Data-driven modelling of nitrous oxide production in wastewater treatment processes using neural ordinary differential equations

A Thesis Submitted for the Degree of Doctor of Philosophy

By

Xiangjun Huang

Department of Civil and Environmental Engineering, Brunel University London

2025

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisors, Professor Evina Katsou and Professor Alireza Mousavi, for their kind guidance, generous offer of this opportunity, and studentship throughout my PhD journey. This support has been invaluable in allowing me to complete this challenging but rewarding academic pursuit.

I am deeply grateful to my entire supervisory team for their insightful guidance, enlightenment, assistance, and understanding throughout my research.

My appreciation extends to all those who offered explanations, engaged in discussions, shared helpful information, or provided any form of support.

My final thanks go to my daughter and family. Their unwavering love and support are always there.

ABSTRACT

Nitrous oxide (N_2O) emissions from wastewater treatment facilities pose a significant environmental challenge. This study proposes a novel data-driven modelling approach using emerging neural ordinary differential equations (NODE) to capture the complex dynamics of N_2O production in typical activated sludge processes.

The author established an experimental simulation platform, based on the BSM1 (benchmark simulation model no.1) plant, with the ASMG1 (activated sludge model for greenhouse gases no.1) mathematical model. This platform generates simulated monitoring data and validates the model. The author then proposes NODE-based models, analogous to traditional biokinetic models, capable of capturing the complex dynamics of N_2O generation through learning from process monitoring data. However, two primary challenges need to be overcome.

First, to address inherent stiffness in the underlying dynamics, the author proposes a **paired normalisation method** for training stability. Additionally, an **incremental training strategy** was introduced, starting from a **collocation method** to establish a robust foundation, followed by refinement using the **direct NODE method** for enhanced accuracy and efficiency.

Second, as monitoring data in wastewater plants typically contain confounding factors from continuous influent variations and operational adjustments, representing **exogenous excitations** to the dynamics to be captured, therefore the training procedures was extended to account for these external influences.

The approaches were validated on the established platform. The results demonstrate the effectiveness of the NODE-based model in capturing the intricate dynamics of N_2O production in wastewater treatment. This research presents a promising new avenue for data-driven modelling of N_2O in wastewater treatment, with the potential to improve process optimisation and emission control strategies.

Keywords: neural ordinary differential equations (NODE), N_2O emissions, wastewater process, data-driven modelling.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

A^2O	Anaerobic-Anoxic-Oxic
ACO	Ant Colony Optimization
AdaBoost	Adaptive Boosting
Adam	Adaptive Moment Estimation
AI	Artificial Intelligence
AMO	Ammonia Monooxygenase
anammox	Anaerobic Ammonia Oxidation
ANN	Artificial Neural Network
ANODE	Augmented Neural ODE
AO	Anoxic-Oxic
AOA	Ammonia Oxidation Archaea
AOB	Ammonium Oxidising Bacteria
ASM	Activated Sludge Model
ASM-EC	Activated Sludge Model with Electron Competition
ASM-ICE	Activated Sludge Model with Indirect Coupling of Electrons
ASMN	Activated Sludge Model for Nitrogen
BOD	Biological Oxygen Demand
BSM	Benchmark Simulation Model
CFD	Computational Fluid Dynamics
CNN	Convolutional Neural Network
COD	Chemical Oxygen Demand
comammox	Complete Ammonia Oxidation
CSTR	Continuous Stirred-tank Reactor
DAE	Differential Algebraic Equation
DGAO	Denitrifying Glycogen-accumulating Organisms
DL	Deep Learning
DNN	Deep Neural Network
DO	Dissolved Oxygen
DPAO	Denitrifying Polyphosphate-accumulating Organisms
DPR	Denitrifying Phosphorus Removal
EF	Emission Factor
EQI	Effluent Quality Index
EVS	Explained Variance Score

FA	Free Ammonia
FLC	Fuzzy Logic Controller
FNA	Free Nitrous Acid
GA	Genetic Algorithm
GAN	Generative Adversarial Network
GBM	Gradient Boosting Machine
GHG	Greenhouse Gas
GPU	Graphics Processing Unit
GSA	Global Sensitivity Analysis
HAO	Hydroxylamine Dehydrogenase
HD	Heterotrophic Denitrification
HDBSCAN	Hierarchical Density-Based Spatial Clustering of Applications with Noise
HN-AD	Heterotrophic Nitrification-Aerobic Denitrification
HRT	Hydraulic Retention Time
IID	Independent and Identically Distributed
IPCC	Intergovernmental Panel on Climate Change
IVP	Initial Value Problem
KN	Kjeldahl Nitrogen
KNN	K-nearest Neighbour
LSTM	Long Short-Term Memory
LULUCF	Land Use, Land-Use Change and Forestry
MADRL	Multi-agent Deep Reinforcement Learning
MAE	Mean Absolute Error
MBR	Membrane Bioreactor
ML	Machine Learning
MLP	Multi-layer Perceptron
MLSS	Mixed Liquor Suspended Solid
Mox	Mediator Oxidized
MPC	Model Predictive Control
Mred	Mediator Reduced
NAR	Nitrate Reductase
NCDE	Neural Controlled Differential Equation
ND	Nitrifier Denitrification

NDHA	Nitrifier Nitrification-Nitrifier Denitrification-Heterotrophic Denitrification-Abiotic
NIR	Nitrite Reductase
NN	Nitrifier Nitrification
NOB	Nitrite Oxidising Bacteria
NODE	Neural Ordinary Differential Equation
NOR	Nitric Oxide Reductase
NOS	Nitrous Oxide Reductase
NOx	Nitrogen Oxide
NPDE	Neural Partial Differential Equation
NSDE	Neural Stochastic Differential Equation
OCI	Operational Cost Index
ODE	Ordinary Differential Equations
OOD	Out-of-distribution
OUR	Oxygen Uptake Rate
PCA	Principal Component Analysis
PCR	Principal Component Regress
PHA	Polyhydroxyalkanoate
PI	Proportion-Integration
PINN	Physics-informed Neural Network
PSO	Particle Swarm Optimization
RAS	Recycle Activated Sludge
ResNet	Residual Neural Network
RL	Reinforcement Learning
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
SA	Simulated Annealing
SCADA	Supervisory Control And Data Acquisition
SCENA	Short Cut Enhanced Nutrient Abatement
SD	Standard Deviation
SGD	Stochastic Gradient Descent
SINDy	Sparse Identification of Nonlinear Dynamics
SRT	Solid Retention Time
SVM	Support Vector Machine
TN	Total Nitrogen

TPU	Tensor Processing Unit
WMO	World Meteorological Organization
WLC	Whole Life Cost
WWTP	Wastewater Treatment Plant

Chapter 1 Introduction

1.1 Greenhouse gas emissions in wastewater treatment

Despite net zero initiatives implemented in many countries, the global warming trend remains alarmingly accelerated. The World Meteorological Organization (WMO) reported that human-induced climate change reached new heights in 2024, with a global mean near-surface temperature of 1.55 ± 0.13 °C above the 1850-1900 average (WMO, 2025). This underscores the urgent need to fast-track commitments to slash greenhouse gas (GHG) emissions and achieve carbon neutrality.

Wastewater treatment, a critical process in modern urban infrastructure, is essential for protecting public health and the environment. These treatment processes involve a complex series of physical, chemical, and biological operations designed to remove contaminants from municipal and industrial wastewater before it is released back into the environment. However, they also contribute to GHG emissions, including direction emissions of carbon dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O) from the treatment processes, and indirect emissions from plant operation, primarily electricity consumptions.

CO_2 emissions stemming from oxidation of wastewater organic matters are excluded from GHG inventory by Intergovernmental Panel on Climate Change (IPCC), due to their biogenic origins and non-anthropogenic contribution (Bartram *et al.*, 2019).

CH_4 is mainly generated under anaerobic conditions where organic contaminants decompose. The pathway involves various types of anaerobic microorganisms through four conversion steps: hydrolysis, acidogenesis, acetogenesis, and methanogenesis, same as processes in anaerobic digester (Zhan, Hu and Wu, 2018).

N_2O is a potent greenhouse gas, with an estimated global warming potential 265 times greater than CO_2 over a 100-year period (Bartram *et al.*, 2019). Furthermore, it is a detrimental substance and the largest donator to ozone depletion. In wastewater treatment, N_2O is a by-product of biological nitrogen removal processes, typically from nitrification and denitrification, with more complicated mechanisms than CO_2 and CH_4 generation (Zhan, Hu and Wu, 2018). N_2O emissions exhibit significant spatiotemporal variations, with magnitudes ranging from negligible to 25% of total influent nitrogen (Ye, Porro and Nopens, 2022).

In wastewater treatment plants (WWTPs), N_2O emissions can account for up to 80% of overall carbon footprint (Daelman *et al.*, 2013), positioning WWTPs as the sixth largest contributor of global emissions (Tchobanoglous *et al.*, 2014). Figure 1.1 illustrates the percentages of the U.S. GHG emissions by gas and N_2O emissions by source based on the data from 1990 to 2022. Similarly, in the UK, the water sector accounts for nearly a third of the industrial and waste process GHG emissions (Water UK, Ricardo and Mott MacDonald, 2020).

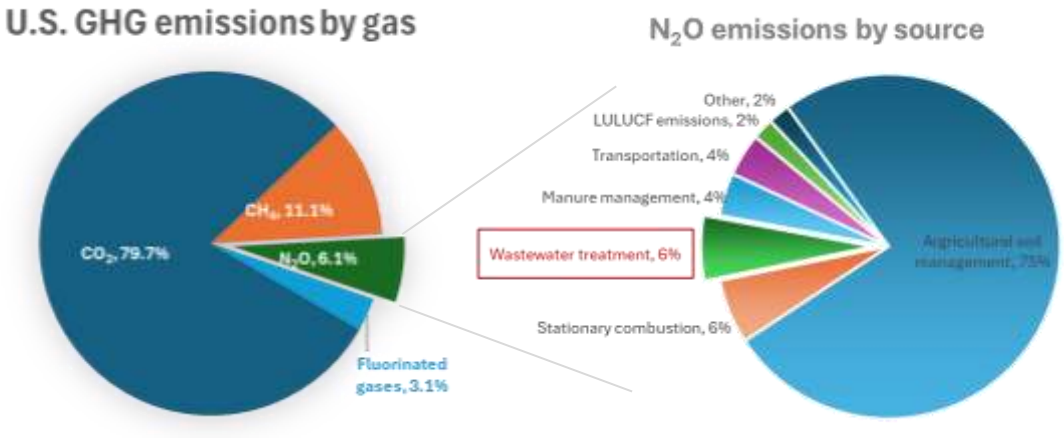


Figure 1.1 U.S. GHG emissions by gas (left) and nitrous oxide emissions by source (right), based on data of 1990-2022, adapted from (U.S. EPA, 1993)

As the development of green electricity advances, direct emissions of N_2O and CH_4, particularly N_2O, will become the dominant source of carbon emissions from WWTPs, presenting a significant environmental concern (Valkova *et al.*, 2020). Therefore, studying and modelling N_2O production is of great significance for minimizing the carbon emissions of WWTPs.

1.2 Conventional N_2O production modelling

Conventional models for wastewater processes, including those that account for N_2O generation, are typically biokinetics based (Ye, Porro and Nopens, 2022), such as well-established Activated Sludge Model Series (ASM1, ASM2, and ASM3) and their extensions. These models employ systems of ordinary differential equations (ODEs), often represented in the form of Gujer matrix by water professionals (Henze *et al.*, 2000). They describe the biochemical dynamics among biomass and substrates based on discovered mechanisms and pathways, utilising numerous kinetic coefficients and

stoichiometric parameters that require calibration and validation to fit the local conditions.

Research advancements have led to mechanistic models incorporating three major pathways of N_2O production: 1) hydroxylamine oxidation, 2) nitrifier denitrification, and 3) heterotrophic denitrification. When properly calibrated, these models can estimate site-specific emissions and inform mitigation strategies.

Despite their advantages, current mathematical models face challenges in fully elucidating the complex biological interactions and operational factors that govern N_2O emissions, due to key knowledge gaps that remain, including:

- The abiotic reaction pathway is still under debate (Stein, 2011; Su *et al.*, 2019),
- The role of archaea and certain microorganism species requires further clarification (Castellano-Hinojosa *et al.*, 2018),
- The extent of contributions from different production pathways under varying conditions is not fully understood.

N_2O production during nitrogen removal is a complex and transient process involving multiple interconnected pathways. These processes are influenced by several factors, including microbial community composition, substrate availability, dissolved oxygen levels, nitrite concentrations, pH, and temperature. However, these factors exhibit significant variability due to ever-changing influent characteristics and periodic operational adjustments, introducing uncertainty into the modelling process. The intricate interplay of these factors makes it challenging to develop comprehensive mathematical models that accurately capture the full range of N_2O dynamics and variations.

Furthermore, applying mathematical models necessitates validation and calibration of numerous parameters, a challenging and time-consuming process due to inherent non-linearity, processes interdependency, and measurement uncertainty (Belia *et al.*, 2009). Collecting reliable and accurate data, particularly from specifically designed lab tests and field measurements, can be difficult and often expensive. Additionally, these calibrations may not be easily transferable when scenarios change.

In summary, the practical application of the conventional mathematical modelling for N_2O prediction requires both accurate elaboration of the underlying mechanisms and

extensive experience and knowledge about process control. However, the increasing complexity of industrial processes often makes such prerequisites difficult to satisfy in real-world settings. Despite ongoing research to address these issues, the limitations discussed hinder the accuracy, reliability and adaptability of conventional mechanistic models, highlighting the need for innovative approaches or alternatives to modelling N_2O in wastewater treatment. Data-driven approaches, for instance, have the capability to overcome some of these limitations by learning complex relationships directly from process monitoring data (Haimi *et al.*, 2013; Pisa, Santín, *et al.*, 2019; Sun and Ge, 2021), potentially capturing dynamics that are difficult to describe using traditional mathematical modelling.

1.3 Data-driven modelling

Data-driven modelling offers a promising alternative for wastewater treatment. These approaches have gained significant traction in recent years due to advancements in data collection, storage capabilities, particularly computing power and sophisticated artificial intelligence (AI) algorithms.

A variety of data-driven techniques can be employed in wastewater treatment, ranging from soft sensor and fuzzy logic systems to earlier machine learning algorithms such as principal component analysis (PCA), random forests, genetic algorithms, support vector machine (SVM), multi-layer perceptron (MLP). More recently, applications of deep learning (DL) have proliferated, including convolutional neural network (CNN), recurrent neural network (RNN), generative adversarial network (GAN), transformers, reinforcement learning (RL), outperforming these earlier methods (Asadi and McPhedran, 2021; Ho *et al.*, 2021; Khalil *et al.*, 2023; Khalil *et al.*, 2024).

Modern WWTPs can collect a wealth of information as online sensors continuously monitor various parameters, including influent composition, dissolved oxygen levels, and effluent quality. Today, many WWTPs are equipped with not only online sensors at different locations of the process stages (see Figure 1.2), but also solenoid actuators which can be controlled remotely. SCADA system samples and collects the readings at a frequency like 5, 10 or 30 minutes, as well control signals are sent to executors from SCADA as necessary. The WWTP data are often formed in time series, reflecting yearly, seasonal, and diurnal cycles with some variations and turbulences brought by environmental and operational changes.

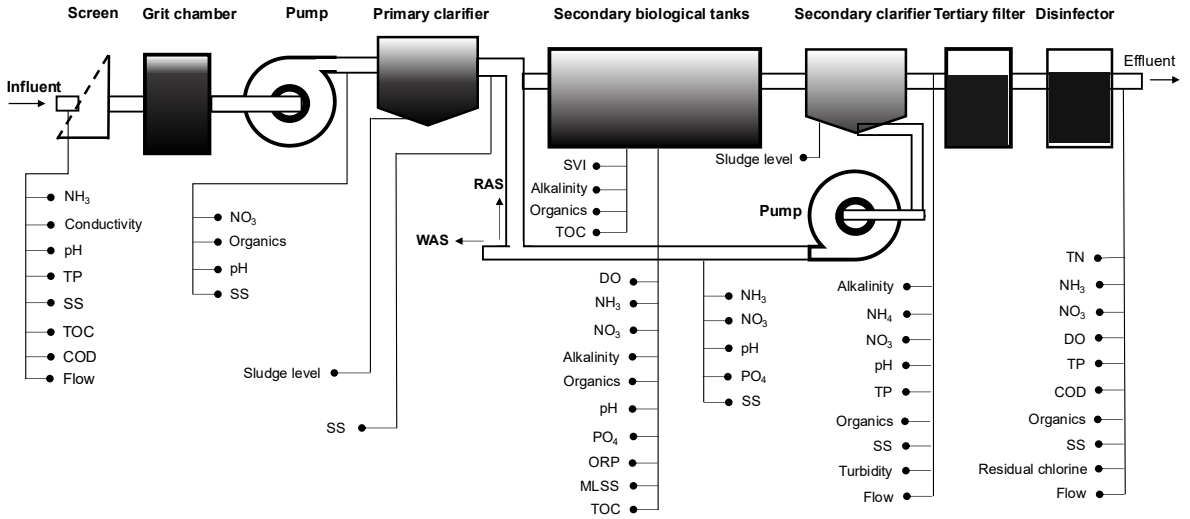


Figure 1.2 Illustration of typical WWTP instrumentation and monitoring parameters

The collected data serve as a rich resource for uncovering hidden patterns and relationships within the treatment process. Data-driven approaches leverage these large datasets, develop predictive models without relying solely on first-principles or mechanistic understanding. They offer several advantages over conventional mechanistic models (Khalil *et al.*, 2023):

- **Reduced reliance on mechanistic understanding**: Data-driven models do not require a complete understanding of the underlying biological and chemical pathways, which can be incomplete in complex systems.
- **Identification of complex relationships**: They can uncover intricate, non-linear relationships between process variables that may be difficult to capture with traditional models.
- **Real-Time Process Monitoring**: Data-driven models have the potential for real-time predictions and control, allowing for proactive adjustments to optimise treatment efficiency and minimize environmental impact.
- **Adaptability**: These models can adapt to changing operational conditions and influent characteristics by continuously learning from new data or through retraining.

While data-driven models hold significant promise, there are challenges to consider (Newhart *et al*., 2019):

- **Data Quality and coverage**: The success of data-driven models hinges on the quality and representativeness of the available data. Ensuring comprehensive and accurate data collection is crucial.
- **Model Interpretability**: Understanding the rationale behind the model's predictions can be difficult due to their black-box nature.

As sensor technology and data management practices continue to evolve, data-driven models are poised to play an increasingly important role in the future of wastewater treatment.

1.4 Neural ODE approach

Neural ordinary differential equations (NODEs) emerge as a powerful technique for modelling dynamic systems within the realm of data-driven approaches. This method leverages the strengths of both deep neural networks (DNNs) and traditional mechanistic models. DNNs provide exceptional expressiveness, allowing NODEs to capture complex non-linear system dynamics. Meanwhile, the mathematical foundation of differential equations within NODEs offers valuable insights into the underlying physical relationships that govern these dynamics. This marriage of DNNs and differential equations positions NODEs as a promising tool for modelling dynamic systems such as those encountered in wastewater treatment.

While neural networks are often criticized for their black-box nature, NODEs offer a level of interpretability through their connection to differential equations. The learned dynamics can be analysed in terms of rate of change and influences between variables, potentially revealing interpretable causal mechanisms within the system (Zou *et al.*, 2024)

Unlike traditional machine learning methods constrained by fixed time steps, NODEs excel at learning and representing complex temporal dynamics, regardless of irregular or variable time intervals frequently encountered in actual data (Kidger *et al*., 2020). The continuous nature allows NODEs to provide solutions at any arbitrary time point. This capability makes them well-suited for modelling wastewater systems with such real-world complexities, allowing them to effectively adapt to the inherent variability and non-linearity present in wastewater treatment processes.

While mechanistic modelling heavily relies on clear understanding of the underlying dynamics and explicit mathematical expressions, conventional machine learning methods like MLPs and RNNs act as black boxes, merely mapping the input and output from large datasets without offering logical or causal interpretations. NODEs bridge this gap by learning and extracting physical laws directly from monitored data, creating continuous time series models that handle data with irregular intervals. Moreover, NODEs often require less data compared to traditional methods and can incorporate prior knowledge, further enhancing the interpretability of their outcomes. Figure 1.3 compares the paradigms of these different modelling approaches.

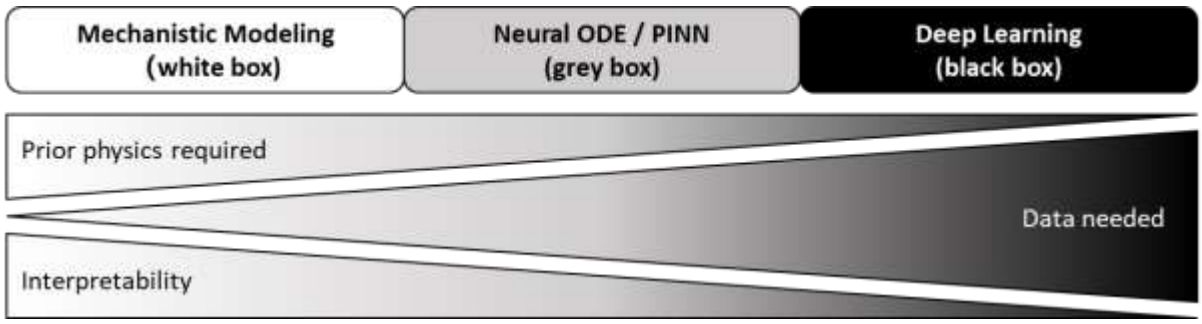


Figure 1.3 paradigms of different modelling approaches

This study explores the capabilities of NODEs for data-driven modelling of wastewater processes, particularly focusing on N_2O production. Given the relative novelty of NODEs, a dedicated chapter later will introduce their fundamental concepts.

1.5 Research objectives

The overarching aim of this research is to develop and implement NODE-based models for data-driven modelling of wastewater treatment processes, with a specific focus on nitrous oxide (N_2O) production dynamics. This aim encompasses investigating and understanding the underlying mechanisms of N_2O emissions and the interactive effects of operational interventions, as well as addressing challenges related to system stiffness, core algorithms, and training stability and efficiency.

To achieve this aim, the following objectives have been established:

- **Objective 1**: Elucidate the mechanisms and pathways of N_2O emissions in wastewater treatment processes and identify the factors influencing N_2O production and emissions. Critically review the progress and limitations of current

mathematical and data-driven modelling approaches for N_2O production in wastewater treatment.

- **Objective 2**: Develop solutions to address the issues of system stiffness, training instability and divergence in NODE models. Ensure that the learning process can be conducted smoothly and successfully.
- **Objective 3**: Develop and train NODE models using the proposed methods and coded algorithms, ensuring they can produce reliable and accurate results.
- **Objective 4**: Validate and test the developed models under various scenarios, evaluate and improve their performance, analyse their limitations, propose improvement methods, and identify future research directions.

Through the accomplishment of these objectives, this research is expected to contribute to the development of a practical and effective solution for modelling N_2O production in wastewater treatment. This will ultimately aid in minimizing GHG emissions and pave the way for the creation of low carbon "smart plants" in water industry.

1.6 Research methodology overview

This research aims to develop and validate a data-driven NODE model for predicting N_2O production in WWTPs. The methodology undertaken in the research encompasses several key phases, each designed to address specific objectives and contribute to the overall goal of improving N_2O modelling in wastewater treatment.

1) Critical literature review

The foundation of this research is built upon an extensive and critical literature review, which serves multiple purposes:

- Understanding N_2O production mechanisms: A thorough exploration of existing literature will be conducted to understand the complex mechanisms and pathways of N_2O production, including the influence of operational factors.
- Mathematical and data-driven modelling assessment: The current state-of-the-art in both mathematical and data-driven modelling approaches for N_2O emissions will be evaluated, identifying their strengths, limitations, and potential for improvement.

- NODE theory and applications: A detailed analysis of NODE theory, its advancements, and existing applications will be undertaken. Given the relative novelty of NODE, a particular focus will be placed on identifying potential gaps and opportunities for contribution.
- Continuous Literature Tracking: To ensure the research remains aligned with the latest developments, a systematic approach to tracking new machine learning algorithms and their potential application to N_2O modelling will be implemented.

2) Experimental simulation platform

Following the literature review, the research will focus on establishing a robust experimental simulation platform.

The decision to conduct research on a simulation platform, rather than using real-world data, not only because lack of real-world data and research time is limited, but also for the reason that NODE is a new approach, necessitating to start from simulation firstly and prove its feasibility before applying to real-world cases. In fact, the use of a simulation platform, rather than real-world data, offers several advantages:

- Provides controlled conditions for testing and validation.
- Allows for the generation of large, diverse datasets that might be impractical to obtain from real WWTPs.
- Enables the exploration of extreme or rare scenarios that are critical for comprehensive model training.
- Facilitates easy comparison with mathematical results to verify performance.

Key aspects of the simulation platform include:

- Simulation platform development: An experimental simulation platform will be constructed to replicate the complex dynamics of WWTPs. This platform will enable controlled experimentation and comparison of model performance against simulated and potentially real-world data.
- Data generation: The simulation platform will generate synthetic datasets that accurately represent the relevant processes and parameters influencing N_2O production. This will provide a robust foundation for model development and testing.

- Validation framework: The developed platform shall be validated by reference database system or assessed by wastewater theory and practice. The platform also provides comparison of simulation results with established mathematical models to ensure accuracy, and metrics for assessing the fidelity of the simulation to real-world scenarios.
- Addressing stiffness issue: It is crucial to overcome the obstacles associated with training NODEs such as stiffness. This is the most challenging and time-consuming phase, requiring significant coding effort and trial-and-error testing and improvement based on the experience built upon failures. The solution will be embedded in the platform to facilitate future experiments.

3) Model development and training

With the stiffness issues addressed, the research will proceed to construct and implement NODE models for N_2O production prediction:

- NODE model construction: Multiple NODE models will be developed, incorporating different training datasets and architectural variations to explore the optimal model configuration.
- Model optimization: The developed NODE models will undergo extensive training and tuning to achieve acceptable levels of accuracy in predicting N_2O production. This process will involve careful selection of hyperparameters, feature engineering, and evaluation metrics.

4) Model Validation and Evaluation

- Scenario Testing: The trained NODE models will be rigorously tested under various simulated scenarios to assess their predictive performance and generalizability.
- Strengths and Weaknesses: A comprehensive analysis of the models' strengths, weaknesses, and limitations will be conducted to identify areas for improvement and potential applications.
- Future Directions: Based on the evaluation results, recommendations for future research and development of NODE-based N_2O modelling will be provided.

The methodological framework and hypothesised workflow for this research are summarised in Figure 1.4.

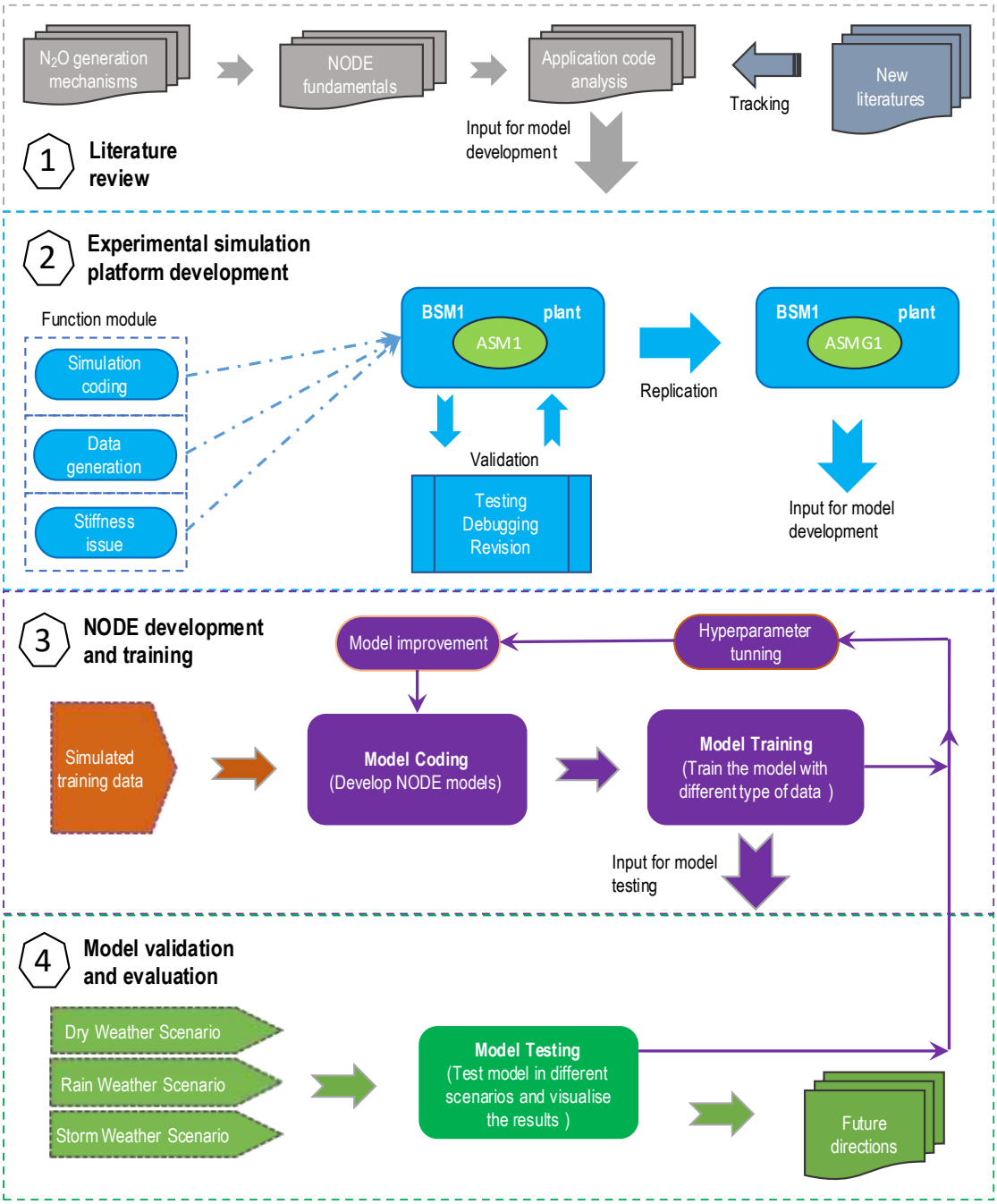


Figure 1.4 Overview of research methodology framework and workflow

1.7 Thesis outline

The thesis comprises seven main chapters and an extensive appendix.

Thesis structure outlining the objectives addressed is summarised in Table 1.1.

Table 1.1 Thesis structure outlining the objectives addressed

Chapter	Paper	Title	Objective met
1		Introduction	
2		Literature review on N_2O production modelling	1
3		NODE fundamentals	1
4		Experimental simulation platform	2
5	Yes	Tackling stiffness for NODE models	2
6	Yes	Implementation of NODE modelling of N_2O in BSM1 plant	3
7		Discussion and conclusion	4
8		Appendix	3
9		Reference	

Each chapter is briefed as follows.

Chapter 1: Introduction

- Highlight the concern over N_2O emissions and their contribution to GHG production in wastewater treatment.
- Brief conventional N_2O modelling methods, introducing data-driven NODE approach.
- State the research objective, methodology and thesis outline.

Chapter 2: Literature review on N_2O production modelling

- Discuss mechanisms and pathways of N_2O emissions in wastewater treatment.
- Review existing knowledge of factors influencing N_2O production and emission.
- Analyse current methods for N_2O production modelling by both mathematical and data-driven approaches.
- Critically evaluate the limitations of existing methods and justify the use of NODE models.

Chapter 3: NODE fundamentals

- Explain the core concepts of NODE models.
- Brief the detailed techniques and training steps of NODE models for dynamic system modelling.

Chapter 4: Experimental simulation platform

- Describe the development or selection of the simulation platform and mathematic model used for your experiments.
- Explain how the platform is configured, and how to organise, process influent data and generate training data including N_2O.
- Validate the data generation process by comparing reference results and wastewater theory.

Chapter 5: Tackling stiffness issue for NODE training

- Define the concept of system stiffness and its challenges in training NODE models.
- Explain the specific methods or techniques proposed and implemented to address stiffness issue.
- Validate the effectiveness of proposed approach and its impact on training stability.

Chapter 6: Implementation of NODE model in BSM1 plant

- Explain how the NODE model is constructed and trained in the settings of BSM1 plant.
- Present the results of model training with different data.
- Evaluate the model's performance and analyse its behaviours in different scenarios.

Chapter 7: Discussion and Conclusion

- Discuss the overall performance of the NODE model in simulating N_2O production.
- Analyse the limitations of the model and potential areas for improvement.
- Summarize the key findings of the research and its contribution to N_2O modelling in wastewater treatment.
- Conclude by suggesting potential future research directions based on your findings.

The appendix contains detailed training loss and logs, extended results and analyses, comprehensive plots and visualisations, additional information and amendment. It complements the main body of the thesis but doesn't directly affect the core arguments or flow of the discussion.

Due to their length, code scripts are not included in the thesis itself but are provided separately in electronic form for accessibility.

1.8 Significance of the research

This research has the potential to significantly advance data-driven modelling of N_2O production in wastewater treatment facilities. Some of its key impacts include:

- **Improved accuracy and generalizability**: By leveraging process monitoring data and the flexibility of NODEs, the model could achieve more accurate predictions of N_2O productions compared to traditional methods. Additionally, the model's ODE-based nature may allow for better generalizability to different plants and operating conditions.
- **Enhanced understanding of N_2O production**: NODEs can capture the complex relationships between various process parameters and N_2O emissions. This deeper understanding will provide valuable insights into the key factors influencing N_2O production, allowing for targeted mitigation strategies.
- **Practical process optimization**: The model's ability can be used to optimize wastewater treatment processes. This optimization can minimize N_2O emissions while maintaining treatment efficiency, leading to more sustainable and environmentally friendly practices.
- **Paving the way for broader applications**: The success of this data-driven approach using NODEs could pave the way for its application to other environmental modelling challenges within wastewater treatment, or beyond the water field, such as chemical engineering, biomedical processes, environmental science, etc.

Chapter 2 Literature review: Modelling of N_2O production in wastewater treatment processes

Nitrous oxide (N_2O) is a byproduct of biological nitrogen removal processes in wastewater treatment. The transformation of nitrogen within these processes is highly complex, extending far beyond the conventional paradigms of nitrification and denitrification. This complexity is largely due to the intricate microbial ecology involved in these processes (Khalil *et al.*, 2024).

While nitrification and denitrification are the two primary routes traditionally recognized in nitrogen removal, recent research has revealed the existence of multiple "side routes" or ancillary bioreactions occurring simultaneously (Ye, Porro and Nopens, 2022). These additional pathways contribute to the overall nitrogen transformation process and N_2O production, making the system more complex than previously understood.

Identifying these complex microbial interactions and the various pathways of nitrogen transformation is crucial for developing effective strategies to mitigate N_2O emissions from wastewater treatment plants.

2.1 Mechanism and pathway

The main methods for identifying N_2O sources include isotope technology, inhibitor methods, and enzyme assays.

- **Isotope technology** involves either adding isotopes or measuring the natural abundance of stable nitrogen isotopes (^{15}N) in wastewater to trace N_2O origins (Wunderlin *et al.*, 2013; Gruber *et al.*, 2022). While the potential of isotope technology is widely recognized, its application in quantifying N_2O generation pathways within wastewater treatment systems requires further refinement, particularly in terms of accuracy and reliability (Duan *et al.*, 2017).
- **Inhibitor methods** involve adding specific denitrification inhibitors to identify N_2O sources (Yang *et al.*, 2022). By selectively inhibiting specific microbial processes, researchers can determine the relative contributions of different pathways to N_2O production.

- **enzyme assays** analyse the activity of denitrifying enzymes to determine N_2O origins (Yang *et al.*, 2022). This technique provides insights into the microbial processes responsible for N_2O generation.

Despite extensive research in the last decades, the mechanisms underlying N_2O generation in wastewater treatment are still not fully understood, with multiple generation pathways being interrelated and often context dependent. There remains ongoing debate regarding the dominant N_2O sources. Additionally, the relative contribution of different N_2O generation pathways can vary significantly under different processes and operating parameters (Ye, Porro and Nopens, 2022). As shown in Figure 2.1, to date, the widely accepted four major pathways are (Chen *et al.*, 2020):

i) hydroxylamine (NH_2OH) oxidation or nitrifier nitrification (NN) pathway,
ii) nitrifier denitrification (ND) pathway,
iii) heterotrophic denitrification (HD) pathway, and
iv) abiotic pathway.

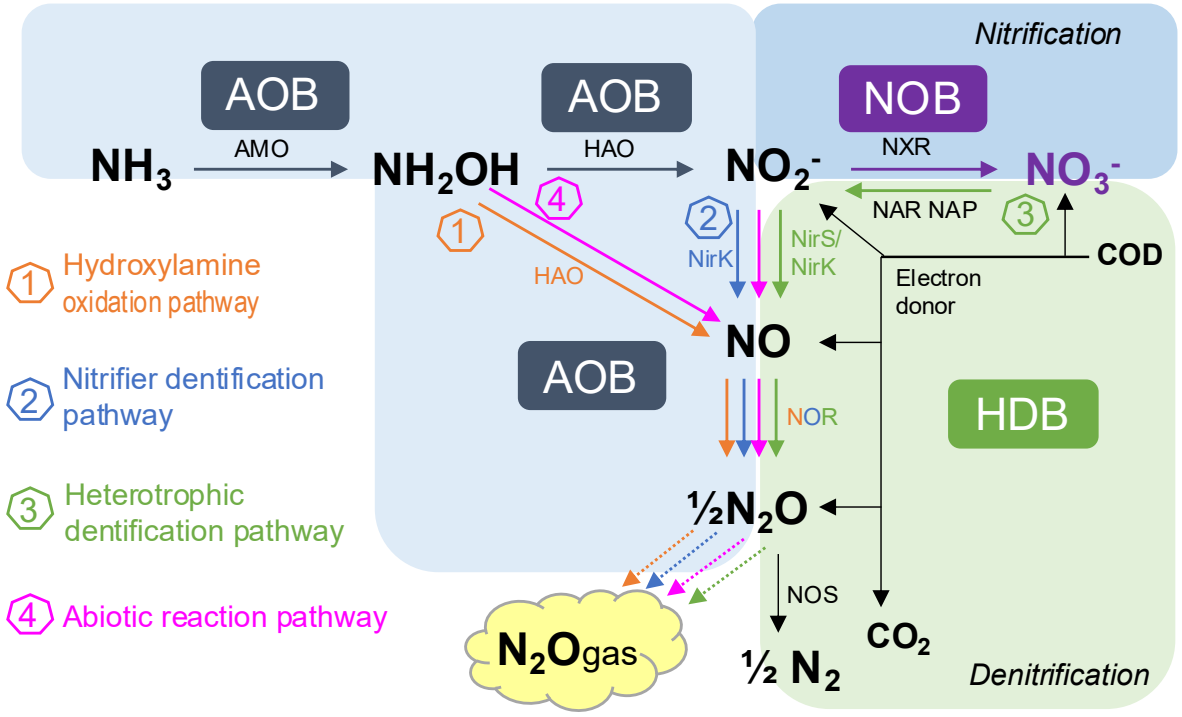


Figure 2.1 possible pathways of N_2O emissions from biological wastewater treatment process

During nitrification, ammonia (NH_3) is oxidized to nitrate (NO_3^-) in two sequential steps: first to nitrite (NO_2^-) via the intermediate hydroxylamine, primarily by ammonia-

oxidizing bacteria (AOB), and then to NO_3^- mainly by nitrite-oxidizing bacteria (NOB). Other microorganisms capable of participating in the nitrification process include ammonia-oxidizing archaea (AOA), heterotrophic nitrification-aerobic denitrification (HN-AD) bacteria, and complete ammonia oxidation (comammox) microorganisms (YU *et al.*, 2019). During denitrification, NO_3^- is converted to nitrogen gas (N_2) through intermediates NO_2^-, NO, and N_2O. Microorganisms involved include heterotrophic denitrifying bacteria (HDB), AOB, anaerobic ammonium oxidation (anammox) bacteria, and HN-AD bacteria (Hao *et al.*, 2023). However, actual reactions might be incomplete, and numerous side reactions can occur, reducing nitrogen removal efficiency and generating N_2O emissions.

2.1.1 Hydroxylamine oxidation pathway

Hydroxylamine is an obligatory intermediate in the initial step of nitrification, where ammonia is converted to nitrite (Tchobanoglous *et al.*, 2014). The production of N_2O via NH_2OH oxidation pathway primarily stems from transient NH_2OH accumulation caused by imbalances in the ammonia oxidation enzyme turnover (Cantera and Stein, 2007). During this process, AOB derive energy by oxidizing NH_3 to NO_2^- through a two-step reaction: first, NH_3 is oxidized to NH_2OH by the enzyme ammonia monooxygenase (AMO), followed by the oxidation of NH_2OH to NO_2^- by the enzyme hydroxylamine oxidoreductase (HAO) (Domingo-Félez and Smets, 2019).

The electron flux generated from NH_2OH promotes NO_2^- reduction. Nevertheless, due to metabolic imbalances, a portion of the accumulated hydroxylamine can undergo further oxidation to nitric oxide (NO), which is subsequently reduced to N_2O by NorS, a homologue of nitric oxide reductases (Stein *et al.*, 2007). This mechanism is widely accepted for the hydroxylamine oxidation pathway of N_2O production. An alternative hypothesis proposes that conversion of NH_2OH by HAO to a nitrosyl radical (NOH), which subsequently decomposes chemically to form N_2O (Poughon, Dussap and Gros, 2001).

Traditionally, NH_2OH was considered the sole intermediate. However, recent findings suggest that NO may also play a role (Caranto and Lancaster, 2017). Notably, while the aforementioned reactions predominantly occur under aerobic conditions during nitrification, NH_2OH oxidation can also take place under anaerobic conditions. The

direct conversion of NH_2OH to N_2O by cytochrome P460 under anaerobic conditions has been reported (Caranto, Vilbert and Lancaster, 2016).

Studies have demonstrated that the NH_2OH oxidation pathway represents a substantial N_2O source in diverse wastewater treatment systems (Law *et al.*, 2012; White and Lehnert, 2016; Wrage-Mönnig *et al.*, 2018), including conventional activated sludge (CAS) processes (Tumendelger, Alshboul and Lorke, 2019), the oxidation tank of anoxic/aerobic processes (Guo *et al.*, 2021), granular partial nitrification-anaerobic ammonia oxidation processes (Liu *et al.*, 2020), and membrane aerated biofilm reactors (Liu *et al.*, 2022). However, the relative contribution of this pathway to overall N_2O emissions can vary significantly depending on operational factors, such as dissolved oxygen concentration, ammonium loading, and pH (Law *et al.*, 2012; Domingo-Félez and Smets, 2019). Further research is essential to elucidate the precise mechanisms governing N_2O formation through this pathway under different operating conditions.

2.1.2 Nitrifier denitrification pathway

Nitrifier denitrification, primarily carried out by ammonia-oxidizing bacteria (AOB), involves the reduction of NO_2^- to N_2O via NO without requiring organic carbon. This pathway is particularly prevalent under low DO conditions, where nitrite substitutes oxygen as the electron acceptor (Zhu *et al.*, 2013). Their biochemical reactions can occur through two primary routes:

- Direct conversion of NO_2^- to N_2O: catalysed by isomeric nitrite reductase (Nir) (Casciotti and Ward, 2001).
- A two-step process: where NO_2^- is first converted to NO by nitrite reductase, followed by the reduction of NO to N_2O by NO reductase (Nor) and cytochrome c554 (Beaumont *et al.*, 2004; Kozlowski, Price and Stein, 2014).

A key distinction of AOB is their lack of N_2O reductase (NOS) in their genomes, preventing the further reduction of N_2O to N_2 (Kozlowski, Kits and Stein, 2016). Consequently, N_2O is the terminal product of this pathway. Nitrifier denitrification occurs concurrently with ammonia oxidation under aerobic conditions and is amplified in microaerobic environments.

The precise role of nitrifier denitrification in AOB remains unresolved. Several hypotheses have been proposed:

- Energy Conservation: Under oxygen-limited conditions, AOB might utilize nitrifier denitrification to generate some energy for survival (Kozlowski, Price and Stein, 2014).
- Electron Dissipation: High NH_4^+ influxes can lead to an excess of electrons. Nitrifier denitrification could act as a mechanism to dissipate these electrons (Domingo-Félez and Smets, 2019).
- Competition Control: By reducing NO_2^-, AOB can gain a competitive advantage by limiting the substrate for NOB (Poth and Focht, 1985).
- Detoxification: Nitrifier denitrification might serve as a detoxification mechanism for AOB to remove excess NO_2^-, which can be toxic at high concentrations (Wrage-Mönnig *et al.*, 2018).
- Intermediate oxidation byproduct: Some studies suggest N_2O production could be an unintended consequence of AOB using NO as an electron sink to accelerate NH_2OH oxidation during aerobic metabolism (Yu *et al.*, 2018).

The relative dominance of nitrifier denitrification versus other N_2O production pathways (e.g., NH_2OH oxidation) is influenced by the concentrations of various nitrogen species in the wastewater. Studies employing nitrogen isotope fractionation analysis have indicated that nitrifier denitrification can be a dominant pathway for N_2O production in systems treating domestic wastewater (Wunderlin *et al.*, 2013). However, under conditions of high ammonia and low nitrite concentrations, the hydroxylamine oxidation pathway may become increasingly relevant (Kozlowski, Kits and Stein, 2016).

Nitrifier denitrification has been identified as a significant source of N_2O emissions in various wastewater treatment processes (Kim *et al.*, 2010; Zhu *et al.*, 2013), including sequencing batch reactors (SBRs) (Li *et al.*, 2019; Liu *et al.*, 2021), oxidation ditches (Zhou *et al.*, 2019), partial nitrification-anammox (Wan and Volcke, 2022), and the anoxic phase of anaerobic/aerobic processes (Guo *et al.*, 2021). Recent research suggests that different operational conditions may favour the growth of specific AOB strains with varying N_2O production pathways (Stein, 2011). For instance, Law et al. (2012) found that in an enriched AOB culture adapted to high levels of NH_4^+ and NO_2^- (~500 mg N/L) and low DO concentrations (0.5-0.8 mg O_2/L), the majority of N_2O

production was attributed to the chemical breakdown of the nitrosyl radical formed during NH_2OH oxidation to NO_2^-.

While the understanding of nitrifier denitrification has improved, further research is imperative to elucidate the precise mechanisms controlling this pathway's activation and its contribution to N_2O production under varying operational scenarios. Additionally, exploring the impact of AOB strain diversity on N_2O production pathways could provide valuable insights for developing more effective mitigation strategies in wastewater treatment systems.

2.1.3 Heterotrophic denitrification pathway

Heterotrophic denitrification is carried out by a diverse group of facultative heterotrophic microorganisms capable of using O_2, NO_2^-, or NO_3^- as electron acceptors to degrade carbon sources and generate energy for growth (Mills, 2019). This pathway involves the sequential reduction of NO_3^- to N_2 through a series of enzymatic reactions, with NO_2^-, NO, and N_2O as intermediate products. The enzymes responsible for these transformations are nitrate reductase (NaR), nitrite reductase (NiR), nitric oxide reductase (NOR), and nitrous oxide reductase (NOS), respectively (Hochstein and Tomlinson, 1988). Different redox active metals catalyse specific enzyme for their reactions, for instance, molybdenum for NO_3^- reduction, iron or copper for NO_2^- reduction, iron for NO reduction, and copper for N_2O reduction (Richardson *et al.*, 2009).

While the complete denitrification pathway culminates in the production of harmless N_2, incomplete denitrification can lead to N_2O accumulation. Several factors contribute to this:

- Specific microbial group: Not all denitrifying bacteria possess all the necessary enzymes for complete denitrification. Some lack NOS, resulting in N_2O as the end product (Hallin *et al.*, 2018; Gao *et al.*, 2019).
- Inhibitive environmental conditions: Factors such as low dissolved oxygen, high organic loading, and pH can inhibit NOS activity, leading to N_2O accumulation (Pan *et al.*, 2012).

- Unbalanced enzyme kinetics: The relative rates of the different enzymatic steps can influence the accumulation of N_2O. For instance, if NOS activity is slower than NOR activity, N_2O can accumulate (Pan, Ni and Yuan, 2013).

Notably, heterotrophic denitrification holds a unique position in wastewater treatment processes as the only known biological sink for N_2O (Chen *et al.*, 2020). This characteristic presents a promising opportunity for developing methods to reduce N_2O emissions by leveraging this natural sink (Zhou *et al.*, 2022). Recent research has focused on identifying and harnessing new bacterial strains with enhanced N_2O reduction potential, such as heterotrophic aerobic denitrifying bacteria (Rajta *et al.*, 2020). These novel strains offer innovative approaches for mitigating N_2O emissions in wastewater treatment systems.

While heterotrophic denitrification serves as a potential sink, studies have also demonstrated that this pathway can be a significant source of N_2O production in certain processes, such as the anoxic zone of aerobic plus anaerobic ammonia oxidation biofilters and nitrifying biofilters (Humbert *et al.*, 2020; J. Li *et al.*, 2022).

2.1.4 Abiotic pathway

In addition to biological processes, abiotic reactions contribute to N_2O production in wastewater treatment, often interacting with biotic mechanisms (Soler-Jofra *et al.*, 2016).

The most prominent abiotic pathway involves the reaction between NH_2OH and nitrous acid (HNO_2) (Falcone, Shug and Nicholas, 1963; Anderson, 1964). This process, along with other generated intermediates, such as HNO and ONNO, through spontaneous reactions or under catalysis of substances like manganese oxide and ferrous ions, can lead to N_2O generation (Yamazaki *et al.*, 2014). Key reactions include:

$$\mathbf{NH_2OH + HNO_2 \rightarrow N_2O + 2H_2O} \qquad \textbf{Equation 2-1}$$

$$\mathbf{2NH_2OH + O_2 \rightarrow N_2O + 3H_2O} \qquad \textbf{Equation 2-2}$$

$$\mathbf{4NH_2OH \rightarrow 2NH_3N_2O + 3H_2O} \qquad \textbf{Equation 2-3}$$

$$\mathbf{2NH_2OH + 2MnO_2 + 4H^+ \rightarrow N_2O + 2Mn^{2+} + 5H_2O} \qquad \textbf{Equation 2-4}$$

Other notable pathways include the reduction of HNO_2 by Fe^{2+} and the oxidation of NH_2OH by Fe^{3+} (Terada *et al.*, 2017). These reactions are significantly influenced by pH and substrate concentration, with N_2O production increasing under acidic conditions and high nitrite concentrations (Zhu-Barker et al., 2015; Su et al., 2019).

$$\mathbf{2HNO_2 + 4Fe^{2+} + 4H^+ \rightarrow 4Fe^{3+} + N_2O + 3H_2O} \qquad \textbf{Equation 2-5}$$

$$\mathbf{4Fe^{3+} + 2NH_2OH \rightarrow 4Fe^{2+} + N_2O + H_2O + 4H^+} \qquad \textbf{Equation 2-6}$$

Intermediates such as nitroxyl (HNO) and hyponitrous acid ($H_2N_2O_2$) can also contribute to N_2O generation through various reactions (Yamazaki *et al.*, 2014):

$$\mathbf{NH_3OH^+ + HNO \rightarrow N_2O + H_2O + H_3O^+} \qquad \textbf{Equation 2-7}$$

$$\mathbf{2HNO \rightarrow N_2O + H_2O} \qquad \textbf{Equation 2-8}$$

$$\mathbf{2H_2N_2O_2 \rightarrow N_2O + H_2O} \qquad \textbf{Equation 2-9}$$

The abiotic pathway, though typically accounting for a smaller portion of total N_2O production, can become substantial under certain conditions, particularly in the presence of heavy metals (Zhu-Barker *et al.*, 2015). Under these conditions, the contribution of abiotic reactions to overall N_2O emissions can increase significantly (Harper *et al.*, 2015).

2.2 Factors influencing N_2O production

Researchers have conducted extensive laboratory testing (Chen *et al.*, 2020; Guo *et al.*, 2021; Lee, Lin and Lei, 2022) and full-scale measurements in WWTPs (Foley *et al.*, 2010; Pan *et al.*, 2016; Song *et al.*, 2020; Khalil *et al.*, 2024) to elucidate the relationships between operational parameters and N_2O production. Various analytical techniques, including "black box" approaches (Song *et al.*, 2020), sensitivity analyses (Lancioni *et al.*, 2024), and principal component analyses (Bellandi *et al.*, 2020), have been employed in these investigations. The studies have identified numerous parameters influencing N_2O production, such as nitrogen load (Song *et al.*, 2020), DO level (Aboobakar *et al.*, 2013; Zhu *et al.*, 2013), carbon-to-nitrogen (C/N) ratio (Yan *et al.*, 2021), nitrite concentration (Cantera and Stein, 2007; Terada *et al.*, 2017),

circulation ratio (Zhang *et al.*, 2024), solids retention time (SRT) (Zhou *et al.*, 2019), pH (Pan *et al.*, 2012), temperature (Li *et al.*, 2019), and salinity (Zhao *et al.*, 2014). Among these, DO levels, carbon source availability, and microbial community composition have been recognized as primary determinants of N_2O production in wastewater treatment processes (Mannina *et al.*, 2018; Lee, Lin and Lei, 2022; Hao *et al.*, 2023).

It is important to note that these factors are intricately interconnected. For example, DO concentrations and the availability of carbon and nitrogen sources affect NO_2^- accumulation, while recycle ratio and feeding regimes influence carbon and nitrogen availability. Collectively, these factors impact microbial community distribution, enzyme activity, and ultimately influencing N_2O generation (Chen *et al.*, 2019; Duan *et al.*, 2020; Ye, Porro and Nopens, 2022; Hao *et al.*, 2023; Khalil *et al.*, 2024). This complex interplay of variables underscores the multifaceted nature of N_2O production in wastewater treatment processes and highlights the need for a comprehensive approach to mitigate its emissions. This section summarizes the impacts on N_2O generation from three categories: wastewater characteristics, process parameters, and microbial populations.

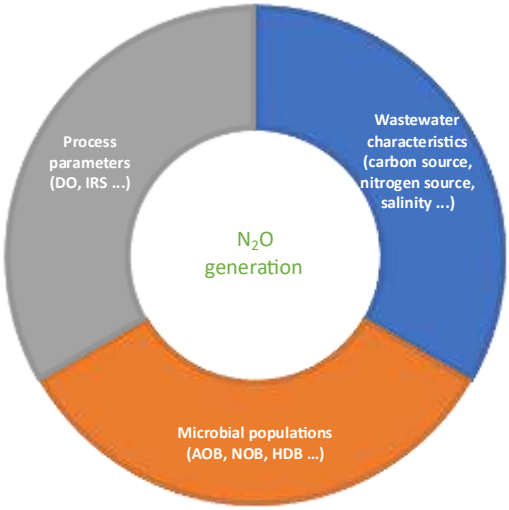


Figure 2.2 Categories of factors influencing N_2O generation in wastewater treatment

2.2.1 Wastewater characteristics

Wastewater characteristics significantly influence N_2O generation and emission. Despite identical treatment processes, substantial variations in N_2O emissions can occur between different wastewaters (Song *et al.*, 2020). Key characteristics include carbon and nitrogen sources, temperature, pH and so on. As essential nutrients for microbial growth, carbon and nitrogen sources impact microbial communities and enzyme activity within the process. Temperature, pH, and other wastewater components affect microbial activity and chemical reactions, indirectly influencing N_2O production.

2.2.1.1 Carbon source

The carbon source in wastewater significantly influences N_2O production by modulating denitrifying microbial growth and enzyme synthesis. Insufficient carbon supply can lead to elevated N_2O emissions due to incomplete denitrification, while excessive organic loading can stimulate AOB activity, thereby increasing N_2O production. The specific carbon source also affects denitrification efficiency and N_2O emissions (Chen *et al.*, 2019).

Studies on anaerobic-anoxic-oxic (A^2O) (Yan *et al.*, 2021), membrane bioreactor (MBR) (Mannina et al., 2018), and SBR (Zhao *et al.*, 2022) processes have identified dissolved inorganic carbon concentration and specific ammonia oxidation activity are two parameters most closely related to N_2O emissions (Song *et al.*, 2020). A strong positive correlation exists between these parameters. High dissolved inorganic carbon levels enhance AOB activity, consequently increasing N_2O emissions. Similarly, elevated organic loads are associated with higher N_2O emissions (Thwaites *et al.*, 2021). Within anaerobic zones, low carbon-to-nitrogen ratios can decrease the heterotrophic denitrification rate, promoting nitrate accumulation and the activity of denitrifying phosphorus-accumulating organisms, thus promoting N_2O production during phosphorus removal (Mannina et al., 2018).

Acetic acid as a carbon source minimizes N_2O emissions while maximizing total nitrogen removal. Conversely, mixed acids promote AOB growth, leading to increased N_2O production (Yan *et al.*, 2021). Although propionic acid results in lower N_2O

emissions compared to acetic acid, denitrification efficiency is reduced by 40% (Pan *et al.*, 2013; Li, Wang and Jia, 2022).

2.2.1.2 Nitrogen source

The concentration of various nitrogenous compounds in wastewater significantly influences N_2O production by affecting multiple N_2O generation pathways and altering microbial populations. The concentrations of nitrogen sources, particularly NH_4^+ and NO_2^-, strongly impact N_2O production (Law *et al.*, 2012; Terada *et al.*, 2017). As substrates for the hydroxylamine oxidation pathway and the nitrifier denitrification pathway, increased concentrations of NH_4^+ and NO_2^- substantially promote N_2O production. These compounds also directly facilitate the abiotic production of N_2O (Harper *et al.*, 2015).

The nitrogen loading of the influent affects the relative abundance of microbial communities, thereby influencing N_2O production. In low-nitrogen wastewater treatment, N_2O-reducing denitrifiers dominate the denitrifying community. Conversely, as influent nitrogen strength increases, N_2O-producing denitrifiers gradually become dominant. Furthermore, the shift from AOA to AOB as the primary ammonia oxidizers with increasing influent strength also promotes N_2O production (Sun *et al.*, 2022).

2.2.1.3 Salinity

Wastewater salinity can induce changes in microbial metabolic enzymes and cell structure, significantly impacting denitrification efficiency and GHG emissions. The relationship between salinity and N_2O production follows a non-linear trend, initially increasing and then decreasing as salinity rises. This trend is significantly negatively correlated with the nitrogen removal rate (Shao *et al.*, 2020).

In salinity shock experiments, increased salinity inhibits the activity of AMO and HAO, thereby slowing the ammonia oxidation rate and NO_2^- formation. Simultaneously, nitrous oxide reductase is inhibited, leading to an increase in N_2O production via the hydroxylamine oxidation pathway and the heterotrophic denitrification pathway (P. Li *et al.*, 2023).

Under long-term salinity acclimation, the abundance of NOB decreases while that of AOB increases. This shift in microbial community composition results in increased N_2O

production through the nitrifier denitrification pathway and decreased N_2O production via the heterotrophic denitrification pathway (Zhao *et al.*, 2014).

2.2.1.4 Others

In addition to carbon and nitrogen sources, N_2O production is influenced by various other factors, including temperature (Li *et al.*, 2019; Humbert *et al.*, 2020), pH (Su *et al.*, 2019; Kemmou and Amanatidou, 2023), and external additives (Caranto and Lancaster, 2017; Terada *et al*., 2017).

Temperature: Within a certain range, microbial activity increases with rising temperature. However, different microorganisms exhibit varying temperature sensitivities, potentially leading to the accumulation of intermediate products in the nitrogen conversion process or altering reaction equilibria, thereby affecting N_2O production(Chen *et al.*, 2019).

pH: The pH level influences both the form of substances in wastewater and microbial activity (Pan *et al.*, 2012; Su *et al.*, 2019). Generally, alkaline conditions are believed to reduce N_2O production.

External Additives:

- Hydroxylamine: The addition of external NH_2OH disrupts the balance between NH_2OH consumption and generation (Soler-Jofra *et al.*, 2016; White and Lehnert, 2016). It strengthens the electron supply to AMO (Zhao *et al.*, 2022), accelerates the conversion of NH_4^+ to NH_2OH, and enhances the contribution of the NH_2OH oxidation pathway to NO and N_2O during the aerobic oxidation of NH_4^+ (Zhao *et al.*, 2021).
- Hydrazine (N_2H_4): The addition of N_2H_4 can reduce N_2O production by inhibiting the activity of AOB and competing with NH_2OH for HAO (Zhao, Lei, *et al.*, 2022).

2.2.2 Microbial populations

The microbial populations significantly influence N_2O production. The relative abundance of various nitrogen-removing microbial groups is a critical factor affecting N_2O generation. Previous studies have reported a positive correlation between N_2O emissions and the abundance of AOB, while a negative correlation exists with AOA and NOB (Duan *et al.*, 2021). High AOB abundance promotes N_2O production,

whereas increased NOB abundance can effectively reduce nitrite accumulation, thereby mitigating N_2O generation. Studies have demonstrated that seasonal increases in N2O emissions are associated with NOB decline, leading to NO_2^- accumulation (Gruber *et al.*, 2021)

Denitrifying bacteria abundance markedly impacts final N_2O emissions. The HD pathway, the sole N_2O-consuming process in biochemical treatment, can eliminate over 80% of N_2O under sufficient carbon availability (Hink, Nicol and Prosser, 2017). Remarkably, the N_2O reduction capacity of the denitrifying microbial community typically exceeds its N_2O production ability by two to ten times. This characteristic makes denitrification a potential N_2O sink in wastewater treatment systems, capable of reducing N_2O not only from denitrification but also from other pathways (Bollon *et al.*, 2016).

N_2O production levels vary significantly among different microorganisms. For instance, AOA (Hink, Nicol and Prosser, 2017; Yin, Bi and Xu, 2018; Jung *et al.*, 2019) and Comammox (Kits *et al.*, 2019) exhibit lower N_2O emission levels compared to AOB. Furthermore, variations exist among different strains of AOB. The AOA genome does not encode typical nitric oxide reductase, and consequently, it cannot produce N_2O through denitrification. It is hypothesized that AOA primarily produces N_2O through coupling with abiotic reactions (Stieglmeier *et al.*, 2014). The chemical reaction between NH_2OH and NO is considered the main source of N_2O in AOA, resulting in an overall lower N_2O yield compared to AOB (Hink, Nicol and Prosser, 2017). Comammox bacteria, which lack genes related to nitrogen oxide (NOx) production, also demonstrate lower levels of N_2O production (Palomo *et al.*, 2018).

2.2.3 Process parameters

Process parameters, particularly DO concentration and recirculation ratio, play crucial roles in N_2O production during wastewater treatment.

2.2.3.1 DO

DO levels influence N_2O formation in both aerobic and anaerobic zones.

Aerobic zone:

- low DO (< 0.5 mg/L): inhibits NOB, leading to nitrite accumulation and enhanced N_2O production through AOB denitrification(Wunderlin *et al.*, 2012; Castellano-Hinojosa *et al.*, 2018).
- Moderate DO (0.5-1.0mg/L): promotes N_2O production via NH_2OH oxidation pathway (Aboobakar *et al.*, 2013; Peng *et al.*, 2014).
- Higher DO (>1.0mg/L): No further promotion of NH_2OH oxidation pathway. Elevated DO also stimulates NOB activity, reducing NO_2^- accumulation and N_2O production from AOB denitrification (Yue *et al.*, 2018).

As AOB denitrification is a more substantial N_2O source than NH_2OH oxidation (Aboobakar *et al.*, 2013; Peng *et al.*, 2014), overall N_2O production tends to decrease with increasing DO within the range of 0.5 to 3.0 mg/L (Yan *et al*., 2019; He *et al*., 2023).

Anaerobic zone:

- Presence of DO inhibits nitrous oxide reductase activity, resulting in incomplete denitrification and increased N_2O production through the HD pathway (Liang *et al.*, 2015; Yue *et al.*, 2018).
- Strict anaerobic conditions allow complete heterotrophic denitrification, promote sufficient reduction of N_2O, and therefore minimize N_2O emissions (Zhu *et al.*, 2013; Yue *et al.*, 2018).

2.2.3.2 Recirculation ratio

Recirculation ratio affects N_2O production by altering carbon source and ammonium (NH_4^+) concentrations, thereby influencing microbial activity. As the ratio gradually increases within a certain range, the N_2O release rate exhibits a trend of first decreasing and then increasing from a certain point (Zhang *et al.*, 2024).

- **Low recirculation ratio**: Insufficient dilution of NH_4^+, leading to high free ammonia (FA) concentrations, which inhibits AOB and NOB activity, with NOB being more severely affected NOB (Law, Lant and Yuan, 2011; Kinh *et al.*, 2017). Consequently, NH_2OH accumulates and is oxidized to N_2O and HNO, in which HNO further reacts to form N_2O under low DO conditions through polymerization and hydrolysis (Duan *et al.*, 2021).

- **High recirculation ratio**: can result in carbon source deficiency and nitrate accumulation, leading to incomplete denitrification and increased N_2O production (Kemmou and Amanatidou, 2023).
- **Optimal recirculation ratio**: achieves low FA and low NO_2^- levels, minimizing N_2O generation through NH_2OH oxidation pathway and ND pathway. Optimal recirculation ratios enhance nitrogen removal while minimizing N_2O release (Zhang *et al.*, 2024).

In summary, maintaining appropriate DO levels in both aerobic and anaerobic zones, while finding the optimal recirculation ratio, can effectively reduce N_2O production through various pathways and enhance overall nitrogen removal efficiency.

2.3 Mechanistic N_2O modelling

Dynamic mechanistic N_2O models based on the different generation pathways are summarised in Table 2.1.

Table 2.1 N_2O mathematical models based on generation pathways

Pathway	Model characteristics	Reference
NN	Three-step NH_3 oxidation via NH_2OH and NOH, NOH chemically decomposes to N_2O.	(Law *et al.*, 2012)
	Three-step NH_3 oxidation via NH_2OH and NO, NO reduced to N_2O; no oxygen inhibition.	(Ni *et al.*, 2013)
ND	Two-step NH_3 oxidation, two-step NO_2^-reduction; oxygen inhibits the reduction of NO_2^- and NO.	(Ni *et al.*, 2011)
	Two-step NH_3 oxidation, two-step NO_2^-reduction; FA and FNA as substrate for AOB growth; no oxygen inhibition, with an added FA inhibition term.	(Pocquet, Queinnec and Spérandio, 2013)
	One-step NH_3 oxidation, two-step NO_2^-reduction; no oxygen inhibition.	(Mampaey *et al.*, 2013)
	One-step NH_3 oxidation, two-step NO_2^-reduction; FA and FNA inhibit NH3 oxidation; oxygen inhibition represented by Haldane function.	(Guo and Vanrolleghem, 2014)
NN + ND	Three-step NH_3 oxidation, one-step NO_2^-reduction; introduction of electron carrier Mred and Mox; simulating electron competition; no oxygen inhibition; not considering cell growth.	(Ni *et al.*, 2014)
	Three-step NH_3 oxidation, one-step NO_2^-reduction; introduction of electron carrier Mred and Mox;	(Peng *et al.*, 2016)

	simulating electron competition; no oxygen inhibition; not considering cell growth; introduction of ATP/ADP linking energy synthesis and consumption processes	
	2-P model, includes five consecutive enzymatic reactions; oxygen inhibition represented by Haldane function; describes the trend of NO/N_2O changes; predicts N_2O emission factors.	(Pocquet *et al.*, 2016)
	Not considering NO production and consumption; including AOB and NOB growth and decay; suitable for very low DO conditions.	(Ding *et al.*, 2018)
HD	ASMN, coupling carbon oxidation and nitrogen oxide reduction; four-step denitrification.	(Hiatt and Grady, 2008)
	ASM-ICE, coupling carbon oxidation and nitrogen oxide reduction; four-step denitrification; introduction of electron carriers.	(Pan, Ni and Yuan, 2013)
	Denitrification reaction rate analogous to electrical current intensity through parallel resistors in a circuit; using fewer parameters.	(Domingo-Félez and Smets, 2020)
	Not considering NO, three-step denitrification; not considering the inhibitory effect of NO.	(Pavissich *et al.*, 2012)
	Extended ASM2d; three-step denitrification; added NO_2^- Inhibition term.	(Wisniewski, Kowalski and Makinia, 2018)
	Introduction of denitrifying polyphosphate-accumulating organisms (DPAOs) and denitrifying glycogen-accumulating organisms (DGAOs) on N_2O production during denitrifying phosphorus removal (DPR) system; four-step denitrification; covering N_2O production, nitrogen oxide reduction, phosphate release and uptake and intracellular polymers turnover.	(Ren *et al.*, 2023)
NN +ND +HD	ASM2d-N_2O, combining the 2-P and ASMN models, expanding the ASM2d model, covering nitrogen, phosphorus, and organic matter removal as well as N_2O stripping; simulating different DO levels for N_2O emissions to determine the optimal aeration strategy.	(Massara *et al.*, 2018)
	Combining the 2-P and ASMN models; adding TIC component to describe the impact of CO_2 absorption on pH; simulating the N_2O production in the SHARON reactor.	(Mampaey *et al.*, 2019)
NN +ND	NDHA model, combining and revising the 2-P and ASMN models; NO as a precursor to N2O	(Domingo-Félez and Smets, 2016)

+HD +abiotic	production; adding two abiotic pathways; calibrated with a specific respiratory stoichiometry.

2.3.1 Modelling of N_2O produced by AOB

Based on the two pathways of N_2O production by AOB, namely the NN pathway and the ND pathway, models of N_2O production during nitrification process can be classified as single-pathway and dual-pathway models.

2.3.1.1 Single pathway model

Several studies have proposed single-pathway models to explain N_2O production during AOB activity (Law *et al.*, 2012; Mampaey *et al.*, 2013; Ni *et al.*, 2013; Pocquet, Queinnec and Spérandio, 2013; Guo and Vanrolleghem, 2014). These models focus on either the hydroxylamine oxidation pathway or nitrifier denitrification pathway, therefore struggling to capture full complexity of the production.

hydroxylamine oxidation pathway

There are two primary models based on the incomplete oxidation of hydroxylamine: the NH_2OH/NOH model (Law *et al.*, 2012) and the NH_2OH/NO model (Ni *et al.*, 2013). In the former, N_2O production results from the spontaneous chemical decomposition of unstable NOH, bypassing the need for biological enzyme catalysis. Conversely, the latter model proposes that NH_2OH serves as an electron donor for NO reduction to N_2O, a process catalysed by biological enzymes. While these models effectively describe N_2O emissions under conditions of high DO and low NO_2^- accumulation, they fall short in predicting the impact of elevated NO_2^- levels on N_2O production.

Law et al. (2012) and Ni and Ye et al. (2013) proposed distinct models for the NH_2OH oxidation pathway. Law et al. (2012) hypothesized that N_2O production resulted from the chemical decomposition of unstable NOH, an intermediate in NH_2OH oxidation. In contrast, Ni and Ye et al. (2013) simulated NH_2OH oxidation as an electron donor process, generating NO, which is subsequently reduced to N_2O under the catalysis of nitric oxide reductase. Notably, Ni and Yuan (2015).assumed that DO did not inhibit NO reduction.

Nitrifier denitrification pathway

Two primary models describe AOB denitrification: the four-step model (Ni *et al.*, 2011) and the three-step model (Mampaey *et al.*, 2013). The former incorporates DO inhibition on AOB-mediated NO_2^- and NO reduction, while the latter employs DO as a substrate to investigate its impact on N_2O and NO emissions but neglects NH_2OH production, limiting its ability to explain N_2O peaks associated with NH_2OH kinetics.

Ni et al. (2011) posits that NO_2^- serves as the final electron acceptor, with NO being an intermediate produced from NH_2OH oxidation. Conversely, Mampaey et al. (2013) propose a coupled ammonia oxidation and denitrification process where NH_3 is the electron donor for NO_2^- reduction.

A key distinction between the two models is the role of DO. Ni et al. (2011) incorporates DO inhibition of AOB denitrification, while Mampaey et al. (2013) does not. Additionally, Ni et al. (2011) distinguishes two steps in ammonia oxidation (NH_3 to NH_2OH, then to NO_2^-), whereas Mampaey et al. (2013) propose a direct conversion of NH_3 to NO_2^-.

Subsequent research has expanded upon these foundational models. Building upon four-step model (Ni *et al.*, 2011), Pocquet, Queinnec and Spérandio (2013) proposed a model that excludes DO inhibition while incorporating pH effects on AOB reaction rates and considering FA and FNA as substrates for AOB denitrification. Guo and Vanrolleghem (2014), in response to Mampaey et al. (2013) , incorporated DO inhibition using the Haldane function for NO_2^- and NO reduction and included FA and FNA inhibition on AOB activity.

Evaluation

Evaluations of single-pathway models by Ni and Yuan et al. (2013) and Spérandio et al. (2016) using batch and long-term wastewater treatment plant data demonstrated their ability to accurately predict NH_4^+, NO_2^-, and NO_3^- concentrations. However, these models consistently failed to reproduce measured N_2O data, suggesting the simultaneous occurrence of both NH_2OH oxidation and AOB denitrification pathways, with their relative contributions varying under different operational conditions (Pocquet *et al.*, 2016).

Peng et al. (2015) further investigated the applicability of single-pathway models. Their findings indicate that the NH_2OH oxidation model is suitable for high DO (>1.5 mg O_2/L) with low NO_2^- accumulation (0-5.0 mg N/L), while the AOB denitrification model is

effective under low DO (<0.5 mg O_2/L) with varying NO_2^- levels or high DO (>0.5 mg O_2/L) with NO_2^- accumulation exceeding 1.0 mg N/L at non-inhibitory concentrations (note: NO_2^- concentrations exceeding 50mg N/L inhibit AOB denitrification).

