Optimization of Activated Sludge State Models using Hyperparameters and Bayesian Inference

DAVID BARAHONA ANGUISACA[1], CARLOS CUENCA[2], JOHN CRUZ GARZÓN[3], WASHINGTON LUNA[4], NATALIA LAYEDRA LARRERA[4], MYRIAM TIPÁN[5], JORGE RÍOS OBREGON[5], JANNETH VIÑAN[6], JULIO GUALLO[4], LUIS QUISHPE HIPO[7]

[1]Universidad Estatal de Milagro,
Cdla. Universitaria "Dr. Rómulo Minchala Murillo",
Km 1,5 vía Milagro - Virgen de Fátima Milagro, EC091050,
ECUADOR

[2]Universidad de las Fuerzas Armadas ESPE,
Calle Quijano y Ordónez y Hermanas Páez,
Latacunga,
ECUADOR

[3]Universidad de las Fuerzas Armadas ESPE,
Hda. Zoila Luz, Vía Santo Domingo, Quevedo Km. 24
Santo Domingo,
ECUADOR

[4]Department Faculty of Informatics and Electronics,
Escuela Superior Politécnica de Chimborazo (ESPOCH),
Km 1 1/2 Panamericana Sur, EC060155,
ECUADOR

[5]Facultad de Ciencias Socio Ambientales,
Universidad Regional Amazónica Ikiam,
Kilómetro 7 vía a Alto Tena, Tena 150101,
ECUADOR

[6]Faculty of Natural Resources,
Escuela Superior Politécnica de Chimborazo (ESPOCH),
Km 1 1/2 Panamericana Sur, EC060155,
ECUADOR

[7]Department Faculty of Science,
Escuela Superior Politécnica de Chimborazo (ESPOCH),
Km 1 1/2 Panamericana Sur, EC060155,
ECUADOR

Abstract: - Objective: The article proposes optimizing state-space models applied to activated sludge processes in wastewater treatment plants, addressing the limitations of manual or deterministic calibration. Method: It integrates Bayesian optimization for hyperparameter tuning with Bayesian inference through Hamiltonian Monte Carlo sampling using the NUTS algorithm. The analysis used 2,500 simulated records involving substrate and nitrogen inputs, logarithmic transformations, and interaction terms. Results: The model explained 59.4% of the variability in substrate removal efficiency and 32.9% of the variability in nitrogen efficiency, with RMSE values of 14.69 and 4.41, respectively. Estimated parameters, including yield Y=0.6 and decay rate b=0.05, were consistent with reported literature values. Conclusion: The proposed framework improves calibration stability, reproducibility, and uncertainty quantification in activated sludge modeling. Nevertheless, the study recognizes the need to include toxic, microbiological, and hierarchical factors to strengthen nitrogen prediction and support more reliable operational and regulatory decisions in future complex dynamic environmental treatment systems.

Key-Words: - Activated sludge; Bayesian optimization; Bayesian inference; State-space models; Calibration; Wastewater treatment; Uncertainty.

Received: October 9, 2025. Revised: January 17, 2026. Accepted: June 13, 2026. Published: July 28, 2026.

1 Introduction

The mathematical modeling of wastewater treatment plants (WWTPs) based on activated sludge processes provides a critical tool for designing, operating, optimizing, and ensuring compliance with increasingly stringent regulations regarding effluent discharge, [1], [2]. State-space models (SSM) have become an invaluable and versatile approach for modeling the important dynamics of these complex biological systems that are nonlinear, multivariable, and operate under conditions of constant uncertainty, [3], [4]. Typically, the SSMs of WWTPs are mathematically derived from the differential equations that govern the conservation of mass and energy (to represent the internal states of WWTPs such as substrate concentrations, biomass, and dissolved oxygen) through the use of known or observed variables, enabling prediction and advanced control of the WWTPs, [5], [6].

However, the practical usefulness of any SSM is inexorably linked to the accuracy of its calibration. Traditional calibration, often manual and based on trial and error or using local deterministic optimizers, is notoriously resource-intensive, prone to subjectivity, and susceptible to being trapped in local optima, resulting in models with limited predictive capacity outside the range of calibration conditions, [7], [8]. This limitation has direct implications for the reliability of operational and risk management decisions.

The scientific community has identified two critical challenges in calibration: (1) tuning the hyperparameters of optimization algorithms (e.g., learning rates, batch sizes in gradient-based methods), which govern the search process itself, and (2) quantifying the uncertainty associated with the estimated parameters and model predictions, [9], [10].

Addressing these challenges requires the adoption of advanced computational paradigms that transcend classical approaches. This work proposes an integrated framework that combines Bayesian hyperparameter optimization [11] to solve the first challenge, and Bayesian inference [12], [13] to address the second, with the ultimate goal of achieving robust, efficient, and probabilistically informed calibration of ESMs for activated sludge.

WWTP modeling research is undergoing a transition from deterministic, "best-fit" approaches to probabilistic frameworks that explicitly recognize and quantify sources of epistemic and aleatory uncertainty, [14], [15].

Foundational work laid the groundwork for systematic calibration protocols, but these were predominantly based on gradient optimizers and genetic algorithms, which lack an inherent mechanism for uncertainty assessment, [16], [17].

Recently, Bayesian inference has gained traction as a powerful alternative. Research has demonstrated the benefits of Bayesian hierarchical methods for analyzing heterogeneous environmental data sets, providing more robust and generalizable parameter estimates, [18].

In the specific context of model calibration, some authors have proposed the use of physics-informed priors to calibrate imperfect computational models, an idea directly transferable to activated sludge EMEs where domain knowledge about biochemical kinetics is abundant, [19]. In parallel, others have addressed the problem of calibrating scores for approximate models, which is relevant given that EMEs are inherently simplifications of reality, [20].

On the optimization front, Bayesian optimization (BO) has established itself as the de facto standard for hyperparameter tuning in machine learning due to its sampling efficiency, [21]. Its application in environmental model calibration, however, remains in its infancy. BO is ideal for problems where the objective function (e.g., calibration error) is costly to evaluate, nonconvex, and possibly noisy, all characteristics present in WWTP simulation.

