

Parameter inference for PDE models based on Physics Informed Neural Networks: applications to modelling problems in plant and animal health with destructive measurements.

Beatrice Laroche & Lorenzo Sala

beatrice.laroche@inrae.fr

lorenzo.sala@inrae.fr

Context and Motivation

There is a strong link between Stochastic Differential Equations (SDEs) and partial differential equations (PDEs). As a very simple example, let us consider individuals described by their position x on the real line \mathbb{R} , whose motion is driven by a constant speed v perturbed by a Gaussian white noise with constant variance σ^2 . This means that the position follows the SDE $dx = vdt + \sigma dB_t$. Denoting $p(x, t)$ the probability for an individual to occupy position x at time t , p satisfies the Fokker-Planck equation (also known as the progressive Kolmogorov equation, see for instance [1,2])

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}(vp) + \frac{\partial^2}{\partial x^2}\left(\frac{\sigma^2}{2}p\right), \quad p(x, 0) = p_0(x)$$

More generally, under suitable regularity conditions, such a PDE system can be associated with an SDE system.

SDEs, which incorporate mechanistic knowledge through the drift term, while capturing intrinsic variability through the noise term, provide a suitable framework for modelling individual behaviours within biological populations. However, in many real biological systems, access to individual trajectories in time is impossible because key measurements - such as microbiota composition, immune status, or gene expression - require individual destruction (e.g., in microbial, plant, or animal systems). This renders classical inference strategies, which rely on rich longitudinal data for each individual, inapplicable. Although the literature on parameter estimation for SDEs is extensive [4,3,7], most contributions address settings with dense longitudinal observations on the same individual. In contrast, situations in which measurements are destructive or time sampling is very sparse have received very limited methodological attention, despite their importance in agricultural, ecological, and biomedical applications.

To overcome the impossibility of observing full SDE trajectories, this project proposes an alternative strategy: instead of simulating individual paths of the SDE we propose to consider the associated Fokker-Planck PDE. The likelihood of an observation then becomes $p(x_i, t)$, making the inference formulation insensitive to the destructive nature of the measurements.

Despite its theoretical appeal, this PDE-based approach has historically been considered computationally infeasible when the system's dimension exceeds 2-3, due to the curse of dimensionality [3]. Recent progress in Physics-Informed Neural Networks (PINNs) offers a promising pathway to overcoming this limitation. PINNs can efficiently approximate high-dimensional PDE solutions, reducing computational costs and enabling likelihood-based inference even in higher-dimensional systems [5,6]. Their integration into parameter estimation—through direct embedding of parameters, inverse PINN formulations, or hybrid approaches combining PINNs with classical methods such as MCMC—constitutes a central methodological challenge of this PhD.

The methodology will be validated on real biological datasets, with possible applications in three complementary domains:

- plant-pathogen interactions (tomato under *Botrytis cinerea* stress), in collaboration with ISA (Sophia Antipolis);
- host-microbiota-pathogen dynamics in poultry infected with *Salmonella*, in collaboration with ISP (Tours);
- biofilm microbiota-pathogen systems, in collaboration with Micalis (Paris-Saclay).

These settings provide rich yet destructive datasets and serve as ideal testbeds for assessing the robustness, scalability, and transferability of the proposed methodology across biological domains.

PhD outline

The PhD will develop a Fokker-Planck PDE-based framework for parameter inference in SDE models using destructive and sparse biological measurements. The work will consist of:

- (i) formulating inference problems under partial or noisy observations;
- (ii) developing PINN-based metamodels to solve Fokker-Planck PDEs efficiently in moderate-to-high dimensions;
- (iii) integrating these metamodels into parameter estimation, either directly or through hybrid strategies (e.g., PINN-MCMC);
- (iv) applying and validating the methodology on real datasets from plant-pathogen interactions, poultry host-microbiota-pathogen dynamics, and biofilm microbiota-pathogen systems.

The doctoral candidate will have access to a dedicated laptop, local cluster computing resources, and GPU facilities at the Lab-IA for large-scale computations, as well as real datasets already available from the SMILE (ISA), MiMoSa (ISP), and B3D (Micalis) teams.

Candidate Skills

The PhD candidate should possess strong mathematical modelling skills (SDEs, PDEs), very good programming experience in Python, and ideally some familiarity with machine learning. A strong interest in interdisciplinary biological applications and collaborative research is essential.

References

- [1] B. Øksendal. "Stochastic differential equations." *Stochastic differential equations: an introduction with applications*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003. 38-50.
- [2] <https://perso.lpsm.paris/~nfournier/PolyCS.pdf>
- [3] P. F. Craigmile et al. Statistical inference for stochastic differential equations. *Wiley Interdisciplinary Reviews: Computational Statistics*, 15(2):e1585, 2023.
- [4] S. Donnet and A. Samson. A review on estimation of stochastic differential equations for pharmacokinetic / pharmacodynamic models. *Advanced Drug Delivery Reviews*, 65:929–939, 2013.
- [5] J. Han, A. Jentzen, and W. E. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences*, 115(34):8505–8510, 2018.
- [6] Z. Hu et al. Tackling the curse of dimensionality with physics-informed neural networks. *Neural Networks*, 176:106369, 2024.
- [7] Q. A. Liu, D. Varshney, and K. B. McAuley. Parameter and uncertainty estimation in stochastic differential equation models with multi-rate data and nonstationary disturbances. *Chemical Engineering Research and Design*, 183:118–133, 2022.