Methodologies in Mechanistic and Machine Learning Modelling of Nitrous Oxide Emissions from Wastewater Treatment Processes

by

Mostafa Wagih Tawfik Khalil

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Environmental Engineering

Department of Civil and Environmental Engineering

University of Alberta

Abstract

Nitrous oxide (N_2O), a potent greenhouse gas, significantly contributes to the carbon footprint of wastewater treatment plants (WWTPs), with a global warming potential 300 times that of CO_2 and also notorious for its ozone-depleting effect. Modelling can be a valuable alternative to quantification of N_2O emissions through monitoring campaigns. Traditional mechanistic models have been limited in their application by impractical calibration processes and a lack of validation for complex treatment systems such as reactors combining biofilm and flocculent biomasses and/or highly dynamic operations. Conversely, machine learning (ML) presents a promising avenue, leveraging abundant WWTP data. However, the scarcity of comprehensive methodological ML frameworks for environmental engineering and the complexity of ML challenge their acceptance among practitioners. This thesis bridges evaluating and refining both modelling approaches to enhance predictability and highlights the areas for future research focus.

In this thesis, two-pathway nitrification and a multi-step denitrification N_2O model was adapted to an Integrated Fixed-Film Activated Sludge (IFAS) system operating within a Sequencing Batch Reactor (SBR), using data from a laboratory-scale experiment. The model, characterized by a one-dimensional (1-D) biofilm, underwent a two-step calibration process informed by sensitivity and identifiability analyses. While it achieved good alignment with experimental data, revealing the model's predictive capacity, its application to different operational conditions in the validation data illustrated limitations in generalization. Further, through extensive simulations (over 1500), the influence of dissolved oxygen (DO), temperature, and ammonia levels was explored. These simulations highlighted the critical role of temperature in

setting optimal DO levels, crucial for balancing N_2O emission reduction with enhanced ammonia removal efficiency.

Parallel to the mechanistic approach, this thesis pioneers a comprehensive ML methodology for N_2O emissions modeling using a long-term dataset from a full-scale WWTP. Recognizing the need for an online monitoring tool that also supports decision making, the proposed approach emphasizes not just model accuracy, but it also considers model complexity, computational speed, and interpretability. Various algorithms, including k-Nearest Neighbors (kNN), decision trees, Deep Neural Networks (DNN), and ensemble learning models such as extreme gradient boosting (XGBoost), adaptive boosting (AdaBoost), and random forest were evaluated. A novel adjustment of a parametric multivariate outlier removal method aligned with data distributions, minimizing data loss. An effective feature selection strategy optimized the balance between data acquisition, model performance, and complexity—cutting feature count by 40% without compromising accuracy. Additionally, an integrated method combining feature selection with hyperparameter optimization (HPO) was introduced, leveraging a Genetic Algorithm (GA), specifically NSGA-II, against the Nelder-Mead algorithm to navigate the intricate, nonlinear data landscape. This comparison underscored effectiveness of GAs in streamlining model complexity and enhancing performance, paving the way for the development of interpretable, computationally efficient ML tools especially for real-time applications.

This thesis reveals that while the applied N_2O mechanistic model offers acceptable predictions within the operational schemes used for calibration, its applicability to varied settings is limited. Despite the complexity of calibration, mechanistic models emerge as indispensable tools for scenario analysis, enhancing design and planning with precise "what if" explorations. In this

context, the study underscores the critical influence of temperature in guiding optimal DO setpoints for effective ammonia removal and emission mitigation. On the ML front, models like kNN and AdaBoost not only demonstrated high accuracy in long-term emission prediction but also challenged the assumed necessity for deep learning, offering simpler, yet effective alternatives.

The developed holistic modelling framework aids the application of ML models as practical tools for ongoing process monitoring, owing to their accuracy, lower complexity, and adaptability.

This thesis contributes to the field by developing methodologies of employing mechanistic and ML models for prediction and mitigation of N_2O emissions. The thesis offers a methodological framework for N_2O emission modelling using ML, with insights that hold potential for broader application in the field of wastewater treatment, equipping practitioners with a robust toolkit for addressing environmental challenges.

Preface

The contents of this thesis are my original work under the supervision of Prof. Yang Liu at the University of Alberta, and co-supervision of Prof. Peter Vanrolleghem at Université Laval, who have made significant contributions to all areas of this research, including conceptualization, methodology, supervision, writing-review & editing, project administration, resources, and funding acquisition. Some colleagues also contributed to the manuscript, and their contributions are listed as follows:

Chapter 4:

A version of this chapter has been published: Khalil, M., AlSayed A., Liu Y., Vanrolleghem P.A. (2023). Machine learning for modeling N_2O emissions from wastewater treatment plants: Aligning model performance, complexity, and interpretability. *Water Research*. 245, 120667.

https://doi.org/10.1016/j.watres.2023.120667

Dr. Ahmed AlSayed contributed to the methodology, investigation, and review and editing of the manuscript.

Chapter 5:

A version of this chapter has been submitted to *Journal of Water Process Engineering* in December 2023: Khalil, M., AlSayed A., Liu Y., Vanrolleghem P.A. An integrated feature selection and hyperparameter optimization algorithm for balanced machine learning models predicting N_2O emissions from wastewater treatment plants.

https://doi.org/10.1016/j.watres.2023.120667

Dr. Ahmed AlSayed contributed to the methodology, investigation, and review and editing of the manuscript.

Acknowledgement

I wish to extend my gratitude to my supervisor, Professor Yang Liu, for her unwavering support and encouragement throughout my doctoral program. My PhD journey was marked by unprecedented challenges, including a global pandemic that disrupted my initial progress, followed by issues related to supply and data. Despite these hurdles, Professor Liu provided me with the freedom to steer my project, trusting my decisions and thereby significantly enriching my management skills. Her support has been invaluable, and I am thankful for her mentorship.

I am profoundly grateful to my co-supervisor, Professor Peter Vanrolleghem, who became the cornerstone of my academic growth. His belief in my potential and guidance have been the beacon that lit my path. Peter is not just a mentor but a remarkable example of what it means to be an exceptional professor—generous with his time, rich in knowledge, and always believing in the possibility of my success. His contributions to my academic pursuits went beyond mere guidance; they were acts of genuine commitment to my growth and achievements. Our collaboration has evolved into a relationship that I treasure deeply, marked by mutual respect and a shared passion for discovery. His unwavering support and encouragement have been instrumental in my journey, and for that, I am eternally thankful.

I also owe a great deal of gratitude to Dr. Ahmed AlSayed, whose friendship and collaboration have been of significant help in my PhD journey. Ahmed has been an unwavering source of encouragement and belief in my capabilities. His enthusiasm for my work and his confidence in my success have been nothing short of inspirational. The hours spent brainstorming with him have enriched my research. His insights and support in a time of disappointment and frustration

have played a pivotal role in my ability to persevere and complete my work. For his generosity, guidance, and genuine belief in me, I am profoundly thankful.

My teaching experience has been a significant highlight of my PhD journey, and for that, I would like to thank all the course instructors I had the privilege of working alongside in teaching at the University of Alberta: Dr. Ian Buchanan, Dr. Bipro Dhar, Dr. Steve Craik, Dr. Yang Liu, Mr. Akinbola George, and Dr. Tong Yu.

This work could not have been completed without the support of my beloved wife, Sara Abdelmotaal. Sara; your enduring love, boundless patience, and constant encouragement have been my source of strength in times of stress and doubt. You believed in me even when I struggled to believe in myself, and your faith in my capabilities has been a source of strength that propelled me forward. Your presence in my life enriches every aspect of it, and this achievement is as much yours as it is mine.

My deepest appreciation to my beloved late father, Wagih, and my dear mother, Nagwa. whose sacrifices have shaped me into the person I have become. To my beloved siblings, Riham, Essam, and Reem, for their unconditional love, guidance, and unwavering belief in me. I am immensely grateful to my friends for their unwavering encouragement and support throughout our journey together. I am grateful for the roles each of you has played in my life, offering me support, wisdom, and companionship at every turn. This achievement is a testament to the love and support that you have all generously shared with me. My heartful thanks go to Taher, Bota, Fakhry, Khalid, Adel, Osama, and Sherif.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1 Background

The increasing effect of climate change resulting from elevated anthropogenic greenhouse gas (GHG) emissions is having a significant impact on the water sector including droughts, extreme weather events, flooding, and water quality issues (Höhne et al., 2020). Given the urgency needed to fight climate change and the paradigm shift towards sustainability, many infrastructure sectors have been extensively studying the pathways towards decarbonization (i.e., net-zero GHG emissions). Meeting the Paris Agreement goal of limiting the temperature increase to 1.5°C requires more sectors to commit to net zero GHG emissions by 2050, including the water sector (Rogelj et al., 2018). However, pathways to net zero GHG emissions in the water sector, particularly the wastewater treatment sector, are different and more complex than other infrastructure sectors. GHG emissions from wastewater treatment plants (WWTPs) are generally classified into three main scopes (Brotto and Lake, 2022). Scope 1 GHG emissions are those that are produced directly during biological wastewater treatment processes, primarily methane (CH_4), and nitrous oxide (N_2O). Scope 2 includes the indirect GHG emissions associated with electricity consumption during the operation of WWTPs. Lastly, scope 3 refers to the GHGs produced indirectly outside the utility's control, such as emissions resulting from purchased chemicals and construction materials. Among others, scope 1 GHG emissions are of significant concern because the non-CO_2 emissions have a stronger global warming potential (GWP) than CO_2 over 100 years. In particular, N_2O is a very potent GHG that is emitted during biological nitrogen remval processes and can contribute up to 80% of the total carbon footprint of a WWTP (Daelman et al., 2013). N_2O has a GWP that is approximately 300 times stronger

than that of CO_2, and is also claimed to be the main ozone depleting substance of the 21st century (Ravishankara et al., 2009).

Estimation of N_2O emissions from WWTPs has thus become necessary as the first step of mitigation, given the temporal and spatial variability shown in previous studies. There is a lack of a practical and low-cost approach for quantification of N_2O emissions. Monitoring campaigns have demonstrated to exhibit notable constraints, including substantial costs and intricate processes (Vasilaki et al., 2020b). Factors such as the campaign's duration, sampling frequency, and measurement techniques have been found to considerably influence the outcomes, thereby further restricting the applicability of monitoring campaigns. Mathematical modelling has been explored as an alternative approach to quantify N_2O emissions and overcome the limitations linked to traditional monitoring campaigns. A mathematical model can also be used to select mitigation strategies and optimize process operation for less N_2O emissions. Depending on the desired application, a model can be used as a stand-alone tool during the design stage of a WWTP, part of a digital twin, deployed on a software (soft) sensor for online monitoring, or be used to implement a model predictive control (MPC).

Over the last decade, N_2O models have been developed based on the knowledge of production, consumption, and mass-transfer mechanisms, often called “mechanistic models”. These mechanistic models varied in complexity as a result of the increasing advancements in the knowledge of N_2O biological pathways. However, uncertainties remain; mechanistic models were not able to accurately depict process dynamics in full-scale applications and showed difficulties in the calibration process (Seshan et al., 2024).

On the other hand, as a large amount of data is becoming available in WWTPs from sensors and laboratories, a data-driven approach that can predict N_2O emissions based on available data holds significant promise in replacing mechanistic models. Machine learning (ML) emerges as a potential modelling approach. However, a limited number of studies focused on the quantitative prediction of N_2O emissions using ML models based on full-scale long-term data (Hwangbo et al., 2020, 2021; Vasilaki et al., 2020a)

This thesis investigates the capabilities of mechanistic models, specifically the two-pathway N_2O model based on Pocquet et al. (2016), to predict N_2O emissions from dynamically operated reactors, uncovering their limitations and challenges. With a focus on enhancing decision-making, it also evaluates the effectiveness of ML models in predicting N_2O emissions. This work offers a methodological framework for future N_2O emission modelling using ML, with insights that hold potential for broader application in the field of wastewater treatment.

1.2 Problem Statement and Research Motivation

The urgency to combat climate change has placed a spotlight on the water sector, especially on minimizing GHG emissions from WWTPs. Notably, N_2O emissions account for the most significant portion of the carbon footprint of WWTPs, emphasizing the need for accurate monitoring and understanding of the factors driving these emissions. Traditional monitoring campaigns, while essential, face challenges such as high costs and complex implementation.

Mathematical modelling presents a promising alternative to N_2O quantification by monitoring. Over the past decade, there have been significant efforts to develop mechanistic models, leveraging advancements in understanding N_2O production mechanisms. However, the field lacks consensus on a definitive mechanistic model for N_2O emission prediction. This is largely

due to the complexity of the biological pathways involved, the remaining uncertainties in the production mechanisms, and the over-parameterization of these models. Furthermore, these models' capability to accurately represent N_2O emissions from different combinations of operational conditions, influent types, and biomass types, has not been fully explored. In particular, dynamic systems with mixed biomass types, like flocculent sludge and biofilm, is yet to be thoroughly investigated. Such an investigation is critical to assess the full potential of N_2O mechanistic models.

Concurrently, WWTPs are generating large volumes of data through laboratory analyses and online sensors and analyzers. This data presents an opportunity for a data-driven approach, employing ML models to predict N_2O emissions. While a few studies have ventured into using ML for this purpose, significant challenges remain in their deployment and successful implementation. For instance, using these models in online monitoring and prediction requires attention not only to accuracy but also to interpretability of the results, which is vital for informed decision-making. Additionally, the complexity of a model often impacts its interpretability, and generally, a less complex model is preferred. This preference aligns with operational needs, as models requiring fewer features are more desirable due to the cost and complexities associated with collecting the data. Addressing these aspects necessitates further investigation and adaptation of the ML workflow to meet the specific requirements of wastewater applications, particularly in predicting N_2O emissions.

1.3 Research Objectives

The primary goal of this thesis is to advance the modelling of N_2O emissions from wastewater treatment processes, aiming to enhance decision-making capabilities. This involves exploration of traditional mechanistic models and a focused effort on improving the accuracy, simplicity, and interpretability of machine learning models to effectively serve this purpose. The main objective was approached through three specific objectives as per the following:

1) **To develop a validated N_2O two-pathways mechanistic model for dynamic operation:** Here, the focus is on assessing the N_2O two-pathways mechanistic model's ability to predict N_2O emission dynamics from an Integrated Fixed-film Activated Sludge Sequencing Batch Reactor (IFAS-SBR). This evaluation will serve as a critical test of the model's applicability and accuracy in real-world scenarios.
2) **To establish a comprehensive machine learning framework for N_2O emission prediction:** This objective focuses on creating an all-encompassing framework for N_2O emission prediction in full-scale WWTPs. It involves exploring various stages of ML project modelling, with an emphasis on balancing model performance against complexity and interpretability.
3) **To optimize the developed machine learning framework through enhanced applicability:** The final objective revolves around enhancing the previously developed ML framework. This will involve optimizing ML model complexity, particularly through focused hyperparameter optimization and strategic input feature selection, to improve the model's overall effectiveness and applicability.

1.4 Thesis outline

The thesis is organized into the following chapters:

Chapter 1 introduces the sustainability challenges faced by the wastewater sector, particularly in terms of GHG emissions. It outlines the research objectives and provides a structural overview of the thesis. Additionally, this chapter includes a concise review of previous modelling efforts for N_2O emissions in wastewater treatment processes, a statement of the problem, and the motivation behind the current research.

Chapter 2 provides a literature review on the production and emission of N_2O from biological wastewater treatment processes, and the current mechanistic and ML models used for N_2O emissions prediction.

Chapter 3 investigates objective 1 to examine the capabilities of current mechanistic models on prediction of the N_2O emissions from a laboratory scale Integrated fixed Film Activated Sludge Sequencing Batch Reactor (IFAS-SBR). The model was calibrated and validated using data representing different operational schemes of the IFAS-SBR. A multi-step calibration protocol, inclusive of sensitivity and identifiability analyses, is employed to refine the model. This chapter also explores scenario analyses using the calibrated model to derive optimal operational conditions for N_2O mitigation and efficient nitrogen removal.

Chapter 4 addresses objective 2 to develop a framework for ML models predicting N_2O emissions from full-scale WWTPs. A 16-month monitoring campaign dataset available in literature (Daelman et al., 2015), is utilized to develop the framework. Various ML models of different categories have been tested including decision trees, k nearest neighbors, random

forests, extreme gradient boosting, adaptive boosting, and deep neural networks. The developed framework focused on balancing the model performance, the complexity of the model, and interpretability of model results.

Chapter 5 details objective 3 to refine the ML framework developed in Chapter 4 by focusing on a coupled feature selection and model hyperparameter optimization. The developed algorithm uses multi-objective optimization using genetic algorithms.

Chapter 6 synthesizes the main findings from the preceding chapters, offering conclusions and drawing broader implications of the research. It also presents recommendations for future work and potential areas for further investigation in the field of N_2O emissions modelling in wastewater treatment.

Chapter 2. Literature Review

2.1 Biological nitrogen removal processes

2.1.1 Conventional nitrogen removal pathways

Biological nitrogen removal from wastewater is traditionally achieved through two primary processes: nitrification and denitrification. Nitrification, an aerobic process, is conducted by two distinct autotrophic groups of organisms: ammonia-oxidizing bacteria (AOB) such as *Nitrosomonas* and *Nitrosospira*, and nitrite-oxidizing bacteria (NOB) such as *Nitrobacter*, *Nitrospina*, and *Nitrospira*. AOB converts ammonia, which is the predominant nitrogen species in wastewater, into nitrite (NO_2^-) in a two-step reaction. Initially, ammonia is oxidized to hydroxylamine (NH_2OH) by the enzyme ammonia monooxygenase (AMO). This reaction is followed by the oxidation of NH_2OH to NO_2^-, catalyzed by hydroxylamine oxidoreductase (HAO). Subsequently, NOB oxidizes NO_2^- to nitrate (NO_3^-). Notably, the nitrification process consumes inorganic carbon (IC) and impacts alkalinity levels (Sin et al., 2008). In contrast, denitrification involves the reduction of NO_3^- to nitrogen gas (N_2) by ordinary heterotrophic organisms (OHO), such as species within the genera *Pseudomonas*, *Paracoccus*, and *Bacillus*. This multi-step process takes place under anoxic conditions and requires organic carbon as both the carbon source and electron donor.

2.1.2 Anammox and Comammox pathways

In addition to conventional nitrogen removal pathways, anammox (anaerobic ammonium oxidation) (van Loosdrecht and Jetten, 1998), and Comammox (complete ammonia oxidation) (Daims et al., 2015), were also discovered. Anammox allows for the direct conversion of NH_4^+

and NO_2^- into N_2 in the absence of oxygen, while Comammox possess the remarkable ability to oxidize NH_4^+ directly to NO_3^- in a single organism.

2.1.3 Shortcut nitrogen removal processes

The engineering of nitrogen removal processes, traditionally reliant on nitrification and denitrification, is characterized by substantial energy and resource demands (Zou et al., 2020). Nitrification necessitates oxygen and alkalinity, whereas denitrification requires an organic carbon source. This challenge can intensify when treating waste streams with high ammonia concentrations and low biodegradable organic carbon content, such as dewatered anaerobically digested sludge, where ammonia levels can exceed 1 g-N/L. Consequently, there is a pressing need to minimize carbon and oxygen consumption. The partial nitritation-denitritation process emerges as a viable solution, significantly reducing the environmental footprint of nitrogen removal.

Through the partial nitritation-denitritation approach, NOB are suppressed, enabling the oxidation of NH_4^+-N to NO_2^--N without progressing to NO_3^--N. This intermediate product, NO_2^--N, serves as a substrate for heterotrophic denitrification, directly converting to N_2, thereby bypassing a portion of the traditional process. Implementing this method can result in up to a 25% reduction in aeration costs and a 40% decrease in external carbon requirements compared to conventional nitrification-denitrification strategies (Daigger, 2014).

The effectiveness of partial nitritation-denitritation hinges on the selective inhibition of NOB and the encouragement of AOB growth. To achieve this, strategies such as maintaining low DO levels and high ammonium concentrations have been employed, exploiting the sensitivity of NOB to these conditions (Kirim et al., 2022). High ammonium concentrations, in particular, act

as a potent inhibitor of NOB (Shao et al., 2019). Additionally, the selective out-competition of NOB is facilitated at elevated temperatures, which disproportionately accelerate AOB growth rates. This differential growth dynamics permit the operation of systems under solids retention times (SRT) conducive to the exclusion of NOB, further optimizing the process efficiency.

2.2 Mechanisms of N_2O Transformation in Biological Nitrogen Removal Processes

2.2.1 Production and removal of N_2O through biological pathways

The production of nitrous oxide (N_2O) during nitrogen conversion processes can occur through mechanisms involving either AOB or OHO (Kampschreur et al., 2009). While N_2O is not an obligate product in ammonia oxidation, AOB is often the predominant source of N_2O, producing it through two main pathways: nitrifier nitrification (NN), and nitrifier denitrification (ND). N2O is also an intermediate in the heterotrophic denitrification process, where it can be both produced and consumed. Figures 2-1 shows the N_2O transformation processes through biological pathways during the conventional biological nitrogen removal, showing the possible intermediates and the enzymes catalyzing biological reactions.

In the NN pathway, also referred to as NH_2OH oxidation pathway, N_2O is produced as a byproduct of the incomplete oxidation of NH_2OH to NO_2^- in the presence of oxygen. However, the exact mechanism of the NN pathway is debated. One suggested mechanism is that N_2O can be produced from the chemical decomposition of NOH that is formed during NH_2OH oxidation, leading to N_2O production (Poughon et al., 2001). The other possible mechanism is the production of N_2O as a result of the reduction of NO, which itself is a product of NH_2OH oxidation (Stein, 2011). Further supporting this mechanism, Caranto and Lancaster (2017) found that NO acts as an obligate intermediate during ammonia oxidation. Consequently, this

implies that, in addition to AMO and HAO, there might be another enzyme involved in the ammonia oxidation reactions, contributing to the production of NO followed by N_2O.

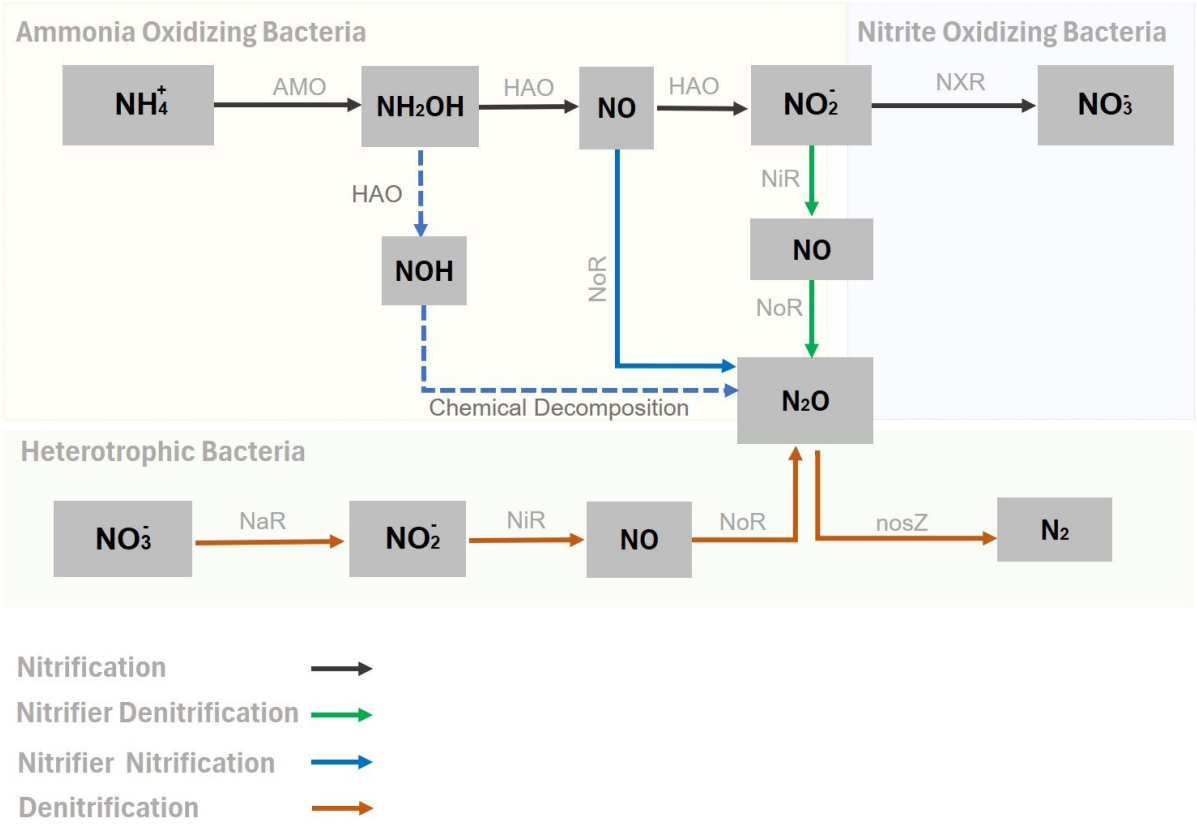


Figure 2-1 N_2O biological pathways

N_2O can also be produced through the ND pathway by reduction of NO_2 to nitric oxide (NO) through nitrite reductase (NirK) followed by its reduction to N_2O through nitric oxide reductase (NOR) (Mampaey et al., 2013). While these reactions occur together with the ammonia oxidation process that requires aerobic conditions, it was shown that the reaction rates increased under microaerobic conditions, however, unlike the denitrification process by OHOs, the ND pathway does not require organic carbon (Pijuan and Zhao, 2022). Although there are various suggested hypotheses about the mechanisms and conditions controlling the activation of the

ND pathway, whether a combination of these mechanisms or only one mechanism is required still remains an open question.

Looking deeper at the metabolic pathways of AOB unveiled the role of electron carriers to transport electrons from oxidation to reduction reactions (Ni et al., 2014). These carriers facilitate electron transfer within the oxidation-reduction reactions, oscillating between reduced (M_{red}) and oxidized (M_{ox}) states to sustain the energy flow necessary for metabolic processes. During the oxidation of NH_3 to NH_2OH, M_{red} acts as an electron donor to the oxygen atom, transitioning to its M_{ox} form in the process. Subsequently, the conversion of NH_2OH to NO_2^- involves a two-step oxidation where M_{ox} accepts four electrons, thereby reverting to M_{red}. NO serves as an intermediate in this series of reactions that is further reduced to N_2O via the NN pathway, where M_{red} donates an electron to NO, once again being oxidized to M_{ox}. Additionally, in the ND pathway, NO_2^- is sequentially reduced to NO and then to N_2O, with M_{red} contributing an electron to each step and being oxidized back to M_{ox}.

Despite the remaining uncertainties in the understanding of N_2O production mechanisms by AOB, there is agreement on the effect of dissolved oxygen (DO). The NN pathway is generally favored under high DO, and the ND pathway is predominant under limited DO conditions (Pijuan and Zhao, 2022). In addition to DO, the contribution of these pathways was shown to be influenced by the concentration of different available nitrogen species in the wastewater. For example, Wunderlin et al. (2013) found that ND was the dominant pathway for N_2O production by AOB during domestic wastewater treatment. However, under high ammonia and low NO_2^- concentration, the NN pathway was more relevant.

The production of N_2O by OHO can also take place during the heterotrophic denitrification reactions. Heterotrophic denitrification consists of four consecutive steps that involve reducing NO_3 to NO_2, then NO, N_2O, and finally N_2 (Zumft, 1997). N_2O is thus both produced and consumed by OHO during denitrification, whereas net production of N_2O can occur when the rate of production is higher than the rate of consumption (Tallec et al., 2008). This discrepancy can take place as a result of various operational and environmental disturbances including chemical oxygen demand (COD) loads and uptake rates, pH and temperature (Richardson et al., 2009). It is important to realize the only means to remove N_2O from the liquid phase is by heterotrophic denitrification and that new process configurations are under development to fully exploit this mechanism (Guo and Vanrolleghem, 2014).

Moreover, in a study by Pan et al. (2013a), the impact of electron competition among reduction of different nitrogen oxides was investigated during denitrification with methanol as a carbon source. The results revealed that the intensity of electron competition governed the N_2O accumulation during denitrification.

2.2.2 Production of N_2O through abiotic pathways

Beyond biological pathways, N_2O production also occurs through chemical pathways, including the chemical decomposition of NH_2OH at high pH, and NH_2OH oxidation by HNO_2, O_2, or Fe^{3+}, alongside HNO_2 reduction by Fe^{2+}. However, Su et al. (2019) found that abiotic N_2O production accounts for less than 3% of total N_2O output in conditions where pH ranges from 6.5 to 8, becoming significant only in acidic environments with a pH below 5.