The limited scope of single-pathway models highlights the need for a unified multi-pathway model to accurately capture the dynamic nature of N_2O production by AOB across different operation conditions.

2.3.1.2 Dual pathway model

The dual-pathway model of AOB addresses the limitations of single-pathway models in comprehensively describing N_2O production. This model incorporates two primary approaches: the decoupling method based on electron carriers and the direct coupling method.

Ni et al. (2014) pioneered a decoupling method that categorizes the complex biochemical reactions and electron transfer processes in AOB metabolism into three oxidation and three reduction reactions. By utilizing reduced mediator (Mred) and oxidized mediator (Mox), this model decouples oxidation and reduction reactions. It effectively predicts the relative contribution of AOB to total N_2O production under varying DO and NO_2^- concentrations, assuming constant inorganic carbon levels.

Building upon the previous work, Peng et al. (2016) proposed an enhanced decoupling method based on electron and energy balance. This model incorporates adenosine triphosphate (ATP) and adenosine diphosphate (ADP), linking biomass growth energy with inorganic carbon fixation. As a result, it elucidates the impact of spatiotemporal changes in inorganic carbon concentration on AOB growth and N_2O production through different pathways.

However, both models have limitations in describing NO production. They assume that NO consumption occurs primarily within AOB cells without accumulation or release, leading to inaccurate predictions of gaseous NO emissions (Ni *et al.*, 2014). Additionally, these models fail to account for the influence of pH on N_2 production.

To address these shortcomings, Pocquet et al. (2016) developed the 2-P model (two-pathway model), which directly couples the two AOB pathways and measures both NO and N_2O emissions. This model accurately predicts the combined effects of free

nitrous acid (FNA) and DO, the impact of NO_2^- concentration changes on N_2O production, and intermediate NO emission trends. It also compensates for the single-pathway model's inability to predict changes in the NO/N_2O ratio. However, the 2-P model cannot describe N_2O production under dynamically changing inorganic carbon concentrations.

While dual-pathway models offer improved N_2O production descriptions compared to single-pathway models, they present challenges due to their numerous parameters and calibration difficulties (Maktabifard *et al.*, 2022). Single-pathway models may be more suitable under specific conditions outlined in the previous chapter. For scenarios outside these conditions, dual-pathway models offer a more accurate representation of N_2O production.

2.3.2 Modelling of N_2O produced by HDB

The ASM1 model simplifies denitrification as a single-step process. While subsequent advancements led to the development of multi-step models. Kotlar et al. (1996) proposed two-step model but did not include the intermediate N_2O. Pavissich et al. (2012) introduced three-step model that incorporate N_2O. However, this model did not consider the inhibitory effects of NO on AOB and NOB. To overcome these limitations and comprehensively understand the accumulation of all denitrification intermediates, four-step denitrification models have been widely adopted.

Hiatt and Grady (2008) developed the activated sludge model for nitrogen (ASMN), coupling nitrogen oxide reduction with organic carbon oxidation through a single redox reaction. This model accounts for pH, temperature, and substrate inhibition (FA and FNA), providing insights into activated sludge performance under high nitrogen conditions. However, it overlooks the critical relationship between electron availability from carbon oxidation and the electron demand of the four denitrification steps. Insufficient electron supply can lead to electron competition, impacting N_2O accumulation.

Pan, Ni and Yuan (2013) introduced the Activated Sludge Model with Indirect Coupling of Electrons (ASM-ICE) model, incorporating electron carriers (Mred and Mox) to indirectly couple carbon oxidation with nitrogen reduction. The model simulates electron competition among denitrification processes by adjusting the affinity

constants of these electron carriers. While this model can predict N_2O accumulation, its accuracy in predicting NO emissions is hindered by parameter uncertainties and limited kinetic data for NO reduction. Although subsequent work by Pan et al. (2015) demonstrated the ASM-ICE model's superiority in capturing electron competition and intermediate accumulation compared to the ASMN model, and its increased complexity due to additional parameters can pose challenges for calibration and may lead to overparameterization.

Domingo- Félez and Smets (2020) proposed the Activated Sludge Model with Electron Competition (ASM-EC) model, drawing an analogy between electron competition in respiratory processes and electron distribution across multiple resistances. This model effectively describes organic carbon oxidation and four-step denitrification with fewer parameters than ASMN and ASM-ICE, providing accurate predictions of denitrification intermediates and enabling optimization of carbon dosage.

All these models employ a four-step denitrification pathway, involving NO_2^-, NO, and N_2O as intermediates. Typically, NO inhibition is incorporated into kinetic rate expressions using a term ($K_{I,NO,H}$) (Mampaey *et al.*, 2019). Recent empirical evidence indicates negligible NO accumulation during anaerobic phases, allowing for the simplification of the four-step model to a three-step process, directly reducing NO_2^- to N_2O (Pavissich *et al.*, 2012; Wisniewski, Kowalski and Makinia, 2018).

While three-step model (Pavissich *et al.*, 2012) addressed NO_2^- and NO reduction, it overlooked NO inhibition. Wisniewski, Kowalski and Makinia (2018) extended this model, incorporating a NO_2^- switch function into the ASM2d framework to account for NO_2^- inhibition. This refined model effectively predicts N_2O and exhibits strong agreement with effluent COD and PO_4^- concentrations (Hongbo *et al.*, 2020).

Ren et al. (2023) developed a model to elucidate the dynamic production of N_2O within denitrifying phosphorus removal (DPR) systems inhabited by denitrifying polyphosphate-accumulating organisms (DPAOs) and denitrifying glycogen-accumulating organisms (DGAOs). The model explores the interplay of competition and cooperation among these microorganisms during the four-step denitrification process, emphasizing the pivotal role of polyhydroxyalkanoate (PHA) and glycogen storage and utilization in N_2O generation. Incorporating four distinct denitrification

pathways for both DPAOs and DGAOs, including (1) anoxic polyphosphate/glycogen storage, (2) anoxic biomass growth, (3) anoxic endogenous respiration, and (4) anoxic PHA respiration, the model accurately reproduced N_2O production dynamics observed in three DPR systems, highlighting the complex interactions between DPAOs and DGAOs (Ren *et al.*, 2023).

2.3.3 Models coupling AOB and HDB pathways

Recent studies have recognised the interplay between AOB and heterotrophic denitrifiers in overall N_2O production (Aboobakar *et al.*, 2013; Ni *et al.*, 2013; Rodriguez-Caballero *et al.*, 2013; Ni *et al.*, 2015). Consequently, integrated models that incorporate both processes have emerged and offer a more accurate description of N_2O production dynamics. These models can be categorized as follows:

1) **ASM1-type Models**: These models couple one of the single-pathway models of AOB with the ASMN model (Ni *et al.*, 2011; Pocquet, Queinnec and Spérandio, 2013; Guo and Vanrolleghem, 2014). They have successfully described N_2O emissions in mixed culture nitrification-denitrification systems and determined the relative contributions of AOB and heterotrophic denitrifying bacteria to N_2O production.
2) **Electron Balance-based Models**: These models combine the dual-pathway model of AOB with the ASMN model (Ni *et al.*, 2015). They have proven effective in describing N_2O emissions in mixed culture systems.
3) **Complete Electron Balance Models**: These models integrate the dual-pathway model of AOB with the ASM-ICE model (Wang *et al.*, 2016). However, further testing is required to validate their effectiveness.
4) **NDHA Model**: The NDHA (Nitrifier Nitrification, Nitrifier Denitrification, Heterotrophic Denitrification, and Abiotic Reaction) model predicts dynamic changes of NO and N_2O under varying conditions of NH_4^+, NO_2^-, and DO (Domingo-Félez *et al.*, 2017). It can qualitatively capture the distribution of NO and N_2O under high or low DO conditions and is calibrated by respiration measurement to assess the uncertainty of N_2O production.
5) **ASM2d-N_2O model**: Massara et al. (2018) expanded the ASM2d model into the ASM2d-N_2O model by incorporating elements from the 2-P and ASMN models. This comprehensive model simulates nitrogen, phosphorus, and organic matter

removal while considering three biological N_2O production pathways and N_2O stripping. It offers a valuable tool for optimizing aeration strategies by predicting N_2O emissions under varying DO concentrations.

While these models provide valuable insights, they often overlook the role of intracellular polymers, particularly PHA, in N_2O accumulation during heterotrophic denitrification under external carbon source limitation (Zhou *et al.*, 2012). To address this gap, updated models have been developed:

1) Liu et al. (2015) addressed this by linking heterotrophic growth with intracellular polymers under external carbon source limitation and N_2O generation and consumption, though autotrophic pathway is not covered.
2) Ding et al. (2016) extended the ASM3 model to encompass N_2O production during both autotrophic nitrification and heterotrophic denitrification, including the influence of intracellular polymers in the A^2O process

In summary, mechanistic modelling of N_2O has advanced considerably, enabling qualitative analysis of production mechanisms and quantitative prediction of emissions from wastewater treatment. Nevertheless, incomplete understanding of N_2O generation processes hinders the development of a unified model structure and limits the models' ability to accurately represent diverse operational conditions. Furthermore, challenges in model calibration and validation compromise the reliability of quantitative N_2O emission predictions (Seshan *et al.*, 2024). To date, the successful full-scale implementation of mathematical N_2O modelling in real WWTPs remains scare and presents significant challenges (Khalil *et al.*, 2024).

2.4 Data-driven N_2O modelling

While mechanistic models have traditionally been employed to simulate N_2O production and emissions in wastewater treatment, data-driven approaches offer a complementary perspective. Unlike biokinetic models, which rely on theoretical underpinnings, data-driven methods extract patterns and relationships directly from process monitoring data, providing a more practical alternative.

2.4.1 Types of data-driven methods used in wastewater

In recent years, advancements in machine learning (ML) and artificial intelligence (AI) have significantly boosted the popularity of AI-based data-driven techniques, many of which have demonstrated state-of-the-art performance (LeCun, Bengio and Hinton, 2015). Industries increasingly leverage 'big data' to construct data-driven models for addressing critical challenges within their respective fields (Zhong *et al.*, 2021).

Given the broad scope of machine learning and artificial intelligence, a wide range of algorithms can be applied to data-driven approaches. To facilitate understanding within this section, a brief overview of various algorithm types is provided below, categorized by their nature and function.

- **Fuzzy logic**

  Fuzzy logic techniques are based on fuzzy set theory, which allows for degrees of truth rather than the classical binary logic. In wastewater treatment, these techniques are often used in risk assessment models (Flores-Alsina *et al.*, 2009). Fuzzy logic controllers (FLC) use linguistic variables and if-then rules to make decisions, mimicking human reasoning. This approach is well-suited for wastewater treatment processes where precise mathematical models are difficult to develop due to the system's complexity (Chiranjivi *et al.*, 2024). Advantages include the ability to handle nonlinearity, incorporate expert knowledge, and operate effectively with noisy or incomplete data. Fuzzy logic has been applied in various aspects of wastewater treatment, including aeration control, sludge bulking prediction, and pH control (Nadiri *et al.*, 2018). It can improve process stability, reduce energy consumption, and enhance overall treatment efficiency. However, limitations exist. Designing effective fuzzy rules requires expert knowledge, which can be subjective; achieving optimal performance for complex systems might be potentially difficult; and there is a lack of learning ability compared to some machine learning techniques (Vijayaraghavan and Jayalakshmi, 2015).

- **Time series forecasting**

  Time series forecasting is a statistical technique used to predict future values based on historical time-ordered data. In wastewater treatment, it can be applied to predict various parameters such as influent flow rates, pollutant concentrations,

or treatment efficiency over time (Berthouex and Box, 1996). The main advantage of time series forecasting is its ability to capture temporal patterns, seasonality, and trends in data, which is particularly useful in wastewater treatment where many processes exhibit cyclical or seasonal variations (Q. Zhang *et al.*, 2019). Common methods include Autoregressive Integrated Moving Average (ARIMA), exponential smoothing, and more advanced techniques like employing Long Short-Term Memory (LSTM) neural networks (Lim and Zohren, 2021). These models can help optimize treatment processes, predict potential system overloads, detect fault or anomaly, and improve resource allocation. However, time series forecasting has limitations: it assumes that past patterns will continue into the future, which may not always hold true in dynamic wastewater systems. Additionally, these models may struggle with sudden, unpredictable events or changes in system behaviour (Kang *et al.*, 2020). Despite these challenges, time series forecasting remains a valuable tool in wastewater treatment, offering insights for operational decision-making and long-term planning when used in conjunction with domain expertise and other modelling approaches (Li and Wang, 2021).

- **Non-neural-network ML**

  Often refer to conventional ML methods, such as PCA, K-means clustering, decision trees, SVM, K-nearest neighbours (KNN), Random Forests (RF), Gradient Boosting Machine (GBM) and Adaptive Boosting (AdaBoost). These algorithms offer diverse approaches to data analysis and prediction in wastewater treatment (Khalil *et al.,* 2023). They excel in different areas: PCA for dimensionality reduction and feature extraction (Tao *et al.*, 2013; Abba, Elkiran and Nourani, 2021); K-means for data clustering (Laili, Indrasti and Wahyudi, 2022); decision trees for interpretable rule-based decisions (Logan, Roberts and Smith, 2024); SVM for robust classification and regression (Ribeiro, Sanfins and Belo, 2013; Cheng *et al.*, 2019); KNN for pattern recognition (Kim *et al.*, 2016); Random forests for ensemble learning and handling complex datasets (P. Zhou *et al.*, 2019); and GBM or AdaBoost for boosting weak learners (Bagherzadeh *et al.*, 2021; Gholizadeh *et al.*, 2024). In wastewater treatment, these techniques have been applied to various tasks such as process optimization, fault detection, effluent quality prediction, and operational parameter estimation (Hafsa, Al-Yaari

and Rushd, 2021). Advantages include their ability to handle non-linear relationships, deal with high-dimensional data, and provide insights into feature importance. Many of these methods also offer good interpretability, which is crucial in process control and decision-making. However, limitations exist. They might be less accurate than deep learning for complex patterns; some require careful feature selection and parameter tuning; others may struggle with highly imbalanced datasets or extrapolation beyond the training data range. Additionally, the performance of these models can be sensitive to the quality and quantity of available data (Hilal *et al.*, 2022; Han *et al.*, 2023).

- **DNN-based models**

  Deep learning employs artificial neural networks with multiple layers to extract complex patterns from data (Alvi *et al.*, 2023). MLPs are foundational (Shen *et al.*, 2024), while Convolutional Neural Networks (CNNs) excel at spatial pattern recognition data (Wenbing and ZHANG, 2020; Y. Li *et al.*, 2023), and Recurrent Neural Networks (RNNs) (Pisa *et al.*, 2018; Wongburi and Park, 2023), including LSTM and its variants (Pisa, Santin, *et al.*, 2019; Yaqub *et al.*, 2020; Farhi *et al.*, 2021; Xu *et al.*, 2023), are particularly effective for time-series forecasting and capturing long-term dependencies in process data. Neural ODEs offer continuous-time modelling in wastewater treatment (Quaghebeur *et al.*, 2022). DNN-based models excel at predicting complex process dynamics, optimizing operations, and detecting anomalies (Mamandipoor *et al.*, 2020; G. Wang *et al.*, 2022; J.-H. Wang *et al.*, 2022; Lin, Hanyue and Bin, 2022; Zhang *et al.*, 2023; Shaban *et al.*, 2024). They can be also used in model predicative control (MPC) (Bernardelli *et al.*, 2020; Wang *et al.*, 2020, 2023; He, Zhang and Li, 2021; Yuting Liu *et al.*, 2023). Their advantages include the ability to automatically extract relevant features, handle large volumes of data, and adapt to changing conditions. However, it requires substantial computational resources, large datasets, and careful hyperparameter tuning. The black-box nature of deep models can hinder interpretability. Additionally, overfitting and out-of-distribution (OOD) generalisation might be an issue if training data are not sufficiently representative (Ng *et al.*, 2020).

- **Generative neural networks**

  Generative neural networks represent cutting-edge techniques in machine learning with emerging applications in wastewater treatment. Generative

Adversarial Networks (GANs) generate synthetic data for augmenting limited datasets, simulating rare events in treatment processes, and improving model performance (Asadi and McPhedran, 2021; Rani *et al.*, 2024). Autoencoders learn efficient data representations for anomaly detection, data denoising and feature extraction (Ba-Alawi *et al.*, 2021, 2022; Peng *et al.*, 2022; Salles *et al.*, 2022; Zhang, Suzuki and Shioya, 2022). Transformers excel in capturing long-range dependencies in time-series data, enabling advanced forecasting and process optimization (Huang *et al.*, 2021; Peng and Fanchao, 2022; Chang, Zhang and Wang, 2023). LLMs can generate human-like text, potentially aiding in report generation, knowledge management, and decision support (B. Xu, Wen, *et al.*, 2024; Liang *et al.*, 2024). While these models offer immense potential for improving wastewater treatment processes, challenges include data quality, computational requirements, and model interpretability (B. Xu, Wen, *et al.*, 2024).

- **Genetic algorithm optimisation**

  Genetic Algorithm (GA) optimization and its variants like Simulated Annealing (SA), Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO) are metaheuristic methods inspired by natural processes (Holenda *et al.*, 2007; Huang *et al.*, 2015). GA mimics biological evolution, using selection, crossover, and mutation to optimize solutions (Beraud, Lemoine and Steyer, 2009; Iqbal and Guria, 2009; Bagheri *et al.*, 2015). SA is inspired by the annealing process in metallurgy, gradually cooling a solution to reach an optimal state (Govindarajan, Kumar and Karunanithi, 2005; Zeferino, Antunes and Cunha, 2009; Cunha and Antunes, 2012). PSO simulates the social behaviour of birds flocking to find the best position (Khoja *et al.*, 2018; Ye *et al.*, 2019; Lu *et al.*, 2021; Su *et al.*, 2022), and ACO models the foraging behaviour of ants to find optimal paths (Verdaguer, Clara and Poch, 2012; Verdaguer *et al.*, 2014; Afshar *et al.*, 2015; Verdaguer, Molinos-Senante and Poch, 2016). They excel at optimising complex, non-linear, and multi-objective problems, such as energy minimization, effluent quality improvement, and process control. However, they can be computationally expensive and sensitive to parameter tuning, requiring careful implementation and problem-specific adaptations (Béraud *et al.*, 2007; Iqbal and Guria, 2009).

- **Reinforcement learning**

Reinforcement learning (RL) is a machine learning paradigm where an agent learns to make decisions by interacting with an environment and receiving rewards or penalties. Key features include the use of value functions, policy gradients, and model-free or model-based approaches (Hernández-del-Olmo, Llanes and Gaudioso, 2012). It has shown promise in wastewater treatment by optimizing complex processes like aeration, chemical dosing, and energy management, leading to enhanced treatment efficiency, reduced operational costs, and improved compliance with environmental regulations (Chen *et al.*, 2021; Yang *et al.*, 2021; Aponte-Rengifo *et al.*, 2023). RL agents can learn optimal control policies without explicit programming, adapting to dynamic conditions and improving performance over time. However, RL requires significant computational resources, data, and careful tuning of hyperparameters, and may suffer from issues such as exploration-exploitation trade-offs and convergence difficulties (Hernández-del-Olmo *et al.*, 2018; Yang *et al.*, 2021).

- **Hybrid models**

  Hybrid models in wastewater treatment combine mechanistic models, grounded in process understanding, with machine learning models for data-driven insights (Cheng *et al.*, 2023). These models leverage the strengths of both approaches and mitigate their weaknesses, offering enhanced predictive accuracy and adaptability compared to standalone approaches (Lotfi *et al.*, 2019; Quaghebeur, Torfs, Baets, *et al.*, 2022). By leveraging mechanistic knowledge, they provide interpretable results, better generalization to new conditions and handle data scarcity effectively (Bagheri *et al.*, 2015; Mahjouri *et al.*, 2017; Asadi and McPhedran, 2021; Mehrani *et al.*, 2022). By fitting data pattern, they enhance the prediction accuracy, identify the key factors and offer practical solutions (Ye et al., 2019; Heo et al., 2021; B. Xu, Pooi, et al., 2024; Lancioni et al., 2024). However, they can be complex to develop and require careful integration of diverse modelling techniques. Applications span process optimization, real-time control, and effluent quality prediction, contributing to improved wastewater treatment efficiency (Li *et al.*, 2021; B. Xu, Pooi, *et al.*, 2024).

- **Auxiliary approaches**

  Auxiliary approaches like Computational Fluid Dynamics (CFD), Global Sensitivity Analysis (GSA), and Monte Carlo simulation enhance data-driven wastewater

treatment modelling by providing insights into process dynamics, parameter sensitivity, and uncertainty quantification (Porro *et al.*, 2019). CFD simulates fluid flow and transport phenomena, offering detailed spatial and temporal information on treatment processes (Le Moullec *et al.*, 2010; Pishnamazi *et al.*, 2012). It is often used for optimizing reactor designs and flow patterns in wastewater (Porro *et al.*, 2019; Patziger, 2021). GSA helps identify key parameters influencing model outputs, improving model understanding and simplification (Sin *et al.*, 2011; Cosenza *et al.*, 2014; Baalbaki *et al.*, 2017; Al *et al.*, 2019). Monte Carlo simulation assesses model robustness by propagating uncertainties through the model, improving prediction reliability (Carrasco and Chang, 2005; Taheriyoun and Moradinejad, 2015; Zhao *et al.*, 2017; Long *et al.*, 2019; Migdał *et al.*, 2022). These methods complement data-driven models by providing mechanistic understanding and enhancing model interpretability, leading to improved process control, optimization, and risk assessment in wastewater treatment (Samstag *et al.*, 2016; Hong *et al.*, 2022).

2.4.2 Data-driven models for N_2O simulation

Data-driven models are capable of solving classification and regression prediction problems by learning implicit associations among variables within large datasets (LeCun, Bengio and Hinton, 2015). Due to the numerous parameters and complex, variable influencing factors inherent to wastewater treatment processes, data-driven models offer substantial advantages for predicting system behaviours (Zhong *et al.*, 2021). In recent years, researchers have explored the application of data-driven models for simulating wastewater treatment processes (Bahramian *et al.*, 2023), enabling model-based optimization of key processes, such as improving pollutant removal efficiency and reducing energy consumption (Newhart *et al.*, 2019).

As an intermediate product of nitrogen transformation in wastewater, N_2O can also be simulated and predicted using data-driven models (Hwangbo *et al.*, 2021), provided N_2O concentration data are available. By predicting N_2O production under varying operating conditions, these models can inform strategies for N_2O reduction or minimization (Lu *et al.*, 2023). Table 2.2 provides an overview of studies that have applied data-driven models to N_2O production and emissions in wastewater treatment over the past decade.

Table 2.2 Summary of studies on data-driven models including N_2O in wastewater treatment in last ten years

No	Reference	Modelling methods	Function	Claimed Results
1	(Liu *et al.*, 2024) preprint	SVR + ANN	Prediction, analysis	Hybrid model achieved accurate N_2O prediction with R^2=99.26% and MAPE=0.49%
2	(Rani *et al.*, 2024) preprint	Generative adversarial wavelet neural operator (GAWNO)	Fault detection	Showcased in the WWTP N_2O dataset, the GAWNO approach holds promise for fault detection.
3	(X. Xu *et al.*, 2024)	Basic RNN, LSTM	Prediction	The optimal LSTM model outperformed basic RNN model with 19% improvement in RMSE for N_2O prediction
4	(Daneshgar *et al.*, 2024)	Flow sheet + CFD biokinetic + risk assessment (fuzzy)	Control	Model based protocol achieved up to yearly 50% reduction primarily in N_2O emission in case study.
5	(Khalil *et al.*, 2024)	mRMR for feature selection, NSGA-II for hyperparameter optimisation	Feature selection and hyperparameter optimisation	Balanced model complexity and performance in a case study with AdaBoost models for N_2O emission prediction.
6	(Lancioni *et al.*, 2024)	ASM2d + MLP, Global sensitivity analysis	Real-time control	The hybrid model support operator to potentially reduce up to 21% GHG emissions while maintain effluent standard
7	(Tejaswini, Maheswari and Ambati, 2024)	PI, MPC Supervisory fuzzy control framework	Control	PI-MPC combination for NO_2^- and DO control showed 25% reduction in total GHG emission compared with literature, while MPC-MPC structure for DO control alone resulted in 49% reduction.
8	(Khalil *et al.*, 2023)	Framework to balance model complexity, performance	Model selection	Showcased in the WWTP N_2O dataset that the best performing models are KNN(R^2=0.88),

		and interpretability		AdaBoost(R^2=0.94) and DNN(R^2=0.90).
9	(B. Szeląg, Zaborowska and Mąkinia, 2023)	K-means	Model selection	XGboost outperformed MARS, and SVM in predicting N_2O emissions.
10	(K. Li *et al.*, 2022)	Modified ASM1 + teacher forcing LSTM	Prediction	Hybrid model outperformed white-box and black-box models, with better capability of predicting low N_2O emission (93% increase in overall performance).
11	(Mehrani *et al.*, 2022)	Mechanistic model (GPS-X) + MLP / SVM / GBM	Prediction	MLP outperformed SVM / GBM with R^2=0.67 and 95% accuracy in N^2O production prediction
12	(Seshan *et al.*, 2022) conference	ASM type biokinetic model (BioWin), ANN	Prediction	ANN model outperformed biokinetic model in N_2O prediction accuracy
13	(Quaghebeur *et al.*, 2022)	Mechanistic + Neural ODE	Prediction	Hybrid model improves predictive performance by combining the strengths of both mechanistic and data-driven model
14	(Asadi and McPhedran, 2021)	GAN + non-linear regression, and GA	Estimation	The hybrid model could reasonably determine GHG emission rate estimator.
15	(Stentoft *et al.*, 2021)	Stochastic differential equations derived from ASM1, MPC	Control / optimisation	Illustrated flexibility of the proposed MPC algorithm through comparison of different control performances
16	(Hwangbo *et al.*, 2021)	GSA, DNN, LSTM	Prediction	LSTM model outperformed DNN model (R^2>94%, 64% reduced RMSE)
17	(Bae *et al.*, 2021)	Random forest	Identification	RF model identified the sOUR-ratio as the most influential trigger of N_2O emissions.
18	(Hwangbo, Al and Sin, 2020)	Integrated framework including deep learning (MLP)	Prediction	Showcased the framework well predicted N_2O concentration (R^2>0.88)

		for process modelling		
19	(Bellandi *et al.*, 2020)	PCA, k-means, agglomerative, and HDBSCAN clustering	Analysis	All identified two main N_2O pathways (NN and ND)
20	(Song *et al.*, 2020)	Random forest	Identification, analysis	Identified the most influential from aerator: dissolved inorganic carbon (DIC) and specific ammonia oxidation activity ($sOUR_{AOB}$); from anoxic: dissolved-organic-carbon to NO_2^-/NO_3^- ratio
21	(Vasilaki, Conca, *et al.*, 2020)	Changepoint detection, SVM	Identification, estimation	Changepoints coincide with changes of N_2O emission range and behaviour. SVM model can detect high risk N_2O emission periods
22	(Vasilaki *et al.* 2020)	Changepoint detection, SVM	Identification, estimation	Changepoints linked with changes of N_2O fluxes; SVM model can detect N_2O emission range.
23	(Porro *et al.*, 2019) conference	CFD + biokinetic model + Fuzzy logic control	Control	Hybrid model can assess the impact of different process control strategies and mixing conditions on reactor GHG production
24	(Vasilaki *et al.*, 2018)	Changepoint detection, Spearman's rank correlation, k-means cluster, PCA	Identification, Analysis	Multivariate analysis revealed correlations between influential factors in N_2O dynamics
25	(Porro *et al.*, 2018) conference	Fuzzy logic and knowledge-based risk assessment model	Online supervision and control	Integrated model predicts risk correlating to actual effluent TSS concentration, and facilitate to make predictive based control scheme
26	(Boiocchi, Gernaey and Sin, 2017)	Fuzzy logic	Control	Drastic reduction of N_2O emission in ND pathway dominant plants.

27	(Sweetapple, Fu and Butler, 2014)	Multi-objective NSGA-II	Optimisation / control	demonstrated the potential of proposed control strategies for the reduction of GHG emissions in a cost-effective manner, with trade-offs and optimised solutions to different problem.

A review of the literature reveals that data-driven models applied to N_2O production and emissions in wastewater treatment primarily serve three functions:

- **Identification**: Uncovering key factors influencing N_2O generation and consumption, while also implementing fault and anomaly detection, dimension reduction, feature selection, and component analysis to support data validation and the development of mitigation strategies (Bellandi *et al.*, 2020).
- **Prediction**: Forecasting N_2O production and/or emissions based on identified input variables using a constructed or trained data-driven model and evaluating N_2O emissions during wastewater treatment system operation (Khalil *et al.*, 2024).
- **Control**: Minimizing N_2O emissions through model-based prediction or direct optimisation algorithms using control methods such as traditional Proportion-Integration (PI) control, FLC, MPC, and advanced multi-objective optimization techniques like multi-agent deep reinforcement learning (MADRL) (Lu *et al.*, 2023).

2.4.2.1 Identifying Key Influencing Factors of N_2O production

Biological process in wastewater treatment plants is a major source of N_2O emissions (Vasilaki *et al.*, 2019; Hongbo *et al.*, 2020). Studies have shown that the release of N_2O in bioreactors is closely related to wastewater properties (such as COD, ammonia, nitrite, nitrate, water temperature) and process operating parameters (such as aeration, recirculation rate) (Kemmou and Amanatidou, 2023; Huang and Liu, 2024). Due to numerous variables involved in the treatment processes, machine learning methods can be used to identify the key influencing factors of N_2O emissions, thereby providing support for subsequent prediction and control.

To quantify the impact of various factors on N_2O emissions, Song et al. (2020) used the random forest method to analyse the emission mechanism and key influencing factors of N_2O in the activated sludge tank of a wastewater treatment plant. The results showed that the dissolved inorganic carbon concentration and specific ammonia

oxidation activity were most significantly correlated with N_2O emissions in the aerobic aeration tank; while in the anoxic tank, the ratio of dissolved organic carbon to nitrate nitrogen had the greatest impact on N_2O emissions. Bae et al. (2021) also used RF method to analyse the correlation between N_2O emissions and wastewater properties in a real wastewater treatment plant. The results showed that microbial changes had little impact on N_2O emissions, while the ratio of oxygen utilization rate of AOB to NOB had the greatest impact on N_2O emissions.

In addition to RF method, Hwangbo et al. (2021) employed GSA based on a DNN to identify the key parameters related to N_2O emissions in WWTPs. GSA, using variance decomposition to reveal how input variables of process conditions impact N_2O emissions. Comparing Sobol and Kucherenko index methods (Cosenza *et al.*, 2014), Hwangbo et al. (2021) found temperature, nitrate, ammonia, and influent flow rate as primary factors affecting liquid-phase N_2O concentration using the Sobol index. The Kucherenko index additionally highlighted the importance of air flow rate and dissolved oxygen setpoint These complementary findings effectively identify crucial parameters for N_2O emission control in wastewater treatment.

Bellandi et al. (2020) developed a PCA model capable of distinguishing between NN and ND pathways using two principal components. To automate the identification of these pathways, three clustering methods were applied to the PCA scatterplot. While K-means clustering provided a reasonable separation of the two primary pathways, it encountered difficulties in classifying boundary points. Conversely, both agglomerative and HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) clustering successfully differentiated between the NN and ND pathways while effectively excluding outliers. Vasilaki et al. (2018) also employed PCA and clustering techniques with multivariate statistical analysis to uncover correlations between influential factors in N_2O dynamics. Furthermore, Vasilaki et al. (2020) applied changepoint detection techniques to SCENA (Short Cut Enhanced Nutrient Abatement) process and Carrousel reactors respectively, demonstrating their capability of pinpoint the N_2O emission “hotspot” period.

The growing prevalence of online monitoring equipment and intelligent control platforms enables continuous data collection on N_2O-related parameters. These datasets serve as a foundation for developing data-driven N_2O models. By employing

data mining techniques, these models can identify crucial operational factors and detect anomalies, thereby supporting the implementation of model predictive control strategies to minimize N_2O emissions in wastewater treatment.

2.4.2.2 Predicting N_2O Emissions

Once key factors influencing N_2O emissions are identified, they can be incorporated as input variables into machine learning prediction models to forecast N_2O emissions.

Earlier studies relied on limited data to estimate emission factors or their ranges. For instance, Vasilaki and Conca et al. (2020) employed SVM to classify N_2O variations and predict N_2O levels in side-stream SBRs. The SVM classifier categorized dissolved N_2O concentration ranges, while the regression model estimated average N_2O levels. In another study, Vasilaki and Danishvar et al. (2020) utilized SVM to identify operational patterns in wastewater treatment systems and predict N_2O emission ranges, thereby providing reliable estimates of emission factors in the absence of real-time N_2O monitoring.

Given the scarcity of online N_2O monitoring in some WWTPs, Asadi and McPhedran (2021) employed GAN to generate additional virtual data related to N_2O emissions. Using temperature, DO, nitrite, nitrate, and ammonium concentrations as input variables, nonlinear regression models provided reasonable N_2O emission rate estimates.

The increasing availability of process time-series data has driven a surge in studies focused on dynamic point prediction using advanced data-driven algorithms like DNN-based models, often achieving state-of-the-art performance (Hwangbo, Al and Sin, 2020; Khalil *et al.*, 2024). Seshan et al. (2022) compared a biokinetic model and an ANN-based model for a real WWTP. The biokinetic model utilized a commercial software for simulation and a full year's real data for calibration. The ANN model employed relevant process parameters, including influent flowrate, and ammonium concentration in the aerobic tank, as inputs to predict the gaseous N_2O concentration. Results indicated that the ANN model outperformed the mechanistic model in terms of prediction accuracy.

Hwangbo, Al and Sin (2020) introduced an integrated framework incorporating a deep learning model and demonstrated its application to an industrial wastewater treatment

plant. The case study revealed that the optimally structured deep learning model significantly outperformed conventional machine learning models, yielding superior accuracy (R^2>97.7%, RMSE=0.032). In subsequent research, Hwangbo et al. (2021) employed the GSA method to identify key factors and subsequently utilized DNN and LSTM models for N_2O emission forecasting. DO, ammonia, nitrate, temperature, influent flow rate, and air flow rate were selected as input variables, with N_2O emissions as the target variable. Owing to its ability to capture long-term dependencies, LSTM model exhibited superior predictive accuracy (R^2> 94%, RMSE reduced by 64%) compared to the DNN-based model (R^2>90%).

Mehrani et al. (2022) developed a hybrid mechanistic and machine learning model for N_2O production forecasting in an experimental nitrifying SBR. The mechanistic model built with a commercial simulation software, generated predicted data, including ammonia, nitrite, nitrate concentrations, and MLSS, which were used as inputs alongside online measurements of DO, pH, and temperature for the machine learning models. Among the three ML models tested (ANN, SVM, and GBM), ANN demonstrated superior performance in predicting dissolved N_2O concentrations.

X. Xu et al. (2024) compared RNN and LSTM models for predicting N_2O emissions from a WWTP. They tuned six key hyperparameter to obtain an optimal model. Results revealed that the LSTM model significantly outperformed the RNN model, achieving an Explained Variance Score (EVS) of 93% compared to 85% for the RNN. Additionally, the LSTM model demonstrated a 19% reduction in RMSE, indicating superior prediction accuracy and robustness to sudden events.

Leveraging advanced algorithms on extensive monitoring data, data-driven models are increasingly precise in predicting N_2O emissions. By capturing the complex relationships between operational parameters and N_2O release, these models provide a robust framework for controlling and mitigating N_2O emissions in wastewater treatment plants.

2.4.2.3 Controlling and mitigating N_2O Emissions

Effective N_2O emission control must be balanced with maintaining effluent quality and minimizing costs, complicating the development of mitigation strategies. Optimization approaches can be model-free or model-based, encompassing techniques such as

traditional PI control, evolutionary algorithms, knowledge or AI-based fuzzy logic control, model predictive control, and advanced reinforcement learning (Lu *et al.*, 2023).

Sweetapple, Fu and Butler (2014) employed the NSGA-II multi-objective evolutionary algorithm to determine Pareto optimal control parameter sets for an activated sludge process. They identified effective solutions to balance emission reduction with competing objectives. Their findings demonstrate that multi-objective optimization can substantially reduce greenhouse gas emissions without plant modifications. However, trade-offs between different objectives must be carefully considered.

Fuzzy logic control is a widely adopted method for mitigating N_2O emissions in wastewater treatment plants (Lu *et al.*, 2023). Boiocchi, Gernaey and Sin (2017) implemented an FLC strategy for a WWTP, utilizing the ammonium-to-nitrate ratio as a control parameter to adjust aeration levels to inhibit N_2O production. While effective in processes with generation of N_2O dominated by the ND pathway, this approach demonstrated limitations in systems primarily influenced by the NN pathway.

Porro et al. (2018, 2019) developed a module adopting FLC as a core to regulate N_2O emissions. This module evolved from knowledge-based rules to an AI driven approach, integrating biokinetic and CFD modelling. Knowledge-based approach effectively elucidates GHG production pathways and key influential factors within wastewater systems. By assigning risk coefficients to those factors, active pathways can be identified, enabling the proposal of targeted mitigation strategies. (Ye, Porro and Nopens, 2022). Hybrid modelling, combining AI-based FLC with mechanistic or CFD models, facilitates the evaluation of various process control strategies and mixing conditions on reactor greenhouse gas production (Porro *et al.*, 2019).

Daneshgar et al. (2024) developed a hybrid model-based protocol integrating flowsheet, CFD, and N_2O risk assessment models. Tests across three case studies demonstrated that their N_2O risk model, grounded in fuzzy decision theory, effectively identified high-risk conditions. This protocol provides a robust framework for WWTPs to minimize carbon footprint while maintaining removal efficiencies.

The enhanced accuracy of data-driven models has significantly improved the performance of model predictive control (MPC) (Lu *et al.*, 2023). By accurately

predicting N_2O emissions under varying operating conditions, MPC enables targeted process control to reduce N_2O emissions. For instance, Stentoft et al. (2021) developed an MPC algorithm optimizing wastewater treatment plants across multiple objectives, including effluent quality, energy consumption, cost, and global warming impact (incorporating direct N_2O emissions and indirect carbon footprint). Prioritizing global warming minimization led to a substantial 35-43% reduction in daily GHG emissions. Lancioni et al. (2024) developed an early warning system based on MPC to mitigate GHG emissions in WWTPs. Their model integrated the ASM2d mechanistic model with a MLP deep learning model, using GSA to identify real-time biokinetic patterns. Successfully applied in a real wastewater plant, the system enabled operators to implement possible mitigation strategies that can reduce direct GHG emissions by up to 21% without compromising effluent standard.

Tejaswini, Maheswari and Ambati (2024) developed an integrated supervisory control framework on the BSM2G platform to evaluate the impact of different control strategies on GHG emissions. Comparing PI, MPC, and FLC controllers, they found that a combined PI-MPC strategy for nitrate and DO control structure achieved the best results, reducing total GHG emissions by 25% compared to literature. Additionally, an MPC-MPC strategy solely for DO control led to a 49% reduction in GHG emissions relative to previous studies.

2.4.3 Selection of data-driven models

The rapid advancement of machine learning has led to a proliferation of data-driven modelling approaches. Selecting the optimal model for accurate performance is crucial. Szeląg, Zaborowska and Mąkinia (2023) proposed an algorithm for selecting machine learning models for N_2O emission prediction in WWTPs. Employing k-means clustering, they compared MARS, SVM, and XGBoost models, finding that prediction accuracy varied based on input data variability. GSA revealed XGBoost as the only model consistently capturing relationships between all input variables and N_2O emissions.

Khalil et al. (2023) developed a comprehensive framework for selecting machine learning algorithms for real-time N_2O emission modelling, prioritising accuracy, model complexity, computational efficiency, and interpretability. Their comparative analysis of KNN, decision trees, ensemble learning, and DNN models identified KNN (R^2=0.88),

AdaBoost (R^2=0.94), and DNN (R^2=0.90) as top performers In subsequent research (Khalil *et al.*, 2024), they introduced an approach employing multi-objective optimization with the NSGA-II genetic algorithm to optimize feature selection and hyperparameters, resulting in computationally efficient online N_2O emission models for WWTPs.

Data-driven models excel at predicting N_2O emission behaviour under specific operating conditions. When provided with sufficient and representative data, these models can generate highly accurate forecasts and adapt to process changes through online updates. Data-driven models offer advantages over traditional mechanistic models by mitigating issues such as over-parameterization, sensitivity to operating conditions, and the need for extensive calibration and validation. Nevertheless, their predictive capabilities are constrained under unforeseen circumstances. In such cases, mechanistic models can be employed to augment the dataset. By harnessing their complementary strengths, the integration of mechanistic models with data-driven approaches offers promising avenues for improving N_2O emission predictions in wastewater treatment systems. This integrated approach has the potential to yield more accurate, adaptable, and comprehensive modelling tools, thereby optimizing wastewater treatment processes and minimizing GHG emissions (Khalil *et al.*, 2024).

Quaghebeur et al. (2022) developed a framework that combines a neural differential equation with a mechanistic ODE model. Their hybrid model, trained on dry weather data but evaluated under rainy conditions, exhibited improved predictive performance (RMSE=0.66) compared to the mechanistic model alone (RMSE=3.06), and enhanced generalisation ability in extrapolating to unobserved rain event, a challenge traditionally encountered by data-driven models (RMSE=5.75).

K. Li et al. (2022) developed a hybrid model for N_2O emissions in wastewater treatment plants, combining a first-principles model and a deep learning model. The first-principles model (white-box) was adapted from the ASM model and implemented in a simulation platform, while the DL model (black-box) employed a teacher-forcing LSTM algorithm. Comparative analysis with black-box and white-box models demonstrated the superior predictive accuracy of the hybrid approach. Furthermore, the study evaluated the hybrid model's applicability and potential for identifying mitigation

strategies, highlighting its promising generalizability and sensitivity to critical factors influencing N_2O emissions.

In summary, the convergence of burgeoning machine learning and AI algorithms with the widespread adoption of online monitoring equipment and the growing volume of process data has significantly advanced data-driven modelling for predicting and understanding N_2O generation and emissions in wastewater treatment plants.

Chapter 3 NODE fundamentals

NODEs represent a novel approach to neural networks that replaces the discrete layers of traditional architectures with a continuous transformation of data. This transformation is governed by a system of ODEs (Chen *et al.*, 2018).

The core concept of NODEs is that the hidden state of the network evolves over a continuous time interval according to a differential equation. The initial state corresponds to the input data, while the final state represents the output. This continuous evolution allows for more expressive and flexible models compared to traditional layer-based architectures (Kidger 2022).

While NODE can replace residual networks for supervised learning, or describe continuous normalizing flows, or generate latent function for time-series forecasting, this study emphasized the particular interest on the potential of NODEs in solving dynamical systems, such as those encountered in wastewater treatment processes.

3.1 Concept and methodology

Using ODEs to describe dynamical systems is a fundamental approach across various scientific and engineering fields (Gear, 1981; Hairer and Wanner, 1996). These equations define the rate of change of key system variables, thereby expressing how the system evolves with respect to time. Complex systems, such as wastewater treatment processes, can be described using a system of mechanistic ODEs, generally formalized in Equation 3-1

$$\frac{\boldsymbol{dY(t)}}{\boldsymbol{dt}} = \boldsymbol{f(Y(t), t)} \qquad \textbf{Equation 3-1}$$

Where t denotes time; $Y(t)$ represents the state vector, which evolves over time. f is the core function defining the relationships between the states and their corresponding derivatives.

Given an initial value, the evolution of the dynamical system can be visualized by solving the ODEs. This scenario is often referred to as an initial value problem (IVP). The solution to an IVP provides a trajectory of the system states over time, starting from the specified initial condition (Postawa, Szczygieł and Kułażyński, 2020).

NODEs extend this concept by bridging the gap between ODEs and deep learning. NODEs can be viewed as a continuous-time extension of residual neural networks (ResNets). While standard neural networks consist of discrete layers that transform the input, NODEs replace these discrete transformations with a continuous transformation defined by ODEs (Chen *et al.*, 2018).

The key innovation of NODEs is modelling the evolution of the hidden state as a continuous-time dynamical system. A neural network is constructed to approximate the function f that computes the derivative of the hidden states. Mathematically, a NODE can be defined as:

$$\frac{\boldsymbol{dY(t)}}{\boldsymbol{dt}} = \boldsymbol{NN(Y(t), t, \theta)} \qquad \textbf{Equation 3-2}$$

where NN denotes the neural network, θ represents the weight and bias parameters of the neural network.

DNN have demonstrated remarkable expressive power in deep learning. These programmable neural networks, with structured architectures and an increased number of hidden layers, performs more accurately in predication tasks compared to traditional machine learning (Wu et al. 2021; Vanrolleghem and Lee 2003). Their exceptional approximation capability is formalised by the *universal approximation theorem* (Kurt, Maxwell and Halbert, 1989; Pinkus, 1999; Elbrächter *et al.*, 2019). This theorem establishes that, under specific conditions, DNN with a single hidden layer can approximate any continuous function to an arbitrary degree of accuracy, given sufficient neurons in that layer. While practical limitations exist due to data availability, quality, and computational constraints, advancements in neural networks advancements continue to enable innovative applications like NODE (Chen *et al.*, 2018; Dupont, Doucet and Teh, 2019; Harry and Howe, 2021), which leverage DNN's function approximation capabilities to tackle complex differential equations.

As a data-driven approach, NODEs offer new opportunities to reveal the underlying function *f* based on observed data. The core distinction between mechanistic ODE and neural ODE lies in how they define the function f. Mechanistic models rely on manually crafted formulas requiring significant domain expertise and tremendous technical efforts. In contrast, NODE models discover this relationship from observed

data $Y(t)$ through a trainable neural network. This approach reduces the dependency on explicit expert-derived formulas, allowing the model to adapt to real-world data.

3.2 Training of NODE

For many dynamical systems, only time-series data of $Y(t)$ are typically measured (Wu *et al.*, 2021). Consequently, training of the neural network is designed to be accomplished through integration of the ODEs. For example, in wastewater processes, the state of wastewater can be monitored by measuring the concentrations of fractioned components at various time points, such as active heterotrophic biomass (X_{BH}), readily biodegradable substrate (S_S), chemical oxygen demand (COD) and ammonia. However, direct measurements of component reaction rates are often unavailable. As a result, NODEs training is designed to utilize only component concentration data. This process involves integrating ODEs from an initial condition $Y(t_0)$, e.g., at the time when influent enters the bioreactor, to a final condition $Y(t_n)$, e.g., at the end of hydraulic retention time (HRT) when the flow leaves the reactor as effluent.

$$\boldsymbol{Y(t_n) = Y(t_0) + \int_{t_0}^{t_n} NN(Y(t))dt} \\ \boldsymbol{= ODESolver(NN, Y(t_0), (t_0, t_n), \theta)} \qquad \textbf{Equation 3-3}$$

Where (t_0, t_n) denotes the range of integration, $Y(t_0)$ represents initial condition, $Y(t_n)$ denotes the states at time t_n.

In practice, the solution is often obtained through numerical methods, with ODE solvers playing a vital role in solving dynamical systems. Thanks to mathematical advancements, various types of ODE solvers are now available (Postawa, Szczygieł and Kułażyński, 2020). However, issues such as stiffness may still be encountered, which will be discussed in subsequent chapters (Kushnir and Rokhlin, 2012).

In machine learning, training a neural network involves iterative update of its weights and bias by reducing the discrepancy between the model's prediction and true data. This process is known as backpropagation (Amari, 1993).

A loss function, also named cost function or error function, is often defined to quantify a model's performance on a given dataset (LeCun, Bengio and Hinton, 2015). It serves

as a guiding metric in the learning process, directing the model toward improved performance through iterative training. In the context of NODEs, the loss function $L(\theta)$ can be defined similarly to conventional machine learning approaches. For example, one common choice is the mean absolute error (MAE):

$$\boldsymbol{L(\theta) = \frac{1}{n}\sum_{i=1}^{n}\left|Y(t_i)^{NN} - Y(t_i)^{observation}\right|}$$
$$\boldsymbol{= \frac{1}{n}\sum_{i=1}^{n}\left|ODESolver(NN, Y(t_0), (t_0, t_i), \theta) - Y(t_i)^{observation}\right|} \quad \textbf{Equation 3-4}$$

Building upon the loss function, gradient descent algorithms can then be employed to iteratively update and optimise the weights and biases of the neural network until achieving the desired accuracy (Kingma and Ba, 2014; Finlay *et al.*, 2020). The primary technical challenge in training NODEs lies in backpropagation within gradient descent algorithms. Gradient calculation can be performed via forward and reverse mode differentiation. While forward pass is straightforward, it often incurs high memory costs and introduces additional numerical errors (Chen *et al.*, 2018). As a remedy, the *adjoint sensitivity* method is thus employed by introducing the hidden states $h(t)$ and the adjoint $a(t)$, which is the gradient of the loss with respect to the hidden state:

$$\boldsymbol{a(t) = \frac{\partial L}{\partial h(t)}} \quad \textbf{Equation 3-5}$$

The weights and biases θ of the neural network can then be optimised by computing the gradient through the integration of $h(t)$ and $a(t)$:

$$\boldsymbol{\frac{dL}{d\theta} = \int_{t_n}^{t_0} a(t)^T \frac{\partial NN(h(t), t, \theta)}{\partial \theta} dt} \quad \textbf{Equation 3-6}$$

The adjoint method reduces memory cost and controls numerical error in exchange of solving a reversing ODE, rather than directly differentiating the accumulated forward pass operations (Chen *et al.*, 2018). However, solver errors can propagate and amplify into gradient flow, rending it ill-conditioned. Even minuscule errors in the forward route can result in a pronounced deviation in the reversing solution (Kim *et al.*, 2021). These issues are exacerbated when the studied system already exhibits stiffness.

By employing modern ODE solvers and gradient descent algorithms, such as the Adam (Adaptive Moment Estimation) optimizer used in the experiments (Kingma and

Ba, 2014), the neural network's weights and biases can be iteratively updated and optimized until the level of desired accuracy is attained.

3.3 Development of NODE

While the concept of using differential equations in modelling has a long history, NODE emerged as a groundbreaking concept with the publication of "Neural Ordinary Differential Equations" by Chen et al. (2018). This work, which received the best paper award at NeurIPS 2018, was hailed as a "radical new design" by MIT Technology Review.

Since its inception, NODEs have rapidly evolved into a diverse family of neural differential equations, spawning various extensions and modifications (Kidger, 2022):

- **Augmented neural ODEs** (ANODEs): Address the limited expressiveness of original NODEs by augmenting the state space with additional dimensions, incorporating latent variables into the ODE system, enabling more complex data modelling (Dupont, Doucet and Teh, 2019).
- **Neural controlled differential equations** (NCDEs): Extend NODEs to handle irregularly sampled time series data through the inclusion of control terms (Kidger et al., 2020).
- **Neural stochastic differential equations** (NSDEs): Introduce stochasticity into NODEs to model uncertainty and noise inherent in real-world systems (Jia and Benson, 2019; Tzen and Raginsky, 2019).
- **Neural partial differential equations** (NPDEs): Apply the NODE framework to partial differential equations, expanding its applicability to a wider range of problems (Ruthotto and Haber, 2020; Sun, Zhang and Schaeffer, 2020).
- **Graph neural ODEs**: Adapt NODEs for graph-structured data, enabling continuous-time graph representation learning (J. Choi *et al.*, 2022; Bergna *et al.*, 2023).
- **Physics-informed neural network** (PINN): Incorporate physical laws and domain knowledge into the NODE framework for improved accuracy and interpretability in scientific modelling (Karniadakis *et al.*, 2021; Kong *et al.*, 2022).
- **Hamiltonian neural networks**: Integrate principles from Hamiltonian mechanics to improve energy conservation in physical systems modelling (Greydanus, Dzamba and Yosinski, 2019).

- **Spatial-temporal neural ODE**: Extend NODEs to model both spatial and temporal dynamics simultaneously (Zhou *et al.*, 2021).
- **Continuous normalizing flows**: Apply NODEs to generative modelling, enabling efficient sampling and density estimation of complex probability distributions (Kobyzev, Prince and Brubaker, 2020).

These variants and extensions have significantly broadened the applicability of NODEs across various domains, including time series analysis (Verma, Heinonen and Garg, 2024), generative modelling (Garsdal, Søgaard and Sørensen, 2022), and physics-informed machine learning (Gabriel S Gusmão *et al.*, 2022; Treibert and Ehrhardt, 2022). The rapid development of NODE-based approaches demonstrates their potential to revolutionize continuous-time modelling in artificial intelligence and scientific computing.

3.4 NODE applications

In the real world, governing laws for most dynamical systems are often incompletely or imperfectly understood. While it can be argued that NODE also lacks explicit closed-form equations at the present, preventing direct symbolic representation of the underlying law (although this might be feasible in the future through methods like SINDy(sparse identification of nonlinear dynamics) (Kaheman, Kutz, and Brunton 2020), they can effectively approximate the practical results achieved by mechanistic models (Xie, Parlikad and Puri, 2019; Esteve-Yagüe and Geshkovski, 2021; Zhou *et al.*, 2021; Garsdal, Søgaard and Sørensen, 2022; Núñez *et al.*, 2023; Verma, Heinonen and Garg, 2024). This capability makes NODEs applicable to processes without established mechanistic models, as many systems exhibit analogous dynamic patterns of growth and decay, such as those observed in bacterial and substrate interactions in wastewater treatment.

Different from conventional machine learning, which merely maps the input and output with discretised time series data at fixed intervals (Zhong *et al.*, 2021), NODE is continuously defined, enabling it to provide solutions at any arbitrary time (Zou *et al.*, 2024). This feature proves particularly valuable when dealing with intermediate products such as N_2O emissions.

NODE presents distinct advantages and holds significant potential for dynamical systems. Their applications have been seen in many science and engineering fields, including:

1) **Prediction:** Flexible output and better generalisation (Dupont et al., 2019; Kidger et al., 2020).
2) **Parameter estimation:** Determination of kinetic parameter solely from operational monitoring data (Bradley and Yr, 2022; Kong *et al.*, 2022).
3) **Modelling**: Accurate predicative modelling, with enhanced robustness through integration with mechanistic model. (Quaghebeur et al., 2022)
4) **Optimal control**: Optimisation of dynamic systems (Sandoval, Petsagkourakis and del Rio-Chanona, 2022).
5) **Knowledge discovery**: Unveiling insights through PINN with certain interpretability and knowledge discovery potentials (Cuomo et al., 2022; Gusmão et al., 2022; Karniadakis et al., 2021; Xue et al., 2021).

Despite rapid advancements in NODEs and their widespread adoption across various fields, their application within the water sector remains notably limited. To date, only two published studies have explored the use of NODEs for water-related systems.

The first study, conducted by De Jaegher et al. (2020), employed a neural differential equation to model colloidal fouling in electrodialysis, a complex process that can be used in both drinking water and wastewater treatment. Their study demonstrated the model's capacity to accurately predict fouling rates even with a limited set of experimental data. Subsequent simulation study and sensitivity analysis validated the model's robustness, quantifying the relative contribution of crossflow velocity (40%), current (20%), and salt concentration (13%) to fouling rates. The study concluded that given sufficient high-quality data with a broad input range. The NODE model can capture the system dynamics and generalize well to unseen conditions. The absence of discontinuities or abrupt changes in model behaviour highlighted the suitability and robustness of NODEs for this complex domain.

The second study by Quaghebeur et al. (2022) proposed a hybrid modelling approach that integrates mechanistic physics-based models with data-driven NODE models. When applied to wastewater treatment, this hybrid model demonstrated the NODE's ability to compensate for knowledge gaps in the mechanistic model, leading to

improved predictive performance. Complementarily, the mechanistic component enhanced the model's capacity to extrapolate to unseen conditions compared to purely data-driven models. The study concluded that this hybrid approach effectively combines the strengths of both modelling paradigms while mitigating their individual limitations, albeit at the cost of reduced model interpretability.

3.5 Opportunities and Challenges

NODE offer a promising approach to modelling complex dynamic systems, such as wastewater treatment processes. By representing the system as a continuous-time differential equation, NODE can capture non-linear relationships between input variables, system states, and output parameters.

Their potential application in wastewater treatment can include:

- **Process modelling**: NODEs can accurately model the dynamics of biological processes in wastewater treatment, such as activated sludge, membrane bioreactors, and anaerobic digestion. This enables better understanding, prediction, and control of these systems (Garsdal, Søgaard and Sørensen, 2022; Zou *et al.*, 2024).
- **Real-time optimization and control**: By integrating sensor data and real-time predictions, NODEs can optimize operational parameters, such as aeration rate, sludge retention time, and chemical dosage, to improve treatment efficiency, reduce energy consumption and mitigate GHG emissions (Sandoval, Petsagkourakis and del Rio-Chanona, 2022).
- **Anomaly detection**: Through continuous monitoring of system behaviour, NODEs can identify abnormal patterns that indicate potential issues, such as equipment failures or process upsets, allowing for early intervention and prevention of treatment failures (Mamandipoor *et al.*, 2020; Salles *et al.*, 2022).
- **Knowledge discovery**: By learning the underlying dynamics of the wastewater treatment process, NODEs can be used to estimate the kinetic parameters or discover knowledge that may still be hidden from professionals (Champion *et al.*, 2019; Fasel *et al.*, 2022).

While NODEs hold great promise, several challenges remain:

- **Computational efficiency**: Training and inference can be computationally expensive due to the need for ODE solvers, particularly for large-scale systems (Golovanev and Hvatov, 2022).
- **Stability**: Ensuring the stability of the ODE system during training is crucial, especially for stiff dynamical systems common in wastewater treatment (Tuor *et al.*, 2020).
- **Interpretability**: Fully unlocking the potential of NODEs for interpretability requires further research to bridge the gap between mathematical representations and real-world phenomena (Golovanev and Hvatov, 2022).
- **Data quality and availability**: The performance of NODEs heavily relies on the quality and quantity of available data, which can be challenging to obtain in some wastewater treatment contexts (Hansen et al. 2024; Bahramian et al. 2023; Newhart et al. 2019).

Despite these challenges, the field of NODEs is rapidly evolving. Ongoing research is dedicated to addressing these limitations and exploring novel applications across various fields such as physics (Lee and Parish, 2021; M. Choi *et al.*, 2022; Zakwan *et al.*, 2023; Kircher, Döppel and Votsmeier, 2024), biology (Giampiccolo *et al.*, 2024; Hossain *et al.*, 2024; Xiang *et al.*, 2024), and finance (Yang *et al.*, 2023). Advances in computational methods, such as adaptive ODE solvers and parallelization techniques (Haque et al. 2023; Bosch et al. 2024), are continually improving the efficiency of NODE implementations.

In summary, NODEs represent a powerful and flexible framework for deep learning, offering new opportunities for modelling complex dynamics and systems. The development of NODE variants and extensions is progressively expanding the potential of this approach, while addressing current challenges will be key to its widespread adoption and integration into practical applications.

Chapter 4 Experimental simulation platform

A self-built simulation platform was developed to conduct experiments and facilitate data-driven modelling of N_2O in wastewater treatment by means of NODEs. This approach was adopted due to the following considerations:

- **Novelty and complexity**: Due to the novelty of NODEs and the complexity of wastewater treatment processes, preliminary experimentation with simulated data was deemed essential before transitioning to real data or real-world applications.
- **Experimental flexibility**: The ability to generate and manipulate diverse datasets is crucial for comprehensive experimentation. Simulated data provide precise control and ground truth validation, which are often challenging with real-world data.
- **Integration requirements**: Embedding NODEs neural network within an ODE modelling framework is essential for this research. This integration is not feasible with commercial simulation software due to lack of source code access.
- **Data availability**: The absence of suitable real-world datasets necessitated the creation of a simulated environment for data generation.

MATLAB (The MathWorks Inc, 2024) served as the programming language for the experiments. Two simulation approaches were employed using MATLAB:

- **Simulink simulation**: Used for generating simulated data intuitively.
- **MATLAB code simulation**: Implemented for Simulink results verification, model validation, and NODE integration due to Simulink's limitations in this regard.

Simulink simulations offer a visual representation closely aligned with process flow, enhancing comprehension. Nevertheless, constructing accurate models, particularly for complex systems, can be challenging due to algebraic loop issues and the complexities of S-function development (Chaturvedi, 2017).

MATLAB language simulation, while requiring programming proficiency, provide greater flexibility and precision. Their higher-level language nature can hinder accessibility for environmental science scholars without coding experience (David, Vasel and Wouwer, 2009).

Given the ease of integrating NODE into modules, MATLAB code simulation is the primary focus of these experiments. Simulink simulations serve as a comparative reference for validation purposes.

The ASMG1 model was selected as the primary mechanistic model for data generation. However, ASM1 would be used initially before full ASMG1 implementation due to its simplicity and available references for code verification and debugging.

The BSM1 plant model was chosen as a fully configured virtual wastewater treatment plant for this study, given its status as a well-established benchmark in wastewater treatment research.

4.1 N_2O mechanistic model used

The ASMG1 mechanistic model was employed to generate observed data including nitrous oxide production. The components of ASMG1 model were then adopted to characterise the NODE model throughout the whole research.

The ASMG1 model extends existing activated sludge models by incorporating GHG emissions, specifically N_2O. It was developed and incrementally refined through a series of advancement by various researchers:

- Hiatt and Grady (2008): formulated ASMN model, introduced N_2O with four-step denitrification for carbon oxidation and nutrient removal and highlighted the role of free ammonia and FNA as true substrates and inhibitors.
- Mampaey et al. (2013): Added an AOB denitrification pathway.
- Guo and Vanrolleghem (2014): Modified DO kinetic terms and validated the model's performance.
- Flores-Alsina et al. (2014): Integrated the model into BSM2 (Corominas *et al.*, 2012), and evaluated its performance in full-scale plants from the viewpoint on balance of effluent quality, economic cost and greenhouse gas emissions.

ASMG1 comprehensively integrates N_2O production and consumption by mixed cultures of AOB and heterotrophic de-nitrifiers. This model is temperature and pH dependant, highlighting the role of free ammonia and free nitrous acid in N_2O generation. Additionally, it refines the nitrifier-denitrification pathway by incorporating Haldane kinetics to represent the influence of DO on N_2O production, a key point of mostly discussed (Guo & Vanrolleghem, 2014). This allows the model to capture

seasonal variations in N_2O emissions. Notably, the ASMG1 maintains simplicity, excluding difficult-to-measure components like hydroxylamine, and comprises only 18 components and 15 reactions.

The ASMG1 model was further improved by the author in two aspects, based on the latest version from the publications (Flores-Alsina *et al.*, 2014; Guo and Vanrolleghem, 2014):

1) **Rationalised Coefficients**: Replaced all decimals with fractions (e.g., 8/7 instead of 1.142857) for better continuity checks.
2) **Corrected misprint or typo**: In equation of "*process rate 7*", the concentration of NOB should multiply right elements (last right bracket (")") should be at the end), according to the appendix of the paper titled "*Comparison of different modelling approaches to better evaluate greenhouse gas emissions from whole wastewater treatment plants*" (Corominas *et al.*, 2012).
3) **Corrected error**: In equation for alkalinity component concentration S_{ALK}, the coefficient of the last adding element for "*process rate 10*": ($-i_{XB}/14$) should be included. It should not stand alone among summary elements.

4.1.1 Gujer matrix and model equations

The details of the modified Gujer matrix and equations for ASMG1 are provided as follows.

1) The intermediate symbols were defined firstly for brevity, shown in Table 4.1

Table 4.1 The intermediate symbols

Symbol	Value
A	= (-24+3·(-16)+8)/14+(-24+2·(-16)+8)/14=64/14-48/14=8/7
B	= 16/7
C	= (-24-16)/14 = -40/14 = -20/7
D	= (-24+2·(-16)+8)/14 = -48/14 = -24/7
E	= $Y_H \cdot \eta_Y$
F	= (1-$Y_H \cdot \eta Y$) / ($Y_H \cdot \eta_Y$)
G	= $Y_{A1} \cdot \eta_{Y\_AOB}$
J	= - i_{XB} / 14
W	= F·7 / 4

2) The number for each process reaction the ASMG1 model is described Table 4.2

Table 4.2 Number of processes

No	Process
r1	Aerobic growth of heterotrophs
r2	Anoxic growth of heterotrophs, reducing NO_3^- to NO_2^-
r3	Anoxic growth of heterotrophs, reducing NO_2^- to NO
r4	Anoxic growth of heterotrophs, reducing NO to N_2O
r5	Anoxic growth of heterotrophs, reducing N_2O to N_2
r6	Aerobic growth of AOB autotrophs
r7	Aerobic growth of NOB autotrophs
r8	Decay of heterotrophs
r9	Decay of AOB autotrophs
r10	Decay of NOB autotrophs
r11	Ammonification of soluble organic nitrogen
r12	Hydrolysis of particulate organics
r13	Hydrolysis of particulate organic nitrogen
r14	Aerobic AOB nitrifier denitrification, reducing NO_2^- to NO
r15	Aerobic AOB nitrifier denitrification, reducing NO to N_2O

3) The process rate equations were defined in the Table 4.3

Table 4.3 Process rate equations

No	**Process Rate ρ**
r1	$\mu_H \cdot S_S/(K_{S1}+S_S)\cdot S_O/(K_{OH1}+S_O)\cdot X_{BH}$
r2	$\mu_H \cdot \eta_{g2} \cdot S_S/(K_{S2}+S_S)\cdot S_{NO3}/(K_{NO3}+S_{NO3})\cdot K_{OH2}/(K_{OH2}+S_O)\cdot X_{BH}$
r3	$\mu_H \cdot \eta_{g3} \cdot S_S/(K_{S3}+S_S)\cdot S_{NO2}/(K_{NO2}+S_{NO2})\cdot K_{OH3}/(K_{OH3}+S_O))\cdot K_{I3NO}/(K_{I3NO}+S_{NO})\cdot X_{BH}$
r4	$\mu_H \cdot \eta_{g4} \cdot S_S/(K_{S4}+S_S)\cdot S_{NO}/(K_{NO}+S_{NO}+S_{NO}^2/K_{I4NO})\cdot K_{OH4}/(K_{OH4}+S_O)\cdot X_{BH}$
r5	$\mu_H \cdot \eta_{g5} \cdot S_S\cdot/(K_{S5}+S_S)\cdot S_{N2O}/(K_{N2O}+S_{N2O})\cdot K_{OH5}/(K_{OH5}+S_O)\cdot K_{I5NO}/(K_{I5NO}+S_{NO})\cdot X_{BH}$
r6	$\mu_{AOB} \cdot S_{FA}/(K_{FA}+S_{FA}+S_{FA}^2/K_{I9FA})\cdot S_O/(K_{O\_AOB}+S_O)\cdot K_{I9FNA}/(K_{I9FNA}+S_{FNA})\cdot X_{AOB}$
r7	$\mu_{NOB} \cdot S_{FNA}/(K_{FNA}+S_{FNA}+S_{FNA}^2/K_{I10FNA})\cdot S_O/(K_{O\_NOB}+S_O)\cdot K_{I10FA}/(K_{I10FA}+S_{FA})\cdot X_{NOB}$
r8	$b_H \cdot X_{BH}$
r9	$b_{AOB} \cdot X_{AOB}$
r10	$b_{NOB} \cdot X_{NOB}$
r11	$k_a \cdot S_{ND} \cdot X_{BH}$
r12	$k_h \cdot (X_S/X_{BH})/(K_X+X_S/X_{BH})\cdot(S_O/(K_{OH}+S_O)+\eta_h \cdot K_{OH}/(K_{OH}+S_O)\cdot S_{NOX}/(K_{NO3}+S_{NOX}))\cdot X_{BH}$
r13	$r12 \cdot X_{ND}/X_S$
r14	$\mu_{AOB} \cdot \eta_{AOB} \cdot S_O/(K_{SO\_AOBden1}+(1-2*(K_{SO\_AOBden1}/K_{IO\_AOBden1})^{0.5})\cdot S_O+S_O^2/K_{IO\_AOBden1})\cdot S_{FA}/(K_{FA\_AOB}+S_{FA})\cdot S_{FNA}/(K_{FNA\_AOB}+S_{FNA})\cdot X_{AOB}$
r15	$\mu_{AOB} \cdot \eta_{AOB} \cdot S_O/(K_{SO\_AOBden2}+(1-2*(K_{SO\_AOBden2}/K_{IO\_AOBden2})^{0.5})\cdot S_O+S_O^2/K_{IO\_AOBden2})\cdot S_{FA}/(K_{FA\_AOB}+S_{FA})\cdot S_{NO}/(K_{NO\_AOB}+S_{NO})\cdot X_{AOB}$

4) The stoichiometric matrix was defined in the Table 4.4

Table 4.4 Stoichiometric matrix of ASMG1 mode

No	S_I	S_S	X_I	X_S	X_{BH}	X_{AOB}	X_P	S_O	S_{NO3}	S_{NH}	S_{ND}	X_{ND}	S_{ALK}	S_{NO2}	S_{NO}	S_{N2O}	S_{N2}	X_{NOB}
r1		$-1/Y_H$			1			$1-1/Y_H$		$-i_{XB}$			J					
r2		-1/E			1				-F/A	$-i_{XB}$			J	F/A				
r3		-1/E			1					$-i_{XB}$			J+F/8	-W	W			
r4		-1/E			1					$-i_{XB}$			J		-W	W		
r5		-1/E			1					$-i_{XB}$			J			-W	W	
r6						1		$1-D/Y_{A1}$		$-i_{XB}-(1/Y_{A1})$			$J-1/(7\cdot Y_{A1})$	$1/Y_{A1}$				
r7								$1-A/Y_{A2}$	$1/Y_{A2}$	$-i_{XB}$			J	$-1/Y_{A2}$				1
r8				$1-f_P$	-1		f_P					$i_{XB}-f_P\cdot i_{XP}$						
r9				$1-f_P$		-1	f_P					$i_{XB}-f_P\cdot i_{XP}$						
r10				$1-f_P$			f_P					$i_{XB}-f_P\cdot i_{XP}$						-1
r11										1	-1		1/14					
r12		1		-1														
r13											1	-1						
r14						1		1-B/G		$-1/G-i_{XB}$			J	-1/G	2/G			
r15						1		1-B/G		$-1/G-i_{XB}$			J-1/(7·G)	1/G	-2/G	2/G		

4.1.2 Calculation of gaseous N_2O flux

The flux of gaseous N_2O emissions for a reactor is calculated using the Henry's law and the striping equation (Foley et al., 2010), based on the liquid N_2O concentration and aeration rate:

$$Flux_{N_2O} = -K_L a_{N_2O} \cdot \left(S_{N_2O}sat - S_{N_2O}\right) \cdot V_{reactor} \quad \textbf{Equation 4-1}$$

K_La_{N2O} is calculated using equations:

$$K_L a_{N_2O} = K_L a_{O_2} \frac{\sqrt{D_{N_2O}}}{\sqrt{D_{O_2}}} \quad \textbf{Equation 4-2}$$

and corrected for the current temperature from reference temperature, using standard factor θ:

$$K_L a_T = K_L a_{Temp\_ref} \cdot \theta^{T-Temp\_ref} \quad \textbf{Equation 4-3}$$

Where,

$Flux_{N2O}$:	the flux of gaseous N_2O,	(g/day)
K_La_{O2}:	the oxygen volumetric mass transfer coefficient,	(/day)
$S_{N2O}sat$:	the saturation concentration of N_2O in water	(g/m^3)
S_{N2O}:	the concentration of dissolved N_2O in water	(g/m^3)
$V_{reactor}$:	the volume of the reactor	(m^3)
K_La_{O2}:	the oxygen transfer coefficient	(/day)
D_{O2}:	the diffusion coefficient of oxygen in water	(m^2/s)
D_{N2O}:	the diffusion coefficient of nitrous oxide in water	(m^2/s)
T:	the current temperature	(K)
Temp_ref:	the reference temperature	(K)

4.1.3 Stoichiometric and kinetic parameters

Stoichiometric parameters as shown in Table 4.4 describe the mass balance of various components during biological processes. They essentially tell us how much of one substance is consumed or produced for a given amount of another substance.

Stoichiometric parameters ensure that the mass of elements (like carbon, nitrogen, phosphorus) is conserved throughout the process, determine the yield coefficients of different microbial groups (e.g., heterotrophs, autotrophs) in converting substrates into

new cells, and predict the quantity of products (e.g., nitrate, nitrite, solids) generated during the process (Hauduc *et al.*, 2013).

Kinetic parameters describe the rate at which biological reactions occur. They quantify the speed of microbial growth and decay, substrate utilization, and product formation. Example of kinetic parameters include maximum specific growth rate (μ_{max}), half-saturation coefficient (K_s), decay rate coefficient (b_d), inhibition coefficients for toxic substances. These parameters are crucial for understanding the dynamic behaviour of wastewater treatment systems (Almeida, Reis and Carrondo, 1997).

In the experiments, the default values of stoichiometric and kinetic parameters for ASMG1 model from literatures (Flores-Alsina *et al.*, 2014) were adopted in the experiments. Please refer to the cited literature or attached electronic file of source code for details as this part is not the focus of this study.

4.2 BSM1 plant configuration

As shown in Figure 4.1, the BSM1 plant consists of five activated sludge reactor tanks connected in series with a total volume of 6000 m^3, followed by a secondary clarifier. The first two tanks are 1000 m^3 anoxic bioreactors with mixing devices, where denitrification reactions mainly occur, while the remaining three are 1333 m^3 aerobic bioreactors with oxygenation devices at the bottom, where carbon oxidation and nitrogen nitrification reactions mainly take place. This configuration represents a typical A/O process that combines nitrification with denitrification for efficient nutrient removal.