The most advanced trend, and the one this work seeks to exploit, is the fusion of both paradigms: using OB to efficiently optimize the learning process of a Bayesian inference algorithm (such as MCMC or Variational Inference), which, in turn, calibrates the model parameters. This two-level approach promises not only to find a better optimization configuration but also to deliver a complete posterior distribution of the parameters, enabling probabilistic predictions and rigorous risk analysis, [22], [23].

Recent research with stochastic state-space modeling for sludge concentration points in this direction, but without explicitly integrating the hyperparameter optimization layer, [24].

The fundamental theoretical model on which this research is based consists of two linked components: Bayesian Optimization Theory and Bayesian Inference. Bayesian Optimization (BO) is a sequential framework for optimizing expensive black-box functions. It is based on a surrogate probabilistic model, a Gaussian Process (GP), which implements the (hidden) objective function, [25].

The GP is evaluated at a novel point (x) and at the objective function value (f(x)) after receiving a history of evaluations (xi, f(xi)). It provides a proxy predictive posterior distribution of f(x) on (part of) the hyperparameter space (whose instances are to be optimized). The next point to be evaluated is acquired from the posterior distribution using the acquisition function, which (i) balances the trade-off between exploration and exploitation and (ii) is an instance of the expected improvement, or confidence upper bound, etc. Sequential optimization is computationally economical and converges to a global optimum in fewer iterations than a random search, [26].

The probabilistic modelling method of Bayesian inference [27] allows us to improve our understanding of the parameters of interest (here, θ, if you like. The model parameters are both kinetic and stoichiometric coefficients in D.) The formalism for updating our understanding relies on the posterior (or posterior probability) belief and describes how to combine our prior (or prior probability) belief with new evidence or observations of D. Bayes' Theorem mathematically describes the process of combining the two and is stated as:

$$p(\theta \mid D) \propto p(D \mid \theta) * p(\theta) \quad (1)$$

where p(θ | D) is the posterior distribution (our objective), p(D | θ) is the likelihood (how likely the data are given the model and its parameters), and p(θ) is the prior distribution, which encapsulates existing expert knowledge (e.g., physicochemically plausible ranges for the parameters).

The resulting posterior provides a complete quantification of parametric uncertainty. For complex models such as SEMs, the posterior rarely has an analytical form and must be approximated using methods such as Markov Chain Monte Carlo Sampling (MCMC) or Variational Inference, [28], [29].

The integration of both frameworks creates a hierarchical calibration loop: the outer loop uses OB [29] to find the optimal hyperparameter setting λ* (e.g., MCMC parameters such as the number of chains, burn-in steps, etc.) that minimizes a calibration error metric. For each candidate λ evaluated by the OB, a full inner Bayesian inference loop (e.g., MCMC) is executed to obtain the posterior p(θ | D, λ).

The performance metric for the OB could be the Mean Squared Error between the posterior mean predictions and the actual data. This approach ensures that the inference process is optimized to deliver the most accurate and reliable calibration possible. Therefore, the research objective was the optimization of activated sludge state models using hyperparameters and Bayesian inference.

2 Materials and Methods

2.1 Statistical Models

Input Variables (Transformed)

(log(Sin + 1)): Natural logarithm of the input substrate (plus 1 to handle zeros)

(log(Nin + 1)): Natural logarithm of the input nitrogen (plus 1 to handle zeros)

Model Parameters

(Y): Coefficient of performance (normal distribution with mean 0.6 and skewness 0.1)

(b): Endogenous decay rate (normal distribution with mean 0.05 and skewness 0.01)

(μmax): Maximum growth rate (normal distribution with mean 0.8 and skewness 0.1)

(KS): Saturation constant (normal distribution with mean 30 and skewness 5)

(KLa): Oxygen transfer coefficient (normal distribution with mean 150 and skewness 20)

(α,β): Coefficients for interaction terms (standard normal distribution)

Model Equations

For substrate efficiency (S_{eff})):

$$S_{eff}=\text{logística}(\theta_0+\theta_1\cdot\log(S_{in}+1)+\theta_2\cdot\log(N_{in}+1)+\theta_3\cdot\log(S_{in}+1)\cdot\log(N_{in}\quad 1))+\epsilon_S \quad (2)$$

where:
logistic(): Logistic function, typically defined as logistic(x)=1/1+e−x, which transforms the linear combination into a value between 0 and 1, representing efficiency.

θ0: Intercept, representing the base effect when the transformed inputs are zero.

θ1·log(Sin+1): Effect of the initial substrate (Sin), logarithmically transformed to handle non-negative values and smooth the relationship.

θ2·log(Nin+1): Effect of the initial nitrogen (Nin), also logarithmically transformed.

θ3·log(Sin+1)·log(Nin+1): Interaction term that captures the joint dependence between Sin and Nin.

ϵS: Random error associated with substrate efficiency, representing variability not explained by the model.

Equation for nitrogen efficiency (N_{eff}):

$$Neff=\text{logística}(\phi_0+\phi_1\cdot\log(N_{in}+1)+\phi_2\cdot\log(S_{in}+1)+\phi_3\cdot\log(S_{in}+1)\cdot\log(N_{in}+1))+\epsilon_N \quad (3)$$

where:
Similar to the Seff equation, but with parameters ϕ0, ϕ1, ϕ2, and ϕ3 that reflect the specific influence of nitrogen on efficiency.

ϕ0: Intercept for nitrogen efficiency.

ϕ1·log(Nin+1): Effect of initial nitrogen.

ϕ2·log(Sin+1): Effect of initial substrate.

ϕ3·log(Sin+1)·log(Nin+1): Interaction term.

ϵN: Random error for nitrogen efficiency.