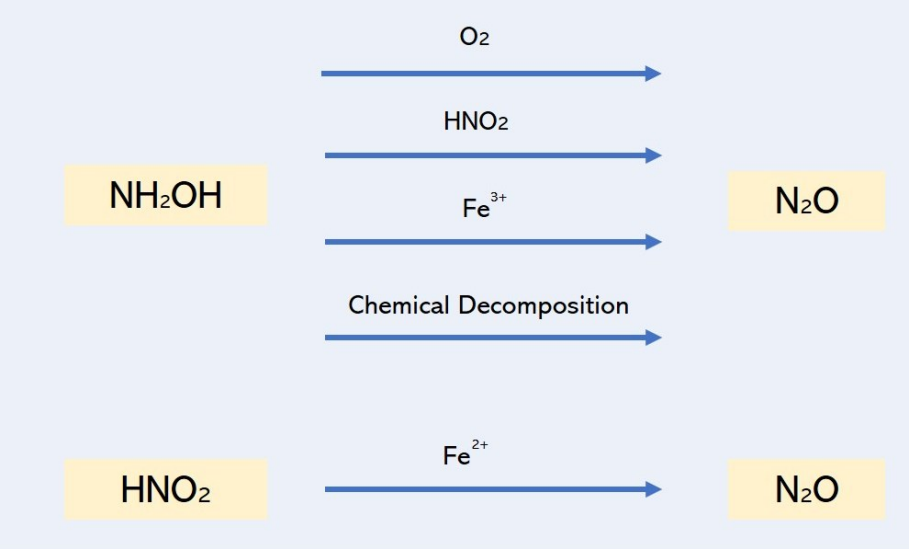


Figure 2-4 N_2O production via abiotic pathways

2.3 Estimation of N_2O Emissions by Emission Factors (EFs) According to IPCC

Emission factors (EFs) are essential tools for empirically estimating N_2O emissions, as outlined in the 2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories (Bartram et al., 2019). This method calculates N_2O emissions using a formula that incorporates both the emission factor and activity data:

$$\textit{Emission Rate (ER)} = \textit{Emission Factor (EF)} \times \textit{Activity Data (AD)}$$

EF and AD values are determined based on three tiers, reflecting data availability. Tier 1 uses default EFs provided by the IPCC, suitable for countries with limited data. Tier 2 involves country-specific EFs and ADs, while Tier 3 applies to facilities with capabilities for plant-specific monitoring, using EFs and ADs derived from plant-specific data. The IPCC's Tier 1 approach, being the only tier with default EFs, is predominantly used as a steady-state empirical model for GHG estimation (Bartram et al., 2019). EFs for N_2O are delineated based on

treatment technology types, including centralized aerobic treatment, anaerobic reactors, and various other technologies. Despite their utility, EFs' effectiveness is limited by their inability to account for the dynamic operational and environmental conditions affecting N_2O emission rates, such as dissolved oxygen (DO) (Kampschreur et al., 2009), and temperature (Adouani et al. 2015). Seasonal changes in temperature can result in temporal variations in N_2O emissions, increasing the need for long-term monitoring campaigns. Vasilaki et al. (2019) showed a relationship between the reported EF and the duration of the monitoring campaign, with higher EFs reported in longer monitoring campaigns. This high sensitivity of N_2O emissions to operational, seasonal and environmental conditions has also been extended to potential variations in N_2O emissions among parallel reactors operated in the same WWTP (Chen et al., 2019). The previously reported N_2O emissions have also demonstrated spatial variations, indicating that the location of a sampling point can influence the results of the measurement campaign (Duan et al., 2020; Gruber et al., 2021).

2.4 Mechanistic Modelling of N_2O Emissions

2.4.1 Introduction to mechanistic modelling

2.4.1.1 Activated Sludge Models (ASMs)

Mechanistic models are mathematical models that are formulated based on the underlying mechanistic understanding of the interactions occurring within treatment systems. Activated Sludge Models (ASMs), developed under the auspices of the International Water Association (IWA), represent a series of models that have evolved over the years (Henze et al., 2000). Beginning with ASM1, introduced in the late 1980s (Henze et al., 1987), these models have progressively incorporated more complex processes such as in ASM2, ASM2d, and ASM3.

The ASMs typically integrate transformation processes in terms of biomass growth, substrate utilization, and the formation of by-products in a mass-balance equation (Gujer, 2008). These biological conversions such as biomass growth and substrate utilization are modelled according to Monod kinetics, typically describing reaction rates using estimated kinetic and yield coefficients using stoichiometric parameters. ASMs model the biological nitrogen removal processes carried out by different microbial species such as OHOs, AOB, and NOB. The influence of environmental parameters like temperature, pH, and dissolved oxygen are also integrated in the models using empirical formulas and switching functions, affecting the rate of biochemical reactions (Henze et al., 2000). When applied to completely mixed reactors to model the changes of the system dynamics over time, this typically results in a set of ordinary differential equations (ODEs), in which they can be solved using one of the available numerical solvers integrated in simulators.

ASMs have been widely applied in the design, optimization, and control of wastewater treatment processes. They enable the simulation of various operational scenarios, and facilitate the prediction of effluent quality. Furthermore, these models aid in understanding the impacts of influent characteristics and operational strategies on the performance of biological treatment processes.

Advanced use of mechanistic models requires deep knowledge of the underlying process mechanisms by the modeller. For a model to be reliably used, calibration is performed to ensure agreement between measured and simulated data (Gillot et al., 2009). The knowledge and expertise of a modeller is thus needed to ensure that the parameters estimated through the calibration process are in an expected range. To further test the model capabilities, validation

is conducted using data other than data used for model calibration. A challenge to mechanistic models is when the model is not able to generalize when the operational and environmental conditions change, requiring another calibration conduct.

2.4.1.2 Biofilm models

A biofilm consists of a solid matrix of extracellular polymeric substances (EPS) that embed microorganisms with pore water that contains both dissolved substances and suspended solids (Morgenroth, 2008). The transport of substrates to microorganisms takes place through molecular diffusion through the biofilm depth, causing a mass transfer limitation. When the ASMs are applied to completely mixed reactors with flocculent biomass, they assume that substrates are locally available to the microorganism and that there is no gradient in concentrations, so Monod kinetics can be directly included into mass balance equations leading to a system of ODEs as mentioned earlier. However, in biofilms, the mass-transfer across biofilm depth leads to inevitable gradients of substrate concentrations across the biofilm depth, leading to the need of both spatially and temporally continuous models that are represented by a set of partial differential equations (PDEs) (Gujer, 2008). In order to predict the reactor performance, the biofilm model should thus be capable of determining the local availability of electron donors, acceptors, or inhibitory compounds as a result of mass transfer. It should also be able to predict the distribution of different types and species of microorganisms over the biofilm given the substrate availability. However, choosing the biofilm model requires a balance between the requirements of the modelling study and the model complexity. Complex models such multi-dimensional models are specifically used when the geometry of substratum is very complex or the modeller is interested to answer questions related to biofilm spatial

distribution (Morgenroth, 2008). However, one-dimensional (1-D) multi-substrate multi-species models can be sufficient in predicting substrate fluxes, multicomponent diffusion, and microbial competition to finally predict reactor performance. They only solve PDEs in the direction perpendicular to the substratum because it has the largest concentration gradients, while averaging values over areas parallel to the substratum.

The 1-D multi-substrate multi-species model was subject to several developments since the 1980s (Wanner and Gujer, 1984, 1986; Wanner and Reichert, 1996; Reichert and Wanner, 1997), where the last two papers were implemented in AQUASIM (Reichert et al., 1998). The biofilm reactor in AQUASIM is split into three zones: solid matrix, pore water, and bulk volume (Reichert et al., 1998). The model relies on four conservation law equations, including an equation for particulate species in the biofilm matrix, particulate species in the biofilm pore water, substances dissolved in the pore water, and the porosity of the biofilm. Traditional Monod kinetic expressions can be integrated into the biofilm model. Nevertheless, it is worth-noting that values of half saturation constants in biofilm models, which explicitly account for mass transfer limitations, are expected to be different from those derived from flocculent biomass experiments and often used in ASMs, which do not consider these limitations and may yield artificially high values (Morgenroth, 2008).

2.4.2 Reviewing N_2O mechanistic models

The existing N_2O mechanistic models that focus on N_2O production by AOB can be classified into single-pathway, and two-pathway models. Initially, only denitrification by OHO was the focus of the developed models (Hiatt and Grady, 2008; Corominas et al., 2012). Furthermore,

other models focused on incorporating the abiotic production of N_2O to have comprehensive models integrating the possible ways for N_2O production.

Single-pathway models are models that included either the NN or ND pathway solely. For the models focusing on the ND pathway, Ni et al. (2011) modelled the ammonia oxidation as two steps with NH_2OH as an intermediate and included DO inhibition function to both the N_2O production reactions. Another model developed by Mampaey et al. (2013) modelled the ammonia oxidation to NO_2 as a one-step direct reaction, excluding NH_2OH as an intermediate. The model also did not include inhibition of the ND pathway reactions due to high DO concentrations. A further modification was mde to this model by Guo and Vanrolleghem (2014) by the addition of a Haldane function to represent DO inhibition and limitation to NO_2 and NO reduction reactions.

Law et al. (2012) modelled the NN pathway as a two-step process with NOH as an intermediate of NH_2OH oxidation and N_2O is finally produced as a result of the chemical decomposition of NOH. In contrast, the model developed by Ni et al. (2013a, 2013b) described N_2O production as a result of NO reduction, which is produced from the oxidation of NH_2OH. This model also assumed that NH_2OH is the electron donor in the NO reduction reaction. The reaction schemes used in single-pathway models describing both NN and ND pathways are presented and compared in figure 2 – 5.

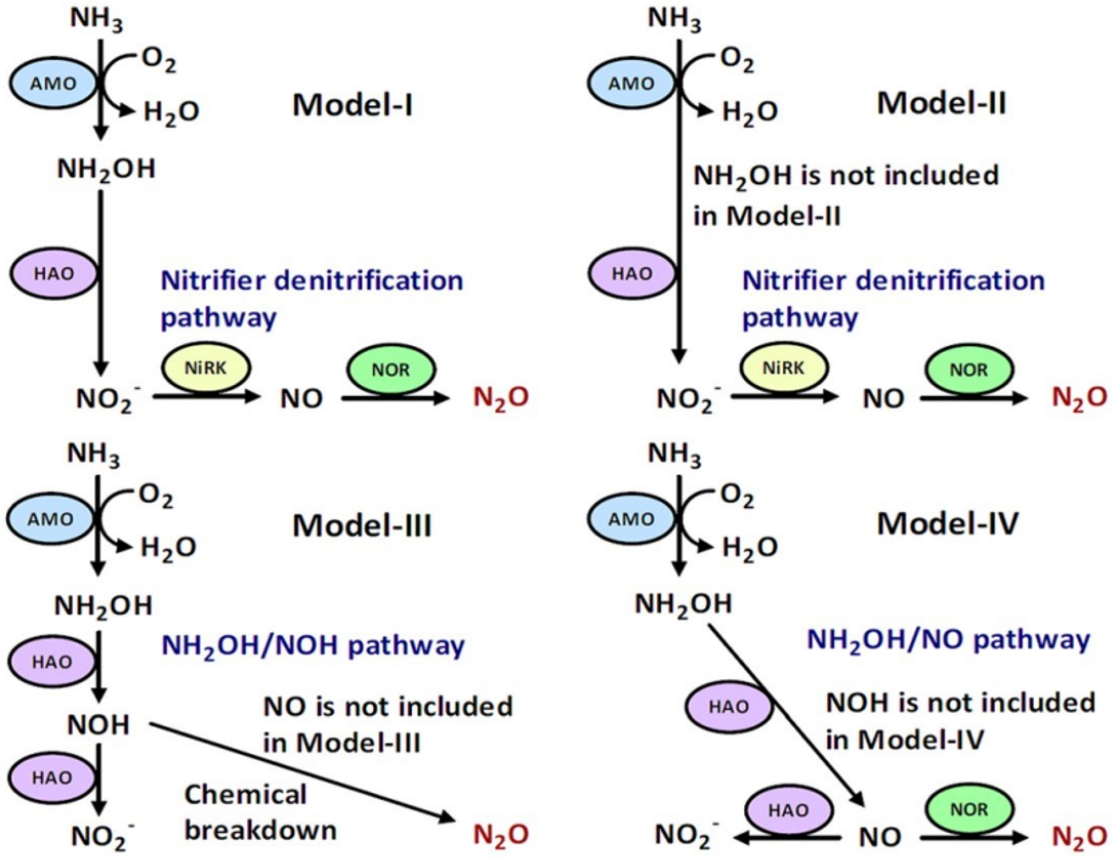


Figure 2-5 Comparison of the reaction schemes used in single-pathway models (Ni et al., 2013b). *Model-I: (Ni et al., 2011), Model-II: Mampaey et al. (2013), Model-III:* (Law et al., 2012*), Model-IV: Ni et al. (2013a, 2013b)*

Two-pathway models integrate both the NN and ND pathways in a single model. A model developed by Ni et al. (2014) decoupled all oxidation-reduction reactions catalyzed by AMO, HAO, NOR, and Nirk, respectively. Electron carriers were introduced as an explicit variable in the model to bridge electron transfer and effectively decoupling the reactions. Another model was developed based on an extension of this model including Adenosine triphosphate (ATP)/Adenosine diphosphate (ADP)-based energy balance (Peng et al., 2015). The decoupling

approach further increases the model structure's complexity and limits its applicability. Thus, oxidation and reduction were coupled and another two-pathway model was proposed by Pocquet et al. (2016) that describes both the NO and N_2O production. The NN pathway was represented in this model as two aerobic processes involving oxidation of NH_2OH to NO, followed by NO to N_2O. The representation of the ND pathway in this model excluded NO as an intermediate to avoid its cycling between the two pathways, and NO_2^- was directly reduced to N_2O in this model. The indirect coupling (Ni et al., 2014; Peng et al., 2015), and direct coupling (Pocquet et al., 2016) modelling approaches were calibrated with experimental N_2O emissions data and compared by Lang et al. (2017). Both approaches were able to adaptively predict N_2O production. Moreover, both approaches could predict the impact of a DO increase on the contribution of NN and ND pathways. However, the direct coupling modelling approach is less complex and more understandable to modelling practitioners thanks to its ASM-based approach. Furthermore, this model considers NO as a state variable allowing for model calibration using NO data, if available.

Unlike the models describing N_2O production by AOB, there is a relatively general agreement on modelling N_2O through the denitrification pathways. The Activated Sludge Model for Nitrogen (ASMN) describes denitrification as four-step independent oxidation-reduction reactions using different kinetics in each step (Hiatt and Grady, 2008). This model can predict the potential N_2O net production due to variable kinetic parameters of each denitrification step. Furthermore, a model by Pan et al. (2013b) named as ASM-indirect coupling of electrons (ICE), used the decoupling approach to model the denitrification kinetics and introducing electron carriers as variables.

Despite the demonstrated insignificance of abiotic pathways at neutral pH (Su et al. 2019), models incorporating the abiotic production of N_2O were also proposed. Harper et al. (2015) proposed a model that describes N-nitrosation using a second order chemical reaction that is function of NH_2OH and NO_2 concentrations. A comprehensive model was then developed to integrate abiotic pathways with biological pathways, named nitrifier nitrification, nitrifier denitrification, heterotrophic denitrification and abiotic pathways (NDHA) (Domingo-Félez and Smets, 2016).

Although there have been significant developments in the mechanistic modelling of N_2O emission, the application of these models suffers from several limitations, specifically regarding their calibration and their ability to depict N_2O dynamics in long-term full-scale applications. Similarly, the calibration of mechanistic N_2O models was also shown to be complicated and time-consuming process due to the large number of parameters included in these models, making them overparameterized with respect to the limited information content of the available data. The overparameterization of these models is evident from the large variation in the values of estimated parameters across different studies (Domingo-Félez and Smets, 2016). Such complicated calibration can be a barrier to the practical application of N_2O mechanistic models, necessitating the need to find an alternative for such applications. In contrast, the knowledge-based nature of mechanistic models allows their utilization for testing hypotheses and scenario analyses. Further exploration of the potential of mechanistic models is thus still required.

2.4.3 Application of N_2O models in biofilm systems

Biofilm models for N_2O applications are less investigated in literature (Nopens et al., 2022). Although the same rate expressions of the discussed N_2O models can be applied to biofilm

treatment processes, biofilm models are generally associated with higher complexity. Diffusion limitations and biofilm stratification are the main factors that can change the behaviour of N_2O production and emission from suspended growth processes (Spérandio et al., 2022). Only a few studies utilized the two pathway AOB model integrated with ASMN for denitrification to describe N_2O emissions from nitrifying biofilters, denitrifying biofilters, and granular sludge reactors using 1-D biofilm models (Fiat et al., 2019; Lang et al., 2019; Zhu et al., 2019). In the reported N_2O emission results by Zhu et al. (2019), the model was able to provide EFs in winter and summer that are in agreement with experimental observations. Moreover, both Zhu et al. (2019) and Fiat et al. (2019) demonstrated that a mass balance on the gaseous phase was needed when modelling a nitrifying biofilter. Other studies focused on understanding the complex nature of biofilms and how it impacts N_2O production and emission (Schreiber et al., 2009; Sabba et al., 2015). In the study by Sabba et al. (2015), the metabolic model developed by Ni et al. (2014) was adapted to explore N_2O emissions from a granular sludge reactor for sidestream nitritation. The results demonstrated that N_2O emissions from nitrifying biofilms could be larger than from suspended growth systems under similar conditions due to the formation and diffusion of NH_2OH. The stratification that exists in biofilms can result in production of NH_2OH by AOB in the outer, aerobic zone and consumption in the inner, anoxic zone, leading to peaks in N_2O production.

2.5 Machine Learning (ML) Modelling

Machine learning (ML) – a critical component of artificial intelligence (AI) - is a data-driven approach where a computer, in the form of a model or algorithm, learns hidden and complex patterns from data without prior requirements of domain knowledge (Mitchell, 1997). These

models can then apply this knowledge to make informed predictions or decisions. Central to ML models is the training process, which involves a dataset comprising numerous data points, typically structured in a table format with rows representing records and columns representing features.

2.5.1 Introduction to ML models

ML models are typically categorized based on how they learn from data into supervised, unsupervised, and reinforcement learning (Géron, 2019). In supervised learning, models are trained on labeled datasets where the outcome for each input is known, making them suitable for predictive tasks like regression and classification. Unsupervised learning models, on the other hand, work with unlabeled data and are adept at uncovering hidden patterns or structures, finding use in clustering and dimensionality reduction. Reinforcement learning is still relatively distinct, involving models that learn to make decisions through trial and error, rewarded or penalized for their actions. It is commonly employed in robotics and navigation.

Among the categories of ML models, supervised learning is the most prevalent and relevant for wastewater treatment predictive models (Zhu et al., 2023). In particular, regression models can be used for quantitative prediction of the value of a variable of interest. For regression-based supervised ML, a variety of models are available varying in their ability to uncover complex and nonlinear patterns in the data to make accurate predictions. Below, some of the relevant models in regression problems are briefly reviewed.

2.5.1.1 Decision trees

Decision trees are widely used for both classification and regression problems. In the context of regression, decision trees can predict continuous values by learning decision rules inferred

from the data features (Breiman, 1984). The core idea is to recursively partition the data space into subsets, where each node in the tree represents a decision rule that splits the data based on feature values. This splitting continues until a stopping criterion is met, typically when further splitting does not significantly reduce prediction error or when nodes contain only a few observations (Quinlan, 1986). The final model is a tree where each leaf represents a predicted value, often the mean of the target variable for the observations in that leaf. One of the key strengths of decision trees is their interpretability; the hierarchical structure of decisions provides a clear and intuitive representation of how the input features affect the predicted outcome (Rokach and Maimon, 2008).

2.5.1.2 k-nearest neighbours (kNN)

The kNN algorithm is a versatile and intuitive method used in machine learning for both classification and regression tasks. In regression problems, kNN predicts the value of a new observation based on the values of its k nearest neighbors in the training set. The algorithm identifies the k closest points to the new observation in the feature space, typically using a distance metric such as Euclidean or Manhattan distance, and then computes the output as the average, or the weighted average of the target values of these neighbors (Altman, 1992). This non-parametric and instance-based approach allows kNN to adapt flexibly to the data without assuming a specific functional form for the relationship between features and the target variable.

2.5.1.3 Ensemble models

Ensemble models in regression represent a sophisticated and highly effective approach to predictive modelling in machine learning. These models combine multiple individual models, typically decision trees, to create a more accurate and robust predictive model than any single

constituent model could provide (Dietterich, 2000). The underlying principle of ensemble methods is that by aggregating the predictions of several models, the errors of individual models are likely to cancel out, reducing the overall prediction error (Maclin and Opitz, 2011). Common ensemble techniques include Bagging, Random Forests, and Boosting. Bagging (Bootstrap Aggregating) involves creating multiple models each trained on a different bootstrap sample of the data and averaging their predictions. Random Forests extend this idea by introducing randomness in the tree-building process to create a diverse set of trees (Breiman, 2001). Boosting, on the other hand, sequentially builds models, each one focusing on the errors of its predecessor, to iteratively improve model performance (Freund and Schapire, 1997). These ensemble methods have been shown to significantly enhance prediction accuracy in regression tasks, especially in complex datasets with nonlinear relationships and interactions among variables.

2.5.1.4 Deep learning models

Neural networks, especially deep neural networks (DNNs), have emerged as a powerful tool for regression analysis in complex and high-dimensional data spaces. Deep neural networks consist of multiple layers of interconnected nodes or neurons, where each layer performs specific transformations of its inputs, thereby capturing complex, non-linear relationships in the data (LeCun et al., 2015). The depth of these networks – the number of hidden layers – enables the learning of hierarchical feature representations, making them exceptionally adept at modelling complex patterns. In regression tasks, the final layer of a DNN is typically designed to output continuous values. The training of DNNs involves optimizing a large number of parameters (weights and biases) using algorithms like backpropagation combined with gradient

descent, which adjusts these parameters to minimize prediction errors (Rumelhart et al., 1986). While DNNs can offer exceptional modelling capabilities, they require substantial data and computational resources, and their 'black box' nature can make interpretability a challenge.

2.5.2 N_2O modelling with ML

The exploration of supervised ML models in N_2O emissions prediction has been relatively limited, including both qualitative assessments using classification models and quantitative evaluations using regression models. For instance, Vasilaki et al. (2020b) combined classification models with unsupervised learning techniques to optimize the frequency of N_2O off-gas sampling. By using support vector machines (SVM) and random forest methods, the study categorized emission levels in different zones, demonstrating the effectiveness of ML in planning monitoring campaigns for emission factor (EF) estimation and suggesting that even with reduced monitoring frequency, accurate annual EF estimations could be feasible. Another study by Vasilaki et al. (2020a) employed both classification and regression ML models alongside abnormal events detection to predict dissolved N_2O concentrations in various phases of a full-scale sequencing batch reactor (SBR) treating anaerobically digested supernatant. The study's "knowledge-discovery framework" addressed the temporal and spatial challenges of SBR operation. The framework initiated with an SVM classifier to predict N_2O consumption during the anoxic phase, followed by another SVM classifier for the anaerobic phase. Based on these classifications, SVM regressors then predicted N_2O concentrations in the aerobic phase. The models exhibited high accuracy, but the complexity of the framework might pose a barrier to widespread adoption. Further, Hwangbo et al. (2020) utilized N_2O emission data from a full-scale activated sludge reactor to train supervised ML regression models. The study highlighted

the superiority of deep neural networks (DNN) over other ML models, claiming that other non-deep learning models would fail to accurately predict N_2O emissions. Subsequent work by Hwangbo et al. (2021) reinforced the superiority of deep learning for complex applications like N_2O emissions prediction, demonstrating the use of long-short term memory (LSTM) models for improved forecasting accuracy. Lastly, Szeląg et al. (2023) developed an algorithm to choose the best-performing ML model for predicting N_2O emissions using data generated from a mechanistic model that is calibrated on a short 4-day measurement campaign. Models included SVM, extreme gradient boosting (XGBoost), and multivariate adaptive regression spline (MARS).

Although the main focus is predictive modelling of N_2O emissions, it is worth noting that unsupervised ML has also been briefly employed in previous studies, aiming to understand N_2O emissions. For instance, methods like clustering, principal component analysis (PCA), and change-point analysis have been used in studies by Vasilaki et al. (2018, 2020b) and Bellandi et al. (2020) to gain insights into N_2O emissions data. In another study, Song et al. (2020) utilized a random forest model to identify key features influencing N_2O emissions.

In addition to ML models, another AI system was used to predict the risk of N_2O emissions. This method, as discussed by Porro et al. (2022), leverages a knowledge-based AI framework augmented by rule-based fuzzy logic. The model incorporates the knowledge from literature about factors and conditions known to influence various N_2O pathways to provide risk assessment of producing or emitting N_2O in the form of a qualitative diagnosis (i.e., low, medium, or high risk). This qualitative risk assessment can be provided for each influencing factor, allowing for insights into active pathways and their mitigation. This approach was

applied to a full-scale WWTP and mitigation action was successfully implemented to reduce N_2O emissions based on the risk due to both low-DO and high-DO conditions, assuming that they correspond to ND, and NN pathways, respectively.

One notable distinction of this knowledge-based system, in comparison to ML models, lies in its reliance on predefined knowledge and rules. While ML models excel at identifying non-linear patterns and relationships within data sets, the knowledge-based system with fuzzy logic operates within the confines of its programmed knowledge base. As a result, although it provides risk assessments based on current understanding, it lacks the capacity to autonomously uncover new insights beyond the scope of its embedded knowledge.

Chapter 3 - Modelling nitrous oxide emissions from nitritation-denitritation IFAS-SBR treating sidestream wastewater *

3.1 Introduction

Wastewater streams that typically contain high levels of nitrogen, with concentrations that may exceed 1 g-N/L, such as anaerobically digested wastewaters, poses a threat to marine life and conflicts with stringent water quality standards (Smith, 2003). Nitritation-denitritation has been recognized as an effective alternative to conventional nitrification-denitrification treatments, which not only aligns with resource conservation efforts by reducing the need for added carbon sources but also saves energy by minimizing the aeration required, offering an integrated solution for nitrogen management in wastewater treatment (Zou et al., 2020). An Integrated Fixed-Film Activated Sludge Sequencing Batch Reactors (IFAS-SBR) can be utilized to improve nitrogen removal efficiency with a minimal reactor volume (Yang et al., 2020). The IFAS technology utilizes biomass carriers to an activated sludge system, thus combining attached and suspended biomass in a single system, referred to as "biofilm", and "flocculent sludge", respectively, and resulting in enhanced biomass retention capabilities, shock loads adaptability, and high contaminants removal. Biofilms are composed from a dense matrix of bacteria glued together by extracellular polymeric substances (EPS), and attach to surfaces (substratum).

---

* A modified version of this chapter has been submitted for journal publication in January, 2024.

In contrast, flocculent sludge consists of loose aggregates or flocs of microorganisms, are suspended in wastewater.

The SBR modes are flexible for various operational conditions such as control of phase durations, and easiness of analyzer installations. However, an exploration of the rate of production and emission of N_2O and how it is related to the choice of operational conditions is required (Liu et al., 2022). Therefore, it is imperative to develop operational strategies that not only optimize nitrogen removal efficiency but also minimize N_2O emissions from IFAS-SBR.

Although the selection and optimization of operational conditions could be conducted using laboratory experiments (Zou et al., 2022), using mechanistic modelling could save time and cost by simulating the effect of these operational conditions on N_2O emissions. However, the application of N_2O mechanistic models to IFAS-SBR reactors performing nitritation/denitritation under dynamic conditions and feature hybrid biomass, has not been explored yet. There is a need evaluate the capabilities of current mechanistic models to describe N_2O emissions from different reactor types and investigate their capabilities for informing operational process improvements to reduce N_2O emissions.

In this chapter, a dynamic model incorporating both AOB N_2O production pathways and heterotrophic denitrification is applied to simulate data from a lab-scale IFAS-SBR reactor performing nitritation/denitritation to treat ammonia-rich wastewater. The model is calibrated and validated on a different operational condition to investigate the generalization ability of the current N_2O mechanistic model to a wider range of operational conditions. The calibrated model is then used to compare the contribution of different pathways and the different types of biomasses (i.e., biofilm and suspended sludge) to production of N_2O. Lastly, a scenario analysis

is conducted to investigate the reactor behaviour in terms of nutrients removal and N_2O emissions under various combinations of operational and environmental conditions. Thus, the model could support a multi-objective optimization that aims to mitigate N_2O emissions while maximizing nitrogen removal efficiency by IFAS-SBR, using the currently available knowledge in mechanistic models by the aid of simulations.

3.2 Approach and methodology

3.2.1 Reactor operation and dataset

The experimental study utilized a 6L bench-scale IFAS-SBR, with 40% of the reactor volume filled with polyethylene carriers to support biofilm growth. The reactor maintained a constant temperature of 20°C and a 50% exchange ratio throughout the study. The experiment was conducted in two distinct stages, characterized primarily by variations in the frequency of the aerobic/anoxic phase alternation, thereby affecting the nitritation/denitritation dynamics. Specifically, the first stage involved operating the reactor with four subcycles, while the second stage increased this number to seven subcycles. This change resulted in different total cycle times and altered durations for both aerobic and anoxic phases to accommodate the increased number of subcycles. The total cycle time was 21.60 hours, while the duration of each aerobic phase was 3.50 hours during four subcycles operation, and 2.00 hours during seven subcycles operation. Moreover, each anoxic phase duration was 1.90 and 1.08 hours during four and seven subcycles operation, respectively. Throughout the aerated periods, the dissolved oxygen (DO) concentration was maintained at 0.6 ± 0.1 mg/L using a gas flowmeter. A concentrated chemical oxygen demand (COD) stock solution (sodium acetate anhydrous) at a concentration of 42.5 ± 2.1 g/L was continuously supplied as an external carbon source during the non-aerated periods

to support anoxic heterotrophic reactions. The timing of COD addition was precisely matched to the duration of the non-aerated periods using a timer, ensuring optimal denitritation conditions.

The ammonia-rich wastewater used in this study originated from the supernatant collected from anaerobically digested primary and secondary sludge obtained from a sludge settling lagoon in a local WWTP situated in Alberta, Canada, with ammonium (NH_4^+) concentrations of 800-1000 mg N/L, 400–600 mg/L chemical oxygen demand (COD), ~ 7.8 pH, 1–3 mg N/L NO_3^-, and 4–6 mg N/L NO_2^-. A detailed description of the experiment and wastewater characteristics can be found in Zou et al. (2022).

3.2.2 Model structure

A dynamic model was developed to simulate the IFAS-SBR implemented in the software package AQUASIM 2.1d (Reichert, 1998). The model structure comprised of a fully mixed compartment representing the flocculent sludge and a biofilm compartment representing the biofilm attached on the polyethylene carriers (substratum). A diffusion link with a high mass transfer coefficient connected the two compartments, ensuring that they shared the same bulk liquid concentrations for all state variables except for N_2O that was excluded from the diffused state variables to accurately quantify N_2O produced from the biofilm and flocculent sludge separately.