The secondary clarifier consists of 10 uniform layers, each 0.4m high, with a total volume of 6000 m^3. Assumed as a non-reactive unit, there are no chemical reactions inside, and it is mainly used for the precipitation of particulate components and sludge recirculation operations. The inflow enters from the 6_{th} layer, the 10_{th} layer (top layer) discharges the treated and settled wastewater, and the 1_{st} layer (bottom layer) enables sludge recirculation and waste sludge discharge. Takacs' model with double exponentially settling velocity method (Takács, Patry and Nolasco, 1991) was adopted for describing the secondary clarifier.

One key modification was made to the original components of the BSM1 plant: the ASM1 (activated sludge model no. 1) was replaced with the ASMG1 model for

bioprocess reactions. This modification allows for comprehensive modelling of greenhouse gas production in the wastewater treatment process.

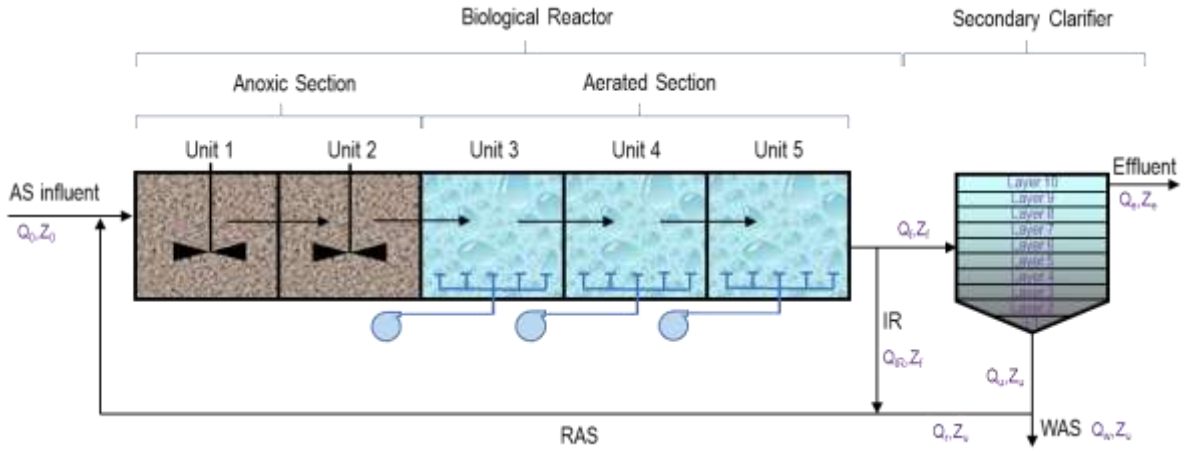


Figure 4.1 Overview of the BSM1 plant. Adapted from(Alex *et al.*, 2018)

The operational parameter settings were adopted from the open-loop control described in the BSM1 technical report (Alex *et al.*, 2018). The recycling activated sludge flowrate was set to match the influent flowrate, while the internal recirculation flowrate was set at three times the influent flowrate. Waste sludge was discharged at a constant rate of 385 m^3/d. Minimal aeration was provided for the first two anoxic tanks, with a K_{La} of 2/d, while the aeration rate for the remaining three oxic tanks was set at a constant value of 240 /d. No external carbon source was added to the tanks. The control parameter settings are summarized in Table 4.5.

Table 4.5 Experimental control settings

No	Controls	Settings
1	RAS flowrate	1 time of influent flowrate
2	IR flowrate	3 times of influent flowrate
3	WAS flowrate	385 (m^3/d)
4	Aeration Kla for unit 1 ~ 5	2, 2, 240, 240, 240 (/d) respectively
5	External carbon addition	none

For a more detailed description of the plant configuration, including specific parameter values and operational ranges, please refer to the official BSM1 report (Alex *et al.*, 2018) and ASMG1 documents (Flores-Alsina et al., 2014; Guo & Vanrolleghem, 2014).

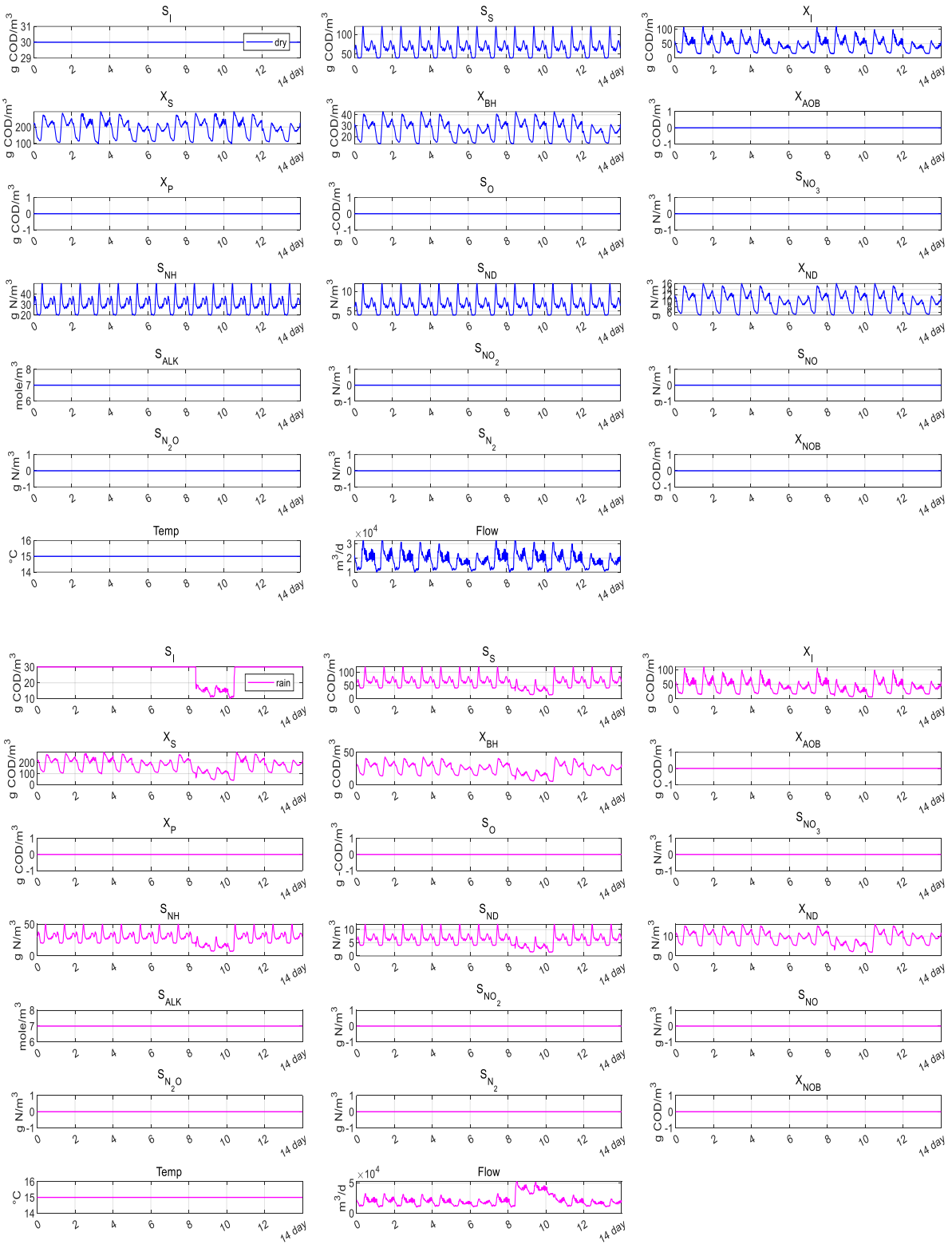
These resources provide in-depth information on the plant structure, process equations, and simulation protocols that form the foundation of this experimental setup.

4.3 Influent data

The BSM1 model provides three types of influent datasets to represent distinct weather conditions, which significantly influence wastewater characteristics and flow rates.

- **Dry weather scenario**: The dataset encompasses two weeks of dynamic dry weather influent data, characterizing relatively smooth flow patterns and stable microbial populations. This simulates routine and normal operations in a wastewater treatment plant.
- **Rain weather scenario**: The dataset presents one week of dynamic dry weather conditions followed by a prolonged rain event in the subsequent week. The rain event results in a rapid increase in flow due to surface runoff and a concomitant decrease in pollutant concentrations. This scenario may induce increased hydraulic loading on treatment units and washout effects on settled solids.
- **Storm weather scenario**: The dataset combines one week of dry weather data with two superimposed storm events. This reflects abrupt flow rate surges and significantly reduced pollutant concentrations due to excessive dilution. Unlike the rain weather scenario, these storm events are characterized by short duration and rapid recovery to normal conditions.

To align with the ASMG1 model's requirements, this study extended the well-established BSM1 influent dataset to include six additional components: nitrite (S_{NO2}), nitric oxide (S_{NO}), nitrous oxide (S_{N2O}), nitrogen (S_{N2}), NOB (X_{NOB}), and Temperature. Zeros were filled for five added components S_{NO2}, S_{NO}, S_{N2O}, S_{N2}, X_{NOB}. The temperature is set at 15°c, and pH is at 7. The extended influent data in dry, rain, and storm weather scenarios for ASMG1 based BSM1 model are visualised in Figure 4.2.


S_I
S_S
X_I
X_S
X_BH
X_AOB
X_P
S_O
S_NO3
S_NH
S_ND
X_ND
S_ALK
S_NO2
S_NO
S_N2O
S_N2
X_NOB
Temp
Flow
dry
rain
g COD/m3
g N/m3
g -COD/m3
mole/m3
°C
m3/d
14 day

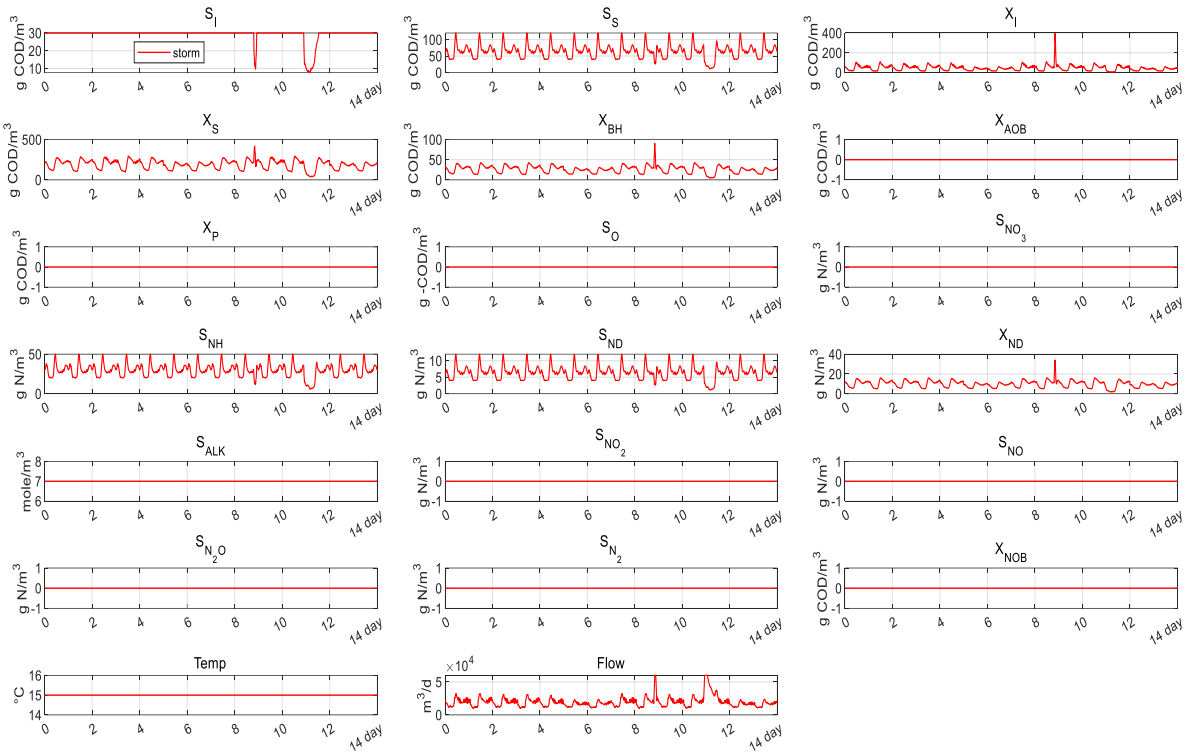


Figure 4.2 ASMG1 influent data of dry, rain and storm weather scenarios

We used extended influent data of 14 days in dry, rain, and storm weather scenarios as input, run the BSM1 plant simulation, and generated five sets of input data and trajectory state data for each reactor tank, mimicking monitored data. Sampling point were set at every 15 minutes.

Prior to initiating any scenario simulations, the system was stabilized for an initial period of 100 days, as specified in BSM1 documentation. This stabilization phase employed a constant input of average dry weather flow rate and flow-weighted average influent concentrations.

4.4 N_2O production simulation with Simulink

Simulink is a graphical programming environment employed for modelling, simulating, and analysing dynamic systems. As part of the MATLAB suite, it provides a visual approach to system design. Simulink models are constructed using interconnected blocks that represent system components. These blocks can encapsulate mathematical operations, physical systems, or other functions. Information is transmitted between blocks via signals. Simulink leverages numerical solvers to integrate differential equations and compute the system's behaviour over time.

Running the model calculates the system's response to inputs within a specified timeframe (Chaturvedi, 2017).

Given the complexity of wastewater process models like ASM1 or ASMG1, custom-built blocks utilizing S-functions are often necessary. Typically programmed in C or C++, S-functions offer custom algorithm, enhanced performance and flexibility for intricate models or real-time applications. The original BSM1 model already incorporated S-functions for the ASM1 process, albeit with simplifications of real-world plant conditions (Alex *et al.*, 2018). Nevertheless, the generated reference results demonstrated reasonable alignment with actual scenarios, establishing it as a suitable benchmark for wastewater process simulation. Commercial simulation software has adopted similar principles and frameworks, while potentially incorporating more complex process reactions.

By creating core blocks with embedded S-functions for individual treatment units and connecting them according to the plant's flow layout, a complete Simulink simulation of the BSM1 plant can be constructed. Once the solver and environmental parameters are defined, the model can be executed to generate results.

The overall simulation diagram of BSM1, incorporating ASMG1 model is presented in Figure 4.3. The modular diagram for sub-system bioreactor, secondary clarifier and hydraulic delay unit are presented in Figure 4.4, Figure 4.5, and Figure 4.6 respectively. Hydraulic delay unit is designed to avoid algebraic loop issue.

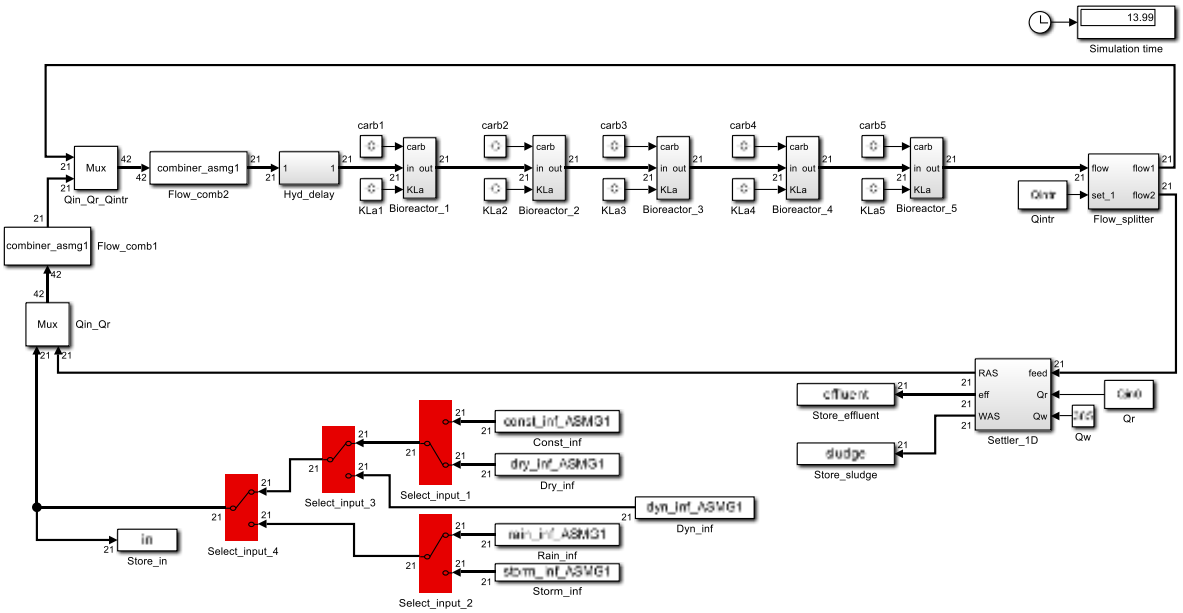


Figure 4.3 Simulation diagram of ASMG1 based BSM1 plant

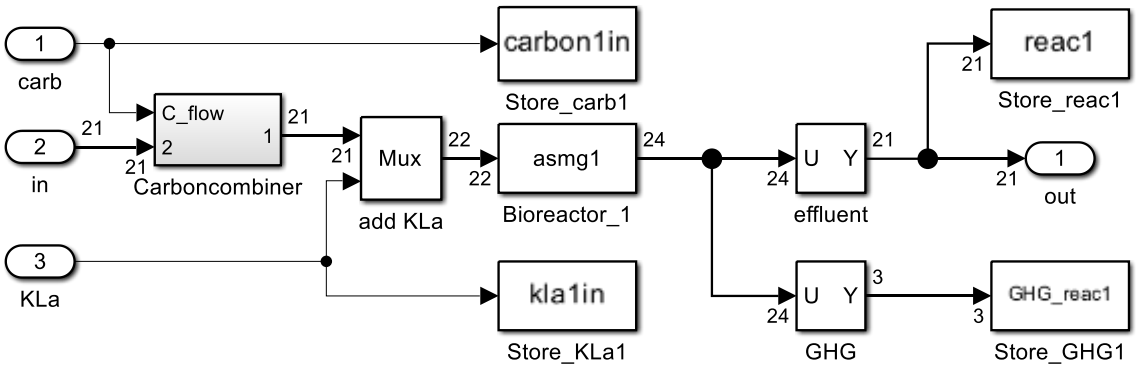


Figure 4.4 Modular diagram of bioreactor

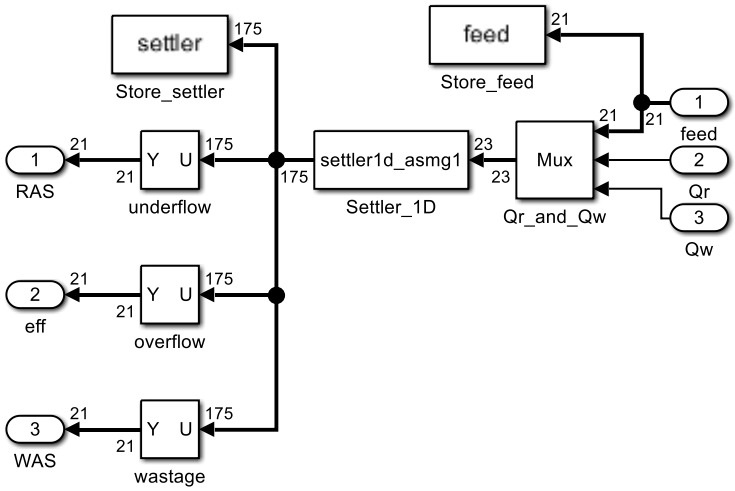


Figure 4.5 Modular diagram of secondary clarifier

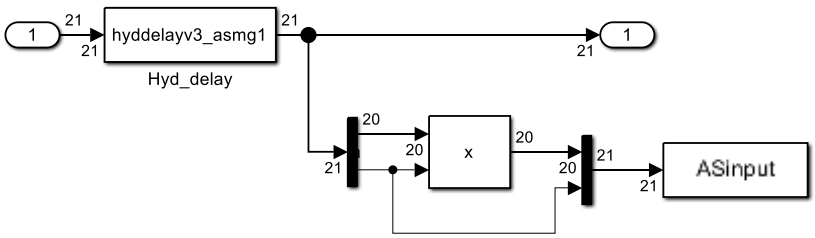


Figure 4.6 Modular diagram of hydraulic delay unit

S-functions were developed for the ASMG1 model and integrated into the BSM1 framework, replacing the ASM1 model S-functions. These new S-functions were embedded within the combiner and bioreactor blocks. The second clarifier component was modified to accommodate ASMG1 components. The code underwent rigorous debugging and was validated against wastewater theory and established treatment practices by comparison with a MATLAB code version.

4.5 N_2O production simulation with MATLAB code

The implementation of N_2O production simulation in the BSM1 plant using MATLAB language follows common wastewater modelling practices (David, Vasel and Wouwer, 2009; Chaturvedi, 2017). The simulation framework comprises three key modules: the ASMG1 module, the BSM1 module, and the plant performance module. These modules serve as the core components and function as the backbone for the simulations.

4.5.1 Key modules

4.5.1.1 ASMG1 module

The ASMG1 module constitutes the core of the model, representing essential bioprocess reactions based on given differential equations. The steps to structure the module are as follows:

1) Specification of basic kinetic parameters.
2) Formulation of temperature- and pH-corrected kinetic parameters.
3) Calculation of free ammonia, free nitrous acid, and total oxidized nitrogen concentrations, which impact N_2O production, and are integrated into the process rate equations.
4) Definition of 15 process rates as shown in Table 4.3, including:
   - Aerobic growth and decay of HDB, AOB and NOB.
   - Four-step heterotrophic denitrification reactions
   - Two-step AOB nitrification reactions
   - Soluble and particulate nitrogen decomposition
5) Calculation of reactions using equations from the stoichiometric matrix, with predefined stoichiometric parameters.

4.5.1.2 BSM1 module

BSM1, a benchmark model for activated sludge wastewater treatment, can be efficiently implemented in MATLAB (The MathWorks Inc., 2024) due to its robust numerical and computational strengths. Additionally, its matrix operations facilitate the representation of complex reactor configurations and interconnections within the BSM1 model. At its core, BSM1 is a system of differential equations describing the dynamics of various components within the activated sludge process.

The mathematical framework of the BSM1 module is established through the following steps:

1) Interpolation of influent flowrate and concentration at the given time ($\boldsymbol{t}$).
2) Calculation of temperature-dependent parameters, particularly for gaseous components NO, N_2O and N_2.
3) Computation of inflow concentrations to activated sludge by mixing influent, RAS, and IR flows. A hydraulic delay unit is added to avoid algebraic loop issue.
4) Core calculations involve evolution of components in five bioreactor tanks by calling the ASMG1 module and determination of gaseous component emissions using the method described in section 4.1.2.
5) Simulation of settling in the secondary clarifier for overflow of the top clean layer as effluent and downflow of the bottom layer for recycling or discharge of the concentrated activated sludge.

Once established, the BSM1 module can simulate the complete evolution trajectory for various influent inputs using numerical solvers. Both steady-state and dynamic simulations are feasible by giving different influent data and time periods. Among various options that MATLAB offered for ODE solvers, the following two are adopted for the experiments:

- *Ode45*, based on the explicit Runge-Kutta (4,5) formula (Dormand-Prince pair), is sufficient for constant influent input during stabilization.
- *Ode15s*, derived from the Gear method, is employed for dynamic simulations. The model's stiffness, resulting from the wastewater treatment system's inherent nature, can be addressed by this solver.

MATLAB's symbolic math toolbox was utilized to derive and manipulate the system's equations accurately, enhancing the model's overall precision and flexibility.

4.5.1.3 Plant performance assessment module

Plant performance assessment is typically conducted during the second week of operation, as the system is presumed to have stabilized after initial fluctuations. Evaluation is based on BSM1 criteria (Alex *et al.*, 2018), with additional metrics for N_2O emissions.

Key performance indicators include:

- Effluent quality index (EQI): As defined in BSM1.
- Operational cost index (OCI): Calculated as the sum of daily pump, aeration, and mixing energy costs, plus five times the daily sludge cost and three times the daily carbon mass cost.
- Violations: Including time in violation, number of violations, and percentage of time in violations, as defined in BSM1.
- 95_{th} percentile effluent concentrations for ammonium, total nitrogen, and total suspended solids.
- N_2O emission rate: Calculated for each bioreactor, with summarized emissions for nitrification and denitrification stages.

4.5.2 Simulated data verification

To validate the code, the ASM1 model was initially used to compare results with the reference data in the BSM1 technical report. A 100-day stabilization with constant influent input for steady state was conducted. The largest error of 0.002 for suspended solids (S_s) in the steady state indicates a strong agreement between the MATLAB code and Simulink simulation results, confirming the reliability of the coded model. Figure 4.7 visually compares the effluent results for both simulations at steady states after stabilization (non-reactive soluble inorganic compounds S_I omitted).

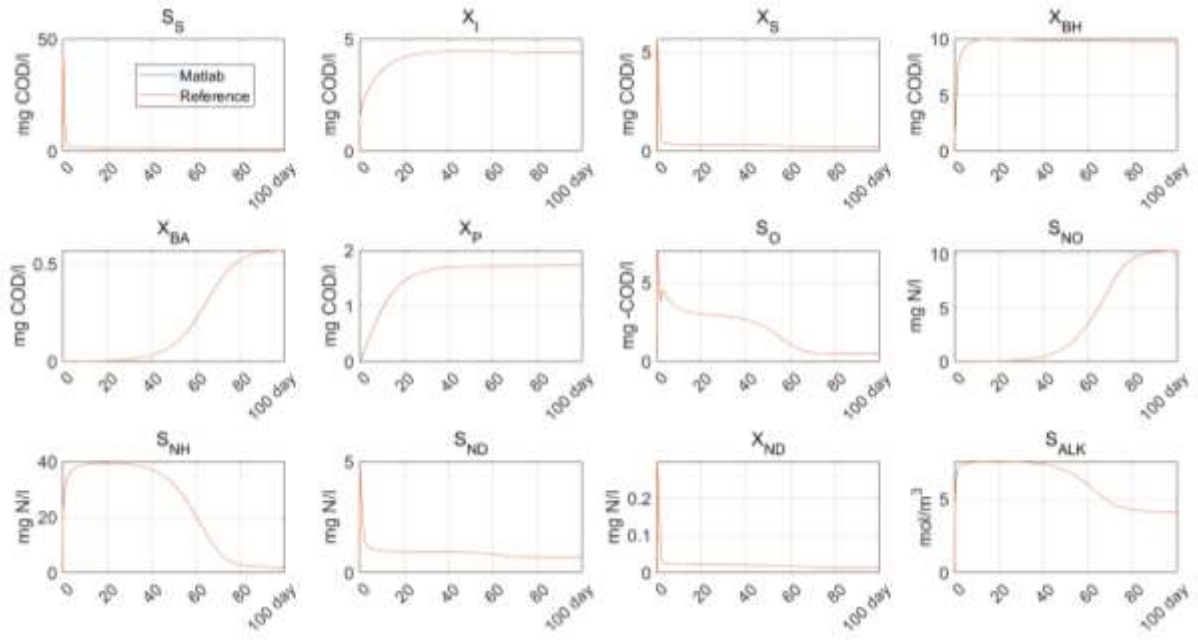


Figure 4.7 Effluent results of 100-day stabilisation obtained from MATLAB code simulation and reference

Table 4.6 presents the steady-state results of five bioreactors after 100 days of stabilization. The comparison with reference data indicates negligible errors, suggesting a high degree of accuracy in the model.

Table 4.6 Steady state results of five bioreactors after 100-day stabilisation

Components	Reactor 1		Reactor 2		Reactor 3		Reactor 4		Reactor 5	
	simulated	reference	simulated	reference	simulated	reference	simulated	reference	simulated	reference
S_I (mg/l)	30	30	30	30	30	30	30	30	30	30
S_S (mg/l)	2.8099	2.8099	1.4598	1.4598	1.1501	1.1501	0.9957	0.9957	0.8899	0.8899
X_I (mg/l)	1148.6	1148.6	1148.6	1148.6	1148.6	1148.6	1148.6	1148.6	1148.6	1148.6
X_S (mg/l)	82.151	82.151	76.4121	76.4121	64.8723	64.8724	55.7056	55.7057	49.3171	49.3171
X_{BH} (mg/l)	2551.2	2551.2	2552.8	2552.8	2556.5	2556.5	2558.6	2558.6	2558.7	2558.7
X_{BA} (mg/l)	147.5533	147.5532	147.4722	147.4721	148.1007	148.1007	148.6848	148.6848	148.9546	148.9545
X_P (mg/l)	448.1451	448.1451	448.8150	448.8150	449.7093	449.7093	450.6044	450.6044	451.4996	451.4996
S_O (mg/l)	0.0043	0.0043	6.3037e-5	6.3037e-5	1.7272	1.7272	2.4331	2.4331	0.4884	0.4884
S_{NO} (mg/l)	5.3355	5.3355	3.6288	3.6288	6.4984	6.4984	9.2540	9.2540	10.3720	10.3720
S_{NH} (mg/l)	7.9586	7.9586	8.3852	8.3852	5.5988	5.5988	3.0210	3.0210	1.7834	1.7834
S_{ND} (mg/l)	1.2166	1.2166	0.8818	0.8818	0.8290	0.8290	0.7670	0.7670	0.6884	0.6884
X_{ND} (mg/l)	5.2858	5.2858	5.0306	5.0306	4.3934	4.3934	3.8796	3.8796	3.5278	3.5278

S_{ALK} (mol/m^3)	4.9331	4.9331	5.0855	5.0855	4.6815	4.6815	4.3005	4.3005	4.1322	4.1322

Table 4.7 presents the steady-state concentrations of solids and soluble components in the secondary clarifier after 100 days of stabilization. A close agreement was observed between these results and the reference data.

Table 4.7 Steady state results in secondary clarifier after 100-day stabilisation

Layer	TSS (mg/l)		S_I (mg/l)		S_S (mg/l)		S_O (mg/l)		S_{NO} (mg/l)		S_{NH} (mg/l)		S_{ND} (mg/l)		S_{ALK} (mol/m^3)	
	sim	ref	sim	ref	sim	ref	sim	ref	sim	ref	sim	ref	sim	ref	sim	ref
10	12.4935	12.5	30	30	0.8899	0.889	0.4883	0.491	10.3709	10.4	1.7846	1.73	0.6884	0.688	4.1324	4.13
9	18.1094	18.1	30	30	0.8899	0.889	0.4883	0.491	10.3711	10.4	1.7844	1.73	0.6884	0.688	4.1324	4.13
8	29.5345	29.5	30	30	0.8899	0.889	0.4883	0.491	10.3713	10.4	1.7842	1.73	0.6884	0.688	4.1323	4.13
7	68.9607	69.0	30	30	0.8899	0.889	0.4883	0.491	10.3715	10.4	1.7839	1.73	0.6884	0.688	4.1323	4.13
6	355.921	356	30	30	0.8899	0.889	0.4883	0.491	10.3717	10.4	1.7837	1.73	0.6884	0.688	4.1323	4.13
5	355.922	356	30	30	0.8899	0.889	0.4883	0.491	10.3719	10.4	1.7835	1.73	0.6884	0.688	4.1323	4.13
4	355.921	356	30	30	0.8899	0.889	0.4883	0.491	10.3717	10.4	1.7837	1.73	0.6884	0.688	4.1323	4.13
3	355.922	356	30	30	0.8899	0.889	0.4883	0.491	10.3715	10.4	1.7840	1.73	0.6884	0.688	4.1323	4.13
2	355.921	356	30	30	0.8899	0.889	0.4883	0.491	10.3713	10.4	1.7842	1.73	0.6884	0.688	4.1323	4.13
1	6390	6394	30	30	0.8899	0.889	0.4883	0.491	10.3711	10.4	1.7844	1.73	0.6884	0.688	4.1324	4.13

**Note*: *sim: simulated; ref: reference*

Although no reference data exists for direct comparison with ASMG1 results, the outcomes were validated against established wastewater theory and treatment practices. The results were deemed to fall within reasonable ranges, rendering them suitable for subsequent experiments.

The visualisation of simulated data generated using the ASMG1 based BSM1 model under various weather conditions can be found in figures of model predications for comparison. Additionally, the appendix 8.1 includes a plant performance assessment for the various weather scenarios achieved with the model.

Chapter 5 Tackling stiffness in NODE models

Wastewater treatment processes inherently exhibit stiffness, which poses challenges for numerical simulation. This issue is exacerbated when such systems are integrated with neural networks. Stiffness has emerged as a primary obstacle to applying NODE in N_2O modelling within wastewater treatment.

5.1 Background

Despite promises in diverse scientific and engineering fields, NODEs have seen limited application in wastewater treatment. Successful NODE implementation hinges on effective training using monitoring data that represent the system's dynamics. However, challenges arose when applying NODEs to a simple wastewater model like ASM1 using the *torchdiffeq* package in Python (Chen, 2018) and following the author's methods (Chen *et al.*, 2018). Training consistently failed (Figure 5.7 and Figure 5.8). Similar difficulties were encountered when capturing the dynamics of rapidly changing components (e.g., N_2O, see Figure 5.15 and Figure 5.16) using the *dlode45* function in MATLAB (The MathWorks Inc, 2023). Stiffness has been recognised as a key culprit behind these setbacks (Kim *et al.*, 2021), and overcoming it is the main objective of this chapter.

Stiffness arises when fast and slow components in the dynamics are presented at largely separated scales. Traditional mathematical modelling approaches utilise adaptive or implicit solvers to address stiffness in ordinary differential equations (ODEs) effectively. However, when these solvers are applied to stiff NODEs trained through gradient-descent based optimisations, their efficacy diminishes. The combination of stiff ODEs and neural networks can lead to two undesirable outcomes: (1) high computational costs due to extremely small time-steps required for numerical stability, and (2) pathological gradients in the loss function, potentially hindering training convergence.

Few existing studies propose methods for mitigating stiffness in NODEs. For example, Kim et al. (2021) illustrated that proper equation and loss function scaling produced good results for two benchmarking stiff problems. However, the experiments utilising this approach for wastewater process modelling frequently encountered underflow errors during training, leading to premature termination.

As stiffness is often peculiar to the studied system, this work is motivated to analyse the underlying cause and find solutions to address the stiffness issue in data-driven wastewater process modelling with NODEs.

5.1.1 Stiffness in mechanistic ODEs

Most real-world systems of ODEs require numerical methods, as analytical solutions are often unavailable or impractical. Stiffness arises when certain numerical methods fail to provide stable solutions unless the step size to be taken is extremely small (Hairer and Wanner, 1996). This phenomenon is an intrinsic property of the differential systems and is surprisingly common in many real-life problems (Kushnir and Rokhlin, 2012). Despite its prevalence, a rigorous mathematical definition for stiffness remains elusive (Kushnir and Rokhlin, 2012).

The “stiffness ratio” is sometimes utilised to quantify system stiffness, defined as the product of the time span and the ratio of the real part of the fastest eigenvalue ($\bar{\lambda}$) and slowest eigenvalue ($\underline{\lambda}$) of the ODE system’s Jacobian:

$$\boldsymbol{Stiffness\ Ratio} = \frac{|\boldsymbol{Re}(\bar{\lambda})|}{|\boldsymbol{Re}(\underline{\lambda})|}(\boldsymbol{t_1} - \boldsymbol{t_0}) \qquad \textbf{Equation 5-1}$$

Empirically, stiff systems often exhibit significant disparities in the rate of change among various components. This disparity manifests as one component evolving slowly over time while another undergoes abrupt or swift changes, attributable to the system’s distinctive chemical or biological kinetics. Apparently, the time span plays a crucial role in the issue. For long-time simulations, the issue can become severely problematic.

Wastewater processes exemplify these disparities in scales and dynamic behaviours. They encompass components with high concentrations that undergo slow changes (e.g., heterogeneous biomass) and transient components or intermediate products with low concentrations that exhibit rapid fluctuations (e.g., dissolved oxygen, soluble substrate, hydroxylamine). Given the typical HRT of bioreactors ranging from 4 to 20 hours, the stiffness of these systems becomes evident.

Given an example from the ASM1 model, to obtain a stable solution within acceptable tolerances, it is advised (Mogens, Willi, Takashi and Mark, 2000) that when applying

numerical methods to the model implementation, the maximum time step size should be less than:

$$\Delta t < \frac{V_k C_{ki}}{O_{ki}+K_{ki}} = \theta_{ki} \qquad \textbf{Equation 5-2}$$

Where V_k is the volume of reactor compartment k. C_{ki}, O_{ki}, and K_{ki} are the concentration, output transport terms, and consumption terms of component i in reactor compartment k respectively. The term θ_{ki} is the mean residence time of component i in compartment k at steady state. With default values, θ_{ki} is of the order of ten minutes for X_{BH}, X_{BA}, X_P, X_S and X_{ND}, of one minute for S_S, S_{ND}, S_{NH}, S_{ALK} but of one second for S_o (Mogens, Willi, Takashi and Mark, 2000). The time step adopted in ODE solver typically ranges from 5-20% of the advised maximum step for a trade-off between sufficient accuracy and acceptable computational cost. If it is large than θ_{ki}, the correctness of the results cannot be guaranteed.

Figure 5.1 illustrates a continuous stirred-tank reactor (CSTR) example modelled using ASM1, evolving from an initial concentration (see Table 5.1) over 6 hours, with consistent DO control at 2 mg/l. It demonstrates the rapid evolution of various components and their first-order derivatives with rates of change ranging from approximately -0.02 mg/(l·d) to 8000 mg/(l·d). Additionally, the figure highlights the asynchronous occurrence of steep and flat segments for each component curve, indicating differing temporal dynamics.

Table 5.1 Initial condition defined in the experiment of ASM1 model

Component	S_S	X_S	X_{BH}	X_{BA}	X_P	S_O	S_{NO}	S_{NH}	S_{ND}	X_{ND}	S_{ALK}	S_{N2}
Unit	mg COD/l	mg COD/l	mg COD/l	mg COD/l	mg COD/l	mg O_2/l	mg N/l	mg N/l	mg N/l	mg N/l	eq ALK/l	mg N/l
Value	59.8	260.1	2552	148	449	2	0	23	1.8	7.8	0.007	0

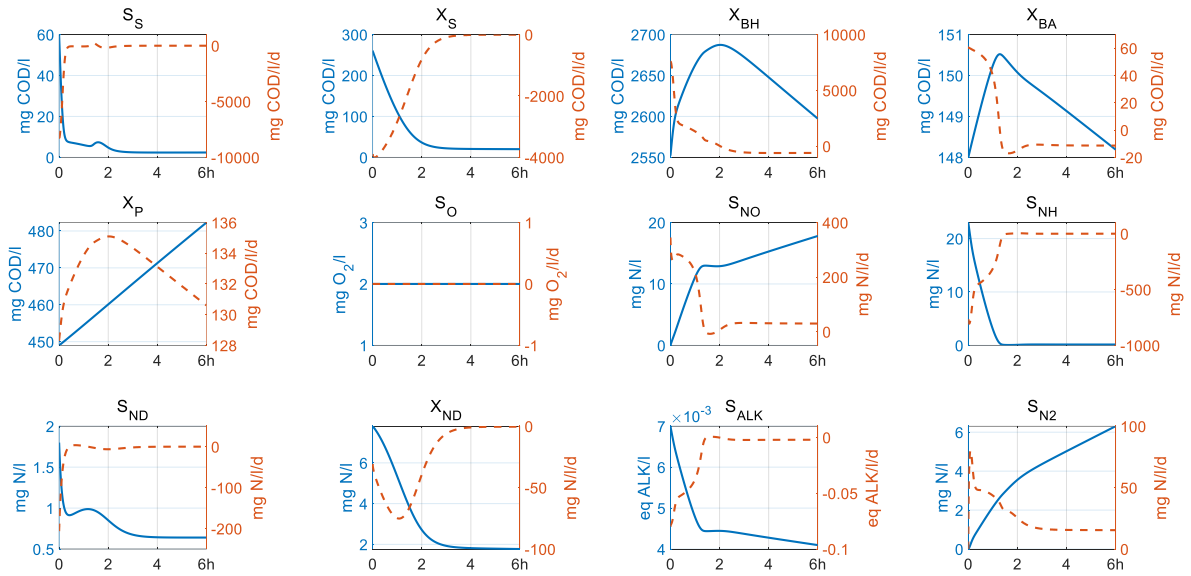


Figure 5.1 Components concentration (solid line) and their reaction rate (dot line) in a CSTR of ASM1 model with constant DO control at 2mg/l

The stiffness issue can be further intensified in more complex models, such as the ASM2d-N_2O model (Massara *et al.*, 2018), which encompasses more volatile and intermediate components and more intricate reactions. For instance, the consumption rate of fermentable substrate (S_f), can rapidly decline from 9000 mg/(l·d) to nearly zero within one minute, while oxygen uptake rate (OUR) may fluctuate around 6000 mg/(l·d). In contrast, nitric oxide (S_{NO}) evolves considerably slowly, ranging between 0 and 5 x10^{-4} mg/l throughout the entire period.

Differential algebraic equations (DAEs) present another source of stiffness. In wastewater modelling, process controls variables, such as aeration and external carbon input, are often expressed in algebraic form, making the use of DAEs inevitable. Recognized as a form of infinite stiffness (Linda, 1982; Hairer and Wanner, 1996), DAEs pose compatibility challenges for certain ODE solvers.

Stiffness remains a significant challenge in numerical analysis of differential equations (Postawa et al., 2020). While advancements in computing power and novel algorithms offer promising solutions, the effectiveness of solvers varies greatly. Modern solvers employing adaptive or implicit methods, such as MATLAB's *ode15s*, can effectively tackle stiff ODEs. However, other solvers may struggle or fail entirely when applied to different stiff systems. This highlights that there is no single "best" algorithm for all stiff problems (Kushnir & Rokhlin, 2012). Instead, solver suitability depends on the

characteristics the specific system. Consequently, researchers often resort to trial-and-error methods to identify an appropriate solution based on solver features and the studied system's behaviour.

5.1.2 Stiffness in data-driven NODEs

Despite employing a carefully chosen solver, known to be effective for the stiff mechanistic ODEs, training of the corresponding NODEs with the same solver remains a challenging task. This discrepancy arises due to the inherent differences in stiffness between NODEs and their mechanistic counterparts.

Similar to traditional DNNs, NODEs are typically initialized with random weights. These random starting points can lead to regions where the neural network - approximated ODEs exhibit vastly different rates of change – especially if not regularised. This disparity in rates can further exacerbate stiffness issues. To mitigate this issue, it is crucial to adopt schemes that reduce the variance in initial gradients, thereby lessening the likelihood of encountering extreme stiffness.

Meanwhile, during each training iteration, the solver interacts with the neural networks, which approximates a callable ODE function. However, this function is constantly evolving due to the changing network parameters (weights and biases) during optimisation. As a result, the Jacobian of the approximated ODEs experiences variations with a degree of randomness throughout training, stemming from the stochastic nature of gradient descent. Essentially, the solver tackles a different ODE with potentially disparate stiffnesses at every iteration. This randomness can lead to high variance in neural network outputs, amplifying the stiffness and creating difficulties for the solver, potentially causing instability, errors or even training divergence.

The adjoint method, commonly used for backpropagation in NODEs, requires multiple ODE solver calls in exchange for reduced memory cost. However, this approach may significantly increase the risk of encountering stiffness in the adjoint calculations, potentially leading to numerical blow-up (Kim *et al.*, 2021).

Moreover, real-world measurement data inevitably contains noise, disrupting smoothness and exacerbating derivative estimation errors. It is well-established that minor discrepancies in an ODE initial state can result in substantial divergences over

time due to the accumulation of truncation and round-off errors (Gear, 1981; Hairer and Wanner, 1996). In the realm of NODEs, data-induced noise can exhibit similar behaviour. A tiny error may propagate and amplify through subsequent steps, potentially causing model instability after a certain period. Consequently, training stiff NODEs with noisy data presents an additional challenge.

In summary, the intrinsic stiffness of the modelled system, compounded by the inherent randomness of neural networks - stemming from initial weight values and/or the stochasticity of calculated gradients during training - poses significant challenges for data-driven modelling using NODEs.

5.2 Methodologies

Stiffness challenges in NODEs are often problem-specific, demanding empirical solutions through experimentation. The crux of successfully training stiff NODEs lies in maintaining stable gradient computations and avoiding ill-conditioned gradients. After extensive exploration and experimentation, the following approaches were proposed to tackle these issues:

1) **Normalisation method**: normalisation, a well-established technique in conventional machine learning, was adapted for seamless integration within the neural network architecture of NODEs. This approach tackles the root cause of the difficulty by scaling the neural network outputs, alleviating the burden on the ODE solver.
2) **Collocation method**: This alternative approach bypasses the stiffness by employing non-ODE-solver-based collocation techniques. Through directly interpolating and regressing the derivatives at desired points, it obviates the need for costly ODE solvers, offering a faster solution.
3) **Incremental training strategy:** This practical strategy firstly employs collocation method for training, followed by direct NODEs training with normalisation based on the previously trained result. Adopting this strategy for training not only largely stabilises the learning process by addressing stiffness but also saves time and refines the results to a higher level compared to utilising the methods alone**.**

5.2.1 Normalisation method

In machine learning regression tasks, data transformation through scaling is often imperative. This is because algorithms used in the training process, such as gradient descent adopted in the NODEs training, are sensitive to feature variance (Amari, 1993). Scaling or normalisation transforms data to be dimensionless and/or have comparable distribution scales. This ensures each feature contributes equally, prevents features with higher magnitudes from dominating, speeds up the convergence of optimisation algorithms, and improves the learning process performance. Lack of data normalisation can lead to slow convergence, inaccurate models, poor generalisability, and even complete failure (Bhanja and Das, 2018; Cabello-Solorzano *et al.*, 2023).

In NODEs, the neural network maps input state variables (let's say *X*) to their time derivatives (*X'*). The proposed normalisation method for NODEs utilises a pair of normalization and de-normalization layers to wrap the deep neural network of NODE architecture (see Figure 5.2).

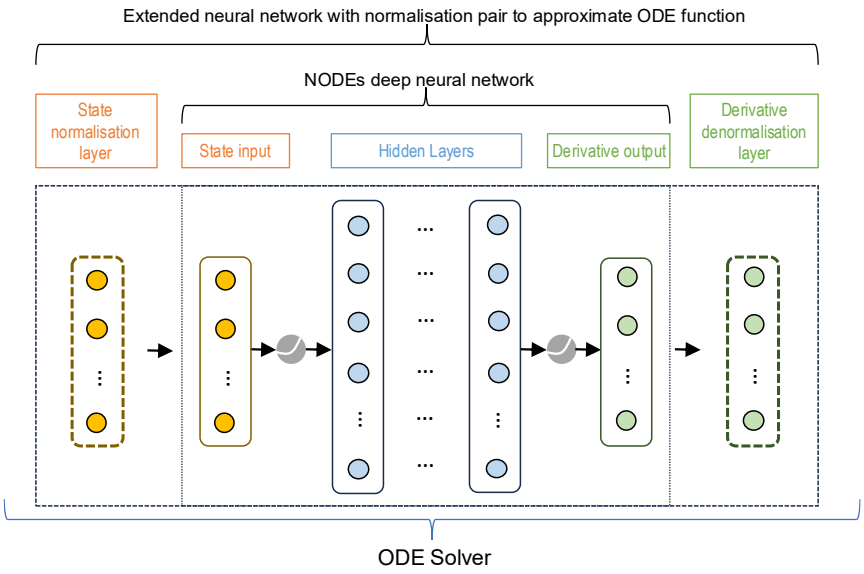


Figure 5.2 Illustration of NODEs normalisation pair layout

The normalisation layer is applied to the neural network input, using standardisation (Z-score normalisation), a common technique in machine learning (Shanker M, Hu M Y and Hung M S, 1996). This transforms the data to have a zero mean and unit standard deviation.

$$X_i^* = \frac{X_i - \mu_X}{\delta_X}$$ **Equation 5-3**

where, μ_x and δ_x denote the mean and standard deviation of the component state data sequence *X*, respectively.

The de-normalisation layer is then applied to the neural network output to restore the state derivative data to its normal range.

$$X_i' = X_i'^* * \delta_{X'} + \mu_{X'}$$ **Equation 5-4**

where, $\mu_{x'}$ and $\delta_{x'}$ denote the mean and standard deviation of the state derivative data sequence *X'*, representing biokinetic rate.

Max-Min normalization is not recommended due to its sensitivity to outliers, which may be present in estimated state derivatives. Standardization offers better robustness in such cases.

It is crucial to note that the normalisation and de-normalisation layers must be applied together within the ODE solver, wrapping the neural network. Unlike conventional machine learning, where data preprocessing occurs at the beginning, it cannot be applied outside the solver. This is because the component states and their derivatives carry physical meaning in the dynamics addressed by the solver. Scaling the data outside the solver would skew the relationships and interplay of the components, substantially distorting the problem to be solved.

From Equation 5-3 and Equation 5-4, it can be seen that four sets of mean value and standard deviation are required. The mean and standard deviation for input can be calculated straightforwardly with the monitored time-series component state data. However, the derivative data sequence *X'* are not available explicitly. To estimate the mean and standard deviation of *X'*, difference quotients can be employed to be applied to the state data sequence *X*, such as the single-sided difference method.

$$X' = (x_2 - x_1,\ x_3 - x_2, \ldots, x_n - x_{n-1}) \cdot / \Delta t$$ **Equation 5-5**

Or central difference:

$$X' = (x_2 - x_1,\ (x_3 - x_2)/2, \ldots, (x_n - x_{n-2})/2, (\ x_n - x_{n-1}))/\Delta t$$

Equation 5-6

Experiments show that single-sided and central differences yield similar results (see Table 5.2). Although the approximation tends to be more accurate with densely sampled state datasets, a relatively small number of discretisation often suffices.

Table 5.2 Comparison of single-sided and central difference quotient method in correlation ratio and accuracy under different numbers of samplings for a 6-hour running CSTR in ASM1 model

Number of samplings	Correlation coefficient		Average accuracy	
	Single-sided	Central	Single-sided	Central
5	75.9%	72.3%	49.6%	50.5%
10	87.1%	83.6%	61.2%	66.3%
20	96.0%	93.1%	71.2%	78.2%
30	98.5%	97.8%	76.8%	83.8%
50	99.0%	99.7%	81.6%	88.0%
100	99.71%	99.98%	86.0%	90.4%
300	99.97%	100%	89.6%	91.5%
500	99.99%	100%	90.4%	91.6%
1000	100%	100%	91.0%	91.6%

In practice, the mean and standard deviation of component reaction rates can also be estimated or corrected by experienced operators from other sources, such as routine operation records, site measurements, established mechanistic models, and digital twin outputs.

As demonstrated in the Experiments and results section, normalisation acts as a preconditioner in NODEs, significantly improving training stability and efficiency. The paired normalisation and de-normalisation layers stabilise gradients, leading to faster convergence and improved training smoothness and efficiency. Notably, the associated computational overhead is minimal, even practically negligible.

5.2.2 Collocation method

While the normalisation method utilises four parameters pre-obtained from the training dataset, the collocation method employs the entire trajectory of the observational training dataset. It estimates the complete trajectory of the corresponding derivatives

using traditional mathematical regression methods, then trains the neural network against the collocated pairs of state input data and estimated state derivatives. This approach eliminates the need for an ODE solver, thereby avoiding the stiffness issue.

In NODEs, the direct approach calls out the ODE solver for derivative calculations at every training step, with derivative computation remaining implicit or hidden to user. Conversely, the collocation method explicitly approximates all the derivatives using kernel functions and interpolation / regression methods prior to the training procedure. Although both methods involve optimisation by gradient descent, training of the neural network in collocation method is simpler as it does not need to go through ODE solver at each step, while direct approach must. Consequently, it can be extremely fast and robust to noise. Figure 5.3 illustrates the different strategies of these two methods.

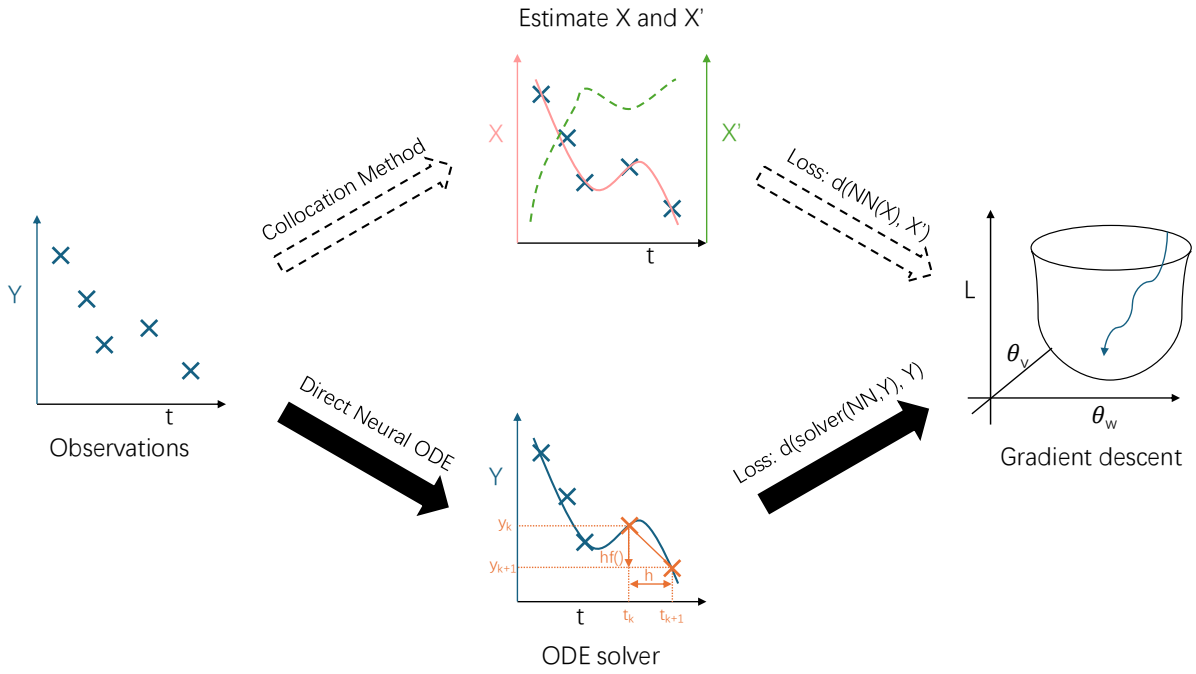


Figure 5.3 Different training strategies by direct NODE and collocation method

The first step of collocation method entails estimating state variables, let's say $X(t)$, and their derivatives, let's say $X'(t)$, from sampled observations (Y_1, Y_2, ..., Y_n), at the time points (t_1, t_2, ..., t_n), with measurement errors (e_1, e_2, …, e_n), then:

$$\boldsymbol{Y_i = X(t_i) + e_i} \qquad \boldsymbol{i = 1, \ldots, n} \qquad \textbf{Equation 5-7}$$

To derive $X(t)$ and $X'(t)$ from Y_i, a common practice is to use non-parametric local linear regression for $X(t)$, and local polynomial (often quadratic) regression for $X'(t)$.

This approach is based on Talay's formular and criterion of minimizing locally weighted least-squares. Liang and Wu (2008) gave a complete deduction process in their paper. The results were briefed as follows.

$$\widehat{\boldsymbol{X}}(\boldsymbol{t}) = \boldsymbol{\varepsilon}_1^T \left(\boldsymbol{T}_{1,t}^T \boldsymbol{W}_t \boldsymbol{T}_{1,t}\right)^{-1} \boldsymbol{T}_{1,t}^T \boldsymbol{W}_t \boldsymbol{Y} \qquad \textbf{Equation 5-8}$$

$$\widehat{\boldsymbol{X}'}(\boldsymbol{t}) = \boldsymbol{\varepsilon}_2^T \left(\boldsymbol{T}_{2,t}^T \boldsymbol{W}_t \boldsymbol{T}_{2,t}\right)^{-1} \boldsymbol{T}_{2,t}^T \boldsymbol{W}_t \boldsymbol{Y} \qquad \textbf{Equation 5-9}$$

where,

$$\boldsymbol{\varepsilon}_1 = \begin{bmatrix} \mathbf{1} \\ \mathbf{0} \end{bmatrix} \quad \boldsymbol{\varepsilon}_2 = \begin{bmatrix} \mathbf{0} \\ \mathbf{1} \\ \mathbf{0} \end{bmatrix} \qquad \textbf{Equation 5-10}$$

$$\boldsymbol{T}_{1,t} = \begin{bmatrix} \mathbf{1} & \boldsymbol{t}_1 - \boldsymbol{t} \\ \mathbf{1} & \boldsymbol{t}_2 - \boldsymbol{t} \\ \vdots & \vdots \\ \mathbf{1} & \boldsymbol{t}_n - \boldsymbol{t} \end{bmatrix} \quad \boldsymbol{T}_{2,t} = \begin{bmatrix} \mathbf{1} & \boldsymbol{t}_1 - \boldsymbol{t} & (\boldsymbol{t}_1 - \boldsymbol{t})^2 \\ \mathbf{1} & \boldsymbol{t}_2 - \boldsymbol{t} & (\boldsymbol{t}_2 - \boldsymbol{t})^2 \\ \vdots & \vdots & \vdots \\ \mathbf{1} & \boldsymbol{t}_n - \boldsymbol{t} & (\boldsymbol{t}_n - \boldsymbol{t})^2 \end{bmatrix} \qquad \textbf{Equation 5-11}$$

$$\boldsymbol{W}_t = \begin{bmatrix} \boldsymbol{K_h}(\boldsymbol{t}_1 - \boldsymbol{t}) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{K_h}(\boldsymbol{t}_2 - \boldsymbol{t}) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{K_h}(\boldsymbol{t}_n - \boldsymbol{t}) \end{bmatrix} \qquad \textbf{Equation 5-12}$$

Where K(·) is a symmetric kernel function. Given an example of Epanechnikov kernel function, so that:

$$\boldsymbol{K_h}(\boldsymbol{t_i} - \boldsymbol{t}) = \boldsymbol{K}\left(\frac{\boldsymbol{t_i} - \boldsymbol{t}}{\boldsymbol{h}}\right)/\boldsymbol{h} \qquad \textbf{Equation 5-13}$$

Where h is a bandwidth:

$$\boldsymbol{h} = \left(\boldsymbol{n}^{-\frac{1}{5}}\right)\left(\boldsymbol{n}^{-\frac{3}{35}}\right)\left((\mathbf{log}(\boldsymbol{n}))^{-\frac{1}{16}}\right) \qquad \textbf{Equation 5-14}$$

The choice of kernel function depends on the observation data characteristics. For instance, cubic spline is preferred for less noisy or relatively sparse data, while B-spline or Epanechnikov kernel is suitable for noisier datasets.

Due to boundary restriction, derivative estimations at both ends are often inaccurate. This can be mended by excluding data at both ends, to achieve smoother results and reduce excessive changes at the boundaries.

Proper data preprocessing can alleviate subsequent burdens and minimise errors in the training process. To enhance the accuracy of estimated X and X', it is advisable to smooth the noisy observation Y before applying the collocation method. This may involve outlier detection, smoothing techniques and cross validation based on wastewater system knowledge.

The next step involves training the neural network with the estimated $\mathrm{X(t)}$ and $\mathrm{X'(t)}$ pairs. This process is straightforward, similar to conventional machine learning. The MAE loss function can be constructed as:

$$\boldsymbol{L(\theta) = \frac{1}{n}\sum_{i=1}^{n}\left|NN_{\theta}(\widehat{X}_i) - \widehat{X'}_i\right|} \qquad \textbf{Equation 5-15}$$

The accuracy of the collocation method depends on the characteristics of observational data and estimation methods. While it can be high if the adopted methods and data align well, the results often require further refinement using more elaborate algorithms.

5.2.3 Incremental training strategy

In practice, the results from the collocation method are often “rough” and not sufficiently accurate, although they may be close to the global minima and less prone to local minimum (Rackauckas *et al.*, 2020). To improve the model fidelity to the optimal level, further training using finer or more elaborate methods, such as local minimum algorithms or direct NODEs approach, is expected. Since the collocation method provides a good initial result, subsequent optimisation will experience reduced stiffness and increased effectiveness. In this way, the model fidelity is incrementally improved.

We refer to this practice - applying a coarse method followed by a finer method that builds upon the results of the previous method - as the incremental training strategy. The idea is to first provide a rough estimation using the collocation method to narrow the approximated range for the result, then refine it locally by the direct NODE method to chieve higher fidelity. Figure 5.4 illustrates the steps of the incremental strategy.

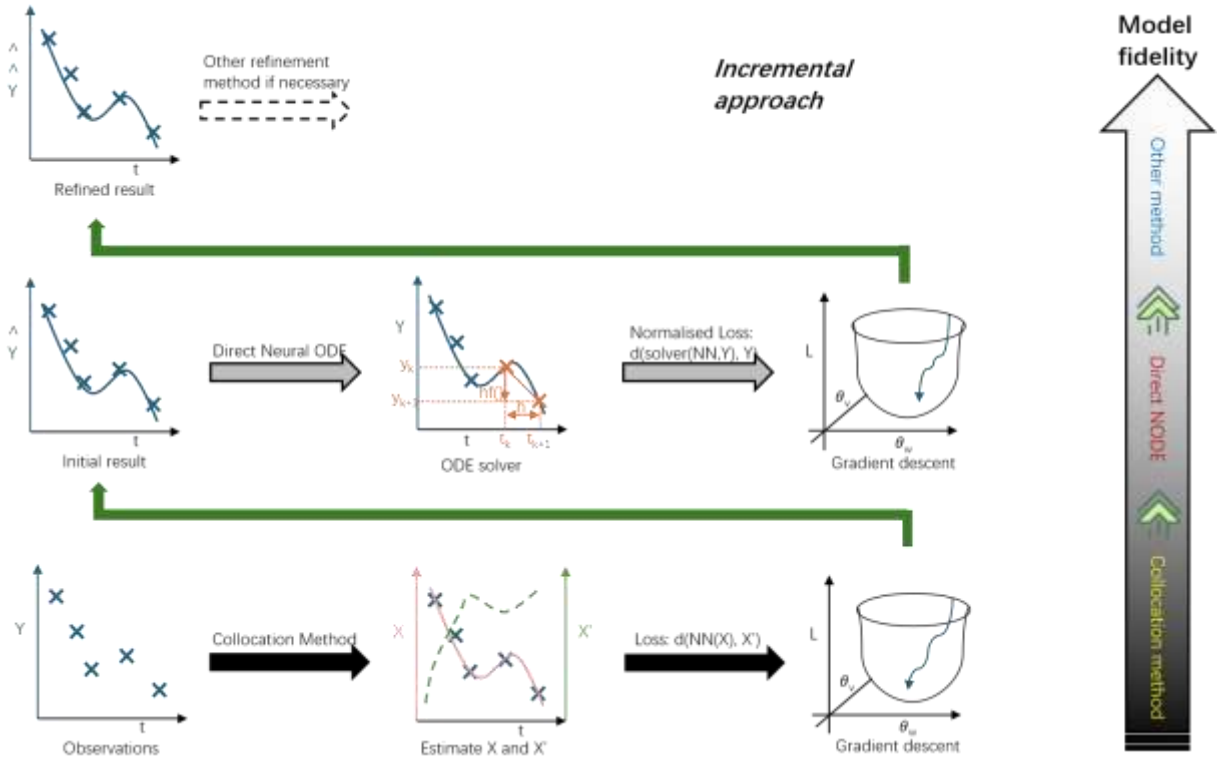


Figure 5.4 incremental strategy

5.3 Experiments and results

Data-driven modelling using NODEs primarily involves the repeated solving of IVPs. To demonstrate the feasibility and efficiency of the proposed methods in training NODE models for wastewater process modelling, two experiments were conducted. Both experiments focused on an IVP in a CSTR using ASM1 and ASM2d-N_2O, respectively.

For comparison purposes, simulated trajectory data generated from these mathematical models were utilised to train the NODEs models. This allows for a direct assessment of the NODE performance against well-established wastewater treatment models.

5.3.1 ASM1 model

The ASM1 is one of the simplest models for wastewater biological process modelling. Introduced in 1987 and revised over the years, it has been widely for simulating organic matter and nitrogen removal in wastewater treatment. The version used consists of 15 components (including two additional components S_{N2} and X_{inorg} for N balance and TSS calculation) and 8 reactions (Mogens, Willi, Takashi and van

Loosdrecht Mark, 2000). Default values for stoichiometric and kinetic parameters were adopted. For detailed model information, please refer to the cited document.

We generated trajectory data of 1000 points using the ASM1 mathematical model, simulating a CSTR from a defined initial state with a fixed dissolved oxygen level of 2 mg O_2/l over 6 hours. The initial values (see Table 5.1) were adapted from the steady state of the bioreactor in BSM1 (Alex *et al.*, 2018). Figure 5.5 illustrate the maximum eigenvalues of the Jacobian over the IVP solution trajectory, showing peak stiffness at approximately 1.7 hours.

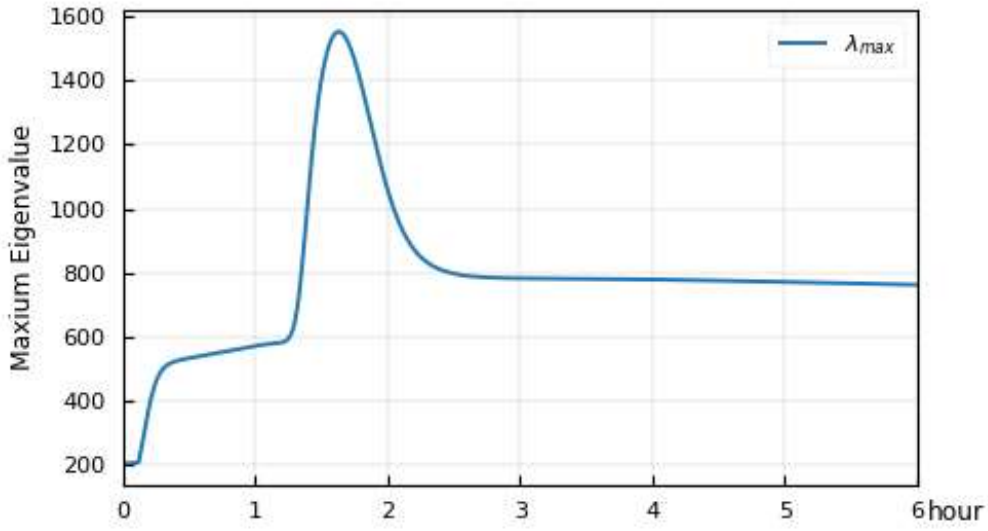


Figure 5.5 Maximum eigenvalues of the Jacobian over the IVP solution trajectory

5.3.1.1 Training options

We conducted experiments using Python 11 and the *torchdiffeq* package (Chen, 2018). After testing various solvers from both *torchdiffeq* and the *SciPy* package (Virtanen *et al.*, 2020), *dopri5* was selected for the experiments. The tests revealed minimal differences between loss functions, with Huber loss performing slightly better. However, MAE was deemed sufficient for process modelling and thus adopted in the experiments.

The neural network for the NODEs was constructed as a multilayer perceptron with four layers and 50 nodes in each hidden layer, using activation functions between the layers. While Chen et al. (2018) used the *Tanh* activation function in most of his NODE examples, *Gelu* activation function was also tested. *Gelu,* a relatively new function,

bridges stochastic regularisers with non-linearities, distinguishing it from other activation functions (Hendrycks and Gimpel, 2016). It has demonstrated higher accuracy compared to *ReLU*, and *ELU* (Devlin *et al.*, 2019).

The experiments showed that *Gelu* outperforms *Relu* and *Tanh* in NODEs training for wastewater modelling. Figure 5.6 compares Gelu and Tanh functions in loss changes for the IVP trajectory training based on ASM1 model, clearly indicating considerably enhanced performance with *Gelu*. It is worth noting that despite nearly 10^2 orders of magnitude loss decline activated by *Gelu* without normalisation as shown in Figure 5.6, the results remained unsatisfactory.

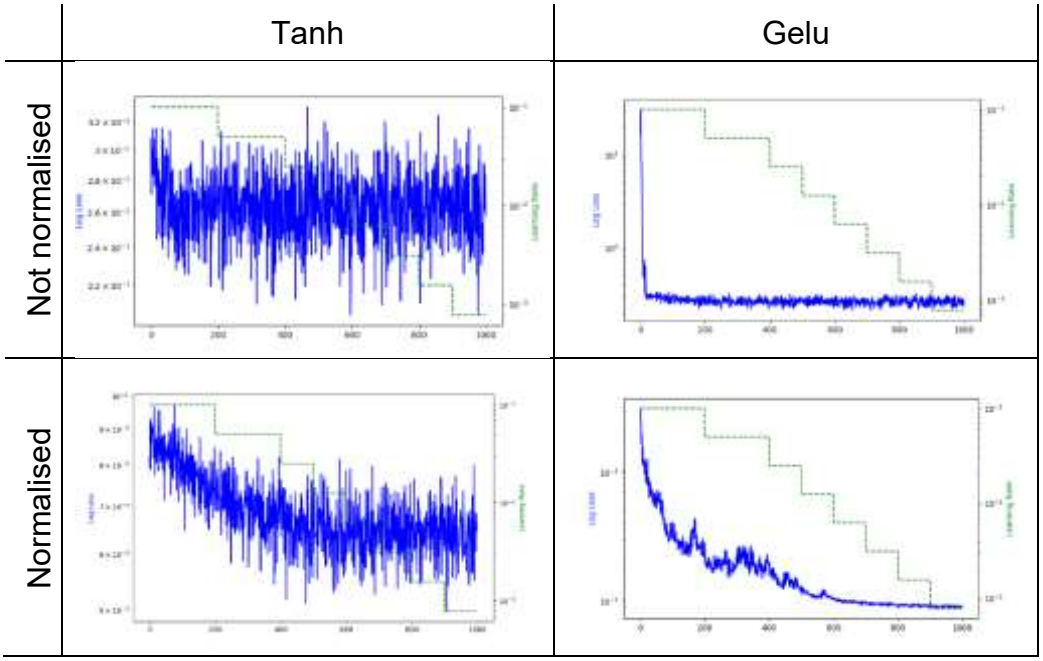


Figure 5.6 Comparison of activation functions in training loss for an ASM1 IVP, with different configurations (top left: Tanh without normalisation; top right: Gelu without normalisation; bottom left: Tanh with normalisation; bottom right: Gelu with normalisation)

The NODEs were trained using the ADAM optimiser (Kingma and Ba, 2014) with a varying learning rate, as shown in the Figure 5.6. The training began with a high learning rate 0.1 to harness speed advantages, then switched to a lower rate 0.001 to refine results in response to loss function changes. Training was conducted for 2000 iterations with a sampling batch size of 512 and 16 steps of the interval calculated each time by the solver. For brevity, three non-reactive components (S_i, X_i, X_{inorg}) and constant DO are not shown in the following results.

5.3.1.2 Normalisation

The efficacy of the proposed method is assessed by comparing the trained model predictions against ground truth trajectories generated by the mathematical model under identical initial conditions. Figure 5.7 illustrates a representative training example without normalisation. The results indicate that the neural network fails to effectively learn from the data, as evidenced by the loss function plateauing after an initial rapid decrease within the first few iterations. The gradient norm exhibits a pathological pattern, remaining at a consistently low level, which reflects the stagnation of the loss throughout the training process. Consequently, the predicted component curves do not align well with the ground truth, resulting in a high overall RMSE of 62.51.

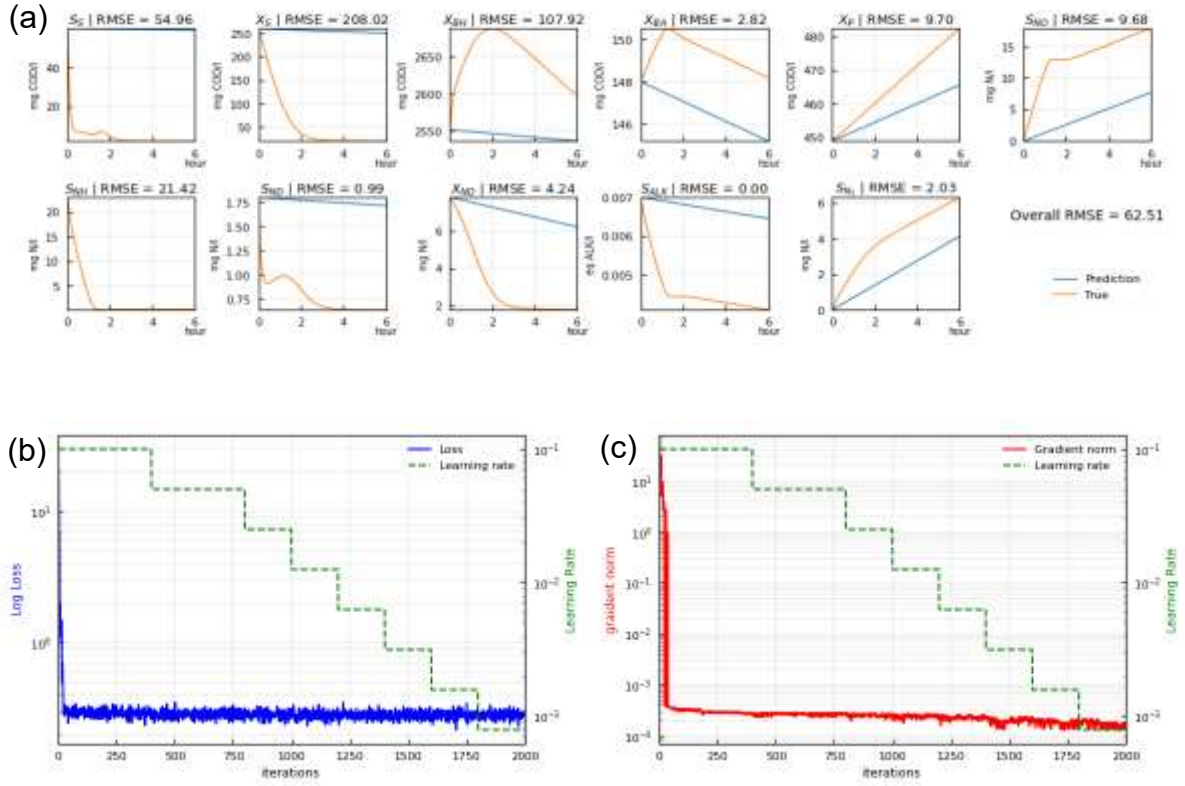


Figure 5.7 NODE training without normalisation (a) training results (b) loss (c) grad norm

The normalisation method is then applied with the estimated mean and standard deviation of the derivatives sequence using single-sided difference quotient. The neural network was wrapped by the normalisation and de-normalisation pair with the estimated parameters. As illustrated in Figure 5.8, this normalisation technique yields significant improvements. The predicted trajectory curves now closely align with the

ground truth, resulting in a substantially reduced overall RMSE of 1.73. Moreover, the loss function exhibits a gradual decrease throughout the training iterations, while the gradient norm demonstrates stable behaviour.