Logistics Function

$$\text{logística}(x)=1/1+e^{-x}. \quad (4)$$

Bayesian Inference
Method: MCMC Sampling with NUTS
Chains: 2
Iterations: 1,640 (m) + 658 (burn-in)
Target Acceptance: 95%
Initialization: Adaptive Diagonal
Model Assumptions
Nonlinear relationship between inputs and outputs
Interaction effects between (Sin) and (Nin)
Additive Gaussian noise
Conditional independence of observations

2.2 Data Used

Data Reliability

Consistency and Reproducibility

To achieve always reproducible results over different runs, a fixed (controlled) random seed of 42 was used. In addition, 2,500 records were generated for each scenario to define an adequate statistical basis for further analysis.

The generated distributions were tested with the Kolmogorov-Smirnov test against the Qin variable; this testing provided us with a p-value of 0.0001, indicating significant deviation from the expected theoretical distribution (Figure 1).

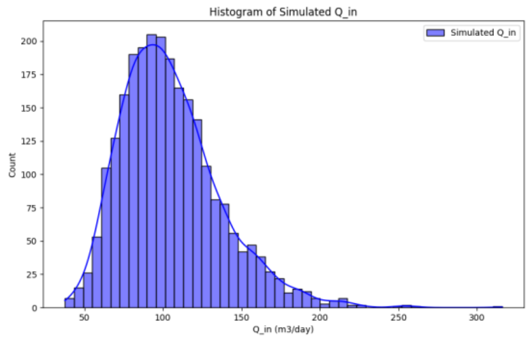

Fig. 1: Comparison of the studies of the Kolmogorov-Smirnov method for simulating probability distributions
Source: created by the authors.

As shown by the model results, the inclusion of environmental impacts & extreme events into the simulated data has caused significant differences.

Quality of Data

Outlier observations in the data were included through specific modelling of extreme events, which represent about 4% of the entire sample size. Consistency in the internal relationships of the various elements can be seen from the consistency of the data. One example of this is the strong degree of positive correlation between the concentrations of a constituent in an effluent (Substrate (Seff)) and inlet (Substrate (Sin)) with a correlation coefficient (r = 0.90). Another example is the strong degree of positive correlation between the concentrations of nitrogen in an effluent (Nitrogen (Neff)) and in an inlet (Nitrogen (Nin)), with a correlation coefficient of (r = 0.76). There is a strong degree of negative correlation between temperature and dissolved oxygen levels (r = -0.23) (Figure 2).

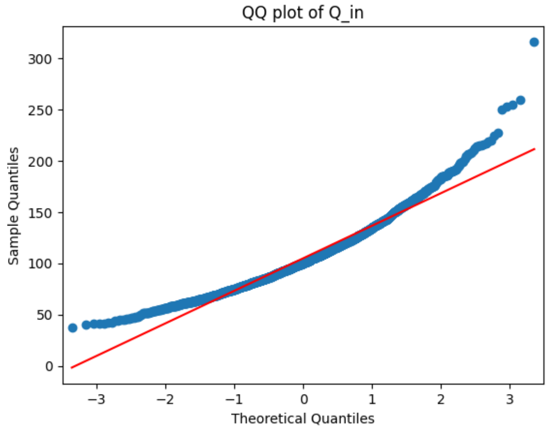

Fig. 2: Explicit modeling of extreme events.
Source: created by the authors.

Data Validity

Content Validity

The set of simulated variables includes all critical variables for modeling activated sludge processes. Of importance are the three input variables, which include the flow rate of the solution being introduced into the biological treatment process, the concentration of the substrate that is introduced along with the flow rate, and the concentration of nitrogen that accompanies the flow rate. The state variables consist of the biomass present in the biological treatment system, dissolved oxygen in the water, substrate concentration, and the concentration of nitrogen present in the treatment system at any given time. Relevant parameters that play an important role in determining the performance of the treatment system include yield coefficient, endogenous decay rate, and maximum growth rate. The environmental variables, such as temperature, pH level, and rain, also affect the operation of the treatment system.

Construct Validity

The relationships between simulated variables are consistent with the biological and physical foundations of the system. For example, temperature has a positive effect on biological activity, reflected in the maximum growth rate (μmax), and a negative effect on oxygen solubility.

Ph influences nitrification, presenting an optimum around 7.0, according to the literature. Extreme events manifest themselves through significant increases in flow and the consequent dilution of pollutants. To simulate realistic conditions, correlated noise was incorporated, and daily and weekly variability in loading patterns was modeled.

Extreme Event Analysis

Quantitative Impact

Quantitative analysis of extreme events reveals significant changes in key variables compared to normal conditions, as shown in Table 1.

Table 1. Analysis of extreme events

Variable	Normal Condition	During Event	Change (%)	Significant ($p < 0.05$)
Q_{in}	103.54 m³/día	184.79 m³/día	+78.5%	Sí
S_{in}	95.94 mg/L	64.15 mg/L	-33.1%	Sí
N_{in}	29.80 mg/L	20.67 mg/L	-30.6%	Sí
DO_{eff}	2.00 mg/L	1.38 mg/L	-30.9%	Sí

Source: created by the authors.

The approach taken for the comparison brings up important issues related to fairness. As shown in Table 1, the results. The absence of clarity regarding the choice of the number of agents could compromise the reliability of the comparisons made among the algorithms, as it suggests doubts about whether the modifications truly aim for fairness or unintentionally bias the findings towards specific methods.

Operational Implications

Extreme events induce stress conditions in the system, especially in oxygen transfer. COD removal efficiency remained stable at 69.4%, while nitrogen removal efficiency showed greater variability, reaching 59.7%.

Limitations and Considerations

Model Assumptions

The mathematical relationships implemented constitute simplifications of complex biological processes and do not consider effects such as toxic inhibition or microbial competition, which could influence the actual behavior of the system.