The simulation of the SBR cycle involved dividing the total volume of the reactor (6.0 L) into a 3.0 L biofilm compartment and a variable volume completely mixed compartment with volume ranging from 0 to 3.0 L while maintaining a 50% exchange ratio of the influent. The SBR cycle encompassed three main phases: feeding, treatment, and decanting. During the

feeding phase, only inflow occurred, expanding the model's variable volume until reaching its maximum value. After feeding, the inflow to the reactor ceased, and the reactor volume remained constant during the treatment period. The decanting phase involved only outflow from the reactor with the same flowrate as the influent, leading to volume decrease until reaching its minimum value. During decanting, solids were preserved in the reactor model using bifurcations, i.e. all solids are returned back to the reactor except for the wasted sludge to maintain the experimental SRT. The treatment phase consisted of a series of aerobic-anoxic alternating subcycles. Each subcycle was implemented by simulating the aerobic and anoxic conditions through turning on and off the air supply to the reactor. For the mathematical representation of oxygen supply, a fully-mixed gas phase reactor was introduced and connected to the SBR through a diffusive link with an on/off time-dependent pattern to simulate the alternating aeration. Figure 3.1 shows the representation of IFAS-SBR in AQUASIM including the used compartments and links.

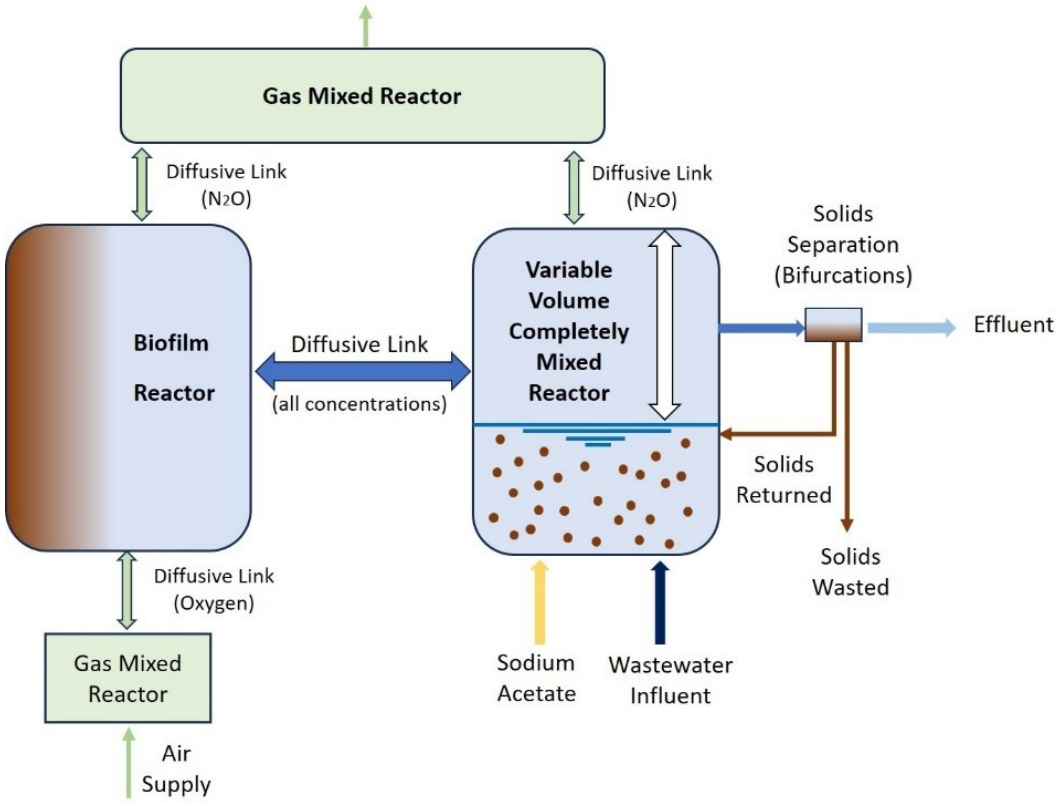


Figure 3-1 Schematic diagram of IFAS-SBR representation in AQUASIM.

3.2.3 Biological conversion and N_2O model

The model included the biological conversion reactions by three species: AOB, nitrite-oxidizing bacteria (NOB), and heterotrophs. Ammonium oxidation processes and N_2O production pathways was modelled according to Pocquet et al. (2016) including the following reactions: (1) ammonia mono-oxygenase (AMO) mediated oxidation: where NH_4^+ is oxidized to NH_2OH in the presence of oxygen; (2) HAO mediated oxidation: NH_2OH is further oxidized to NO; (3) another HAO mediated oxidation: NO is oxidized to NO_2^-; (4) NN pathway: This pathway involves the production of N_2O through the reduction of NO, mediated by NOR, coupled with the oxidation of NH_2OH to NO_2^- (reaction 2); (5) ND pathway: The reduction of

NO_2^- to N_2O occurs with NO as an intermediate step, utilizing NH_2OH as an electron donor. This reaction is mediated by NOR and NirK. Reactions 1, 2, and 3 require the presence of oxygen to proceed. Although reaction 4 does not explicitly require oxygen, it is dependent on the availability of oxygen as it is coupled with reaction 2. Therefore, reaction 4 cannot occur in the absence of oxygen. Reaction 5, on the other hand, is the only reaction performed by AOB that does not require the presence of oxygen. It is defined as an anoxic process inhibited by oxygen. The inhibition of reaction 5 by oxygen was incorporated into the model using a non-competitive term, as described in Ni et al. (2011). To simplify the model representation and avoid the inclusion of an intermediate step involving NO, reaction 5 was lumped into a single reaction that directly reduces NO_2^- to N_2O. This approach effectively bypasses the NO loop in the model (Pocquet et al., 2016). The oxidation of NO_2^- to NO_3 by nitrite-oxidizing bacteria (NOB) was modelled according to (Hiatt and Grady, 2008), considering inhibitory effects of NO_2^- and NH_4^+. By considering these five reactions, the developed mathematical model is able to capture the key nitrogen transformation processes occurring within the IFAS-SBR flocculent sludge and biofilm. Furthermore, heterotrophic biomass included both aerobic growth on organics and anoxic heterotrophic denitrification. The denitrification process was modelled as four sequential steps reducing the NO_3^- to N_2 gas with NO_2^-, NO, N_2O as intermediates (Hiatt and Grady, 2008). Total ammonium and total nitrite were used as the substrates in the kinetic rate expressions and half-saturation constants were modified accordingly, following the approach adopted by Wan and Volcke (2022). The list of model state variables is included in Table 3.1, and the stoichiometric matrix is shown in Table 3.2, and the kinetic rate expressions used in the model are shown in Table 3.3. The decay of all species in the model was described

as a first order reaction producing both biodegradable and non-biodegradable particulates. The biodegradable particulates are then hydrolyzed to soluble organics. The hydrolysis process was modelled as a simplified rate as suggested by Lackner et al. (2008).

Table 3-1 Description of model state variables

Variable	Description	Unit
X_{AOB}	Ammonia oxidizing bacteria biomass concentration	g COD. m^{-3}
X_H	Heterotrophic biomass concentration	g COD. m^{-3}
X_{NOB}	Nitrite oxidizing bacteria biomass concentration	g COD. m^{-3}
X_I	Inert particulate organics	g COD. m^{-3}
X_s	Slowly biodegradable organics concentration	g COD. m^{-3}
S_{NH}	Total ammonia nitrogen concentration	g $N.m^{-3}$
S_{NO2}	Total nitrite nitrogen concentration	g $N.m^{-3}$
S_{NO3}	Nitrate concentration	g $N.m^{-3}$
S_{NH_2OH}	Hydroxylamine concentration	g $N.m^{-3}$
S_{NO}	Nitric oxide concentration	g $N.m^{-3}$
S_{N2}	Nitrogen gas concentration	g $N.m^{-3}$
S_O	Dissolved oxygen concentration	g $O_2.m^{-3}$
S_S	Readily biodegradable organics concentration	g COD. m^{-3}
S_{N2O}	Nitrous oxide concentration	g $N.m^{-3}$
S_{N2O_NN}	Nitrous oxide concentration produced from nitrifier nitrification	g $N.m^{-3}$
S_{N2O_ND}	Nitrous oxide concentration produced from nitrifier denitrification	g $N.m^{-3}$

Table 3-2 Process stoichiometric matrix of the model variables used in this chapter

Process	S_{NH}	S_{NH_2OH}	S_{N2}	S_{N2O}	S_{NO}	S_{NO2}	S_{NO3}	S_O	S_S	S_{N2O_NN}	S_{N2O_ND}	X_{AOB}	X_{NOB}	X_{ND}	X_S	X_H	X_I
AOB																	
R1	-1	1						-1.14									
R2	$-i_{NXB}$	$\frac{-1}{Y_{AOB}}$			$\frac{1}{Y_{AOB}}$			$1-\frac{1.71}{Y_{AOB}}$				1					
R3					-1	1		-0.57									
R4		-1		2		-1					2						
R5		-1		4	-4	1				4							
R6												-1		$i_{NXB}-f_{XI}i_{NXI}$	$1-f_{XI}$		f_{XI}
NOB																	
R7	$-i_{NXB}$					$\frac{-1}{Y_{NOB}}$	$\frac{1}{Y_{NOB}}$	$1-\frac{1.14}{Y_{NOB}}$					1				
R8													-1	$i_{NXB}-f_{XI}i_{NXI}$	$1-f_{XI}$		f_{XI}
Heterotrophs																	
R9	$\frac{i_{NSS}}{Y_H}-i_{NXB}$							$1-\frac{1}{Y_H}$	$-\frac{1}{Y_H}$							1	
R10	$\frac{i_{NSS}}{Y_H\eta_Y}-i_{NXB}$					$\frac{1-Y_H\eta_Y}{1.14Y_H\eta_Y}$	$\frac{-1-Y_H\eta_Y}{1.14Y_H\eta_Y}$		$\frac{-1}{Y_H\eta_Y}$							1	
R11	$\frac{i_{NSS}}{Y_H\eta_Y}-i_{NXB}$				$\frac{1-Y_H\eta_Y}{0.57Y_H\eta_Y}$	$-\frac{1-Y_H\eta_Y}{0.57Y_H\eta_Y}$			$\frac{-1}{Y_H\eta_Y}$							1	

R12	$\frac{i_{NSS}}{Y_H\eta_Y} - i_{NXB}$			$\frac{1-Y_H\eta_Y}{0.57Y_H\eta_Y}$	$-\frac{1-Y_H\eta_Y}{0.57Y_H\eta_Y}$				$\frac{-1}{Y_H\eta_Y}$							1	
R13	$\frac{i_{NSS}}{Y_H\eta_Y} - i_{NXB}$		$\frac{1-Y_H\eta_Y}{0.57Y_H\eta_Y}$	$-\frac{1-Y_H\eta_Y}{0.57Y_H\eta_Y}$					$\frac{-1}{Y_H\eta_Y}$							1	
R14												$i_{NXB} - f_{XI}i_{NXI}$	$1-f_{XI}$	-1	f_{XI}		
Hydrolysis																	
R15									1					-1			
R16	1												-1				
g COD/unit component	0	-1.14	-1.71	-2.29	-2.86	-3.43	-4.57	-1	1	-2.29	-2.29	1	1	1	1	1	1
g N/unit component	1	1	1	1	1	1	1	0	i_{NSS}	1	1	i_{NXB}	i_{NXB}	1	i_{NXB}	i_{NXB}	i_{NXB}

Table 3-3 Process rate expressions.

Process		Kinetic rate expressions
AOB		
R1	Ammonia oxidation (AMO)	$q_{max,AOB}\frac{S_O}{K_{AOB,1,O}+S_O}\frac{S_{NH}}{K_{AOB,NH}+S_{NH}}X_{AOB}$
R2	Hydroxylamine oxidation (HAO)	$\mu_{AOB}\frac{S_O}{K_{AOB,2,O}+S_O}\frac{S_{NH2OH}}{K_{AOB,NH2OH}+S_{NH2OH}}\frac{S_{NH}}{S_{NH}+10^{-12}}X_{AOB}$
R3	Nitrite oxidation (HAO)	$q_{max,AOB}\frac{S_O}{K_{AOB,2,O}+S_O}\frac{S_{NO}}{K_{AOB,HAO,NO}+S_{NO}}X_{AOB}$
R4	N_2O production (ND pathway)	$q_{max,AOB}\,\eta_{ND}\frac{S_{NO2}}{K_{AOB,NO2}+S_{NO2}}\frac{S_{NH2OH}}{K_{AOB,NH2OH}+S_{NH2OH}}\frac{K_{AOB,I,O}}{K_{AOB,I,O}+S_O}X_{AOB}$
R5	N_2O production (NN pathway)	$q_{AOB,,NN}\,\eta_{NN}\frac{S_{NO}}{K_{AOB,NN,NO}+S_{NO}}\frac{S_{NH2OH}}{K_{AOB,NH2OH}+S_{NH2OH}}X_{AOB}$
R6	Decay	$b_{AOB}X_{AOB}$
NOB		
R7	Nitrite oxidation	$\mu_{NOB}\frac{S_O}{K_{NOB,O}+S_O}\frac{K_{NOBH,NH}}{K_{NOBH,NH}+S_{NH}}\frac{S_{NO2}}{K_{NOB,NO2}+S_{NO2}+S_{NO2}^{\,2}/K_{NOB,I,NO2}}X_{NOB}$
R8	Decay	$b_{NOB}X_{NOB}$
Heterotrophs		
R9	Aerobic growth	$\mu_H\frac{S_O}{K_{H,O}+S_O}\frac{S_{NH}}{K_{H,NH}+S_{NH}}\frac{S_S}{K_{H,S}+S_S}X_H$
R10	Nitrate denitrification	$\mu_H\eta_{g2}\frac{S_S}{K_{H,S,1}+S_S}\frac{S_{NO3}}{K_{H,NO3}+S_{NO3}}\frac{K_{H,O}}{K_{H,O}+S_O}\frac{S_{NH}}{K_{H,NH}+S_{NH}}X_H$
R11	Nitrite denitrification	$\mu_H\eta_{g3}\frac{S_S}{K_{H,S,2}+S_S}\frac{S_{NO2}}{K_{H,NO2}+S_{NO2}}\frac{K_{H,I,O,2}}{K_{H,I,O,2}+S_O}\frac{K_{H,I3,NO}}{K_{H,I3,NO}+S_{NO}}\frac{S_{NH}}{K_{H,NH}+S_{NH}}X_H$

R12	Nitric oxide denitrification	$\mu_H \eta_{g4} \frac{S_S}{K_{H,S,3} + S_S} \frac{S_{NO}}{K_{H,NO} + S_{NO} + {S_{NO}}^2 / K_{H,I4,NO}} \frac{K_{H,I,O,3}}{K_{H,I,O,3} + S_O} \frac{S_{NH}}{S_{NH} + K_{H,NH}} X_H$
R13	Nitrous oxide denitrification	$\mu_H \eta_{g5} \frac{S_S}{K_{H,S,4} + S_S} \frac{S_{N2O}}{K_{H,N2O} + S_{N2O}} \frac{K_{H,I,O,4}}{K_{H,I,O,4} + S_O} \frac{K_{H,I5,NO}}{K_{H,I5,NO} + S_{NO}} \frac{S_{NH}}{K_{H,NH} + S_{NH}} X_H$
R14	Decay	$b_H X_H$

3.2.4 Liquid-gas mass transfer

Transfer of N_2O from liquid to gas phase was modelled by diffusive links that connects the model compartments to a completely mixed gas compartment, representing the reactor head space. Uniform hydraulic pressure and gas phase composition were assumed across the reactor height. Therefore, the liquid gas transfer rate was calculated by:

$$R_{N_2O} = K_L a_{N_2O} \left(C_L - \frac{C_G}{H} \right) \qquad 3.1$$

where R_{N_2O} is the N_2O transfer rate (g $m^{-3}d^{-1}$), $K_L a_{N_2O}$ is N_2O volumetric liquid-gas transfer coefficient (d^{-1}), C_G is the concentration of N_2O in the gas phase, H is Henry's law constant, and $C_L$the concentration of N_2O in the liquid phase. The mass transfer coefficient of N_2O was calculated based on the gas-liquid oxygen transfer as follows:

$$K_L a_{N_2O} = K_L a_o \sqrt{\frac{D_{N_2O}}{D_o}} \qquad 3.2$$

where $K_L a_o$ represents the volumetric gas-liquid oxygen transfer, and D represents the molecular diffusivity of gas in water (m^2d^{-1}).

$K_L a_o$ was defined with an on/off time-dependent pattern to simulate the alternating aeration. During aerated periods, the mass transfer coefficient of oxygen was assumed to be correlated

to the superficial gas velocity calculated from the air supply flow rate and the reactor cross sectional area (Wan and Volcke, 2022), using the following expression:

$$K_La = \frac{0.6\, Q_{air}\, P_{atm}}{A\left(P_{atm} + \frac{pg\, H}{2}\right)} \qquad 3.3$$

where Q_{air} ($m^3\ d^{-1}$) denotes the airflow rate, P_{atm} is the atmospheric pressure in Pa, A (m^2) is the cross-sectional area of the reactor, H (m) is the reactor height, p (kg m^{-3}) is the density of water, and g (m s^{-2}) in is the gravitational acceleration. During non-aerated periods, K_La_o was assumed to be Zero, simulating no superficial gas velocity and no oxygen dissolution to the liquid.

3.2.5 Temperature effect

The effect of temperature on growth rates was modelled using a modified Arrhenius equation (Henze et al., 2000). Using this expression, the effect of temperature varies between bacterial species depending on their activation energies. The effect of temperature on the diffusion coefficients was also modelled as suggested in Venard and Street (1975). For mass transfer, both the temperature (T) impact on Henry's law constant and the mass-transfer coefficient were included according to the following equations:

$$H_T = H_{ref} e^{\alpha\left(\frac{1}{T} - \frac{1}{T_{ref}}\right)} \qquad 3.4$$

$$K_La_T = K_La_{Tref} 1.024^{(T - T_{ref})} \qquad 3.5$$

Where H_T and H_{ref} are the Henry's law constants ($M.atm^{-1}$) at temperature T and reference temperature T_{ref} (293.15 K), respectively; α represents the temperature dependency constant

(1500 and 2600 for O_2 and N_2O, respectively). K_La_T and K_La_{Tref} are the mass transfer coefficients at temperature T and at reference temperature T_{ref} (293.15 K), respectively. Furthermore, the effect of temperature on the diffusion coefficient has been included in the model using the following expression as in:

$$D_{i\,(T)} = D_{i\,Tref}\frac{\nu_{Tref}}{\nu_T}\frac{T}{T_{ref}} \qquad 3.6$$

Where $D_{i\,(T)}$ is the diffusion coefficient of substance (i) in water at temperature T; ν_T and ν_{Tref} (Pa.s) represent the dynamic viscosity of water at temperature and Tref, respectively.

3.2.6 Simulation initialization

The model underwent a simulation spanning 150 days with each cycle operated with four subcycles. The determination of the initialization simulation's duration was achieved through a trial-and-error method, wherein the primary metric of focus was the removal rates of NH_4^+ and NO_2^-. The objective of initialization was to ensure the model's stable performance over time in terms of NH_4^+ and NO_2^-. To ensure accurate simulation of the actual experimental reactor, the experimental SRT was used to calculate the sludge volume wasted from the completely mixed compartment of the SBR model. The simulation was run multiple times to determine the total simulation time at which the removal rates were relatively stable. Additionally, to ascertain the adequacy of the initialization simulation's total duration, it was cross-referenced with the standard benchmark of three times the estimated SRT.

3.2.7 Sensitivity analysis

A series of subsequent local sensitivity analyses was conducted to identify and select the model parameters that will be estimated in the calibration process based on their influence on model results. In each sensitivity analysis, simulations were based on the four subcycles

model and performed with a single parameter value changed at a time, while keeping all other parameters at their default values. To obtain a comprehensive sensitivity ranking across model parameters, the absolute-relative (AR) sensitivity was calculated at each time step and then averaged, resulting in a single AR sensitivity value for each parameter with respect to the selected state variables (NH_4^+, NO_2^-, and N_2O), which also has the unit of the state variable. AR sensitivity of each parameter was estimated by measuring the absolute changes in variables for a relative change in parameter values. A parameter was considered influential if its calculated AR sensitivity exceeded 5% of the highest influential parameter's sensitivity for the corresponding state variable. Consequently, parameters with lower AR sensitivity values were excluded from the calibration process since their influence on the model outputs was comparatively minimal. During sensitivity analysis, parameters were set to default values as shown in Table 3.4. The sensitivity analyses conducted prior to the calibration of NO_2^- and N_2O also involved the default values of parameters except for the parameters estimated in the previous calibration.

Furthermore, the form of the AR sensitivity functions of the state variables were examined in relation to the most influential parameters. If the form of a parameter's sensitivity function resembled that of another parameter, indicating a proportionality between their estimates, the parameter with the lower AR sensitivity value was considered non-identifiable and excluded from parameter estimation. This proportionality indeed implies that changes in one parameter's value can be compensated by changes in the other parameter (Gujer, 2008). The proportionality was evaluated by visual inspection of the correlation of the sensitivity functions of model parameters. For each variable, pairwise plots of sensitivity functions for

the most influential parameters were utilized to quantify the proportionality between the sensitivity functions of parameters with respect to the same variable.

Table 3-4 Default values used for model kinetic and stoichiometric parameters

Parameter	Description	Unit	Value	Source
μ_{AOB}	Maximum growth rate of AOB	d^{-1}	0.78	(Hiatt and Grady, 2008)
$K_{AOB,NH}$	AOB affinity constant for total ammonia nitrogen	mg-N. L^{-1}	0.20	(Wan and Volcke, 2022) (Vannecke and Volcke, 2015)
$K_{AOB,1,O}$	AOB affinity constant for O_2 (AMO reaction)	mg-O_2. L^{-1}	1.0	(Pocquet et al., 2016)
$K_{AOB,2,O}$	AOB affinity constant for O_2 (HAO reaction)	mg-O_2. L^{-1}	0.3	(Lang et al., 2017)
$K_{AOB,NH2OH}$	AOB affinity constant for NH_2OH	mg-N. L^{-1}	0.90	(Pocquet et al., 2016)
$K_{AOB,NN,NO}$	AOB affinity constant for NO (NN pathway reaction)	mg-N. L^{-1}	0.008	(Pocquet et al., 2016)
$K_{AOB,NO2}$	AOB affinity constant for total nitrite nitrogen	mg-N. L^{-1}	75.8	(Wan and Volcke, 2022) (Vannecke and Volcke, 2015)
$K_{AOB,I,O}$	AOB inhibition constant by O_2 (N_2O production via ND pathway)	mg-O_2. L^{-1}	0.1	(Wan and Volcke, 2022)
μ_H	Growth rate of heterotrophs	d^{-1}	6.25	(Hiatt and Grady, 2008)
$K_{H,S}$	Heterotrophs affinity constant for COD	mg-COD.L^{-1}	40	(Hiatt and Grady, 2008)
$K_{H,S,1}$	Heterotrophs affinity constant for COD (anoxic NO_3 denitrification)	mg-COD.L^{-1}	20	(Hiatt and Grady, 2008)

$K_{H,S,2}$	Heterotrophs affinity constant for COD (anoxic NO_2^- denitrification)	mg-COD.L^{-1}	20	(Hiatt and Grady, 2008)
$K_{H,S,3}$	Heterotrophs affinity constant for COD (anoxic NO denitrification)	mg-COD.L^{-1}	20	(Hiatt and Grady, 2008)
$K_{H,S,4}$	Heterotrophs affinity constant for COD (anoxic N_2O denitrification)	mg-COD.L^{-1}	20	(Hiatt and Grady, 2008)
$K_{H,I2,NO}$	NO inhibition of heterotrophs (anoxic NO_2^- denitrification)	g $N.m^{-3}$	0.5	(Hiatt and Grady, 2008)
$K_{H,I3,NO}$	NO inhibition of heterotrophs (anoxic NO denitrification)	g $N.m^{-3}$	0.3	(Hiatt and Grady, 2008)
$K_{H,I4,NO}$	NO inhibition of heterotrophs (anoxic N_2O denitrification)	g $N.m^{-3}$	0.075	(Hiatt and Grady, 2008)
K_{N2O}	Heterotrophs affinity constant for N_2O	g $N.m^{-3}$	0.05	(Hiatt and Grady, 2008)
η_{ND}	Reduction factor applied for the N_2O production through ND pathway	___	0.250	(Pocquet et al., 2016)
η_{NN}	Reduction factor applied for the N_2O production through NN pathway	___	0.0015	(Pocquet et al., 2016)
$K_{H,I,O,2}$	Heterotrophs O_2 inhibition constant (anoxic NO_2^- denitrification)	mg-O_2. L^{-1}	0.1	(Hiatt and Grady, 2008)
$K_{H,I,O,4}$	Heterotrophs O_2 inhibition constant (anoxic N_2O denitrification)	mg-O_2. L^{-1}	0.1	(Hiatt and Grady, 2008)

f_{XI}	Fraction of inert COD in biomass	g-COD. g-COD^{-1}	0.08	(Henze et al., 2000)
i_{NSS}	Nitrogen content of soluble organic substrate	g-N. g-COD^{-1}	0.03	(Henze et al., 2000)
i_{NXB}	Nitrogen content of biomass	g-N. g- COD^{-1}	0.07	(Henze et al., 2000)
Y_{AOB}	Yield of AOB	g-COD. g- N^{-1}	0.15	(Henze et al., 2000)
Y_{NOB}	Yield of NOB	g-COD. g- N^{-1}	0.057	(Henze et al., 2000)
Y_H	Yield of heterotrophs	g-COD. g- N^{-1}	0.6	(Hiatt and Grady, 2008)
η_Y	Yield reduction factor under anoxic conditions	___	0.9	(Hiatt and Grady, 2008)
η_{g2}	Reduction factor for μ_H under anoxic conditions in process 10	___	0.28	(Hiatt and Grady, 2008)
η_{g3}	Reduction factor for maximum growth rate under anoxic conditions in process 11	___	0.16	(Hiatt and Grady, 2008)
η_{g4}	Reduction factor for maximum growth rate under anoxic conditions in process 12	___	0.35	(Hiatt and Grady, 2008)
η_{g5}	Reduction factor for maximum growth rate under anoxic conditions in process 13	___	0.35	(Hiatt and Grady, 2008)
η_Y	Reduction factor for yield under anoxic conditions	___	0.9	(Hiatt and Grady, 2008)

3.2.8 Model calibration and validation

The model was calibrated using the data of four subcycles. Parameters estimated for calibration were chosen based on the sensitivity analysis results. In addition to parameters that were deemed non-identifiable, parameters delineating the elemental biomass composition, yield values, decay rates, and temperature coefficients were retained at their default values and excluded from calibration given the high agreement on their values in literature. The calibration commenced by estimation of the final chosen parameter set with the model being adjusted to align with the NH_4^+ data, followed by the NO_2^- data. Those parameters that had significant influence on NO_2^- concentrations were subsequently assessed for their impact on NH_4^+. If any such parameter exhibited an influence on NH_4^+, it was incorporated into the calibration process of NO_2^-. This sequential calibration protocol was pioneered by Corominas et al. (2011), and further developed by Mannina et al. (2011). The calibration aimed at minimizing the chi square (X^2) objective function according to Gujer (2008) as follows:

$$X^2 = \sum_{i=1}^{n} \left[\frac{y_{m.i} - y_i}{\sigma_{m.i}} \right]^2 \qquad 3.7$$

where X^2 denotes the sum of squared weighted deviations between the measured and simulated effluent concentrations, $y_{m.i}$ the measured value of the state variable in the experiment, y_i the corresponding model prediction of the state variable at the same time, $\sigma_{m.i}$ the standard error of the measurement $y_{m.i}$. The standard errors of state variables were assumed as 1% of the average measurements. Following the successful calibration for effluent NH_4^+ and NO_2^- concentrations, attention was directed toward calibrating the N_2O concentration in the gas phase. The choice of parameters for this calibration was again

driven by the results of the sensitivity analysis. The step-by-step calibration approach employed in this study is illustrated in Figure 3.2. Post-calibration, the model underwent validation, employing data derived from experiments with seven operational sub-cycles, i.e. obtained under quite different operational conditions, providing a challenging validation test.

3.2.9 Scenario analysis of reactor operational conditions

The calibrated model was used in over 1500 scenarios simulating a wide range of DO, temperatures, and influent NH_4^+ to investigate the combined effect of environmental and operational conditions, and influent nitrogen load on both N_2O emissions and nitrogen removal. Specifically, the temperature varied from 5°C to 25°C, while DO during the aerated time was adjusted between the limits of 0.3 to 2 mg/L, and influent NH_4^+ concentrations fluctuated between 700 to 1000 mg/L to represent cases of both regular and higher influent loading. The simulation results were aggregated and processed to show the average N_2O concentration in the headspace during the simulation and the final effluent concentration of NH_4^+. The results were then plotted into heatmaps to show the average concentration of either N_2O or NH_4^+ of every single combination of DO and temperature when simulated under different influent loads.