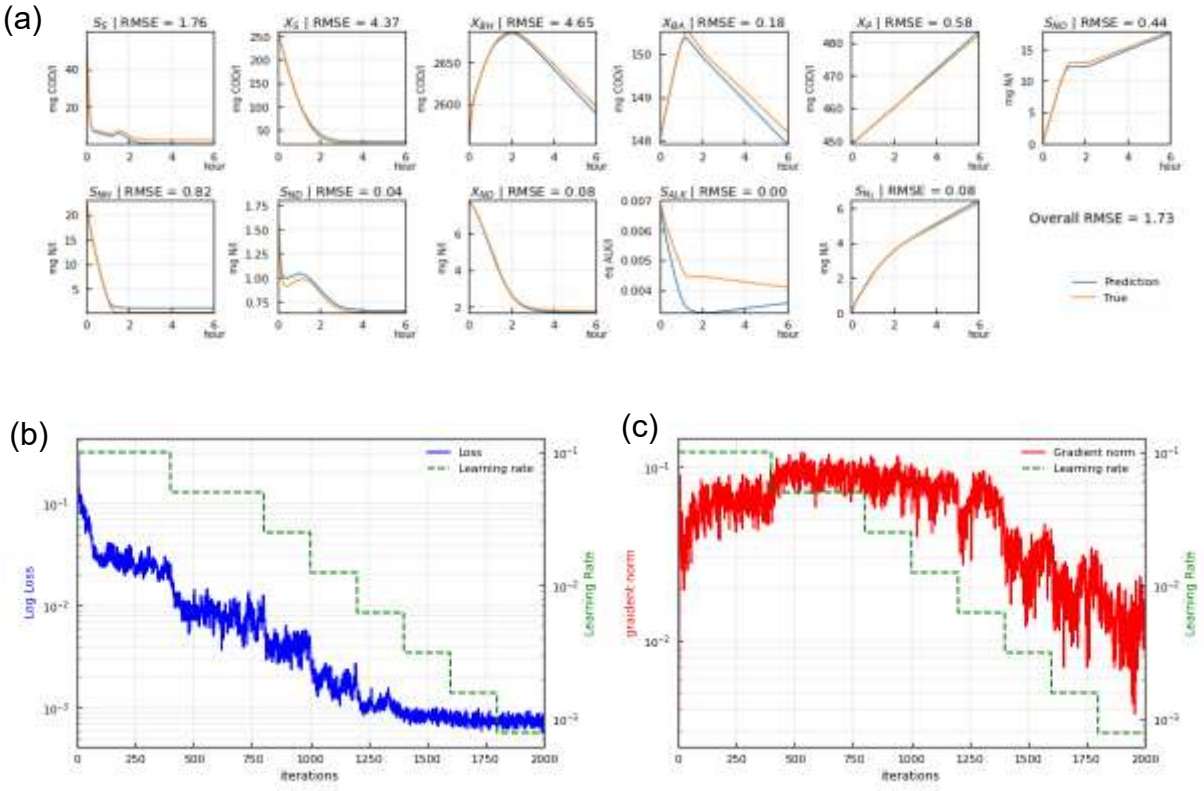


Figure 5.8 NODE training with normalisation (a) training results (b) loss (c) grad norm

5.3.1.3 Incremental strategy

The incremental strategy was implemented by first training the normalised model using the collocation method, followed by direct NODE training. Figure 5.9 and Figure 5.10 illustrate the results of the collocation method utilising the Epanechnikov kernel function, comparing the ground truth with the collocated trajectory and its derivatives. The smoothed trajectory demonstrated a close fit to the ground truth, as evidenced by a low RMSE of 6.39. However, the derivatives exhibited significant disparity, with a high RMSE of 325.5, indicating challenges in accurate estimation.

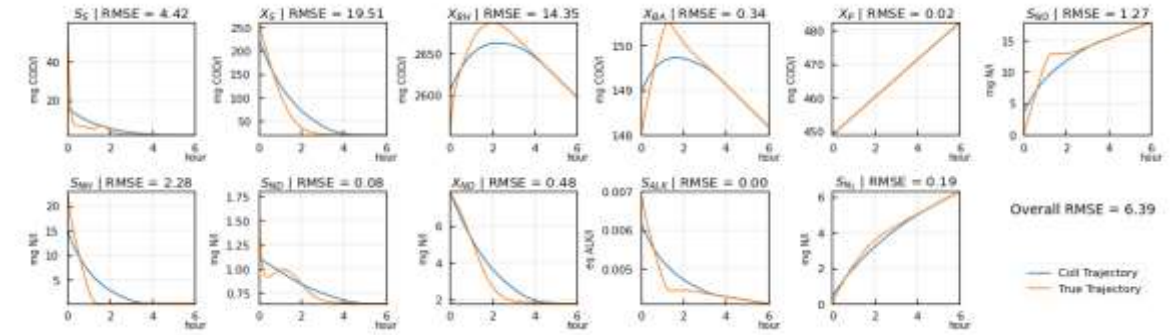


Figure 5.9 comparison of collocated data and ground truth of the trajectory

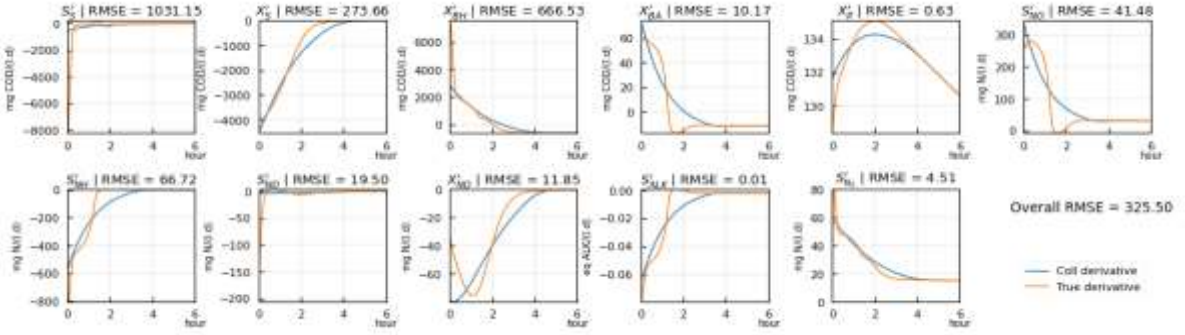
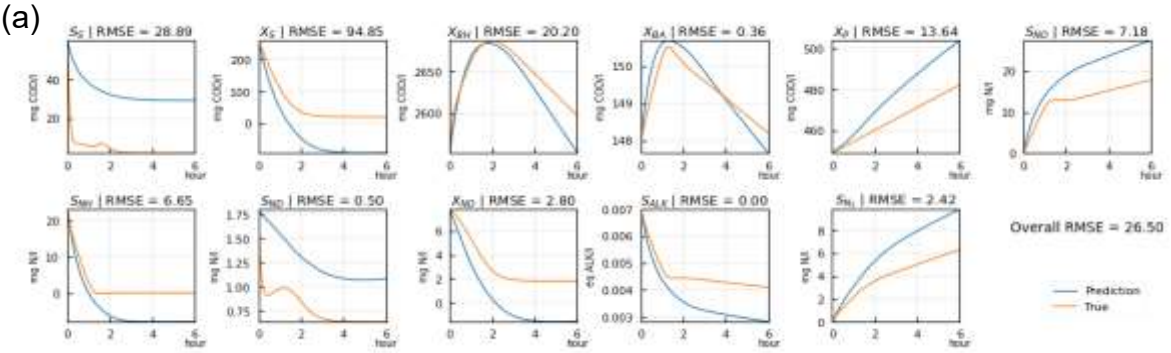


Figure 5.10 comparison of collocated derivative and ground truth derivative

Figure 5.11 illustrates a representative result from the model through NODE prediction after trained on collocated dataset in the collocation training stage. As the number of iterations increases, the loss consistently decreases, while the gradient norm remains stable. Despite this apparent progress, the RMSE remains high at 31.44 after 2,000 iterations. This persistent discrepancy can be attributed to substantial errors in derivative estimation using the collocation method.

(a)



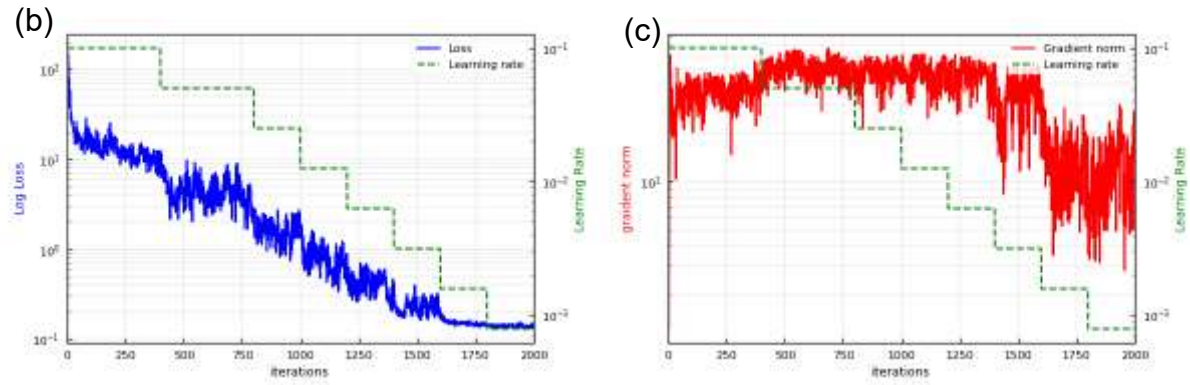


Figure 5.11 collocation training stage by incremental strategy (a) training results (b) loss (c) grad norm

Figure 5.12 displays the results of the subsequent direct NODEs training stage. Throughout this stage, the loss exhibits a generally consistent decrease, while the gradient norm maintains stability at a relatively low level compared to collocation training phase.

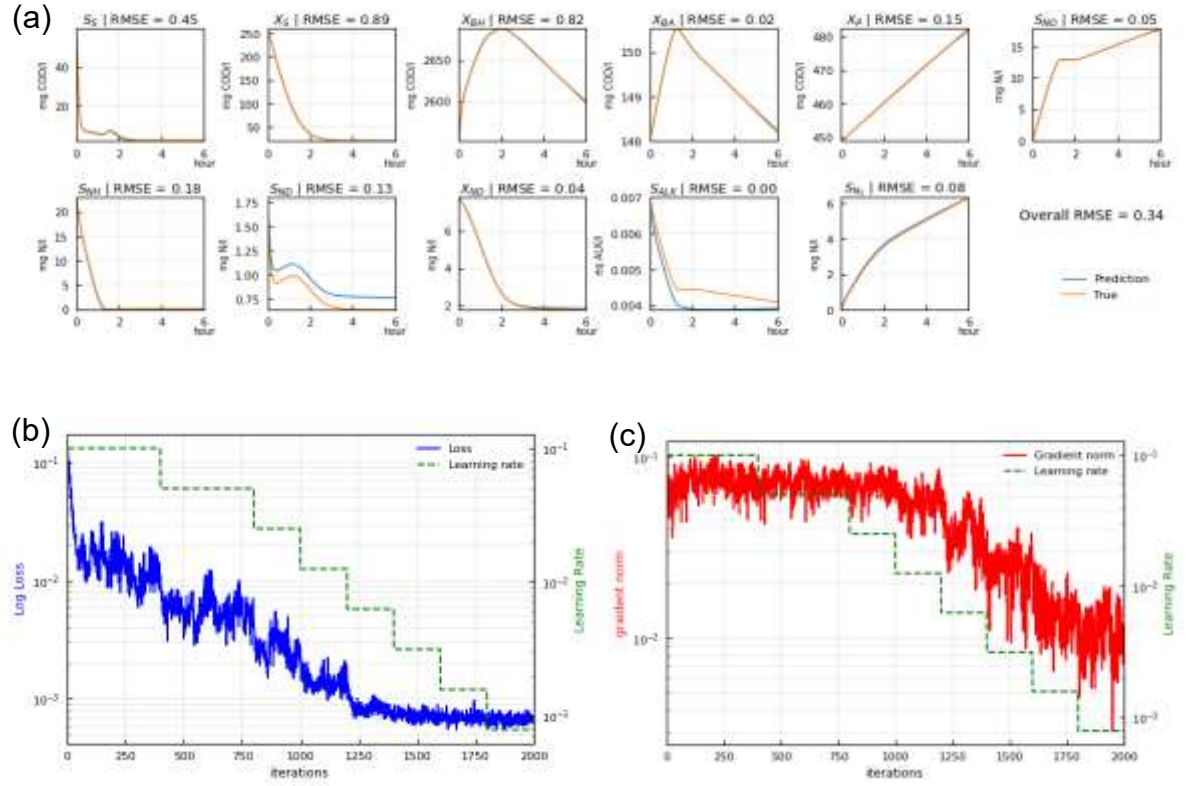


Figure 5.12 direct NODE training stage by incremental strategy (a) training results (b) loss (c) grad norm

Due to the stochastic nature of neural networks, results may vary slightly across different training runs. To assess efficiency, this study conducted 100 trials of each training method, each comprising 2,000 iterations under identical conditions. The tests were performed on a computer with an Intel® Core™ i7 CPU (2.8 GHz), 16 MB RAM, without a dedicated GPU. The test program was executed within the VSCode IDE on a Windows 10 64-bit operating system.

Figure 5.13 presents the results of this efficiency test (detailed data available on the project's GitHub repository). The analysis reveals that the incremental training strategy, compared to the NODE-only training, consumes 24.3% less time on average and yields a 24.7% lower RMSE. Notably, when collocation training precedes NODE training, the resulting RMSE demonstrates a smaller standard deviation (1.2) compared to the method without collocation integration (1.4), suggesting enhanced training stability.

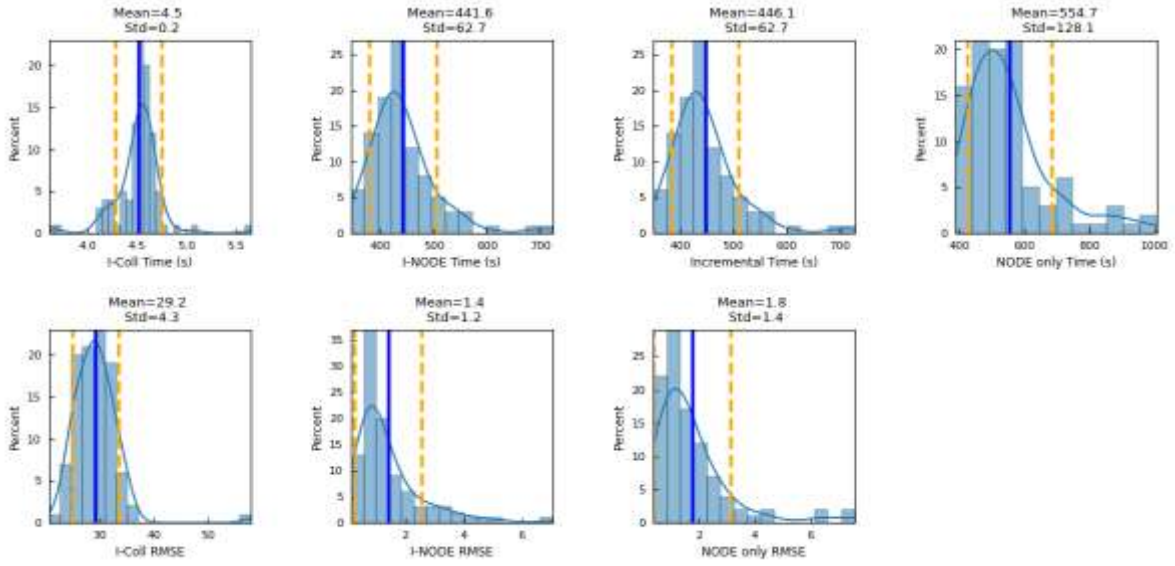


Figure 5.13 Distribution of time consumption and RMSE for 100 times running by different training methods. (I-Coll: Incremental collocation training part, I-NODE: Incremental NODE training part)

5.3.2 ASM2d-N_2O model

The growing concern over climate change has intensified focus on greenhouse gas emissions from wastewater treatment plants, particularly nitrous oxide (N_2O). Recent research (Ye, Porro and Nopens, 2022) identifies four potential pathways for N_2O

generation during biological nitrogen removal in wastewater treatment: (i) hydroxylamine oxidation, (ii) nitrifier denitrification, (iii) heterotrophic denitrification, and (iv) abiotic reactions. While these pathways may coexist at varying ratios, the significance of the fourth pathway remains under debate.

This experiment employs the ASM2d-N_2O model, an extension of the ASM2d model that incorporates N_2O emissions. This comprehensive model describes 40 reactions involving 24 fractionated components, encompassing the biological removal of carbon, nitrogen, and phosphorus, including the three major N_2O emission pathways. The experiments utilised the stoichiometric and kinetic parameters reported in the original paper (Massara et al., 2018).

The model's complexity arises from its inclusion of greenhouse gas emissions like N_2O and other transient, low-concentration byproducts such as nitric oxide. Additionally, it represents complex biochemical reactions. These factors collectively contribute to a significant degree of stiffness in the system due to the vast differences in scales and magnitudes between various components. This characteristic makes the ASM2d-N_2O model a suitable test case for evaluating the proposed solutions.

MATLAB was employed as the programming language for this experiment. This choice was motivated by the desire to leverage the latest NODE techniques, as commercial platforms typically update their products more frequently than open-source alternatives. The data originated from an ASM2d-N_2O model simulation of a CSTR for six hours under constant dissolved oxygen control at 2 mg/L. The simulation began from an initial condition detailed in Table 5.3. The generated trajectory data for the IVP was discretised into 1000 points.

Table 5.3 Initial condition defined in experiment of ASM2d-N2O model

Component	Unit	Value	Component	Unit	Value
S_{O2}	mg O2/l	2	S_{N2}	mg N/l	0
S_F	mg COD/l	48.5	X_I	mg COD/l	40.4
S_A	mg COD/l	32.3	X_S	mg COD/l	202
S_{NH4}	mg N/l	20	X_H	mg COD/l	2500
S_{NH2OH}	mg N/l	0	X_{PAO}	mg COD/l	250
S_{N2O}	mg N/l	0	X_{PP}	mg P/l	70
S_{NO}	mg N/l	0	X_{PHA}	mg COD/l	100

S_{NO2}	mg N/l	0	X_{AOB}	mg COD/l	200
S_{NO3}	mg N/l	2.6	X_{NOB}	mg COD/l	100
S_{PO4}	mg P/l	9	X_{TSS}	mg TSS/l	189
S_I	mg COD/l	48.5	X_{MeOH}	mg TSS/l	50
S_{ALK}	mol HCO_3^-/m^3	5	X_{MeP}	mg TSS/l	220

To simulate the real situation and evaluate how the proposed methods behave on noisy monitoring data, four datasets were prepared for this experiment. One set contained the original, noise-free data. The remaining three datasets were contaminated with varying levels of white noise, each with a different standard deviation (SD) amplitude: 0.01, 0.05, and 0.1. To mitigate the effect of noise on training, a Gaussian filter was applied with a window size of 50 for smoothing before training the models. Figure 5.14 illustrates the generated ground truth data, the data corrupted with noise, and the smoothed data.

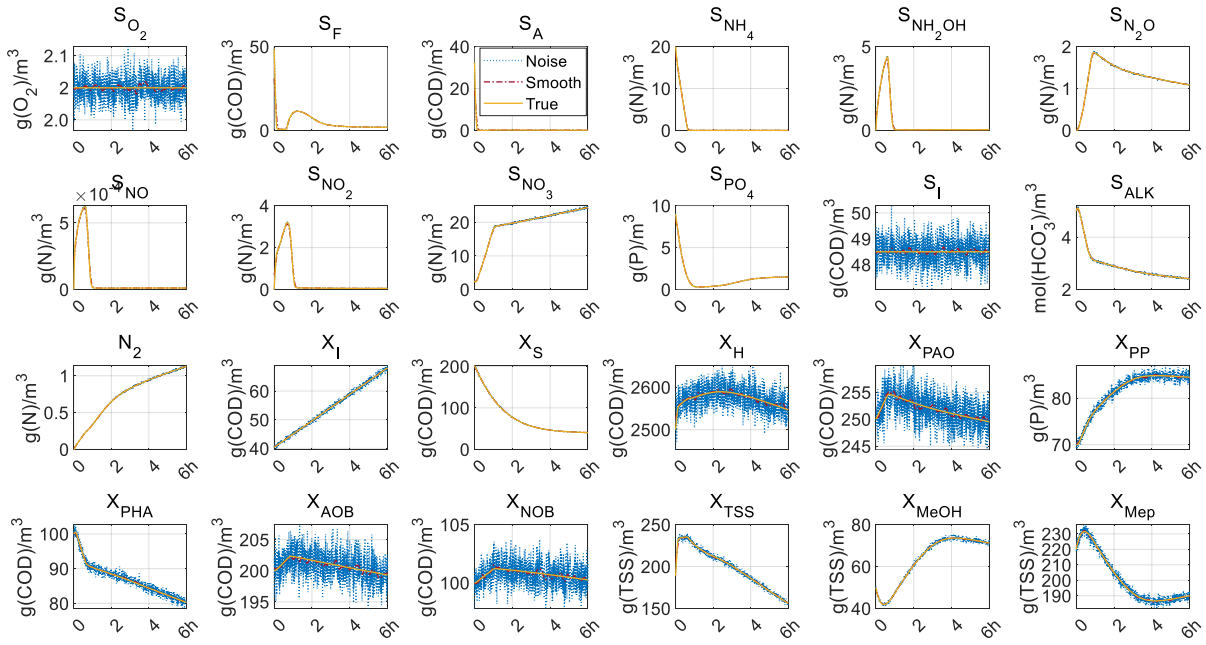


Figure 5.14 Ground truth, noised and smoothed trajectory data of ASM2d-N_2O model

5.3.2.1 Training options

MATLAB offers a comprehensive suite of ODE solvers yet currently provides only one solver specifically designed for Neural ODE (NODE) problems. This limitation reflects the relative novelty of the NODE approach, underscoring the need for further development in this area. Despite the *dlode45* solver being documented as well-suited

for non-stiff problems (MATLAB, 2024), it was employed in this specific case for experimentation.

For the NODE model, this study constructed a MLP architecture with two hidden layers, each containing 50 nodes. The Gelu activation function was employed between layers for improved performance. Xavier Glorot initialization was utilised for the weight matrices within the neural network to address vanishing/exploding gradients. The MAE served as the loss function throughout the experiments, aiming to minimise the absolute difference between predicted and ground truth values.

The NODEs were trained using the *Adamupdate* optimiser (Kingma and Ba, 2014) with a gradient decay factor of 0.9, a squared gradient decay factor of 0.999, and a global learning rate of 0.01. Custom loops were implemented to manage the training process. The collocation training stage utilised 3000 iterations, followed by 1000 iterations for direct NODE training (or 3000 iterations if used independently). A batch size of 200 and time steps of 800 were employed during training.

5.3.2.2 Normalisation

We initially evaluated the performance of direct NODE training without normalization. To validate the results, the predictions from the trained model were compared with solutions generated by the mathematical model for the same IVP. Figure 5.15 presents a typical example of training with data containing 0.05 SD noise and without normalization. While the model successfully describes the trajectory of most components, it struggles to capture the trajectories of low-valued scaled components, such as S_{NO}, which have magnitudes on the order of 10^{-4} mg/L.

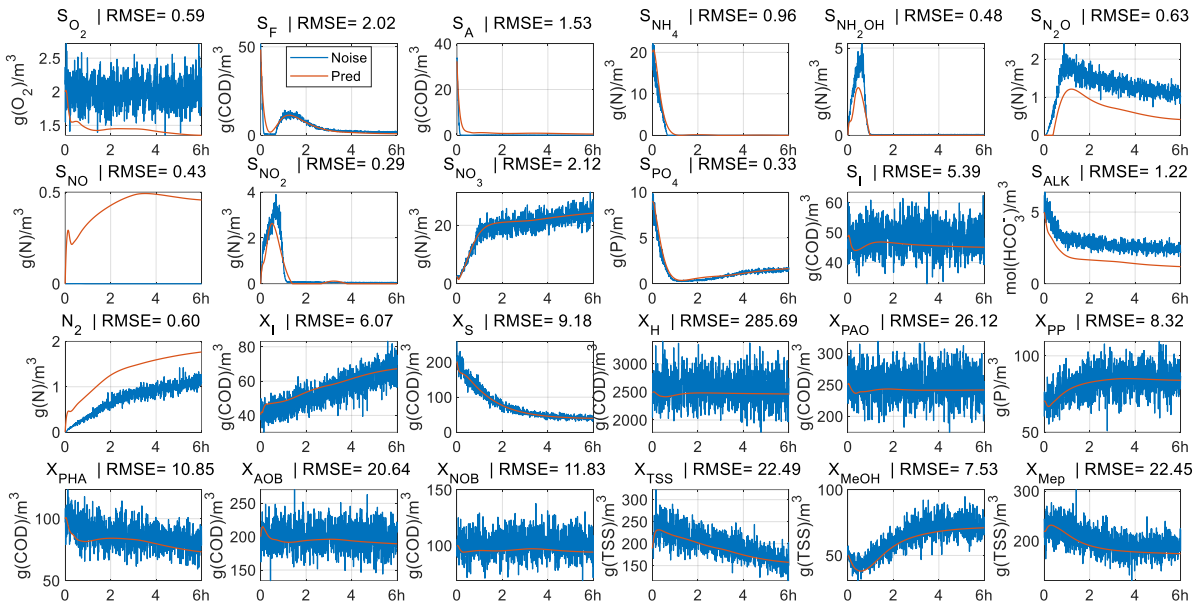


Figure 5.15 result of training without normalisation with data containing 0.05 SD noise

Following this observation, the normalisation method was implemented. Figure 5.16 illustrates the results of direct NODE training with normalisation-denormalisation pair. For the normalisation layer, the mean and standard deviation were directly calculated from the smoothed component state data sequence. Differently, the denormalisation layer employed the differential quotient method to estimate the mean and standard deviation of the derivative data sequence from the smoothed dataset. Although some turbulences are still evident, the predicted trajectory progressively improves in smoothness with increasing training iterations. Importantly, the model now captures the trajectories of low-valued scaled components like S_{NO}, leading to more satisfactory overall results.

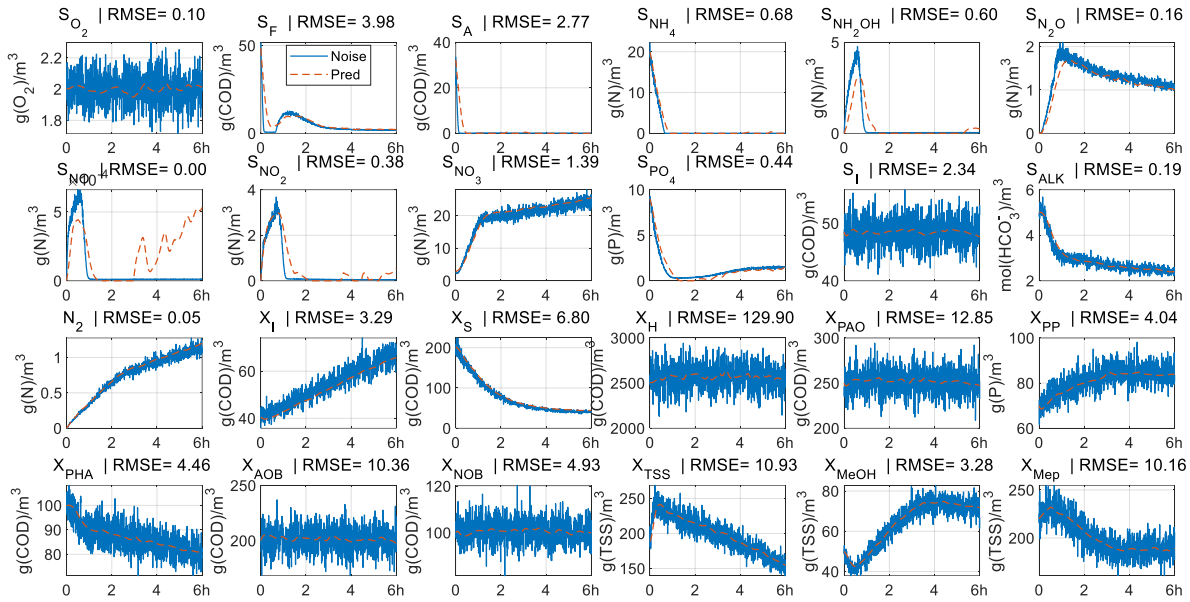


Figure 5.16 result of training with normalisation with data containing 0.05 SD noise

5.3.2.3 Incremental strategy

We evaluated the performance of incremental training strategies using data contaminated with normally distributed noise at three SD levels: 0.01, 0.05, and 0.1. The training process consisted of two stages: first, the model was trained using the collocation method, followed by further training with the direct NODE method. The neural network employed Z-score normalisation at the input layer and de-normalisation at the output layer. An Epanechnikov kernel function was chosen for data collocation.

The collocation training exhibited a fast and stable convergence process, requiring 3000 iterations. Subsequent direct NODE training also demonstrated good convergence, achieving satisfactory results with only 300 iterations, although 1000 iterations were used for further refinement.

As illustrated in Figure 5.19, Figure 5.18 and Figure 5.19, the predictions were compared to the observations after training with the same number of iterations on data with noise levels of 0.01, 0.05, and 0.1 SD, respectively. The results clearly demonstrate that the accuracy deteriorates as the noise level increases. Lower noise levels yield better results, as evidenced by the RMSE values of 5.56, 27.28, and 55.52 for noise levels of 0.01, 0.05, and 0.1 SD, respectively. This reinforces the importance

of minimising noise through data smoothing techniques before training to avoid training instability and ensure optimal performance.

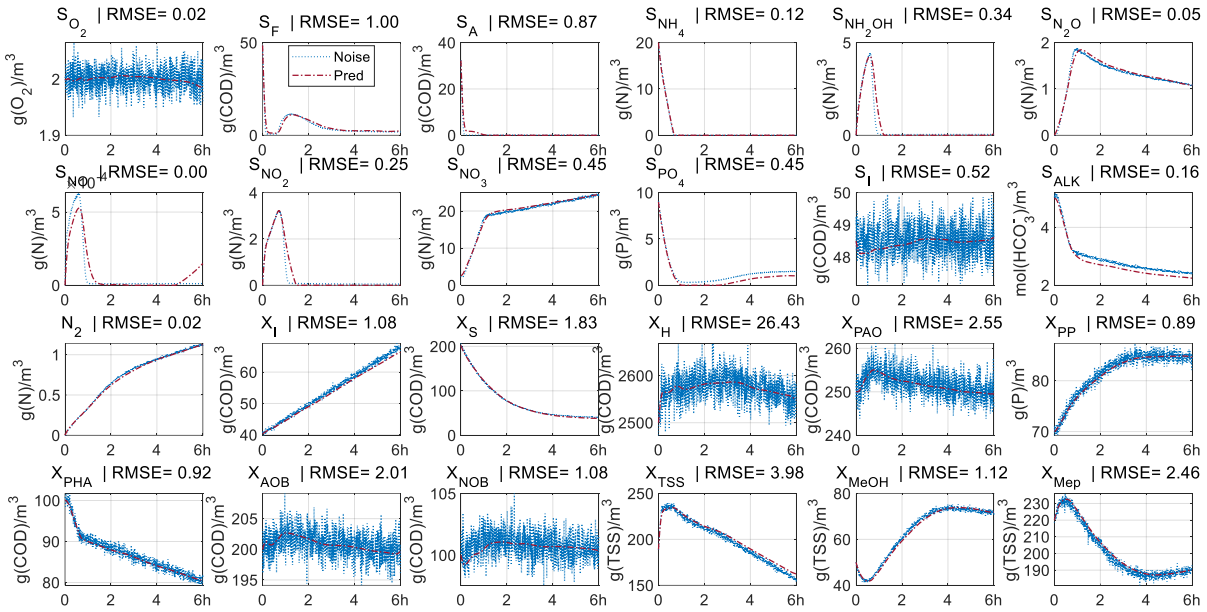


Figure 5.17 Validation of training by incremental strategy with data containing 0.01 SD noise

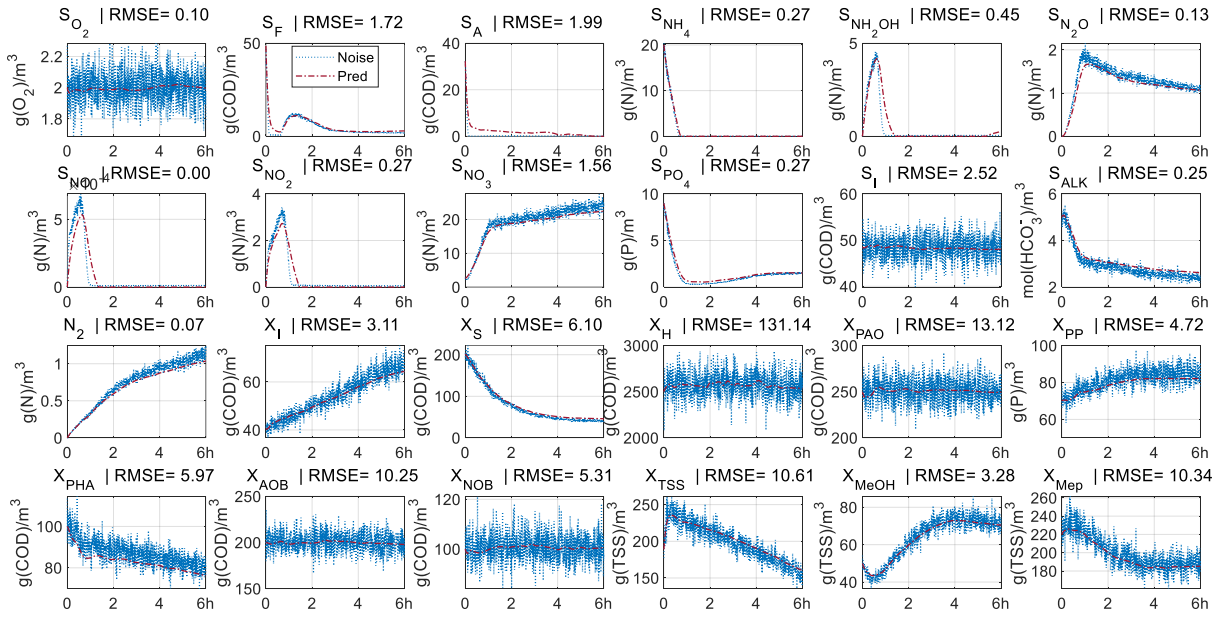


Figure 5.18 Validation of training by incremental strategy with data containing 0.05 SD noise

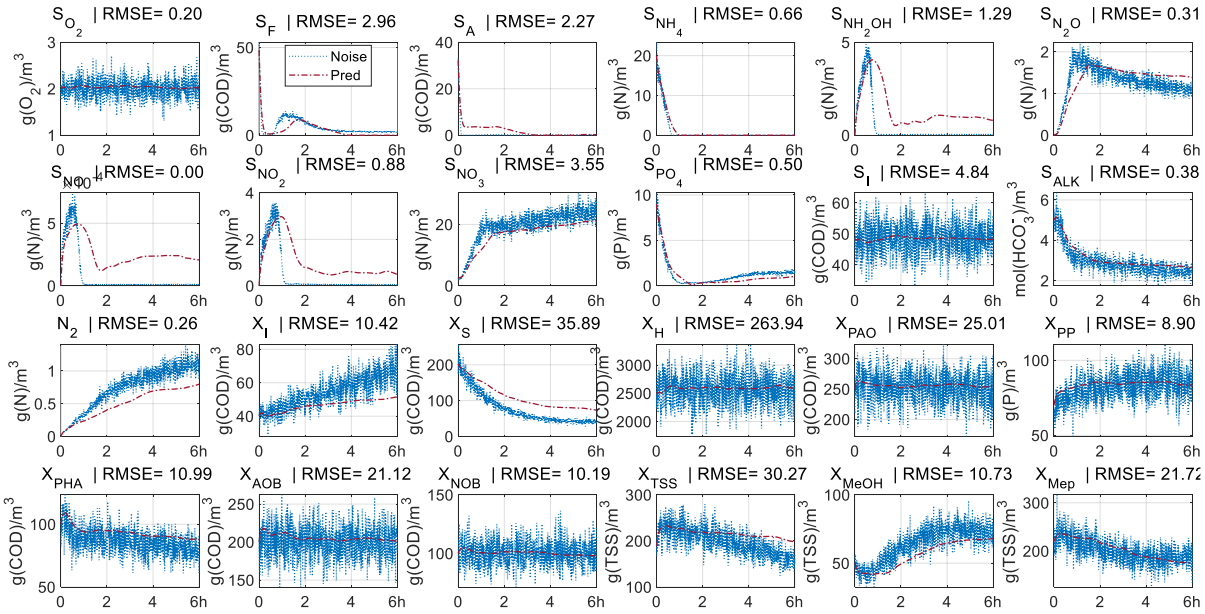


Figure 5.19 Validation of training by incremental strategy with data containing 0.1 SD noise

The comparison demonstrates that the incremental approach of collocation training followed by the direct NODE method improved the result and enhanced the training stability and speed. Regardless of the chosen approach, normalisation remains crucial for successful training of stiff neural ODE models for wastewater process modelling.

5.4 Summary

Training NODEs on the stiff dynamics of wastewater treatment processes presents substantial challenges. This study proposed a novel normalization method to stabilize the training process by addressing the disparate scales inherent in wastewater data. By facilitating smoother gradient descent, this approach enables more accurate data-driven wastewater process modelling.

To expedite training, this study developed an incremental strategy combining the efficiency of collocation methods with the precision of direct NODE training. This hybrid approach effectively overcomes initial stiffness and improves overall model accuracy.

However, the sensitivity of NODEs to noise necessitates careful data preprocessing. Data smoothing is highly recommended to attenuate noise amplification and enhance model robustness.

The findings demonstrate the potential of NODEs for extracting underlying mechanisms from wastewater monitoring data. The proposed normalization and incremental training methods offer practical solutions for overcoming stiffness challenges, thus expanding the applicability of NODEs in wastewater treatment plants. These advancements contribute to more efficient and data-driven wastewater management.

Chapter 6 Implementation of NODE modelling of N_2O in BSM1 plant

Nitrous oxide (N_2O) emissions from wastewater treatment plants (WWTPs) pose a significant environmental threat due to their 265-fold greater global warming potential than carbon dioxide and detrimental effects on the ozone layer depletion (WMO, 2024). Accurate modelling of N_2O production is crucial for improved understanding, predicting, and ultimately mitigating emissions.

6.1 Background

Data-driven modelling, fuelled by advancements in computing and sophisticated AI algorithms, offers a promising alternative (Guo *et al.*, 2020; Ji *et al.*, 2021; Cuomo, Cola, *et al.*, 2022; Kong *et al.*, 2022). This study focuses on emerging NODEs (Chen *et al.*, 2018), a novel method that combines the expressive power of neural networks with the continuous integration capabilities of ODEs. This combination allows NODE to capture the intrinsic dynamics that of mathematical models attempt to describe. Consequently, NODEs can effectively adapt to the inherent variability and non-linearity in wastewater treatment systems. Additionally, unlike traditional machine learning methods constrained by fixed time steps, NODEs excel at learning and representing complex temporal dynamics regardless of irregular or variable time intervals frequently encountered in wastewater treatment data (Kidger et al., 2020). This makes NODEs a well-suited tool for modelling wastewater systems with such real-world complexities.

The core objective is to train a DNN that can approximate the dynamics in the form of ODEs from process monitoring data. However, two challenges must be overcome. The first lies in learning the intrinsic dynamics from the monitoring data that contain external influences, such as changes from the continuous influent input or/and operational control adjustment. Mathematically, this translates to solving a system of NODE with exogenous excitement (Böttcher and Asikis, 2022). To address this an updated training algorithm that can separate the intrinsic dynamics from the data with external factors was developed. Secondly wastewater mechanistic ODEs often exhibit stiffness, a property that can hinder training stability (Kim *et al.*, 2021). Preceding section already tested the proposed normalisation method and incremental training strategy and proved their effectiveness in tackling the stiffness.

The proposed methods successfully trained NODE models using data simulated with the ASMG1 model (Guo and Vanrolleghem, 2014) on BSM1 plant scenarios (Alex *et al.*, 2018). Validation results were impressive, particularly regarding prediction of minutely scaled nitrous oxide, indicating potential for real-world applications.

6.2 Methodologies

BSM1 plant time series data, simulated using a self-built experimental platform, replicates real-world monitoring data from a plug-flow A/O process. However, the data is a composite of underlying intrinsic reactions and exogenous factors. Consequently, the initial step involves developing an algorithm to address exogenous excitation, followed by handling the stiffness issue during the training process.

6.2.1 NODE with exogenous excitation

NODEs have demonstrated effectiveness in learning dynamics across various systems, including time series prediction tasks such as weather forecasting (Verma, Heinonen and Garg, 2024), electricity demand prediction (Xie, Parlikad and Puri, 2019), and COVID-19 spread modelling (Berkhahn and Ehrhardt, 2022). They excel at identifying dynamical patterns and trends within data, regardless of the specific context or the underlying driving forces. Even for complex systems like wastewater treatment, NODE models can be powerful tools. For instance, they could predict future effluent quality based solely on historical time-series effluent data, assuming a consistent influent pattern.

Intrinsic dynamics represent the inherent relationships with the system, enabling generalisation. These dynamics capture the autonomous rise and fall of biomass and substrates within a reactor, driven by their intrinsic biochemical properties and physical propensity towards equilibrium (Mogens, Willi, Takashi and van Loosdrecht Mark, 2000).

Real WWTPs experience continuous influent flow and operational adjustments for optimal efficiency. Observed monitoring data reflects a combination of:

- **Intrinsic biochemical reactions**: Represented by the function f in mechanistic models.
- **External influences**: Continuous influent input and operational control settings.

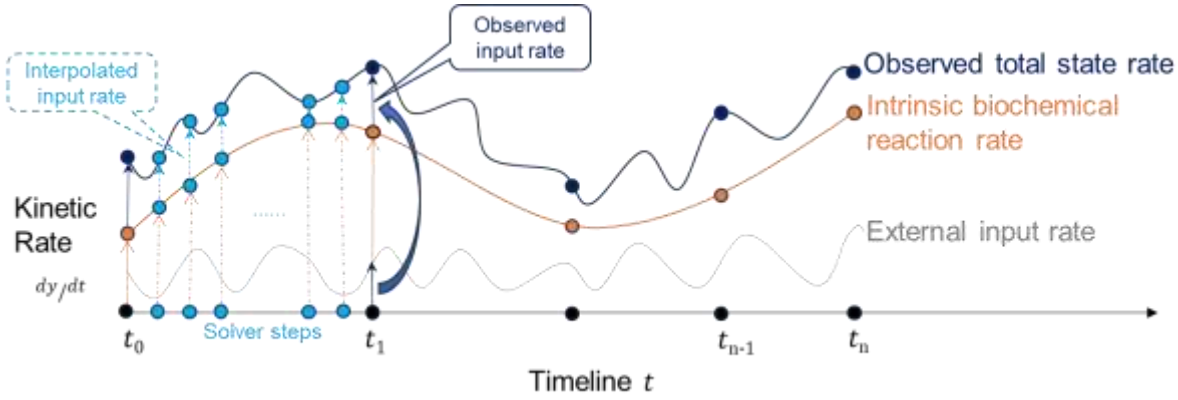


Figure 6.1 Derivation of intrinsic biochemical reaction rates from observations including exogenous inputs for NODE model training.

As illustrated in Figure 6.1, it can be generalised in equations.

$$\frac{dY_{total}(t)}{dt} = \frac{dY_{intrinsic}(t)}{dt} + \frac{dY_{input}(t)}{dt} \qquad \textbf{Equation 6-1}$$

where $\frac{dY_{total}(t)}{dt}$ represents the total of rate of change observed, $\frac{dY_{intrinsic}(t)}{dt}$ represents biochemical intrinsic dynamics, and $\frac{dY_{input}(t)}{dt}$ captures exogenous perturbations.

The aim is to capture the intrinsic biochemical dynamics, represented by function f in mechanistic ODE model. This necessitates differentiating between intrinsic dynamics and variations caused by influent changes in the observed data. Mathematically, this translates to solving a NODE with an exogenous excitation term.

The training procedure must be extended to distinguish between external input effects and intrinsic dynamics. At each step of the ODE solver, the influence of the external inputs is subtracted from the data, ensuring that only the intrinsic time-series trajectory data are utilised for training the neural network. This approach aligns with established practices in mathematical modelling.

This approach enables the NODE model to learn intrinsic biochemical dynamics while accounting for exogenous excitations, potentially improving its generalisation capability and applicability to real-world WWTP scenarios.

6.2.2 Tackling stiffness issue

Stiffness arises in dynamical systems when processes occur at vastly disparate time scales (Hairer and Wanner, 1996). This phenomenon is commonly observed in wastewater treatment processes. For instance, heterotrophic biomass, operating on a magnitude of 10^3, exhibits relatively steady behaviour, while N_2O generation, a transient intermediate product, occurs on a much smaller scale around 10^{-3} and changes rapidly.

In NODE models, the inherent stiffness in wastewater mechanistic ODE, is further amplified by the stochastic nature of neural networks. This amplification originates from the randomness in weight and bias initialisation, as well as their updates during gradient descent algorithms used for training (Kim *et al.*, 2021). Conventional stiff ODE solvers cannot effectively handle this combined stiffness, posing a significant challenge for training NODE.

To address this issue, a novel normalisation method and an incremental training strategy were proposed and tested in previous chapter. This implementation adopted these methodologies to tackle stiffness issue.

The NODE model training consists of two main stages:

1. **Collocation Method**: Initially, employing collocation method (Roesch, Rackauckas and Stumpf, 2021) , which uses local polynomial regression to generate time series pairs of smoothed states Y and estimated derivatives Y'. This method provides a preliminary result and helps optimize the neural network parameters in the first stage. By training on these collocated data, this method effectively mitigates stiffness during the initial training phase.
2. **Direct NODE Training**: Following the collocation stage, the model was refined through direct NODE training with proposed normalisation method (Finlay *et al.*, 2020). This second stage builds upon the preliminary results, incrementally improving model fidelity.

This incremental training approach offers several advantages:

- Reduced computational load: The collocation method provides a computationally efficient starting point, reducing the overall training time.

- Mitigated stiffness: By avoiding stiffness in the initial stage, a more stable foundation was created for subsequent direct NODE training.
- Enhanced convergence: The pre-trained model from the collocation stage often leads to faster and more reliable convergence during direct NODE training.

The results demonstrate that this incremental strategy significantly improves the training process, particularly for stiff systems prevalent in N_2O emission modelling. It allows for more efficient use of computational resources while maintaining model accuracy.

6.3 Implementation settings

The implementation was setup on the self-built environment as described in Chapter 4. The experimental environment was based on the BSM1 plant, a well-established platform in wastewater treatment research. ASMG1 mathematical model was employed to replace the ASM1 to generate observed data including nitrous oxide production. The core components of ASMG1 were then adopted to characterise the NODE model of N_2O production implemented in BSM1 plant. MATLAB (The MathWorks Inc., 2024) served as the programming language for the implementation.

6.3.1 Training options

The selection of the training hyperparameters was based on empirical experimentation. The choices are detailed as follows.

ODE solver: *dlode45* function was selected for the experiments, as it is the only available solver for NODE in MATLAB (MATLAB, 2024) so far, despite it is not recommended for stiff problems.

Structure of neural network: The neural network architecture plays a crucial role in encapsulating the model's dimensional complexity and the reaction dynamics. The test used a MLP with four hidden layers, each containing 50 nodes. While deeper structures (e.g. 5-6 layers, 100-200 nodes) could enhance representation, they are computationally expensive. Conversely, a two-layer model was considered insufficiently expressive.

Activation function: After evaluating Tanh and ReLU functions, the GELU (Hendrycks and Gimpel, 2016) was chosen for its efficiency in propagating gradients between layers.

Initiation of the network parameters: Xavier Glorot initialization (Glorot and Bengio, 2010) was used for weight matrices to mitigate vanishing or exploding gradients and therefore enhance the speed and stability of the training process. Kaiming He initialization (He *et al.*, 2015) is also a viable option, but it is particularly fit for ReLU activations.

Normalisation parameters: In some experiments, the mean and standard deviation used for normalization are calculated directly from the training data and through difference quotient method. Other experiments define these parameters as fixed values (see Table 6.1) based on observations of the training dataset and considerations from wastewater treatment theory and practice.

Table 6.1 fixed values for normalisation parameters for BSM1 plant

Component	Unit	yMean	yStd	Unit	dyMean	dyStd
S_I	mg COD/l	30	10	mg COD/l/d	0	0
S_S	mg COD/l	1.6	2	mg COD/l/d	-200	400
X_I	mg COD/l	1160	300	mg COD/l/d	0	0
X_S	mg COD/l	60	100	mg COD/l/d	-500	500
X_{BH}	mg COD/l	2400	500	mg COD/l/d	150	500
X_{AOB}	mg COD/l	128	20	mg COD/l/d	20	20
X_P	mg COD/l	420	80	mg COD/l/d	58	5
S_O	mg COD/l	2	2	mg COD/l/d	-800	800
S_{NO3}	mg N/l	12	5	mg N/l/d	0	200
S_{NH}	mg N/l	6	6	mg N/l/d	-80	100
S_{ND}	mg N/l	0.6	0.2	mg N/l/d	-20	60
X_{ND}	mg N/l	5	3	mg N/l/d	-40	50
S_{ALK}	mol /m^3	4.5	1	mol/m^3/d	0	30
S_{NO2}	mg N/l	0.1	0.2	mg N/l/d	5	20
S_{NO}	mg N/l	0.005	0.005	mg N/l/d	0.2	1
S_{N2O}	mg N/l	0.004	0.002	mg N/l/d	0.2	1
S_{N2}	mg N/l	13.5	1	mg N/l/d	60	100
X_{NOB}	mg COD/l	44	2	mg N/l/d	5	5

Loss function: The mean absolute error (MAE) served as the loss function throughout the experiments, guiding the optimization process towards minimising the absolute difference between predicted and actual values.

Gradient descent optimisation: Training was conducted using the adaptive moment estimation (Adam) optimiser (Kingma and Ba, 2014), an advanced variant of stochastic gradient descent (SGD) known for its adaptive learning rate capabilities. The configuration of the Adam optimizer was set with a gradient decay factor of 0.9, a squared gradient decay factor of 0.999, and a learning rate of 0.01, consistent with common practices.

Iteration, batch size and NODE steps: The number of training iterations depend on the algorithm, computational efficiency, and desired error threshold**.** In this case, custom loops were implemented to manage the training process. The training consisted of two stages: a collocation training stage utilising 3000 iterations, followed by a direct NODE training stage for 1000 iterations. Batch size is not suggested below the 10% of the total discretised trajectory points. Longer NODE steps improve prediction accuracy but increase solver difficulty. A batch size of 20% was employed and a single time step was used within the NODE solver, balancing accuracy with computational limitations. While larger batch sizes and extended time steps could improve accuracy, they exponentially increase computational demands.

6.3.2 Experiments

Four sets of experiments were conducted to evaluate the proposed algorithms and methods:

1) Prediction ability assessment: This study examined the prediction accuracy of the trained model on identical dry weather scenario.
2) Cross-scenario performance: This study assessed the model's generalisation ability by testing its performance on three different weather scenarios after training it with only the first twelve days of rainy data.
3) Reverse scenario validation: This study tested how a model, which were trained with all dry weather data, performed in rain weather scenario, and vice versa, how a model, which were trained with all rain weather data, performed in dry weather scenario.

4) Reduced dimension modelling: This study explored the feasibility of using the model with potentially incomplete data. A NODE model with reduced dimensions was trained and tested, acknowledging that not all component data in the model may be available in real-world scenarios.

Selecting an optimal metric for model evaluation presents challenges, and there is no single perfect metric. While this study employed Root Mean Squared Error (RMSE) and coefficient of determination (R-squared) for assessment, these metrics have limitations. They can be informative and provide valuable insight, but also misleading, particularly when comparing results derived from datasets with different lengths.

6.4 Results

6.4.1 Prediction on identical scenario

A NODE model was trained using data from the first seven days of the dry weather scenario. Figure 6.2 and Figure 6.3 depict the training loss curves for the collocation and direct NODE training stages, respectively, showcasing a steady decrease in error.

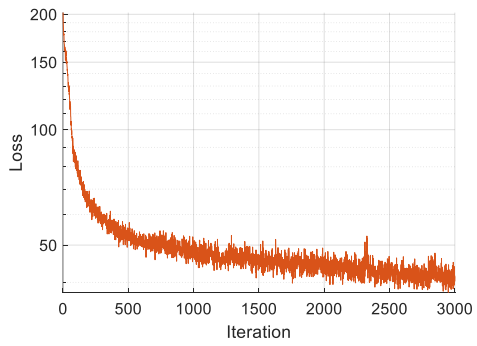


Figure 6.2 Collocation training loss curve

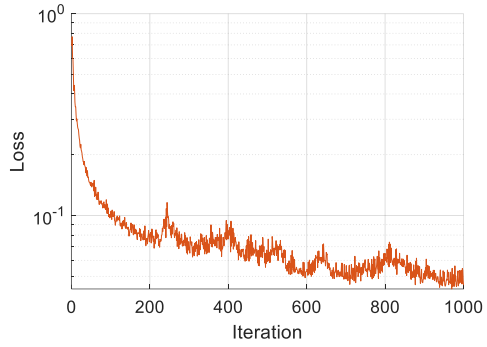


Figure 6.3 NODE training loss curve

The trained model was then used to predict system behaviour for the subsequent seven days (days 7-14) on the same dry weather scenario. Figure 6.4 (see full results in appendix 8.2) compares the predictions with the original simulated data, showing that the model generally captured the dynamic trends, including N_2O generations across all five reactors. Notably the model performed best for reactors four and five.

Figure 6.4 Testing in predication of day 7 - 14 in dry weather scenario

Table 6.2 summarise the RMSE for both training and testing stages. The overall RMSE for the training stage (days 0-7) was 1.18, while for the testing stage (days 7-14), it was 1.11, indicating good agreement between predictions and actual data and consistent performance across both phases.

Table 6.2 Summary of RMSE in dry weather scenario

Stage	Day	Reactor 1	Reactor 2	Reactor 3	Reactor 4	Reactor 5	Overall
Training	0-7	2.58	0.36	0.31	0.20	0.15	1.18
Testing	7-14	2.43	0.39	0.32	0.21	0.16	1.11

We then used the trained NODE model to assess plant performance for days 7-14 under dry weather conditions. Table 6.3 presents the results for N_2O emissions during

nitrification/denitrification, comparing the NODE model predictions with those of the ASMG1 model. The NODE model predicted slightly lower N_2O emissions in anoxic reactors and higher emissions in oxic reactors, resulting in a 2.44% higher total N_2O emission prediction compared to the ASMG1 model.

Table 6.3 Assessment of N_2O emissions for day 7-14 of dry weather scenario

Reactor	N_2O emissions (kg N-N_2O/d)		N_2O emissions (kg N_2O/d)		difference
	ASMG1 model	NODE model	ASMG1 model	NODE model	
Anoxic 1	0.010455	0.010418	0.016429	0.01637	-0.35%
Anoxic 2	0.011148	0.011089	0.017518	0.017426	-0.53%
Oxic 1	1.2195	1.2514	1.9164	1.9665	2.62%
Oxic 2	0.60874	0.62763	0.9566	0.98628	3.10%
Oxic 3	0.2772	0.27856	0.4356	0.43774	0.49%
Total	2.1271	2.1791	3.3426	3.4243	2.44%

It's important to note that the dissolved oxygen (S_O) levels in anoxic reactors one and two did not visually align well with the observed data. This discrepancy stems from the significant scale difference in dissolved oxygen levels between anoxic and oxic tanks. The S_O in reactor two ranged from 0 to 6 x 10^{-3} mg/L, while reactor five exhibited levels from 0 to 6 mg/L – a thousandfold difference. This scale disparity exceeded the precision achievable by the loss function, given the limited number of training iterations used in the experiments.

In practical wastewater treatment applications, on-site dissolved oxygen measurements are typically around to a precision of 0.1 mg/L due to instrument limitations and wastewater inhomogeneity (Roman M. D., 2014). Given these real-world constraints, this study did not prioritise improving dissolved oxygen level predictions in anoxic tanks for this study, as further refinement would offer minimal practical benefit.

6.4.2 Cross-scenario performance

To evaluate the model's generalization ability, a new model was trained using the first 12 days of data from rain weather scenario, which included a rain event between days

8 and 12. Then this model's performance was assessed on days 12 to 14 across three weather scenarios.

Figure 6.5, Figure 6.6 and Figure 6.7 (see full results in appendix 8.3) illustrate the model's performance on day 12 to 14 in dry, rain, and storm weather scenarios, respectively. The results demonstrate that the trained model generalizes well across various conditions, indicating its broad applicability to unseen data. Notably, the model effectively captured the underlying system dynamics, including accurate N_2O production, across all scenarios.

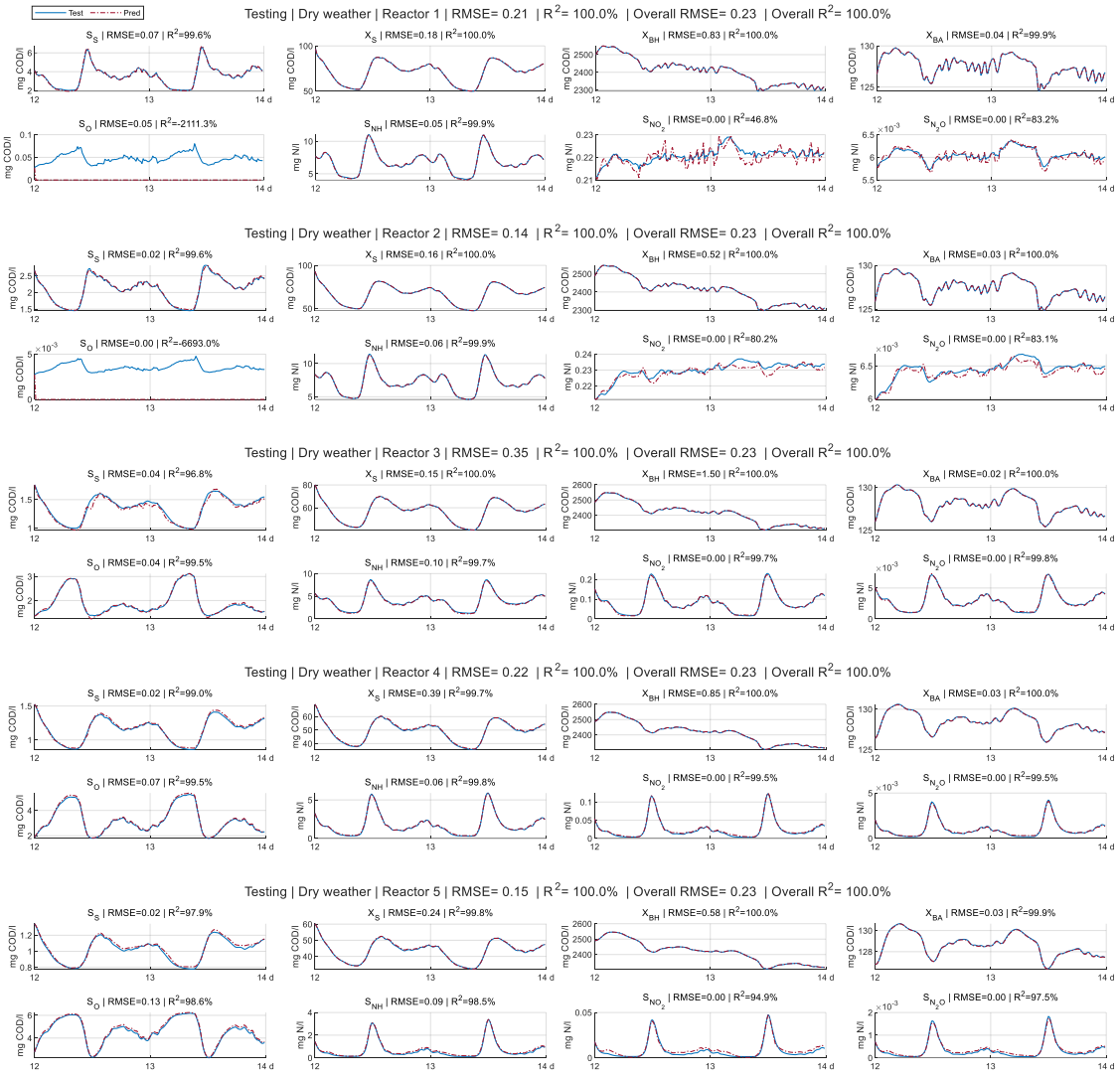


Figure 6.5 testing in dry weather scenario

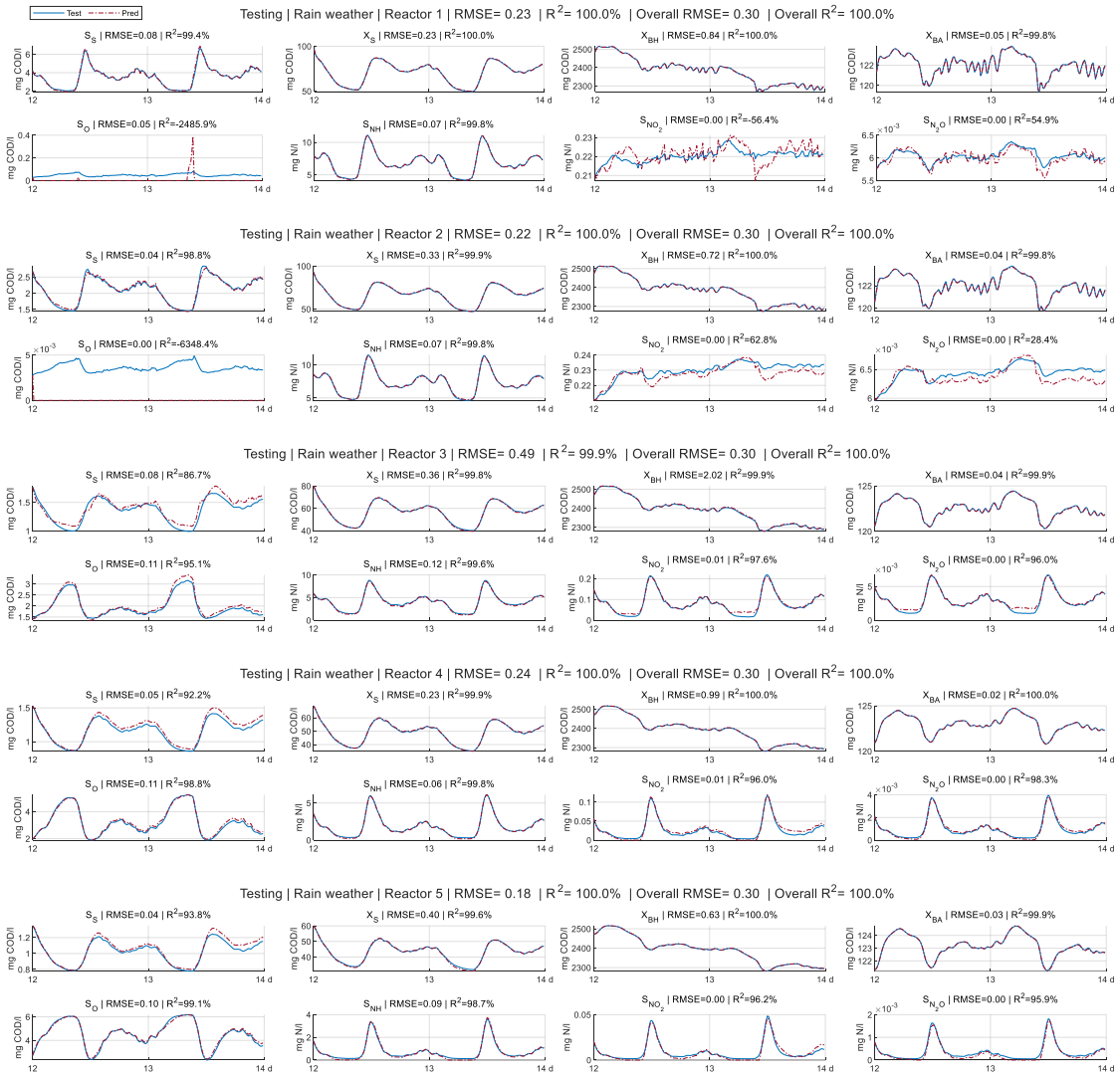
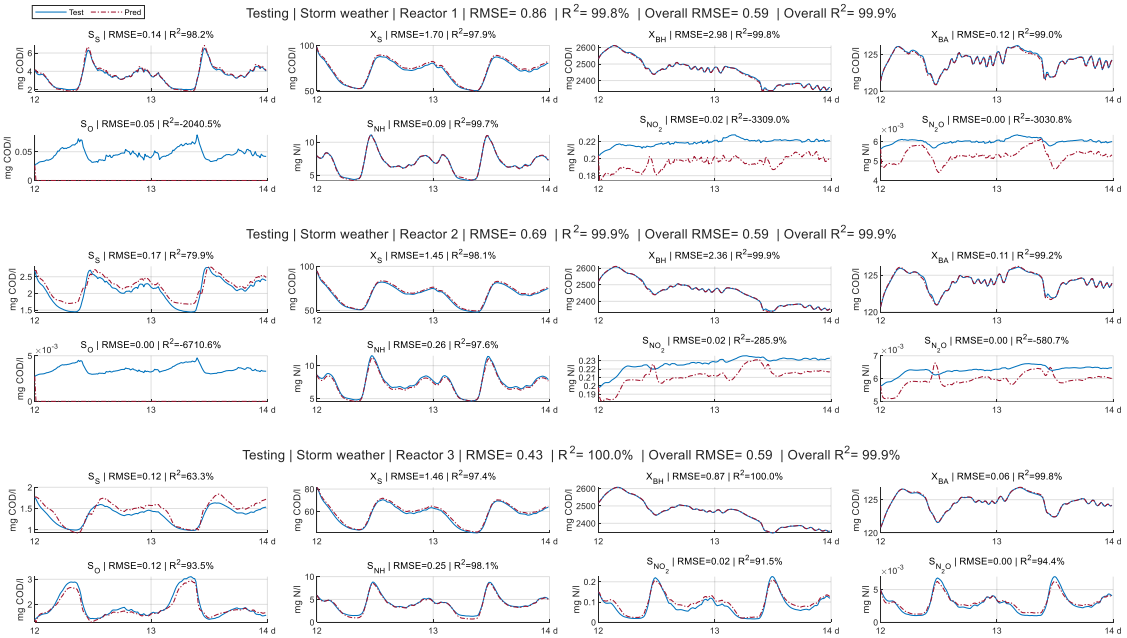


Figure 6.6 testing in rain weather scenario



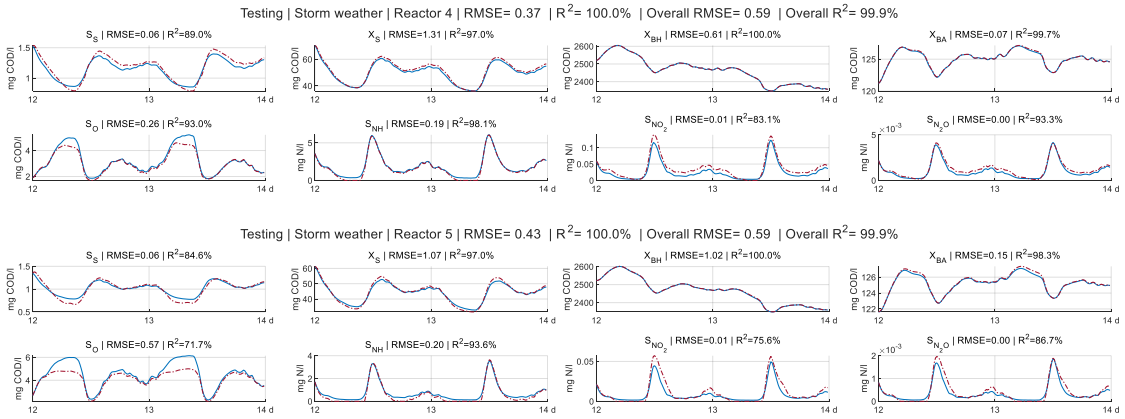


Figure 6.7 testing in storm weather scenario

Table 6.4 summarises the RMSE values for the model's performance under dry, rainy, and stormy weather conditions. The general trend indicates that the model performs best in reactors four and five across all scenarios. However, there are no statistically significant differences between reactors or scenarios. This indicates the model's broad applicability to various weather conditions.

The overall RMSE values for the dry, rain, and storm scenarios were 0.23, 0.30, and 0.59, respectively. While the model's performance slightly degraded in more extreme weather conditions, it maintained reasonable accuracy across all scenarios. This robust performance across varied weather conditions underscores the model's potential for practical applications in wastewater treatment processes, where adaptability to changing environmental conditions is crucial.

Table 6.4 RMSE of the model testing in dry, rain and storm scenarios

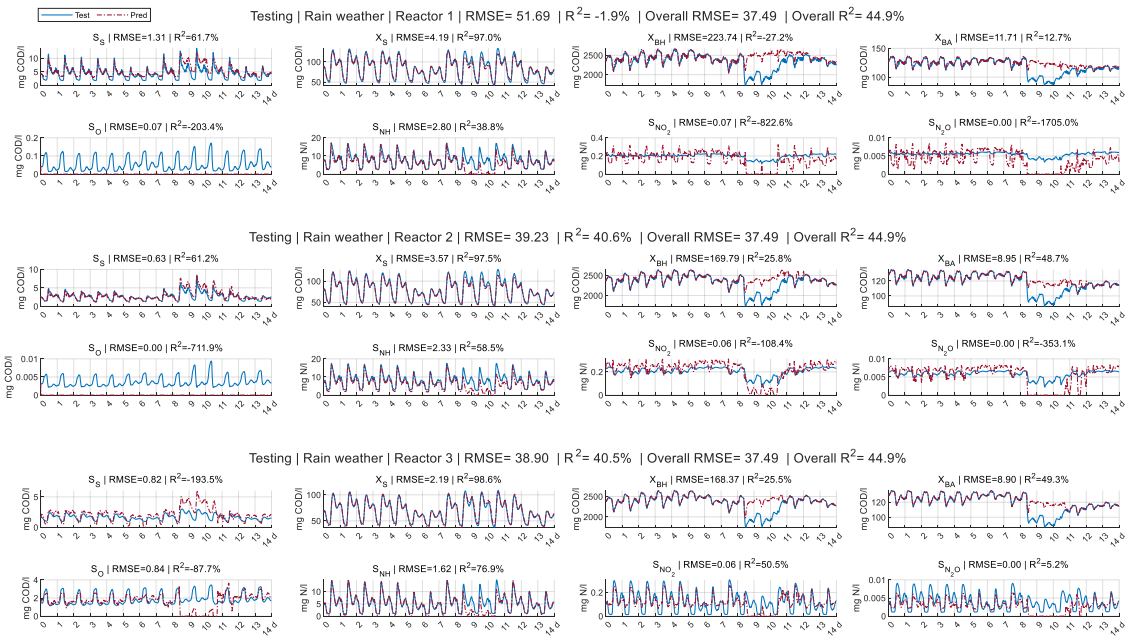
Scenario	Day	Reactor 1	Reactor 2	Reactor 3	Reactor 4	Reactor 5	Overall
Dry	12-14	0.21	0.14	0.35	0.22	0.15	0.23
Rain	12-14	0.24	0.22	0.49	0.24	0.18	0.30
Storm	12-14	0.86	0.69	0.43	0.37	0.43	0.59

6.4.3 Reverse-scenario validation

We evaluated the model's generalisation ability through reverse scenario validation. Two separate models were trained: a dry weather model tested in rain weather scenario, and a rain-weather model tested in dry weather scenario.

- **Dry Weather Model**: This model was trained using data from all 14 days in the dry weather scenario, representing the system's typical operation. These data excluded disruptions caused by extreme events such as rain or storms. The model was then tested on rain weather data, particularly focusing on its performance during the rain event period between day 8 to day 12.
- **Rain Weather Model:** This model was trained using data from all 14 days in the rain weather scenario. These data encompassed both normal operating conditions (day 1 to 8) and interruptions caused by rain (day 8 to 12). Subsequently, the model was tested on dry weather data to assess its prediction accuracy.