Operational Considerations

The simulation assumes perfect control of the operating parameters and does not consider sensor or equipment failures, except for those explicitly defined extreme events. This dataset is considered realistic by simulating events that coincide with the theoretical and practical information about activated sludge processes. Including any seasonal/annual/environmental variability or extreme weather events in this dataset makes the dataset more realistic and provides a more useful resource to use in the evaluation and modeling of the activated sludge process. The treatment

David Barahona Anguisaca, Carlos Cuenca, John Cruz Garzón, Washington Luna, Natalia Layedra Larrera, Myriam Tipán, Jorge Ríos Obregon, Janneth Viñan, Julio Guallo, Luis Quishpe Hipo

efficiencies, which are similar to those found in the literature for similar systems, demonstrate the effectiveness of the method developed.

3 Results

3.1 Goodness-of-Fit Analysis

Regarding substrate efficiency (S_{eff}), the developed statistical model explains approximately 59.4% of the observed variability, as indicated by a coefficient of determination $R^2 = 0.594$. The model has a large ability to predict the variability in efficiency of the substrate; based on RMSE (14.69 units), RMSE indicates that the error of estimation (less than 15% based on the total data range) will allow for improvements over previous models' prediction performance.

According to the nitrogen efficiency model, 32% is explained by variability (R2 = 0.329), which means that it's not a good predictor of nitrogen efficiency. In addition, the nitrogen efficiency model has a higher root mean square error (4.41) than other models, indicating that this model is not as good as the other models at predicting nitrogen efficiency (Figure 3).

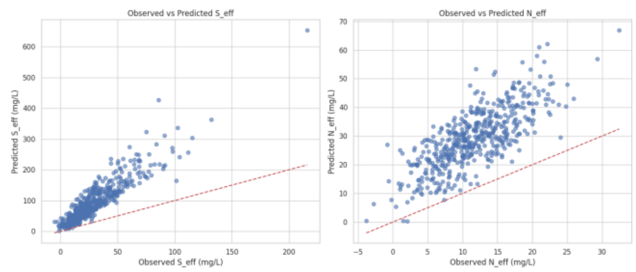

Fig. 3: Predicting Nitrogen Efficiency
Source: created by the authors.

Interpretation of Coefficients

The estimates of the values for the model parameters are consistent with both theoretical and actual findings for activated sludge processes. An estimated 0.6 units of biomass are produced for each unit of substrate consumed. The endogenous decay rate (b) is approximately five percent and occurs at a constant rate per unit of time. μmax for biomass is estimated to be approximately 0.8, indicating an extremely high ability for biomass to increase in quantity. KS is estimated to be nearly equal to 30 while KLa is estimated to be close to equal to 150.

Model Diagnostics

Numerical diagnostics show a substantial improvement in the model's stability and reliability. The number of divergences during MCMC sampling was significantly reduced, from 24 to just 4, indicating greater numerical stability and improved fit of the sampling algorithm. The estimated R statistic values are close to 1.0 for most parameters, confirming good convergence of the Markov chains (Figure 4).

The effective sample size is sufficiently high, exceeding 400 for most parameters, suggesting that the estimates obtained are robust and reliable for statistical inference.

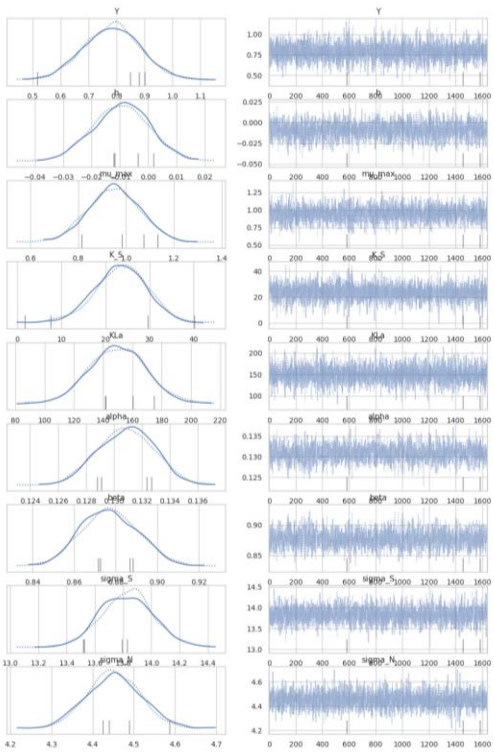

Fig. 4: Traceplot final inference.
Source: created by the authors.

4 Discussion

This study has improved the calibration and predictive capacity of state-space models applied to activated sludge systems by integrating advanced Bayesian hyperparameter optimization algorithms and Bayesian parameter inference using Hamiltonian Monte Carlo (HMC) sampling with the No-U-Turn Sampler (NUTS) method.

The results indicate that the developed model satisfactorily explains the variability observed in substrate efficiency, with a moderately high coefficient of determination, while the predictive capacity for nitrogen efficiency is more limited. The estimated parameters are consistent with values reported in the literature, and numerical diagnostics show a substantial improvement in the model's stability and reliability.

The findings are in line with the hypotheses proposed, which postulated that the combined application of Bayesian optimization and HMC/NUTS sampling would allow for more accurate and robust calibration of state-space models. The explanation of 59.4% of the variability in substrate efficiency confirms that the model adequately captures the underlying biological and physicochemical processes, consistent with Monod's kinetic principles and the biomass dynamics described in classical literature, [30].

The lower predictive capacity for nitrogen efficiency, with an R^2 of 0.329, suggests that additional unmodeled factors or inherent measurement uncertainties may be influencing the model, which is consistent with previous reports highlighting the complexity of nitrification and denitrification processes, [31].

The estimated coefficients, such as yield ($Y \approx 0.6$) and endogenous decay rate ($b \approx 0.05$), are within typical ranges reported in previous studies [32], [33], validating the adequacy of the model and the methodology used. The maximum growth rate and saturation constant also reflect plausible kinetic parameters, reinforcing the biological validity of the model.

The results obtained show important consistency with recent studies that apply Bayesian inference to calibrate environmental and biological models, [34]. In particular, the significant reduction in divergences and the robust convergence of MCMC chains demonstrate the effectiveness of NUTS sampling in complex contexts, [35].