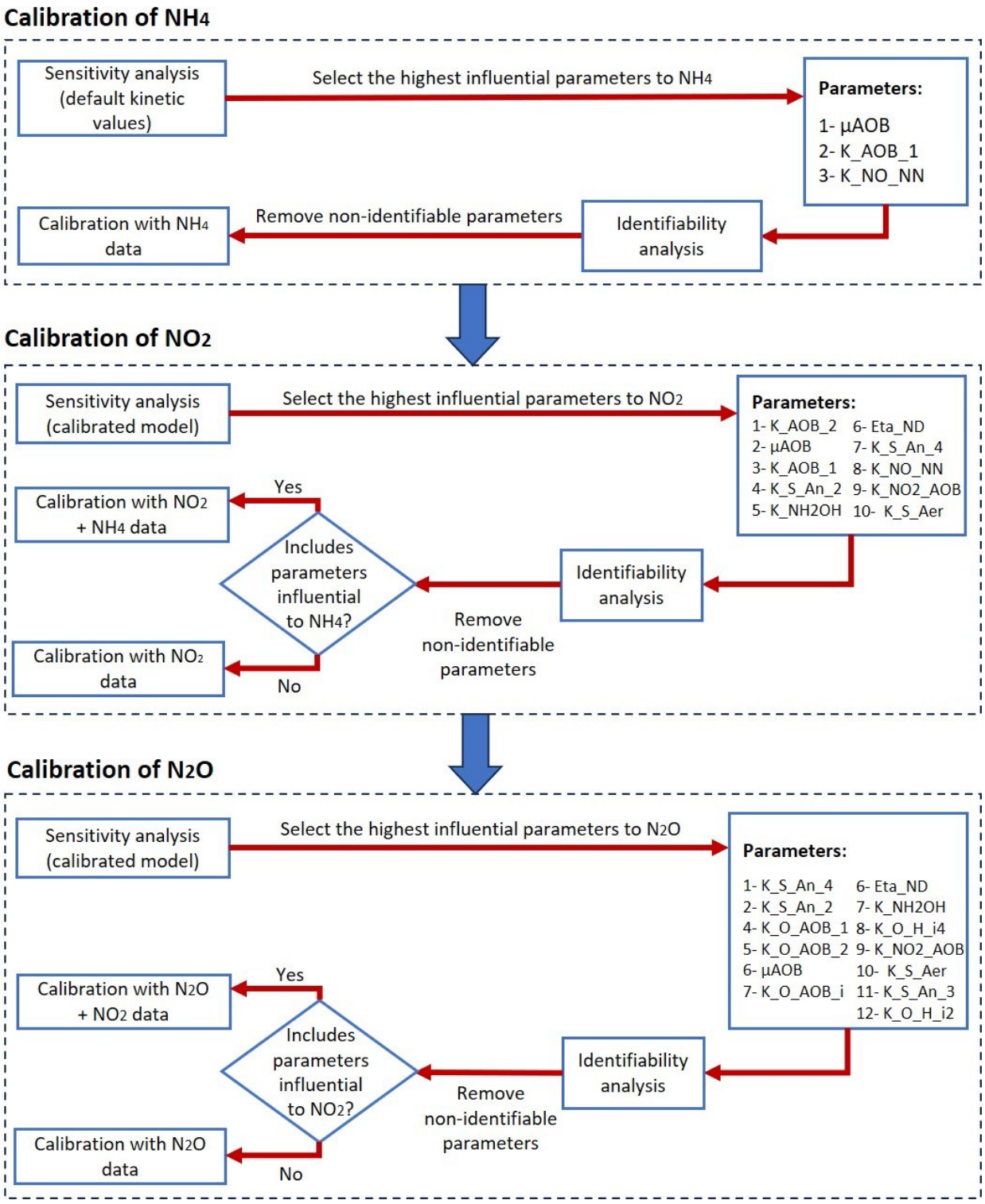


Figure 3-2 The procedure used for calibration of mechanistic model in the study.

3.3 Results and Discussion

3.3.1 Sensitivity analysis and model calibration

An initial simulation of the model was performed, operating under four subcycles with the model parameters set to reference values. Figure 3.3 captures the sensitivity ranking of the model parameters, providing insights into how various model parameters impact the values of state variables NH_4^+, NO_2^-, and N_2O. Foremost, NH_4^+ concentrations are majorly influenced by parameters linked to AMO-mediated reaction kinetics, where the maximum specific growth rate (μ_{AOB}) and the oxygen affinity constant ($K_{AOB,1,O}$) were the most influential. Regarding NO_2^-, the analysis demonstrated that its concentration is influenced not just by the μ_{AOB} of AOB but also by a collection of other factors, including oxygen affinity constants pertinent to both AMO and HAO-mediated AOB reactions ($K_{AOB,1,O}$ and $K_{AOB,2,O}$) organic substrate affinity constants associated with heterotrophic anoxic reactions, and the ND pathway kinetics. NO_2^- plays a pivotal role in both aerobic and anoxic metabolic pathways, in addition to its function as an electron acceptor for the ND pathway producing N_2O, explaining its influence by a larger set of parameters compared to NH_4^+ (Hiatt and Grady, 2008). The sensitivity results also show that N_2O concentrations are heavily influenced by the affinity constants of anoxic heterotrophic organic substrate, in addition to parameters related to both AMO and HAO-mediated AOB reactions. This aligns with the current knowledge on N_2O production pathways, as the reactions involving those parameters are the sources of NO_2^- and NH_2OH formations, which ultimately steer the production of N_2O through the ND and NN pathways (Ni et al., 2014). Meanwhile, the organic substrate affinity constants mark the anoxic denitrification reactions, with N_2O featuring as an intermediate. The elevated influence of these parameters refers to the

significant contribution of anoxic heterotrophic pathways to the N_2O sink in the IFAS-SBR reactor. This can be potentially attributed to the predominant low DO operation of the reactor, due to both the aerated/non-aerated alternate operation and the low DO even during the phases when aeration is active. Figures 3.4, 3.5, and 3.6 show the pairwise relationship between the sensitivity functions of the model most influential parameters with respect to NH_4^+, NO_2^-, and N_2O, respectively.

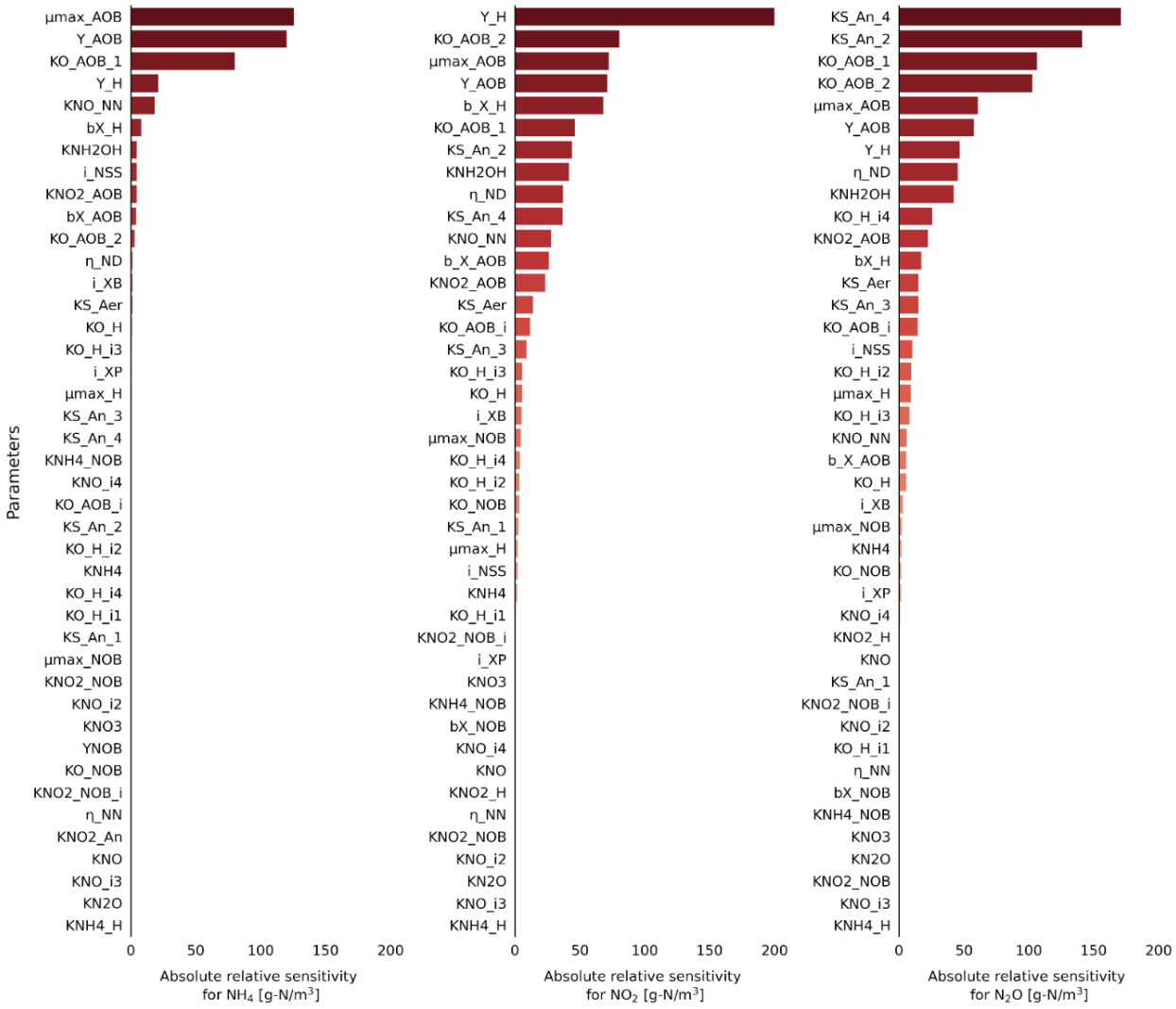


Figure 3-3 Ranked absolute relative (AR) sensitivity results of model parameters towards model state variables (NH_4^+, NO_2^-, and N_2O concentrations). The cutoff line is based on 5% of the maximum AR sensitivity for each state variable.

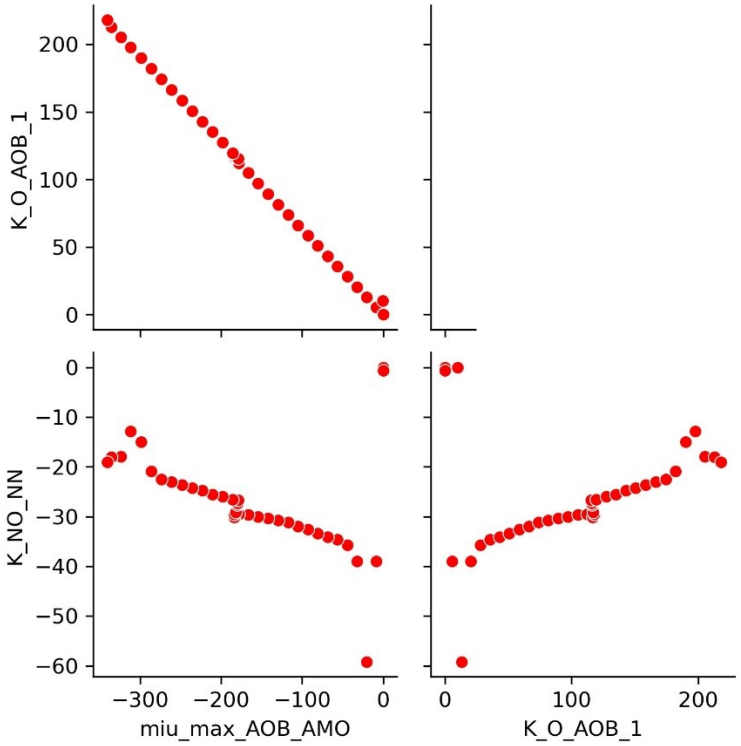


Figure 3-4 The pairwise relationship between the sensitivity functions of model most influential parameters with respect to NH_4^+

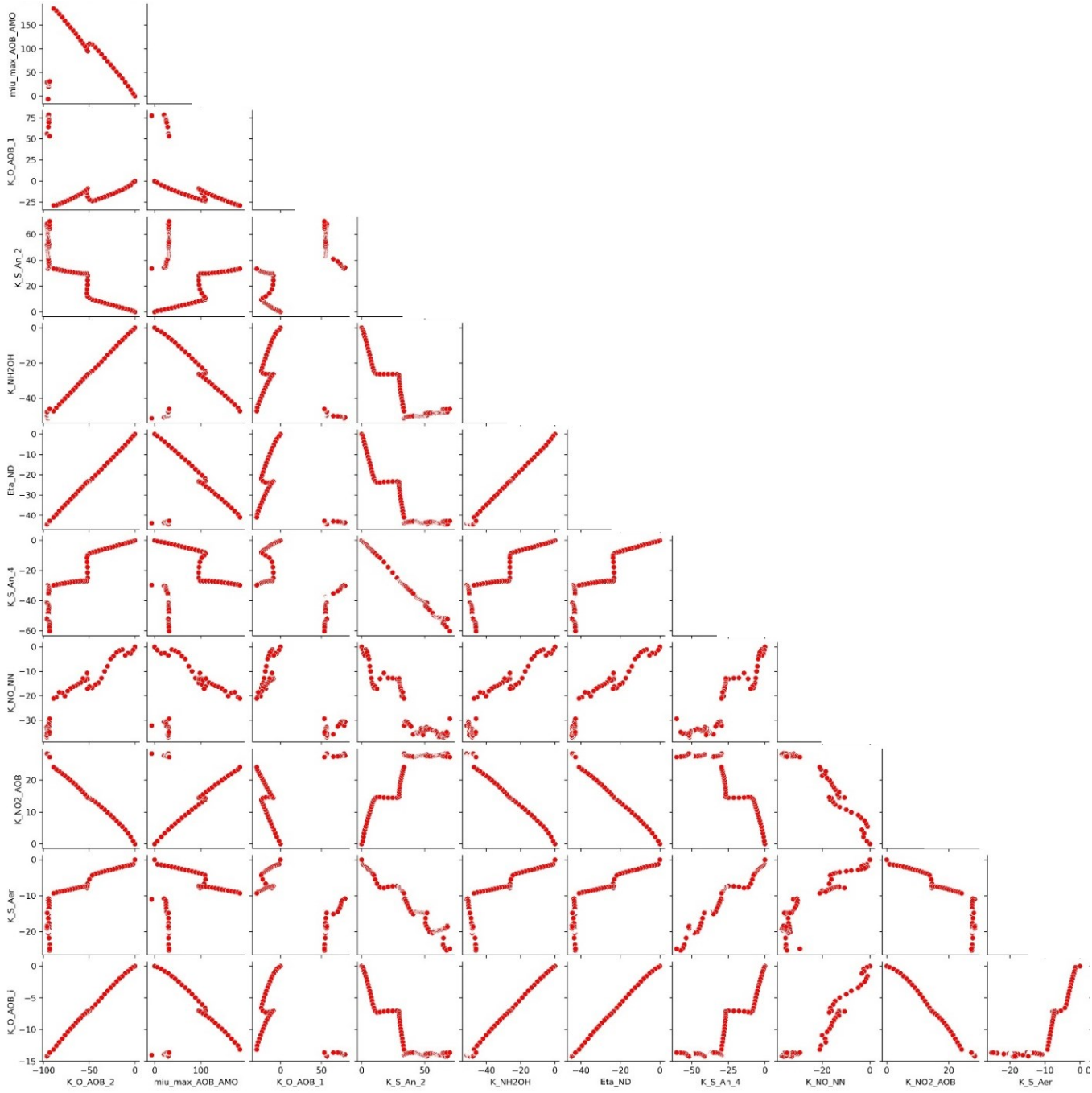


Figure 3-5 The pairwise relationship between the sensitivity functions of model most influential parameters with respect to NO_2^-

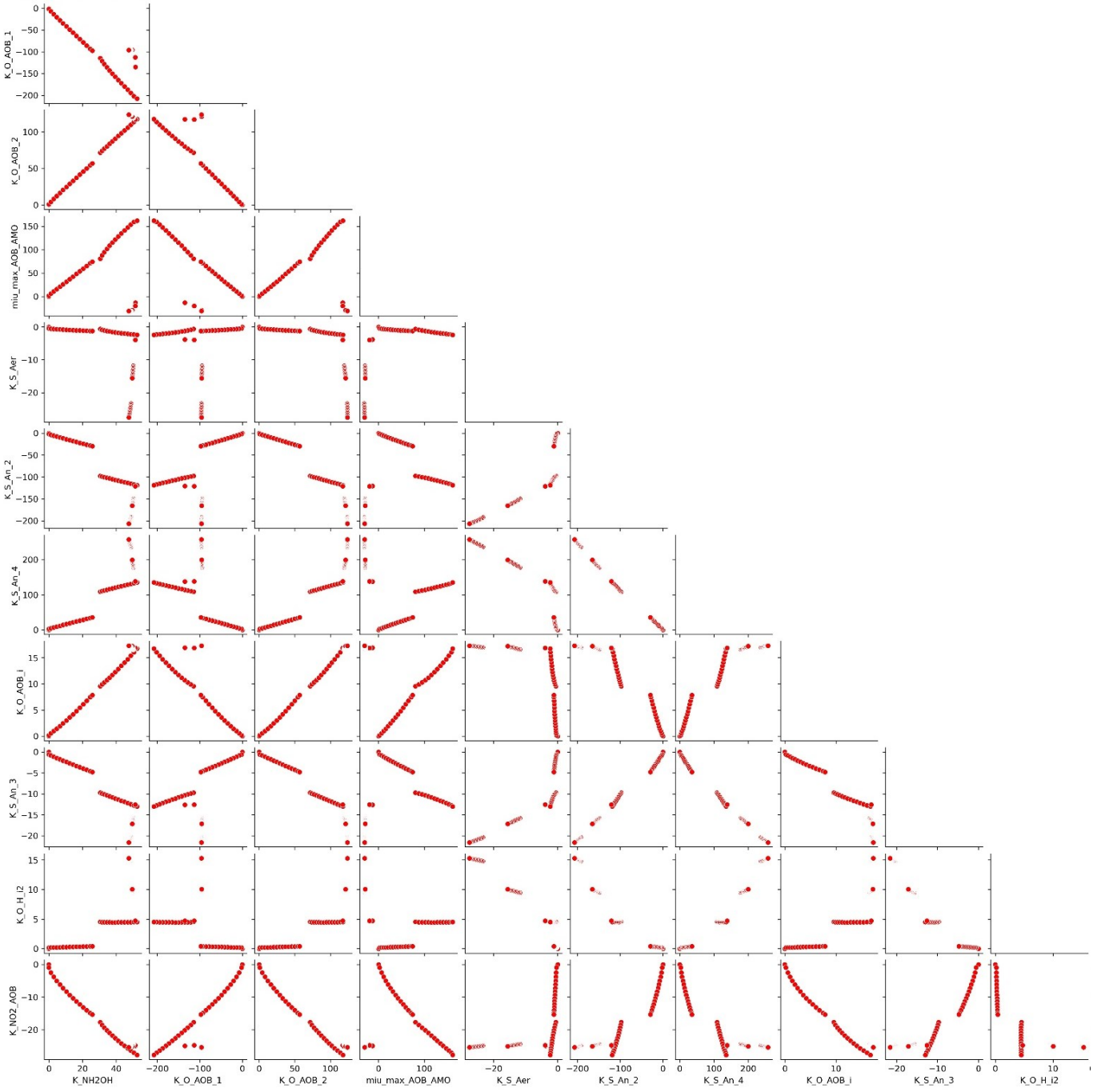


Figure 3-6 The pairwise relationship between the sensitivity functions of model most influential parameters with respect to N_2O

The model was calibrated with data obtained from the IFAS-SBR reactor operation with four subcycles. The numerical values of the parameters obtained from calibration in comparison with default values are shown in Table 3.5. Figure 3.7 shows the experimental data and model simulations for the dynamics of the NH_4^+ and NO_2^- concentrations in a typical full cycle operated with 4 subcycles of aerobic/anoxic alternations. The model was able to predict the dynamics of NH_4^+ and NO_2^- during a full cycle. The rate of NH_4^+ oxidation in the model was similar to the oxidation rate obtained from the experimental data. However, the NO_2^- production rate in the simulation was a bit lower than the rate calculated from the experimental data. Moreover, simulation results showed that nitrogen (N_2) production taking place during aeration (results not shown) indicating that NO_2^- was still denitrified during aerated phases. Thus, during aeration denitrifiers have more contribution in the simulation than the actual experiment. This is a result of the relatively low DO concentration (0.6 mg/L) during aerated phases, which is not high enough to completely inhibit NO_2^- denitrification in the model given their calibrated oxygen inhibition constants ($K_{H,I,O,2} = 0.17$). Similarly, the rate of NO_2^- denitrification in the simulation looks lower than the rate in the experimental data because the DO during the non-aeration time did not completely drop to 0 mg/L in the experiment, and DO up to 0.3 mg/L was measured during the non-aeration time (Zou et al., 2022). The estimated value of $K_{H,I,O,2}$ provided the best balance for the model to capture the high and low peaks of the experimental NO_2^-. Experimentation with higher and lower values of $K_{H,I,O,2}$ resulted in an increased ability of the model to solely capture either the high or low NO_2^- peaks while missing the other peak (results not shown). This is potentially an indication of the limited ability of the model to predict target variables with higher accuracy. The model was neither

able to depict the full quantitative range of NH_4^+ nor the NO_2^- dynamics potentially due to using a multi-step calibration procedure. Although this procedure was used in previous studies (Lang et al., 2017), it has the disadvantage of potentially resulting in a nonoptimal N_2O formation during calibration of NH_4^+ and NO_2^- which could lead to deviations in the modelled NO_2^- or NH_4^+ (Wan and Volcke, 2022).

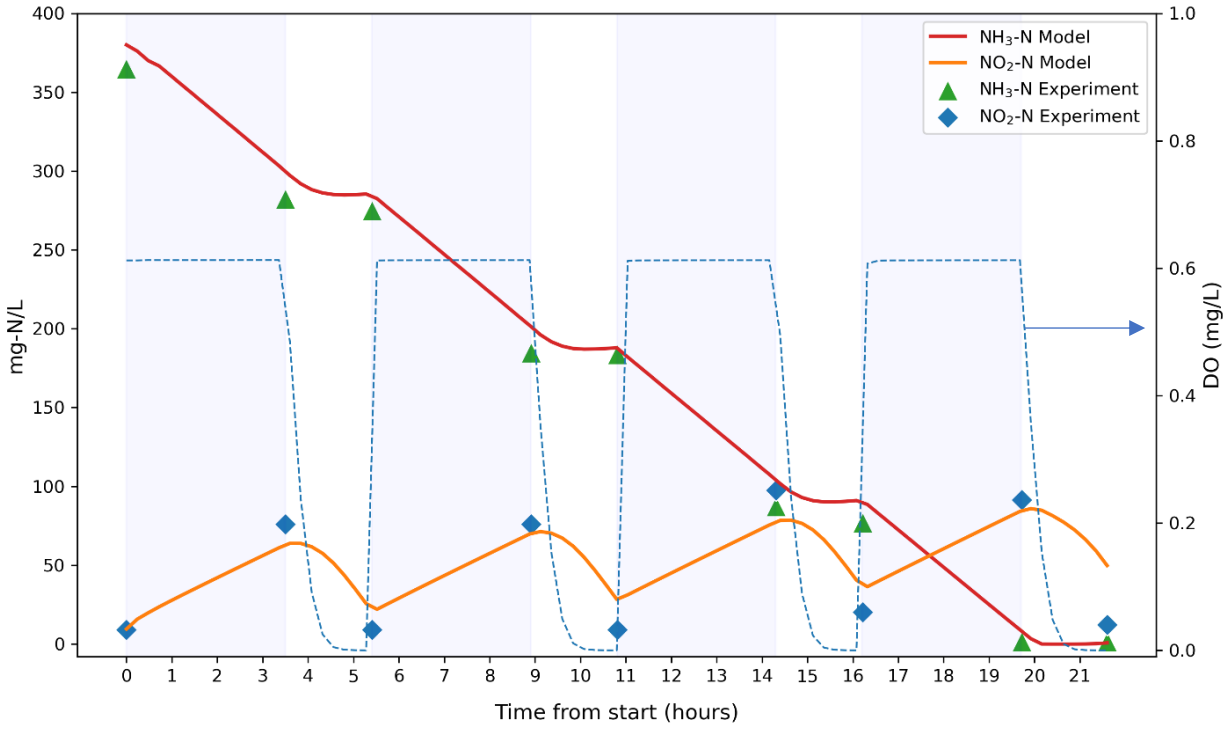


Figure 3-7 Profiles of NH_4^+ and NO_2^- in full cycle (four subcycles) comparing the measured experimental data (points) with model predictions (lines) Shaded areas represent times when the aeration is turned on and the dotted line shows the simulated DO in the bulk liquid.

Table 3-5 Numerical values of the parameters estimated in this study and the corresponding default values used before calibration.

Parameter	Description	Unit	Estimated value	Default value	Source
μ_{AOB}	Maximum growth rate of AOB	d^{-1}	0.39	0.78	(Hiatt and Grady, 2008)
$K_{AOB,NH2OH}$	AOB affinity for NH_2OH	mg $N.L^{-1}$	0.4	0.90	(Pocquet et al., 2016)
$K_{AOB,NN,NO}$	AOB affinity for NO (NN pathway reaction)	mg $N.L^{-1}$	0.005	0.008	(Pocquet et al., 2016)
η_{ND}	Reduction factor applied for the N_2O production through ND pathway	Unitless	0.35	0.250	(Pocquet et al., 2016)
$K_{H,S}$	Heterotrophs affinity for COD (aerobic)	mg $O_2.L^{-1}$	34.6	40	(Hiatt and Grady, 2008)
$K_{H,S,2}$	Heterotrophs constant for COD (anoxic NO_2^- denitrification)	mg $O_2.L^{-1}$	20.6	20	(Hiatt and Grady, 2008)
$K_{H,S,3}$	Heterotrophs affinity for COD (anoxic NO denitrification)	mg $O_2.L^{-1}$	19.2	20	(Hiatt and Grady, 2008)
$K_{H,S,4}$	Heterotrophs affinity for COD (anoxic N_2O denitrification)	mg $O_2.L^{-1}$	23.5	20	(Hiatt and Grady, 2008)

$K_{H,I,O,2}$	Heterotrophs O_2 inhibition (anoxic NO_2^- denitrification)	mg $O_2.L^{-1}$	0.17	0.1	(Hiatt and Grady, 2008)
$K_{H,I,O,4}$	Heterotrophs O_2 inhibition (anoxic N_2O denitrification)	mg $O_2.L^{-1}$	0.088	0.1	(Hiatt and Grady, 2008)

Figure 3.8a compares the actual and simulated N_2O concentration dynamics during a full cycle and the percentage of emissions produced during both anoxic and aerobic phases with model results. The model was able to predict the overall N_2O concentration dynamics during the entire cycle in the headspace with R^2, and RMSE of 0.82 and 34.9, respectively. During the time of aeration, the N_2O concentration in the headspace increased until reaching its peak by the end of the aeration time before it stays constant after the aeration stops. The constant N_2O concentration during the non-aerated periods is due to the absence of air flow, as the mass transfer coefficient in the model was linked to the air flow as described in equation (3.3).

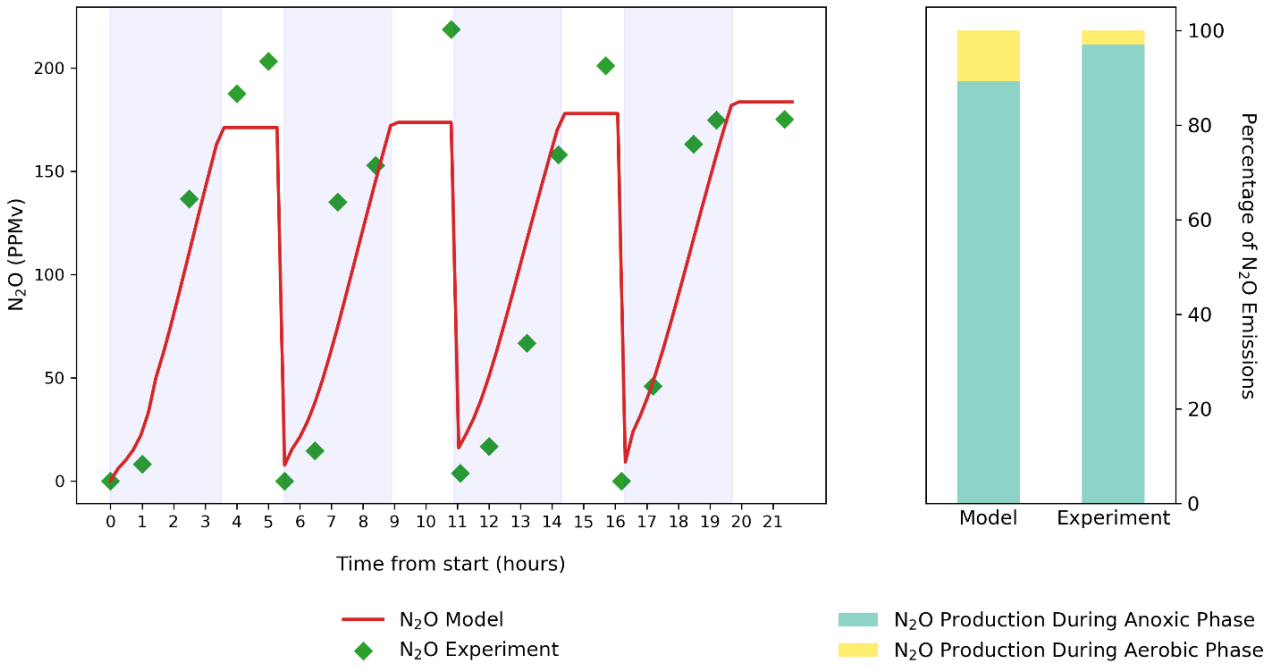


Figure 3-8 Profile of N_2O concentrations in the headspace (ppm) in a full cycle (four subcycles) experiment (right) comparison of the percentage of N_2O produced during aerated and non-aerated periods as calculated from the model and experimental observations.

3.3.2 N_2O production pathways and quantification

The model results from the calibrated model indicate that ND is the dominant pathway for N_2O production in the IFAS-SBR reactor with 95% of the produced N_2O coming from ND pathway. Although experimental findings suggested ND pathway dominance, direct quantification of each pathway's contribution remained elusive due to the lack of a method for measuring them explicitly (Zou et al., 2020). The ND pathway is often more significant at higher NO_2^- and low DO concentrations (Peng et al., 2014). These conditions are unavoidable in case of nitritation-denitritation given that NOB are suppressed and thus, NO_2^- will accumulate (Fux et al., 2006). Furthermore, the low DO during the aerated phase

of the currently studied reactor makes the ND pathway contribution even more significant (Yang et al., 2009).