Figure 6.8 (see full results in appendix 8.4) illustrates that the dry weather data trained model performed well for normal operation conditions (day 1 to 8) in the rain weather scenario. However, its performance faltered during the rain event period (day 8 to 12). This is reflected by a drop in the overall R-squared value to 44.9% and an increase in the overall RMSE to 37.49. Conversely, as shown in Figure 6.9 (see full results in appendix 8.4), the rain weather data trained model maintained good performance throughout the entire dry weather scenario. This is evident from the high overall R-squared value of 99.7% and the low RMSE of 1.44.



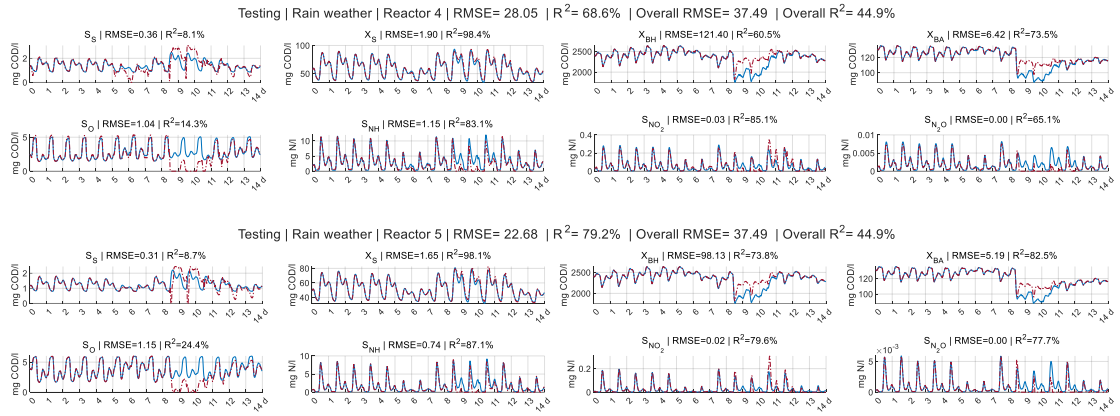


Figure 6.8 testing dry weather data-trained model in rain weather scenario

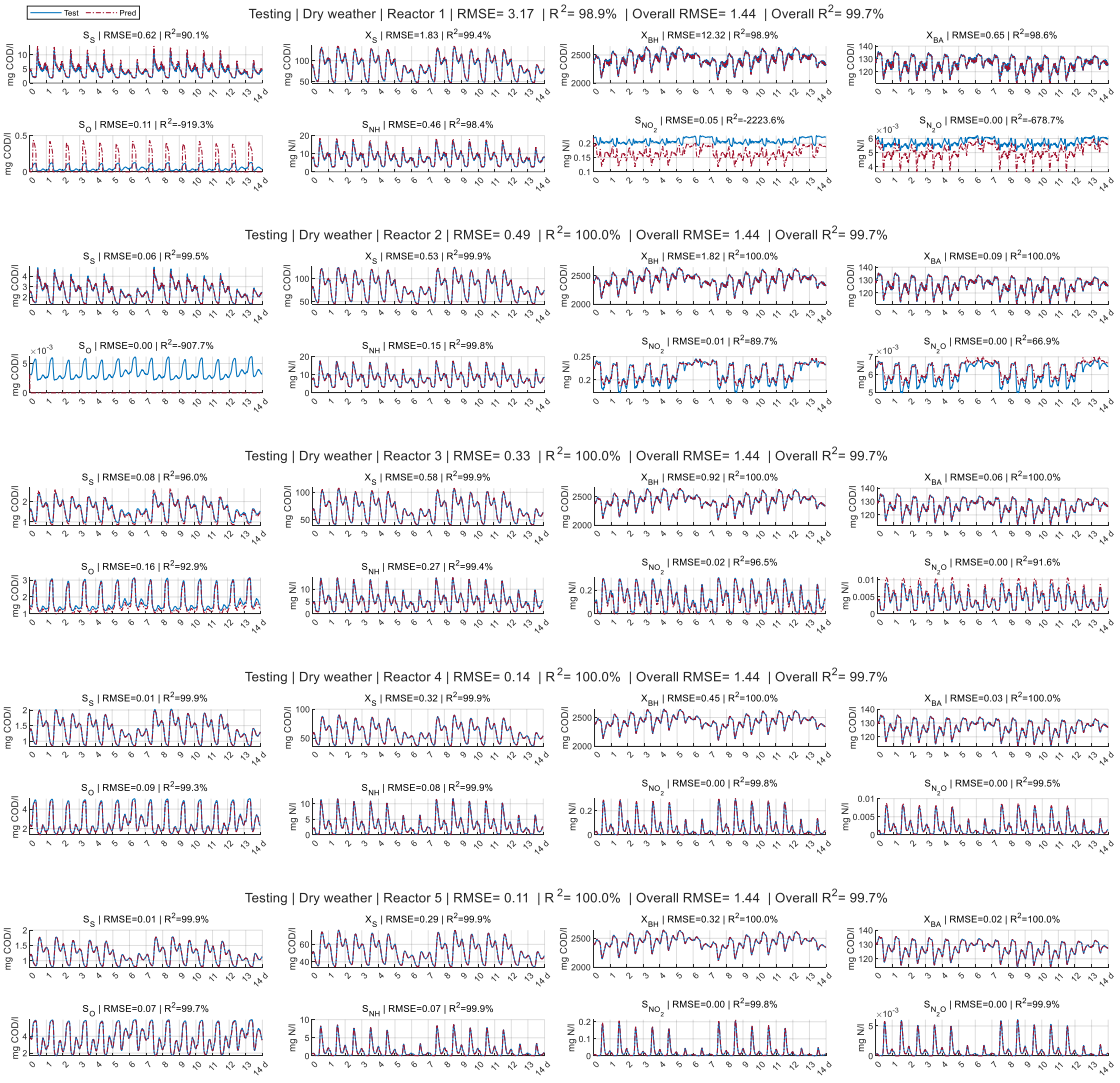
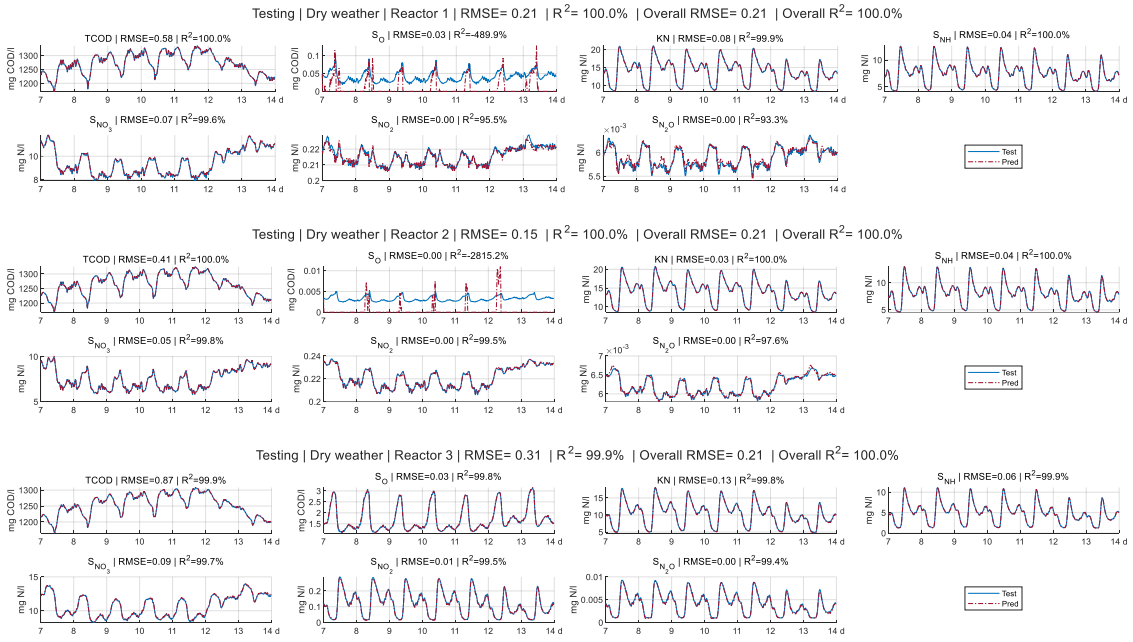


Figure 6.9 testing rain weather data trained model in dry weather scenario

6.4.4 NODE model with reduced dimensions

In real-world applications, directly measuring all components used in the ASMG1 model can be challenging. To address this limitation, a NODE model with reduced dimensions was explored. Four carbonaceous substrates were consolidated into total COD (T_{COD}), and ammonia, both soluble and particulate organic nitrogen, were combined into Kjeldahl nitrogen (KN). Key components crucial to N_2O generation (Tchobanoglous *et al.*, 2014), including dissolved oxygen (S_O), ammonium (S_{NH_4}), nitrate (S_{NO_3}), and nitrite (S_{NO_2}), were retained in the model. Components deemed less relevant (such as alkalinity) or those exhibiting relative stability (e.g., biomass) were excluded from the model.

The lower-dimensional model was trained using the first seven days of dry weather data and tested for predictions in the subsequent week. Figure 6.10 (see full results in appendix 8.5) demonstrates that the model effectively captured the dynamic trends of N_2O emissions despite the significant reduction in input dimensions. This suggests the potential of using such simplified models for practical applications where complete data acquisition might be difficult.



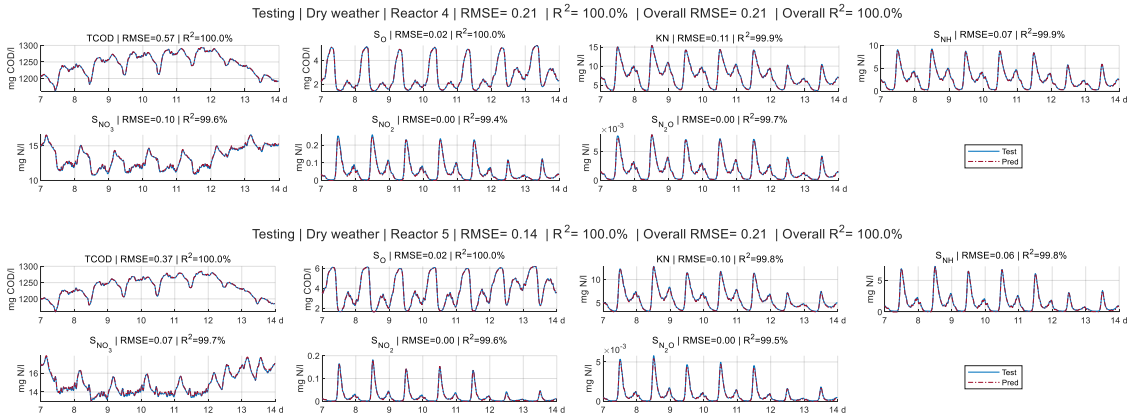


Figure 6.10 testing of model with reduced dimensions

While this study utilised a limited 14-day data timeframe, a small fraction of a typical wastewater treatment plant's operational timeline, it effectively demonstrates the feasibility of the NODE approach for data-driven modelling of wastewater processes. Notably, the model captured the dynamic behaviour of N_2O productions across different treatment stages, despite their significant variations.

Further refinement of the model is possible by exploring tuning hyperparameters like batch size, number of node steps, and training iterations. Additionally, estimation of normalisation parameters (mean and standard deviation) can be enhanced by combining domain knowledge with site-specific data from the wastewater treatment plant. These enhancements could potentially lead to improved model accuracy.

These findings suggest that the NODE approach, particularly with reduced dimensions, holds promise for practical applications in wastewater treatment process modelling. By balancing model complexity with data availability, this method could provide valuable insights into N_2O productions and other critical parameters, even in scenarios where comprehensive measurements are challenging to obtain.

6.5 Summary

A NODE-based model for N_2O production was implemented and tested under various operational scenarios at the BSM1 plant. The model demonstrated robust predictive capabilities for N_2O behaviour across both anoxic and aerobic phases, accurately capturing even minute and transient N_2O variations. Moreover, the model proved resilience to short-term operational disturbances and effectively predicated system responses under varying weather scenarios. These achievements are attributed to the

novel algorithm developed for handling exogenous factors, and the data normalization and incremental training strategy proposed for tackling model stiffness.

The results underscore the potential of NODE models to accurately represent the complex dynamics of wastewater treatment processes and highlight its promise for optimising N_2O mitigation strategies. Nevertheless, given the multifaceted nature of N_2O production pathways and its spatial and temporal variability, comprehensive real-world validation is imperative to refine the model and facilitate its broader application in practical wastewater treatment.

Chapter 7 Discussion and conclusion

The preceding chapters comprehensively explored data-driven N_2O modelling using NODEs, with a particular focus on addressing training instability arising from wastewater process stiffness. Through rigorous experimentation and analysis, valuable insights have been gained into NODE performance for capturing N_2O production and emission patterns from monitoring data. This chapter delves deeper into the implications of these findings, situating them within the broader research context. By critically examining the strengths, limitations, and potential impact of this study, it is expected to provide a comprehensive discussion and draw meaningful conclusions, with summary of key contributions and outline of future research directions.

7.1 Discussion

The development and successful deployment of robust and effective machine learning models, such as NODEs, are contingent upon a confluence of factors that collectively influence model performance. While advancements in algorithm design and computational resources have propelled the field forward, several critical challenges persist. This section discusses four key areas significantly impacting NODE training and generalization: training instability, unforeseen data generalization, data quality and availability, and computational cost. A comprehensive understanding of these challenges is essential for developing effective strategies to mitigate their effects and ultimately enhance model performance.

7.1.1 Training instability

NODEs offer a promising approach to modelling complex systems by representing the dynamics as a continuous process. However, their training is notoriously unstable. This instability arises from several key factors:

- **Stiffness**: Stiff systems exhibit widely varying timescales, making them challenging to solve numerically. NODEs often encounter stiffness, particularly when modelling complex systems. Standard solvers, like Euler or Runge-Kutta, struggle with stiff problems, leading to instability (Kim *et al.*, 2021).

- **Gradient explosion/vanishing**: Similar to traditional neural networks, gradients can explode or vanish during backpropagation through the ODE solver. This can hinder convergence and destabilize training (Amari, 1993).
- **NODE solver choice**: The choice of NODE solver significantly impacts stability. Implicit solvers are generally more stable for stiff problems but can be computationally expensive, while explicit solvers are faster but prone to instability for stiff systems. Unfortunately, NODE solvers are less developed compared to traditional ODE solvers (Kushnir and Rokhlin, 2012; Postawa, Szczygieł and Kułażyński, 2020).
- **Initialization**: Poor initialization of the neural network parameters can lead to unstable trajectories and divergent solutions (Glorot and Bengio, 2010).

While sources of instability are multifaceted, stiffness is the most prominent and challenging issue in wastewater systems (Brown *et al.*, 2021; Bradley *et al.*, 2022). The results of this study demonstrate the proposed normalisation method effectively addresses instability and is easily implementable. Similar scaling method have also shown promise in the literature (Ji *et al.*, 2021; Kim *et al.*, 2021). However, other approaches with potential to mitigate stiffness exist, two have been identified for future exploration:

- **Gradient clipping**:

  Gradient clipping is a regularization technique that addresses training instability by controlling the magnitude of gradients during backpropagation. If the norm of the gradient exceeds a predefined threshold, it's scaled down to meet that threshold. By imposing a threshold on the gradient norm, it prevents gradients from exploding, thereby enhancing stability. While commonly used in neural networks, this method is equally applicable to NODEs (J. Zhang *et al.*, 2019).

  However, implementing gradient clipping requires careful parameter tuning. The clipping threshold, frequency, and choice of norm (e.g., L1, L2) significantly impact performance. Finding optimal values for these parameters can be challenging. Additionally, aggressive clipping may lead to information loss (Qian *et al.*, 2021; Koloskova, Hendrikx and Stich, 2023).

- **Specialized solver**:

While NODE and ODE solvers share similarities, they also exhibit distinct characteristics. Unfortunately, the development of NODE solvers, especially for stiff systems, remains limited. Developing specialized solvers tailored for NODEs, such as those incorporating adaptive step size methods (e.g. Dormand-Prince) (Kimura, 2009; Seen, Gobithaasan and Miura, 2014; Chalvidal *et al.*, 2020; Zhuang *et al.*, 2020; Kloberdanz and Le, 2023) or implicit schemes (Poli *et al.*, 2021; Baker *et al.*, 2022; Pal, Edelman and Rackauckas, 2022), can significantly enhance stability.

Ideal NODE solvers should prioritize the following objectives:

1) **Efficiency**: The solver should be computationally efficient to handle large datasets and complex models.
2) **Stability**: The solver should be robust to the inherent instability of neural networks.
3) **Accuracy**: The solver should accurately approximate the solution of the ODE to ensure correct gradient computation.
4) **Differentiability**: The solver should be differentiable to enable backpropagation.

Based on the experience from ODE solver, the types of potential NODE solver may include:

1) **Adaptive step size solvers**: Dynamically adjust the step size based on the estimated error, improving efficiency and stability. Examples include Dormand-Prince (Kimura, 2009), Runge-Kutta-Fehlberg (Seen, Gobithaasan and Miura, 2014).
2) **Implicit solvers**: Effectively handle stiff ODEs, which can be applied in NODEs. However, they typically require iterative solution methods, increasing computational cost. Examples include Backward Euler (Skelboe and Andersen, 1989; Biswas *et al.*, 2013), Implicit Runge-Kutta methods (Cartwright and Piro, 1992).
3) **Symplectic solvers**: Preserve the symplectic structure of Hamiltonian systems (Leimkuhler and Skeel, 1994; Zhong, Dey and Chakraborty, 2019), which can be beneficial for certain types of physical systems. Examples

include Leapfrog method (Shampine, 2009), Verlet integration (Sanz-Serna, 1992; Murua, 1999).

4) **Hybrid solvers**: Combine the strengths of different solvers to achieve better performance. For example, using an implicit solver for stiff regions and an explicit solver for non-stiff regions (Shi *et al.*, 2012; Tutueva *et al.*, 2020).

While gradient clipping offers a partial solution, specialized NODE solvers hold the most promise. By providing theoretical guarantees for stability and convergence of learning process, they can significantly advance the field of NODEs and expand their applicability.

7.1.2 Out-of-distribution generalization

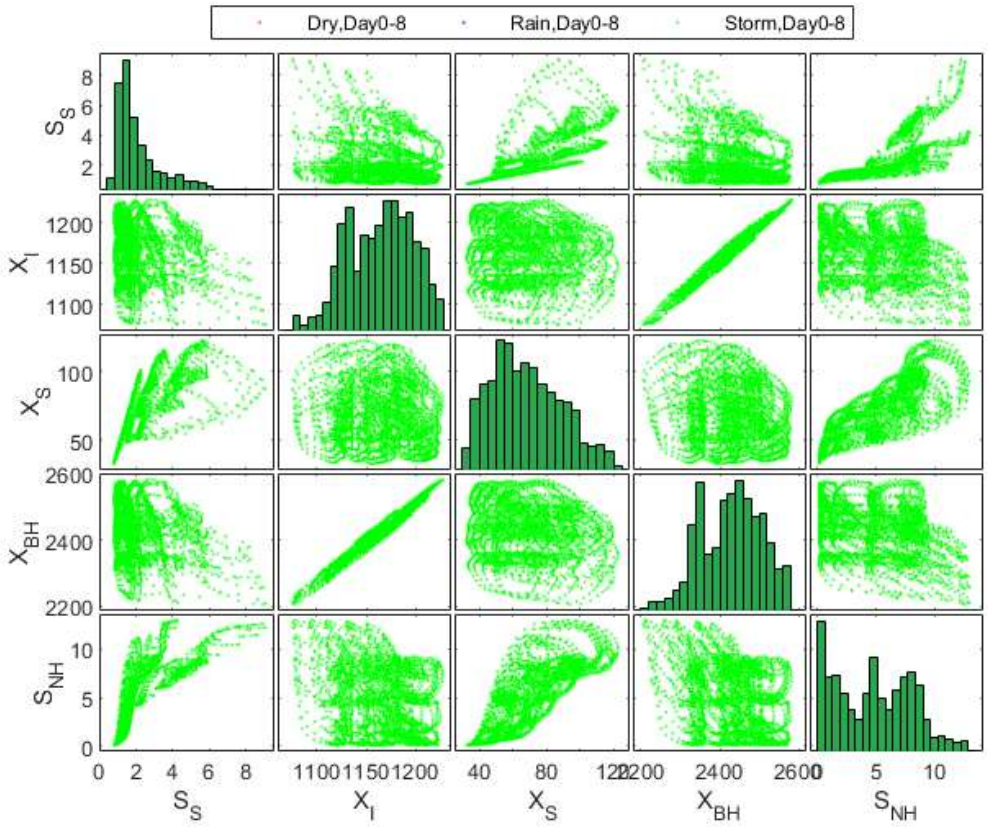
Traditional machine learning models assume that training and test data are statistically similar, a condition known as independent and identically distributed (IID) (Cao, 2022). However, real-world data often deviates from this assumption, undergoing unexpected distributional changes. As a result, deployed models frequently exhibit significant performance degradation, a phenomenon known as out-of-distribution generalization problem (Hendrycks *et al.*, 2021; Ye *et al.*, 2021). NODEs, which rely on DNN to approximate dynamics, are similarly vulnerable to OOD challenges.

This problem arises from distribution shift, where the statistical properties of training and test data differ (Fang *et al.*, 2020). This discrepancy can be attributed to factors such as data scarcity or non-representative data selection during training.

Leveraging neural networks as their core component, NODEs inherit both their capabilities and limitations (Elbrächter *et al.*, 2019). Like other data-driven models, NODEs heavily rely on the quality and diversity of training data. Comprehensive data coverage is more critical than sheer data volume, as diverse scenarios within the data enhance the model's ability to generalise to unseen situations. Insufficient data coverage, where training data doesn't encompass the full range of possibilities, can lead to poor generalizability and unreliable predictions (Zhang *et al.*, 2021).

The results of this study exemplify this. Figure 7.1 illustrate the joint distribution of key components for dry, rain, and storm weather scenarios during two distinct periods: day 0-8 and day 8-12, respectively. While the initial 8-day period shows no discernible differences, a clear divergence among weather scenarios emerges during days 8-12,

coinciding with rain and storm events. the reverse scenario validation results (section 6.4.3) corroborate these findings. The dry weather data-trained model struggled to handle actual rain events due to the absence of rain information in its training data. Conversely, the rain weather data-trained model performed adequately in dry weather scenarios, as it encompassed dry weather conditions.



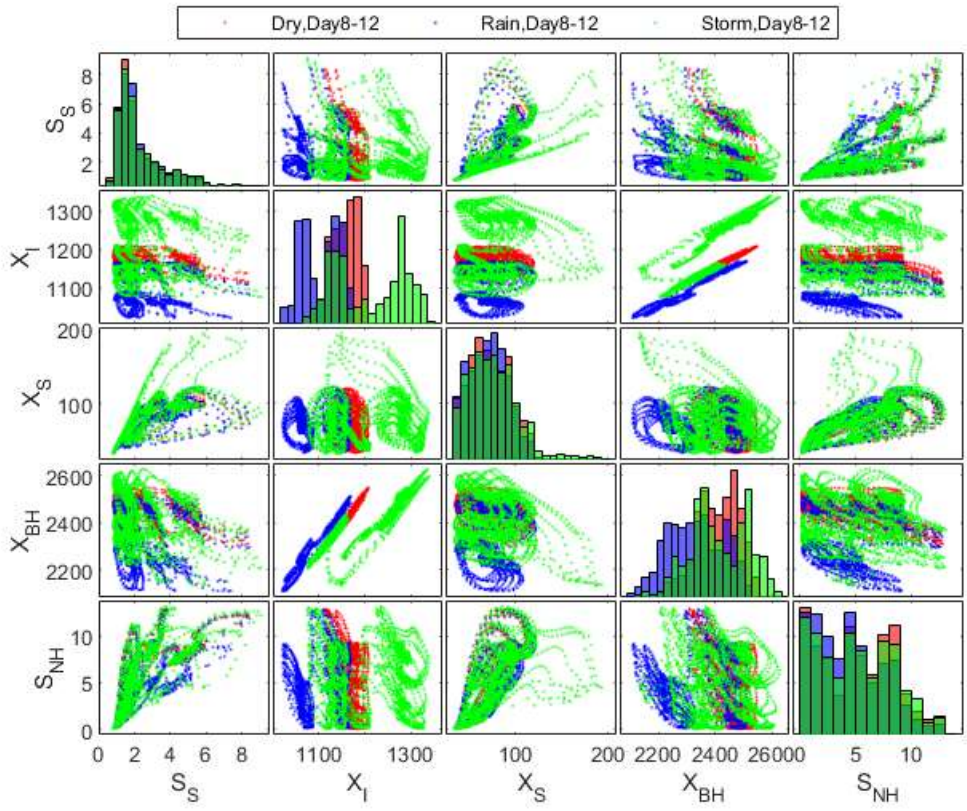


Figure 7.1 Joint distribution of dry, rain, and storm weather data for day 0-8 (left) and day 8-12 (right)

Solutions for OOD generalization issue can be from different aspects:

- **Data-centric approaches**

  Augmenting training data with synthetic samples can mitigate data scarcity (Dupont, Doucet and Teh, 2019). Expanding the training data distribution to encompass diverse scenarios improves model robustness. OOD detection techniques, which identify data points outside the training distribution, can aid in data augmentation(Cui and Wang, 2022). Integrating wastewater theory and practice is crucial for ensuring that training data covers the full spectrum of potential wastewater conditions. Rigorous data cleaning and preprocessing are essential for data quality and consistency (Chu *et al.*, 2016).

- **Model-centric approaches**

  Exploring network architectures inherently resistant to distribution shifts is vital. Regularization techniques such as L1/L2 regularization (Girosi, Jones and Poggio, 1995), dropout (Baldi and Sadowski, 2013), and early stopping (Mahsereci *et al.*,

2017) can help prevent overfitting. Adversarial training enhances model robustness by exposing it to adversarial examples (Rani *et al.*, 2024). Meta-learning (Finn *et al.*, 2019; Hospedales *et al.*, 2021), continuous learning (Liu, 2017), and transfer learning (Weiss, Khoshgoftaar and Wang, 2016) improve long-term generalization by enabling the model to adapt to evolving data distributions. For instance, in wastewater treatment, equipment degradation, microbial changes, and operational variations can induce distribution shifts. Regularly updating the model with fresh data is essential to maintain performance.

The optimal strategy for OOD generalization depends on the specific problem and dataset. A combination of data-centric and model-centric approaches is often necessary to achieve optimal performance.

7.1.3 Data availability and quality

Data is the fuel that powers machine learning models, including NODEs. Both data quantity and quality significantly impact model performance, accuracy, and reliability (McDonald, 2021; Budach *et al.*, 2022; Yue Liu *et al.*, 2023). The adage "garbage in, garbage out" is particularly applicable to data-driven models. High-quality data enables models to learn accurate patterns and relationships, leading to reliable and trustworthy predictions. Conversely, poor data quality can introduce biases, resulting in inaccurate or unfair outcomes.

Obtaining qualified data in wastewater treatment is crucial for effective process monitoring, control, and optimization. Data collection methods in WWTP include:

1) **Sensor-based monitoring**: Modern WWTPs increasingly utilize sensors to measure various parameters such as pH, temperature, dissolved oxygen, solids, and chemical concentrations. Regular sensor calibration and continuous data logging are essential for accurate monitoring. To account for rapidly changing wastewater characteristics due to factors like rainfall, industrial discharges, and seasonal variations, outlier and anomaly detection techniques can be employed (Zamora and Torres, 2014; Haimi *et al.*, 2016). Given the potential for sensor drift over time, implementing calibration schedules and sensor replacement plans is crucial.
2) **Laboratory testing**: Regular laboratory testing of wastewater samples is crucial for WWTPs to monitor parameters such as BOD, intermediate products, inhibitory

compounds, and fractionated components, which are often difficult to measure using online sensors (Pagga, Bachner and Strotmann, 2006). In addition, parallel testing data for normal parameters provide a benchmark for sensor calibration and validation. Adherence to stringent quality control procedures is essential to ensure the reliability of laboratory analysis results.

3) **Process control systems**: Process control systems provide valuable data on air flow rates, pump speed, and valve positions, enabling in-depth analysis of process behaviour and optimization of control strategies (Moles *et al.*, 2003).

Data-driven approaches often demand substantial data volumes, which can be challenging to obtain solely through laboratory testing (Schneider *et al.*, 2021). Consequently, the reliability of online sensors and the quality of generated data are critical for successful data-driven initiatives in WWTPs. Real-world plants often encounter sensor discrepancies, drifts, and malfunctions due to technical limitations, fouling, or harsh operating conditions, compromising data accuracy. While techniques such as mass balance calculations, statistical analysis, additional laboratory validation, and anomaly detection can mitigate these issues, prioritizing sensor maintenance, calibration, and testing is fundamental to ensure data integrity (Luca *et al.*, 2019; Zhang, Tooker and Mueller, 2020).

Surrogate parameters offer an alternative avenue for addressing the data requirements of emerging AI applications in wastewater treatment (Edzwald, Becker and Wattier, 1985). By developing cost-effective and reliable sensors employing technologies such as light, electronics, and wave-based measurements, it is possible to generate a richer dataset for wastewater characterization. While these surrogate parameters may not directly correspond to traditional analytical components, they can provide valuable insights for enhancing operational efficiency in WWTPs.

Data preprocessing and cleaning are essential to minimize the impact of noise and outliers on model reliability (Chu *et al.*, 2016). Techniques such as outlier detection (Corominas *et al.*, 2011), noise reduction (Moravec *et al.*, 2021), and smoothing methods (Ilyas and Rekatsinas, 2022) should be employed to ensure the model learns true underlying patterns rather than anomalies.

By following these steps and addressing the specific challenges in data collection, high-quality data can be obtained to support effective data-driven process optimization.

7.1.4 Model reliability

Model reliability is influenced by multiple factors. In real WWTPs, it is primarily influenced by data quality and coverage challenges. For example, although wastewater theories highlight the critical role of microbial communities and substrate composition in reaction kinetics, technological limitations prevent real-time measurements of these parameters. While surrogate parameters offer partial solutions, the inability to fully characterize the wastewater system state hinders the application of NODE models (Sundui *et al.*, 2021).

The complex nature of wastewater treatment processes demands a comprehensive dataset encompassing diverse influent characteristics, biological factors, and plant operational parameters to accurately capture system dynamics. To fully account for seasonal fluctuations and temperature-dependent processes, a dataset spanning at least one year is recommended. Practical implementation requires careful consideration of operational profiles and potential model limitations under specific conditions, aiding by investigation and data distribution analysis to identify possible edge cases where the model might struggle (Wang *et al.*, 2021).

Assuming optimal data quality, coverage, and out-of-distribution robustness, NODE models can demonstrate reliable performance. While model predictions may still be a subset of the mechanistic model due to data availability, they can suffice for routine operations under stable conditions. To enhance model reliability, integrating domain knowledge is essential. By assigning physical meaning to model parameters and incorporating physical laws and constraints, the model can improve the plausibility of predictions (Faisal *et al.*, 2023). A combination of comprehensive data coverage, domain expertise, extended temporal datasets, operational considerations, and robust data preprocessing empowers NODE models to deliver reliable nitrous oxide production predictions in complex and dynamic wastewater treatment environments.

7.1.5 Computational cost

NODE models, due to their continuous dynamics modelling approach, incur significant computational costs. The experiments in this study using 14 days of data for the BSM plant modelling demonstrated NODE models exhibit an averagely 50-fold increase in training time compared to a collocation training under identical computational resources.

Several factors contribute to the computational expense of NODEs:

- **Numerical integration**: Solving ODEs numerically requires iterative computations, which can be computationally intensive, especially for complex ODE systems. The choice of solver (e.g. Runge-Kutta) and step size significantly impacts the computational cost (Chalvidal *et al.*, 2020; Dong *et al.*, 2020).
- **Backpropagation**: The backpropagation process through the ODE solver is complex and computationally demanding due to the non-differentiable nature of solver operations (Baker *et al.*, 2022).
- **Memory consumption**: Storing intermediate states of the ODE solver for backpropagation can lead to high memory usage, especially for long time horizons or complex systems (Finlay *et al.*, 2020).

Solutions to reduce NODE computational cost include:

- **Efficient NODE solvers**: Adaptive step size solvers can optimize computational efficiency by adjusting step size based on solution behaviour (Zhuang *et al.*, 2020). Alternatively, less accurate but computationally cheaper solvers can be employed during initial training stages or when precision is less critical (Roesch, Rackauckas and Stumpf, 2021). Specialized solvers designed for stiff systems can also improve efficiency (Kushnir and Rokhlin, 2012; Kloberdanz and Le, 2023).
- **Approximation techniques**: Checkpoint-based saving of intermediate solver states can reduce memory consumption during backpropagation (Zhuang *et al.*, 2020). Discretizing the continuous NODE into a discrete-time system through finite differences can simplify computations but may introduce approximation errors (Kloberdanz and Le, 2023).
- **Hardware acceleration**: Leveraging the parallel processing capabilities of GPUs (graphics processing units) or specialized hardware like TPUs (tensor processing units) can significantly accelerate numerical integration and backpropagation calculations (Shi *et al.*, 2012; Seen, Gobithaasan and Miura, 2014).
- **Model architecture optimization**: Simplify the neural network component of the ODE to reduce computational overhead (Golovanev and Hvatov, 2022).

The choice of solutions depends on the specific problem and desired accuracy. Often, trade-off have to be considered between computational efficiency and the accuracy of

the solution (Khalil *et al.*, 2024). Certainly, continuous research is expected for new techniques to improve the efficiency to be developed.

By carefully considering these factors and applying appropriate strategies, it is possible to mitigate the computational cost of Neural ODEs and make them more practical for real-world applications.

7.2 Benefits of NODE approach

The NODE approach offers several key advantages for modelling wastewater treatment processes:

1. **Physical interpretation**: NODE models maintain compatibility with the physical meaning inherent in traditional mathematical models. This allows for a more intuitive understanding of the modelled processes (Zou *et al.*, 2024).
2. **Data-driven dynamics learning**: NODEs can implicitly capture complex system dynamics directly from monitoring data without requiring explicit knowledge of underlying physical equations. This makes them particularly well-suited for complex wastewater systems where deriving these equations might be challenging (Zakwan *et al.*, 2023).
3. **Enhanced generalization**: By learning the underlying differential equations, NODE models can generalize better to unseen data compared to models trained only on discrete time points. This allows for more accurate predictions under varying conditions (Garsdal, Søgaard and Sørensen, 2022; Kircher, Döppel and Votsmeier, 2024).
4. **Unveiling system insights**: By analysing the trained neural network, researchers can gain valuable insights into the key factors influencing the system's dynamics. This knowledge can be used to optimize wastewater treatment processes and improve overall efficiency (Zakwan *et al.*, 2023).
5. **Continuous-time modelling**: NODE models naturally handle data that evolve continuously over time, offering more accurate and realistic representation of real-world wastewater treatment processes compared to discrete-time models (Garsdal, Søgaard and Sørensen, 2022).
6. **Flexibility**: NODE models easily handle irregularly sampled time series data, making them adaptable to practical scenarios where uniform measurement is not feasible or available (Esteve-Yagüe and Geshkovski, 2021).

These benefits position the NODE approach as a powerful tool for wastewater process modelling, offering an optimal balance between data-driven insights and physical interpretability. By leveraging these capabilities, researchers and practitioners can develop more accurate, flexible, and insightful models for complex wastewater treatment systems, potentially leading to improved process understanding and optimisation.

7.3 Challenges of NODE approach

Despite its potential, NODE approach faces several significant challenges, particularly when dealing with complex processes like N_2O production:

1. **Training Instability**: The most significant challenge in implementing the NODE approach is the instability of the training process. This issue is particularly pronounced in stiff, complex, and chaotic systems, as well as over long-time horizons. This instability can stem from various factors, including poor initializations, gradient noise, and the inherent complexity of the system itself (Finlay et al., 2020; Golovanev & Hvatov, 2022). Maintaining stability in the numerical integration of ODEs is crucial, as instabilities can lead to divergent solutions and unreliable models. Careful selection of integration methods and step sizes is necessary to ensure stability while balancing accuracy and computational efficiency (Dikeman, Zhang and Yang, 2022).
2. **Computational Cost**: Training NODE models can be computationally expensive compared to traditional machine learning methods. This is partly linked to the need for smaller integration steps to maintain stability, leading to a higher number of calculations (Golovanev and Hvatov, 2022).
3. **Solver Limitations**: While numerous mature ODE algorithms and solver libraries exist for mathematical dynamical systems, NODE solvers capable of handling massive stiffness are still rare. This limitation is particularly evident in real-world applications with complex dynamics (Baker *et al.*, 2022).
4. **Immature Ecosystem**: While recent advancements have seen libraries like PyTorch, TensorFlow, and MATLAB incorporate support for NODEs, the surrounding ecosystem and tooling remain less mature compared to those for conventional neural networks. This relative infancy can result in instability and computational inefficiency, particularly when tackling real-world problems

involving stiff equations or high-dimensional systems within current library frameworks. These limitations can translate into practical implementation hurdles, hindering the practical application of NODEs despite their theoretical promise.

These challenges highlight the need for continued research in several areas:

- Development of more robust and efficient NODE solvers, particularly for stiff and high-dimensional systems.
- Improvement of training stability techniques for complex, chaotic systems.
- Optimization of computational methods to reduce resource requirements.
- Expansion and maturation of the NODE ecosystem and tooling.

The NODE approach holds promise for wastewater treatment process modelling, but addressing these challenges is crucial for its practical implementation. Overcoming these hurdles will pave the way for the broader adoption of NODE models in real-world wastewater treatment applications.

7.4 Conclusion

This study demonstrates the potential of NODE models in capturing the complex and nonlinear dynamics of wastewater treatment processes, with a specific focus on the challenging area of N_2O production. Despite the inherent difficulties in modelling N_2O due to its diverse production pathways and significant spatiotemporal variations, the proposed methodologies empowered the NODE model to effectively represent these intricate and disparate system behaviours.

1) To address the critical challenge of stiffness, this study proposes a novel normalisation method that effectively stabilizes the training process. By enabling smoother gradient descent and balancing optimization across disparate system scales, this readily implementable method facilitates accurate data-driven modelling of wastewater processes.
2) This study introduces an incremental training strategy for NODE models that leverages the efficiency and noise resilience of the collocation method to bypass the initial stiffness hurdle. The initial solution is subsequently refined using the direct NODE approach to enhance accuracy. This combined methodology underscores the importance of adapting modelling techniques to specific problem stages.

3) The developed algorithm effectively addresses the impact of exogenous factors and external influences on the data. This enables the model to learn the underlying intrinsic dynamics from real-world composite data that incorporates both external perturbations and internal process interactions.
4) This study emphasizes the critical role of data quality and comprehensive data coverage in successful data-driven modelling of wastewater processes using NODEs. Additionally, it underscores the necessity of developing specialized solvers tailored to this specific application.

This study demonstrated that the model effectively predicts system behaviour under both anoxic and aerobic conditions, accurately capturing even subtle and transient N_2O fluctuations. Its robustness is further evidenced by its resilience to short-term disturbances and its ability to predict system responses under varying weather conditions. These findings underscore the potential of data-driven NODE modelling for optimizing N_2O mitigation strategies in wastewater treatment plants.

In summary, while acknowledging the challenges of training and hyperparameter tuning, NODE models demonstrate significant potential for wastewater treatment modelling due to their ability to handle complex, continuous-time processes. This study introduces novel methodologies to address key challenges in NODE training, including normalization, incremental training strategy, and algorithms handling exogenous factors. These contributions represent substantial advancements and facilitate broader adoption of NODE models in wastewater treatment plants, enabling more efficient and data-driven management strategies. Future research with extensive real-world validation is essential for further refining and promoting the widespread application of NODE models in practical settings.

Chapter 8 Appendix

8.1 Plant performance under varying weather scenarios

Table 8.1 Plant performance assessment under varying weather scenarios

Performance	Unit	Dry	Rain	Storm
Overall plant performance assessment during period	day	7-14	7-14	7-14
Influent average concentrations				
Flow rate	m^3/d	18446.3875	24201.7981	21047.0253
S_I	g COD/m^3	30	22.8657	26.2931
S_S	g COD/m^3	69.5046	52.9758	60.9164
X_I	g COD/m^3	51.2085	39.0307	52.5648
X_S	g COD/m^3	202.3243	154.2097	185.4121
X_{BH}	g COD/m^3	28.1703	21.4711	26.4419
X_{AOB}	g COD/m^3	0	0	0
X_P	g COD/m^3	0	0	0
S_O	g -COD/m^3	0	0	0
S_{NO3}	g N/m^3	0	0	0
S_{NH}	g N/m^3	31.5563	24.052	27.6571
S_{ND}	g N/m^3	6.9505	5.2976	6.0916
X_{ND}	g N/m^3	10.5903	8.0719	9.9406
S_{ALK}	mol HCO_3/m^3	7	7	7
S_{NO2}	g N/m^3	0	0	0
S_{NO}	g N/m^3	0	0	0
S_{N2O}	g N/m^3	0	0	0
S_{N2}	g N/m^3	0	0	0
X_{NOB}	g COD/m^3	0	0	0
Kjeldahl nitrogen	g N/m^3	54.5923	41.6097	49.1172
Total nitrogen	g N/m^3	54.5923	41.6097	49.1172
Total COD	g COD/m^3	381.2077	290.553	351.6284
BOD	g BOD/m^3	193.5346	147.5103	175.9258
Influent average load				
S_I	kg COD/day	553.3916	553.3916	553.3916
S_S	kg COD/day	1282.1096	1282.1096	1282.1096
X_I	kg COD/day	944.6124	944.6124	1106.3336
X_S	kg COD/day	3732.1519	3732.1519	3902.3736
X_{BH}	kg COD/day	519.6401	519.6401	556.5226

X_{AOB}	kg COD/day	0	0	0
X_P	kg COD/day	0	0	0
S_O	kg -COD/day	0	0	0
S_{NO3}	kg N/day	0	0	0
S_{NH}	kg N/day	582.1005	582.1005	582.1005
S_{ND}	kg N/day	128.211	128.211	128.211
X_{ND}	kg N/day	195.3534	195.3534	209.2192
S_{ALK}	kmol HCO_3/day	129.1247	169.4126	147.3292
S_{NO2}	kg N/day	0	0	0
S_{NO}	kg N/day	0	0	0
S_{N2O}	kg N/day	0	0	0
S_{N2}	kg N/day	0	0	0
X_{NOB}	kg COD/day	0	0	0
Kjeldahl nitrogen	kg N/day	1007.0307	1007.0307	1033.7716
Total nitrogen	kg N/day	1007.0307	1007.0307	1033.7716
Total COD	kg COD/day	7031.9057	7031.9057	7400.7312
BOD	kg BOD/day	3570.0148	3570.0148	3702.7147
Effluent average concentrations				
Flow rate	m^3/d	18061.3875	23816.7981	20662.0253
S_I	g COD/m^3	30	22.8334	26.302
S_S	g COD/m^3	1.2306	1.3296	1.3382
X_I	g COD/m^3	4.7093	6.0951	5.9409
X_S	g COD/m^3	0.22347	0.32654	0.31418
X_{BH}	g COD/m^3	9.8367	12.865	11.7737
X_{AOB}	g COD/m^3	0.52118	0.66315	0.59528
X_P	g COD/m^3	1.7191	2.1841	1.9625
S_O	g -COD/m^3	3.5196	3.4175	3.3628
S_{NO3}	g N/m^3	14.7321	12.1225	13.0126
S_{NH}	g N/m^3	1.6763	1.7814	1.8529
S_{ND}	g N/m^3	0.56057	0.59055	0.59629
X_{ND}	g N/m^3	0.016472	0.023513	0.023094
S_{ALK}	mol HCO_3/m^3	3.8121	4.5443	4.2251
S_{NO2}	g N/m^3	0.027773	0.027325	0.031487
S_{NO}	g N/m^3	0.001912	0.0019047	0.0021444
S_{N2O}	g N/m^3	0.0010255	0.001018	0.0011531
S_{N2}	g N/m^3	13.3997	13.3094	13.3283
X_{NOB}	g COD/m^3	0.18092	0.23132	0.20673

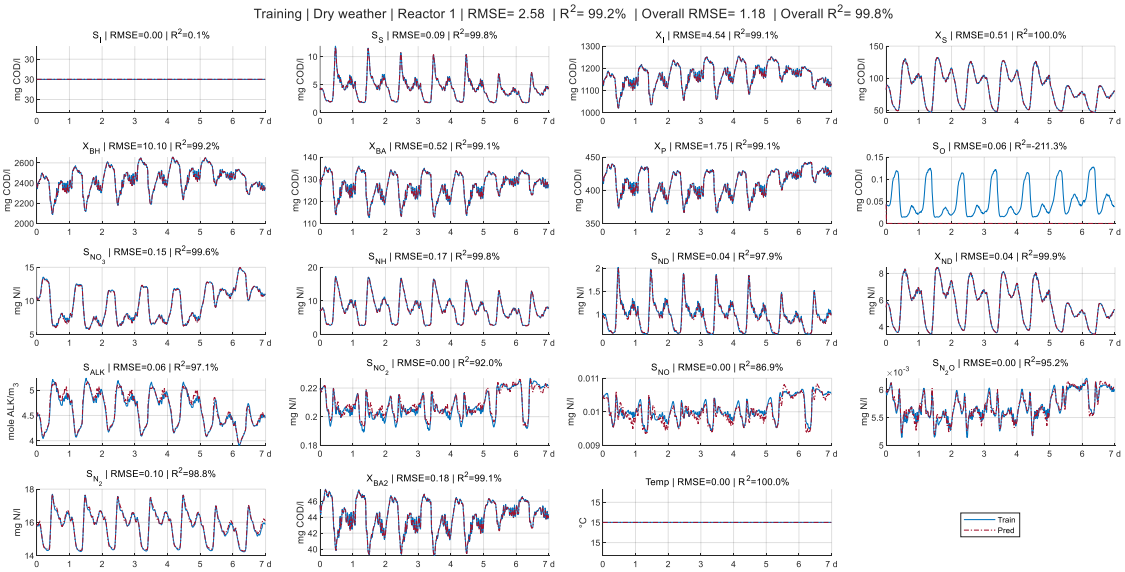
Kjeldahl nitrogen	g N/m^3	3.5454	4.0756	4.028
Total nitrogen (limit = 18 g N/m^3)	g N/m^3	18.3082	16.2283	17.0753
Total COD (limit = 100 g COD/m^3)	g COD/m^3	48.4213	46.5281	48.4336
BOD (limit = 10 g/m^3)	g BOD/m^3	2.7874	3.5787	3.3055
Effluent average load				
S_I	kg COD/day	541.8415	543.8175	543.4534
S_S	kg COD/day	22.2272	31.6675	27.65
X_I	kg COD/day	85.0563	145.1649	122.7518
X_S	kg COD/day	4.0362	7.7772	6.4916
X_{BH}	kg COD/day	177.6636	306.4021	243.2688
X_{AOB}	kg COD/day	9.4133	15.794	12.2997
X_P	kg COD/day	31.0494	52.0177	40.5501
S_O	kg -COD/day	63.5685	81.3935	69.4823
S_{NO3}	kg N/day	266.0821	288.7198	268.8662
S_{NH}	kg N/day	30.2762	42.4283	38.2837
S_{ND}	kg N/day	10.1247	14.0649	12.3205
X_{ND}	kg N/day	0.29751	0.56001	0.47718
S_{ALK}	kmol HCO_3/day	68.8526	108.2298	87.2981
S_{NO2}	kg N/day	0.50162	0.6508	0.65059
S_{NO}	kg N/day	0.034533	0.045365	0.044307
S_{N2O}	kg N/day	0.018522	0.024245	0.023825
S_{N2}	kg N/day	242.0169	316.9865	275.39
X_{NOB}	kg COD/day	3.2676	5.5092	4.27
Kjeldahl nitrogen	kg N/day	64.0343	97.0668	83.2257
Total nitrogen	kg N/day	330.6711	386.507	352.8106
Total COD	kg COD/day	874.5551	1108.15	1000.7369
BOD	kg BOD/day	50.3451	85.2334	68.2986
Sludge average concentrations				
Flow rate	m^3/d	385	385	385
S_I	g COD/m^3	30.0003	24.8983	26.966
S_S	g COD/m^3	1.2397	1.2846	1.3219
X_I	g COD/m^3	2261.7603	2179.1299	2398.219
X_S	g COD/m^3	100.8052	103.3403	107.6986
X_{BH}	g COD/m^3	4730.8012	4612.8182	4774.5945
X_{AOB}	g COD/m^3	250.317	237.9036	243.2574

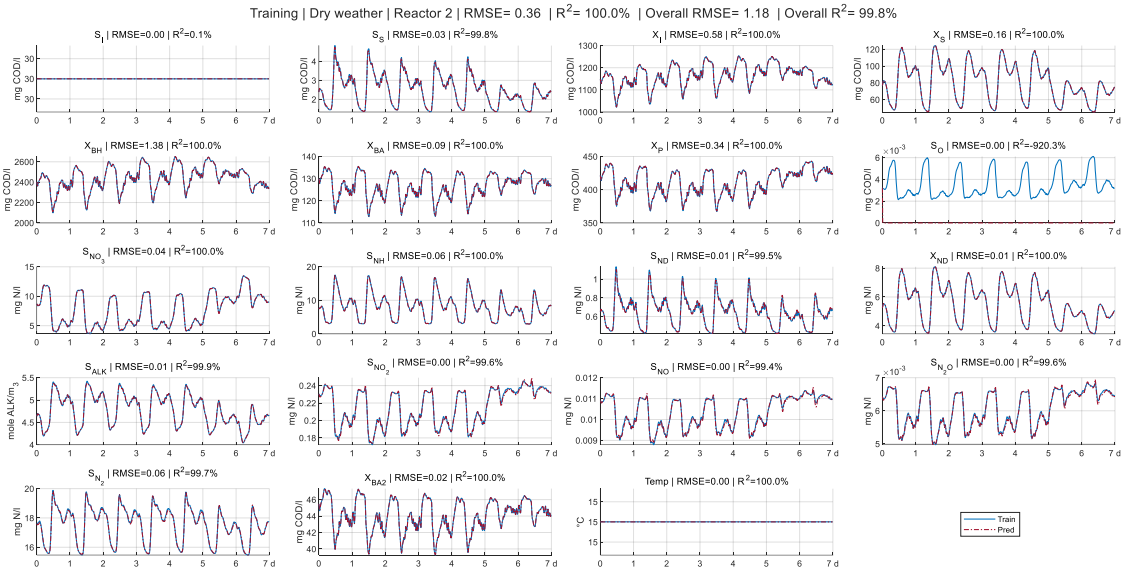
X_P	g COD/m^3	825.4722	784.8628	804.4082
S_O	g -COD/m^3	3.4699	3.4899	3.369
S_{NO3}	g N/m^3	14.8122	13.13	13.4838
S_{NH}	g N/m^3	1.6373	1.6677	1.8303
S_{ND}	g N/m^3	0.56304	0.57617	0.59202
X_{ND}	g N/m^3	7.5158	7.6141	8.0248
S_{ALK}	mol HCO_3/m^3	3.7975	4.3061	4.1337
S_{NO2}	g N/m^3	0.026558	0.025106	0.031012
S_{NO}	g N/m^3	0.0018433	0.0017588	0.0021052
S_{N2O}	g N/m^3	0.00098837	0.00093729	0.0011295
S_{N2}	g N/m^3	13.3989	13.3256	13.3427
X_{NOB}	g COD/m^3	86.8904	83.0029	84.5029
Kjeldahl nitrogen	g N/m^3	636.0123	616.9779	646.4774
Total nitrogen	g N/m^3	650.8538	630.1357	659.9955
Total COD	g COD/m^3	8287.2861	8027.2406	8440.9684
BOD	g BOD/m^3	1191.1532	1160.9129	1200.7967
Sludge average load				
S_I	kg COD/day	11.5501	9.5858	10.3819
S_S	kg COD/day	0.47728	0.49459	0.50892
X_I	kg COD/day	870.7777	838.965	923.3143
X_S	kg COD/day	38.81	39.786	41.464
X_{BH}	kg COD/day	1821.3585	1775.935	1838.2189
X_{AOB}	kg COD/day	96.372	91.5929	93.6541
X_P	kg COD/day	317.8068	302.1722	309.6971
S_O	kg -COD/day	1.3359	1.3436	1.2971
S_{NO3}	kg N/day	5.7027	5.055	5.1913
S_{NH}	kg N/day	0.63038	0.64205	0.70467
S_{ND}	kg N/day	0.21677	0.22182	0.22793
X_{ND}	kg N/day	2.8936	2.9314	3.0895
S_{ALK}	kmol HCO_3/day	1.462	1.6578	1.5915
S_{NO2}	kg N/day	0.010225	0.009666	0.01194
S_{NO}	kg N/day	0.0007097	0.00067716	0.00081051
S_{N2O}	kg N/day	0.00038052	0.00036085	0.00043484
S_{N2}	kg N/day	5.1586	5.1304	5.1369
X_{NOB}	kg COD/day	33.4528	31.9561	32.5336
Kjeldahl nitrogen	kg N/day	242.8576	235.6191	246.9416

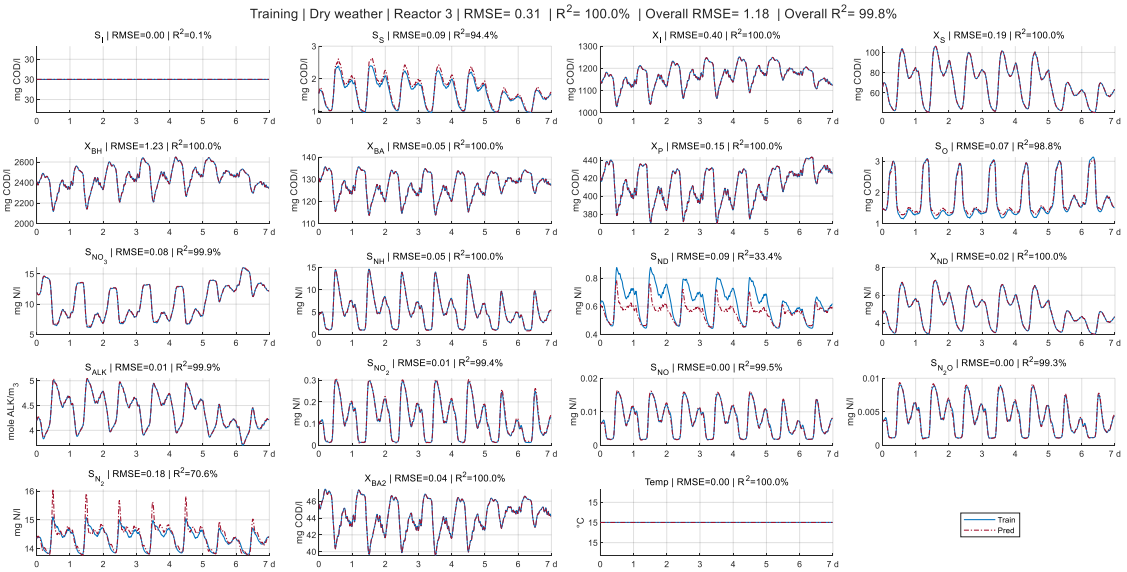
Total nitrogen	kg N/day	248.5716	240.6849	252.1462
Total COD	kg COD/day	3190.6052	3090.4876	3249.7728
BOD	kg BOD/day	458.594	446.9515	462.3067
Quality index				
Influent quality index (IQI)	kg poll.units/d	52177.4614	52177.4613	54167.1531
Effluent quality index (EQI)	kg poll.units/d	6028.3727	7884.0201	6974.4052
Sludge production				
Sludge production for disposal	kg SS	16462.1091	15536.2085	17433.8839
Average sludge production for disposal per day	kg SS/day	2355.2347	2222.766	2494.2665
Sludge production released into effluent	kg SS	1627.6277	2792.33	2252.2199
Average sludge production released into effluent / day	kg SS/day	232.8648	399.4988	322.2252
Total sludge production	kg SS	18089.7369	18328.5384	19686.1038
Total average sludge production per day	kg SS/day	2588.0995	2622.2648	2816.4918
Energy/chemical consumption				
Average aeration energy	kWh/day	4283.3778	4283.3778	4283.3778
Average pumping energy	kWh/day	388.1777	503.286	440.1905
Average carbon source addition	kg COD/day	0	0	0
Average mixing energy	kWh/day	0.35768	0.35768	0.35768
Operational cost index				
Sludge production cost index		11776.1734	11113.8302	12471.3327
Aeration energy cost index		4283.3778	4283.3778	4283.3778
Pumping energy cost index		388.1777	503.286	440.1905
Carbon source dosage cost index		0	0	0
Mixing energy cost index		0.35768	0.35768	0.35768
Total operational cost index		16448.0866	15900.8516	17175.2587
N_2O emissions during nitrification/denitrification				
Anoxic tank 1	kg N_2O/day	0.010774	0.0099702	0.01027
Anoxic tank 2	kg N_2O/day	0.011361	0.010932	0.010878
Aeration tank 1	kg N_2O/day	1.1921	1.0593	1.18
Aeration tank 2	kg N_2O/day	0.54236	0.4993	0.57496
Aeration tank 3	kg N_2O/day	0.23629	0.23716	0.26871
Total	kg N_2O/day	1.9929	1.8167	2.0448
95% percentile for effluent				

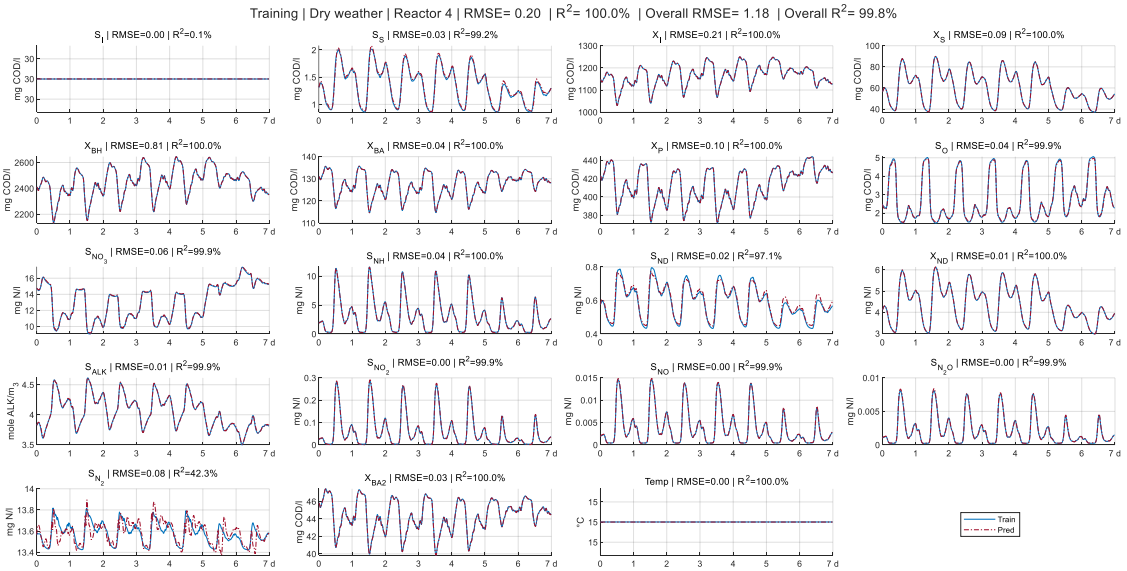
S_{NH} (Ammonia95)	g N/m^3	4.7423	4.8048	4.9255
TN (TN95)	g N/m^3	20.7636	20.214	20.6124
TSS (TSS95)	g SS/m^3	15.8608	23.2378	22.5545
Violation of effluent total nitrogen level (18 g N/m^3)				
Operating time length	day	3.3333	2.6771	2.9687
Percentage of the operation time	%	47.69%	38.301%	42.4739%
Number of occasions that violated the limit	pcs	8	6	10
Violation of effluent ammonia level (4 g N/m^3)				
Operating time length	day	0.70833	0.67708	0.75
Percentage of the operation time	%	10.1341%	9.687%	10.7303%
Number of occasions that violated the limit	pcs	5	5	5

8.2 Results of prediction in dry weather scenario




Training | Dry weather | Reactor 2 | RMSE= 0.36 | R2= 100.0% | Overall RMSE= 1.18 | Overall R2= 99.8%
SI | RMSE=0.00 | R2=0.1%
SS | RMSE=0.03 | R2=99.8%
XI | RMSE=0.58 | R2=100.0%
XS | RMSE=0.16 | R2=100.0%
XBH | RMSE=1.38 | R2=100.0%
XBA | RMSE=0.09 | R2=100.0%
XP | RMSE=0.34 | R2=100.0%
SO | RMSE=0.00 | R2=-920.3%
SNO3 | RMSE=0.04 | R2=100.0%
SNH | RMSE=0.06 | R2=100.0%
SND | RMSE=0.01 | R2=99.5%
XND | RMSE=0.01 | R2=100.0%
SALK | RMSE=0.01 | R2=99.9%
SNO2 | RMSE=0.00 | R2=99.6%
SNO | RMSE=0.00 | R2=99.4%
SN2O | RMSE=0.00 | R2=99.6%
SN2 | RMSE=0.06 | R2=99.7%
XBA2 | RMSE=0.02 | R2=100.0%
Temp | RMSE=0.00 | R2=100.0%
Train
Pred


Training | Dry weather | Reactor 3 | RMSE= 0.31 | R2= 100.0% | Overall RMSE= 1.18 | Overall R2= 99.8%
SI | RMSE=0.00 | R2=0.1%
SS | RMSE=0.09 | R2=94.4%
XI | RMSE=0.40 | R2=100.0%
XS | RMSE=0.19 | R2=100.0%
XBH | RMSE=1.23 | R2=100.0%
XBA | RMSE=0.05 | R2=100.0%
XP | RMSE=0.15 | R2=100.0%
SO | RMSE=0.07 | R2=98.8%
SNO3 | RMSE=0.08 | R2=99.9%
SNH | RMSE=0.05 | R2=100.0%
SND | RMSE=0.09 | R2=33.4%
XND | RMSE=0.02 | R2=100.0%
SALK | RMSE=0.01 | R2=99.9%
SNO2 | RMSE=0.01 | R2=99.4%
SNO | RMSE=0.00 | R2=99.5%
SN2O | RMSE=0.00 | R2=99.3%
SN2 | RMSE=0.18 | R2=70.6%
XBA2 | RMSE=0.04 | R2=100.0%
Temp | RMSE=0.00 | R2=100.0%
Train
Pred


Training | Dry weather | Reactor 4 | RMSE= 0.20 | R2= 100.0% | Overall RMSE= 1.18 | Overall R2= 99.8%
SI | RMSE=0.00 | R2=0.1%
SS | RMSE=0.03 | R2=99.2%
XI | RMSE=0.21 | R2=100.0%
XS | RMSE=0.09 | R2=100.0%
XBH | RMSE=0.81 | R2=100.0%
XBA | RMSE=0.04 | R2=100.0%
XP | RMSE=0.10 | R2=100.0%
SO | RMSE=0.04 | R2=99.9%
SNO3 | RMSE=0.06 | R2=99.9%
SNH | RMSE=0.04 | R2=100.0%
SND | RMSE=0.02 | R2=97.1%
XND | RMSE=0.01 | R2=100.0%
SALK | RMSE=0.01 | R2=99.9%
SNO2 | RMSE=0.00 | R2=99.9%
SNO | RMSE=0.00 | R2=99.9%
SN2O | RMSE=0.00 | R2=99.9%
SN2 | RMSE=0.08 | R2=42.3%
XBA2 | RMSE=0.03 | R2=100.0%
Temp | RMSE=0.00 | R2=100.0%
Train
Pred

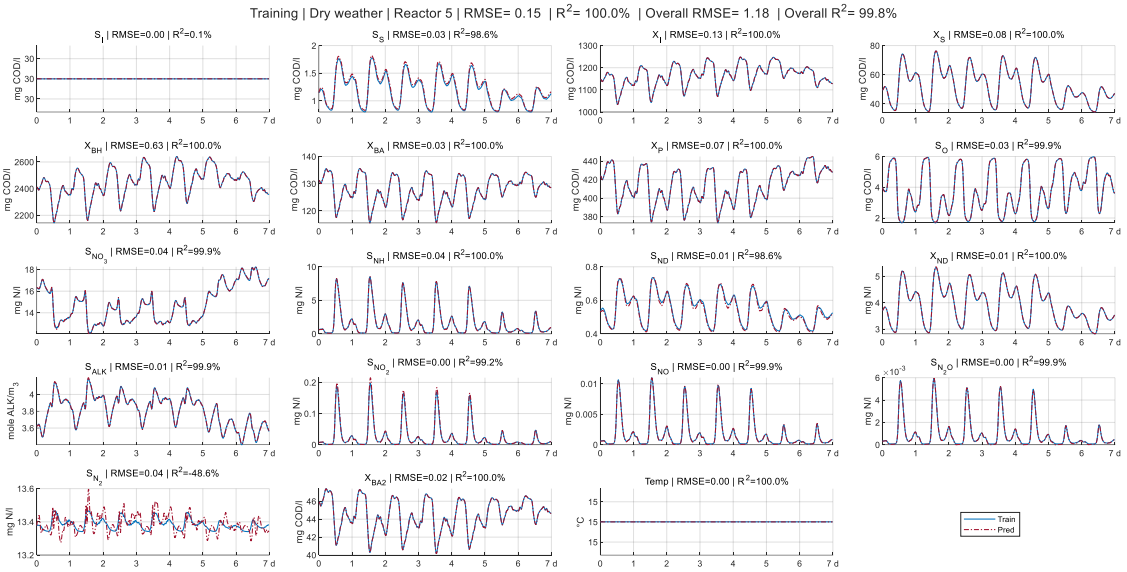
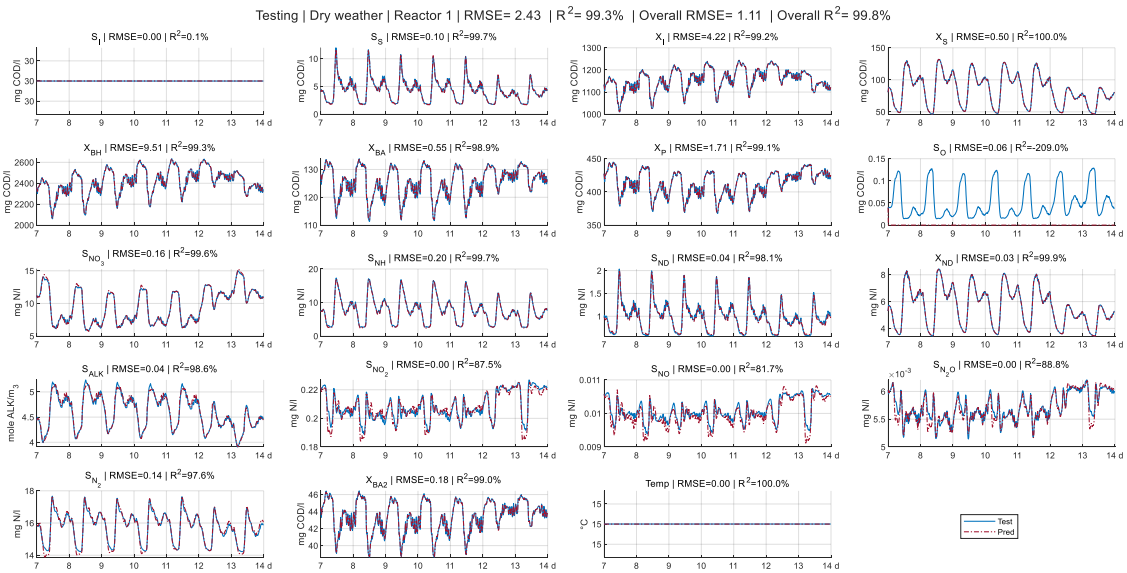
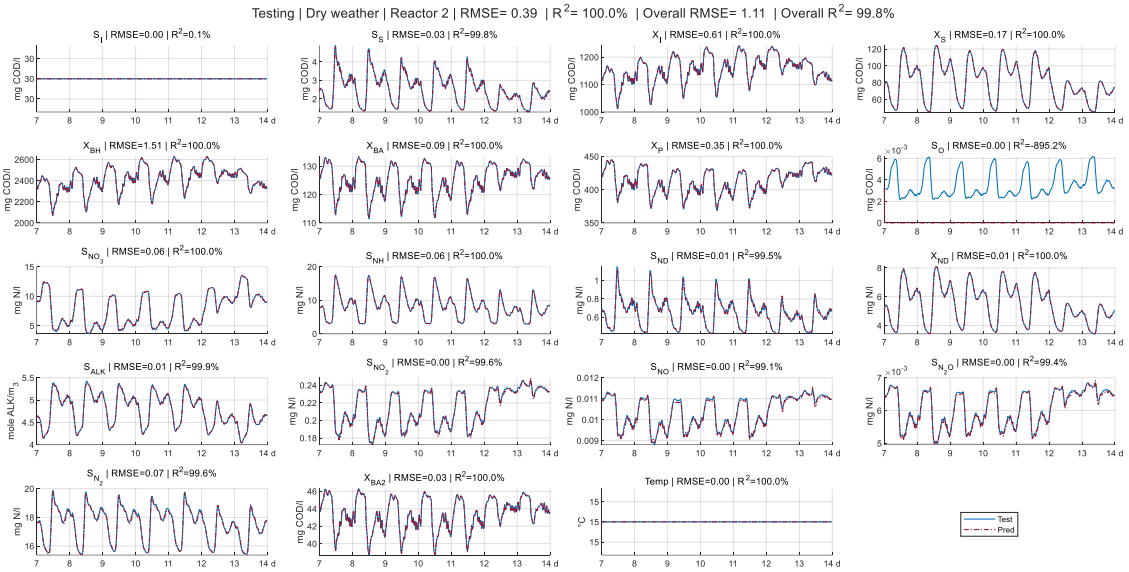
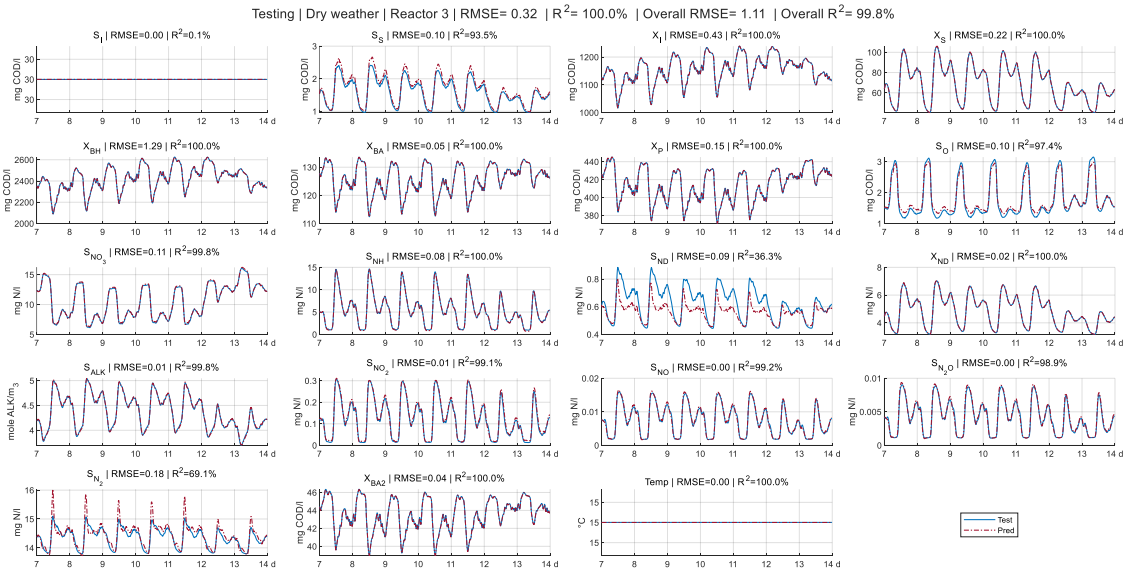
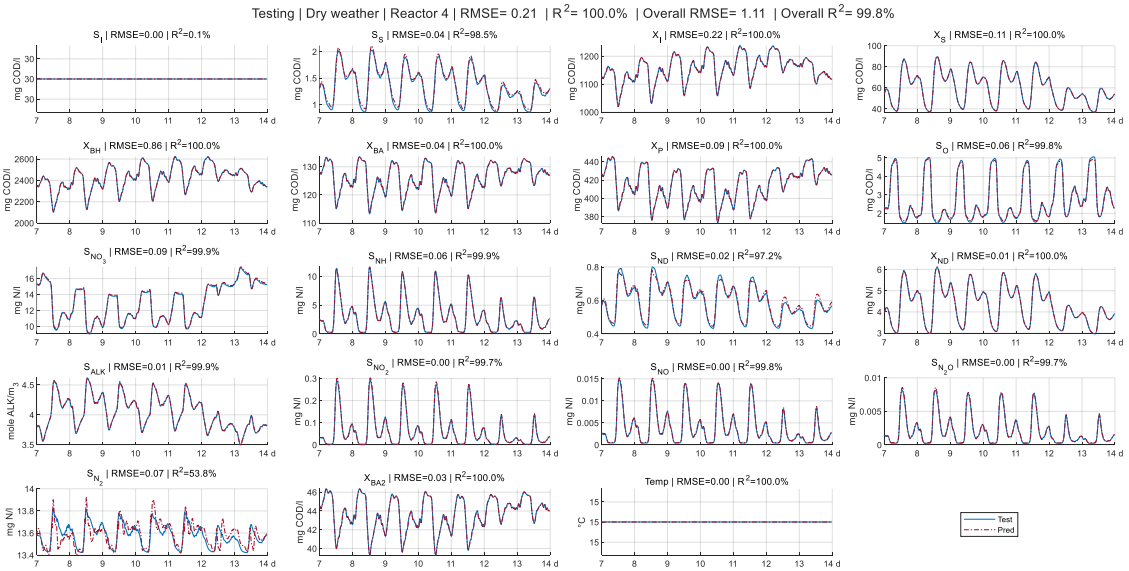