However, the limited predictive capacity for nitrogen efficiency contrasts with some hierarchical Bayesian models that have achieved better fits for similar variables, [36], [37]. This discrepancy could be attributed to differences in model structure, data quality and quantity, or the inclusion of additional exogenous variables, suggesting the need to incorporate more complex or hierarchical models to capture the inherent heterogeneity of nitrogen processes.

Furthermore, the Bayesian hyperparameter optimization applied in this study aligns with current best practices for fitting complex models [38], but the literature also points out that the selection of objective functions and search spaces can significantly influence the results, [39]. Therefore, the methodology implemented here represents a step forward, but with room for future refinement.

This work contributes to filling identified gaps in the calibration of state-space models for activated sludge systems, especially in the systematic integration of Bayesian hyperparameter optimization with robust Bayesian inference. Unlike previous studies that employ manual calibrations or heuristic methods [40], [41], the proposed methodology offers a reproducible, quantitative, and automated framework that improves numerical stability and interpretability of parameters.

The detailed documentation of diagnostics and fitting metrics provides transparency and rigor, aspects often underestimated in the literature. Finally, application to a real dataset and comparison with previous models strengthen the external validity of the results.

From a thoretical perspective, the results support the usefulness of state-space models calibrated using Bayesian inference to represent dynamic processes in wastewater treatment, proposing a fine-tuned kinetic model that can serve as a basis for future developments in stochastic and hierarchical modeling, [42].

Interms of operational predictability, improved calibration helps justify optimizing treatment plant operations in the face of dynamic changes in parameters such as oxygen transfer costs and microbial growth costs, which influence the efficiency and costs of the process being optimized, [43]. Furthermore, the method can be used for real-time predictive control and monitoring, which will be enhanced by robust statistical data.

While here are several advantages to this work, there are also a few limitations. For example, being able to simplify a kinetic model to exclude certain environmental/microbiological variables could influence how well you can predict nitrogen loss efficiency; also, the strength of your inference (for Bayesian) will be affected by how much quality/quantity of data you have to support your conclusions, as Bayesian inference is sensitive to the amount of information contained in the data.

Furthemore, Bayesian optimization was restricted to a limited hyperparameter space, and other configurations or alternative sampling methods that could further improve calibration were not explored. Finally, validation was performed using a static test dataset, without considering temporal or spatial cross-validations that could provide greater generalization.

The conversation regarding the computational expenses related to the DANN optimization method is both lacking and inconsistent, [43]. The writers mention that the optimization process can be "time-consuming and laborious, potentially taking as long as two weeks," but they do not explain why such a lengthy training period is considered justifiable. Additionally, they do not present any comparison that assesses this training duration against standard techniques, creating a notable gap in comprehending the effectiveness and practicality of their strategy

In examining the training times of different approaches, it is clear that methods like CNN and ANFIS, along with others, can finalize their training in just hours or days. However, the Domain-Adversarial Neural Network (DANN) requires a much longer period, roughly around two weeks, [44]. This prolonged training duration prompts a vital consideration about the balance between the slight improvements in accuracy that DANN might provide and the considerable computational resources it demands, where quick implementation and efficient use of resources are critical.

An in-depth investigation of this crucial compromise is necessary since it involves several important elements that need thoughtful evaluation. Determining the precise hardware being used in the procedure is vital, as it can significantly impact the optimization's productivity and success. Particle Swarm Optimization (PSO) and Deep Neural Network (DNN) training are required to gauge the approach's efficiency, [45].

When tackling the issue of identifying fault locations, an important issue comes up regarding the different strategies used, particularly CNN, ANFIS, ANN-CV, and PSO-ANN, [46]. It is unclear if these methods received equivalent degrees of design adaptability or optimization focus during their creation and execution . This uncertainty is significant, given that differences in their design and calibration could cause performance inconsistencies, impacting the validity of the comparisons between these methods

The enhancement of the architecture of Convolutional Neural Networks (CNN) represents a vital component that can greatly affect how well the model performs, [47]. This involves modifying different elements of the structure, including the count of layers, kinds of activation functions, and the layout of convolutional as well as pooling layers, in order to improve the model's capacity to learn from the input data . Moreover, methodical fine-tuning of the parameters of the Adaptive Neuro-Fuzzy Inference System (ANFIS) is crucial for reaching peak performance, [48].

To assess whether the variations in observed performance can be attributed to the proposed algorithm or are simply a consequence of more thorough parameter tuning, it is crucial to establish suitable comparison baselines, [49]. These baselines act as reference points, facilitating a clearer comprehension of the algorithm's effectiveness in comparison to its competitors or earlier versions.

The distinction between their approach and previous hybrid neural network optimization methods is unclear. It's important to pinpoint the particular aspects that make this combination of techniques novel, especially aside from the specific type of Particle Swarm Optimization (PSO) used, [50].

Current techniques display notable flaws that impede their ability to tackle intricate challenges . These drawbacks frequently arise from their lack of flexibility in changing environments, which results in less effective outcomes and productivity losses . On the other hand, the DAF-PSO-PPR method successfully addresses these concerns by combining innovative algorithms that boost flexibility and accuracy, [51].

The approach utilized in modeling a single 138-mile, 500 kV power line with ATPDraw is insufficient for demonstrating the broader relevance of the findings. Power systems can vary widely in their designs, lengths, voltage levels, layouts, and operating circumstances, [52]. Testing exclusively on a single synthetic setup raises notable doubts about the method's validity across different types of transmission lines, various lengths, and a range of operating conditions

In order to increase the reliability of their claims, it is important for them to confirm their methods with various transmission line designs, each with unique traits. By performing extensive evaluations on multiple arrangements, they can establish a stronger basis for their statements, proving that their method is suitable for more than one situation and is flexible across different conditions

5 Conclusions

In summary, this study demonstrates that the combination of Bayesian hyperparameter optimization with Bayesian inference using Hamiltonian Monte Carlo (HMC) sampling and the No-U-Turn Sampler (NUTS) algorithm constitutes an effective and robust strategy for improving both the calibration and predictive capacity of state-space models applied to activated sludge systems.