Figure 3.8b demonstrates that approximately 90% of N_2O was produced through nitritation (i.e., when aeration was turned on) according to the model results, compared to 97% as deduced from the experimental observations. Further analysis of the model results, as shown in Figure 3.6, reveals that for the total N_2O produced via the ND pathway, 89% occurred during aerated periods. Similarly, 96% of the N_2O produced through the NN pathway also originated during these aerated periods. This finding aligns with the mechanistic understanding of the NN pathway, where the reaction requires oxygen, with NH_2OH serving as the electron donor (Ni et al., 2013a). Interestingly, this observation applies to the ND pathway as well, even though their reactions are typically described as anoxic. The significant amount of N_2O emissions by the ND pathway during the aerated phase can be explained by the operational conditions of the IFAS-SBR, where the DO concentration during the aerated phase is around 0.6 mg/L, which is not high enough for complete inhibition of the ND pathway. The fact that lower N_2O was produced during the anoxic phase is due to the lack of competition for NO_2^- by heterotrophs and AOBs during this phase, as heterotrophs, with their NO_2^- affinity constant two orders of magnitude lower than that of AOB, can easily outcompete them.

The calibrated model, which simulates the operation of four sub-cycles, was utilized to provide a detailed analysis of N_2O production. The analysis revealed that 99.5% of N_2O production during the simulation was carried by the ND pathway compared to only 0.5% by NN pathway, while no net production took place by HD pathway. In terms of biomass contribution, biofilm contributed with about 90% of the produced N_2O compared to 10%

produced at the flocculent sludge. Figure 3.9 shows the distribution of N_2O produced in both aerated and non-aerated phases within the biofilm and flocculent sludge by both NN and ND pathways, accounting for the total N_2O production. The results show that most of the N_2O production was carried out during the aeration phases with about 90% and 70% of the N_2O produced in biofilm and flocculent sludge was produced during the aerated phase, respectively. This can be attributed to the production of both NO_2^- and NH_2OH during aeration, in which NO_2^- serves as the electron acceptor for the ND pathway and NH_2OH as the electron donor for both ND and NN pathways. During non-aerated conditions, the competition is higher for NO_2^- as anoxic heterotrophs consume it at higher rates. Furthermore, the NH_2OH would deplete due to the absence of its production in anoxic conditions. It should be highlighted that the simulation revealed no net N_2O production through the HD pathways. Instead, during non-aerated periods, HD pathways functioned as a sink, partially consuming the N_2O produced by the NN and ND pathways. This observation may be attributed to the supplementary COD during non-aerated periods, which likely reduced the competition for electrons that typically results in N_2O production via the HD pathway. The figure also demonstrates that N_2O produced during the aerated phase contributes a higher portion of the total N_2O produced from the biofilm than of the total N_2O produced in the flocculent sludge. This can be attributed to mass transfer limitation in the biofilm, where the low concentration of NH_2OH can be even lower inside the biofilm matrix, causing limited availability of electrons for the ND pathway to take place in the biofilm.

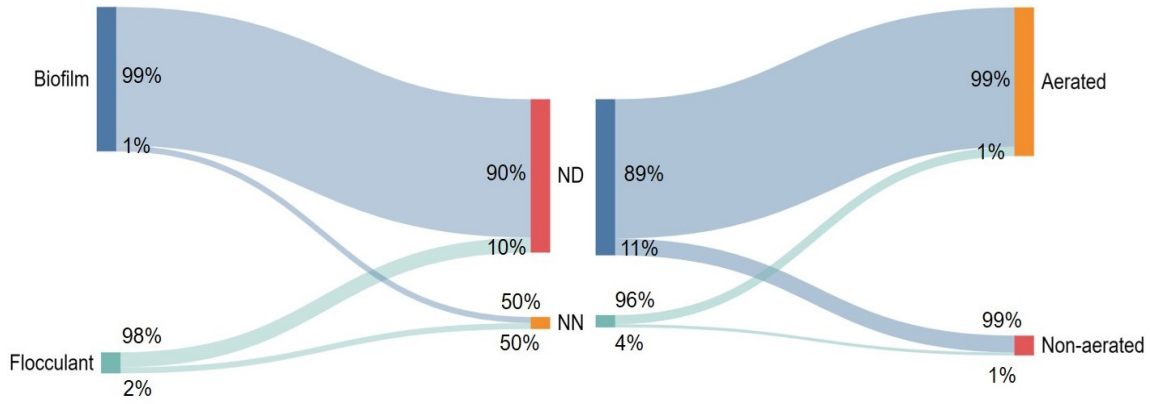


Figure 3-9 Quantification of N_2O production sources in terms of the two pathways (ND or NN), aeration conditions (aerated or non-aerated periods, right), and the type of biomass (biofilm or flocculant sludge, left) based on the simulations with 4 subcycles operation. The percentages represent the percentage of N_2O produced with the condition of the corresponding node with respect to the other node. The heterotrophic denitrification pathway did not contribute to N_2O production.

3.3.3 Model validation

Model validation was conducted with data of reactor operation with 7 subcycles; a different aeration alternation frequency from the data in which the model has been calibrated. Figure 3.10 shows both the actual and modelled ammonia and nitrite concentration during a full cycle of operation. The model was able to generalize and reproduce the ammonia dynamics during the operation of one cycle. Overall, while the experimental data demonstrated complete removal of the initial ammonia concentration of 400 mg/L, the model predicted a residual concentration of 20 mg/L. This discrepancy equates to a 5% error in the calculated ammonia removal efficiency, reflecting the deviation due to a marginally lower

simulated nitrification rate compared to actual experimental conditions. This discrepancy may be attributed to potential differences in biomass concentration and composition observed between the calibration and validation experiments. Additionally, the adaptation of Ammonia-Oxidizing Bacteria (AOB) to the reactor's operational conditions, such as low dissolved oxygen (DO) levels, could also play a significant role. Notably, the transition to a seven-subcycle operation occurred one month after a period of operation with four subcycles, suggesting temporal changes in biological activity and reactor performance. This can result into higher growth rate or lower oxygen affinity constants leading to higher ammonia consumption than the model predicts. For nitrite, the model predicted the final concentration by the end of cycle as 60 mg-N/L, while the actual concentration in the reactor was 35 mg-N/L, corresponding approximately to the non-nitrified ammonia. Similar to the performance with the calibration data, the model also predicts a higher rate of NO_2^- accumulation than the observed data used for validation, leading to over-estimation of both NO_2^- high and low peaks. This can be attributed to a potential adaptation of the heterotrophic denitrifiers to the DO concentration in the reactor during the experiment, given the time difference between 4 subcycles operation (calibration data) and 7 subcycles operation (validation data). This potential adaptation can be simulated by changing the values of the affinity constants from the values estimated based on the calibration. This allows avoiding the elevated NO_2^- denitrification rates during aerated phases that result in reduced NO_2^- concentration at the end of aeration. Noteworthy, this modelling challenge is inherent to the nature of mechanistic models of biological wastewater treatment processes represented mainly by the activated sludge models (ASMs). Indeed, in these models parameters are assumed to be constant over the simulation time, limiting the model's ability

to capture the variability of subprocesses such as the microbial community changes (Ni and Yu, 2010). An interesting solution to overcome this challenge is supplementing ASM models with a data-driven part to capture the unobserved dynamics, forming a "hybrid model" (Schneider et al., 2022).

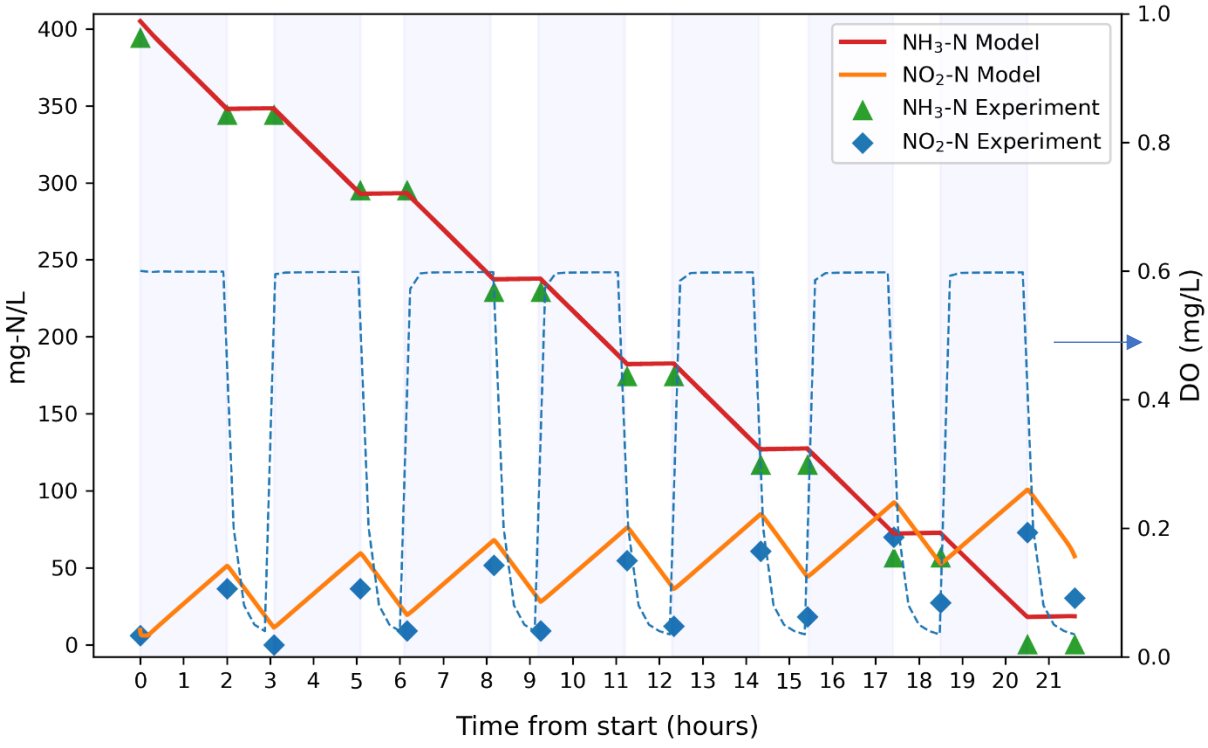


Figure 3-10 Profiles of NH_4^+ and NO_2^- during seven-subcycles operation of the reactor (validation data). Shaded areas represent times when the aeration is turned on and the dotted line shows the simulated DO in the bulk liquid.

In Figure 3.11, a comparison is presented between the observed N_2O concentrations and the corresponding predictions from the simulation over a complete cycle operated with seven subcycles. The simulation provided a qualitative match to the actual N_2O emissions observed in the reactor but failed to capture the full range of N_2O dynamics within the

reactor's headspace. The experimental observations showed that the flux of N_2O to the headspace during aerated periods declined in the last two subcycles, and the concentration N_2O in the headspace was significantly lower than in previous subcycles; a behaviour that was not captured by the model. While one may relate this to the aforementioned difficulty of the model to sufficiently remove NH_4^+ in the last cycle, the discrepancy occurs in the last two subcycles, leading to continued production of N_2O all the way till the end of the seventh subcycle, indicating that ND was the dominant pathway producing N_2O in the simulation. The decline of N_2O production in the last two subcycles in the experimental observations does not appear to be related to the exhaustion of NH_4^+ in the mixed liquor, over 50 mg-N/L remaining at the end of the sixth subcycle, at which point the N_2O production has already reduced significantly. This observation warrants further study. This discrepancy suggests that a more refined model could better represent the N_2O flux under varying operational conditions. A similar limitation was noted by Wan and Volcke, (2022) in their study of a granular sludge partial nitritation-anammox reactor, which also demonstrated the model shortcoming to capture the full dynamics. One potential explanation for the model's inability to fully replicate the observed N_2O dynamics is the initial conditions set for the IFAS-SBR model, especially the biomass composition, which were based on a four-subcycle simulation while the validation data represents the seven-subcycle operational strategy. Considering that the experiment spanned several months, the interim allowed for substantial microbial community changes. These microbial community changes are evidenced by the shift in the functional gene ratio *nosZ/(nirS+nirK)* from 0.63 to 0.58 reported by (Zou et al., 2020). The *nirS* and *nirK* genes are responsible for reducing NO_2^- to NO, while the *nosZ* gene codes for reducing N_2O to

N_2. This shift suggests a diminished N_2O reduction capacity, which could account for the inability of the model to predict the higher N_2O levels. Additionally, variations in biofilm thickness over time could alter the concentration profiles within the biofilm matrix, further influencing reactor performance. To address these limitations, future model developments could benefit from continuous multi-cycle simulations, but this will need data collection over several cycles.

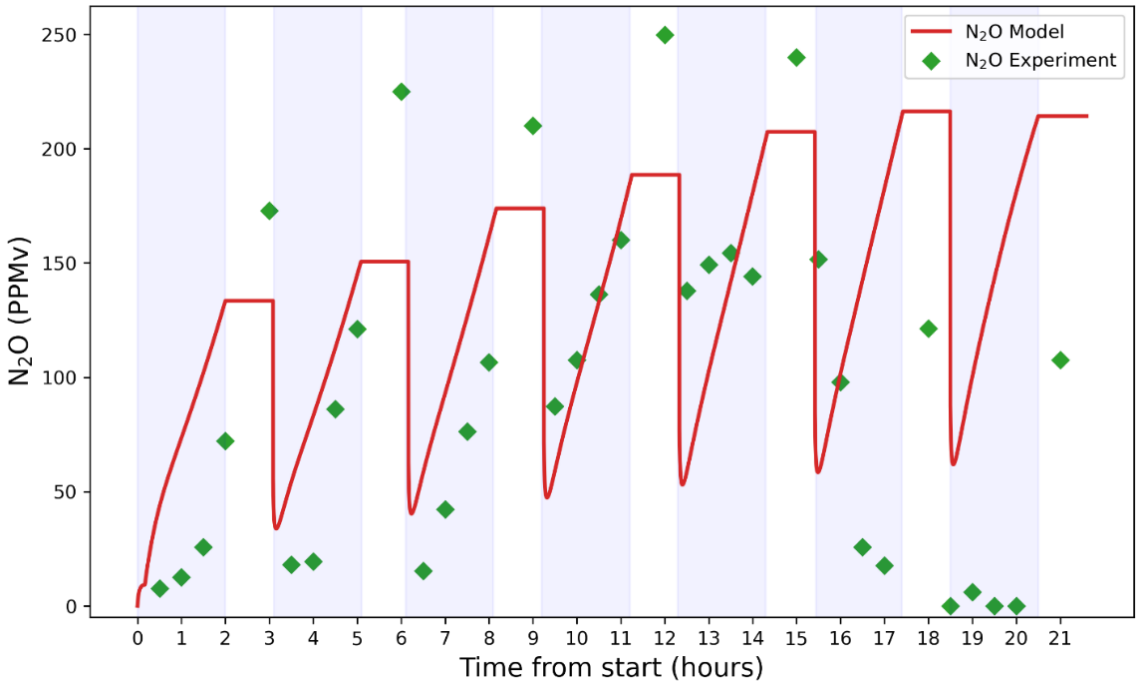


Figure 3-11 Profile of N_2O concentrations in the headspace (ppm) in a seven subcycles experiment (validation data) comparing the measured experimental data shown as points with the model predictions shown as lines. Shaded areas represent times when the aeration is turned on.

3.3.4 Reactor performance and N_2O emissions under various operational conditions

The combined impact of both DO setpoint and liquid temperature on N_2O emissions and NH_4^+ removal was simulated under a wide range of influent NH_4^+ concentration using the 4 subcycles calibrated model. Figure 3.12 and 3.13 represent the average values of both the average N_2O in the headspace and effluent NH_4^+, respectively. The results were obtained from over 1500 simulations covering DO ranges from 0.2 to 2.0 mg/L, temperatures from 5 to 25 ^oC, and influent ammonia concentrations ranging from 700 to 1000 mg/L. Figure 3.12 shows that the effluent NH_4^+ concentration is inversely related to the temperature and DO, while Figure 3.13 shows that the N_2O emissions increase with both temperature and DO. It is a well-established knowledge that NH_4^+ removal gets better at higher DO, given the expected faster nitrification rate. However, the purpose of the heatmap was to find the optimal DO under varying temperature that achieves NH_4^+ removal with minimal possible emissions. Note that Figure 3.11 and 3.12 show the average effluent concentrations over a range of influent NH_4^+ concentrations.

N_2O concentration in the headspace exhibited an increase with increasing DO in the aeration phase despite the fact that the majority of the N_2O produced in this experiment occurred by the anoxic ND pathway. This can be explained that by the fact that higher DO will increase the reaction rate of the NN pathway during aerated periods regardless of the already-high production taking place during the anoxic phase by the ND pathway. This also matches with the correlation results shown in Figure 3.14. Indeed, the DO showed moderate to strong correlation with average NH_2OH and NO over the reactor operation time. The N_2O production by the ND pathway can also increase with DO during the aerated period as the increased DO will lead to higher nitrification rates, which means a higher

NO_2^- and NH_2OH accumulation, which are the electron acceptor and donor of the ND pathway, respectively (Kampschreur et al., 2009). Furthermore, because of the mass-transfer limitation and the concentration gradient created by the biofilm of the IFAS reactor, the increase of DO may not lead to decreased N_2O production by the ND pathway during the aeration periods, as there would exist areas inside the biofilm with low enough DO for the ND pathway to function. However, this will be completely dependent on the biofilm thickness, as an extremely low biofilm thickness may not create sufficient gradient to deplete DO at the biofilm base.

The effect of temperature on N_2O emissions is two-fold, as it affects both liquid-gas mass-transfer of N_2O and the overall kinetics of biological transformation reactions (Baeten et al., 2020). The current model has included the impact of temperature on mass-transfer by including its effect on both the saturation concentration of N_2O in the liquid phase and the mass transfer coefficient. The saturation concentration of N_2O in the liquid is expected to decrease with increased temperature, leading to a higher chance of N_2O to escape to the gas phase. Furthermore, higher temperature results in lower viscosity and higher diffusion rate, increasing the mass transfer coefficient. The overall behaviour is that an increase in temperature leads to an increase in N_2O emissions in addition to the increased enzymatic activities of both nitrifiers and denitrifiers at higher temperatures (Ahn et al., 2010). In contrast, it was demonstrated by Adouani et al. (2015) that a lower temperature slows down the activity of the N_2O reductase, which could lead to an increase in N_2O emissions. However, lower temperatures were associated with higher emissions in other studies, such as in Daelman et al. (2013), where the emissions were significantly higher when temperatures were as low as 9°C compared to 18°C. However, it was not mentioned

whether the DO was changed during this lower temperature or not. A higher DO in the aerated zones might have been required during lower temperature to achieve the same nitrogen removal as in summer, causing the higher emissions.

The impact of changing DO under several temperatures on NH_4^+ removal and N_2O emissions is contradictory. Under lower temperatures, a higher DO during the aeration phases may be needed to achieve the same ammonia removal. However, this may cause higher N_2O emissions. Therefore, the heatmaps in Figure 3.12 and 3.13 could be used to identify the optimal operational zone. For the current reactor with the same other operational conditions, the optimal DO needs to be well-above 1.4 mg/L during aeration time to achieve NH_4^+ concentration in the effluent below 25 mg-N/L when temperature is below 15ºC. For the same temperature, DO during aeration phase needs to be controlled below 0.8 mg/L for the least N_2O emissions. Under lower temperature the DO could be further increased without significant increase in N_2O emissions. For example, at a temperature of 5ºC, the DO could be increased up to 1.1 mg/L with almost the same N_2O emissions rate. Moreover, the increase in NH_4^+ removal as a response to DO increase is less significant under low temperature, making it more challenging to change the DO set point to satisfy both NH_3 removal and N_2O emissions. This result is in alignment with conclusions of other studies that suboptimal DO conditions are a key factor contributing to elevated N_2O emissions and that multi-objective control is necessary, instead of single objective control that is only based on NH_4^+, for instance, as in Hwangbo et al. (2021).

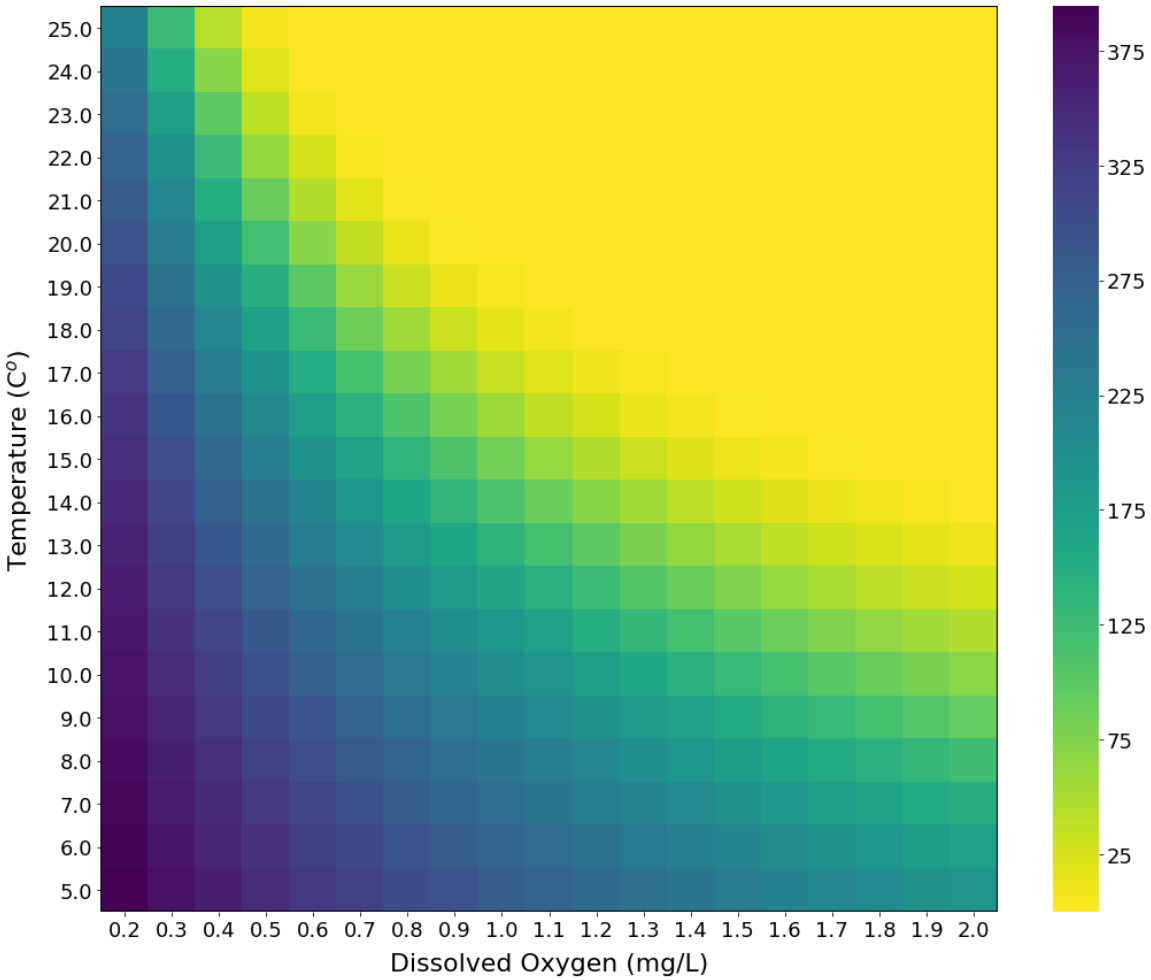


Figure 3-12 The combined effect of dissolved oxygen and temperature on NH_4^+ effluent concentrations (mg N.L-1) as simulated with the calibrated model for an IFAS-SBR reactor under 4 subcycles operation at an influent NH_4^+ concentration of influent NH_4^+ ranging from (700 – 1000 mg $N.L^{-1}$).

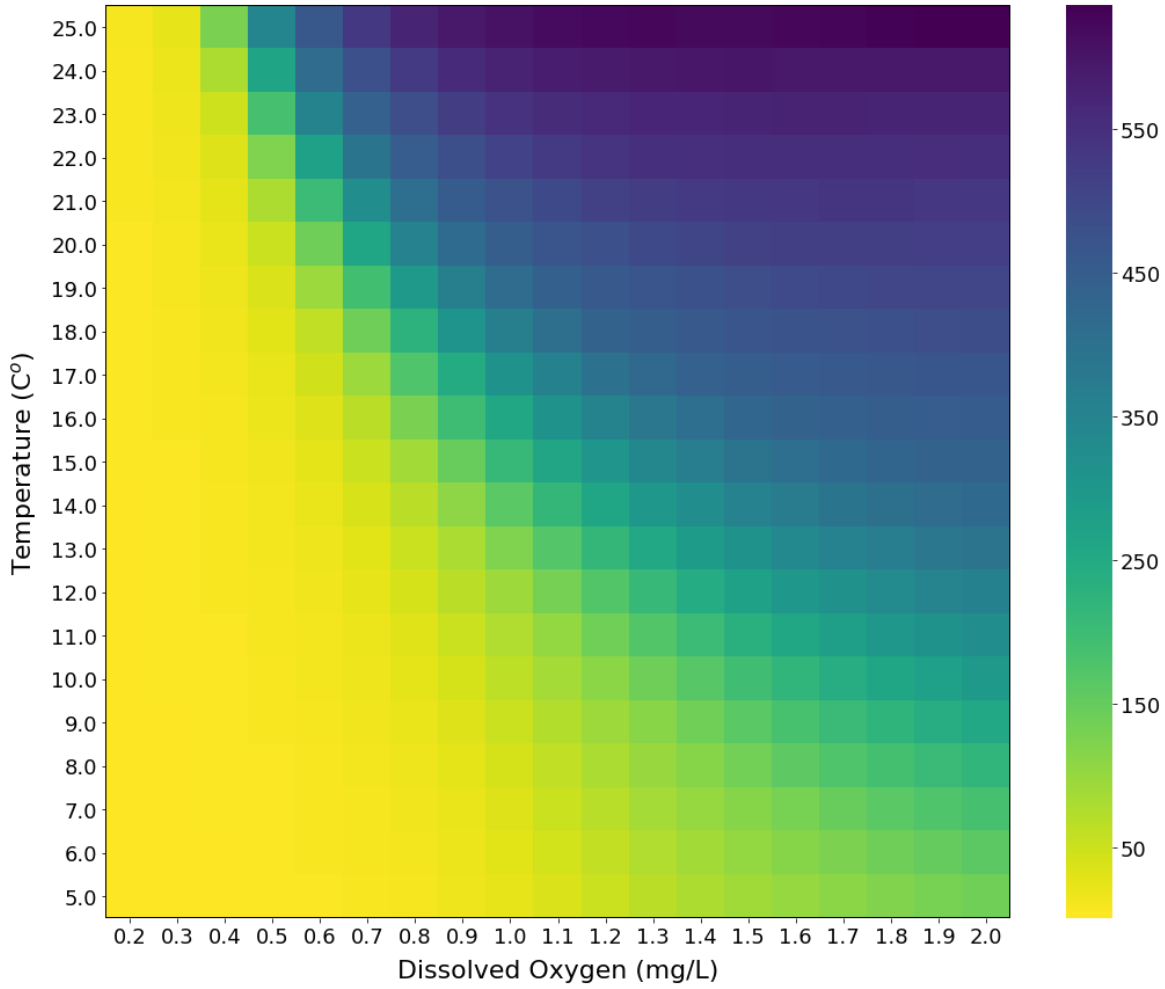


Figure 3-13 The combined effect of dissolved oxygen and temperature on the average N_2O concentration in the headspace (ppm) as simulated with the calibrated model for an IFAS-SBR reactor under 4 subcycles operation at an influent NH_4^+concentration of influent NH_4^+ranging from (700 – 1000 mg N.L-1).

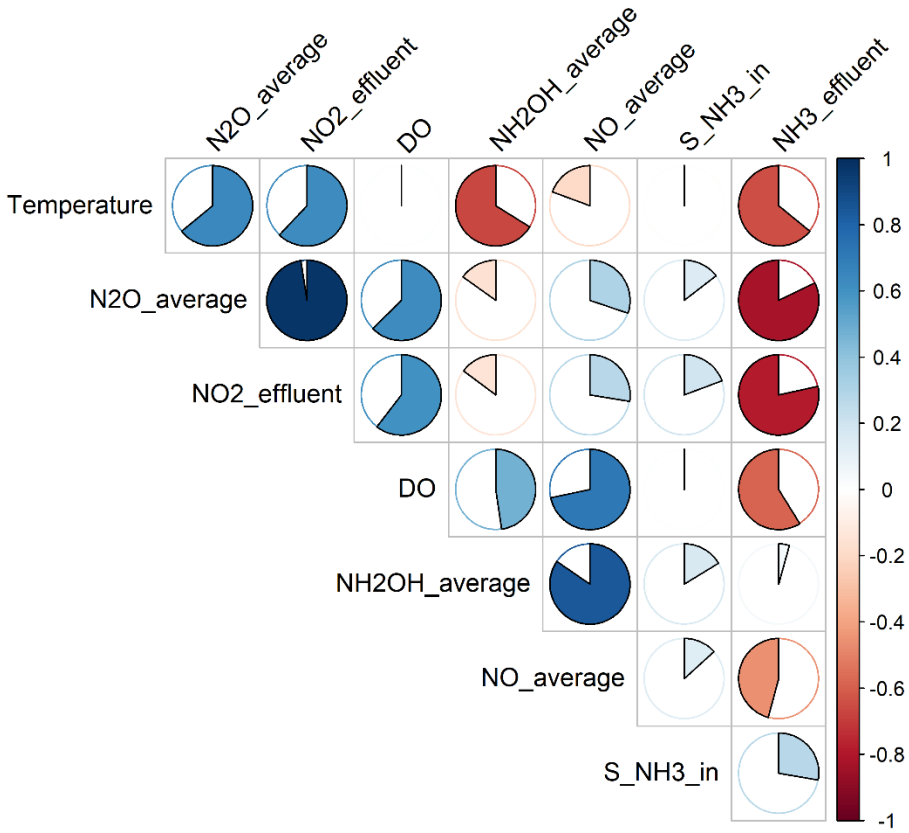


Figure 3-14 Correlation between variables obtained from scenario analysis simulations

3.4 Summary

A mechanistic model for IFAS-SBR reactor was calibrated with detailed dynamic data, including N_2O emissions. The model was able to predict the N_2O dynamics in the calibration data in which the SBR cycle was composed of four aerated/non-aerated alternating subcycles, and partially qualitatively described the N_2O emission dynamics during the validation data when the SBR cycle was operated with seven alternating aerated/non-aerated subcycles. The model has been used to reveal insights of N_2O emissions dynamics and pathway contributions given the hybrid nature of the reactor biomass (biofilm/suspended) and the dynamic operation. Most of the N_2O was produced during the aerated phase of the reactor operation through the nitrifier denitrification

pathway given the low dissolved oxygen concentration (≈ 0.60 mg/L) during aerated periods. The low levels of dissolved oxygen were sufficient to slow down heterotrophic anoxic reactions during aerated periods. Consequently, unlike in non-aerated periods, this reduction in activity decreased the competition for nitrite, facilitating its consumption via the nitrifier denitrification pathway and producing N_2O. Moreover, biofilm was shown to be more contributing to N_2O production than flocculent sludge in the studied reactor.