Figure 8.1 Training result using first 7-day dry weather data




Testing | Dry weather | Reactor 2 | RMSE= 0.39 | R2= 100.0% | Overall RMSE= 1.11 | Overall R2= 99.8%
SI | RMSE=0.00 | R2=0.1%
SS | RMSE=0.03 | R2=99.8%
XI | RMSE=0.61 | R2=100.0%
XS | RMSE=0.17 | R2=100.0%
XBH | RMSE=1.51 | R2=100.0%
XBA | RMSE=0.09 | R2=100.0%
XP | RMSE=0.35 | R2=100.0%
SO | RMSE=0.00 | R2=-895.2%
SNO3 | RMSE=0.06 | R2=100.0%
SNH | RMSE=0.06 | R2=100.0%
SND | RMSE=0.01 | R2=99.5%
XND | RMSE=0.01 | R2=100.0%
SALK | RMSE=0.01 | R2=99.9%
SNO2 | RMSE=0.00 | R2=99.6%
SNO | RMSE=0.00 | R2=99.1%
SN2O | RMSE=0.00 | R2=99.4%
SN2 | RMSE=0.07 | R2=99.6%
XBA2 | RMSE=0.03 | R2=100.0%
Temp | RMSE=0.00 | R2=100.0%
Test
Pred


Testing | Dry weather | Reactor 3 | RMSE= 0.32 | R2= 100.0% | Overall RMSE= 1.11 | Overall R2= 99.8%
SI | RMSE=0.00 | R2=0.1%
SS | RMSE=0.10 | R2=93.5%
XI | RMSE=0.43 | R2=100.0%
XS | RMSE=0.22 | R2=100.0%
XBH | RMSE=1.29 | R2=100.0%
XBA | RMSE=0.05 | R2=100.0%
XP | RMSE=0.15 | R2=100.0%
SO | RMSE=0.10 | R2=97.4%
SNO3 | RMSE=0.11 | R2=99.8%
SNH | RMSE=0.08 | R2=100.0%
SND | RMSE=0.09 | R2=36.3%
XND | RMSE=0.02 | R2=100.0%
SALK | RMSE=0.01 | R2=99.8%
SNO2 | RMSE=0.01 | R2=99.1%
SNO | RMSE=0.00 | R2=99.2%
SN2O | RMSE=0.00 | R2=98.9%
SN2 | RMSE=0.18 | R2=69.1%
XBA2 | RMSE=0.04 | R2=100.0%
Temp | RMSE=0.00 | R2=100.0%
Test
Pred


Testing | Dry weather | Reactor 4 | RMSE= 0.21 | R2= 100.0% | Overall RMSE= 1.11 | Overall R2= 99.8%
SI | RMSE=0.00 | R2=0.1%
SS | RMSE=0.04 | R2=98.5%
XI | RMSE=0.22 | R2=100.0%
XS | RMSE=0.11 | R2=100.0%
XBH | RMSE=0.86 | R2=100.0%
XBA | RMSE=0.04 | R2=100.0%
XP | RMSE=0.09 | R2=100.0%
SO | RMSE=0.06 | R2=99.8%
SNO3 | RMSE=0.09 | R2=99.9%
SNH | RMSE=0.06 | R2=99.9%
SND | RMSE=0.02 | R2=97.2%
XND | RMSE=0.01 | R2=100.0%
SALK | RMSE=0.01 | R2=99.9%
SNO2 | RMSE=0.00 | R2=99.7%
SNO | RMSE=0.00 | R2=99.8%
SN2O | RMSE=0.00 | R2=99.7%
SN2 | RMSE=0.07 | R2=53.8%
XBA2 | RMSE=0.03 | R2=100.0%
Temp | RMSE=0.00 | R2=100.0%
Test
Pred

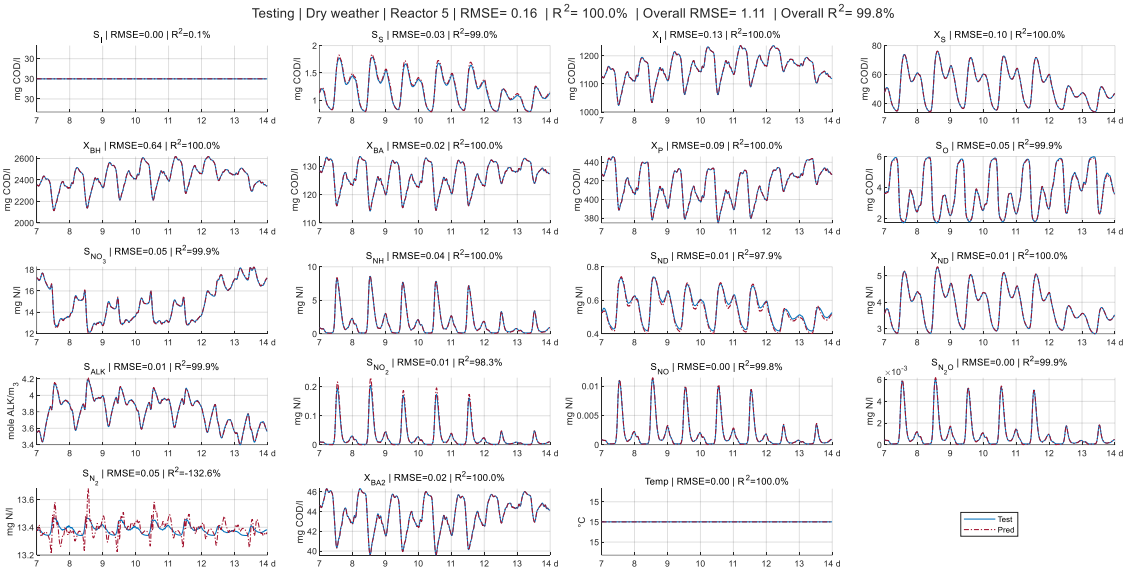


Figure 8.2 Testing results for day 7-14 of dry weather scenario

8.3 Results of cross scenario validation

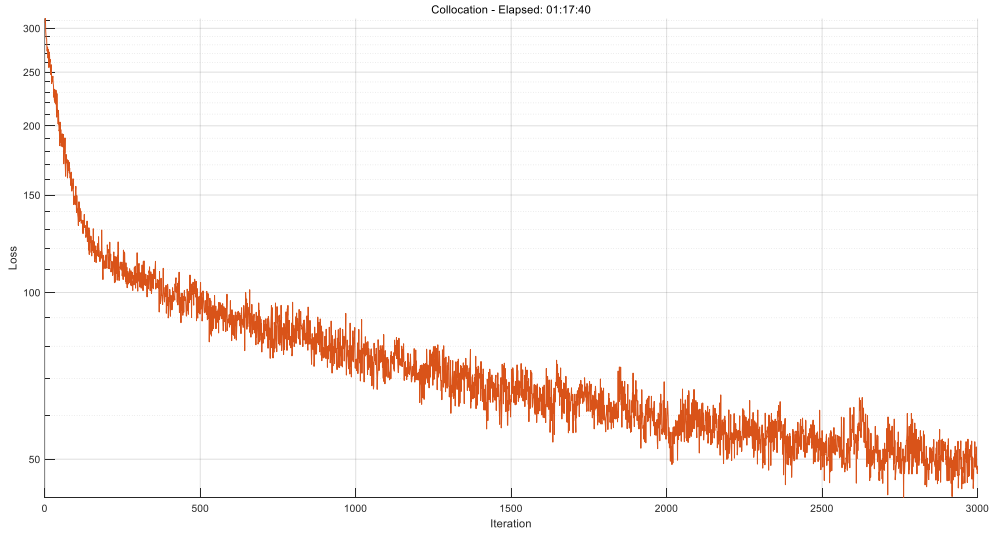


Figure 8.3 Collocation training loss on first 12-day rain weather data

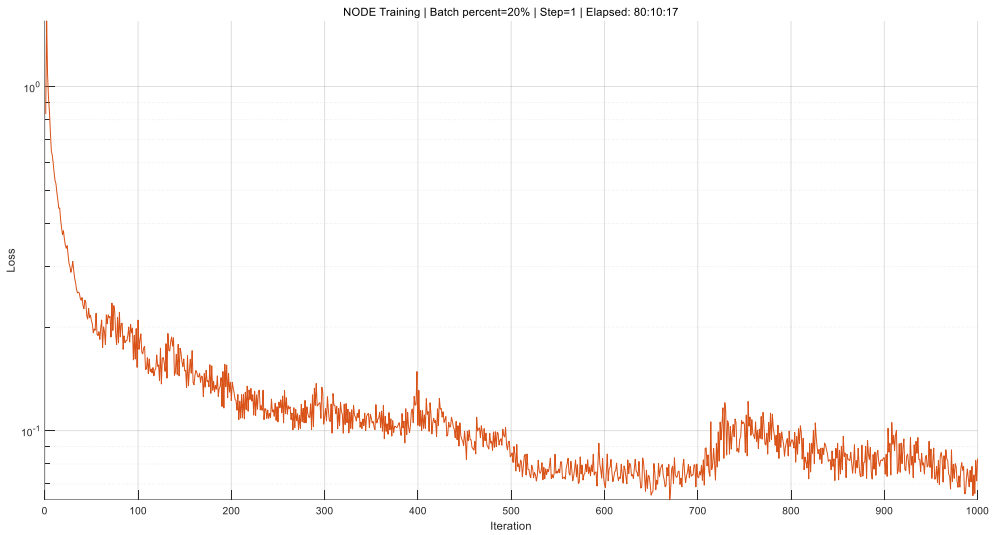
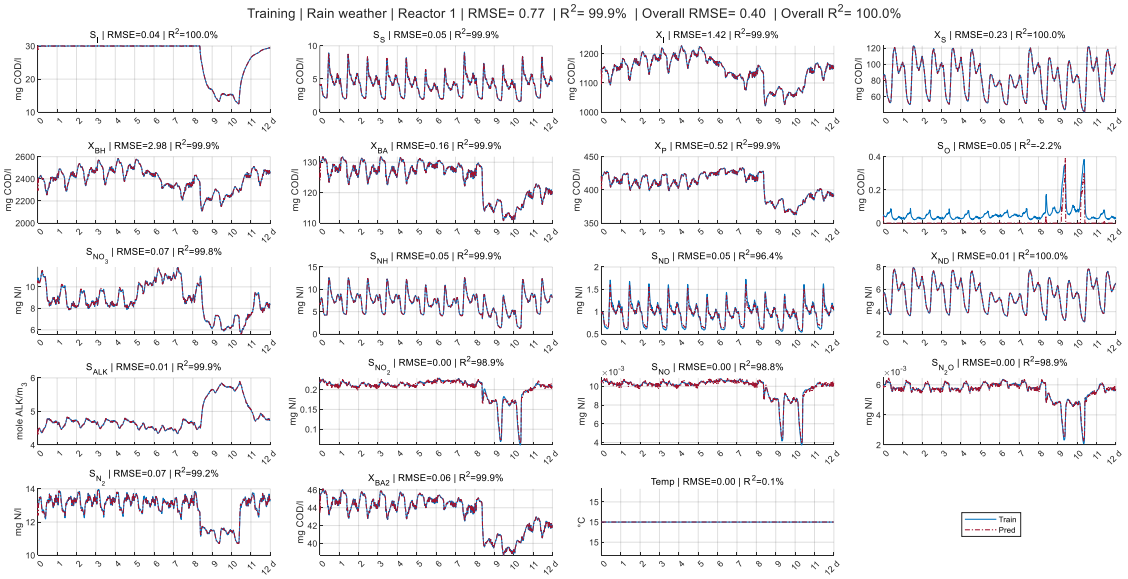
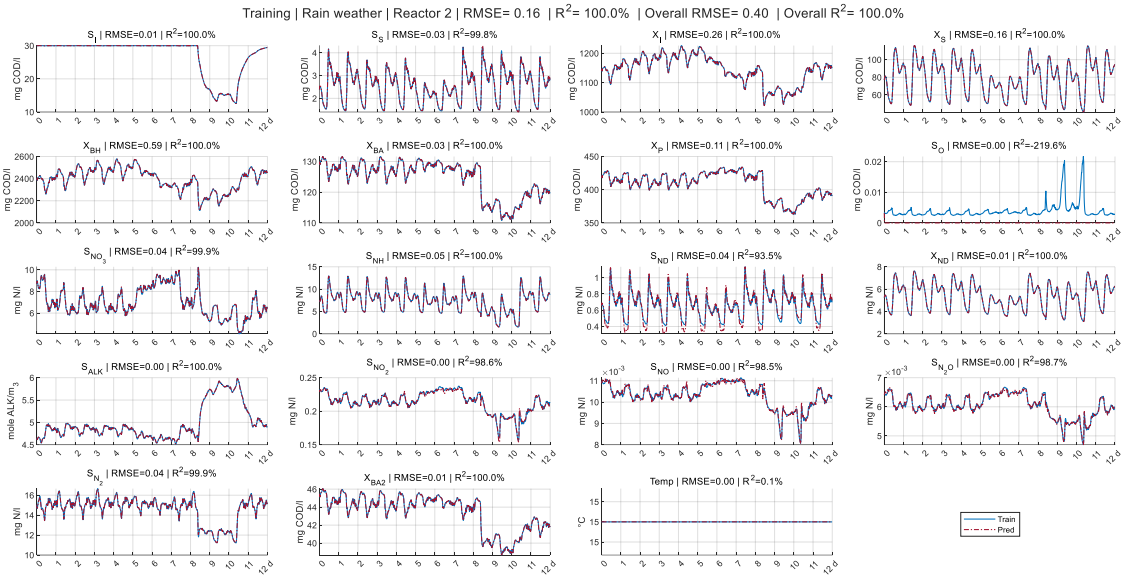
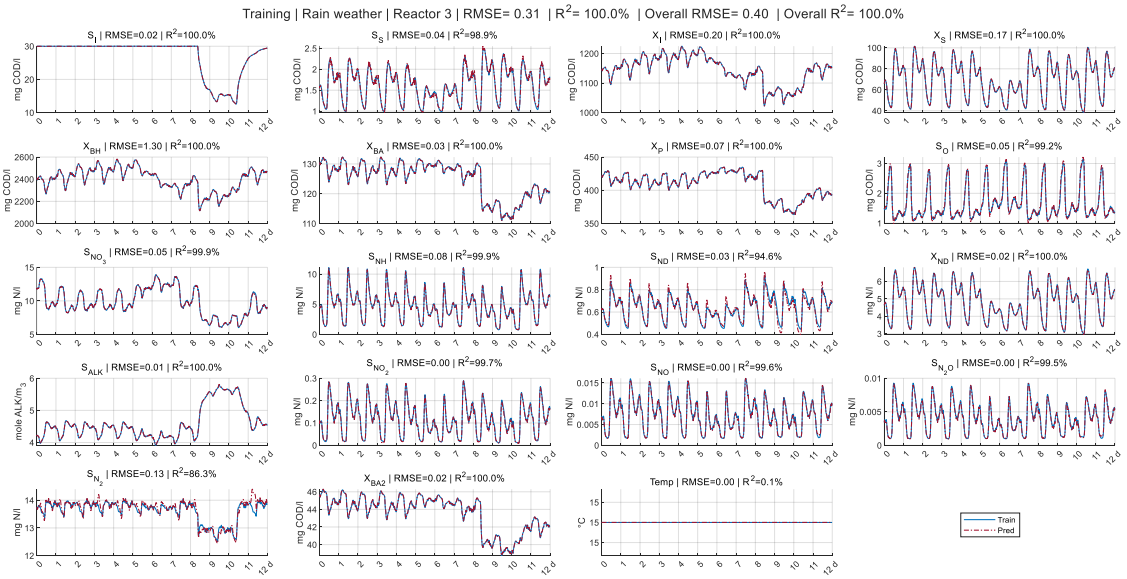
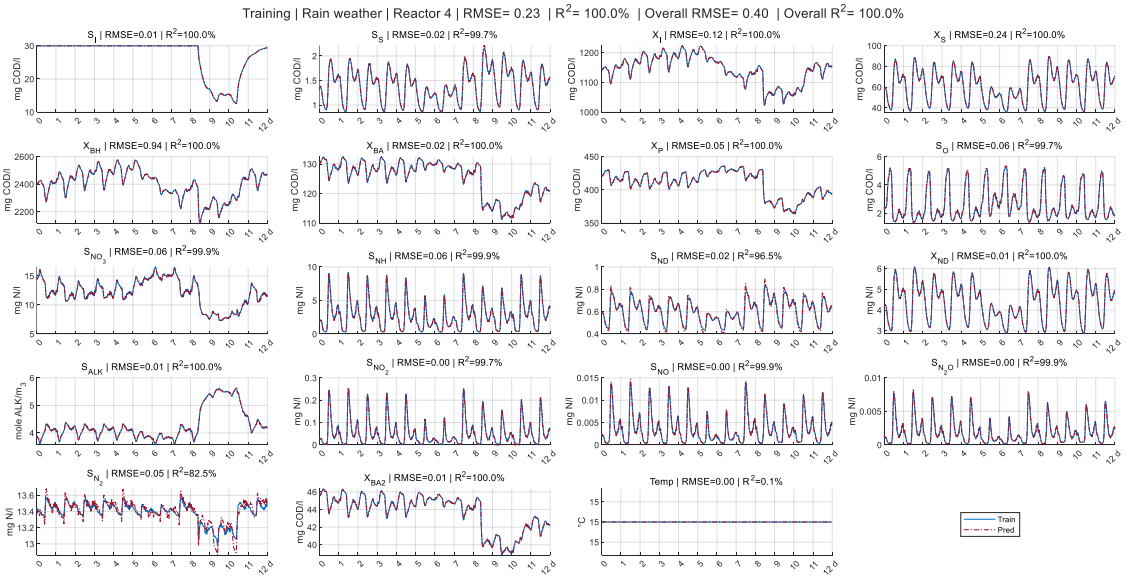


Figure 8.4 NODE training loss on first 12-day rain weather data




Training | Rain weather | Reactor 2 | RMSE= 0.16 | R2= 100.0% | Overall RMSE= 0.40 | Overall R2= 100.0%
SI | RMSE=0.01 | R2=100.0%
SS | RMSE=0.03 | R2=99.8%
XI | RMSE=0.26 | R2=100.0%
XS | RMSE=0.16 | R2=100.0%
XBH | RMSE=0.59 | R2=100.0%
XBA | RMSE=0.03 | R2=100.0%
XP | RMSE=0.11 | R2=100.0%
SO | RMSE=0.00 | R2=-219.6%
SNO3 | RMSE=0.04 | R2=99.9%
SNH | RMSE=0.05 | R2=100.0%
SND | RMSE=0.04 | R2=93.5%
XND | RMSE=0.01 | R2=100.0%
SALK | RMSE=0.00 | R2=100.0%
SNO2 | RMSE=0.00 | R2=98.6%
SNO | RMSE=0.00 | R2=98.5%
SN2O | RMSE=0.00 | R2=98.7%
SN2 | RMSE=0.04 | R2=99.9%
XBA2 | RMSE=0.01 | R2=100.0%
Temp | RMSE=0.00 | R2=0.1%
mg COD/l
mg N/l
mole ALK/m3
°C
Train
Pred


Training | Rain weather | Reactor 3 | RMSE= 0.31 | R2= 100.0% | Overall RMSE= 0.40 | Overall R2= 100.0%
SI | RMSE=0.02 | R2=100.0%
SS | RMSE=0.04 | R2=98.9%
XI | RMSE=0.20 | R2=100.0%
XS | RMSE=0.17 | R2=100.0%
XBH | RMSE=1.30 | R2=100.0%
XBA | RMSE=0.03 | R2=100.0%
XP | RMSE=0.07 | R2=100.0%
SO | RMSE=0.05 | R2=99.2%
SNO3 | RMSE=0.05 | R2=99.9%
SNH | RMSE=0.08 | R2=99.9%
SND | RMSE=0.03 | R2=94.6%
XND | RMSE=0.02 | R2=100.0%
SALK | RMSE=0.01 | R2=100.0%
SNO2 | RMSE=0.00 | R2=99.7%
SNO | RMSE=0.00 | R2=99.6%
SN2O | RMSE=0.00 | R2=99.5%
SN2 | RMSE=0.13 | R2=86.3%
XBA2 | RMSE=0.02 | R2=100.0%
Temp | RMSE=0.00 | R2=0.1%
mg COD/l
mg N/l
mole ALK/m3
°C
Train
Pred


Training | Rain weather | Reactor 4 | RMSE= 0.23 | R2= 100.0% | Overall RMSE= 0.40 | Overall R2= 100.0%
SI | RMSE=0.01 | R2=100.0%
SS | RMSE=0.02 | R2=99.7%
XI | RMSE=0.12 | R2=100.0%
XS | RMSE=0.24 | R2=100.0%
XBH | RMSE=0.94 | R2=100.0%
XBA | RMSE=0.02 | R2=100.0%
XP | RMSE=0.05 | R2=100.0%
SO | RMSE=0.06 | R2=99.7%
SNO3 | RMSE=0.06 | R2=99.9%
SNH | RMSE=0.06 | R2=99.9%
SND | RMSE=0.02 | R2=96.5%
XND | RMSE=0.01 | R2=100.0%
SALK | RMSE=0.01 | R2=100.0%
SNO2 | RMSE=0.00 | R2=99.7%
SNO | RMSE=0.00 | R2=99.9%
SN2O | RMSE=0.00 | R2=99.9%
SN2 | RMSE=0.05 | R2=82.5%
XBA2 | RMSE=0.01 | R2=100.0%
Temp | RMSE=0.00 | R2=0.1%
mg COD/l
mg N/l
mole ALK/m3
°C
Train
Pred

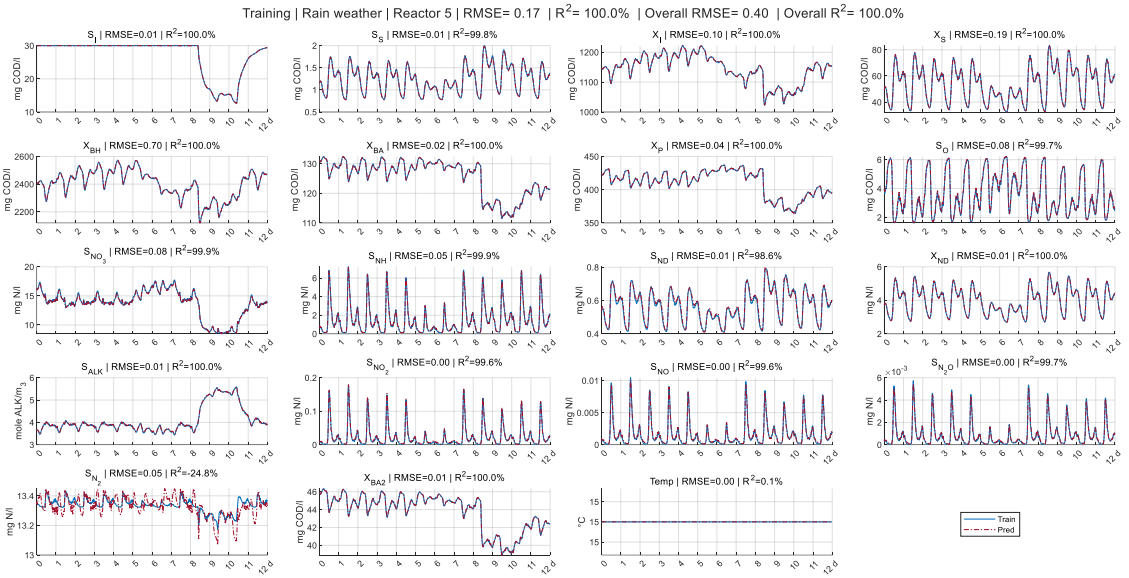
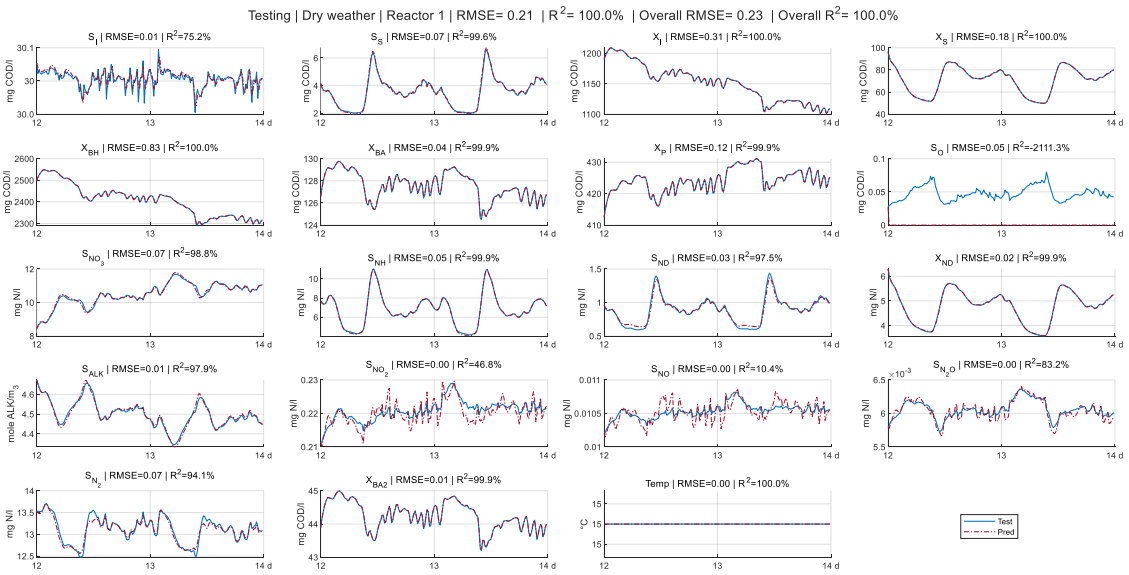
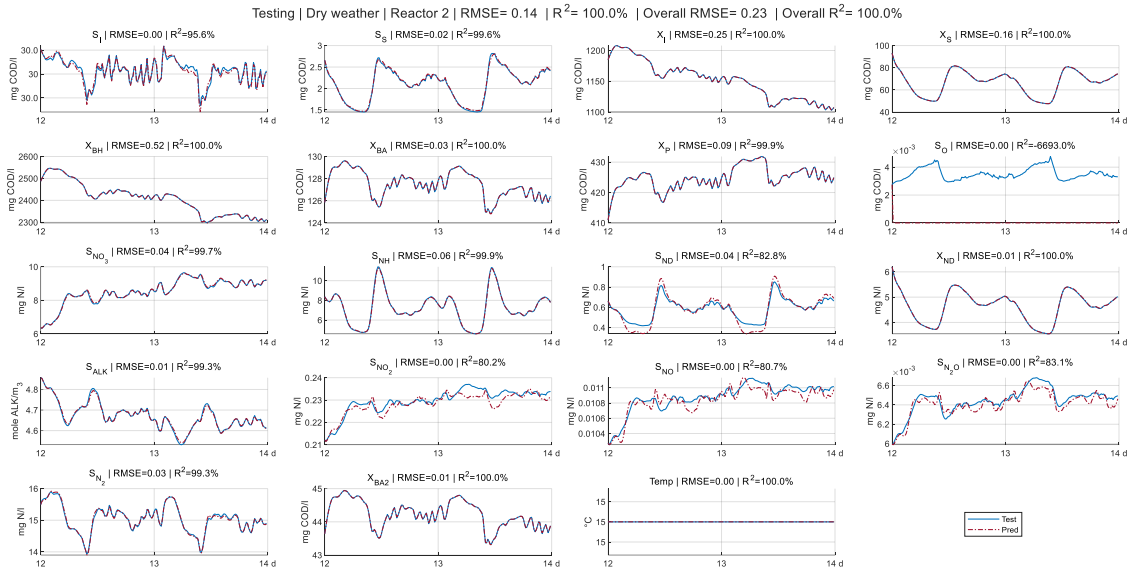
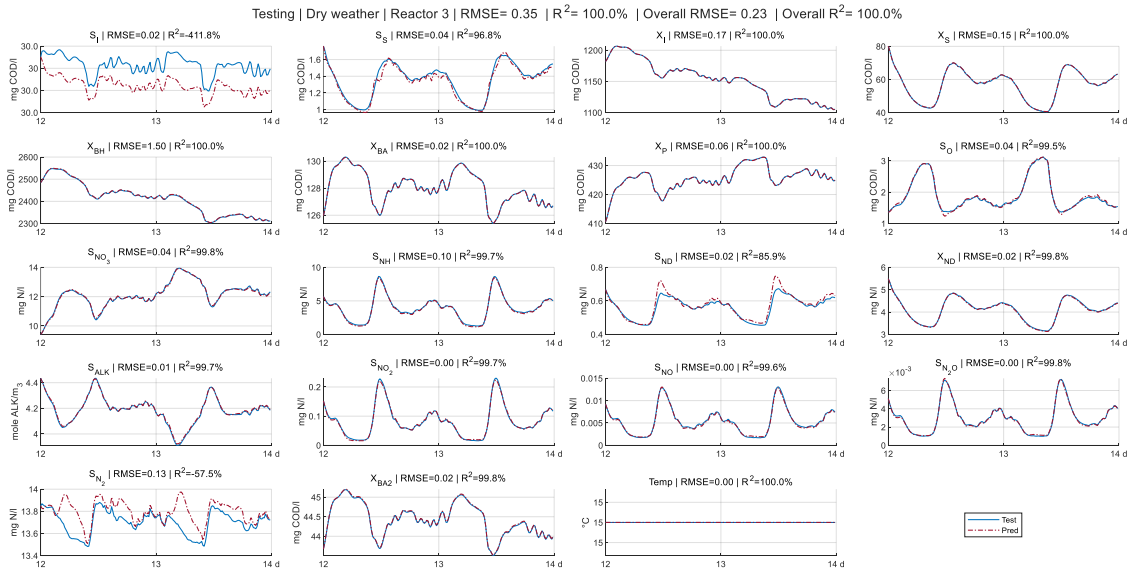
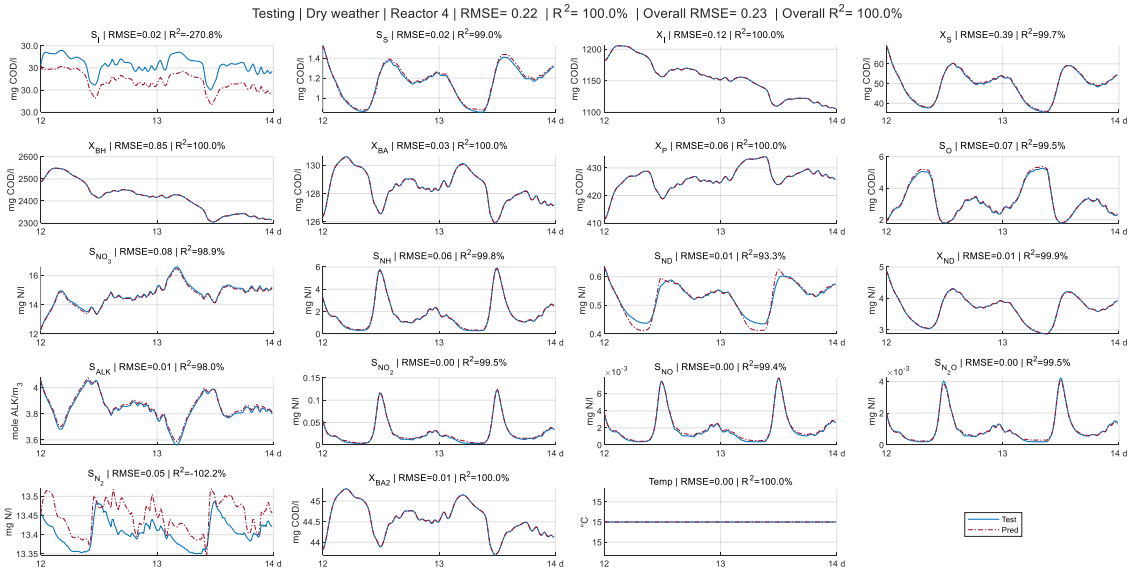


Figure 8.5 Training results using first 12-day rain weather data




Testing | Dry weather | Reactor 2 | RMSE= 0.14 | R2= 100.0% | Overall RMSE= 0.23 | Overall R2= 100.0%
SI | RMSE=0.00 | R2=95.6%
SS | RMSE=0.02 | R2=99.6%
XI | RMSE=0.25 | R2=100.0%
XS | RMSE=0.16 | R2=100.0%
XBH | RMSE=0.52 | R2=100.0%
XBA | RMSE=0.03 | R2=100.0%
XP | RMSE=0.09 | R2=99.9%
SO | RMSE=0.00 | R2=-6693.0%
SNO3 | RMSE=0.04 | R2=99.7%
SNH | RMSE=0.06 | R2=99.9%
SND | RMSE=0.04 | R2=82.8%
XND | RMSE=0.01 | R2=100.0%
SALK | RMSE=0.01 | R2=99.3%
SNO2 | RMSE=0.00 | R2=80.2%
SNO | RMSE=0.00 | R2=80.7%
SN2O | RMSE=0.00 | R2=83.1%
SN2 | RMSE=0.03 | R2=99.3%
XBA2 | RMSE=0.01 | R2=100.0%
Temp | RMSE=0.00 | R2=100.0%
Test
Pred


Testing | Dry weather | Reactor 3 | RMSE= 0.35 | R2= 100.0% | Overall RMSE= 0.23 | Overall R2= 100.0%
SI | RMSE=0.02 | R2=-411.8%
SS | RMSE=0.04 | R2=96.8%
XI | RMSE=0.17 | R2=100.0%
XS | RMSE=0.15 | R2=100.0%
XBH | RMSE=1.50 | R2=100.0%
XBA | RMSE=0.02 | R2=100.0%
XP | RMSE=0.06 | R2=100.0%
SO | RMSE=0.04 | R2=99.5%
SNO3 | RMSE=0.04 | R2=99.8%
SNH | RMSE=0.10 | R2=99.7%
SND | RMSE=0.02 | R2=85.9%
XND | RMSE=0.02 | R2=99.8%
SALK | RMSE=0.01 | R2=99.7%
SNO2 | RMSE=0.00 | R2=99.7%
SNO | RMSE=0.00 | R2=99.6%
SN2O | RMSE=0.00 | R2=99.8%
SN2 | RMSE=0.13 | R2=-57.5%
XBA2 | RMSE=0.02 | R2=99.8%
Temp | RMSE=0.00 | R2=100.0%
Test
Pred


Testing | Dry weather | Reactor 4 | RMSE= 0.22 | R2= 100.0% | Overall RMSE= 0.23 | Overall R2= 100.0%
SI | RMSE=0.02 | R2=-270.8%
SS | RMSE=0.02 | R2=99.0%
XI | RMSE=0.12 | R2=100.0%
XS | RMSE=0.39 | R2=99.7%
XBH | RMSE=0.85 | R2=100.0%
XBA | RMSE=0.03 | R2=100.0%
XP | RMSE=0.06 | R2=100.0%
SO | RMSE=0.07 | R2=99.5%
SNO3 | RMSE=0.08 | R2=98.9%
SNH | RMSE=0.06 | R2=99.8%
SND | RMSE=0.01 | R2=93.3%
XND | RMSE=0.01 | R2=99.9%
SALK | RMSE=0.01 | R2=98.0%
SNO2 | RMSE=0.00 | R2=99.5%
SNO | RMSE=0.00 | R2=99.4%
SN2O | RMSE=0.00 | R2=99.5%
SN2 | RMSE=0.05 | R2=-102.2%
XBA2 | RMSE=0.01 | R2=100.0%
Temp | RMSE=0.00 | R2=100.0%
Test
Pred

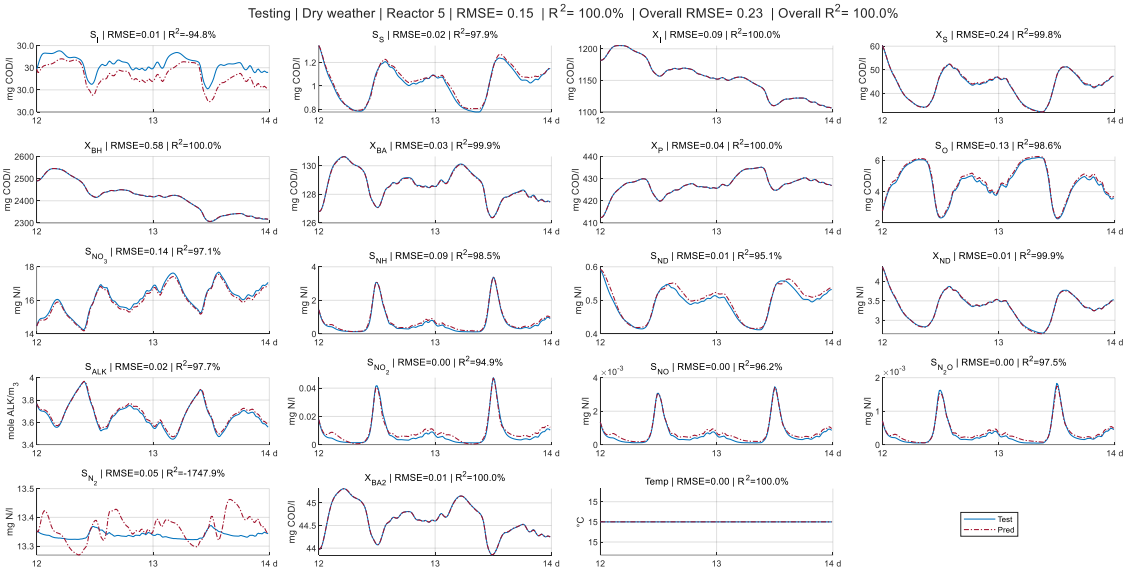
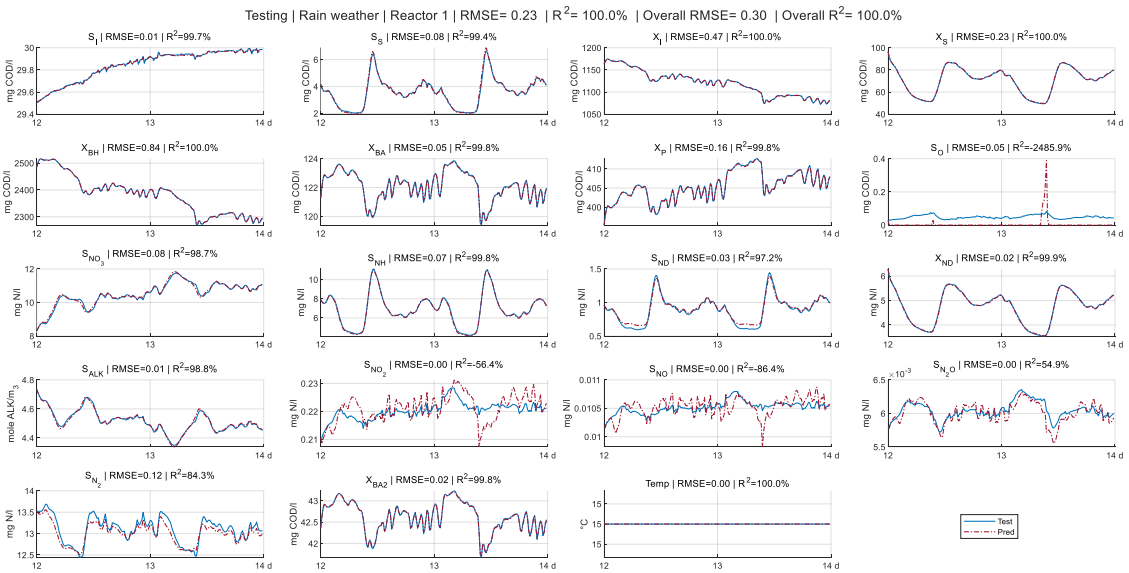
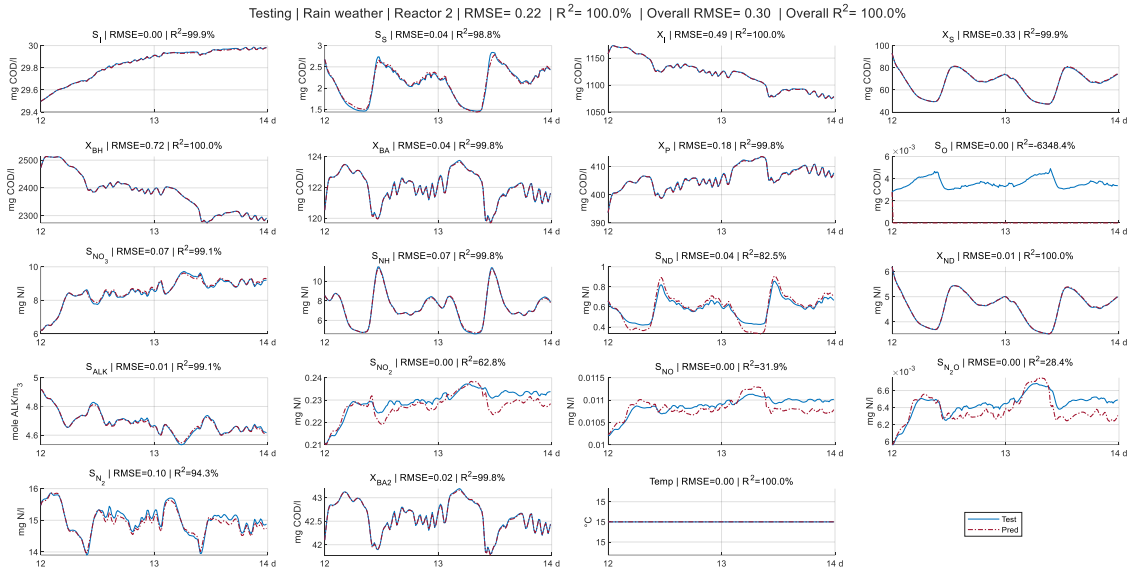
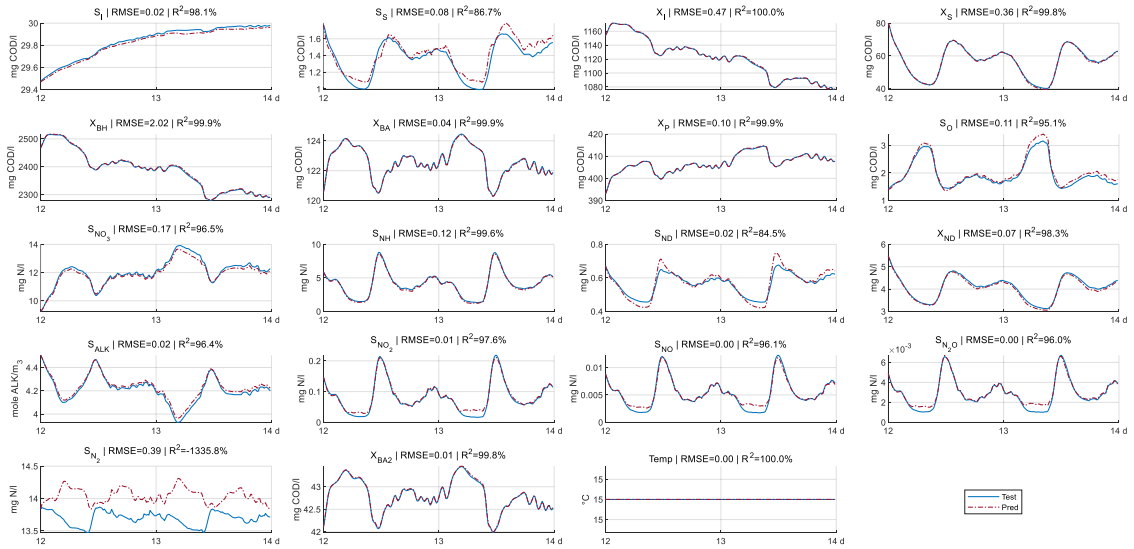
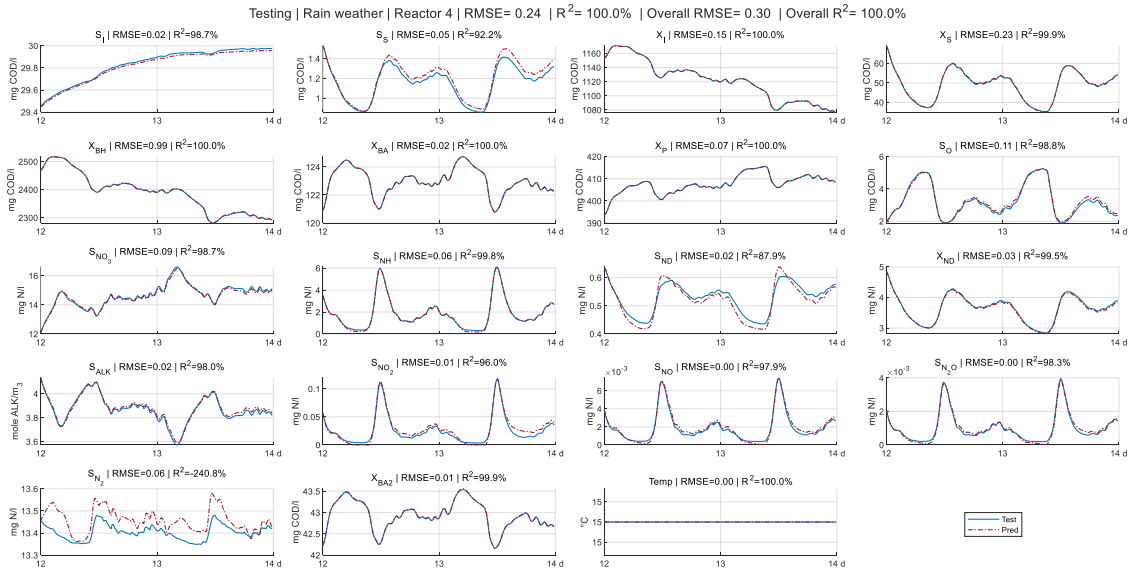


Figure 8.6 Testing results for day 12-14 of dry weather scenario




Testing | Rain weather | Reactor 2 | RMSE= 0.22 | R2= 100.0% | Overall RMSE= 0.30 | Overall R2= 100.0%
SI | RMSE=0.00 | R2=99.9%
SS | RMSE=0.04 | R2=98.8%
XI | RMSE=0.49 | R2=100.0%
XS | RMSE=0.33 | R2=99.9%
XBH | RMSE=0.72 | R2=100.0%
XBA | RMSE=0.04 | R2=99.8%
XP | RMSE=0.18 | R2=99.8%
SO | RMSE=0.00 | R2=-6348.4%
SNO3 | RMSE=0.07 | R2=99.1%
SNH | RMSE=0.07 | R2=99.8%
SND | RMSE=0.04 | R2=82.5%
XND | RMSE=0.01 | R2=100.0%
SALK | RMSE=0.01 | R2=99.1%
SNO2 | RMSE=0.00 | R2=62.8%
SNO | RMSE=0.00 | R2=31.9%
SN2O | RMSE=0.00 | R2=28.4%
SN2 | RMSE=0.10 | R2=94.3%
XBA2 | RMSE=0.02 | R2=99.8%
Temp | RMSE=0.00 | R2=100.0%
mg COD/l
mg N/l
mole ALK/m3
°C
Test
Pred


Testing | Rain weather | Reactor 3 | RMSE= 0.49 | R2= 99.9% | Overall RMSE= 0.30 | Overall R2= 100.0%
SI | RMSE=0.02 | R2=98.1%
SS | RMSE=0.08 | R2=86.7%
XI | RMSE=0.47 | R2=100.0%
XS | RMSE=0.36 | R2=99.8%
XBH | RMSE=2.02 | R2=99.9%
XBA | RMSE=0.04 | R2=99.9%
XP | RMSE=0.10 | R2=99.9%
SO | RMSE=0.11 | R2=95.1%
SNO3 | RMSE=0.17 | R2=96.5%
SNH | RMSE=0.12 | R2=99.6%
SND | RMSE=0.02 | R2=84.5%
XND | RMSE=0.07 | R2=98.3%
SALK | RMSE=0.02 | R2=96.4%
SNO2 | RMSE=0.01 | R2=97.6%
SNO | RMSE=0.00 | R2=96.1%
SN2O | RMSE=0.00 | R2=96.0%
SN2 | RMSE=0.39 | R2=-1335.8%
XBA2 | RMSE=0.01 | R2=99.8%
Temp | RMSE=0.00 | R2=100.0%
mg COD/l
mg N/l
mole ALK/m3
°C
Test
Pred


Testing | Rain weather | Reactor 4 | RMSE= 0.24 | R2= 100.0% | Overall RMSE= 0.30 | Overall R2= 100.0%
SI | RMSE=0.02 | R2=98.7%
SS | RMSE=0.05 | R2=92.2%
XI | RMSE=0.15 | R2=100.0%
XS | RMSE=0.23 | R2=99.9%
XBH | RMSE=0.99 | R2=100.0%
XBA | RMSE=0.02 | R2=100.0%
XP | RMSE=0.07 | R2=100.0%
SO | RMSE=0.11 | R2=98.8%
SNO3 | RMSE=0.09 | R2=98.7%
SNH | RMSE=0.06 | R2=99.8%
SND | RMSE=0.02 | R2=87.9%
XND | RMSE=0.03 | R2=99.5%
SALK | RMSE=0.02 | R2=98.0%
SNO2 | RMSE=0.01 | R2=96.0%
SNO | RMSE=0.00 | R2=97.9%
SN2O | RMSE=0.00 | R2=98.3%
SN2 | RMSE=0.06 | R2=-240.8%
XBA2 | RMSE=0.01 | R2=99.9%
Temp | RMSE=0.00 | R2=100.0%
mg COD/l
mg N/l
mole ALK/m3
°C
Test
Pred

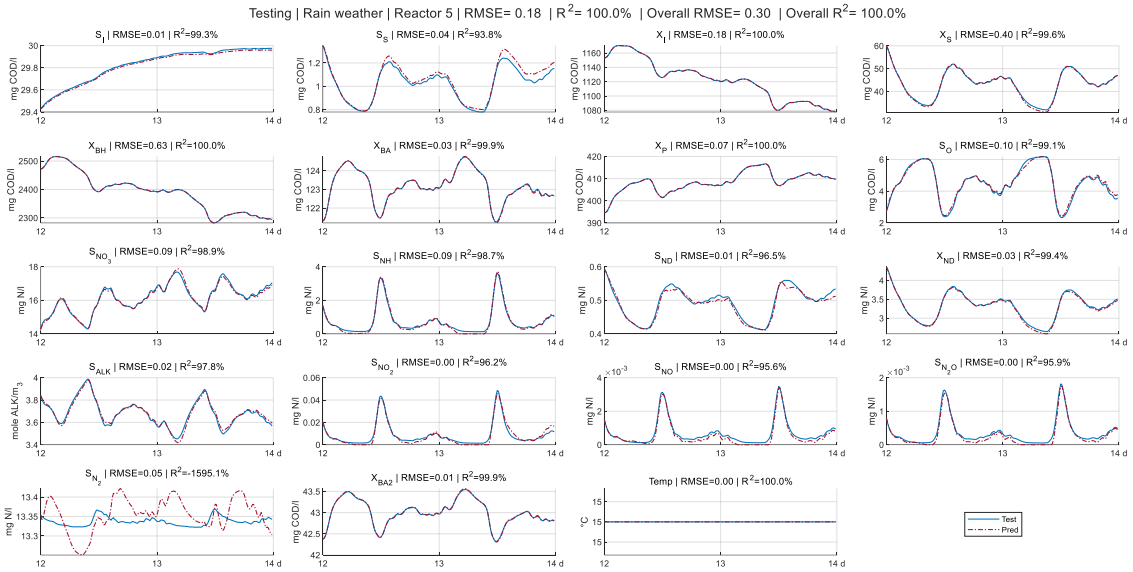
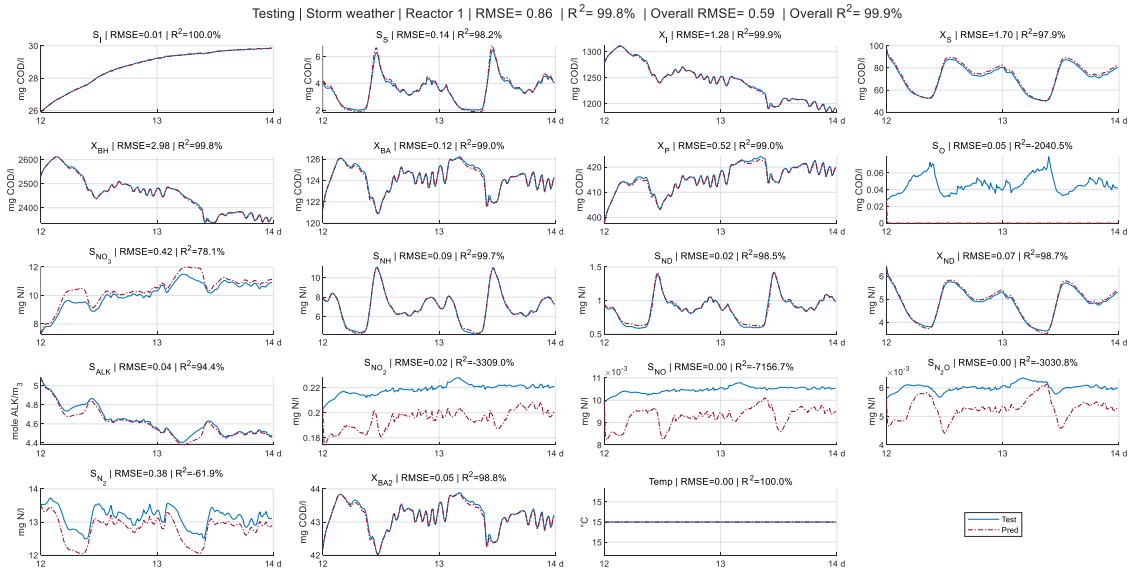
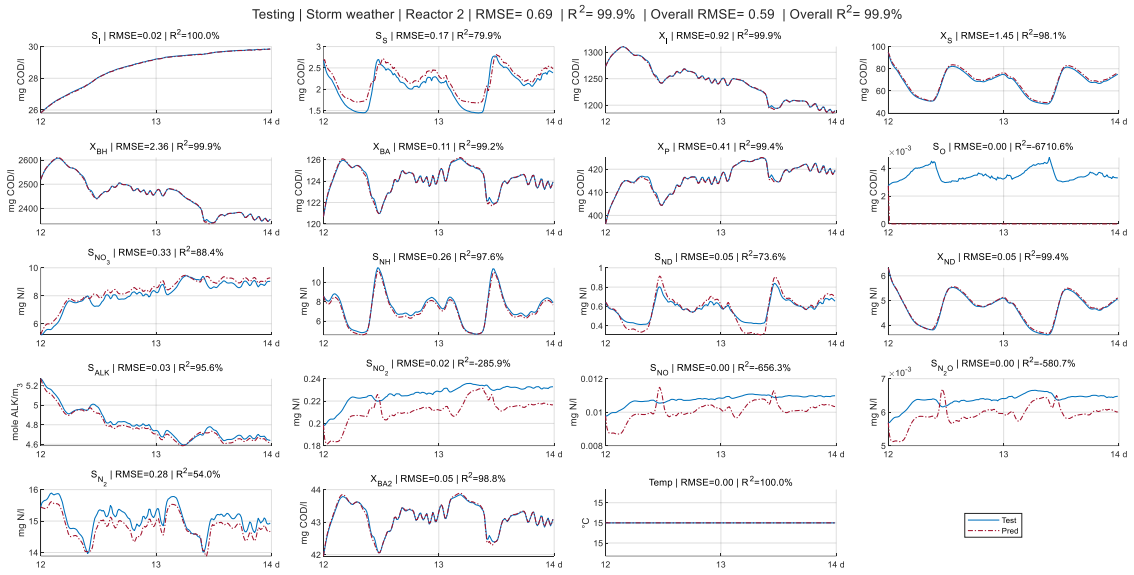
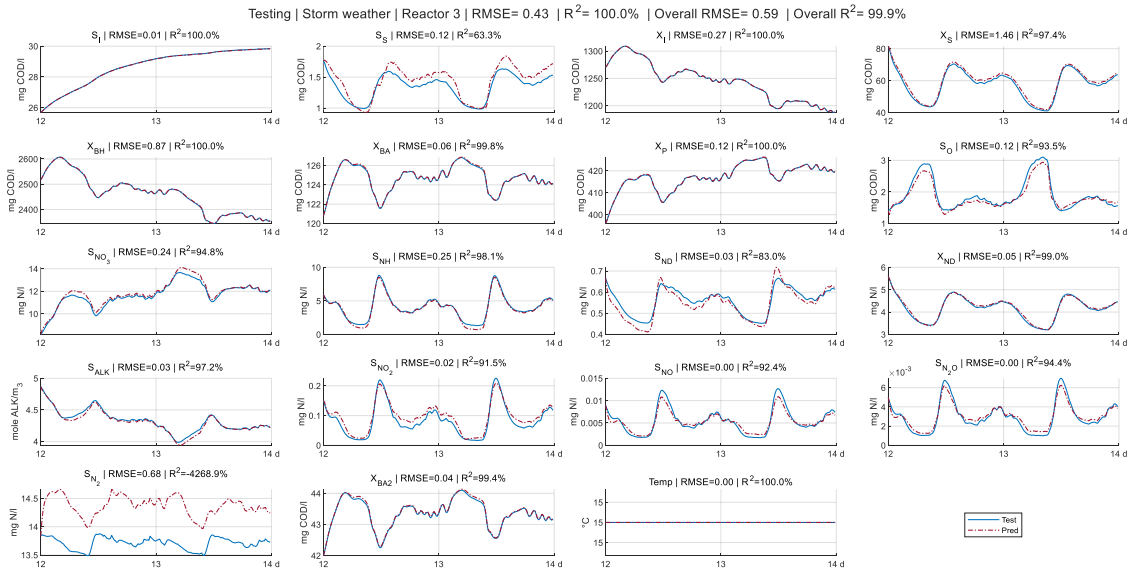
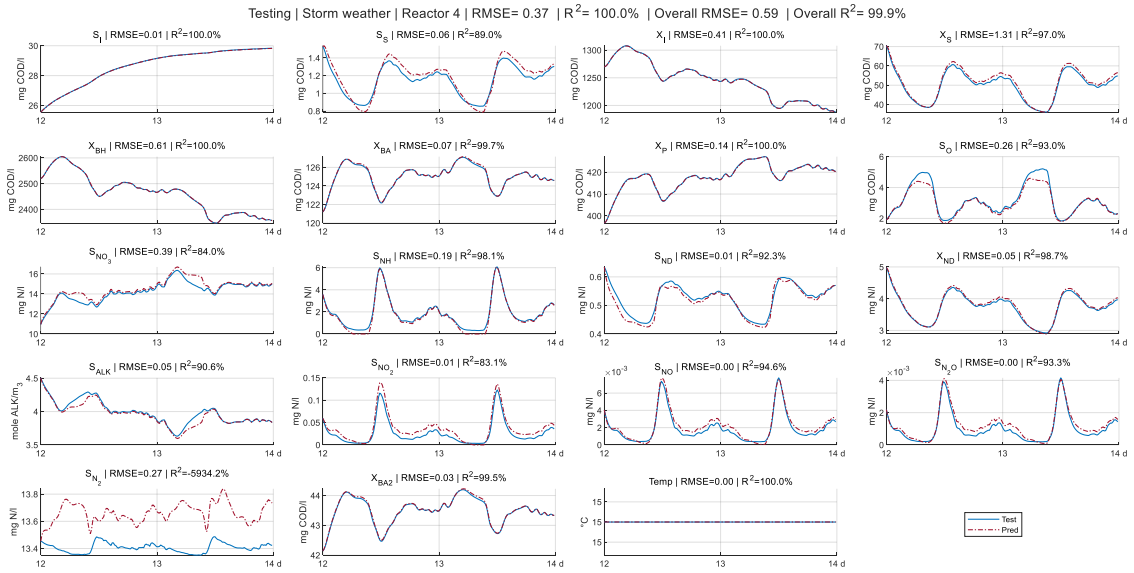


Figure 8.7 Testing results for day 12-14 of rain weather scenario




Testing | Storm weather | Reactor 2 | RMSE= 0.69 | R2= 99.9% | Overall RMSE= 0.59 | Overall R2= 99.9%
SI | RMSE=0.02 | R2=100.0%
SS | RMSE=0.17 | R2=79.9%
XI | RMSE=0.92 | R2=99.9%
XS | RMSE=1.45 | R2=98.1%
XBH | RMSE=2.36 | R2=99.9%
XBA | RMSE=0.11 | R2=99.2%
XP | RMSE=0.41 | R2=99.4%
SO | RMSE=0.00 | R2=-6710.6%
SNO3 | RMSE=0.33 | R2=88.4%
SNH | RMSE=0.26 | R2=97.6%
SND | RMSE=0.05 | R2=73.6%
XND | RMSE=0.05 | R2=99.4%
SALK | RMSE=0.03 | R2=95.6%
SNO2 | RMSE=0.02 | R2=-285.9%
SNO | RMSE=0.00 | R2=-656.3%
SN2O | RMSE=0.00 | R2=-580.7%
SN2 | RMSE=0.28 | R2=54.0%
XBA2 | RMSE=0.05 | R2=98.8%
Temp | RMSE=0.00 | R2=100.0%
Test
Pred


Testing | Storm weather | Reactor 3 | RMSE= 0.43 | R2= 100.0% | Overall RMSE= 0.59 | Overall R2= 99.9%
SI | RMSE=0.01 | R2=100.0%
SS | RMSE=0.12 | R2=63.3%
XI | RMSE=0.27 | R2=100.0%
XS | RMSE=1.46 | R2=97.4%
XBH | RMSE=0.87 | R2=100.0%
XBA | RMSE=0.06 | R2=99.8%
XP | RMSE=0.12 | R2=100.0%
SO | RMSE=0.12 | R2=93.5%
SNO3 | RMSE=0.24 | R2=94.8%
SNH | RMSE=0.25 | R2=98.1%
SND | RMSE=0.03 | R2=83.0%
XND | RMSE=0.05 | R2=99.0%
SALK | RMSE=0.03 | R2=97.2%
SNO2 | RMSE=0.02 | R2=91.5%
SNO | RMSE=0.00 | R2=92.4%
SN2O | RMSE=0.00 | R2=94.4%
SN2 | RMSE=0.68 | R2=-4268.9%
XBA2 | RMSE=0.04 | R2=99.4%
Temp | RMSE=0.00 | R2=100.0%
Test
Pred


Testing | Storm weather | Reactor 4 | RMSE= 0.37 | R2= 100.0% | Overall RMSE= 0.59 | Overall R2= 99.9%
SI | RMSE=0.01 | R2=100.0%
SS | RMSE=0.06 | R2=89.0%
XI | RMSE=0.41 | R2=100.0%
XS | RMSE=1.31 | R2=97.0%
XBH | RMSE=0.61 | R2=100.0%
XBA | RMSE=0.07 | R2=99.7%
XP | RMSE=0.14 | R2=100.0%
SO | RMSE=0.26 | R2=93.0%
SNO3 | RMSE=0.39 | R2=84.0%
SNH | RMSE=0.19 | R2=98.1%
SND | RMSE=0.01 | R2=92.3%
XND | RMSE=0.05 | R2=98.7%
SALK | RMSE=0.05 | R2=90.6%
SNO2 | RMSE=0.01 | R2=83.1%
SNO | RMSE=0.00 | R2=94.6%
SN2O | RMSE=0.00 | R2=93.3%
SN2 | RMSE=0.27 | R2=-5934.2%
XBA2 | RMSE=0.03 | R2=99.5%
Temp | RMSE=0.00 | R2=100.0%
Test
Pred

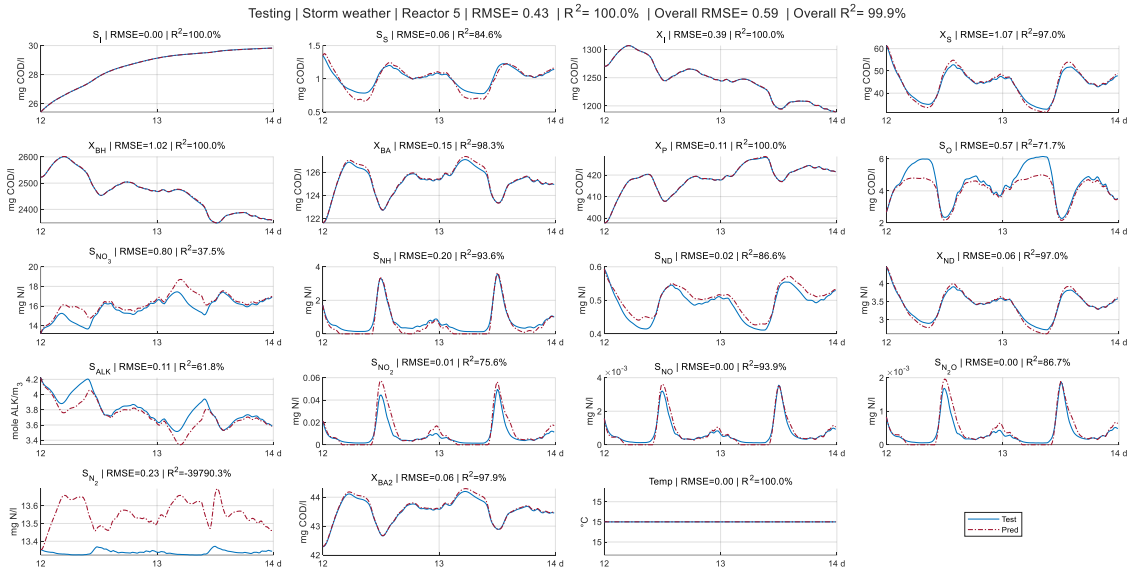
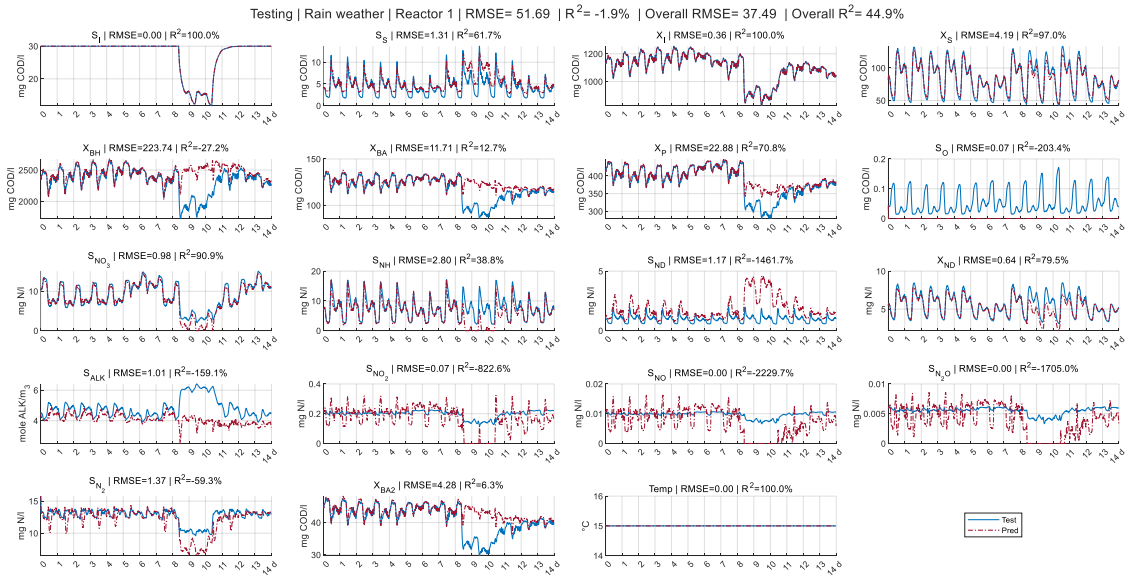
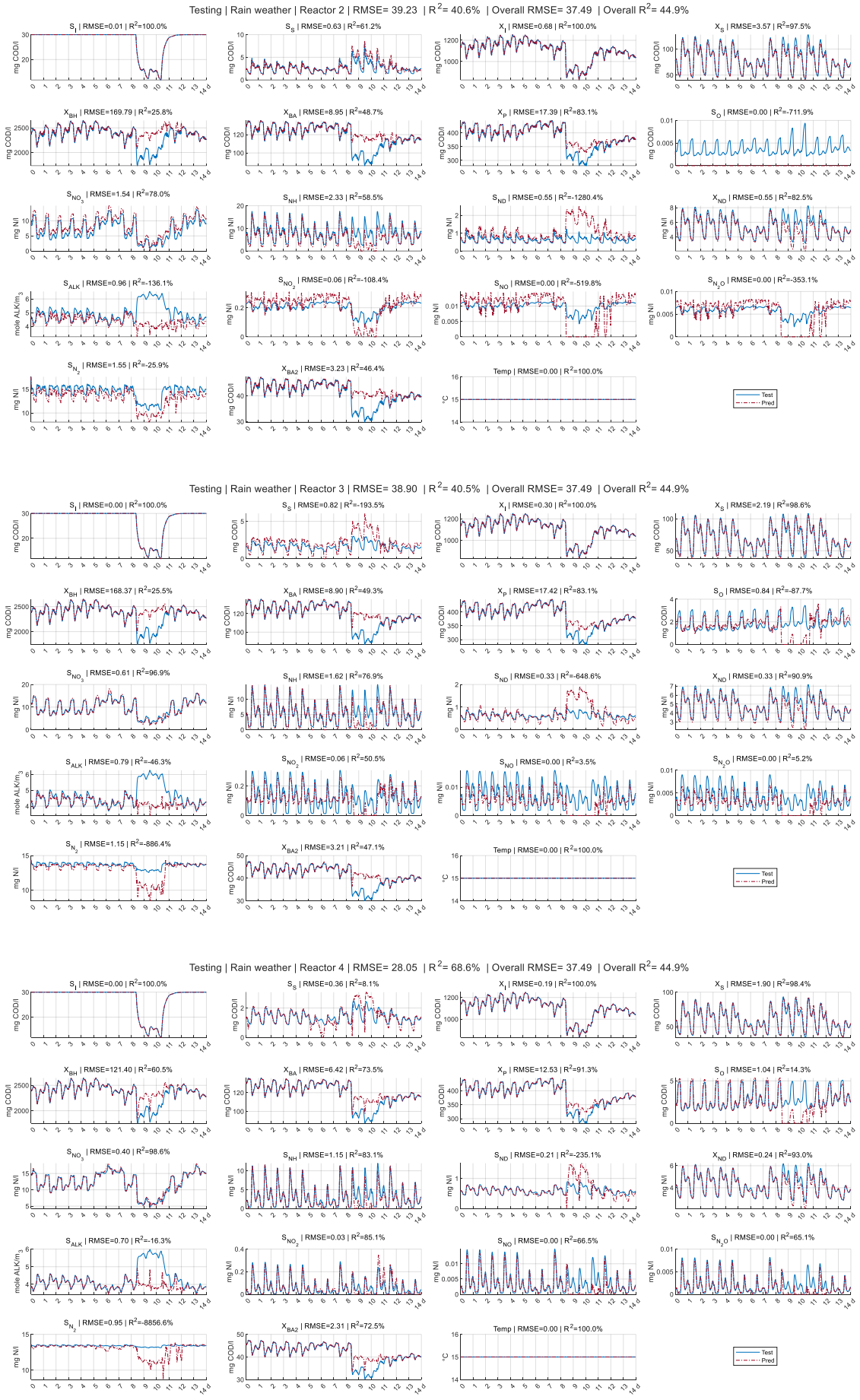