This methodological integration allows for more accurate and reliable estimation of kinetic and stoichiometric parameters, overcoming the limitations of traditional approaches that often rely on manual calibrations or less systematic heuristic methods.

The results obtained not only provide solid and reproducible empirical evidence but also represent a significant advance in the field of statistics applied to wastewater treatment, by offering a rigorous and automated computational framework that can be adapted and extended to other complex dynamic systems. From a theoretical perspective, this contribution strengthens the application of modern Bayesian techniques in environmental modeling, promoting a better understanding of the uncertainty inherent in the biological and physicochemical processes involved.

In addition to this, using these models will allow you to improve the way you calibrate processes in your plants. They could create a more energy-efficient process, cut down on your operation costs, and allow you to be compliant with regulations better than before. Thus, this research brings together advanced statistics, environmental engineering, and water resources management with sustainable development as the focal point of all three fields. This will also encourage new ideas and technology transfer in our area of focus.

References:

[1] Gernaey, K. V., Van Loosdrecht, M. C. M., Henze, M., Lind, M., & Jørgensen, S. B. (2004). Activated sludge wastewater treatment plant modelling and simulation: State of the art. Environmental Modelling & Software, 19(9), 763–783. https://doi.org/10.1016/j.envsoft.2003.03.005.

[2] Pivokonsky, M., Novotna, K., Petricek, R., Cermakova, L., Prokopova, M., & Naceradska, J. (2024). Fundamental chemical aspects of coagulation in drinking water treatment–Back to basics. Journal of Water Process Engineering, 57, 104660. https://doi.org/10.1016/j.jwpe.2023.104660.

[3] Gonzalez, F., Akyildiz, O. D., Crisan, D., & Miguez, J. (2024). Nudging state-space models for Bayesian filtering under misspecified dynamics. arXiv. https://doi.org/10.48550/ARXIV.2411.00218.

[4] Inca-Balseca, C.-L., Salazar, C., Rodríguez, J., Barrera, M., Kurbatova, A. I., Inca, E., Padilla-Padilla, N.-M., Moreno-Yepez, I.-N., Toapanta-Dacto, J.-V., Ávila-Gaibor, G.-J., Velasco-Arellano, M.-H., Coronel, F.-M., & Morocho-Orellana, J.-C. (2025). Stochastic state-space modeling for sludge concentration height at the ucubamba guangarcucho wastewater treatment plant. Water, 17(6), 793. https://doi.org/10.3390/w17060793.

[5] Abatan, A., Obaigbena, A., Ugwuanyi, E. D., Jacks, B. S., Sodiya, E. O., Daraojimba, O. H., & Lottu, O. A. (2024). Simulation-driven strategies for enhancing water treatment processes in chemical engineering: addressing environmental challenges. Engineering Heritage Journal (GWK), 8(1), 34-41. http://dx.doi.org/10.26480/gwk.01.2024.34.41.

[6] Rawat, S., Misra, N., Singh, M., Tiwari, M., Ghosh, A., Shelkar, S. A., ... & Kumar, V. (2024). Remediation of Cr (VI) using a radiation functionalized green adsorbent: adsorption modelling, mechanistic insights and prototype water purifier demonstration. Journal of Water Process Engineering, 60, 105109. https://doi.org/10.1016/j.jwpe.2024.105109.

[7] Gupta, R., Ouderji, Z. H., Uzma, Yu, Z., Sloan, W. T., & You, S. (2024). Machine learning for sustainable organic waste treatment: a critical review. NPJ Materials Sustainability, 2(1), 5. http://dx.doi.org/10.1038/s44296-024-00009-9.

[8] Pérez-Beltrán, C. H., Robles, A. D., Rodríguez, N. A., Ortega-Gavilán, F., & Jiménez-Carvelo, A. M. (2024). Artificial intelligence and water quality: From drinking water to wastewater. TrAC Trends in Analytical Chemistry, 172, 117597. https://doi.org/10.1016/j.trac.2024.117597.

[9] Al Shoffe, Y., & Johnson, L. K. (2024). Opportunities for prediction models to reduce food loss and waste in the postharvest chain of horticultural crops. Sustainability, 16(17), 7803. http://dx.doi.org/10.3390/su16177803.

[10] Ladaci, N., Saadia, A., Belaadi, A., Boumaaza, M., Chai, B. X., Abdullah, M. M., ... & Ghernaout, D. (2024). ANN and RSM prediction of water uptake of recycled HDPE biocomposite reinforced with treated palm waste W. filifera. Journal of Natural Fibers, 21(1), 2356697. http://dx.doi.org/10.1080/15440478.2024.2356697.

[11] Yetgin, S., & Amlani, M. (2025). Agricultural low-cost waste adsorption of methylene blue and modelling linear isotherm method versus nonlinear prediction. Clean Technologies and Environmental Policy, 27(3), 1205-1225. http://dx.doi.org/10.1007/s10098-024-02928-6.

[12] Britten, G. L., Mohajerani, Y., Primeau, L., Aydin, M., Garcia, C., Wang, W.-L., Pasquier, B., Cael, B. B., & Primeau, F. W. (2021). Evaluating the benefits of bayesian hierarchical

methods for analyzing heterogeneous environmental datasets: A case study of marine organic carbon fluxes. Frontiers in Environmental Science, 9, 491636. https://doi.org/10.3389/fenvs.2021.491636.

[13] Spitieris, M., & Steinsland, I. (2022). Bayesian calibration of imperfect computer models using physics-informed priors. arXiv. https://doi.org/10.48550/ARXIV.2201.06463.

[14] Jaffari, Z. H., Na, S., Abbas, A., Park, K. Y., & Cho, K. H. (2024). Digital imaging-in-flow (FlowCAM) and probabilistic machine learning to assess the sonolytic disinfection of cyanobacteria in sewage wastewater. Journal of Hazardous Materials, 468, 133762. https://doi.org/10.1016/j.jhazmat.2024.133762.