As the model included the effect of temperature on both mass-transfer and microbial activity, it was used to provide insights on the optimization of operational strategies to achieve both N_2O mitigation and NH_4^+ removal during seasonal temperature changes. The reason of the potential increase in N_2O emissions during winter time might be attributed to the potential increase of the DO setpoint necessary for achieving nutrient removal. This makes it challenging to achieve optimal operation for both NH_4^+ removal and N_2O emission mitigation in nitritation-denitritation reactors treating nitrogen-rich wastewater. A balance between NH_4^+ removal and N_2O emissions should be focused on during the operation of reactors with alternating aeration instead of controlling DO solely based on NH_4^+ removal. Thus, a multi-objective controller is recommended.

Chapter 4 - Machine learning for modelling nitrous oxide emissions from wastewater treatment plants: Aligning model performance, complexity, and interpretability *

4.1 Introduction

4.1.1 Limitations of mechanistic models

The advancements of mechanistic models describing N_2O emissions from biological nitrogen removal processes were driven by the hypothesis that a comprehensive model structure would be capable of capturing complex patterns in full-scale applications. However, this often results in extremely complex and over-parametrized models (Domingo-Félez and Smets, 2016), leading to a challenging calibration and validation process with high variability in the estimated parameters across different studies (Vasilaki et al., 2018). For example, reported values in literature for affinity constants of two key variables in N_2O production pathways: nitric oxide (NO) and nitrite (NO_2^-), varied with two orders of magnitude (Domingo-Félez and Smets, 2016). Furthermore, efficient calibration of mechanistic models may require monitoring of pathway intermediates that are either difficult or expensive to measure in full-scale WWTPs such as hydroxylamine (NH_2OH), and nitric oxide (NO), in addition to ammonia (NH_4-N), nitrite (NO_2^-), nitrate (NO_3^-), and nitrous oxide (N_2O).

---

* A modified version of this chapter has been published in September, 2023 in *Water Research* Journal

Furthermore, efficient calibration of mechanistic models would ideally require monitoring of pathway intermediates that are either difficult or expensive to measure in full-scale WWTPs such as hydroxylamine (NH_2OH), and nitric oxide (NO), in addition to ammonia (NH_4-N), nitrite (NO_2^-), nitrate (NO_3^-), and nitrous oxide (N_2O). This can result in increased complexity and cost to the calibration process of mechanistic models to get a reliable quantitative emissions estimation. These challenges limit the usage of mechanistic models for preliminary estimation of emissions in future applications as they fail to provide process modelling that accurately describes N_2O emissions, and supports a mitigation plan. Although mechanistic models are favorable because the interpretability and trust in model results that are generally inherent in the knowledge-based models, mechanistic models for N_2O are still struggling to provide accurate predictions, still limiting their applicability.

4.1.2 Machine learning as a potential candidate

Given the limitations of mechanistic modelling, and the large operational data available in WWTPs from both sensors and laboratories, machine learning holds significant promise for N_2O emissions prediction. Machine learning models are versatile and powerful, providing the ability to capture non-linear and complex relationships in the data. However, accuracy is not often enough for process models, especially when insights into the root causes of the problem are needed such as in the case of N_2O emissions modelling. By choosing an interpretable model, machine learning can be utilized for process modelling serving not only to describe N_2O emissions but also to identify crucial factors associated with emissions, and suggest appropriate corrective actions during dynamic conditions as a part of N_2O emissions mitigation plan. Furthermore, opting for a computationally efficient and less complex ML model offers enables real-time online monitoring and prediction of

N_2O emissions, which could be deployed in soft sensors. A lack of collaboration with statisticians and computer scientists in the development of research in the area of wastewater data or information technology were also noted (Corominas et al., 2018). This increases the needs for further investigation in the machine learning methods to create methods more tailored to data generated from WWTPs. Thus, ML models should be seen through a more comprehensive vision that combines process modelling and understanding under the same framework.

A more holistic evaluation of ML models is still needed to yield a model selection for N_2O prediction application. The evaluation criteria should help achieving a high accuracy using relatively simple models to be used in the application of N_2O emission soft sensors. Previous studies lack such a comprehensive approach that includes not only a comprehensive prediction accuracy, but also interpretability, model complexity, generalization ability, and computational speed. This can help use ML efficiently to provide reasonable insights on the primary sources of N_2O emissions, contribute to providing a mitigation plan, in addition to the regular task of N_2O emissions prediction. Only limited number of studies focused on the quantitative prediction of N_2O emissions using ML-supervised regression models based on full-scale long-term (Hwangbo et al., 2020, 2021). Hwangbo et al. (2021) used a deep neural network (DNN) for process modelling of the N_2O emissions, and long short-term memory (LSTM) for N_2O forecasting, based on full-scale facility data. Despite the high reported accuracy for process modelling in this study, DNN is still a highly complex and a low interpretable model (Montavon et al., 2017).

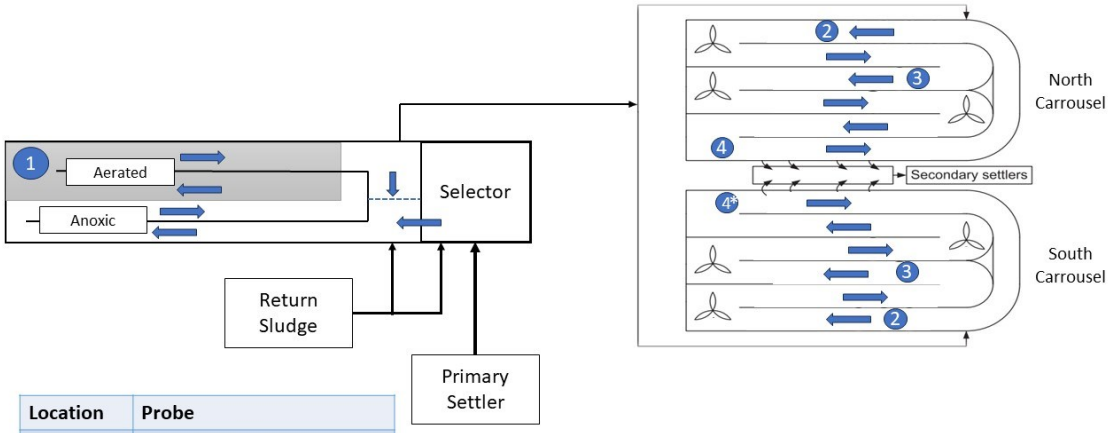
The overall objective of this chapter is to introduce a comprehensive ML-based modelling framework that encompasses efficient data preprocessing, ML model selection, and input feature reduction. Rather than advocating for a universal model, this paper offers a tailored, fit-for-purpose approach that adapts to the specific context of each case study. Data characteristics are therefore incorporated through exploratory data analysis and a robust data pre-treatment methodology. Model evaluation is approached from a holistic perspective, encompassing accuracy metrics, model complexity, computational power, and interpretability. Additionally, the trade-off between data acquisition associated costs and model performance is addressed by implementing an input feature selection technique. This study applies the framework to various models grounding a dataset obtained from a full-scale N_2O monitoring campaign conducted at a reactor performing BNR with available sensor data.

4.2 Approach and Methodology

4.2.1 Raw data and process description

This study utilized data gathered by Daelman et al., (2015) from a 16-month monitoring campaign at the Kralingseveer WWTP in the Netherlands. The plant treats 80,000 cubic meters of domestic wastewater daily, and features a plug-flow reactor followed by two parallel carrousel reactors. The focus of this work is solely on the northern carrousel reactor that included alternating anoxic/aerobic zones, with aeration achieved through surface aerators (Figure 4.1). The reactor was covered, and the off-gas was funneled into ducts, directed to a gas analyzer, and the N_2O levels were measured. The reactor was also equipped with NH_4^+-N [g/m^3], NO_3^--N [g/m^3], NO_2^--N [g/m^3], temperature [^oC], total suspended solids (TSS) [g/m^3], along with three dissolved oxygen (DO) [g/m^3] probes at

different locations (DO1, DO2, DO3). Concentrations of NH_4^+-N [g/m^3], NO_3^--N [g/m^3], and DO [g/m^3] at the end of the plug-flow reactor were also available, representing the corresponding influent concentrations to the north carrousel. Moreover, influent flow rate [m^3/d] and NH_4^+-N loading [kg/d] to the north carrousel were available.



Location	Probe
1	NH_4-N, NO_3-N, DO
2	DO1, TSS
3	DO2, Temperature
4	DO3, NH_4-N, NO_3-N,NO_2-N
4*	DO3, NH_4-N, NO_3-N

Figure 4-1 Layout of Kralingseveer WWTP modified from Vasilaki et al. (2018)

Descriptive analysis (Table 4.1) was employed to ascertain the attributes of each examined feature concerning their central inclination, to reveal the position of the distribution for each variable, spread to quantify the dataset's variation, and distribution to portray the symmetry of each feature.

Table 4-1 Statistical analysis of the dataset

Variable	Unit	Mean	Median	Min	Max	Q_{25}	Q_{75}	Skewness	Kurtosis	Variance Inflation Factor
NH_4- N	$gN\,m^{-3}$	1.67	1.12	0.00	24.62	0.85	1.39	4.96	29.36	2.75
NO_3^--N	$gN\,m^{-3}$	5.28	4.35	0.00	22.60	1.27	8.22	0.84	0.14	4.99
NO_2^--N	$gN\,m^{-3}$	1.13	0.74	0.00	11.51	0.29	1.59	1.89	4.59	4.885
DO1	$gN\,m^{-3}$	0.55	0.01	0.00	4.99	0.00	1.12	1.90	4.06	3.142
DO2	$gN\,m^{-3}$	0.63	0.14	0.00	4.99	0.00	0.95	1.67	2.46	3.460
DO3	$gN\,m^{-3}$	2.03	2.06	0.00	3.71	1.52	2.49	0.10	-0.32	17.804
TSS	$gN\,m^{-3}$	4.13	4.03	1.04	9.99	3.71	4.32	2.89	15.05	23.749
TEMP	℃	17.53	18.59	9.30	21.40	15.40	19.88	-0.69	-0.75	50.874
NO_3^-_In	$gN\,m^{-3}$	2.77	2.19	0.08	12.99	0.84	3.98	1.18	1.05	5.834
DO_In	$gN\,m^{-3}$	2.45	2.45	0.14	4.99	2.23	2.57	0.91	4.24	16.137
NH_4_loading	$kg\,N\,d^{-1}$	99.87	80.46	0.62	770.5	46.47	125	2.48	9.34	6.040
Flow rate	m^3h^{-1}	3850	3174.04	280	14371	2484	4096	2.24	5.02	8.641
NH_4_In	$gN\,m^{-3}$	11.60	11.25	0.12	33.64	6.94	15.55	0.37	-0.28	12.155
N_2O-N	$kg\,N\,h^{-1}$	68.68	10.67	0.00	556	0.03	108.7	1.68	2.14	__

4.2.2 Data pre-processing

4.2.2.1 Data synchronization and splitting

Data were collected from sensors with varying frequencies across different features, ranging from 10 to 30 minutes. To ensure that all the features are reported at the same time stamp, resampling and merging steps were employed. Firstly, each feature's dataset was

individually resampled to have a uniform frequency of 30 minutes, representing the lowest common frequency. This resampling process harmonized the time intervals across all features, facilitating further merging of features using an outer join method. Moreover, data instances with missing values were removed from the final dataset. The methods employed utilized all available features, avoided data imputation and interpolation, and maintained a practical frequency for N_2O emission reporting.

The pre-processed dataset was then randomly split into training and testing sets using a standard ratio of 75% for training and 25% for testing. The training data was further divided into k-fold cross-validation (CV) splits to optimize model hyperparameters effectively. Lastly, the dataset was scaled using normalization, keeping all the data in the range [0,1].

4.2.2.2 Outlier detection

The Mahalanobis distance was utilized for multivariate outlier detection, taking into account the correlations between variables by incorporating their covariance matrix. The process involves fitting the most appropriate statistical distribution model to the distribution of calculated distances, which is determined by comparing various distribution models and selecting the one with the least residual sum of squares (RSS).

To enhance the robustness of the selection process, a range of statistical distributions is tested, and the distribution yielding the lowest RSS is chosen as the most suitable model. After selecting the best distribution, outliers are identified based on their statistical significance in relation to the chosen distribution. Data points exhibiting a significance level of less than 0.01 are considered potential outliers.

The Mahalanobis distance calculation involves the following equations:

$$S = \frac{1}{n-1}\sum(X-M)(X-M)^T \quad 4.1$$

$$D(x) = \sqrt{(X-M)^T S^{-1}(X-M)} \quad 4.2$$

Here, S is the covariance matrix, n represents the number of observations, X is the vector of the observations for a particular variable, M is the mean vector of the dataset, and T signifies the transpose operation. $D(x)$ represents the Mahalanobis distances calculated to the mean of the dataset. By incorporating the covariance matrix, the Mahalanobis distance effectively accounts for the relationships between variables, thus providing a more accurate identification of multivariate outliers.

4.2.2.3 Feature scaling

Model input features have different scales that can result in a different contribution to the regression model output depending on the specific scale of each feature, which can affect the accuracy of several models. Therefore, the entire dataset features and output have been scaled using the minimum-maximum normalization method that results in the scaled data being in the range [0, 1], according to the following equation:

$$\frac{x_i - x_{min}}{x_{max} - x_{min}} \quad 4.3$$

where $x_{min}$and x_{max} are the minimum and maximum value of each feature.

4.2.3 Machine learning models

4.2.3.1 Model development and optimization

This study included ML models of varying complexity, including the intuitive k-Nearest Neighbors (kNN) algorithm that measures distances between data points and identifies

similarities. A Decision Tree (DT) was also developed for its simplicity as a branching structure that recursively divides data based on selected feature conditions. Additionally, ensemble models based on decision trees were also employed using several aggregation algorithms: bagging (Random Forest), gradient boosting (XGBoost), and adaptive boosting (AdaBoost). A benchmark Deep Neural Network was developed for comparison. Hyperparameters of the employed models were optimized using a grid search algorithm. Grid search explores a pre-defined grid space of hyperparameters to find the set of values that maximizes the model mean performance on the k-fold CV subsets. The list of optimized values of hyperparameters for each model, along with the corresponding ranges used in the grids as input for the algorithm, is provided Tables 4.2, 4.3, 4.4, 4.5, 4.6, and 4.7. All modelling steps in this study were implemented in Python 3.9 using open-source packages such as NumPy, Pandas, Scikit-learn, Matplotlib, Tensorflow, and SciPy.

Table 4-2 Hyperparameters of decision tree model

Parameter	**Minimum**	**Maximum**	**Selected**
Tree depth	1	40	30
Minimum number of samples in leaf	2	15	4
Minimum number of samples to split node	2	15	9

Table 4-3 Hyperparameters of k-NN model

Parameter	**Minimum**	**Maximum**	**Selected**
N_neighbors	1	10	2

Table 4-4 Hyperparameters of random forest model

Parameter	Minimum	Maximum	Selected
Number of estimators (trees)	50	1500	1250
Splitting criterion	(Squared error, absolute error)		Squared error
Max depth of base trees	1	40	20

Table 4-5 Hyperparameters of XGBoost model

Parameter	Minimum	Maximum	Selected
Number of estimators (trees)	50	1500	1200
Learning rate	0.001	1.5	0.01
Minimum child weight	3	6	5
Subsample	0.5	0.8	0.7
Max depth of trees	1	40	18
Splitting criterion	(Linear error, squared error, exponential error)		squared error

Table 4-6 Hyperparameters of AdaBoost model

Parameter	Minimum	Maximum	Selected
Number of estimators (trees)	50	1500	800
Learning rate	0.001	1	1
Max depth of trees	1	40	18

Loss function	(Linear error, squared error, exponential error)	square

Table 4-7 Hyperparameters of Deep Neural Netwo model

Parameter	Minimum	Maximum	Selected
Number of hidden layers	1	8	5
Number of neurons in layer 1	20	200	160
Number of neurons in layer 2	20	200	120
Number of neurons in layer 3	20	200	80
Number of neurons in layer 4	20	200	40
Number of neurons in layer 4	20	200	20
Learning rate	0.0001	1	0.001
Loss function			Mean-squarer error
Activation function			Rectified linear unit
Optimization algorithm			Adam
Number of epochs			1000 (with early stopping)

4.2.3.2 Preliminary model filtration

The Akaike Information Criterion (AIC) was used to select a single model from each of the three groups, to focus on the most efficient models. AIC assesses the trade-off between the goodness of fit and the complexity of the model, making it a reasonable criterion for the preliminary filtration of models.

4.2.3.3 Model performance evaluation

The accuracy of N_2O estimations has been comprehensively evaluated using both statistical metrics and dynamic process evaluation using the out-of-sample test set. Regression-based metrics included the coefficient of determination (R^2) – used to estimate the proportion of variance explained by the model – along with mean absolute error (MAE) and root mean squared error (RMSE), both of which were employed to provide information about the magnitude of error across different range of N_2O emissions. This was supported by a detailed residual analysis to provide deeper insights on the model performance and detect potential areas of improvement. Furthermore, to evaluate the ability of the models to depict process dynamics, a zoomed time series plots were used to compare model predictions with the real N_2O emissions. A cross-correlation analysis was also conducted to ensure that there is no time lag between the real ad estimated profiles of N_2O emissions. Lastly, the degree of potential overfitting - where the model cannot generalize to new unseen data - was tested using the learning curves to gain insights about the potential improvement of model ability when data volume increases.

In addition to accuracy evaluation, a comparative analysis of training and prediction times across models was conducted. The experiments were conducted on a system running Windows 11 with an Intel(R) Core (TM) i7-9750H CPU, operating at a base frequency of 2.60 GHz. The system was equipped with 16.0 GB of RAM, and it operated on a 64-bit architecture.

4.2.3.4 Feature selection

A two-stage feature selection algorithm has been utilized to select the optimal set of features. The initial phase involves an adjusted version of the maximum relevance

minimum redundancy (mRMR) algorithm (Ding and Peng, 2003), which identifies the top m features that exhibit the highest relevance to the target variable while maintaining minimal redundancy among the previously chosen features, regardless of the used model. Relevance was assessed using mutual information based on the following formula:

$$MI(x;y) = \iint p(x,y)\, log \frac{p(x,y)}{p(x)p(y)} dxdy \quad 4.4$$

Where $p(x,y)$ is the joint probability function, $p(x)$ and $p(y)$ are the marginal probability distribution functions of X and Y. Redundancy, on the other hand, was defined using the variance inflation factor (VIF) - a statistical metric that evaluates the severity of multicollinearity by quantifying the degree to which the variance of a specific independent variable's estimated coefficient is inflated due to multicollinearity. For each feature, a linear regression was performed using this feature as the dependent variable and all other features as independent variables. The VIF of the feature is then calculated using the formula:

$$VIF = \frac{1}{1 - R^2} \quad 4.5$$

To pinpoint the optimal number of features m, the second phase of this algorithm incorporates a wrapper approach that applies the modified mRMR technique within every model to determine the Pareto Front solutions. These solutions embody the ideal balance between model performance and number of input features.

4.2.3.5 Permutation Feature importance

Permutation feature importance is a technique utilized to evaluate the significance of input features by permuting the values of each feature and measuring the resulting impact on

model performance. The degradation in model performance, quantified by mean squared error, is then attributed to the importance of the feature. Thus, feature importance is a model-agnostic measure, as it does not rely on any specific model characteristics or assumptions.

4.3 Results and Discussion

4.3.1 Data pre-processing

The study is based on a final dataset with more than 10,000 datapoints over a time window of 300 days, which is the outcome of the data synchronization process applied to 13 features and the target variable (N_2O emissions). Figure 4.2 presents the normalized box plots of all features, facilitating a common x-axis comparison. Overall, a heavy-tailed distribution can be observed for the majority of these features, a finding that is supported by the descriptive analysis of the dataset shown in Table 4.1. For instance, the skewness of NH_4, TSS, DO1, and NO_2^- features were 4.96, 2.89, 1.90, and 1.89, respectively. This is not uncommon in WWTPs where data collected from complex biological treatment processes may exhibit a skewed or bimodal distribution due to operational or seasonal changes in the process. Widely-adopted univariate statistical methods for outlier detection such as the interquartile range (IQR) presented by box plots, or Z-score methods would remove a large portion of the data (about 10% of the dataset in this study) with valuable information. This does not agree with the purpose of modelling in the wastewater domain, where the goal is modelling of the entire spectrum including extreme values that are likely to happen. Thus, these methods must be avoided when the data is highly-skewed (Hubert and Vandervieren, 2008).

A multivariate outlier detection method was implemented using the Mahalanobis distance. The detected outliers comprised less than 1% of the dataset. Among various statistical distributions, the distribution of the calculated Mahalanobis distances was found to fit the generalized extreme value distribution as shown in Figure 4.3. Modelling the statistical distribution of the calculated distances allowed a parametric detection of the outliers through selection of a threshold based on statistical significance. Data instances with Mahalanobis distance higher than the specified threshold (significance = 0.01) was identified as outliers. Using the Mahalanobis distance also accounted for dependence between features in the dataset, which is aligned with the nature of data collected from complex biological treatment processes, where correlations between variables are present. This multivariate approach can help to identify unusual or abnormal events from a process perspective, where traditional univariate outlier detection methods fail. Data points deemed as outliers in the multivariate context were not necessarily extreme values for the individual variables, showing the discrepancies between univariate and multivariate outlier detection methods.

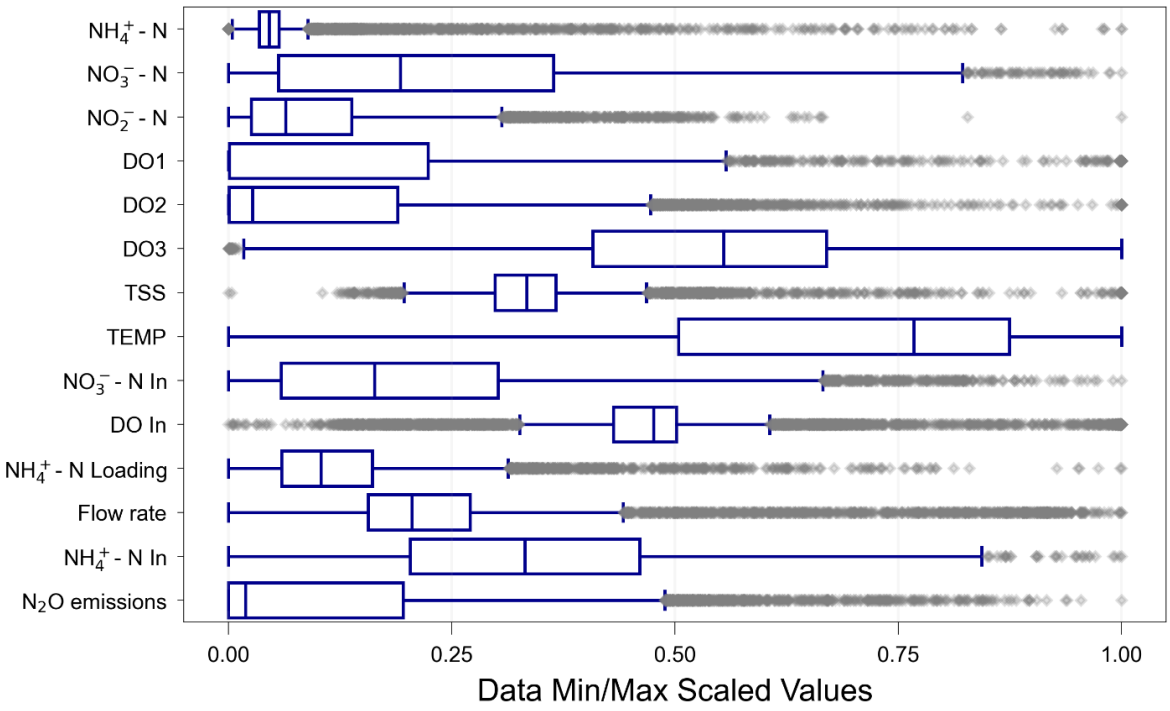


Figure 4-2 Box plots of the normalized dataset features

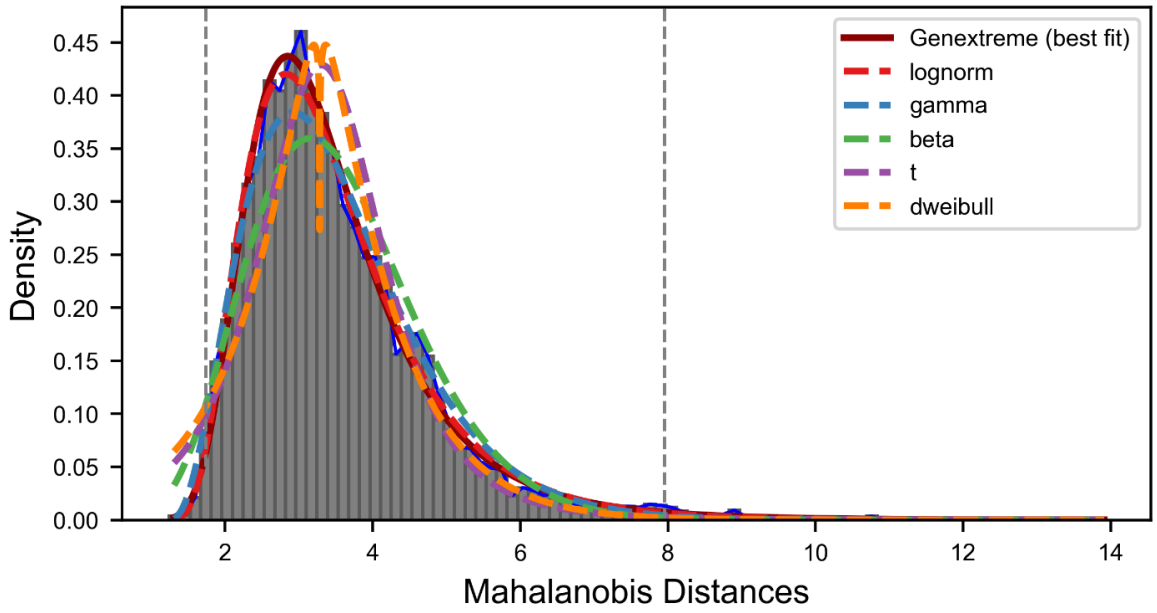


Figure 4-3 Distribution of the calculated Mahalanobis distances with the threshold line (at significance = 0.01) and the best statistical distributions

4.3.2 A holistic approach for modelling N_2O emissions from WWTPs

Previous efforts on modelling N_2O emissions using ML models have focused solely on prediction accuracy. Quite to the contrary, the present study proposes a holistic approach that strikes a balance between model interpretability, complexity, needed computational power, in addition to a comprehensive performance evaluation of the model accuracy. The urge for such a holistic approach stems from the uncertainty around N_2O production pathways, the complexity of wastewater treatment operations and multiple objectives of process modelling. The developed model shall serve not only as an online soft sensor that can accurately and efficiently predict N_2O emissions with optimal computational complexity and speed, but also be interpretable to inform operators with key contributors, and support mitigation activities.

A comprehensive accuracy evaluation is needed to provide a better understanding of model performance. Reported models in literature were evaluated majorly using the coefficient of determination (R^2) which can be misleading, as illustrated in the work of Hwangbo et al. (2021), where R^2 was only used for reporting model performance, neglecting potential high prediction errors in some output ranges. In this study, the accuracy of models was evaluated using R^2, MAE, RMSE, a residual analysis. R^2 was used to estimate the model's ability to explain variance in predicted N_2O, while MAE, RMSE, and residual analysis were used to provide a clear overview of the model prediction errors across the entire range of N_2O values. Furthermore, including time series plots and cross-correlation analysis were conducted to evaluate the performance of models in terms of process dynamics. For that purpose, detailed time series plots were used to evaluate the ability of models to capture emission patterns, including local minima and maxima. This was supported by a cross

correlation analysis, ensuring there is no lag between real emissions and model estimations. Furthermore, the extent of overfitting was evaluated, in which a model may adhere too closely to the training data's noise and fail to generalize to unseen data. This was particularly considered during model development by employing the k-fold CV in hyperparameter optimization.

