E. Papageorgiou and P. Kolias, "A novel epidemiologically informed particle filter for assessing epidemic phenomena. Application to the monkeypox outbreak of 2022," Inverse Probl., vol. 40, no. 3, p. 035006, 2024. https://doi.org/10.1088/1361-6420/ad1e2f

[226] M. Backenköhler, L. Bortolussi, and V. Wolf, (2020, August). Bounding mean first passage times in population continuous-time Markov chains. In International Conference on Quantitative Evaluation of Systems (pp. 155-174). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-59854-9_13

[227] S. T. Radev, F. Graw, S. Chen, N. T. Mutters, V. M. Eichel, T. Bärnighausen, et al., "OutbreakFlow: Model-based Bayesian inference of disease outbreak dynamics with invertible neural networks and its application to the COVID-19 pandemics in Germany," PLOS Comput. Biol., vol. 17, no. 10, p. e1009472, Oct. 25 2021. https://doi.org/10.1371/journal.pcbi.1009472

[228] E. Kharazmi, M. Cai, X. Zheng, Z. Zhang, G. Lin, and G. E. Karniadakis, "Identifiability and predictability of integer- and fractional-order epidemiological models using physics-informed neural networks," Nat. Comput. Sci., vol. 1, no. 11, pp. 744–753, Nov. 2021. https://doi.org/10.1038/s43588-021-00158-0

[229] C. Millevoi, D. Pasetto, and M. Ferronato, "A Physics-Informed Neural Network approach for compartmental epidemiological models," PLOS Comput. Biol., vol. 20, no. 9, p. e1012387, Sep. 5 2024. https://doi.org/10.1371/journal.pcbi.1012387

[230] B. G. de Figueiredo, J. M. Calabrese, W. F. Fagan, and R. Martinez-Garcia, (2024). Ergodicity shapes inference in biological reactions driven by a latent trajectory. arXiv preprint arXiv:.

[231] S. Picault, Y. L. Huang, V. Sicard, T. Hoch, E. Vergu, F. Beaudeau, et al., (2018). A generic multi-level stochastic modelling framework in computational epidemiology. bioRxiv, 491605.

[232] K. K. L. Hwang, C. J. Edholm, O. Saucedo, L. J. S. Allen, and N. Shakiba, "A hybrid epidemic model to explore stochasticity in COVID-19 dynamics," Bull. Math. Biol., vol. 84, no. 9, p. 91, Jul. 20 2022. https://doi.org/10.1007/s11538-022-01030-6

[233] N. W. Lenner, (2020). Reverse-time inference of biological dynamics (Doctoral dissertation, Georg-August-Universität Göttingen).

[234] S. Rogowski, Parameter Estimation Techniques for Nonlinear Biological Models. The Florida State University, 2024.