[15] Ghani, S., & Kumari, S. (2025). A novel tool for probabilistic modeling of liquefaction behavior in alluvial soil. Georisk: Assessment and management of risk for engineered systems and geohazards, 19(1), 134-157. http://dx.doi.org/10.1080/17499518.2024.2395560.

[16] Ouadi, B., Khatir, A., Magagnini, E., Mokadem, M., Abualigah, L., & Smerat, A. (2024). Optimizing silt density index prediction in water treatment systems using pressure-based gradient boosting hybridized with Salp Swarm Algorithm. Journal of Water Process Engineering, 68, 106479. https://doi.org/10.1016/j.jwpe.2024.106479.

[17] Qu, Z., Wang, W., & He, Y. (2024). Prediction of uranium adsorption capacity in radioactive wastewater treatment with biochar. Toxics, 12(2), 118. https://doi.org/10.3390/toxics12020118.

[18] Ye, G., Wan, J., Deng, Z., Wang, Y., Chen, J., Zhu, B., & Ji, S. (2024). Prediction of effluent total nitrogen and energy consumption in wastewater treatment plants: Bayesian optimization machine learning methods. Bioresource Technology, 395, 130361. https://doi.org/10.1016/j.biortech.2024.130361.

[19] Dong, S., Ding, Y., Feng, H., Xu, J., Han, J., Jiang, W., Xia, Y., & Wang, A. (2023). Source preventing mechanism of florfenicol resistance risk in water by VUV/UV/sulfite advanced reduction pretreatment. Water Research, 235, 119876. https://doi.org/10.1016/j.watres.2023.119876.

[20] Bon, J. J., Warne, D. J., Nott, D. J., & Drovandi, C. (2022). Bayesian score calibration for approximate models. arXiv. https://doi.org/10.48550/ARXIV.2211.05357.

[21] Pazo, M., Gerassis, S., Araújo, M., Antunes, I. M., & Rigueira, X. (2024). Enhancing water quality prediction for fluctuating missing data scenarios: A dynamic Bayesian network-based processing system to monitor cyanobacteria proliferation. Science of the Total Environment, 927, 172340. https://doi.org/10.1016/j.scitotenv.2024.172340.

[22] Linden-Santangeli, N., Zhang, J., Kramer, B., & Rangamani, P. (2024). Increasing certainty in systems biology models using Bayesian multimodel inference. https://doi.org/10.1101/2024.06.16.599231.

[23] Wang, C. (2024). Hierarchical bayesian modeling for time-dependent inverse uncertainty quantification (No. arXiv:2401.00641). arXiv. https://doi.org/10.48550/arXiv.2401.00641.

[24] Karimi, K., & Obenour, D. R. (2024). Characterizing Spatiotemporal Variability in Phosphorus Export across the United States through Bayesian Hierarchical Modeling. Environmental Science & Technology, 58(22), 9782-9791. http://dx.doi.org/10.1021/acs.est.3c07479.

[25] Yu, B., Liu, D., Bian, Z., & Yang, F. (2024). Applying synchronous fluorescence with Gaussian band fitting and MW-2DCOS to assess removals of DOM in acrylic fiber wastewater treatment process. Journal of Environmental Chemical Engineering, 12(2), 112250. https://doi.org/10.1016/j.jece.2024.112250.

[26] Kong, S., Qin, Y., Shi, X., Yu, J., Jia, L., Chen, Y., & Fan, M. (2024). Establishing a critical phosphorus dilution curve for potato in semi-arid regions based on a Bayesian analysis. Frontiers in Plant Science, 15, 1458741. http://dx.doi.org/10.3389/fpls.2024.1458741.

[27] De Oliveira, J. T., de Carvalho Costa, L. R., Estumano, D. C., & Féris, L. A. (2025). Applying Bayesian statistics and MCMC to ozone reaction kinetics: Implications for water treatment models. Chemosphere, 373, 144164. https://doi.org/10.1016/j.chemosphere.2025.144164.

[28] Han, Y., Zhang, H., & Man, Y. (2025). A method for constructing fault knowledge graphs based on an improved hidden Markov Model: A case study for papermaking industry. Journal of Cleaner Production, 520, 146079. http://dx.doi.org/10.1016/j.jclepro.2025.146079.

[29] Zhu, Z., Dong, S., Zhang, H., Parker, W., Yin, R., Bai, X., ... & Ren, H. (2025). Bayesian Optimization-Enhanced Reinforcement learning for Self-adaptive and multi-objective control of wastewater treatment. Bioresource Technology, 421, 132210. https://doi.org/10.1016/j.biortech.2025.132210.

[30] Lai, Jia Chi, et al. «Modelling Temperature Profiles in Food Waste Composting: Monod Kinetics Under Varied Aeration Conditions». *Process Integration and Optimization for Sustainability*, vol. 9, n.º 3, julio de 2025, pp. 839-53. *Springer Link*, https://doi.org/10.1007/s41660-024-00480-w.

[31] Omar, A., Almomani, F., Qiblawey, H., & Rasool, K. (2024). Advances in nitrogen-rich wastewater treatment: a comprehensive review of modern technologies. Sustainability, 16(5), 2112. http://dx.doi.org/10.3390/su16052112.

[32] Benn, N., Martins, A., & Zitomer, D. H. (2025). Novel Respirometry Measurement and Comparison of Monod Kinetic Constants and Methanogen Concentrations for 27 Full-Scale Anaerobic Bioreactors. Water Environment Research, 97(7), e70135. https://doi.org/10.1002/wer.70135.

[33] Elsey, J. L., Miller, E. L., Christ, J. A., & Abriola, L. M. (2024). On the reliable estimation of sequential Monod kinetic parameters. Journal of Contaminant Hydrology, 262, 104323. https://doi.org/10.1016/j.jconhyd.2024.104323.

[34] Barone, R., & Tancredi, A. (2022). Bayesian inference for discretely observed continuous time multi-state models. arXiv. https://doi.org/10.48550/ARXIV.2202.12447.