To provide operators and decision-makers with action plan based on a soft sensor, trust in the model predictions is needed. As a prerequisite to trust, inclusion of interpretability to model selection and evaluation criteria is necessary. According to Lipton, (2018), model properties that enable interpretability can be categorized under model transparency, and post-hoc interpretability. Unlike black-box models, a transparent model provides an understanding of the underlined mechanism. Models such as deep neural networks (DNN) can be seen as black-box and less transparent models in comparison with simpler models such as decision trees, where the model can be presented to the user through visual artifacts (Ribeiro et al., 2016). However, a decision tree, for instance, may grow too deep that makes the user inference much challenging. Such cases are likely to happen when the data are high-dimensional or the model is allowed to become too complex. This necessitates the importance of considering both model complexity, and feature reduction while adoption of ML in WWTP applications.

Furthermore, using the model in a soft-sensor application would require an attention to the computational power of the model, generally referring to the resources required to perform training and prediction using the model. These resources can include processing time, memory, storage, and the ability to leverage parallelism. In this study, processing time was used as a measurable metric for model computational power. A faster model is generally

more desirable for real-time monitoring applications, allowing for more frequent updates and high-resolution predictions (Torfs et al., 2022).

In addition to the advantages of a lower data dimensionality – in terms of using fewer model features- to model complexity and interpretability, there is also a need to reduce the number of input features from a wastewater process perspective. An overlooked fact in the previous studies is that increasing the number of input features needs more intensive data collection through either new routine measurements and/or sensor installations. This translates into higher cost, operational complexity, and higher implied uncertainty. Thus, a careful consideration of the number of model input features and data dimensionality should be part of a data-driven model development. A trade-off between model performance and the number of features is required, to ensure there are no unnecessary, redundant, or less-important features used in the model.

Therefore, in this study, a new definition for model comparison and selection, considering performance, generalization, complexity, data acquisition associated costs, and computational power, was employed, rather than solely focusing on accuracy metrics. First, a preliminary selection of models was performed. ML models were classified into three categories based on their varying degrees of transparency, and complexity. The first category encompasses simple models like kNN and Decision Trees, which provide easily understandable results and require relatively low computational resources. The second category includes moderately complex, partially transparent ensemble models such as AdaBoost, XGBoost, and Random Forest, which can achieve higher performance while maintaining a degree of interpretability and insight into the models' inner workings. The third category consists of complex, high-performance models like Deep Neural Networks

(DNN), which offer powerful learning capabilities at the cost of lower interpretability, transparency, and increased computational demands. Then, another model selection step was employed by choosing only a single model from each category based on a combination of accuracy and model complexity (using AIC), in which then a comprehensive evaluation was applied on the selected three models.

4.3.3 Model selection and performance comparison

Figure 4.4 shows the actual and predicted N_2O values based on the test data using ML regression models across the three different categories. For the first group, DT achieved a lower accuracy potentially due to the piecewise constant approximation, where the model divides the input space into regions and assigns a constant value to each region. This leads to relatively low model performance when the data shows complex patterns, which is the case in the present dataset. However, the problem can be solved by using ensemble tree-based models, which were adopted in the second category of models in this study. Among the three tested ensemble methods, AdaBoost gave both a better fit to the data and showed lower errors over the different ranges of the data than XGBoost, and Random Forest models. For the third group, the DNN validated the relatively high performance previously reported by Hwangbo et al. (2020). However, it did not demonstrate the highest performance among the experimented models in this study.

Table 4.8 shows the AIC scores comparing models across each group to focus on the most effective models, finding the optimal compromise between model accuracy and complexity. For the first group, the kNN scored better than the decision tree. For the second group, AdaBoost scored a better AIC than RF and XGboost, indicating that both algorithms

needed a higher structural complexity to performance ratio, increasing the needed computational power for online applications.

Therefore, moving forward, a single model representing each group (kNN, AdaBoost, and DNN) was used in the next steps of the present study. The selected models were not compared using AIC. Models with vastly different structures can yield unreliable results, as AIC's validity hinges on the assumption of similar underlying assumptions of models.

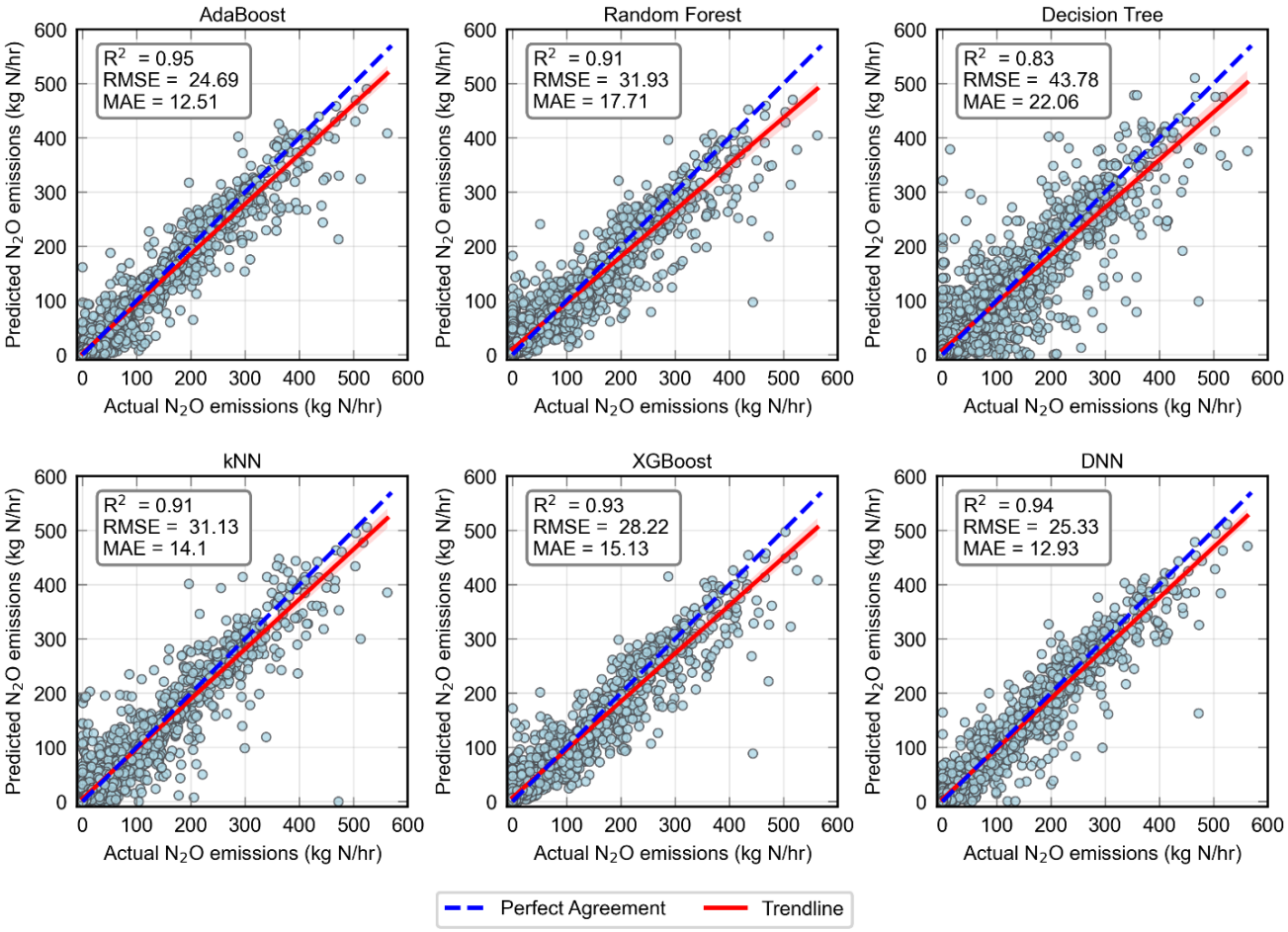


Figure 4-4 Performance of the six models on the test dataset. The graphs show the actual vs the modelled N_2O values with the actual trendline compared to the perfect agreement line ($R^2 = 1$).

Table 4-8 Comparison of AIC of models among different groups

Group	Model	AIC
Group 1	kNN	23,512
	Decision Tree	32,433
	AdaBoost	311,771
Group 2	Random Forest	961,766
	XGBoost	501,770
Group 3	DNN	121,814

4.3.4 Balancing data acquisition and model performance through feature selection

Figure 4.5 compares the kNN and AdaBoost performances for different feature subsets selected by the modified mRMR algorithm with selection based on MI. The key difference between both approaches is that mRMR considers features relevance to the target output and redundancy between features, while MI only considers the former. It can be seen that the mRMR approach exhibited a consistent increase in model performance as more features are added, which tend to level off beyond eight features.

On the other hand, the MI approach resulted in a fluctuation in the performance after adding new features in some cases, such is in kNN model (Figure 4.5b). This observation highlights the limitations of using relevance alone (represented by MI) for feature selection in datasets characterized by complex relationships and multicollinearity. While the widely-used relevance-based methods rely on calculating relevance for each feature to the target variable using various criteria (MI, Pearson correlation, F-regression), this approach may

not be optimal for developing low-cost, less-complex, and highly-interpretable models due to the existence of redundancy in the selected features. This issue is especially pronounced when dealing with complex datasets like wastewater data, where multicollinearity is frequently present (Ching et al., 2021).

Table 4.1 shows the VIF calculated for each feature in the dataset, highlighting the high multicollinearity (redundancy) in the data that needs to be treated. Including redundant features can lead to decreased interpretability and increased data collection, storage, and processing costs. This necessitates the consideration of redundancy while omitting features from the model, in which the modified version of the mRMR algorithm was utilized to address these challenges through considering both features relevance with the target variable and redundancy with other features.

As evident from Figure 4.5a and b, utilizing the modified mRMR method demonstrated a more efficient and reliable approach for feature selection. Indeed, it can be seen as considering both maximizing relevance and reducing redundancy within the dataset. For instance, in the case of the AdaBoost model (Figure 4.5a), the features selected by the modified mRMR consistently contributed to better model performance than relevance-based method (using only MI), indicating the method's efficiency. As a modification to mRMR, redundancy was considered among the data features using the VIF to provide a more comprehensive representation of multicollinearity. The algorithm begins by selecting the feature with the highest mutual information with the target variable and recursively selects features that offer the most relevance compared to redundancy, as determined by the VIF. The highest VIF values in the present dataset were attributed to Temperature, TSS, and DO3. However, the modified mRMR algorithm favored the selection of these features

over others with lower VIF such as influent flow rate, ammonia loading, and nitrate concentration in the influent.

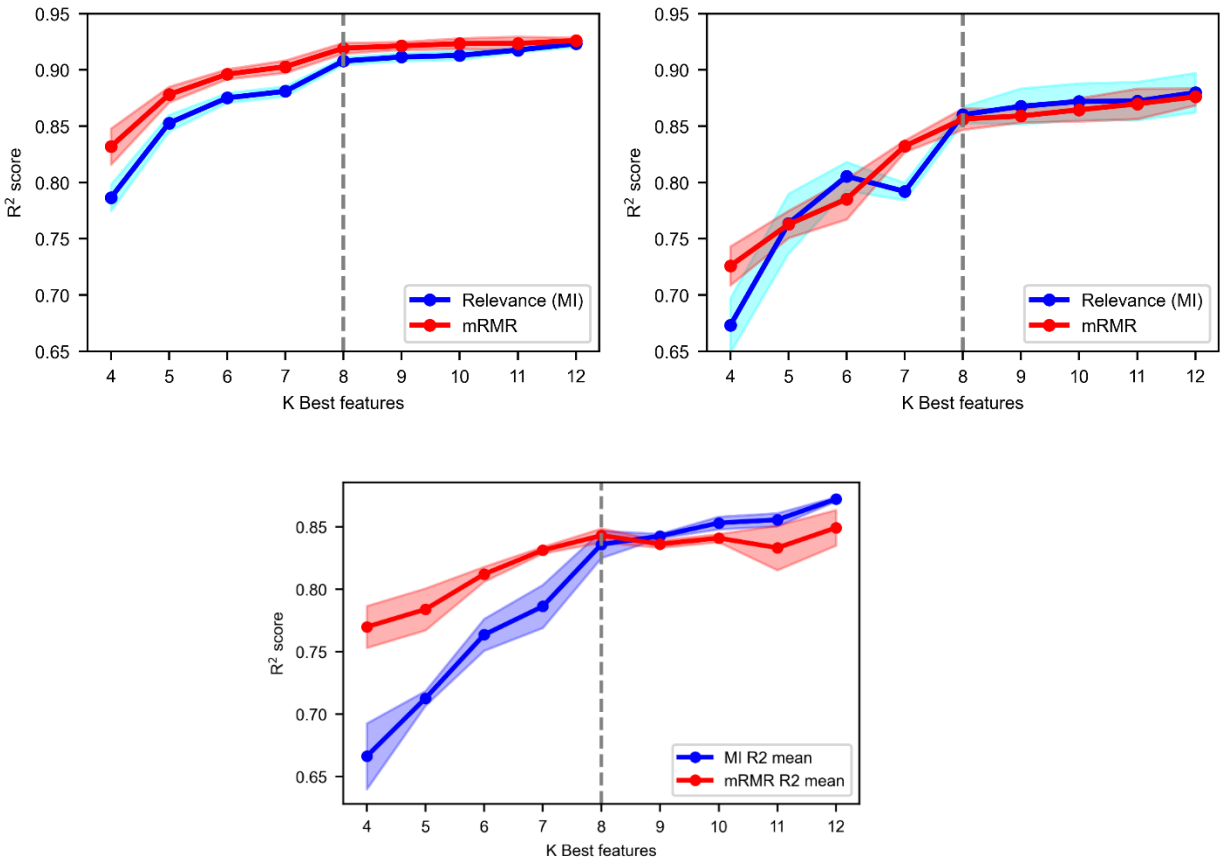


Figure 4-5 The performance of models as a result of the K best features selection: A (on the left): AdaBoost model, B (on the right): kNN model, C (to the bottom): DNN model

Table 4.9 presents the features selected by both methods (mRMR and MI) for varying numbers of features. The input feature subset chosen by mRMR not only improved model performance but was also more relevant from a process knowledge standpoint, offering a more insightful and convincing selection method for complex datasets. For instance, the top 5 selected features by the modified mRMR method were (NO_2^-, NO_3^-, NH_4, DO_1, TEMP), whereas the relevance-based method omitted NH_4 from the top five important

features. Including NH_4 aligns with the process knowledge as the concentration of NH_4 is a crucial determinant of N_2O concentration because it serves as the primary source of nitrogen in the influent. This agrees with Hwangbo et al. (2021)'s global sensitivity analysis results listing NH_4 as one of the most influential feature on N_2O emissions. Moreover, the residual NH_4 concentration offers an approximation of the proportion of nitrogen that has undergone conversion in the biological nutrient removal (BNR) process. In addition, NH_4 is identified as a precursor to hydroxylamine, which is the principal electron donor in the nitrifier nitrification pathway for N_2O production (Ni et al., 2011). This underlines the significance of NH_4 in both N_2O production pathways, suggesting that its inclusion as an input feature is essential for a data-driven model with a high level of interpretability and accuracy. The relevance-based chose the DO at two intermediate locations in the carrousel while mRMR approach selected the DO at the end of the carrousel. An indication of the optimized selection by mRMR approach is obvious when using a powerful model such as AdaBoost that is able to extract most of the possible information from the available data. At this number of features, the AdaBoost performance increased from ($R^2 = 0.85$) when using the relevance-based method to ($R^2 = 0.88$) when using mRMR.

To select the best number of features, the modified mRMR was used as a wrapper method, where each model was fed with an increasing number of the best features selected by the mRMR approach and the model performance was recorded. As shown in Figure 4.5a and b, the performance of kNN and AdaBoost has leveled off after the incorporation of eight features. Similar conclusions can be derived for the DNN, despite that mRMR exhibited a better performance than MI only for eight or less features.

Table 4-9 Comparison of the features selected based on both the relevance-based method using mutual information (MI) and the maximum relevance minimum redundancy (mRMR) method

No	Best features (mRMR)	Best features (MI)
1	TEMP	TEMP
2	TEMP, NH_4	TEMP, DO3
3	TEMP, NH_4, NO_3	TEMP, DO3, NO2
4	TEMP, NH_4, NO_3, DO1	TEMP, DO3, NO2, DO2
5	TEMP, NH_4, NO_3, DO1, DO2	TEMP, DO3, NO2, DO2, TSS
6	TEMP, NH_4, NO_3, DO1, DO2, DO3	TEMP, DO3, NO2, DO2, TSS, DO1
7	TEMP, NH_4, NO_3, DO1, DO2, DO3, TSS	TEMP, DO3, NO2, DO2, TSS, DO1, NH_4_In
8	TEMP, NH_4, NO_3, DO1, DO2, DO3, TSS, NO3_In	TEMP, DO3, NO2, DO2, TSS, DO1, NH_4_In, NO_3
9	TEMP, NH_4, NO_3, DO1, DO2, DO3, TSS, NO3_In, DO_In	TEMP, DO3, NO2, DO2, TSS, DO1, NH_4_In, NO_3, NO_3_In
10	TEMP, NH_4, NO_3, DO1, DO2, DO3, TSS, NO3_In, DO_In, NH4_loading	TEMP, DO3, NO2, DO2, TSS, DO1, NH_4_In, NO_3, NO_3_In, DO_In
11	TEMP, NH_4, NO_3, DO1, DO2, DO3, TSS, NO3_In, DO_In, NH4_loading, Influent	TEMP, DO3, NO2, DO2, TSS, DO1, NH_4_In, NO_3, NO_3_In, DO_In, NH_4
12	TEMP, NH_4, NO_3, DO1, DO2, DO3, TSS, NO3_In, DO_In, NH_4_loading, Influent, NH_4_In	TEMP, DO3, NO2, DO2, TSS, DO1, NH_4_In, NO_3, NO_3_In, DO_In, NH_4, Influent

Thus, eight features were selected as the optimal number of features to allow a trade-off between enhancing model performance and minimizing data cost. This method allowed obtaining the Pareto front solution- an ensemble of solutions offering an ideal equilibrium

for this multi-objective optimization, defined as a balance between model performance and the number of features. As shown in Table 4.10, employing an effective feature selection method resulted in about 40% reduction in the amount of data needed with reduction of model performance in terms of R^2 by only 1%. Also, a decrease by 50% was achieved in the model computational time, allowing less resources and better opportunities for incorporation in a more frequent real-time monitoring of N_2O emissions and utilization in digital twins. It is noteworthy that this does not assume generalization of specific input features to other cases. Rather, it demonstrates the framework - through application to a full-scale dataset - that researchers and engineers can employ to curtail input features across other case studies.

Table 4-10 Accuracy metrics and processing time of models before and after feature reduction (best performance highlighted in bold)

Model		Accuracy metrics			Processing time (s)	
		R^2	MAE	RMSE	Training	Prediction
kNN	Full features	0.91	14.10	31.13	**0.010**	**0.09**
	Feature reduced	0.88	18.00	37.23	**0.008**	**0.04**
AdaBoost	Full features	**0.95**	**12.51**	**24.69**	64	0.85
	Feature reduced	**0.94**	**13.12**	**26.27**	33	0.49
DNN	Full features	0.94	12.93	25.33	368	0.20
	Feature reduced	0.90	21.91	33.92	227	0.15

4.3.5 Model comprehensive evaluation

4.3.5.1 Model accuracy

Figure 4.6 depicts the dynamic evaluation of N_2O emission process modelling by the feature-reduced kNN, AdaBoost, and DNN models on the test data. The presented plots focus on two zoomed windows of five- and three-days from two different periods, highlighting distinct profiles of N_2O emissions. Figure 4.7, 4.8, and 4.9 show the N_2O prediction profile during the entire time window for the test dataset. Table 4.10 shows the relevant accuracy metrics for these models. AdaBoost exhibited the best performance across all metrics ($R^2 = 0.94$) and fitted most of the extreme high and low N_2O emissions. Although the DNN demonstrated a better overall performance than the kNN, achieving R^2 values of 0.90 and 0.88, and RMSE values of 33.92 and 37.23, respectively, MAE scores of 21.91 were recorded for DNN and 18.00 for kNN. Upon analyzing the dynamic performance and residual characteristics of both models, the discrepancy between RMSE and MAE can be attributed to the distinct overestimation patterns exhibited by each model (refer to Figure 4.6). DNN predictions exhibited higher emissions compared to actual data, albeit with smaller magnitude. These elevated emissions extended over longer time periods when contrasted to the kNN model's predictions. In contrast, the kNN model displayed fewer spikes but produced higher residual errors, which contributed more notably to RMSE than to MAE. The analysis of residuals for the three models also reveals that the most significant errors appear in instances of extremely low or high N_2O emissions. This can be attributed to the high skewness of the actual N_2O emissions, which suggests the potential for improvement through training the models with more balanced data. This pattern was less pronounced in the case of AdaBoost, which thus demonstrated the best overall

performance. Nonetheless, a potential area for improvement is in the accuracy of predicting actual peak values. Additionally, the cross-correlation analysis revealed that no lag could be observed between the actual and estimated N_2O emissions across the three models (results now shown). Both the statistical and dynamic evaluation of the three models demonstrated that AdaBoost outperformed both DNN and kNN for the overall performance, especially for describing extreme N_2O emissions. This finding agrees with the most recent survey on machine learning models performance on tabular data indicating that DNN is mostly outperformed by gradient-boosted tree ensembles for tabular supervised learning tasks (Borisov et al., 2022). Lastly, the learning curves presented in Figure 4.10 show that, while the models show perfect performance on training sets, they also performed very-well on the cross-validation sets, indicating no significant overfitting and showing that the models will still likely benefit from getting trained on more data.

The analysis for model performance conducted in this study highlights the importance of a comprehensive model evaluation, as each metric provides a unique perspective on model performance and no single metric should be used in isolation. Evaluation of model performance should include both various statistical metrics and dynamic process evaluation over the entire data window, instead of viewing only zoomed time windows. Such a comprehensive approach is absent from existing literature for N_2O emissions.

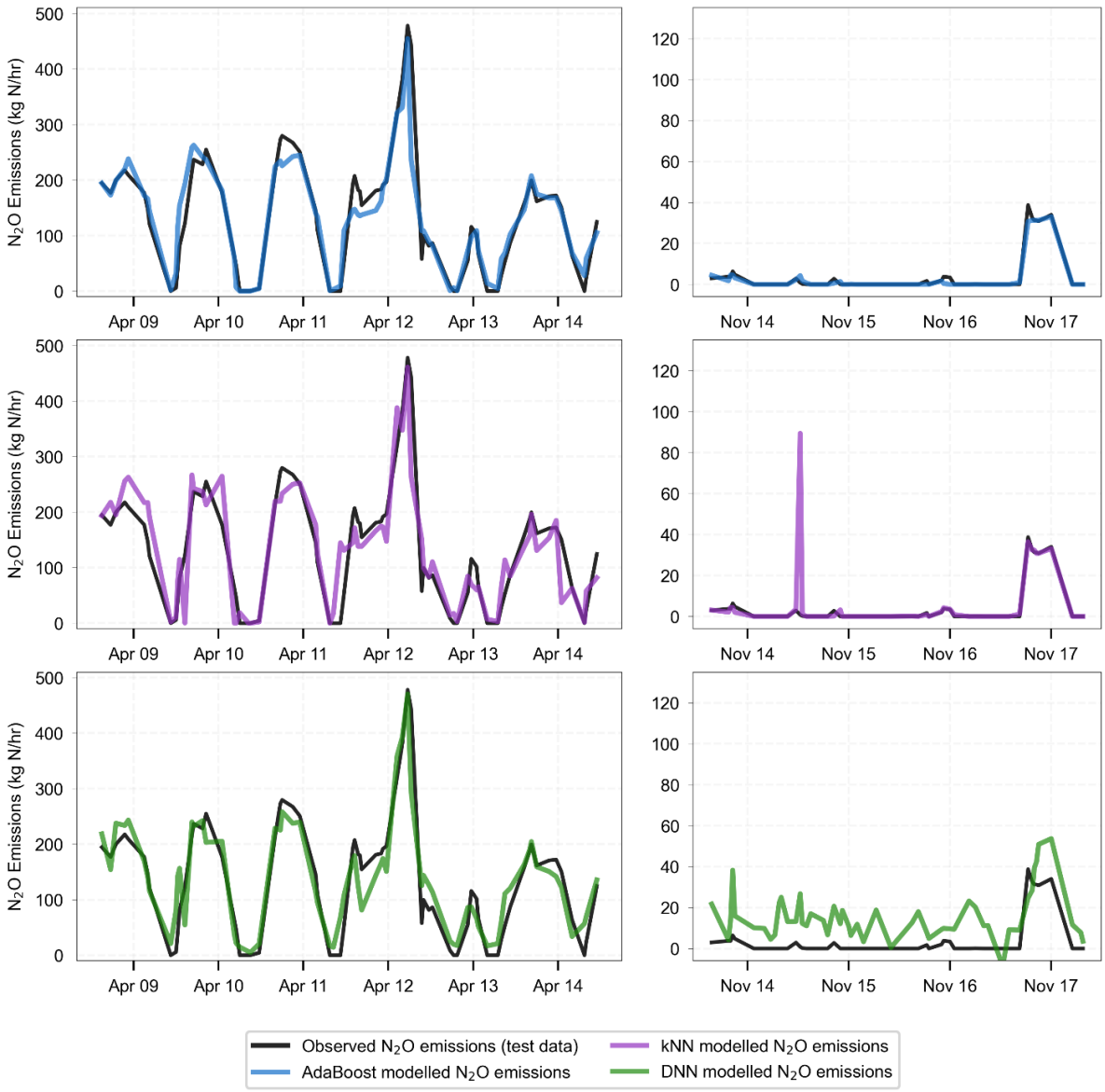
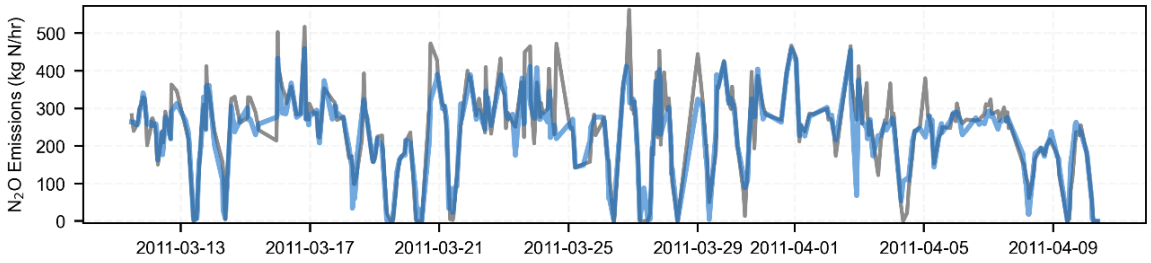
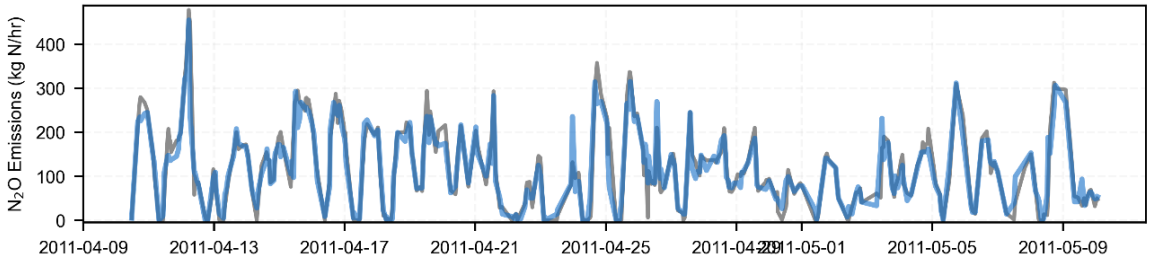
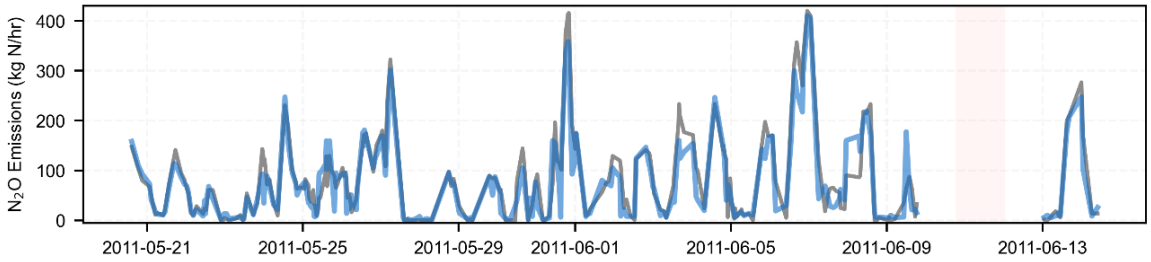
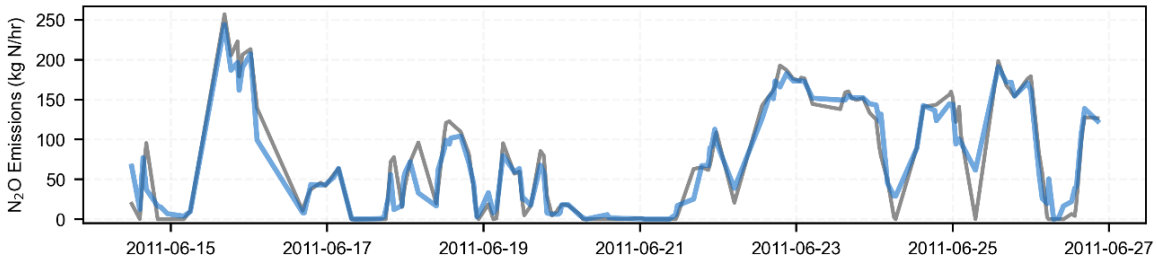