Figure 8.8 Testing results for day 12-14 of storm weather scenario

8.4 Results of reverse-scenario validation




Testing | Rain weather | Reactor 2 | RMSE= 39.23 | R2= 40.6% | Overall RMSE= 37.49 | Overall R2= 44.9%
SI | RMSE=0.01 | R2=100.0%
SS | RMSE=0.63 | R2=61.2%
XI | RMSE=0.68 | R2=100.0%
XS | RMSE=3.57 | R2=97.5%
XBH | RMSE=169.79 | R2=25.8%
XBA | RMSE=8.95 | R2=48.7%
XP | RMSE=17.39 | R2=83.1%
SO | RMSE=0.00 | R2=-711.9%
SNO3 | RMSE=1.54 | R2=78.0%
SNH | RMSE=2.33 | R2=58.5%
SND | RMSE=0.55 | R2=-1280.4%
XND | RMSE=0.55 | R2=82.5%
SALK | RMSE=0.96 | R2=-136.1%
SNO2 | RMSE=0.06 | R2=-108.4%
SNO | RMSE=0.00 | R2=-519.8%
SN2O | RMSE=0.00 | R2=-353.1%
SN2 | RMSE=1.55 | R2=-25.9%
XBA2 | RMSE=3.23 | R2=46.4%
Temp | RMSE=0.00 | R2=100.0%
Test
Pred
Testing | Rain weather | Reactor 3 | RMSE= 38.90 | R2= 40.5% | Overall RMSE= 37.49 | Overall R2= 44.9%
SI | RMSE=0.00 | R2=100.0%
SS | RMSE=0.82 | R2=-193.5%
XI | RMSE=0.30 | R2=100.0%
XS | RMSE=2.19 | R2=98.6%
XBH | RMSE=168.37 | R2=25.5%
XBA | RMSE=8.90 | R2=49.3%
XP | RMSE=17.42 | R2=83.1%
SO | RMSE=0.84 | R2=-87.7%
SNO3 | RMSE=0.61 | R2=96.9%
SNH | RMSE=1.62 | R2=76.9%
SND | RMSE=0.33 | R2=-648.6%
XND | RMSE=0.33 | R2=90.9%
SALK | RMSE=0.79 | R2=-46.3%
SNO2 | RMSE=0.06 | R2=50.5%
SNO | RMSE=0.00 | R2=3.5%
SN2O | RMSE=0.00 | R2=5.2%
SN2 | RMSE=1.15 | R2=-886.4%
XBA2 | RMSE=3.21 | R2=47.1%
Temp | RMSE=0.00 | R2=100.0%
Test
Pred
Testing | Rain weather | Reactor 4 | RMSE= 28.05 | R2= 68.6% | Overall RMSE= 37.49 | Overall R2= 44.9%
SI | RMSE=0.00 | R2=100.0%
SS | RMSE=0.36 | R2=8.1%
XI | RMSE=0.19 | R2=100.0%
XS | RMSE=1.90 | R2=98.4%
XBH | RMSE=121.40 | R2=60.5%
XBA | RMSE=6.42 | R2=73.5%
XP | RMSE=12.53 | R2=91.3%
SO | RMSE=1.04 | R2=14.3%
SNO3 | RMSE=0.40 | R2=98.6%
SNH | RMSE=1.15 | R2=83.1%
SND | RMSE=0.21 | R2=-235.1%
XND | RMSE=0.24 | R2=93.0%
SALK | RMSE=0.70 | R2=-16.3%
SNO2 | RMSE=0.03 | R2=85.1%
SNO | RMSE=0.00 | R2=66.5%
SN2O | RMSE=0.00 | R2=65.1%
SN2 | RMSE=0.95 | R2=-8856.6%
XBA2 | RMSE=2.31 | R2=72.5%
Temp | RMSE=0.00 | R2=100.0%
Test
Pred

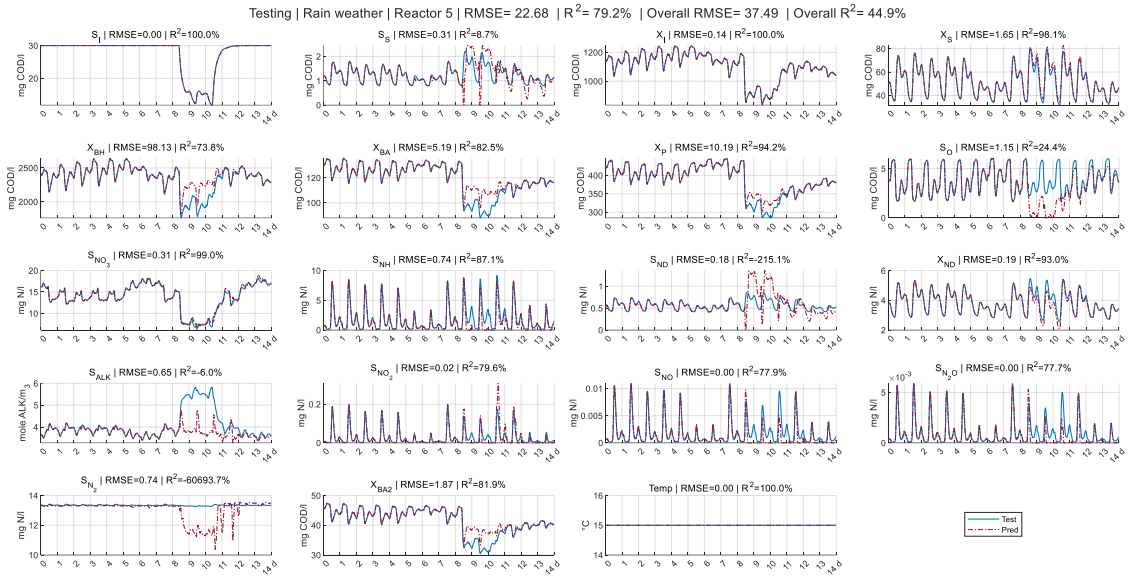
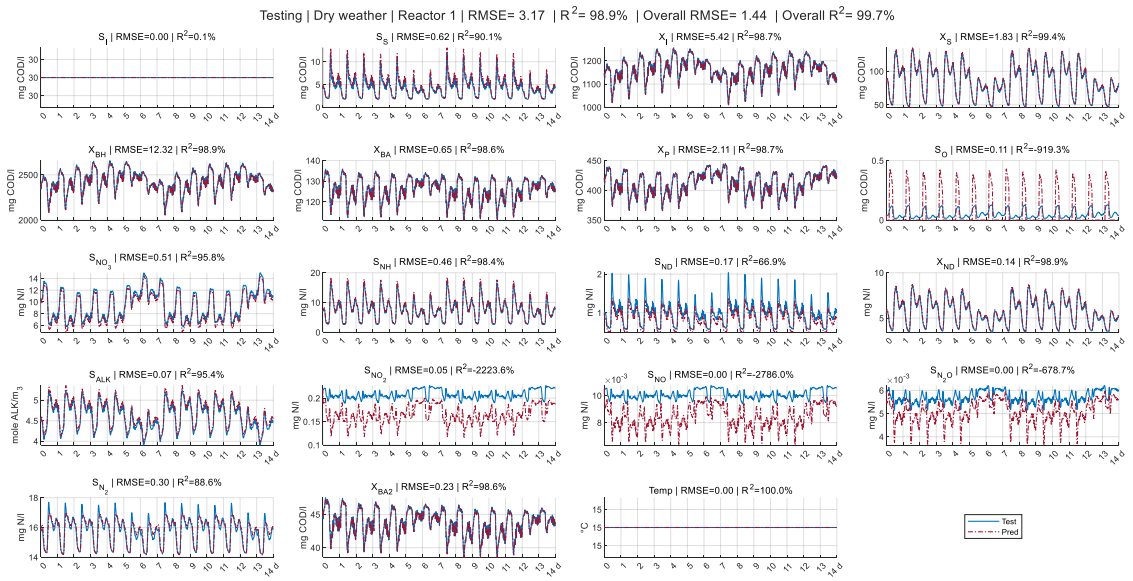
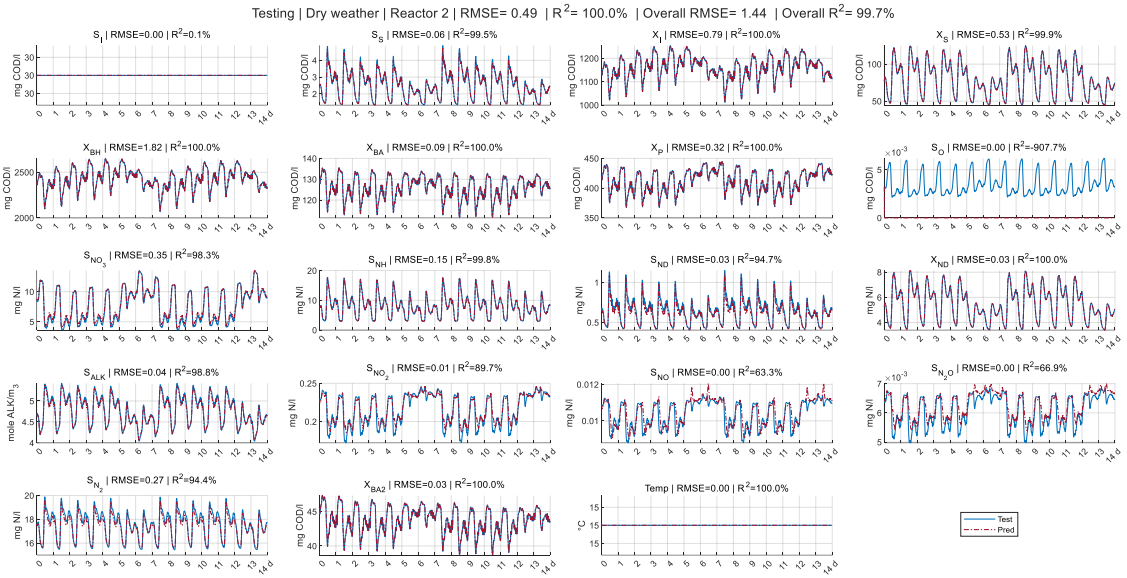
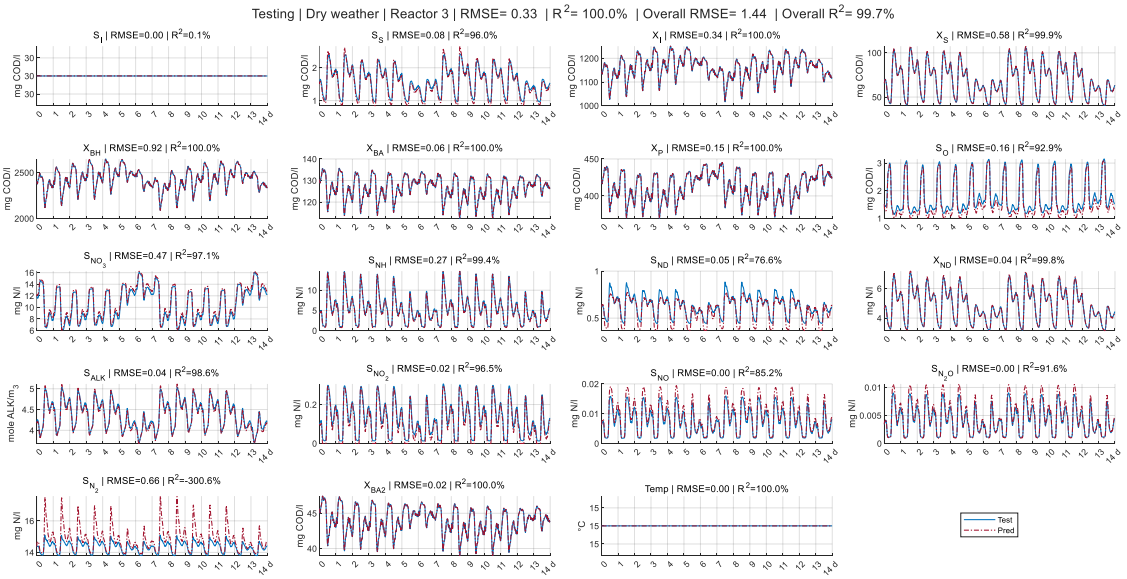
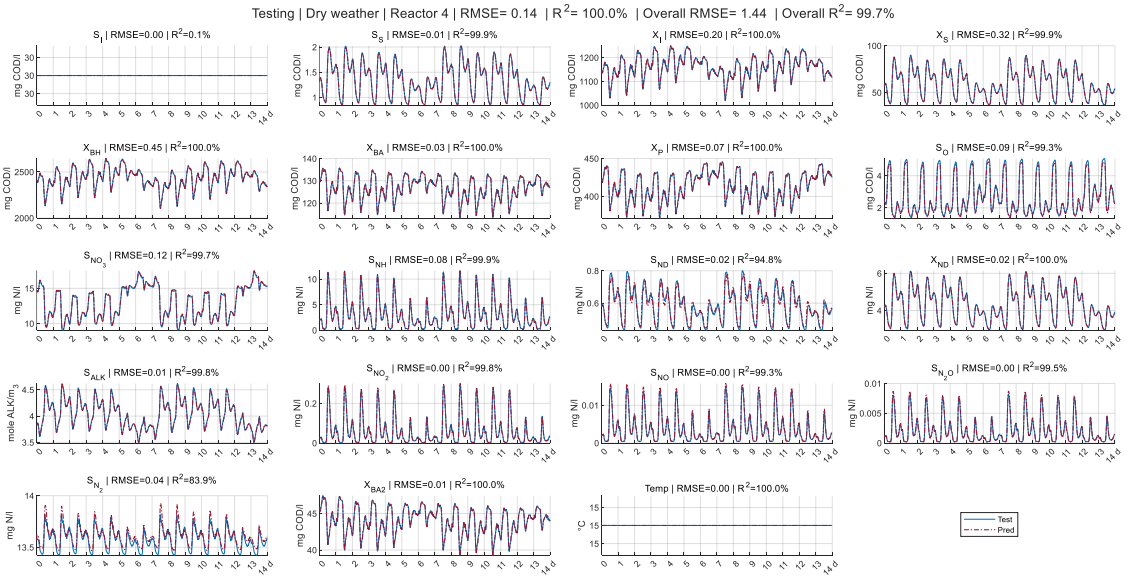


Figure 8.9 Testing dry weather data-trained model in rain weather scenario




Testing | Dry weather | Reactor 2 | RMSE= 0.49 | R2= 100.0% | Overall RMSE= 1.44 | Overall R2= 99.7%
SI | RMSE=0.00 | R2=0.1%
SS | RMSE=0.06 | R2=99.5%
XI | RMSE=0.79 | R2=100.0%
XS | RMSE=0.53 | R2=99.9%
XBH | RMSE=1.82 | R2=100.0%
XBA | RMSE=0.09 | R2=100.0%
XP | RMSE=0.32 | R2=100.0%
SO | RMSE=0.00 | R2=-907.7%
SNO3 | RMSE=0.35 | R2=98.3%
SNH | RMSE=0.15 | R2=99.8%
SND | RMSE=0.03 | R2=94.7%
XND | RMSE=0.03 | R2=100.0%
SALK | RMSE=0.04 | R2=98.8%
SNO2 | RMSE=0.01 | R2=89.7%
SNO | RMSE=0.00 | R2=63.3%
SN2O | RMSE=0.00 | R2=66.9%
SN2 | RMSE=0.27 | R2=94.4%
XBA2 | RMSE=0.03 | R2=100.0%
Temp | RMSE=0.00 | R2=100.0%
Test
Pred


Testing | Dry weather | Reactor 3 | RMSE= 0.33 | R2= 100.0% | Overall RMSE= 1.44 | Overall R2= 99.7%
SI | RMSE=0.00 | R2=0.1%
SS | RMSE=0.08 | R2=96.0%
XI | RMSE=0.34 | R2=100.0%
XS | RMSE=0.58 | R2=99.9%
XBH | RMSE=0.92 | R2=100.0%
XBA | RMSE=0.06 | R2=100.0%
XP | RMSE=0.15 | R2=100.0%
SO | RMSE=0.16 | R2=92.9%
SNO3 | RMSE=0.47 | R2=97.1%
SNH | RMSE=0.27 | R2=99.4%
SND | RMSE=0.05 | R2=76.6%
XND | RMSE=0.04 | R2=99.8%
SALK | RMSE=0.04 | R2=98.6%
SNO2 | RMSE=0.02 | R2=96.5%
SNO | RMSE=0.00 | R2=85.2%
SN2O | RMSE=0.00 | R2=91.6%
SN2 | RMSE=0.66 | R2=-300.6%
XBA2 | RMSE=0.02 | R2=100.0%
Temp | RMSE=0.00 | R2=100.0%
Test
Pred


Testing | Dry weather | Reactor 4 | RMSE= 0.14 | R2= 100.0% | Overall RMSE= 1.44 | Overall R2= 99.7%
SI | RMSE=0.00 | R2=0.1%
SS | RMSE=0.01 | R2=99.9%
XI | RMSE=0.20 | R2=100.0%
XS | RMSE=0.32 | R2=99.9%
XBH | RMSE=0.45 | R2=100.0%
XBA | RMSE=0.03 | R2=100.0%
XP | RMSE=0.07 | R2=100.0%
SO | RMSE=0.09 | R2=99.3%
SNO3 | RMSE=0.12 | R2=99.7%
SNH | RMSE=0.08 | R2=99.9%
SND | RMSE=0.02 | R2=94.8%
XND | RMSE=0.02 | R2=100.0%
SALK | RMSE=0.01 | R2=99.8%
SNO2 | RMSE=0.00 | R2=99.8%
SNO | RMSE=0.00 | R2=99.3%
SN2O | RMSE=0.00 | R2=99.5%
SN2 | RMSE=0.04 | R2=83.9%
XBA2 | RMSE=0.01 | R2=100.0%
Temp | RMSE=0.00 | R2=100.0%
Test
Pred

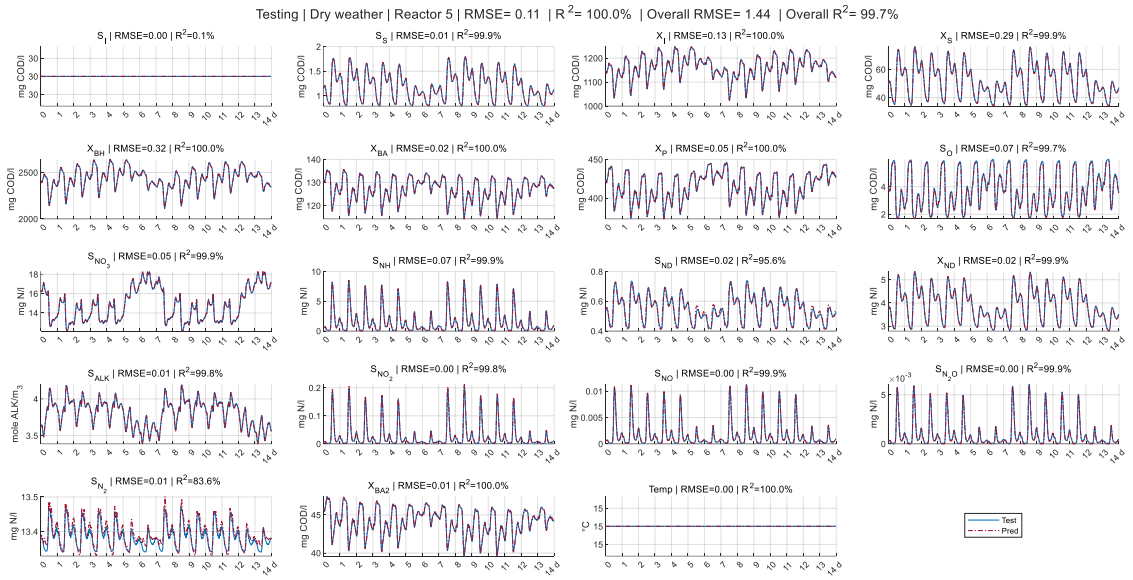


Figure 8.10 Testing rain weather data-trained model in dry weather scenario

8.5 Results of NODE with reduced dimensions

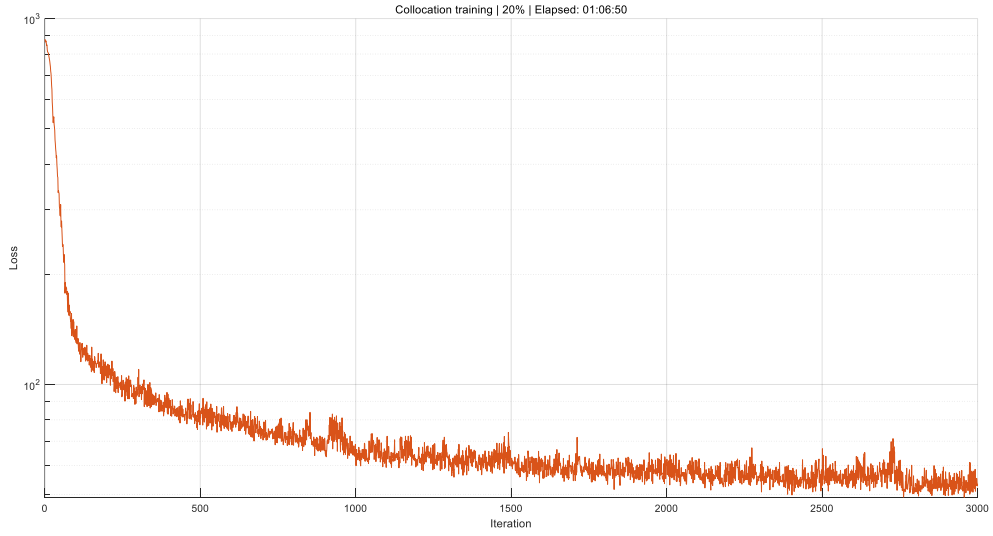


Figure 8.11 Collocation training loss on first 7-day dry weather scenario with reduced dimensions

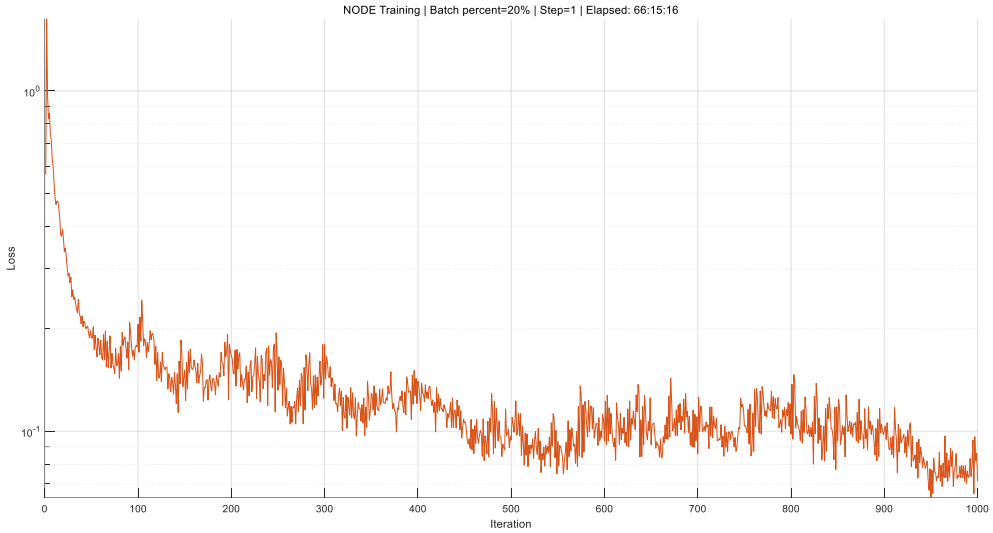
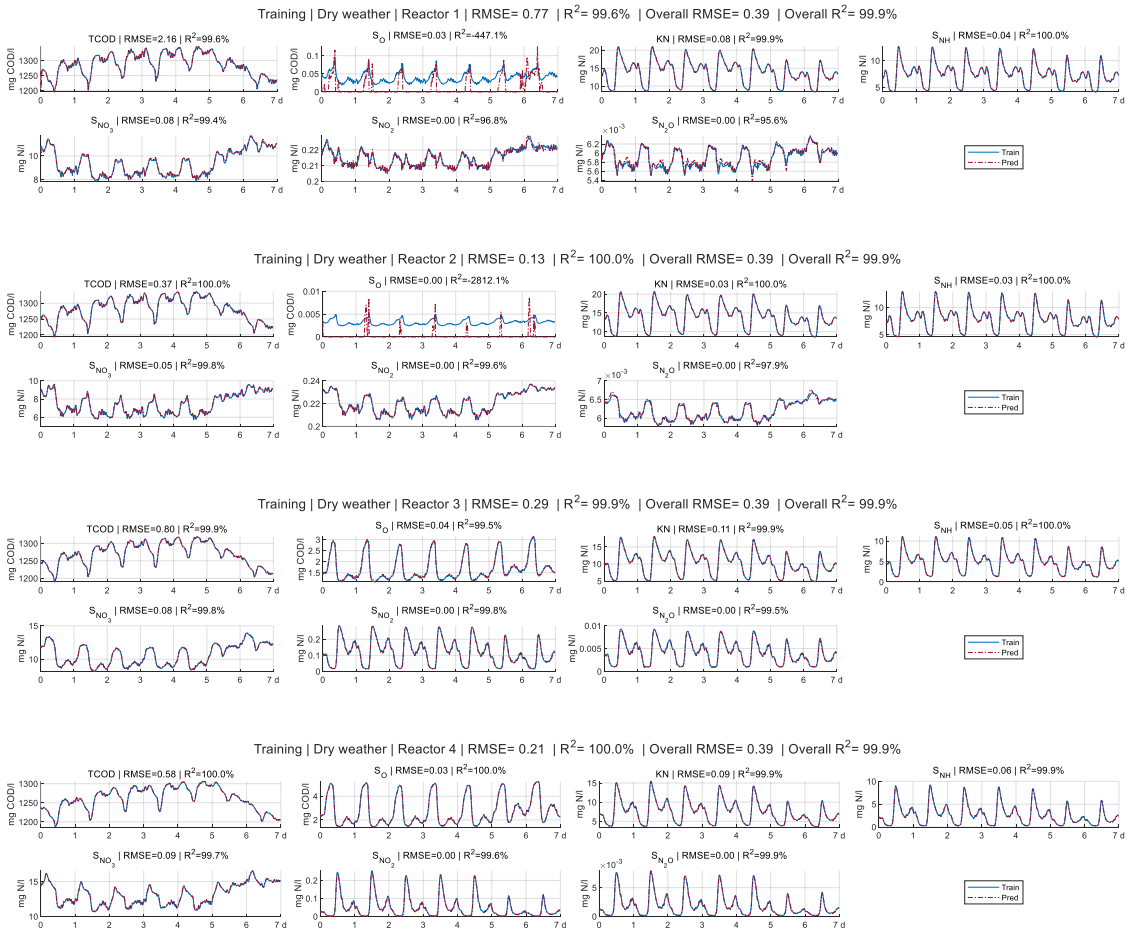


Figure 8.12 NODE training loss on first 7-day dry weather scenario with reduced dimensions



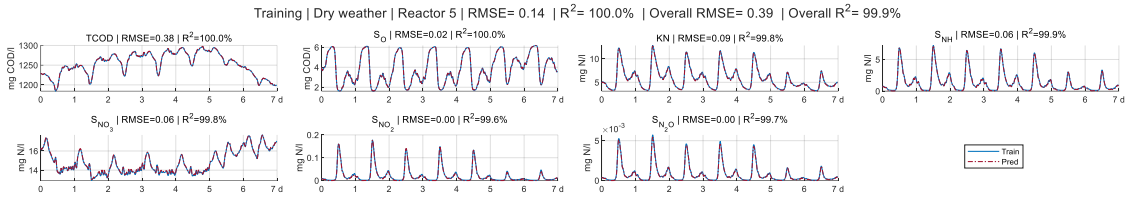


Figure 8.13 Training results using first 7-day dry weather data with reduced dimensions

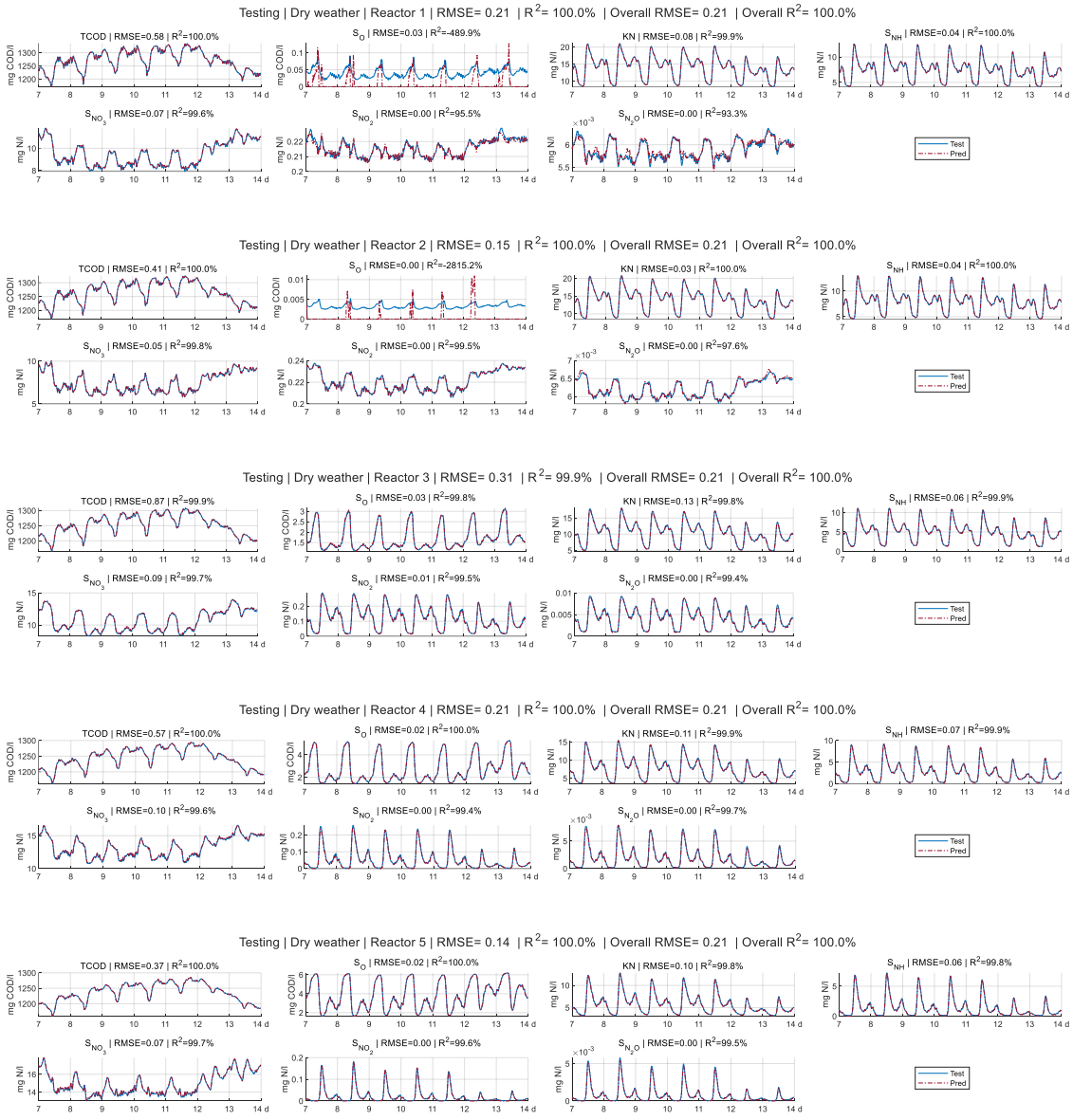


Figure 8.14 Testing results for day 7-14 of dry weather scenario with reduced dimensions

Chapter 9 References

A., Castellano-Hinojosa, Maza-Marquez P., Melero-Rubio Y., Gonzalez-Lopez J., and Rodelas B. 2018. 'Linking N2O Emissions to Population Dynamics of Nitrifying and Denitrifying Prokaryotes in 4 Full-Scale WWTP'. *Chemosphere* 200:57–66. https://doi.org/10.1016/j.chemosphere.2018.02.102.

Abba, S I, Gozen Elkiran, and Vahid Nourani. 2021. 'Improving Novel Extreme Learning Machine Using PCA Algorithms for Multi-Parametric Modeling of the Municipal Wastewater Treatment Plant'. *Desalin. Water Treat* 215:414–26.

Aboobakar, Amina, Elise Cartmell, Tom Stephenson, Mark Jones, Peter Vale, and Gabriela Dotro. 2013. 'Nitrous Oxide Emissions and Dissolved Oxygen Profiling in a Full-Scale Nitrifying Activated Sludge Treatment Plant'. *Water Research* 47 (2): 524–34. https://doi.org/10.1016/j.watres.2012.10.004.

Afshar, Abbas, Fariborz Massoumi, Amin Afshar, and Miquel A Mariño. 2015. 'State of the Art Review of Ant Colony Optimization Applications in Water Resource Management'. *Water Resources Management* 29:3891–3904.

Al, Resul, Chitta Ranjan Behera, Alexandr Zubov, Krist V Gernaey, and Gürkan Sin. 2019. 'Meta-Modeling Based Efficient Global Sensitivity Analysis for Wastewater Treatment Plants–An Application to the BSM2 Model'. *Computers & Chemical Engineering* 127:233–46.

Alex, J, L Benedetti, J Copp, K V Gernaey, U Jeppsson, I Nopens, M N Pons, J P Steyer, and P Vanrolleghem. 2018. 'Benchmark Simulation Model No. 1 (BSM1)'.

Almeida, J S, M A M Reis, and M J T Carrondo. 1997. 'A Unifying Kinetic Model of Denitrification'. *Journal of Theoretical Biology* 186 (2): 241–49.

Alvi, Maira, Damien Batstone, Christian Kazadi Mbamba, Philip Keymer, Tim French, Andrew Ward, Jason Dwyer, and Rachel Cardell-Oliver. 2023. 'Deep Learning in Wastewater Treatment: A Critical Review'. *Water Research*. Elsevier Ltd. https://doi.org/10.1016/j.watres.2023.120518.

Amari, Shun ichi. 1993. ‘Backpropagation and Stochastic Gradient Descent Method’. *Neurocomputing* 5 (4–5): 185–96. https://doi.org/10.1016/0925-2312(93)90006-O.

Anderson, J H. 1964. ‘The Metabolism of Hydroxylamine to Nitrite by Nitrosomonas’. *Biochemical Journal* 91 (1): 8.

Aponte-Rengifo, Oscar, Mario Francisco, Ramón Vilanova, Pastora Vega, and Silvana Revollar. 2023. ‘Intelligent Control of Wastewater Treatment Plants Based on Model-Free Deep Reinforcement Learning’. *Processes* 11 (8): 2269.

Asadi, Mohsen, and Kerry Neil McPhedran. 2021. ‘Greenhouse Gas Emission Estimation from Municipal Wastewater Using a Hybrid Approach of Generative Adversarial Network and Data-Driven Modelling’. *Science of The Total Environment* 800:149508. https://doi.org/10.1016/j.scitotenv.2021.149508.

Ba-Alawi, Abdulrahman H, Jorge Loy-Benitez, SangYun Kim, and ChangKyoo Yoo. 2022. ‘Missing Data Imputation and Sensor Self-Validation towards a Sustainable Operation of Wastewater Treatment Plants via Deep Variational Residual Autoencoders’. *Chemosphere* 288:132647.

Ba-Alawi, Abdulrahman H, Paulina Vilela, Jorge Loy-Benitez, SungKu Heo, and ChangKyoo Yoo. 2021. ‘Intelligent Sensor Validation for Sustainable Influent Quality Monitoring in Wastewater Treatment Plants Using Stacked Denoising Autoencoders’. *Journal of Water Process Engineering* 43:102206.

Baalbaki, Zeina, Elena Torfs, Viviane Yargeau, and Peter A Vanrolleghem. 2017. ‘Predicting the Fate of Micropollutants during Wastewater Treatment: Calibration and Sensitivity Analysis’. *Science of the Total Environment* 601:874–85.

Bae, Wo Bin, Yongeun Park, Kartik Chandran, Jingyeong Shin, Sung Bong Kang, Jinhua Wang, and Young Mo Kim. 2021. ‘Temporal Triggers of N2O Emissions during Cyclical and Seasonal Variations of a Full-Scale Sequencing Batch Reactor Treating Municipal Wastewater’. *Science of The Total Environment* 797:149093.

Bagheri, Majid, Sayed Ahmad Mirbagheri, Zahra Bagheri, and Ali Morad Kamarkhani. 2015. ‘Modeling and Optimization of Activated Sludge Bulking for a Real

Wastewater Treatment Plant Using Hybrid Artificial Neural Networks-Genetic Algorithm Approach'. *Process Safety and Environmental Protection* 95:12–25.

Bagherzadeh, Faramarz, Mohamad-Javad Mehrani, Milad Basirifard, and Javad Roostaei. 2021. 'Comparative Study on Total Nitrogen Prediction in Wastewater Treatment Plant and Effect of Various Feature Selection Methods on Machine Learning Algorithms Performance'. *Journal of Water Process Engineering* 41:102033.

Bahramian, Majid, Recep Kaan Dereli, Wanqing Zhao, Matteo Giberti, and Eoin Casey. 2023. 'Data to Intelligence: The Role of Data-Driven Models in Wastewater Treatment'. *Expert Systems with Applications* 217:119453.

Baker, Justin, Hedi Xia, Yiwei Wang, Elena Cherkaev, Akil Narayan, Long Chen, Jack Xin, Andrea L Bertozzi, Stanley J Osher, and Bao Wang. 2022. 'Proximal Implicit ODE Solvers for Accelerating Learning Neural ODEs'. *ArXiv Preprint ArXiv:2204.08621*.

Baldi, Pierre, and Peter J Sadowski. 2013. 'Understanding Dropout'. *Advances in Neural Information Processing Systems* 26.

Beaumont, Hubertus J E, Bas van Schooten, Sylvia I Lens, Hans V Westerhoff, and Rob J M van Spanning. 2004. 'Nitrosomonas Europaea Expresses a Nitric Oxide Reductase during Nitrification'. *Journal of Bacteriology* 186 (13): 4417–21.

Belia, E, Y Amerlinck, L Benedetti, B Johnson, G Sin, P A Vanrolleghem, K V Gernaey, et al. 2009. 'Wastewater Treatment Modelling: Dealing with Uncertainties'. *Water Science and Technology* 60 (8): 1929–41. https://doi.org/10.2166/wst.2009.225.

Bellandi, Giacomo, Stefan Weijers, Riccardo Gori, and Ingmar Nopens. 2020. 'Towards an Online Mitigation Strategy for N2O Emissions through Principal Components Analysis and Clustering Techniques'. *Journal of Environmental Management* 261:110219. https://doi.org/10.1016/j.jenvman.2020.110219.

Beraud, Beno\^\it, Cyrille Lemoine, and Jean-Philippe Steyer. 2009. 'Multiobjective Genetic Algorithms for the Optimisation of Wastewater Treatment Processes'. In *Computational Intelligence Techniques for Bioprocess Modelling, Supervision and Control*, 163–95. Springer.

Béraud, Benoit, J P Steyer, C Lemoine, E Latrille, G Manic, and C Printemps-Vacquier. 2007. 'Towards a Global Multi Objective Optimization of Wastewater Treatment Plant Based on Modeling and Genetic Algorithms'. *Water Science and Technology* 56 (9): 109–16.

Bergna, Richard, Felix Opolka, Pietro Liò, and Jose Miguel Hernandez-Lobato. 2023. 'Graph Neural Stochastic Differential Equations'. *ArXiv Preprint ArXiv:2308.12316*.

Berkhahn, Sarah, and Matthias Ehrhardt. 2022. 'A Physics-Informed Neural Network to Model COVID-19 Infection and Hospitalization Scenarios'. *Advances in Continuous and Discrete Models* 2022 (1). https://doi.org/10.1186/s13662-022-03733-5.

Bernardelli, A, S Marsili-Libelli, A Manzini, S Stancari, G Tardini, D Montanari, G Anceschi, P Gelli, and S Venier. 2020. 'Real-Time Model Predictive Control of a Wastewater Treatment Plant Based on Machine Learning'. *Water Science and Technology* 81 (11): 2391–2400.

Berthouex, P M, and George E Box. 1996. 'Time Series Models for Forecasting Wastewater Treatment Plant Performance'. *Water Research* 30 (8): 1865–75.

Bhanja, Samit, and Abhishek Das. 2018. 'Impact of Data Normalization on Deep Neural Network for Time Series Forecasting'.

Biswas, B N, Somnath Chatterjee, S P Mukherjee, and Subhradeep Pal. 2013. 'A Discussion on Euler Method: A Review'. *Electronic Journal of Mathematical Analysis and Applications* 1 (2): 2090–2792.

Boiocchi, Riccardo, Krist V. Gernaey, and Gürkan Sin. 2017. 'A Novel Fuzzy-Logic Control Strategy Minimizing N2O Emissions'. *Water Research* 123:479–94. https://doi.org/10.1016/j.watres.2017.06.074.

Bollon, Julien, Ahlem Filali, Yannick Fayolle, Sabrina Guerin, Vincent Rocher, and Sylvie Gillot. 2016. 'Full-Scale Post Denitrifying Biofilters: Sinks of Dissolved N2O?' *Science of the Total Environment* 563:320–28.

Bosch, Nathanael, Adrien Corenflos, Fatemeh Yaghoobi, Filip Tronarp, Philipp Hennig, and Simo Särkkä. 2024. ‘Parallel-in-Time Probabilistic Numerical ODE Solvers’. *Journal of Machine Learning Research* 25 (206): 1–27.

Böttcher, Lucas, and Thomas Asikis. 2022. ‘Near-Optimal Control of Dynamical Systems with Neural Ordinary Differential Equations’. *Machine Learning: Science and Technology* 3 (4). https://doi.org/10.1088/2632-2153/ac92c3.

Bradley, William, Gabriel S Gusmão, Andrew J Medford, and Fani Boukouvala. 2022. ‘Training Stiff Dynamic Process Models via Neural Differential Equations’. In *Computer Aided Chemical Engineering*, 49:1741–46. Elsevier.

Bradley, William, and Th Yr. 2022. ‘Parameter Estimation of Mechanistic Differential Equations via Neural Differential Equations’. https://www.machinedesign.com/learning.

Brown, Thomas S, Harbir Antil, Rainald Löhner, Fumiya Togashi, and Deepanshu Verma. 2021. ‘Novel Dnns for Stiff Odes with Applications to Chemically Reacting Flows’. In *High Performance Computing: ISC High Performance Digital 2021 International Workshops, Frankfurt Am Main, Germany, June 24–July 2, 2021, Revised Selected Papers 36*, 23–39.

Brunton, Steven L, Joshua L Proctor, and J Nathan Kutz. 2016. ‘Sparse Identification of Nonlinear Dynamics with Control (SINDYc)’. *IFAC-PapersOnLine* 49 (18): 710–15.

Budach, Lukas, Moritz Feuerpfeil, Nina Ihde, Andrea Nathansen, Nele Noack, Hendrik Patzlaff, Felix Naumann, and Hazar Harmouch. 2022. ‘The Effects of Data Quality on Machine Learning Performance’. *ArXiv Preprint ArXiv:2207.14529*.

Cabello-Solorzano, Kelsy, Isabela Ortigosa de Araujo, Marco Peña, Luís Correia, and Antonio J. Tallón-Ballesteros. 2023. ‘The Impact of Data Normalization on the Accuracy of Machine Learning Algorithms: A Comparative Analysis’. In *18th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2023)*, edited by Pablo García Bringas, Hilde Pérez García, Francisco Javier Martínez de Pisón, Francisco Martínez Álvarez,

Alicia Troncoso Lora, Álvaro Herrero, José Luis Calvo Rolle, Héctor Quintián, and Emilio Corchado, 344–53. Cham: Springer Nature Switzerland.

Cantera, J Jason L, and Lisa Y Stein. 2007. 'Role of Nitrite Reductase in the Ammonia-Oxidizing Pathway of Nitrosomonas Europaea'. *Archives of Microbiology* 188:349–54.

Cao, Longbing. 2022. 'Beyond Iid: Non-Iid Thinking, Informatics, and Learning'. *IEEE Intelligent Systems* 37 (4): 5–17.

Caranto, Jonathan D, and Kyle M Lancaster. 2017. 'Nitric Oxide Is an Obligate Bacterial Nitrification Intermediate Produced by Hydroxylamine Oxidoreductase'. *Proceedings of the National Academy of Sciences* 114 (31): 8217–22.

Caranto, Jonathan D, Avery C Vilbert, and Kyle M Lancaster. 2016. 'Nitrosomonas Europaea Cytochrome P460 Is a Direct Link between Nitrification and Nitrous Oxide Emission'. *Proceedings of the National Academy of Sciences* 113 (51): 14704–9.

Carrasco, Ivan Jose, and Shoou-Yuh Chang. 2005. 'Random Monte Carlo Simulation Analysis and Risk Assessment for Ammonia Concentrations in Wastewater Effluent Disposal'. *Stochastic Environmental Research and Risk Assessment* 19:134–45.

Cartwright, Julyan H E, and Oreste Piro. 1992. 'The Dynamics of Runge–Kutta Methods'. *International Journal of Bifurcation and Chaos* 2 (03): 427–49.

Casciotti, Karen L, and Bess B Ward. 2001. 'Dissimilatory Nitrite Reductase Genes from Autotrophic Ammonia-Oxidizing Bacteria'. *Applied and Environmental Microbiology* 67 (5): 2213–21.

Castellano-Hinojosa, A, P Maza-Márquez, Y Melero-Rubio, J González-López, and B Rodelas. 2018. 'Linking Nitrous Oxide Emissions to Population Dynamics of Nitrifying and Denitrifying Prokaryotes in Four Full-Scale Wastewater Treatment Plants'. *Chemosphere* 200:57–66. https://doi.org/10.1016/j.chemosphere.2018.02.102.

Chalvidal, Mathieu, Matthew Ricci, Rufin VanRullen, and Thomas Serre. 2020. 'Go with the Flow: Adaptive Control for Neural ODEs'. *ArXiv Preprint ArXiv:2006.09545*.

Champion, Kathleen, Bethany Lusch, J Nathan Kutz, and Steven L Brunton. 2019. 'Data-Driven Discovery of Coordinates and Governing Equations'. *Proceedings of the National Academy of Sciences* 116 (45): 22445–51.

Chang, Peng, Shirao Zhang, and Zichen Wang. 2023. 'Soft Sensor of the Key Effluent Index in the Municipal Wastewater Treatment Process Based on Transformer'. *IEEE Transactions on Industrial Informatics*.

Chaturvedi, Devendra K. 2017. *Modeling and Simulation of Systems Using MATLAB and Simulink*. CRC press.

Chen, Hongbo, Long Zeng, Dongbo Wang, Yaoyu Zhou, and Xiao Yang. 2020. 'Recent Advances in Nitrous Oxide Production and Mitigation in Wastewater Treatment'. *Water Research* 184:116168.

Chen, Kehua, Hongcheng Wang, Borja Valverde-Pérez, Siyuan Zhai, Luca Vezzaro, and Aijie Wang. 2021. 'Optimal Control towards Sustainable Wastewater Treatment Plants Based on Multi-Agent Reinforcement Learning'. *Chemosphere* 279:130498.

Chen, Ricky, Yulia Rubanova, Jesse Bettencourt, and David Duvenaud. 2018. 'Neural Ordinary Differential Equations'.

Chen, Ricky T. Q. 2018. 'Torchdiffeq'. Github. https://github.com/rtqichen/torchdiffeq.

Chen, Xueming, Artur Tomasz Mielczarek, Kirsten Habicht, Mikkel Holmen Andersen, Dines Thornberg, and Gurkan Sin. 2019. 'Assessment of Full-Scale N2O Emission Characteristics and Testing of Control Concepts in an Activated Sludge Wastewater Treatment Plant with Alternating Aerobic and Anoxic Phases'. *Environmental Science & Technology* 53 (21): 12485–94.

Chen-di, Y U, H O U Li-jun, ZHENG Yan-ling, L I U Min, Y I N Guo-yu, G A O Juan, L I U Cheng, and CHANG Yong-kai. 2019. 'Community Structure and Gene Expression Analysis for Nitrifier Enrichment Cultures'. *Journal of East China Normal University (Natural Science)* 2019 (3): 164.

Cheng, Hongchao, Yiqi Liu, Daoping Huang, and Bin Liu. 2019. ‘Optimized Forecast Components-SVM-Based Fault Diagnosis with Applications for Wastewater Treatment’. *IEEE Access* 7:128534–43.

Cheng, Xuhong, Zhiwei Guo, Yu Shen, Keping Yu, and Xu Gao. 2023. ‘Knowledge and Data-Driven Hybrid System for Modeling Fuzzy Wastewater Treatment Process’. *Neural Computing and Applications*, 1–22.

Chiranjivi, M, K Suresh, A Anand Kumar, M Siddartha, M Ravi Teja, and M Rani. 2024. ‘Forecasting of Effluent from Wastewater Treatment Industries Using a Novel Fuzzy Logic System’. In *E3S Web of Conferences*, 547:2008.

Choi, Jeongwhan, Hwangyong Choi, Jeehyun Hwang, and Noseong Park. 2022. ‘Graph Neural Controlled Differential Equations for Traffic Forecasting’. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 36:6367–74.

Choi, Matthew, Daniel Flam-Shepherd, Thi Ha Kyaw, and Alán Aspuru-Guzik. 2022. ‘Learning Quantum Dynamics with Latent Neural Ordinary Differential Equations’. *Physical Review A* 105 (4): 42403.

Chu, Xu, Ihab F Ilyas, Sanjay Krishnan, and Jiannan Wang. 2016. ‘Data Cleaning: Overview and Emerging Challenges’. In *Proceedings of the 2016 International Conference on Management of Data*, 2201–6.

Corominas, Ll, Manel Garrido-Baserba, Kris Villez, Gustaf Olsson, Ulises Cortés, and Manel Poch. 2018. ‘Transforming Data into Knowledge for Improved Wastewater Treatment Operation: A Critical Review of Techniques’. *Environmental Modelling & Software* 106:89–103.

Corominas, Lluís, Xavier Flores-Alsina, Laura Snip, and Peter A Vanrolleghem. 2012. ‘Comparison of Different Modeling Approaches to Better Evaluate Greenhouse Gas Emissions From Whole Wastewater Treatment Plants’. *Biotechnol. Bioeng* 109:2854–63. https://doi.org/10.1002/bit.24544/abstract.

Corominas, Llu\’\is, Kris Villez, Daniel Aguado, Leiv Rieger, Christian Rosén, and Peter A Vanrolleghem. 2011. ‘Performance Evaluation of Fault Detection Methods for Wastewater Treatment Processes’. *Biotechnology and Bioengineering* 108 (2): 333–44.

Cosenza, Alida, Giorgio Mannina, Peter A Vanrolleghem, and Marc B Neumann. 2014. 'Variance-Based Sensitivity Analysis for Wastewater Treatment Plant Modelling'. *Science of the Total Environment* 470:1068–77.

Cui, Peng, and Jinjia Wang. 2022. 'Out-of-Distribution (OOD) Detection Based on Deep Learning: A Review'. *Electronics* 11 (21): 3500.

Cunha, M C, and A Antunes. 2012. 'Simulated Annealing Algorithms for Water Systems Optimization'. *WIT Transactions on State-of-the-Art in Science and Engineering* 56.

Cuomo, Salvatore, Vincenzo Schiano Di Cola, Fabio Giampaolo, Gianluigi Rozza, Maziar Raissi, and Francesco Piccialli. 2022. 'Scientific Machine Learning Through Physics–Informed Neural Networks: Where We Are and What's Next'. *Journal of Scientific Computing* 92 (3). https://doi.org/10.1007/s10915-022-01939-z.

Daelman, M. R.J., E. M. van Voorthuizen, L. G.J.M. van Dongen, E. I.P. Volcke, and M. C.M. van Loosdrecht. 2013. 'Methane and Nitrous Oxide Emissions from Municipal Wastewater Treatment - Results from a Long-Term Study'. *Water Science and Technology* 67 (10): 2350–55. https://doi.org/10.2166/wst.2013.109.

Daneshgar, S., Y. Amerlinck, A. Amaral, C. De Mulder, A. Di Nisio, G. Bellandi, R. Gori, et al. 2024. 'An Innovative Model-Based Protocol for Minimisation of Greenhouse Gas (GHG) Emissions in WRRFs'. *Chemical Engineering Journal* 483 (March). https://doi.org/10.1016/j.cej.2023.148327.

David, Robert, J-L Vasel, and A Vande Wouwer. 2009. 'Settler Dynamic Modeling and MATLAB Simulation of the Activated Sludge Process'. *Chemical Engineering Journal* 146 (2): 174–83.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. 'BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding'. In *North American Chapter of the Association for Computational Linguistics*. https://api.semanticscholar.org/CorpusID:52967399.

Dikeman, Henry E, Hongyuan Zhang, and Suo Yang. 2022. 'Stiffness-Reduced Neural ODE Models for Data-Driven Reduced-Order Modeling of Combustion Chemical Kinetics'. In *AIAA SCITECH 2022 Forum*, 226.

Ding, Xiaoqian, Jianqiang Zhao, Kun Gao, Bo Hu, Xiaoling Li, Guanghuan Ge, Yangyang Yu, and Jinna Wu. 2018. 'Modeling of Nitrous Oxide Production by Ammonium-Oxidizing Bacteria'. *Environmental Engineering Science* 35 (1): 1–10.

Ding, Xiaoqian, Jianqiang Zhao, Bo Hu, Ying Chen, Guanghuan Ge, Xiaoling Li, Sha Wang, Kun Gao, and Xiaolei Tian. 2016. 'Mathematical Modeling of Nitrous Oxide Production in an Anaerobic/Oxic/Anoxic Process'. *Bioresource Technology* 222:39–48.

Domingo-Félez, C, and Barth F Smets. 2016. 'A Consilience Model to Describe N2O Production during Biological N Removal'. *Environmental Science: Water Research & Technology* 2 (6): 923–30.

Domingo-Félez, Carlos, Maria Calderó-Pascual, Gürkan Sin, Benedek G Plósz, and Barth F Smets. 2017. 'Calibration of the Comprehensive NDHA-N2O Dynamics Model for Nitrifier-Enriched Biomass Using Targeted Respirometric Assays'. *Water Research* 126:29–39.

Domingo-Félez, Carlos, and Barth F. Smets. 2019. 'Regulation of Key N2O Production Mechanisms during Biological Water Treatment'. *Current Opinion in Biotechnology* 57:119–26.

Domingo-Félez, Carlos, and Barth F Smets. 2020. 'Modeling Denitrification as an Electric Circuit Accurately Captures Electron Competition between Individual Reductive Steps: The Activated Sludge Model–Electron Competition Model'. *Environmental Science & Technology* 54 (12): 7330–38.

Dong, Chengyu, Liyuan Liu, Zichao Li, and Jingbo Shang. 2020. 'Towards Adaptive Residual Network Training: A Neural-Ode Perspective'. In *International Conference on Machine Learning*, 2616–26.

Duan, Haoran, Ben van den Akker, Benjamin J Thwaites, Lai Peng, Caroline Herman, Yuting Pan, Bing-Jie Ni, Shane Watt, Zhiguo Yuan, and Liu Ye. 2020. 'Mitigating

Nitrous Oxide Emissions at a Full-Scale Wastewater Treatment Plant'. *Water Research* 185:116196.

Duan, Haoran, Liu Ye, Dirk Erler, Bing-Jie Ni, and Zhiguo Yuan. 2017. 'Quantifying Nitrous Oxide Production Pathways in Wastewater Treatment Systems Using Isotope Technology–a Critical Review'. *Water Research* 122:96–113.

Duan, Haoran, Yingfen Zhao, Konrad Koch, George F Wells, Min Zheng, Zhiguo Yuan, and Liu Ye. 2021. 'Insights into Nitrous Oxide Mitigation Strategies in Wastewater Treatment and Challenges for Wider Implementation'. *Environmental Science & Technology* 55 (11): 7208–24.

Dupont, Emilien, Arnaud Doucet, and Yee Whye Teh. 2019. 'Augmented Neural ODEs', May. http://arxiv.org/abs/1904.01681.

Edzwald, James K, William C Becker, and Kevin L Wattier. 1985. 'Surrogate Parameters for Monitoring Organic Matter and THM Precursors'. *Journal-American Water Works Association* 77 (4): 122–32.

Elbrächter, Dennis, Dmytro Perekrestenko, Philipp Grohs, and Helmut Bölcskei. 2019. 'Deep Neural Network Approximation Theory', May. http://arxiv.org/abs/1901.02220.

Esteve-Yagüe, Carlos, and Borjan Geshkovski. 2021. 'Sparsity in Long-Time Control of Neural ODEs', May. http://arxiv.org/abs/2102.13566.

Faisal, M, Kashem M Muttaqi, Danny Sutanto, Ali Q Al-Shetwi, Pin Jern Ker, and M A Hannan. 2023. 'Control Technologies of Wastewater Treatment Plants: The State-of-the-Art, Current Challenges, and Future Directions'. *Renewable and Sustainable Energy Reviews* 181:113324.

Falcone, A B, A L Shug, and D J D Nicholas. 1963. 'Some Properties of a Hydroxylamine Oxidase from Nitrosomonas Europaea'. *Biochimica et Biophysica Acta* 77:199–208.

Fang, Tongtong, Nan Lu, Gang Niu, and Masashi Sugiyama. 2020. 'Rethinking Importance Weighting for Deep Learning under Distribution Shift'. *Advances in Neural Information Processing Systems* 33:11996–7.

Farhi, Nitzan, Efrat Kohen, Hadas Mamane, and Yuval Shavitt. 2021. 'Prediction of Wastewater Treatment Quality Using LSTM Neural Network'. *Environmental Technology & Innovation* 23:101632.

Fasel, Urban, J Nathan Kutz, Bingni W Brunton, and Steven L Brunton. 2022. 'Ensemble-SINDy: Robust Sparse Model Discovery in the Low-Data, High-Noise Limit, with Active Learning and Control'. *Proceedings of the Royal Society A* 478 (2260): 20210904.

Finlay, Chris, Jörn-Henrik Jacobsen, Levon Nurbekyan, and Adam M Oberman. 2020. 'How to Train Your Neural ODE: The World of Jacobian and Kinetic Regularization'.

Finn, Chelsea, Aravind Rajeswaran, Sham Kakade, and Sergey Levine. 2019. 'Online Meta-Learning'. In *International Conference on Machine Learning*, 1920–30.

Flores-Alsina, Xavier, Magnus Arnell, Youri Amerlinck, Lluís Corominas, Krist V. Gernaey, Lisha Guo, Erik Lindblom, et al. 2014. 'Balancing Effluent Quality, Economic Cost and Greenhouse Gas Emissions during the Evaluation of (Plant-Wide) Control/Operational Strategies in WWTPs'. *Science of the Total Environment* 466–467:616–24. https://doi.org/10.1016/j.scitotenv.2013.07.046.

Flores-Alsina, Xavier, Joaquim Comas, Ignasi Rodriguez-Roda, Krist V Gernaey, and Christian Rosen. 2009. 'Including the Effects of Filamentous Bulking Sludge during the Simulation of Wastewater Treatment Plants Using a Risk Assessment Model'. *Water Research* 43 (18): 4527–38.

Foley, Jeffrey, David de Haas, Zhiguo Yuan, and Paul Lant. 2010. 'Nitrous Oxide Generation in Full-Scale Biological Nutrient Removal Wastewater Treatment Plants'. *Water Research* 44 (3): 831–44. https://doi.org/10.1016/j.watres.2009.10.033.

Gao, Han, Yanping Mao, Xiaotian Zhao, Wen-Tso Liu, Tong Zhang, and George Wells. 2019. 'Genome-Centric Metagenomics Resolves Microbial Diversity and Prevalent Truncated Denitrification Pathways in a Denitrifying PAO-Enriched Bioprocess'. *Water Research* 155:275–87.

Garsdal, M L, V Søgaard, and S M Sørensen. 2022. ‘Generative Time Series Models Using Neural ODE in Variational Autoencoders’. *ArXiv Preprint ArXiv:2201.04630*.

Gear, C W. 1981. ‘Numerical Solution of Ordinary Differential Equations: Is There Anything Left to Do?’ *SIAM Review* 23 (1): 10–24. https://doi.org/10.1137/1023002.

Gholizadeh, Mahdi, Reza Saeedi, Amin Bagheri, and Mohammad Paeezi. 2024. ‘Machine Learning-Based Prediction of Effluent Total Suspended Solids in a Wastewater Treatment Plant Using Different Feature Selection Approaches: A Comparative Study’. *Environmental Research* 246:118146.

Giampiccolo, Stefano, Federico Reali, Anna Fochesato, Giovanni Iacca, and Luca Marchetti. 2024. ‘Robust Parameter Estimation and Identifiability Analysis with Hybrid Neural Ordinary Differential Equations in Computational Biology’. *BioRxiv*, 2024–26.

Girosi, Federico, Michael Jones, and Tomaso Poggio. 1995. ‘Regularization Theory and Neural Networks Architectures’. *Neural Computation* 7 (2): 219–69.

Glorot, Xavier, and Yoshua Bengio. 2010. ‘Understanding the Difficulty of Training Deep Feedforward Neural Networks’. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, 249–56.

Golovanev, Yakov, and Alexander Hvatov. 2022. ‘On the Balance between the Training Time and Interpretability of Neural ODE for Time Series Modelling’, May. http://arxiv.org/abs/2206.03304.

Govindarajan, L, S Krishna Kumar, and T Karunanithi. 2005. ‘Optimal Design of Wastewater Treatment Plant Using Adaptive Simulated Annealing’. *Journal of Applied Sciences and Environmental Management* 9 (1): 107–13.

Greydanus, Samuel, Misko Dzamba, and Jason Yosinski. 2019. ‘Hamiltonian Neural Networks’. *Advances in Neural Information Processing Systems* 32.

Gruber, Wenzel, Paul M. Magyar, Ivan Mitrovic, Kerstin Zeyer, Michael Vogel, Luzia von Känel, Lucien Biolley, et al. 2022. ‘Tracing N2O Formation in Full-Scale Wastewater Treatment with Natural Abundance Isotopes Indicates Control by

Organic Substrate and Process Settings'. *Water Research X* 15 (May). https://doi.org/10.1016/j.wroa.2022.100130.

Gruber, Wenzel, Robert Niederdorfer, Jörg Ringwald, Eberhard Morgenroth, Helmut Bürgmann, and Adriano Joss. 2021. 'Linking Seasonal N2O Emissions and Nitrification Failures to Microbial Dynamics in a SBR Wastewater Treatment Plant'. *Water Research X* 11:100098.

Guo, Jingbo, Qiwei Cong, Jun Zhang, Lanhe Zhang, Lingwei Meng, Mingwei Liu, and Fang Ma. 2021. 'Nitrous Oxide Emission in a Laboratory Anoxic-Oxic Process at Different Influent PHs: Generation Pathways and the Composition and Function of Bacterial Community'. *Bioresource Technology* 328:124844.

Guo, Lisha, and Peter A. Vanrolleghem. 2014. 'Calibration and Validation of an Activated Sludge Model for Greenhouse Gases No. 1 (ASMG1): Prediction of Temperature-Dependent N2O Emission Dynamics'. *Bioprocess and Biosystems Engineering* 37 (2): 151–63. https://doi.org/10.1007/s00449-013-0978-3.

Guo, Zhiwei, Boxin Du, Jianhui Wang, Yu Shen, Qiao Li, Dong Feng, Xu Gao, and Heng Wang. 2020. 'Data-Driven Prediction and Control of Wastewater Treatment Process through the Combination of Convolutional Neural Network and Recurrent Neural Network'. *RSC Advances* 10 (23): 13410–19. https://doi.org/10.1039/d0ra00736f.

Gusmão, Gabriel S, Adhika P Retnanto, Shashwati C.da Cunha, and Andrew J Medford. 2022. 'Kinetics-Informed Neural Networks'. *Catalysis Today*. https://doi.org/10.1016/j.cattod.2022.04.002.

Hafsa, Noor, Mohammed Al-Yaari, and Sayeed Rushd. 2021. 'Prediction of Arsenic Removal in Aqueous Solutions with Non-Neural Network Algorithms'. *The Canadian Journal of Chemical Engineering* 99:S135–S146.

Haimi, Henri, Michela Mulas, Francesco Corona, Stefano Marsili-Libelli, Paula Lindell, Mari Heinonen, and Riku Vahala. 2016. 'Adaptive Data-Derived Anomaly Detection in the Activated Sludge Process of a Large-Scale Wastewater Treatment Plant'. *Engineering Applications of Artificial Intelligence* 52:65–80.

Haimi, Henri, Michela Mulas, Francesco Corona, and Riku Vahala. 2013. 'Data-Derived Soft-Sensors for Biological Wastewater Treatment Plants: An Overview'. *Environmental Modelling and Software* 47:88–107. https://doi.org/10.1016/j.envsoft.2013.05.009.

Hairer, Ernst, and Gerhard Wanner. 1996. 'Solving Ordinary Differential Equations II - Stiff and Differential Algebraic Problems'. Springer.

Hallin, Sara, Laurent Philippot, Frank E Löffler, Robert A Sanford, and Christopher M Jones. 2018. 'Genomics and Ecology of Novel N2O-Reducing Microorganisms'. *Trends in Microbiology* 26 (1): 43–55.

Han, Honggui, Meiting Sun, Fangyu Li, Zezhong Liu, and Chen Wang. 2023. 'Self-Supervised Deep Clustering Method for Detecting Abnormal Data of Wastewater Treatment Process'. *IEEE Transactions on Industrial Informatics* 20 (2): 1155–66.

Hansen, Laura Debel, Peter Alexander Stentoft, Daniel Ortiz-Arroyo, and Petar Durdevic. 2024. 'Exploring Data Quality and Seasonal Variations of N2O in Wastewater Treatment: A Modeling Perspective'. *Water Practice and Technology* 19 (3): 1016–31. https://doi.org/10.2166/wpt.2024.045.

Hao, X D, Z L Yang, W B Yu, and R B Liu. 2023. 'N2O Emission from the Processes of Wastewater Treatment: Mechanisms and Control Strategies'. *Huan Jing Ke Xue= Huanjing Kexue* 44 (2): 1163–73.

Haque, Mirazul, Simin Chen, Wasif Haque, Cong Liu, and Wei Yang. 2023. 'AntiNODE: Evaluating Efficiency Robustness of Neural ODEs'. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 1507–17.

Harper Jr, Willie F, Yuki Takeuchi, Shohei Riya, Masaaki Hosomi, and Akihiko Terada. 2015. 'Novel Abiotic Reactions Increase Nitrous Oxide Production during Partial Nitrification: Modeling and Experiments'. *Chemical Engineering Journal* 281:1017–23.

Harry, Nikolaus, and Reginald Howe. 2021. 'Learning Neural Ordinary Differential Equations for Optimal Control'.

Hauduc, H, L Rieger, A Oehmen, M C M van Loosdrecht, Y Comeau, A Héduit, P A Vanrolleghem, and S Gillot. 2013. 'Critical Review of Activated Sludge Modeling: State of Process Knowledge, Modeling Concepts, and Limitations'. *Biotechnology and Bioengineering* 110 (1): 24–46. https://doi.org/10.1002/bit.24624.

He, K, X Zhang, S Ren, and J Sun. 2015. 'Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification'. In *2015 IEEE International Conference on Computer Vision (ICCV)*, 1026–34. https://doi.org/10.1109/ICCV.2015.123.

He, Ning, Mengrui Zhang, and Ruoxia Li. 2021. 'An Improved Approach for Robust MPC Tuning Based on Machine Learning'. *Mathematical Problems in Engineering* 2021 (1): 5518950.

He, Yanying, Yingrui Liu, Xuecheng Li, Tingting Zhu, and Yiwen Liu. 2023. 'Unveiling the Roles of Biofilm in Reducing N2O Emission in a Nitrifying Integrated Fixed-Film Activated Sludge (IFAS) System'. *Water Research* 243:120326.

Hendrycks, Dan, Steven Basart, Norman Mu, Saurav Kadavath, Frank Wang, Evan Dorundo, Rahul Desai, et al. 2021. 'The Many Faces of Robustness: A Critical Analysis of out-of-Distribution Generalization'. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 8340–49.

Hendrycks, Dan, and Kevin Gimpel. 2016. 'Gaussian Error Linear Units (GELUs)', May. http://arxiv.org/abs/1606.08415.

Heo, SungKu, KiJeon Nam, Shahzeb Tariq, Juin Yau Lim, Junkyu Park, and ChangKyoo Yoo. 2021. 'A Hybrid Machine Learning–Based Multi-Objective Supervisory Control Strategy of a Full-Scale Wastewater Treatment for Cost-Effective and Sustainable Operation under Varying Influent Conditions'. *Journal of Cleaner Production* 291:125853.

Hernández-del-Olmo, Félix, Elena Gaudioso, Raquel Dormido, and Natividad Duro. 2018. 'Tackling the Start-up of a Reinforcement Learning Agent for the Control of Wastewater Treatment Plants'. *Knowledge-Based Systems* 144:9–15.

Hernández-del-Olmo, Félix, Félix H Llanes, and Elena Gaudioso. 2012. 'An Emergent Approach for the Control of Wastewater Treatment Plants by Means of Reinforcement Learning Techniques'. *Expert Systems with Applications* 39 (3): 2355–60.

Hiatt, William C., and C. P. Leslie Grady. 2008. 'An Updated Process Model for Carbon Oxidation, Nitrification, and Denitrification'. *Water Environment Research* 80 (11): 2145–56. https://doi.org/10.2175/106143008x304776.

Hilal, Anwer Mustafa, Maha M Althobaiti, Taiseer Abdalla Elfadil Eisa, Rana Alabdan, Manar Ahmed Hamza, Abdelwahed Motwakel, Mesfer Al Duhayyim, and Noha Negm. 2022. 'An Intelligent Carbon-Based Prediction of Wastewater Treatment Plants Using Machine Learning Algorithms'. *Adsorption Science & Technology* 2022:8448489.

Hink, Linda, Graeme W Nicol, and James I Prosser. 2017. 'Archaea Produce Lower Yields of N2O than Bacteria during Aerobic Ammonia Oxidation in Soil'. *Environmental Microbiology* 19 (12): 4829–37.

Ho, Long, Ruben Jerves-Cobo, Marie Anne Eurie Forio, Ans Mouton, Ingmar Nopens, and Peter Goethals. 2021. 'Integrated Mechanistic and Data-Driven Modeling for Risk Assessment of Greenhouse Gas Production in an Urbanized River System'. *Journal of Environmental Management* 294:112999. https://doi.org/10.1016/j.jenvman.2021.112999.

Hochstein, Lawrence I, and Geraldine A Tomlinson. 1988. 'The Enzymes Associated with Denitrification.'

Holenda, B, Eu Domokos, A Rédey, and J Fazakas. 2007. 'Aeration Optimization of a Wastewater Treatment Plant Using Genetic Algorithm'. *Optimal Control Applications and Methods* 28 (3): 191–208.

Hong, Feng, Hailin Tian, Xi Yuan, Shuchang Liu, Qintian Peng, Yan Shi, Lei Jin, et al. 2022. 'CFD-Assisted Modeling of the Hydrodynamic Cavitation Reactors for Wastewater Treatment—A Review'. *Journal of Environmental Management* 321:115982.

Hongbo, Chen, Zeng Long, Wang Dongbo, Zhou Yaoyu, and Yang Xiao. 2020. 'Recent Advances in Nitrous Oxide Production and Mitigation in Wastewater Treatment'. *Water Research* 184:116168. https://doi.org/10.1016/j.watres.2020.116168.

Hospedales, Timothy, Antreas Antoniou, Paul Micaelli, and Amos Storkey. 2021. 'Meta-Learning in Neural Networks: A Survey'. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44 (9): 5149–69.

Hossain, Intekhab, Viola Fanfani, Jonas Fischer, John Quackenbush, and Rebekka Burkholz. 2024. 'Biologically Informed NeuralODEs for Genome-Wide Regulatory Dynamics'. *Genome Biology* 25 (1): 127.

Huang, Mingzhi, Yongwen Ma, Jinquan Wan, and Xiaohong Chen. 2015. 'A Sensor-Software Based on a Genetic Algorithm-Based Neural Fuzzy System for Modeling and Simulating a Wastewater Treatment Process'. *Applied Soft Computing* 27:1–10.

Huang, Songqing, and Xiuhong Liu. 2024. 'Nitrous Oxide Generation, Influencing Factors and Emission Reduction in Anaerobic-Anoxic-Oxic Process Wastewater Treatment Plant'. In *E3S Web of Conferences*. Vol. 520. EDP Sciences. https://doi.org/10.1051/e3sconf/202452001011.

Huang, Zi-lan, Luo-yi Zhang, Yi Zhang, Shu-wei Qian, and Chong-jun Wang. 2021. 'Transformer Based Multi-Output Regression Learning for Wastewater Treatment'. In *2021 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI)*, 698–703.

Humbert, Guillaume, Mathieu Sebilo, Justine Fiat, Longqi Lang, Ahlem Filali, Véronique Vaury, Mathieu Spérandio, and Anniet M Laverman. 2020. 'Isotopic Evidence for Alteration of Nitrous Oxide Emissions and Producing Pathways' Contribution under Nitrifying Conditions'. *Biogeosciences* 17 (4): 979–93.

Hurst, Christon J. 2019. *Understanding Terrestrial Microbial Communities*. Springer.

Hwangbo, Soonho, Resul Al, Xueming Chen, and Gürkan Sin. 2021. 'Integrated Model for Understanding N2O Emissions from Wastewater Treatment Plants: A Deep

Learning Approach'. *Environmental Science and Technology* 55 (3): 2143–51. https://doi.org/10.1021/acs.est.0c05231.

Hwangbo, Soonho, Resul Al, and Gürkan Sin. 2020. 'An Integrated Framework for Plant Data-Driven Process Modeling Using Deep-Learning with Monte-Carlo Simulations'. *Computers & Chemical Engineering* 143:107071.

Ilyas, Ihab F, and Theodoros Rekatsinas. 2022. 'Machine Learning and Data Cleaning: Which Serves the Other?' *ACM Journal of Data and Information Quality (JDIQ)* 14 (3): 1–11.

IPCC. 2014. 'IPCC Fifth Assessment Report, 2014 (AR5)'.

Iqbal, Jawed, and Chandan Guria. 2009. 'Optimization of an Operating Domestic Wastewater Treatment Plant Using Elitist Non-Dominated Sorting Genetic Algorithm'. *Chemical Engineering Research and Design* 87 (11): 1481–96.

Jaegher, Bram De, Eneko Larumbe, Wim De Schepper, Arne Verliefde, and Ingmar Nopens. 2020. 'Colloidal Fouling in Electrodialysis: A Neural Differential Equations Model'. *Separation and Purification Technology* 249:116939.

Ji, Weiqi, Weilun Qiu, Zhiyu Shi, Shaowu Pan, and Sili Deng. 2021. 'Stiff-PINN: Physics-Informed Neural Network for Stiff Chemical Kinetics'.