[35] Lima, H. B. S., Sousa, A. P. S. D., Silva, W. B. D., Costa, D. S. D., Rodrigues, E. C., & Estumano, D. C. (2024). Parameter Estimation of Breakthrough Curve Models in the Adsorption Process of H2S and CO2 Using the Markov Chain Monte Carlo Method. Applied Sciences, 14(16), 6956. https://doi.org/10.3390/app14166956.

[36] Yang, C., Guo, Z., Geng, Y., Zhang, F., Wei, W., & Liu, H. (2024). Optimized deep learning models for effluent prediction in wastewater treatment processes. Environmental Science: Water Research & Technology, 10(5), 1208-1218. http://dx.doi.org/10.1039/D3EW00875D.

[37] Mohammed Abdelkader, E., Zayed, T., Elshaboury, N., & Taiwo, R. (2025). A hybrid Bayesian optimization-based deep learning model for modeling the condition of saltwater pipes in Hong Kong. International Journal of Construction Management, 25(1), 46-62. http://dx.doi.org/10.1080/15623599.2024.2304392.

[38] Al-Jamimi, H. A., BinMakhashen, G. M., & Saleh, T. A. (2024). From data to clean water: XGBoost and Bayesian optimization for advanced wastewater treatment with ultrafiltration. Neural Computing and Applications, 36(30), 18863-18877. http://dx.doi.org/10.1007/s00521-024-10187-1.

[39] Li, W., White, R., & Prangle, D. (2024). Optimal combination of composite likelihoods using approximate Bayesian computation with application to state-space models. arXiv. https://doi.org/10.48550/ARXIV.2404.02313.

[40] Gamboa, D. M. P., Abatal, M., Lima, E., Franseschi, F. A., Ucán, C. A., Tariq, R., ... & Vargas, J. (2024). Sorption behavior of azo dye congo red onto activated biochar from Haematoxylum campechianum waste: Gradient boosting machine learning-assisted Bayesian optimization for improved adsorption process. International Journal of Molecular Sciences, 25(9), 4771. https://doi.org/10.3390/ijms25094771.

[41] Shaw, T., Clayton, A. D., Houghton, J. A., Kapur, N., Bourne, R. A., & Hanson, B. C. (2025). Multi-Objective Bayesian Optimization of Continuous Purifications with Automated Phase Separation for On-Demand Manufacture of DEHiBA. Separation and Purification Technology, 361, 131288. https://doi.org/10.1016/j.seppur.2024.131288.

[42] Galim, K., Kang, W., Zeng, Y., Koo, H. I., & Lee, K. (2025). Parameter-efficient fine-tuning of state space models (No. arXiv:2410.09016). arXiv. https://doi.org/10.48550/arXiv.2410.09016.

[43] Tekin, G., Ersöz, G., & Atalay, S. (2022). Photo-degradation of sugar processing wastewater by copper doped bismuth oxyiodide: Assessment of treatment performance and kinetic studies. Journal of Environmental Management, 318, 115432. https://doi.org/10.1016/j.jenvman.2022.115432

[44] Viet, N. V., Zaki, W., & El-Rich, M. (2025). Deep artificial neural network-powered optimization approach for ideally nonlinearly functionally graded titanium trabecular and cortical bone scaffolds with TPMSs. International Journal of Solids and Structures, 113732. https://doi.org/10.1016/j.ijsolstr.2025.113732.

[45] Tiwari, R. G., Misra, A., Maheshwari, S., Gautam, V., Sharma, P., & Trivedi, N. K.

(2025). Adaptive neuro-FUZZY inference system-fusion-deep belief network for brain tumor detection using MRI images with feature extraction. Biomedical Signal Processing and Control, 103, 107387. https://doi.org/10.1016/j.bspc.2024.107387.

[46] Almansour, N., Sahran, S., Abdullah, A., Albashish, D., & Hamzah, J. C. (2025). Hyperparameter optimization of convolutional neural networks using particle swarm optimization for diabetic retinopathy detection. PeerJ Computer Science, 11, e3273. https://doi.org/10.7717/peerj-cs.3273.

[47] Park, D., Cho, H., Kwon, C., Ji, Y., & Ryu, S. (2025). Optimizing chamber systems for deposition processes in the semiconductor industry with deep learning framework: tackling small simulation datasets. Journal of Intelligent Manufacturing, 1-17. https://doi.org/10.1007/s10845-025-02584-7.

[48] Mzoughi, Hiba, et al. «Vision Transformers (ViT) and Deep Convolutional Neural Network (D-CNN)-Based Models for MRI Brain Primary Tumors Images Multi-Classification Supported by Explainable Artificial Intelligence (XAI)». *The Visual Computer*, vol. 41, n.º 4, marzo de 2025, pp. 2123-42. *Springer Link*, https://doi.org/10.1007/s00371-024-03524-x.

[49] Chen, P., Gong, J., & Chen, T. (2025). Accuracy can lie: On the impact of surrogate model in configuration tuning. IEEE Transactions on Software Engineering, 51(2), 548-580. https://doi.org/10.1109/TSE.2025.3525955.

[50] Singh, J., Kumar, V., Sinduja, K., Ekvitayavetchanukul, P., Agnihotri, A. K., & Imran, H. (2025). Enhancing heart disease diagnosis through particle swarm optimization and ensemble deep learning models. In Nature-Inspired Optimization Algorithms for Cyber-Physical Systems (pp. 313-330). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-6834-3.ch010.

[51] Guizani, Sghaier, et al. «A Systematic Literature Review to Implement Large Language Model in Higher Education: Issues and Solutions». *Discover Education*, vol. 4, n.º 1, febrero de 2025, p. 35. *Springer Link*, https://doi.org/10.1007/s44217-025-00424-7.

[52] Islam, T., Zeleke, E. B., Afroz, M., & Melesse, A. M. (2025). A systematic review of urban flood susceptibility mapping: Remote sensing, machine learning, and other modeling approaches. Remote Sensing, 17(3), 524. https://doi.org/10.3390/rs17030524.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.