Jia, Junteng, and Austin R Benson. 2019. 'Neural Jump Stochastic Differential Equations'. *Advances in Neural Information Processing Systems* 32.

Jung, Man-Young, Joo-Han Gwak, Lena Rohe, Anette Giesemann, Jong-Geol Kim, Reinhard Well, Eugene L Madsen, Craig W Herbold, Michael Wagner, and Sung-Keun Rhee. 2019. 'Indications for Enzymatic Denitrification to N2O at Low PH in an Ammonia-Oxidizing Archaeon'. *The ISME Journal* 13 (10): 2633–38.

Kaheman, Kadierdan, J Nathan Kutz, and Steven L Brunton. 2020. 'SINDy-PI: A Robust Algorithm for Parallel Implicit Sparse Identification of Nonlinear Dynamics'. *Proceedings of the Royal Society A* 476 (2242): 20200279.

Kang, Hoon, Seunghyeok Yang, Jianying Huang, and Jeill Oh. 2020. 'Time Series Prediction of Wastewater Flow Rate by Bidirectional LSTM Deep Learning'. *International Journal of Control, Automation and Systems* 18:3023–30.

Karniadakis, George Em, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. 2021. 'Physics-Informed Machine Learning'. *Nature Reviews Physics* 3 (6): 422–40. https://doi.org/10.1038/s42254-021-00314-5.

Kemmou, Liana, and Elisavet Amanatidou. 2023. 'Factors Affecting Nitrous Oxide Emissions from Activated Sludge Wastewater Treatment Plants—A Review'. *Resources* 12 (10). https://doi.org/10.3390/resources12100114.

Khalil, Mostafa, Ahmed AlSayed, Ahmed Elsayed, Mohamed Sherif Zaghloul, Katherine Y Bell, Ahmed Al-Omari, Farokh Laqa Kakar, et al. 2024. 'Advances in GHG Emissions Modelling for WRRFs: From State-of-the-Art Methods to Full-Scale Applications'. *Chemical Engineering Journal* 494:153053. https://doi.org/10.1016/j.cej.2024.153053.

Khalil, Mostafa, Ahmed AlSayed, Yang Liu, and Peter A. Vanrolleghem. 2023. 'Machine Learning for Modeling N2O Emissions from Wastewater Treatment Plants: Aligning Model Performance, Complexity, and Interpretability'. *Water Research* 245 (October). https://doi.org/10.1016/j.watres.2023.120667.

Khalil, Mostafa, Ahmed AlSayed, Yang Liu, and Peter A Vanrolleghem. 2024. 'An Integrated Feature Selection and Hyperparameter Optimization Algorithm for Balanced Machine Learning Models Predicting N2O Emissions from Wastewater Treatment Plants'. *Journal of Water Process Engineering* 63:105512. https://doi.org/10.1016/j.jwpe.2024.105512.

Khoja, Intissar, Taoufik Ladhari, Anis Sakly, and Faouzi M'sahli. 2018. 'Parameter Identification of an Activated Sludge Wastewater Treatment Process Based on Particle Swarm Optimization Method'. *Mathematical Problems in Engineering* 2018 (1): 7823930.

Kidger, Patrick. 2022. 'On Neural Differential Equations', May. http://arxiv.org/abs/2202.02435.

Kidger, Patrick, James Morrill, James Foster, and Terry Lyons. 2020. 'Neural Controlled Differential Equations for Irregular Time Series', May. http://arxiv.org/abs/2005.08926.

Kim, Minsoo, Yejin Kim, Hyosoo Kim, Wenhua Piao, and Changwon Kim. 2016. 'Evaluation of the K-Nearest Neighbor Method for Forecasting the Influent Characteristics of Wastewater Treatment Plant'. *Frontiers of Environmental Science & Engineering* 10:299–310.

Kim, Sang-Wan, Morio Miyahara, Shinya Fushinobu, Takayoshi Wakagi, and Hirofumi Shoun. 2010. 'Nitrous Oxide Emission from Nitrifying Activated Sludge Dependent on Denitrification by Ammonia-Oxidizing Bacteria'. *Bioresource Technology* 101 (11): 3958–63.

Kim, Suyong, Weiqi Ji, Sili Deng, Yingbo Ma, and Christopher Rackauckas. 2021. 'Stiff Neural Ordinary Differential Equations', May. https://doi.org/10.1063/5.0060697.

Kimura, Toshinori. 2009. 'On Dormand-Prince Method'. *Jpn. Malaysia Tech. Instit* 40:1–9.

Kingma, Diederik P., and Jimmy Ba. 2014. 'Adam: A Method for Stochastic Optimization', December. http://arxiv.org/abs/1412.6980.

Kinh, Co Thi, Johwan Ahn, Toshikazu Suenaga, Nakanya Sittivorakulpong, Pongsak Noophan, Tomoyuki Hori, Shohei Riya, Masaaki Hosomi, and Akihiko Terada. 2017. 'Free Nitrous Acid and PH Determine the Predominant Ammonia-Oxidizing Bacteria and Amount of N 2 O in a Partial Nitrifying Reactor'. *Applied Microbiology and Biotechnology* 101:1673–83.

Kircher, Tim, Felix A Döppel, and Martin Votsmeier. 2024. 'Embedding Physics into Neural ODEs to Learn Kinetics from Integral Reactors'. In *Computer Aided Chemical Engineering*, 53:817–22. Elsevier.

Kits, K D, M Y Jung, J Vierheilig, P Pjevac, C J Sedlacek, S Liu, C Herbold, et al. 2019. 'Low Yield and Abiotic Origin of N2O Formed by the Complete Nitrifier Nitrospira Inopinata'. *Nat Commun* 10 (1836).

Kloberdanz, Eliska, and Wei Le. 2023. 'S-SOLVER: Numerically Stable Adaptive Step Size Solver for Neural ODEs'. In *International Conference on Artificial Neural Networks*, 388–400.

Kobyzev, Ivan, Simon J D Prince, and Marcus A Brubaker. 2020. 'Normalizing Flows: An Introduction and Review of Current Methods'. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43 (11): 3964–79.

Koloskova, Anastasia, Hadrien Hendrikx, and Sebastian U Stich. 2023. 'Revisiting Gradient Clipping: Stochastic Bias and Tight Convergence Guarantees'. In *International Conference on Machine Learning*, 17343–63.

Kong, Xianghao, Koji Yamashita, Brandon Foggo, and Nanpeng Yu. 2022. 'Dynamic Parameter Estimation with Physics-Based Neural Ordinary Differential Equations'.

Kotlar, E, B Tartakovsky, Y Argaman, and M Sheintuch. 1996. 'The Nature of Interaction between Immobilized Nitrification and Denitrification Bacteria'. *Journal of Biotechnology* 51 (3): 251–58.

Kozlowski, Jessica A, K Dimitri Kits, and Lisa Y Stein. 2016. 'Comparison of Nitrogen Oxide Metabolism among Diverse Ammonia-Oxidizing Bacteria'. *Frontiers in Microbiology* 7:1090.

Kozlowski, Jessica A, Jennifer Price, and Lisa Y Stein. 2014. 'Revision of N2O-Producing Pathways in the Ammonia-Oxidizing Bacterium Nitrosomonas Europaea ATCC 19718'. *Applied and Environmental Microbiology* 80 (16): 4930–35.

Kurt, Hornik, Tinchcombe Maxwell, and White Halbert. 1989. 'Multilayer Feedforward Networks Are Universal Approximators-1989'.

Kushnir, Dan, and Vladimir Rokhlin. 2012. 'A Highly Accurate Solver for Stiff Ordinary Differential Equations'. *SIAM Journal on Scientific Computing* 34 (3): A1296–1315. https://doi.org/10.1137/100810216.

Laili, N, N S Indrasti, and D Wahyudi. 2022. 'Design of Sustainable Coffee Processing Wastewater Treatment System Using K-Means Clustering Algorithm'. In *IOP Conference Series: Earth and Environmental Science*, 1063:12032.

Lancioni, Nicola, Bartosz Szelag, Massimiliano Sgroi, Krzysztof Barbusiński, Francesco Fatone, and Anna Laura Eusebi. 2024. 'Novel Extended Hybrid Tool for Real Time Control and Practically Support Decisions to Reduce GHG

Emissions in Full Scale Wastewater Treatment Plants'. *Journal of Environmental Management* 365 (August). https://doi.org/10.1016/j.jenvman.2024.121502.

Law, Yingyu, Paul Lant, and Zhiguo Yuan. 2011. 'The Effect of PH on N2O Production under Aerobic Conditions in a Partial Nitritation System'. *Water Research* 45 (18): 5934–44.

Law, Yingyu, Bing-Jie Ni, Paul Lant, and Zhiguo Yuan. 2012. 'N2O Production Rate of an Enriched Ammonia-Oxidising Bacteria Culture Exponentially Correlates to Its Ammonia Oxidation Rate'. *Water Research* 46 (10): 3409–19.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. 'Deep Learning'. *Nature* 521 (7553): 436–44. https://doi.org/10.1038/nature14539.

Lee, Kookjin, and Eric J Parish. 2021. 'Parameterized Neural Ordinary Differential Equations: Applications to Computational Physics Problems'. *Proceedings of the Royal Society A* 477 (2253): 20210162.

Lee, Yu-Jen, Bin-le Lin, and Zhongfang Lei. 2022. 'Nitrous Oxide Emission Mitigation from Biological Wastewater Treatment – A Review'. *Bioresource Technology* 362:127747. https://doi.org/10.1016/j.biortech.2022.127747.

Leimkuhler, Benedict J, and Robert D Skeel. 1994. 'Symplectic Numerical Integrators in Constrained Hamiltonian Systems'. *Journal of Computational Physics* 112 (1): 117–25.

Li, Can, Shufeng Liu, Tao Ma, Maosheng Zheng, and Jinren Ni. 2019. 'Simultaneous Nitrification, Denitrification and Phosphorus Removal in a Sequencing Batch Reactor (SBR) under Low Temperature'. *Chemosphere* 229:132–41.

Li, Cong, Qian Wang, and Wenlin Jia. 2022. 'N2O Reduction during Denitrifying Phosphorus Removal with Propionate as Carbon Source'. *Environmental Science and Pollution Research* 29 (9): 12390–98. https://doi.org/10.1007/s11356-021-14629-4.

Li, Jianmin, Xiuhong Liu, Pengchao Gu, Bin Cui, Qing Yang, and Wei Zeng. 2022. 'N2O Production and Emission Pathways in Anammox Biofilter for Treating Wastewater with Low Nitrogen Concentrations'. *Science of The Total Environment* 852:158282.

Li, Kaili, Haoran Duan, Linfeng Liu, Ruihong Qiu, Ben van den Akker, Bing Jie Ni, Tong Chen, Hongzhi Yin, Zhiguo Yuan, and Liu Ye. 2022. 'An Integrated First Principal and Deep Learning Approach for Modeling Nitrous Oxide Emissions from Wastewater Treatment Plants'. *Environmental Science and Technology* 56 (4): 2816–26. https://doi.org/10.1021/acs.est.1c05020.

Li, Pengzhang, Yun Wang, Yue Liu, Shuying Wang, and Yongzhen Peng. 2023. 'The Effect of Salinity on N2O Emissions during Domestic Wastewater Partial Nitrification Treatment in a Sequencing Batch Reactor'. *Water* 15 (19). https://doi.org/10.3390/w15193502.

Li, Wenjing, and Xiaoxiao Wang. 2021. 'Time Series Prediction Method Based on Simplified LSTM Neural Network'. *Journal of Beijing University of Technology* 47 (5): 480–88.

Li, Xiaoyong, Xiaohui Yi, Zhenghui Liu, Hongbin Liu, Tao Chen, Guoqiang Niu, Bo Yan, Chen Chen, Mingzhi Huang, and Guangguo Ying. 2021. 'Application of Novel Hybrid Deep Leaning Model for Cleaner Production in a Paper Industrial Wastewater Treatment System'. *Journal of Cleaner Production* 294:126343.

Li, Yue, Bin Kong, Weiwei Yu, and Xingliang Zhu. 2023. 'An Attention-Based CNN-LSTM Method for Effluent Wastewater Quality Prediction'. *Applied Sciences* 13 (12): 7011.

Liang, Hua, and Hulin Wu. 2008. 'Parameter Estimation for Differential Equation Models Using a Framework of Measurement Error in Regression Models'. *Journal of the American Statistical Association* 103 (484): 1570–83. https://doi.org/10.1198/016214508000000797.

Liang, Jiheng, Xiangyang Yu, Weibin Hong, and Yefan Cai. 2024. 'Information Extraction of UV-NIR Spectral Data in Waste Water Based on Large Language Model'. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 318:124475.

Liang, Weihao, Chao Yu, Hongqiang Ren, Jinju Geng, Lili Ding, and Ke Xu. 2015. 'Minimization of Nitrous Oxide Emission from CASS Process Treating Low

Carbon Source Domestic Wastewater: Effect of Feeding Strategy and Aeration Rate'. *Bioresource Technology* 198:172–80.

Lim, Bryan, and Stefan Zohren. 2021. 'Time-Series Forecasting with Deep Learning: A Survey'. *Philosophical Transactions of the Royal Society A* 379 (2194): 20200209.

Lin, Wei, Yu Hanyue, and Li Bin. 2022. 'Prediction of Wastewater Treatment System Based on Deep Learning'. *Frontiers in Ecology and Evolution* 10:1064555.

Linda, Petzoid. 1982. 'Differential/Algebraic Equations Are Not ODE's'. *SIAM Journal on Scientific and Statistical Computing*.

Liu, Bing. 2017. 'Lifelong Machine Learning: A Paradigm for Continuous Learning'. *Frontiers of Computer Science* 11 (3): 359–61.

Liu, Fangqin, Ning Ding, Guanghua Zheng, and Jiangrong Xu. 2024. 'Carbon Emission Prediction and Emission Reduction Analysis of Wastewater Treatment Plant Based on Machine Learning and Deep Learning'. *Research Square*. https://doi.org/10.21203/rs.3.rs-4561438/v1.

Liu, Tang, Shufeng Liu, Shishi He, Zhichao Tian, and Maosheng Zheng. 2021. 'Minimization of N2O Emission through Intermittent Aeration in a Sequencing Batch Reactor (SBR): Main Behavior and Mechanism'. *Water* 13 (2): 210.

Liu, Yingrui, Tingting Zhu, Shuqi Ren, Tianhang Zhao, Hongxiang Chai, Yifeng Xu, Lai Peng, and Yiwen Liu. 2022. 'Contribution of Nitrification and Denitrification to Nitrous Oxide Turnovers in Membrane-Aerated Biofilm Reactors (MABR): A Model-Based Evaluation'. *Science of The Total Environment* 806:151321.

Liu, Yiwen, Lai Peng, Xueming Chen, and Bing-Jie Ni. 2015. 'Mathematical Modeling of Nitrous Oxide Production during Denitrifying Phosphorus Removal Process'. *Environmental Science & Technology* 49 (14): 8595–8601.

Liu, Yiwen, Tianhang Zhao, Zhongxian Su, Tingting Zhu, and Bing-Jie Ni. 2020. 'Evaluating the Roles of Coexistence of Sludge Flocs on Nitrogen Removal and Nitrous Oxide Production in a Granule-Based Autotrophic Nitrogen Removal System'. *Science of the Total Environment* 730:139018.

Liu, Yue, Zhengwei Yang, Xinxin Zou, Shuchang Ma, Dahui Liu, Maxim Avdeev, and Siqi Shi. 2023. 'Data Quantity Governance for Machine Learning in Materials Science'. *National Science Review* 10 (7): nwad125.

Liu, Yutian, Shoubin Zhang, Liping Qiu, and Yuanyuan Zhang. 2016. *Progress and Prospect of Research on N2O Dynamics Model in Wastewater Biological Treatment Process*. https://doi.org/10.2991/icseee-15.2016.74.

Liu, Yuting, Wenchong Tian, Jun Xie, Weizhong Huang, and Kunlun Xin. 2023. 'LSTM-Based Model-Predictive Control with Rationality Verification for Bioreactors in Wastewater Treatment'. *Water* 15 (9): 1779.

Logan, Justin, Nicole Roberts, and Michael Smith. 2024. 'Improved Operation Forecasting of Air Blowers in a Wastewater Treatment Plant via Machine Learning-Based Models'. In *SoutheastCon 2024*, 165–70.

Long, Sha, Lin Zhao, Hongbo Liu, Jingchen Li, Xia Zhou, Yunfeng Liu, Zhi Qiao, Yingxin Zhao, and Yongkui Yang. 2019. 'A Monte Carlo-Based Integrated Model to Optimize the Cost and Pollution Reduction in Wastewater Treatment Processes in a Typical Comprehensive Industrial Park in China'. *Science of the Total Environment* 647:1–10.

Lotfi, Khadije, Hossein Bonakdari, Isa Ebtehaj, Farouq S Mjalli, Mohammad Zeynoddin, Robert Delatolla, and Bahram Gharabaghi. 2019. 'Predicting Wastewater Treatment Plant Quality Parameters Using a Novel Hybrid Linear-Nonlinear Methodology'. *Journal of Environmental Management* 240:463–74.

Lu, Hao, Huazhe Wang, Qinglian Wu, Haichao Luo, Qi Zhao, Banghai Liu, Qishi Si, Shanshan Zheng, Wanqian Guo, and Nanqi Ren. 2023. 'Automatic Control and Optimal Operation for Greenhouse Gas Mitigation in Sustainable Wastewater Treatment Plants: A Review'. *Science of the Total Environment*. Elsevier B.V. https://doi.org/10.1016/j.scitotenv.2022.158849.

Lu, Lu, Hui Zheng, Jing Jie, Miao Zhang, and Rui Dai. 2021. 'Reinforcement Learning-Based Particle Swarm Optimization for Sewage Treatment Control'. *Complex & Intelligent Systems* 7 (5): 2199–2210.

Luca, Laurentiu, Ramon Vilanova, George Adrian Ifrim, Emil Ceanga, Sergiu Caraman, and Marian Barbu. 2019. 'Control Strategies of a Wastewater Treatment Plant'. *IFAC-PapersOnLine* 52 (1): 257–62.

Mahjouri, Maryam, Mohd Bakri Ishak, Ali Torabian, Latifah Abd Manaf, and Normala Halimoon. 2017. 'The Application of a Hybrid Model for Identifying and Ranking Indicators for Assessing the Sustainability of Wastewater Treatment Systems'. *Sustainable Production and Consumption* 10:21–37.

Mahsereci, Maren, Lukas Balles, Christoph Lassner, and Philipp Hennig. 2017. 'Early Stopping without a Validation Set'. *ArXiv Preprint ArXiv:1703.09580*.

Maktabifard, Mojtaba, Kati Blomberg, Ewa Zaborowska, Anna Mikola, and Jacek Makinia. 2022. 'Model-Based Identification of the Dominant N2O Emission Pathway in a Full-Scale Activated Sludge System'. *Journal of Cleaner Production* 336:130347.

Mamandipoor, Behrooz, Mahshid Majd, Seyedmostafa Sheikhalishahi, Claudio Modena, and Venet Osmani. 2020. 'Monitoring and Detecting Faults in Wastewater Treatment Plants Using Deep Learning'. *Environmental Monitoring and Assessment* 192 (2): 148.

Mampaey, K. E., B. Beuckels, M. J. Kampschreur, R. Kleerebezem, M. C.M. Van Loosdrecht, and E. I.P. Volcke. 2013. 'Modelling Nitrous and Nitric Oxide Emissions by Autotrophic Ammonia-Oxidizing Bacteria'. *Environmental Technology (United Kingdom)* 34 (12): 1555–66. https://doi.org/10.1080/09593330.2012.758666.

Mampaey, Kris E, Mathieu Spérandio, Mark C M van Loosdrecht, and Eveline I P Volcke. 2019. 'Dynamic Simulation of N2O Emissions from a Full-Scale Partial Nitritation Reactor'. *Biochemical Engineering Journal* 152:107356.

Mannina, Giorgio, Marco Capodici, Alida Cosenza, and Daniele Di Trapani. 2018. 'Nitrous Oxide from Integrated Fixed-Film Activated Sludge Membrane Bioreactor: Assessing the Influence of Operational Variables'. *Bioresource Technology* 247:1221–27.

Massara, Theoni Maria, Borja Solís, Albert Guisasola, Evina Katsou, and Juan Antonio Baeza. 2018. 'Development of an ASM2d-N2O Model to Describe Nitrous Oxide Emissions in Municipal WWTPs under Dynamic Conditions'. *Chemical Engineering Journal* 335 (May):185–96. https://doi.org/10.1016/j.cej.2017.10.119.

MATLAB. 2024. 'Documentation: Dlode45 Function'. 2024. https://uk.mathworks.com/help/deeplearning/ref/dlarray.dlode45.html.

McDonald, Andrew. 2021. 'Data Quality Considerations for Petrophysical Machine-Learning Models'. *Petrophysics* 62 (06): 585–613.

Mehrani, Mohamad-Javad, Faramarz Bagherzadeh, Min Zheng, Przemyslaw Kowal, Dominika Sobotka, and Jacek Makinia. 2022. 'Application of a Hybrid Mechanistic/Machine Learning Model for Prediction of Nitrous Oxide (N2O) Production in a Nitrifying Sequencing Batch Reactor'. *Process Safety and Environmental Protection* 162:1015–24.

Migdał, Karolina, Agnieszka Operacz, Iryna Vaskina, Paulina Śliz, Jorge Tavares, Adelaide Almeida, and Michał Migdał. 2022. 'Assessment of the Reliability of the Operation of a Sewage Treatment Plant Using Monte Carlo Simulation'. *Journal of Water and Land Development*.

Mogens, Henze, Gujer Willi, Mino Takashi, and van Loosdrecht Mark. 2000. *Activated Sludge Models ASM1,ASM2, ASM2d, and ASM3*. IWA Publishing.

Moles, Carmen G, G Gutierrez, Antonio A Alonso, and Julio R Banga. 2003. 'Integrated Process Design and Control via Global Optimization: A Wastewater Treatment Plant Case Study'. *Chemical Engineering Research and Design* 81 (5): 507–17.

Moravec, Marek, Miroslav Badida, Nikoleta Mikušová, L\`ydia Sobotová, Jozef Švajlenka, and Tibor Dzuro. 2021. 'Proposed Options for Noise Reduction from a Wastewater Treatment Plant: Case Study'. *Sustainability* 13 (4): 2409.

Moullec, Y Le, C Gentric, O Potier, and J P Leclerc. 2010. 'Comparison of Systemic, Compartmental and CFD Modelling Approaches: Application to the Simulation of

a Biological Reactor of Wastewater Treatment'. *Chemical Engineering Science* 65 (1): 343–50.

Murua, Ander. 1999. 'Formal Series and Numerical Integrators, Part I: Systems of ODEs and Symplectic Integrators'. *Applied Numerical Mathematics* 29 (2): 221–51.

Nadiri, Ata Allah, Sima Shokri, Frank T-C Tsai, and Asghar Asghari Moghaddam. 2018. 'Prediction of Effluent Quality Parameters of a Wastewater Treatment Plant Using a Supervised Committee Fuzzy Logic Model'. *Journal of Cleaner Production* 180:539–49.

Newhart, Kathryn B, Ryan W Holloway, Amanda S Hering, and Tzahi Y Cath. 2019. 'Data-Driven Performance Analyses of Wastewater Treatment Plants: A Review'. *Water Research* 157:498–513. https://doi.org/10.1016/j.watres.2019.03.030.

Ng, Kim Hoong, Y S Gan, Chin Kui Cheng, Kun-Hong Liu, and Sze-Teng Liong. 2020. 'Integration of Machine Learning-Based Prediction for Enhanced Model's Generalization: Application in Photocatalytic Polishing of Palm Oil Mill Effluent (POME)'. *Environmental Pollution* 267:115500.

Ni, Bing-Jie, Yuting Pan, Ben van den Akker, Liu Ye, and Zhiguo Yuan. 2015. 'Full-Scale Modeling Explaining Large Spatial Variations of Nitrous Oxide Fluxes in a Step-Feed Plug-Flow Wastewater Treatment Reactor'. *Environmental Science & Technology* 49 (15): 9176–84. https://doi.org/10.1021/acs.est.5b02038.

Ni, Bing-Jie, Lai Peng, Yingyu Law, Jianhua Guo, and Zhiguo Yuan. 2014. 'Modeling of Nitrous Oxide Production by Autotrophic Ammonia-Oxidizing Bacteria with Multiple Production Pathways'. *Environmental Science & Technology* 48 (7): 3916–24.

Ni, Bing-Jie, Maël Ruscalleda, Carles Pellicer-Nacher, and Barth F Smets. 2011. 'Modeling Nitrous Oxide Production during Biological Nitrogen Removal via Nitrification and Denitrification: Extensions to the General ASM Models'. *Environmental Science & Technology* 45 (18): 7768–76.

Ni, Bing-Jie, Liu Ye, Yingyu Law, Craig Byers, and Zhiguo Yuan. 2013. 'Mathematical Modeling of Nitrous Oxide (N2O) Emissions from Full-Scale Wastewater Treatment Plants'. *Environmental Science & Technology* 47 (14): 7795–7803.

Ni, Bing-Jie, and Zhiguo Yuan. 2015. 'Recent Advances in Mathematical Modeling of Nitrous Oxides Emissions from Wastewater Treatment Processes'. *Water Research* 87:336–46.

Ni, Bing-Jie, Zhiguo Yuan, Kartik Chandran, Peter A. Vanrolleghem, and Sudhir Murthy. 2013. 'Evaluating Four Mathematical Models for Nitrous Oxide Production by Autotrophic Ammonia-Oxidizing Bacteria'. *Biotechnology and Bioengineering* 110 (1): 153–63.

Núñez, Matías, Nadia L. Barreiro, Rafael A. Barrio, and Christopher Rackauckas. 2023. 'Forecasting Virus Outbreaks with Social Media Data via Neural Ordinary Differential Equations'. *Scientific Reports* 13 (1). https://doi.org/10.1038/s41598-023-37118-9.

Pagga, Udo, Jürgen Bachner, and Uwe Strotmann. 2006. 'Inhibition of Nitrification in Laboratory Tests and Model Wastewater Treatment Plants'. *Chemosphere* 65 (1): 1–8.

Pal, Avik, Alan Edelman, and Christopher Rackauckas. 2022. 'Mixing Implicit and Explicit Deep Learning with Skip Deqs and Infinite Time Neural Odes (Continuous Deqs)'. *Training* 4:5.

Palomo, Alejandro, Anders G Pedersen, S Jane Fowler, Arnaud Dechesne, Thomas Sicheritz-Pontén, and Barth F Smets. 2018. 'Comparative Genomics Sheds Light on Niche Differentiation and the Evolutionary History of Comammox Nitrospira'. *The ISME Journal* 12 (7): 1779–93.

Pan, Yuting, Ben Van Den Akker, Liu Ye, Bing-Jie Ni, Shane Watts, Katherine Reid, and Zhiguo Yuan. 2016. 'Unravelling the Spatial Variation of Nitrous Oxide Emissions from a Step-Feed Plug-Flow Full Scale Wastewater Treatment Plant'. *Scientific Reports* 6 (1): 20792.

Pan, Yuting, Bing-Jie Ni, Philip L Bond, Liu Ye, and Zhiguo Yuan. 2013. 'Electron Competition among Nitrogen Oxides Reduction during Methanol-Utilizing Denitrification in Wastewater Treatment'. *Water Research* 47 (10): 3273–81.

Pan, Yuting, Bing-Jie Ni, Huijie Lu, Kartik Chandran, David Richardson, and Zhiguo Yuan. 2015. 'Evaluating Two Concepts for the Modelling of Intermediates Accumulation during Biological Denitrification in Wastewater Treatment'. *Water Research* 71:21–31.

Pan, Yuting, Bing-Jie Ni, and Zhiguo Yuan. 2013. 'Modeling Electron Competition among Nitrogen Oxides Reduction and N2O Accumulation in Denitrification'. *Environmental Science & Technology* 47 (19): 11083–91.

Pan, Yuting, Liu Ye, Bing-Jie Ni, and Zhiguo Yuan. 2012. 'Effect of PH on N2O Reduction and Accumulation during Denitrification by Methanol Utilizing Denitrifiers'. *Water Research* 46 (15): 4832–40.

Pascal, Wunderlin, Siegrist Hansruedi, Tuzson Béla, Joss Adriano, Emmenegger Lukas, Mohn Joachim, and others. 2013. 'Isotope Signatures of N2O in a Mixed Microbial Population System: Constraints on N2O Producing Pathways in Wastewater Treatment'.

Patziger, M. 2021. 'Improving Wastewater Treatment Plant Performance by Applying CFD Models for Design and Operation: Selected Case Studies'. *Water Science and Technology* 84 (2): 323–32.

Pavissich, J P, B L Read-Daily, K Sandberg, F Sabba, and R Nerenberg. 2012. 'Nitrous Oxide (N2O) Reduction by Denitrifying Bacteria: Relating Kinetics and Gene Expression'. In *WEFTEC 2012*.

Peng, Chang, and Meng Fanchao. 2022. 'Fault Detection of Urban Wastewater Treatment Process Based on Combination of Deep Information and Transformer Network'. *IEEE Transactions on Neural Networks and Learning Systems*.

Peng, Chang, Wang Kai, Zheng Kun, and Meng Fanchao. 2022. 'Monitoring of Wastewater Treatment Process Based on Multi-Stage Variational Autoencoder'. *Expert Systems with Applications* 207:117919.

Peng, Lai, Bing-Jie Ni, Dirk Erler, Liu Ye, and Zhiguo Yuan. 2014. 'The Effect of Dissolved Oxygen on N2O Production by Ammonia-Oxidizing Bacteria in an Enriched Nitrifying Sludge'. *Water Research* 66:12–21.

Peng, Lai, Bing-Jie Ni, Yingyu Law, and Zhiguo Yuan. 2016. 'Modeling N2O Production by Ammonia Oxidizing Bacteria at Varying Inorganic Carbon Concentrations by Coupling the Catabolic and Anabolic Processes'. *Chemical Engineering Science* 144:386–94.

Peng, Lai, Bing-Jie Ni, Liu Ye, and Zhiguo Yuan. 2015. 'Selection of Mathematical Models for N2O Production by Ammonia Oxidizing Bacteria under Varying Dissolved Oxygen and Nitrite Concentrations'. *Chemical Engineering Journal* 281:661–68.

Pinkus, Allan. 1999. 'Approximation Theory of the MLP Model in Neural Networks'. *Acta Numerica* 8:143–95. https://doi.org/10.1017/S0962492900002919.

Pisa, Ivan, Ignacio Sant\'\in, José Lopez Vicario, Antoni Morell, and Ramon Vilanova. 2018. 'A Recurrent Neural Network for Wastewater Treatment Plant Effuents' Prediction'. *Actas de Las XXXIX Jornadas de Automática, Badajoz, 5-7 de Septiembre de 2018*.

Pisa, Ivan, Ignacio Santin, Antoni Morell, Jose Lopez Vicario, and Ramon Vilanova. 2019. 'LSTM-Based Wastewater Treatment Plants Operation Strategies for Effluent Quality Improvement'. *IEEE Access* 7:159773–86.

Pisa, Ivan, Ignacio Santín, Jose Lopez Vicario, Antoni Morell, and Ramon Vilanova. 2019. 'ANN-Based Soft Sensor to Predict Effluent Violations in Wastewater Treatment Plants'. *Sensors (Switzerland)* 19 (6): 1–26. https://doi.org/10.3390/s19061280.

Pishnamazi, Mahboobeh, Azam Marjani, Saeed Shirazian, and Mohammad Samipurgiri. 2012. 'Mathematical Modeling and Numerical Simulation of Wastewater Treatment Unit Using CFD'. *Oriental Journal of Chemistry* 28 (1): 51.

Pocquet, M, I Queinnec, and M Spérandio. 2013. 'Adaptation and Identification of Models for Nitrous Oxide (N2O) Production by Autotrophic Nitrite Reduction'. In

Proceedings 11th IWA Conference on Instrumentation, Control and Automation (ICA2013). Narbonne, France, September, 18–20.

Pocquet, Mathieu, Z Wu, Isabelle Queinnec, and M Spérandio. 2016. 'A Two Pathway Model for N2O Emissions by Ammonium Oxidizing Bacteria Supported by the NO/N2O Variation'. *Water Research* 88:948–59.

Poli, Michael, Stefano Massaroli, Atsushi Yamashita, Hajime Asama, Jinkyoo Park, and Stefano Ermon. 2021. 'TorchDyn: Implicit Models and Neural Numerical Methods in PyTorch'. In *Neural Information Processing Systems, Workshop on Physical Reasoning and Inductive Biases for the Real World*. Vol. 2.

Porro, J, U Rehman, T Flameling, G Bellandi, A Deeke, W Audenaert, S Weijers, and I Nopens. 2019. 'An Integrated CFD/Biokinetic/N2O Risk Modelling Approach for Mitigating N2O Emissions and Optimal WWTP Design'. In *10th IWA Symposium on Modelling and Integrated Assessment,(Watermatex 2019), Copenhagen, Denmark*.

Porro, Jose, Chaïm De Mulder, Youri Amerlinck, Elena Torfs, Sophie Balemans, Alexandra Deeke, Tony Flameling, et al. 2018. 'Integrating Artificial Intelligence and Mathematical Modelling for Online Supervision and Control of Water Resource Recovery Facilities'. In *WRRmod 2018*.

Postawa, Karol, Jerzy Szczygieł, and Marek Kułażyński. 2020. 'A Comprehensive Comparison of ODE Solvers for Biochemical Problems'. *Renewable Energy* 156 (August):624–33. https://doi.org/10.1016/J.RENENE.2020.04.089.

Poth, Mark, and Dennis D Focht. 1985. '15N Kinetic Analysis of N2O Production by Nitrosomonas Europaea: An Examination of Nitrifier Denitrification'. *Applied and Environmental Microbiology* 49 (5): 1134–41.

Poughon, Laurent, Claude-Gilles Dussap, and Jean-Bernard Gros. 2001. 'Energy Model and Metabolic Flux Analysis for Autotrophic Nitrifiers'. *Biotechnology and Bioengineering* 72 (4): 416–33.

Qian, Jiang, Yuren Wu, Bojin Zhuang, Shaojun Wang, and Jing Xiao. 2021. 'Understanding Gradient Clipping in Incremental Gradient Methods'. In *International Conference on Artificial Intelligence and Statistics*, 1504–12.

Quaghebeur, Ward, Elena Torfs, Bernard De Baets, and Ingmar Nopens. 2022. 'Hybrid Differential Equations: Integrating Mechanistic and Data-Driven Techniques for Modelling of Water Systems'. *Water Research* 213 (April). https://doi.org/10.1016/j.watres.2022.118166.

Rackauckas, Christopher, Yingbo Ma, Julius Martensen, Collin Warner, Kirill Zubov, Rohit Supekar, Dominic Skinner, Ali Ramadhan, and Alan Edelman. 2020. 'Universal Differential Equations for Scientific Machine Learning'. *ArXiv Preprint ArXiv:2001.04385*.

Rajta, Ankita, Ranjana Bhatia, Hema Setia, and Priyanka Pathania. 2020. 'Role of Heterotrophic Aerobic Denitrifying Bacteria in Nitrate Removal from Wastewater'. *Journal of Applied Microbiology* 128 (5): 1261–78.

Rani, Jyoti, Tapas Tripura, Hariprasad Kodamana, and Souvik Chakraborty. 2024. 'Generative Adversarial Wavelet Neural Operator: Application to Fault Detection and Isolation of Multivariate Time Series Data'. *ArXiv*, January. http://arxiv.org/abs/2401.04004.

Ren, Shuqi, Yingrui Liu, Yanying He, Tingting Zhu, Xueming Chen, and Yiwen Liu. 2023. 'Mathematical Modeling of the Dynamic Effect of Denitrifying Glycogen-Accumulating Organisms on Nitrous Oxide Production during Denitrifying Phosphorus Removal'. *Chemical Engineering Journal* 453:139802.

Ribeiro, Daniel, António Sanfins, and Orlando Belo. 2013. 'Wastewater Treatment Plant Performance Prediction with Support Vector Machines'. In *Advances in Data Mining. Applications and Theoretical Aspects: 13th Industrial Conference, ICDM 2013, New York, NY, USA, July 16-21, 2013. Proceedings 13*, 99–111.

Richardson, David, Heather Felgate, Nick Watmough, Andrew Thomson, and Elizabeth Baggs. 2009. 'Mitigating Release of the Potent Greenhouse Gas N2O from the Nitrogen Cycle–Could Enzymic Regulation Hold the Key?' *Trends in Biotechnology* 27 (7): 388–97.

Rodriguez-Caballero, Adrián, A Ribera, José Luis Balcázar, and M Pijuan. 2013. 'Nitritation versus Full Nitrification of Ammonium-Rich Wastewater: Comparison

in Terms of Nitrous and Nitric Oxides Emissions'. *Bioresource Technology* 139:195–202.

Roesch, Elisabeth, Christopher Rackauckas, and Michael P H Stumpf. 2021. 'Collocation Based Training of Neural Ordinary Differential Equations'. *Statistical Applications in Genetics and Molecular Biology* 20 (2): 37–49. https://doi.org/10.1515/sagmb-2020-0025.

Roman M. D., & Felseghi R A. 2014. 'Analysis of Oxygen Transfer and Dissolved Oxygen Concentration Measurement Tests in a Wastewater Treatment Plant'. *Applied Mechanics and Materials*. https://doi.org/10.4028/www.scientific.net/amm.656.486.

Ruthotto, Lars, and Eldad Haber. 2020. 'Deep Neural Networks Motivated by Partial Differential Equations'. *Journal of Mathematical Imaging and Vision* 62 (3): 352–64.

Salles, Rodrigo, Jérôme Mendes, Rita P Ribeiro, and João Gama. 2022. 'Fault Detection in Wastewater Treatment Plants: Application of Autoencoders Models with Streaming Data'. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, 55–70.

Samstag, R W, J J Ducoste, Alonso Griborio, Ingmar Nopens, Damien J Batstone, Jim D Wicks, Stephen Saunders, E A Wicklein, Genevieve Kenny, and Julien Laurent. 2016. 'CFD for Wastewater Treatment: An Overview'. *Water Science and Technology* 74 (3): 549–63.

Sandoval, Ilya Orson, Panagiotis Petsagkourakis, and Ehecatl Antonio del Rio-Chanona. 2022. 'Neural ODEs as Feedback Policies for Nonlinear Optimal Control', May. http://arxiv.org/abs/2210.11245.

Sanz-Serna, Jesus M. 1992. 'Symplectic Integrators for Hamiltonian Problems: An Overview'. *Acta Numerica* 1:243–86.

Schneider, Mariane Yvonne, Hidenori Harada, Kris Villez, and Max Maurer. 2021. 'How Well Can Inaccurate Sensors Quantify and Improve the Performance of a Fleet of On-Site Wastewater Treatment Plants?'

Seen, Wo Mei, R U Gobithaasan, and Kenjiro T Miura. 2014. 'GPU Acceleration of Runge Kutta-Fehlberg and Its Comparison with Dormand-Prince Method'. In *AIP Conference Proceedings*, 1605:16–21.

Seshan, Siddharth, Johann Poinapen, Jules B. Van Lier, and Zoran Kapelan. 2022. 'Biokinetic and Artificial Intelligence Models for the Simulation of Nitrous Oxide Emissions from Wastewater Treatment Plants'. In *WDSA CCWI 2022*. Universitat Politecnica de Valencia. https://doi.org/10.4995/wdsa-ccwi2022.2022.14158.

Seshan, Siddharth, Johann Poinapen, Marcel H. Zandvoort, Jules B. van Lier, and Zoran Kapelan. 2024. 'Limitations of a Biokinetic Model to Predict the Seasonal Variations of Nitrous Oxide Emissions from a Full-Scale Wastewater Treatment Plant'. *Science of the Total Environment* 917 (March). https://doi.org/10.1016/j.scitotenv.2024.170370.

Shaban, Wafaa Mohamed, Dongxi Xie, Khalid Elbaz, and Shui-Long Shen. 2024. 'Real-Time Water Quality Prediction of Wastewater Treatment Plants Using Advanced Deep Learning Networks'. *Journal of Water Process Engineering* 65:105775.

Shampine, Lawrence F. 2009. 'Stability of the Leapfrog/Midpoint Method'. *Applied Mathematics and Computation* 208 (1): 293–98.

Shanker M, Hu M Y, and Hung M S. 1996. 'Effect of Data Standardization on Neural Network Training'. *Int. J. Mgmt Sci*. Vol. 24.

Shao, Xuexin, Linli Zhao, Xuancai Sheng, and Ming Wu. 2020. 'Effects of Influent Salinity on Water Purification and Greenhouse Gas Emissions in Lab-Scale Constructed Wetlands'. *Environmental Science and Pollution Research* 27 (17): 21487–96. https://doi.org/10.1007/s11356-020-08497-7.

Shen, Yu, Xiaogang Zhu, Zhiwei Guo, Keping Yu, Osama Alfarraj, Victor C M Leung, and Joel J P C Rodrigues. 2024. 'A Deep Learning-Based Data Management Scheme for Intelligent Control of Wastewater Treatment Processes Under Resource-Constrained IoT Systems'. *IEEE Internet of Things Journal*.

Shi, Yu, William H Green, Hsi-Wu Wong, and Oluwayemisi O Oluwole. 2012. 'Accelerating Multi-Dimensional Combustion Simulations Using GPU and Hybrid Explicit/Implicit ODE Integration'. *Combustion and Flame* 159 (7): 2388–97.

Sin, Gürkan, Krist V Gernaey, Marc B Neumann, Mark C M van Loosdrecht, and Willi Gujer. 2011. 'Global Sensitivity Analysis in Wastewater Treatment Plant Model Applications: Prioritizing Sources of Uncertainty'. *Water Research* 45 (2): 639–51.

Skelboe, Stig, and Per Ulfkjaer Andersen. 1989. 'Stability Properties of Backward Euler Multirate Formulas'. *SIAM Journal on Scientific and Statistical Computing* 10 (5): 1000–1009.

Soler-Jofra, Aina, Berber Stevens, Maaike Hoekstra, Cristian Picioreanu, Dmitry Sorokin, Mark C M van Loosdrecht, and Julio Pérez. 2016. 'Importance of Abiotic Hydroxylamine Conversion on Nitrous Oxide Emissions during Nitritation of Reject Water'. *Chemical Engineering Journal* 287:720–26.

Song, Min Joon, Sangki Choi, Wo Bin Bae, Jaejin Lee, Heejoo Han, Daehyun D Kim, Miye Kwon, Jaewook Myung, Young Mo Kim, and Sukhwan Yoon. 2020. 'Identification of Primary Effecters of N2O Emissions from Full-Scale Biological Nitrogen Removal Systems Using Random Forest Approach'. *Water Research* 184:116144. https://doi.org/10.1016/j.watres.2020.116144.

Spérandio, Mathieu, Mathieu Pocquet, Lisha Guo, Bing-Jie Ni, Peter A Vanrolleghem, and Zhiguo Yuan. 2016. 'Evaluation of Different Nitrous Oxide Production Models with Four Continuous Long-Term Wastewater Treatment Process Data Series'. *Bioprocess and Biosystems Engineering* 39:493–510.

Stein, Lisa Y. 2011. 'Heterotrophic Nitrification and Nitrifier Denitrification'. *Nitrification*, 95–114.

Stein, Lisa Y, Daniel J Arp, Paul M Berube, Patrick S G Chain, Loren Hauser, Mike S M Jetten, Martin G Klotz, et al. 2007. 'Whole-Genome Analysis of the Ammonia-Oxidizing Bacterium, Nitrosomonas Eutropha C91: Implications for Niche Adaptation'. *Environmental Microbiology* 9 (12): 2993–3007.

Stentoft, Peter Alexander, T Munk-Nielsen, Jan Kloppenborg Møller, Henrik Madsen, Borja Valverde-Pérez, Peter Steen Mikkelsen, and Luca Vezzaro. 2021. 'Prioritize Effluent Quality, Operational Costs or Global Warming?–Using Predictive Control of Wastewater Aeration for Flexible Management of Objectives in WRRFs'. *Water Research* 196:116960.

Stieglmeier, Michaela, Maria Mooshammer, Barbara Kitzler, Wolfgang Wanek, Sophie Zechmeister-Boltenstern, Andreas Richter, and Christa Schleper. 2014. 'Aerobic Nitrous Oxide Production through N-Nitrosating Hybrid Formation in Ammonia-Oxidizing Archaea'. *The ISME Journal* 8 (5): 1135–46.

Su, Bingqin, Yuting Lin, Jian Wang, Xiaohui Quan, Zhankun Chang, and Chuangxue Rui. 2022. 'Sewage Treatment System for Improving Energy Efficiency Based on Particle Swarm Optimization Algorithm'. *Energy Reports* 8:8701–8.

Su, Qingxian, Carlos Domingo-Félez, Marlene Mark Jensen, and Barth F Smets. 2019. 'Abiotic Nitrous Oxide (N2O) Production Is Strongly PH Dependent, but Contributes Little to Overall N2O Emissions in Biological Nitrogen Removal Systems'. *Environmental Science & Technology* 53 (7): 3508–16. https://doi.org/10.1021/acs.est.8b06193.

Sun, Haishu, Shanghua Wu, Shugeng Feng, Cancan Jiang, Rui Wang, Shengjun Xu, Lijuan Cui, and Xuliang Zhuang. 2022. 'Impact of Influent Strengths on Nitrous Oxide Emission and Its Molecular Mechanism in Constructed Wetlands Treating Swine Wastewater'. *Environmental Research* 210:112957. https://doi.org/10.1016/j.envres.2022.112957.

Sun, Qingqiang, and Zhiqiang Ge. 2021. 'A Survey on Deep Learning for Data-Driven Soft Sensors'. *IEEE Transactions on Industrial Informatics* 17 (9): 5853–66. https://doi.org/10.1109/TII.2021.3053128.

Sun, Yifan, Linan Zhang, and Hayden Schaeffer. 2020. 'NeuPDE: Neural Network Based Ordinary and Partial Differential Equations for Modeling Time-Dependent Data'. In *Mathematical and Scientific Machine Learning*, 352–72.

Sundui, Batsuren, Olga Alejandra Ramirez Calderon, Omar M Abdeldayem, Jimena Lázaro-Gil, Eldon R Rene, and Uyanga Sambuu. 2021. 'Applications of Machine

Learning Algorithms for Biological Wastewater Treatment: Updates and Perspectives'. *Clean Technologies and Environmental Policy* 23:127–43.

Sweetapple, Christine, Guangtao Fu, and David Butler. 2014. 'Multi-Objective Optimisation of Wastewater Treatment Plant Control to Reduce Greenhouse Gas Emissions'. *Water Research* 55 (May):52–62. https://doi.org/10.1016/j.watres.2014.02.018.

Szeląg, B., E. Zaborowska, and J. Mąkinia. 2023. 'An Algorithm for Selecting a Machine Learning Method for Predicting Nitrous Oxide Emissions in Municipal Wastewater Treatment Plants'. *Journal of Water Process Engineering* 54 (August). https://doi.org/10.1016/j.jwpe.2023.103939.

Taheriyoun, Masoud, and Saber Moradinejad. 2015. 'Reliability Analysis of a Wastewater Treatment Plant Using Fault Tree Analysis and Monte Carlo Simulation'. *Environmental Monitoring and Assessment* 187:1–13.

Takács, Imre, Gilles G Patry, and Daniel Nolasco. 1991. 'A Dynamic Model of the Clarification-Thickening Process'. *Water Research* 25 (10): 1263–71.

Tao, E P, W H Shen, T L Liu, and X Q Chen. 2013. 'Fault Diagnosis Based on PCA for Sensors of Laboratorial Wastewater Treatment Process'. *Chemometrics and Intelligent Laboratory Systems* 128:49–55.

Tchobanoglous, G, H D Stensel, R Tsuchihashi, and F Burton. 2014. *Wastewater Engineering: Treatment and Resource Recovery*. 5th Intern. McGraw-Hill Education.

Tejaswini, E. S.S., P. Maheswari, and Seshagiri Rao Ambati. 2024. 'Integrated Supervisory Fuzzy Control Framework for Biological Wastewater Treatment Plants Operation and Their Effect on GHG Emissions'. *Chemical Engineering Science* 292 (June). https://doi.org/10.1016/j.ces.2024.119915.

Terada, Akihiko, Sho Sugawara, Keisuke Hojo, Yuki Takeuchi, Shohei Riya, Willie F Harper Jr, Tomoko Yamamoto, et al. 2017. 'Hybrid Nitrous Oxide Production from a Partial Nitrifying Bioreactor: Hydroxylamine Interactions with Nitrite'. *Environmental Science & Technology* 51 (5): 2748–56.

The MathWorks Inc. 2023. ‘Dynamical System Modeling Using Neural ODE’. 2023. https://uk.mathworks.com/help/deeplearning/ug/dynamical-system-modeling-using-neural-ode.html.

The MathWorks Inc. 2024. ‘MATLAB Version: 24.1.0.2628055 (R2024a) Update 4’. Natick, Massachusetts, United States: The MathWorks Inc. https://www.mathworks.com.

Thwaites, Benjamin J, Richard Stuetz, Michael Short, Petra Reeve, Juan-Pablo Alvarez-Gaitan, Nirmala Dinesh, Renae Philips, and Ben van den Akker. 2021. ‘Analysis of Nitrous Oxide Emissions from Aerobic Granular Sludge Treating High Saline Municipal Wastewater’. *Science of The Total Environment* 756:143653.

Treibert, Sarah, and Matthias Ehrhardt. 2022. ‘PINN Training Using Biobjective Optimization: The Trade-off between Data Loss and Residual Loss’. http://www.imacm.uni-wuppertal.de.

Tumendelger, Azzaya, Zeyad Alshboul, and Andreas Lorke. 2019. ‘Methane and Nitrous Oxide Emission from Different Treatment Units of Municipal Wastewater Treatment Plants in Southwest Germany’. *PloS One* 14 (1): e0209763.

Tuor, Aaron R, Ján Drgoňa, Drgoˇ Drgoňa, and Draguna L Vrabie. 2020. ‘Constrained Neural Ordinary Differential Equations with Stability Guarantees’. https://github.com/pnnl/neural_ODE_ICLR2020.

Tutueva, Aleksandra V, Ekaterina A Rodionova, Dmitrii O Pesterev, Artur I Karimov, and Maria N Kozak. 2020. ‘High-Performance Method-Switching ODE Solver’. In *2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, 1446–48.

Tzen, Belinda, and Maxim Raginsky. 2019. ‘Neural Stochastic Differential Equations: Deep Latent Gaussian Models in the Diffusion Limit’. *ArXiv Preprint ArXiv:1905.09883*.

U.S. EPA. 2024. ‘Overview of Greenhouse Gases’. 13 July 2024. https://www.epa.gov/ghgemissions/overview-greenhouse-gases.

Valkova, Tanya, Vanessa Parravicini, Ernis Saracevic, Joseph Tauber, Karl Svardal, and Jörg Krampe. 2020. ‘A Method to Estimate the Direct Nitrous Oxide Emissions of Municipal Wastewater Treatment Plants Based on the Degree of Nitrogen Removal’. *Journal of Environmental Management*, no. August. https://doi.org/10.1016/j.jenvman.2020.111563.

Vasilaki, V, V Conca, N Frison, A L Eusebi, F Fatone, and E Katsou. 2020. ‘A Knowledge Discovery Framework to Predict the N2O Emissions in the Wastewater Sector’. *Water Research* 178:115799. https://doi.org/10.1016/j.watres.2020.115799.

Vasilaki, V, Sebelan Danishvar, Ali Mousavi, and Evina Katsou. 2020. ‘Data-Driven versus Conventional N2O EF Quantification Methods in Wastewater; How Can We Quantify Reliable Annual EFs?’ *Computers & Chemical Engineering* 141:106997.

Vasilaki, V, T M Massara, P Stanchev, F Fatone, and E Katsou. 2019. ‘A Decade of Nitrous Oxide (N2O) Monitoring in Full-Scale Wastewater Treatment Processes: A Critical Review’. *Water Research* 161:392–412. https://doi.org/10.1016/j.watres.2019.04.022.

Vasilaki, V, E I P Volcke, A K Nandi, M C M van Loosdrecht, and E Katsou. 2018. ‘Relating N2O Emissions during Biological Nitrogen Removal with Operating Conditions Using Multivariate Statistical Techniques’. *Water Research* 140:387–402. https://doi.org/10.1016/j.watres.2018.04.052.

Verdaguer, M, N Clara, O Gutiérrez, and M Poch. 2014. ‘Application of Ant-Colony-Optimization Algorithm for Improved Management of First Flush Effects in Urban Wastewater Systems’. *Science of the Total Environment* 485:143–52.

Verdaguer, Marta, Narcis Clara, and Manel Poch. 2012. ‘Ant Colony Optimization-Based Method for Managing Industrial Influents in Wastewater Systems’. *AIChE Journal* 58 (10): 3070–79.

Verdaguer, Marta, Mar\’\ia Molinos-Senante, and Manel Poch. 2016. ‘Optimal Management of Substrates in Anaerobic Co-Digestion: An Ant Colony Algorithm Approach’. *Waste Management* 50:49–54.

Verma, Yogesh, Markus Heinonen, and Vikas Garg. 2024. 'ClimODE: Climate and Weather Forecasting with Physics-Informed Neural ODEs'.

Vijayaraghavan, G, M Jayalakshmi, and others. 2015. 'A Quick Review on Applications of Fuzzy Logic in Waste Water Treatment'. *Int. J. Res. Appl. Sci. Eng. Technol* 3 (5): 421–25.

Virtanen, Pauli, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, et al. 2020. 'SciPy 1.0: Fundamental Algorithms for Scientific in Python'. *Nature Methods* 17:261–72. https://doi.org/10.1038/s41592-019-0686-2.

Wan, Xinyu, and Eveline I P Volcke. 2022. 'Dynamic Modelling of N2O Emissions from a Full-Scale Granular Sludge Partial Nitritation-Anammox Reactor'. *Biotechnology and Bioengineering* 119 (6): 1426–38.

Wang, Dong, Sven Thunéll, Ulrika Lindberg, Lili Jiang, Johan Trygg, Mats Tysklind, and Nabil Souihi. 2021. 'A Machine Learning Framework to Improve Effluent Quality Control in Wastewater Treatment Plants'. *Science of the Total Environment* 784:147138.

Wang, Gongming, Jing Bi, Qing-Shan Jia, Junfei Qiao, and Lei Wang. 2022. 'Event-Driven Model Predictive Control with Deep Learning for Wastewater Treatment Process'. *IEEE Transactions on Industrial Informatics* 19 (5): 6398–6407.

Wang, Gongming, Qing-Shan Jia, Junfei Qiao, Jing Bi, and MengChu Zhou. 2020. 'Deep Learning-Based Model Predictive Control for Continuous Stirred-Tank Reactor System'. *IEEE Transactions on Neural Networks and Learning Systems* 32 (8): 3643–52.

Wang, Gongming, Yidi Zhao, Caixia Liu, and Junfei Qiao. 2023. 'Data-Driven Robust Adaptive Control with Deep Learning for Wastewater Treatment Process'. *IEEE Transactions on Industrial Informatics* 20 (1): 149–57.

Wang, Jian-Hui, Xiao-Long Zhao, Zhi-Wei Guo, Peng Yan, Xu Gao, Yu Shen, and You-Peng Chen. 2022. 'A Full-View Management Method Based on Artificial Neural Networks for Energy and Material-Savings in Wastewater Treatment Plants'. *Environmental Research* 211:113054.

Wang, Qilin, Bing-Jie Ni, Romain Lemaire, Xiaodi Hao, and Zhiguo Yuan. 2016. 'Modeling of Nitrous Oxide Production from Nitritation Reactors Treating Real Anaerobic Digestion Liquor'. *Scientific Reports* 6 (1): 25336.

Water UK, Ricardo, and Mott MacDonald. 2020. 'Net Zero 2030 Routemap', 90.

Weiss, Karl, Taghi M Khoshgoftaar, and DingDing Wang. 2016. 'A Survey of Transfer Learning'. *Journal of Big Data* 3:1–40.

Wenbing, F A N, and Zhenzhen ZHANG. 2020. 'A CNN-SVR Hybrid Prediction Model for Wastewater Index Measurement'. In *2020 2nd International Conference on Advances in Computer Technology, Information Science and Communications (CTISC)*, 90–94.

White, Corey J, and Nicolai Lehnert. 2016. 'Is There a Pathway for N2O Production from Hydroxylamine Oxidoreductase in Ammonia-Oxidizing Bacteria?' *Proceedings of the National Academy of Sciences* 113 (51): 14474–76.

Wisniewski, K, Maciej Kowalski, and J Makinia. 2018. 'Modeling Nitrous Oxide Production by a Denitrifying-Enhanced Biologically Phosphorus Removing (EBPR) Activated Sludge in the Presence of Different Carbon Sources and Electron Acceptors'. *Water Research* 142:55–64.

WMO. 2024. 'WMO Global Annual to Decadal Climate Update'. https://library.wmo.int/records/item/68910-wmo-global-annual-to-decadal-climate-update.

WMO (2025) *State of the Global Climate 2024 (WMO-No. 1368)*. Geneva, Switzerland: World Meteorological Organization. Available at: https://wmo.int/publication-series/state-of-global-climate-2024.

Wongburi, Praewa, and Jae K Park. 2023. 'Prediction of Wastewater Treatment Plant Effluent Water Quality Using Recurrent Neural Network (RNN) Models'. *Water* 15 (19): 3325.

Wrage-Mönnig, Nicole, Marcus A Horn, Reinhard Well, Christoph Müller, Gerard Velthof, and Oene Oenema. 2018. 'The Role of Nitrifier Denitrification in the Production of Nitrous Oxide Revisited'. *Soil Biology and Biochemistry* 123:A3–A16.

Wu, Bing, Rudraksha D Majumdar, Daniel H Lysak, Rajshree Ghosh Biswas, Maryam Tabatabaei-Anaraki, Amy Jenne, Xiang You, et al. 2021. 'Towards Real-Time Kinetic Monitoring of Wastewater Treatment: A Case Study of Sunlight and Ozone Treatment of Unconcentrated Wastewater Using Flow NMR'. *Chemical Engineering Journal* 405 (July 2020): 126696. https://doi.org/10.1016/j.cej.2020.126696.

Wunderlin, Pascal, Moritz F Lehmann, Hansruedi Siegrist, Béla Tuzson, Adriano Joss, Lukas Emmenegger, and Joachim Mohn. 2013. 'Isotope Signatures of N2O in a Mixed Microbial Population System: Constraints on N2O Producing Pathways in Wastewater Treatment'. *Environmental Science & Technology* 47 (3): 1339–48.

Wunderlin, Pascal, Joachim Mohn, Adriano Joss, Lukas Emmenegger, and Hansruedi Siegrist. 2012. 'Mechanisms of N2O Production in Biological Wastewater Treatment under Nitrifying and Denitrifying Conditions'. *Water Research* 46 (4): 1027–37. https://doi.org/10.1016/j.watres.2011.11.080.

Xiang, Jinlin, Bozhao Qi, Marc Cerou, Wei Zhao, and Qi Tang. 2024. 'DN-ODE: Data-Driven Neural-ODE Modeling for Breast Cancer Tumor Dynamics and Progression-Free Survivals'. *Computers in Biology and Medicine* 180:108876.

Xie, Xiang, Ajith Kumar Parlikad, and Ramprakash Srinivasan Puri. 2019. 'A Neural Ordinary Differential Equations Based Approach for Demand Forecasting within Power Grid Digital Twins'. In *2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids, SmartGridComm 2019*. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/SmartGridComm.2019.8909789.

Xu, Boyan, Ching Kwek Pooi, Kar Ming Tan, Shujuan Huang, Xueqing Shi, and How Yong Ng. 2023. 'A Novel Long Short-Term Memory Artificial Neural Network (LSTM)-Based Soft-Sensor to Monitor and Forecast Wastewater Treatment Performance'. *Journal of Water Process Engineering* 54:104041.

Xu, Boyan, Ching Kwek Pooi, Tsuey Shan Yeap, Kwok Yii Leong, Xi Yee Soh, Shujuan Huang, Xueqing Shi, Giorgio Mannina, and How Yong Ng. 2024. 'Hybrid Model Composed of Machine Learning and ASM3 Predicts Performance of Industrial Wastewater Treatment'. *Journal of Water Process Engineering* 65:105888.

Xu, Boyan, Liang Wen, Zihao Li, Yuxing Yang, Guanlan Wu, Xiongpeng Tang, Yu Li, et al. 2024. ‘Unlocking the Potential: Benchmarking Large Language Models in Water Engineering and Research’. *ArXiv Preprint ArXiv:2407.21045*.

Xu, Xiaozhen, Anlei Wei, Songjun Tang, Qi Liu, Hanxiao Shi, and Wei Sun. 2024. ‘Prediction of Nitrous Oxide Emission of a Municipal Wastewater Treatment Plant Using LSTM-Based Deep Learning Models’. *Environmental Science and Pollution Research* 31 (2): 2167–86. https://doi.org/10.1007/s11356-023-31250-9.

Xue, Yexiang, Md Nasim, Maosen Zhang, Cuncai Fan, Xinghang Zhang, and Anter El-Azab. 2021. ‘Physics Knowledge Discovery via Neural Differential Equation Embedding’.

Yamazaki, T, T Hozuki, K Arai, S Toyoda, K Koba, T Fujiwara, and N Yoshida. 2014. ‘Isotopomeric Characterization of Nitrous Oxide Produced by Reaction of Enzymes Extracted from Nitrifying and Denitrifying Bacteria’. *Biogeosciences* 11 (10): 2679–89.

Yan, Xu, Jie Yang, Dongli Guo, Jiahui Ma, Xianfa Su, and Jianhui Sun. 2021. ‘Effect of Carbon Source on Nitrous Oxide Emission Characteristics and Sludge Properties during Anoxic/Aerobic Wastewater Treatment Process’. *Environmental Science and Pollution Research* 28 (41): 57557–68. https://doi.org/10.1007/s11356-021-14713-9.

Yan, Xu, Shikan Zheng, Dezhi Qiu, Jie Yang, Yunping Han, Zhaoman Huo, Xianfa Su, and Jianhui Sun. 2019. ‘Characteristics of N2O Generation within the Internal Micro-Environment of Activated Sludge Flocs under Different Dissolved Oxygen Concentrations’. *Bioresource Technology* 291:121867.

Yang, Luxuan, Ting Gao, Yubin Lu, Jinqiao Duan, and Tao Liu. 2023. ‘Neural Network Stochastic Differential Equation Models with Applications to Financial Data Forecasting’. *Applied Mathematical Modelling* 115:279–99.

Yang, Qinmin, Weiwei Cao, Wenchao Meng, and Jennie Si. 2021. ‘Reinforcement-Learning-Based Tracking Control of Waste Water Treatment Process under

Realistic System Conditions and Control Performance Requirements'. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 52 (8): 5284–94.

Yang, Rui, Lin-jiang Yuan, Ru Wang, Zhi-xian He, Lin Lei, and Yan-chen Ma. 2022. 'Analyzing the Mechanism of Nitrous Oxide Production in Aerobic Phase of Anoxic/Aerobic Sequential Batch Reactor from the Perspective of Key Enzymes'. *Environmental Science and Pollution Research* 29 (26): 39877–87.

Yaqub, Muhammad, Hasnain Asif, Seongboem Kim, and Wontae Lee. 2020. 'Modeling of a Full-Scale Sewage Treatment Plant to Predict the Nutrient Removal Efficiency Using a Long Short-Term Memory (LSTM) Neural Network'. *Journal of Water Process Engineering* 37:101388.

Ye, Haotian, Chuanlong Xie, Tianle Cai, Ruichen Li, Zhenguo Li, and Liwei Wang. 2021. 'Towards a Theoretical Framework of Out-of-Distribution Generalization'. *Advances in Neural Information Processing Systems* 34:23519–31.

Ye, Liu, Jose Porro, and Ingmar Nopens. 2022. *Quantification and Modelling of Fugitive Greenhouse Gas Emissions from Urban Water Systems*. First Edition. IWA Publishing.

Ye, Xudong, Bing Chen, Liang Jing, Baiyu Zhang, and Yong Liu. 2019. 'Multi-Agent Hybrid Particle Swarm Optimization (MAHPSO) for Wastewater Treatment Network Planning'. *Journal of Environmental Management* 234:525–36.

Yi, Zhang. 2023. 'NmODE: Neural Memory Ordinary Differential Equation'. *Artificial Intelligence Review* 56 (12): 14403–38.

Yin, Zhixuan, Xuejun Bi, and Chenlu Xu. 2018. 'Ammonia-Oxidizing Archaea (AOA) Play with Ammonia-Oxidizing Bacteria (AOB) in Nitrogen Removal from Wastewater'. *Archaea* 2018 (1): 8429145.

Ying, Lv, Zhang Shoubin, Xie Kang, Liu Guicai, Qiu Liping, Liu Yutian, and Zhang Yuanyuan. 2022. 'Establishment of Nitrous Oxide (N2O) Dynamics Model Based on ASM3 Model during Biological Nitrogen Removal via Nitrification'. *Environmental Technology* 43 (8): 1170–80. https://doi.org/10.1080/09593330.2020.1822447.

Yu, Ran, Octavio Perez-Garcia, Huijie Lu, and Kartik Chandran. 2018. 'Nitrosomonas Europaea Adaptation to Anoxic-Oxic Cycling: Insights from Transcription Analysis, Proteomics and Metabolic Network Modeling'. *Science of The Total Environment* 615:1566–73.

Yue, Xiu, Guangping Yu, Yuqian Lu, Zhuhan Liu, Qianhua Li, Jiali Tang, and Jian Liu. 2018. 'Effect of Dissolved Oxygen on Nitrogen Removal and the Microbial Community of the Completely Autotrophic Nitrogen Removal over Nitrite Process in a Submerged Aerated Biological Filter'. *Bioresource Technology* 254:67–74.

Zakwan, Muhammad, Loris Di Natale, Bratislav Svetozarevic, Philipp Heer, Colin N Jones, and G Ferrari Trecate. 2023. 'Physically Consistent Neural ODEs for Learning Multi-Physics Systems'. *IFAC-PapersOnLine* 56 (2): 5855–60.

Zamora, D, and A Torres. 2014. 'Method for Outlier Detection: A Tool to Assess the Consistency between Laboratory Data and Ultraviolet–Visible Absorbance Spectra in Wastewater Samples'. *Water Science and Technology* 69 (11): 2305–14.

Zeferino, João A, António P Antunes, and Maria C Cunha. 2009. 'An Efficient Simulated Annealing Algorithm for Regional Wastewater System Planning'. *Computer-Aided Civil and Infrastructure Engineering* 24 (5): 359–70.

Zhan, Xinmin, Zhenhu Hu, and Guangxue Wu. 2018. *Greenhouse Gas Emission and Mitigation in Municipal Wastewater Treatment Plants*. IWA publishing.

Zhang, Jingzhao, Tianxing He, Suvrit Sra, and Ali Jadbabaie. 2019. 'Why Gradient Clipping Accelerates Training: A Theoretical Justification for Adaptivity'. *ArXiv Preprint ArXiv:1905.11881*.

Zhang, Qianqian, Zhong Li, Spencer Snowling, Ahmad Siam, and Wael El-Dakhakhni. 2019. 'Predictive Models for Wastewater Flow Forecasting Based on Time Series Analysis and Artificial Neural Network'. *Water Science and Technology* 80 (2): 243–53.

Zhang, Wenjin, Nicholas B Tooker, and Amy V Mueller. 2020. 'Enabling Wastewater Treatment Process Automation: Leveraging Innovations in Real-Time Sensing,

Data Analysis, and Online Controls'. *Environmental Science: Water Research & Technology* 6 (11): 2973–92.

Zhang, Xingxuan, Peng Cui, Renzhe Xu, Linjun Zhou, Yue He, and Zheyan Shen. 2021. 'Deep Stable Learning for Out-of-Distribution Generalization'. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5372–82.

Zhang, Yituo, Chaolin Li, Hengpan Duan, Kefen Yan, Jihong Wang, and Wenhui Wang. 2023. 'Deep Learning Based Data-Driven Model for Detecting Time-Delay Water Quality Indicators of Wastewater Treatment Plant Influent'. *Chemical Engineering Journal* 467:143483.

Zhang, Yue, Genki Suzuki, and Hiroyuki Shioya. 2022. 'Prediction and Detection of Sewage Treatment Process Using N-BEATS Autoencoder Network'. *IEEE Access* 10:112594–608.

Zhang, Zexi, Wei Xing, Jia Lu, Xinyu Gao, Fangxu Jia, and Hong Yao. 2024. 'Nitrogen Removal and Nitrous Oxide Emission in the Partial Nitritation/Anammox Process at Different Reflux Ratios'. *Science of The Total Environment* 906:167520.

Zhao, Junkai, Shuhan Lei, Guangwei Cheng, Ju Zhang, Bingfeng Shi, Shuting Xie, and Jianqiang Zhao. 2022. 'Comparison of Inhibitory Roles on Nitrite-Oxidizing Bacteria by Hydroxylamine and Hydrazine during the Establishment of Partial Nitrification'. *Bioresource Technology* 355:127271.

Zhao, Junkai, Jianqiang Zhao, Shuting Xie, and Shuhan Lei. 2021. 'The Role of Hydroxylamine in Promoting Conversion from Complete Nitrification to Partial Nitrification: NO Toxicity Inhibition and Its Characteristics'. *Bioresource Technology* 319:124230.

Zhao, Junkai, Jianqiang Zhao, Wenjuan Yang, Bo Hu, Ting Huang, Shuting Xie, Shuhan Lei, and Wei Hou. 2022. 'Mechanisms of NO and N2O Production by Enriched Nitrifying Sludge in a Sequencing Batch Reactor: Effects of Hydroxylamine'. *Journal of Environmental Management* 316:115237.

Zhao, Lin, Yi Ji, Jiayi Yao, Sha Long, Dong Li, and Yongkui Yang. 2017. 'Quantifying the Fate and Risk Assessment of Different Antibiotics during Wastewater

Treatment Using a Monte Carlo Simulation'. *Journal of Cleaner Production* 168:626–31.

Zhao, Wei, Yayi Wang, Ximao Lin, Dong Zhou, Mianli Pan, and Jian Yang. 2014. 'Identification of the Salinity Effect on N2O Production Pathway during Nitrification: Using Stepwise Inhibition and 15N Isotope Labeling Methods'. *Chemical Engineering Journal* 253:418–26.

Zhong, Shifa, Kai Zhang, Majid Bagheri, Joel G Burken, April Gu, Baikun Li, Xingmao Ma, et al. 2021. 'Machine Learning: New Ideas and Tools in Environmental Science and Engineering'. *Environmental Science & Technology* 55 (19): 12741–54.

Zhong, Yaofeng Desmond, Biswadip Dey, and Amit Chakraborty. 2019. 'Symplectic Ode-Net: Learning Hamiltonian Dynamics with Control'. *ArXiv Preprint ArXiv:1909.12077*.

Zhou, Fan, Liang Li, Kunpeng Zhang, and Goce Trajcevski. 2021. 'Urban Flow Prediction with Spatial–Temporal Neural ODEs'. *Transportation Research Part C: Emerging Technologies* 124:102912.

Zhou, Nan, Chenyuan Dang, Zhirong Zhao, Shishi He, Maosheng Zheng, Wen Liu, and Xiangke Wang. 2019. 'Role of Sludge Retention Time in Mitigation of Nitrous Oxide Emission from a Pilot-Scale Oxidation Ditch'. *Bioresource Technology* 292:121961.

Zhou, Pengxiao, Zhong Li, Spencer Snowling, Brian W Baetz, Dain Na, and Gavin Boyd. 2019. 'A Random Forest Model for Inflow Prediction at Wastewater Treatment Plants'. *Stochastic Environmental Research and Risk Assessment* 33:1781–92.

Zhou, Yan, Melvin Lim, Soekendro Harjono, and Wun Jern Ng. 2012. 'Nitrous Oxide Emission by Denitrifying Phosphorus Removal Culture Using Polyhydroxyalkanoates as Carbon Source'. *Journal of Environmental Sciences* 24 (9): 1616–23.

Zhou, Yiwen, Shuting Zhao, Toshikazu Suenaga, Megumi Kuroiwa, Shohei Riya, and Akihiko Terada. 2022. ‘Nitrous Oxide-Sink Capability of Denitrifying Bacteria Impacted by Nitrite and PH’. *Chemical Engineering Journal* 428:132402.

Zhu, Xia, Martin Burger, Timothy A Doane, and William R Horwath. 2013. ‘Ammonia Oxidation Pathways and Nitrifier Denitrification Are Significant Sources of N2O and NO under Low Oxygen Availability’. *Proceedings of the National Academy of Sciences* 110 (16): 6328–33.

Zhuang, Juntang, Nicha Dvornek, Xiaoxiao Li, Sekhar Tatikonda, Xenophon Papademetris, and James Duncan. 2020. ‘Adaptive Checkpoint Adjoint Method for Gradient Estimation in Neural Ode’. In *International Conference on Machine Learning*, 11639–49.

Zhu-Barker, Xia, Amanda R. Cavazos, Nathaniel E. Ostrom, William R. Horwath, and Jennifer B. Glass. 2015. ‘The Importance of Abiotic Reactions for Nitrous Oxide Production’. *Biogeochemistry* 126 (3): 251–67. https://doi.org/10.1007/s10533-015-0166-4.

Zou, Bob Junyi, Matthew E. Levine, Dessi P. Zaharieva, Ramesh Johari, and Emily B. Fox. 2024. ‘Hybrid^2 Neural ODE Causal Modelling’, February. http://arxiv.org/abs/2402.17233.