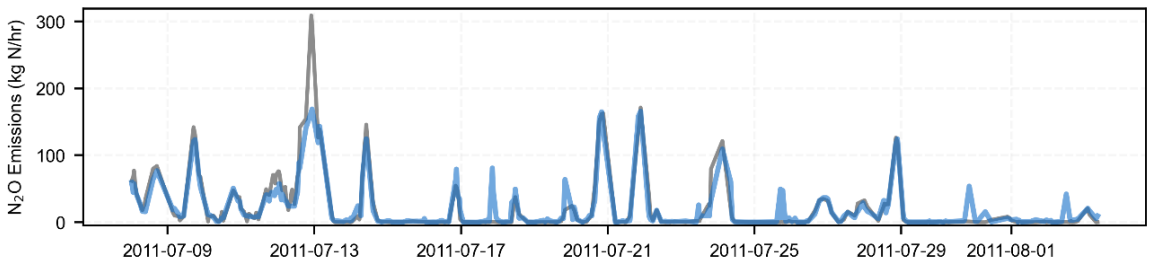
Figure 4-6 The performance of AdaBoost, kNN, and DNN models compared to the test dataset during two zoomed windows


N2O Emissions (kg N/hr)
0
100
200
300
400
500
2011-03-13
2011-03-17
2011-03-21
2011-03-25
2011-03-29
2011-04-01
2011-04-05
2011-04-09


N2O Emissions (kg N/hr)
0
100
200
300
400
2011-04-09
2011-04-13
2011-04-17
2011-04-21
2011-04-25
2011-04-29
2011-05-01
2011-05-05
2011-05-09


N2O Emissions (kg N/hr)
0
100
200
300
400
2011-05-21
2011-05-25
2011-05-29
2011-06-01
2011-06-05
2011-06-09
2011-06-13


N2O Emissions (kg N/hr)
0
50
100
150
200
250
2011-06-15
2011-06-17
2011-06-19
2011-06-21
2011-06-23
2011-06-25
2011-06-27


N2O Emissions (kg N/hr)
0
100
200
300
2011-07-09
2011-07-13
2011-07-17
2011-07-21
2011-07-25
2011-07-29
2011-08-01

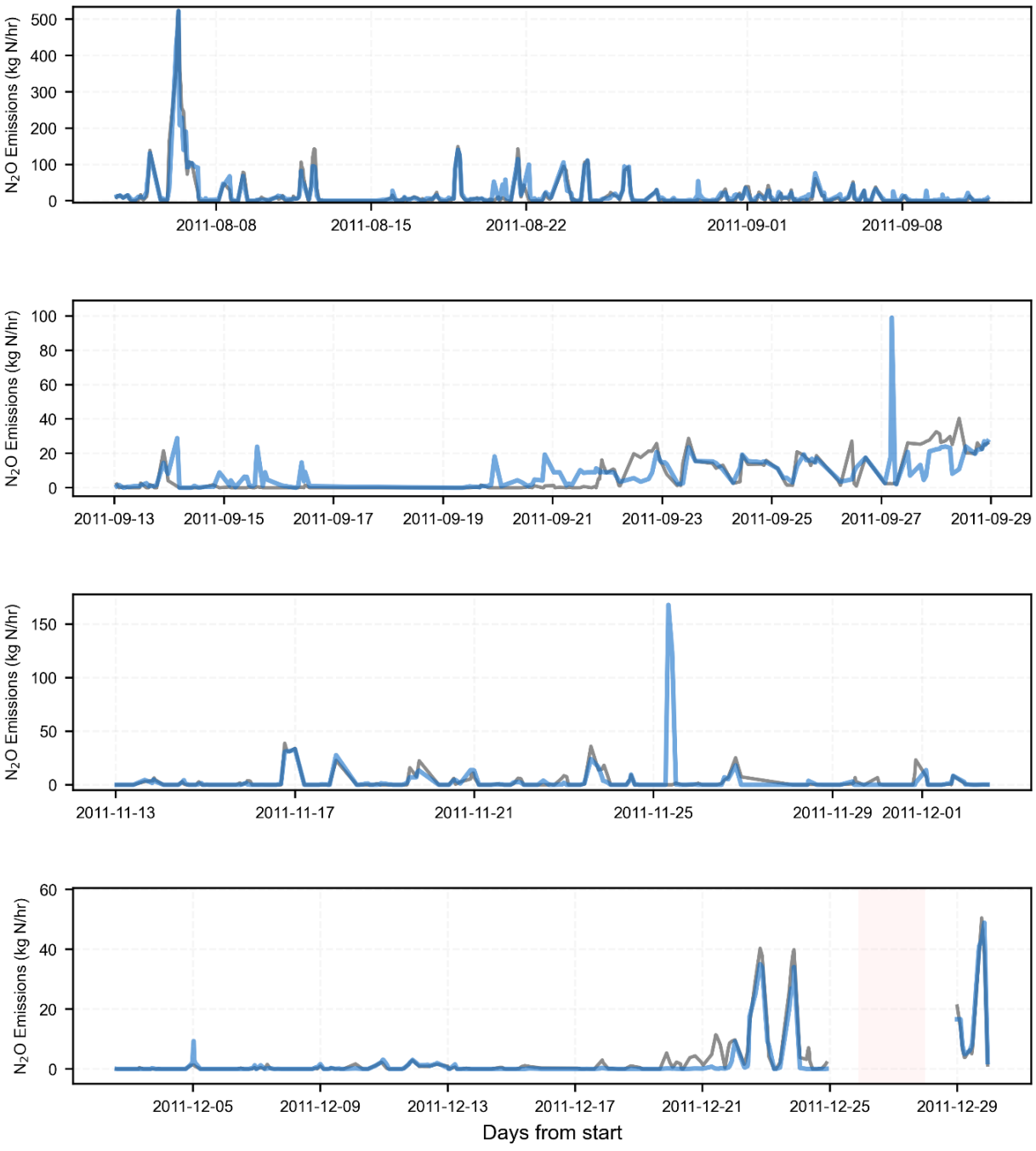
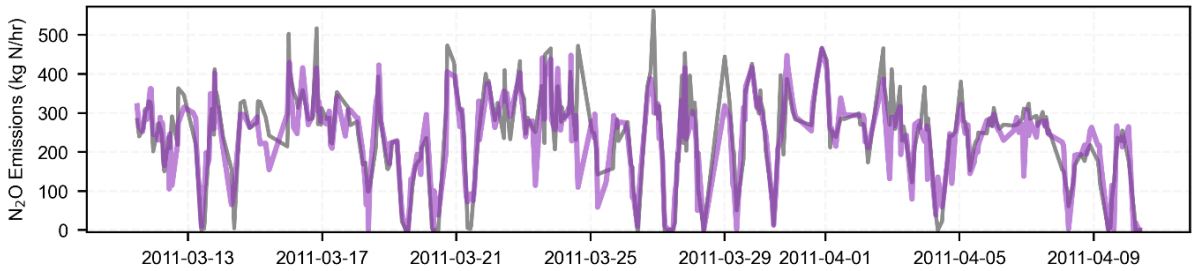
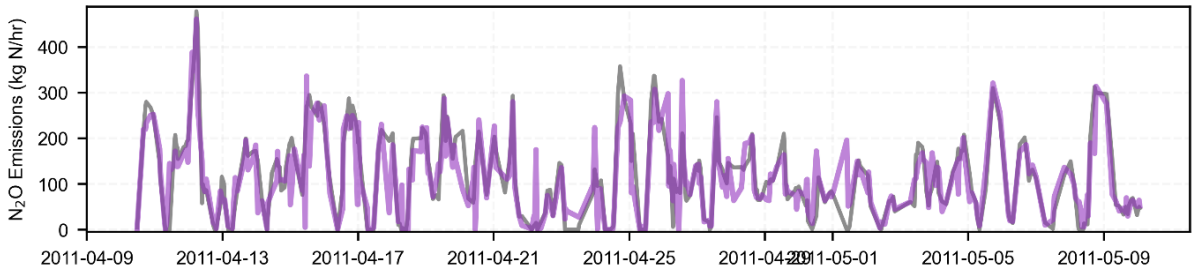
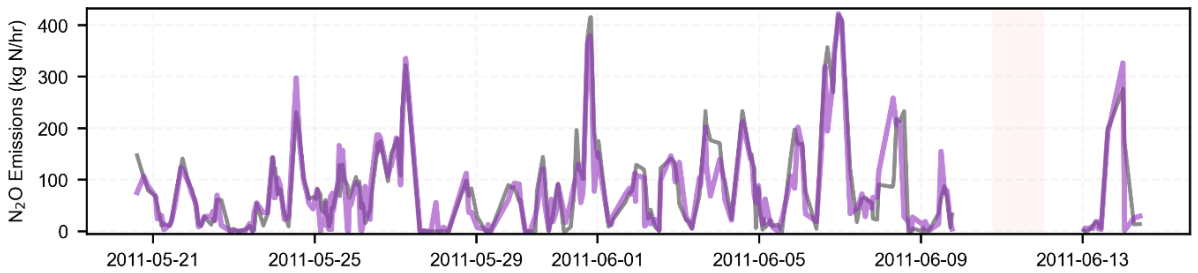
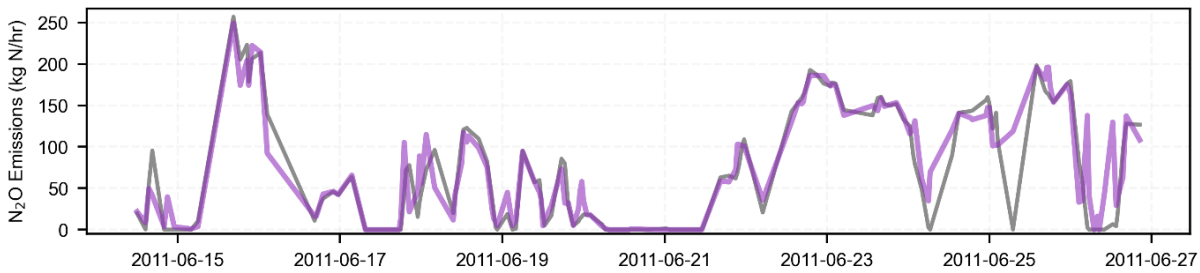


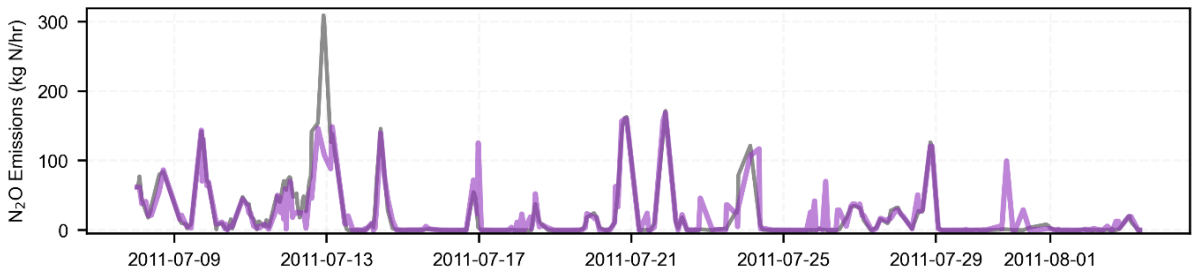
Figure 4-7 The performance of AdaBoost compared to the test dataset during the full-time window


N2O Emissions (kg N/hr)
0
100
200
300
400
500
2011-03-13
2011-03-17
2011-03-21
2011-03-25
2011-03-29
2011-04-01
2011-04-05
2011-04-09


N2O Emissions (kg N/hr)
0
100
200
300
400
2011-04-09
2011-04-13
2011-04-17
2011-04-21
2011-04-25
2011-04-29
2011-05-01
2011-05-05
2011-05-09


N2O Emissions (kg N/hr)
0
100
200
300
400
2011-05-21
2011-05-25
2011-05-29
2011-06-01
2011-06-05
2011-06-09
2011-06-13


N2O Emissions (kg N/hr)
0
50
100
150
200
250
2011-06-15
2011-06-17
2011-06-19
2011-06-21
2011-06-23
2011-06-25
2011-06-27


N2O Emissions (kg N/hr)
0
100
200
300
2011-07-09
2011-07-13
2011-07-17
2011-07-21
2011-07-25
2011-07-29
2011-08-01

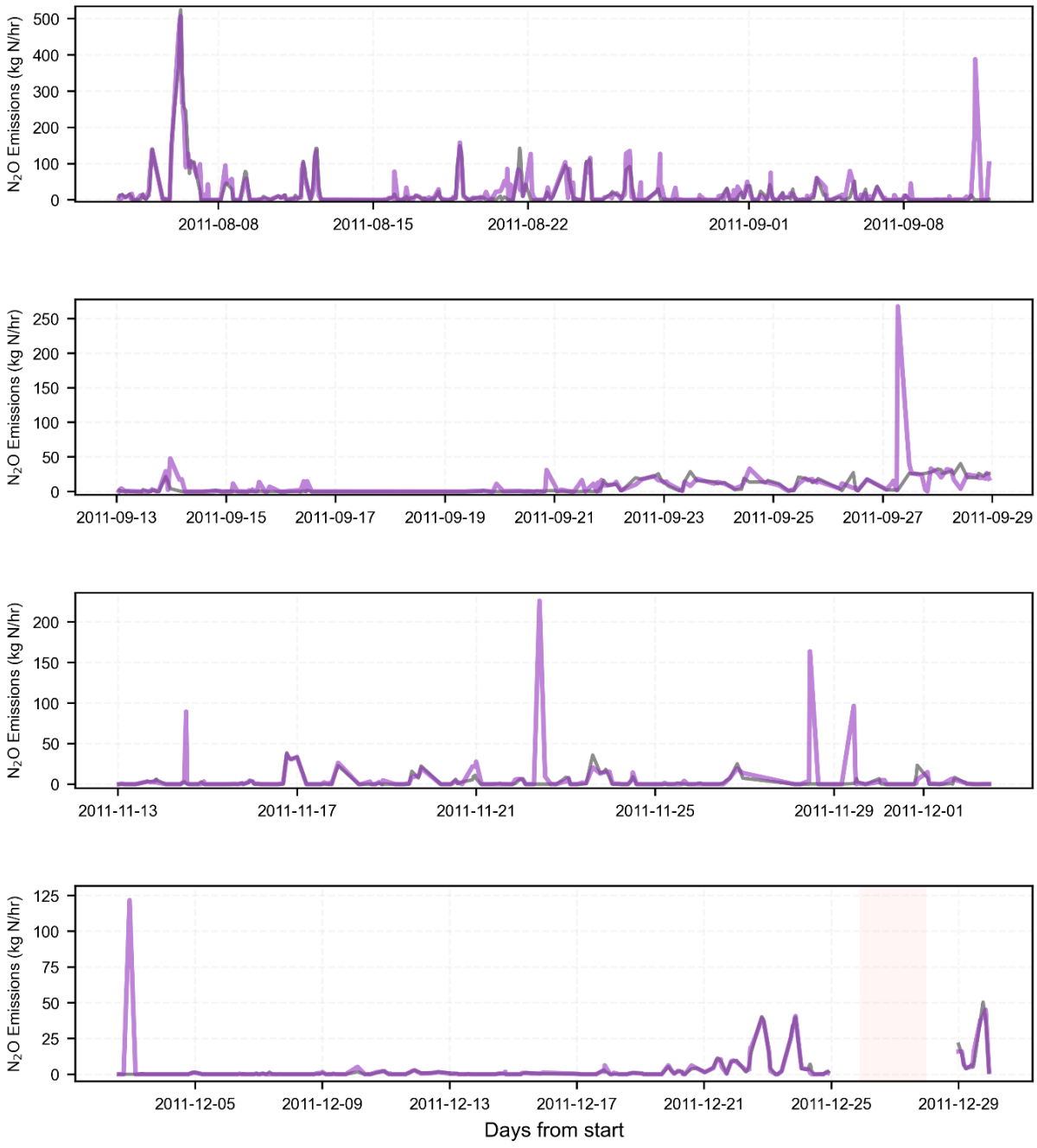
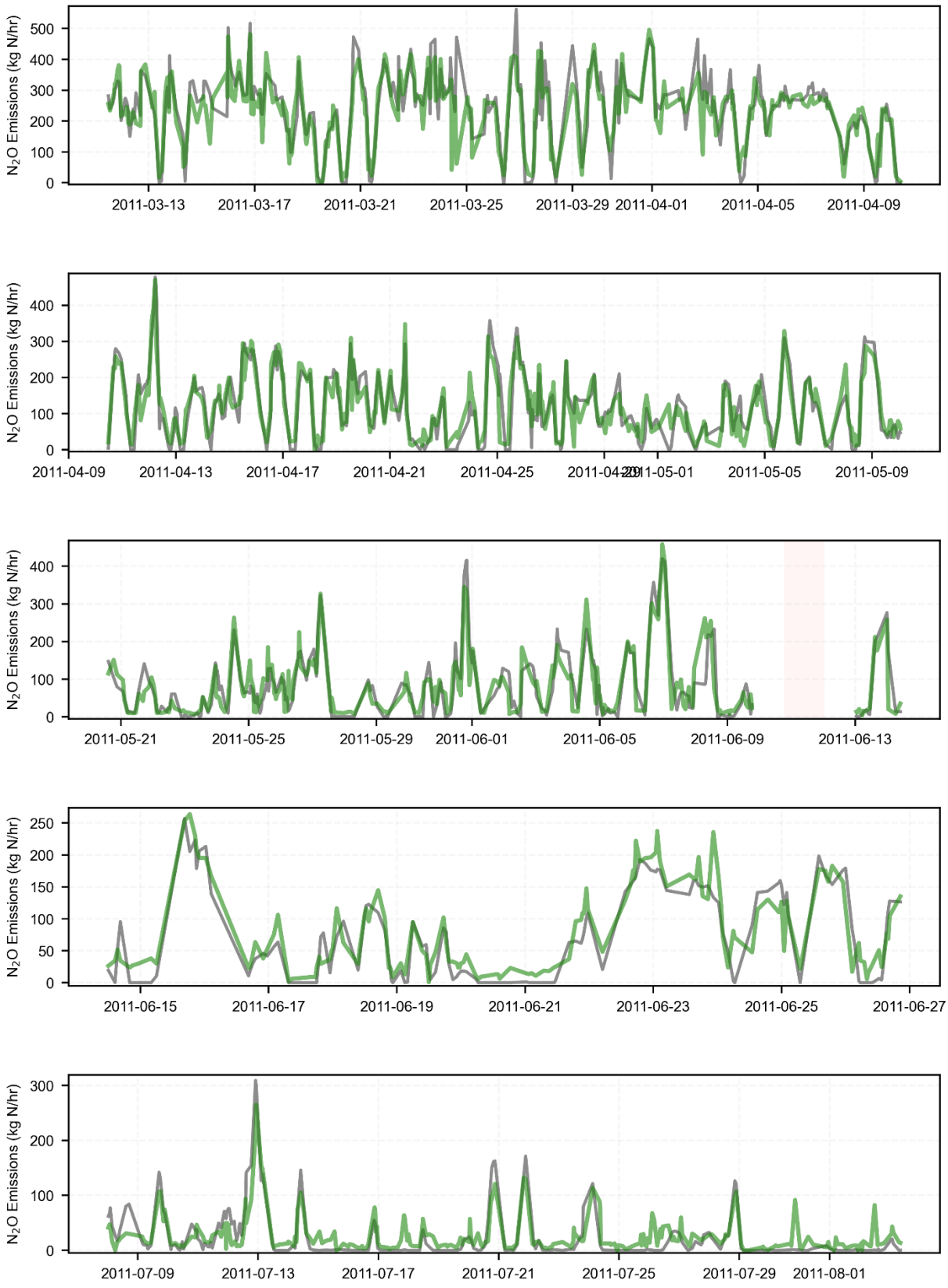


Figure 4-8 The performance of kNN compared to the test dataset during the full-time window


N2O Emissions (kg N/hr)
0
100
200
300
400
500
2011-03-13
2011-03-17
2011-03-21
2011-03-25
2011-03-29
2011-04-01
2011-04-05
2011-04-09
N2O Emissions (kg N/hr)
0
100
200
300
400
2011-04-09
2011-04-13
2011-04-17
2011-04-21
2011-04-25
2011-05-01
2011-05-05
2011-05-09
N2O Emissions (kg N/hr)
0
100
200
300
400
2011-05-21
2011-05-25
2011-05-29
2011-06-01
2011-06-05
2011-06-09
2011-06-13
N2O Emissions (kg N/hr)
0
50
100
150
200
250
2011-06-15
2011-06-17
2011-06-19
2011-06-21
2011-06-23
2011-06-25
2011-06-27
N2O Emissions (kg N/hr)
0
100
200
300
2011-07-09
2011-07-13
2011-07-17
2011-07-21
2011-07-25
2011-07-29
2011-08-01

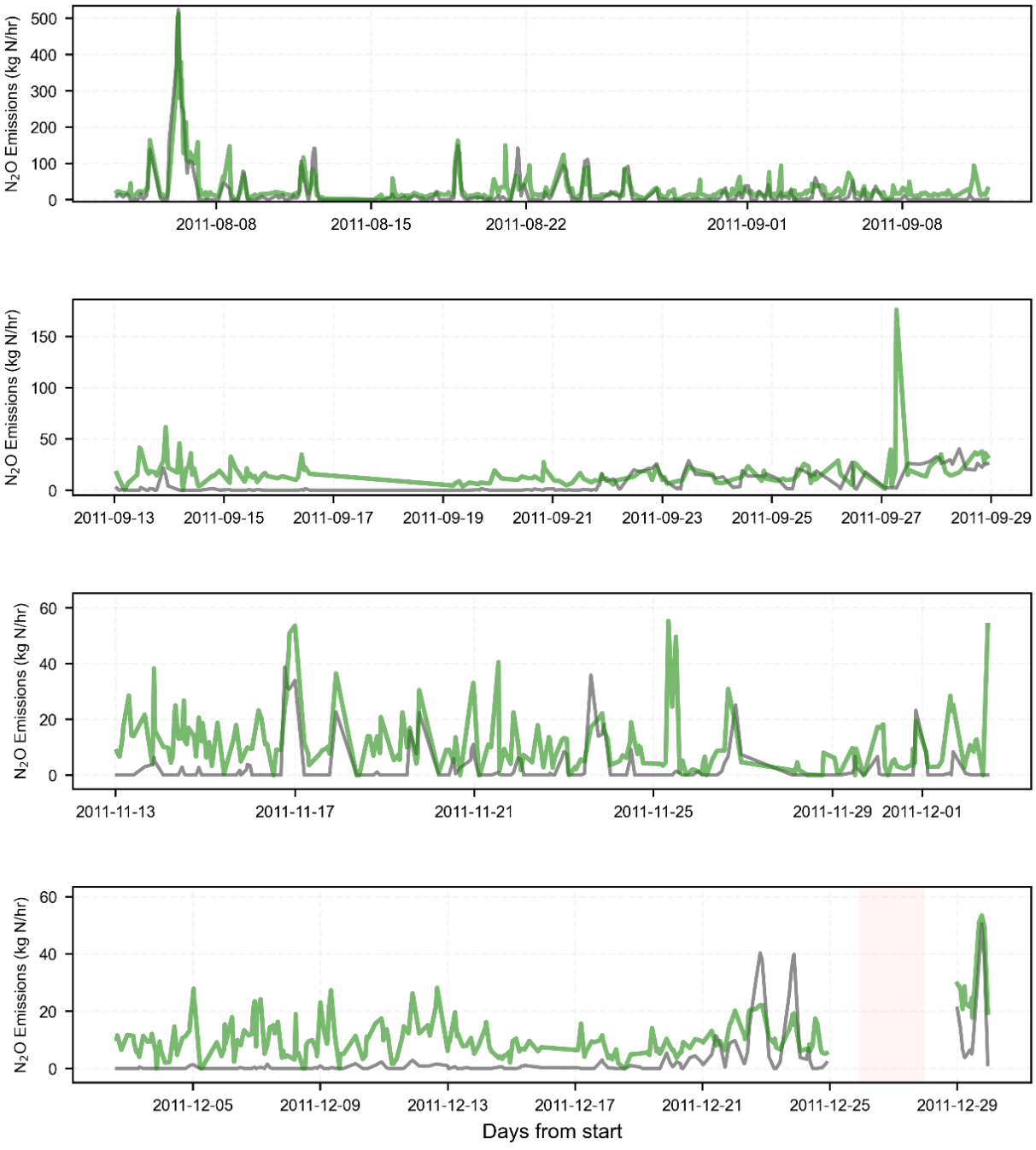


Figure 4-9 The performance of AdaBoost compared to the test dataset during the full-time window

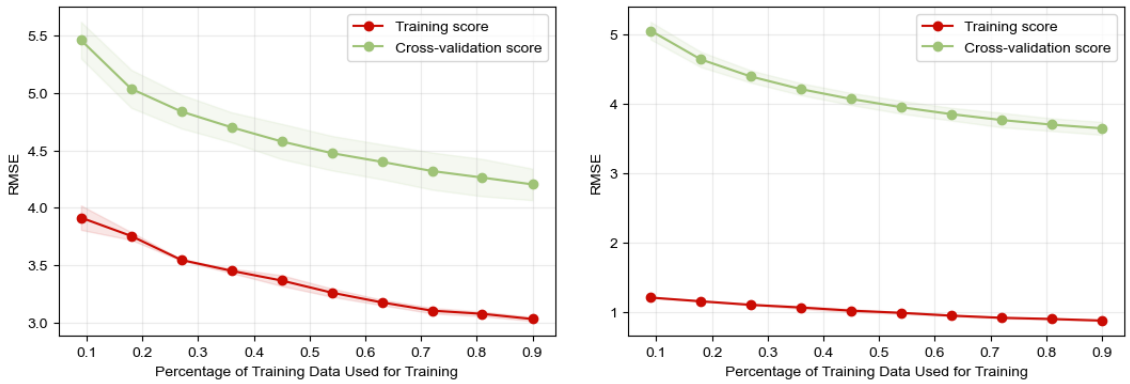
4.3.5.2 Processing time

The processing times were also measured given their importance in an online monitoring application such as soft sensors. Therefore, both training and prediction times have been measured and compared across kNN, AdaBoost, and DNN models as shown in Table 2. The results highlight the superiority of simpler models such as kNN in facilitating exceptionally rapid processing to generate predictions, while maintaining a relatively good prediction performance. However, it is imperative to conduct holistic assessment of algorithmic suitability for online applications, encompassing factors such as memory efficiency, batch vs. online updates, scalability, and resource constraints. Furthermore, DNN showed the longest training time among the three models, confirming its complexity. It is noteworthy that while the differences in processing time may appear negligible, these times are based on the current dataset; their amplification is plausible with expanding data volumes, an aspect crucial in online applications. Notably, incremental learning, a technique where a model gets updated with new data batches without retraining the model from scratch, holds promise for the AdaBoost model (Zhang et al., 2019). This potential for model adaptation over time could further contribute to reducing the processing times as the model becomes more refined.

4.3.5.3 Extent of overfitting

Figure 4.10 shows the model's performance for both the training and cross-validation datasets when calculated using various sizes of the training data (as a proportion of the total training data). These plots – referred to as the learning curves – are used to diagnose potential overfitting, where models perform exceptionally well on the training data but fail to generalize effectively to new, unseen data, leading to poor performance on validation or

test datasets. The AdaBoost model showed a gap between the training and validation dataset curves, raising a concern about overfitting. However, achieving the highest performance on the test dataset indicates that AdaBoost is indeed generalizing well to new, unseen data, and that the overfitting concerns may not be as severe as initially perceived. Moreover, the fact that the validation score improves with an increase in data volume further suggests that potential overfitting issues in the model could be mitigated over time by providing more training data. While DNN showed less severe overfitting problems with a smaller gap between learning and validation curves, the lower performance on the test data than AdaBoost and the fluctuations on the learning curves might be attributed that the model is sensitive to specific data points, causing their performance to fluctuate as they encounter different subset of the data. Finally, the kNN model showed intermediate gap between the learning and validation curve with both curves improve with larger data volume, indicating the possibility of benefiting from larger dataset and online training of the model.



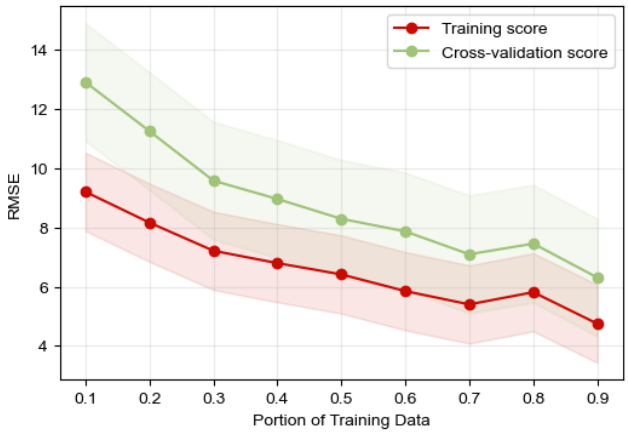


Figure 4-10 (on the left): kNN, (on the right): AdaBoost, (bottom): DNN

4.3.5.4 Residual analysis of models

Residual analysis across models involved assessing their behavior concerning observed N_2O concentrations and input features. The residual behavior against observed N_2O concentration showed that AdaBoost consistently displayed lower error magnitudes, while the kNN model exhibited higher errors for extreme N_2O values and the DNN model for lower N_2O values compared to AdaBoost. It was observed that all models displayed random deviations from zero error in the mid-range N_2O values. However, a non-uniform trend emerged with respect to extreme low and high N_2O concentrations, revealing increased negative and positive errors respectively. This non-uniform trend does not span the entire spectrum of N_2O values. Rather, it is primarily evident in the presence of a relatively high number of zero N_2O values within the dataset. This unique characteristic poses challenges due to dataset imbalance and requires careful consideration in interpreting model performance. Furthermore, residuals analysis against input features was conducted by plotting residuals against each of the input features. The analysis showed that no correlation exists between model errors and model input features as presented in Figures

4.11, 4.12, 4.13, and 4.14. This confirms the absence of bias or heteroscedasticity in the models.

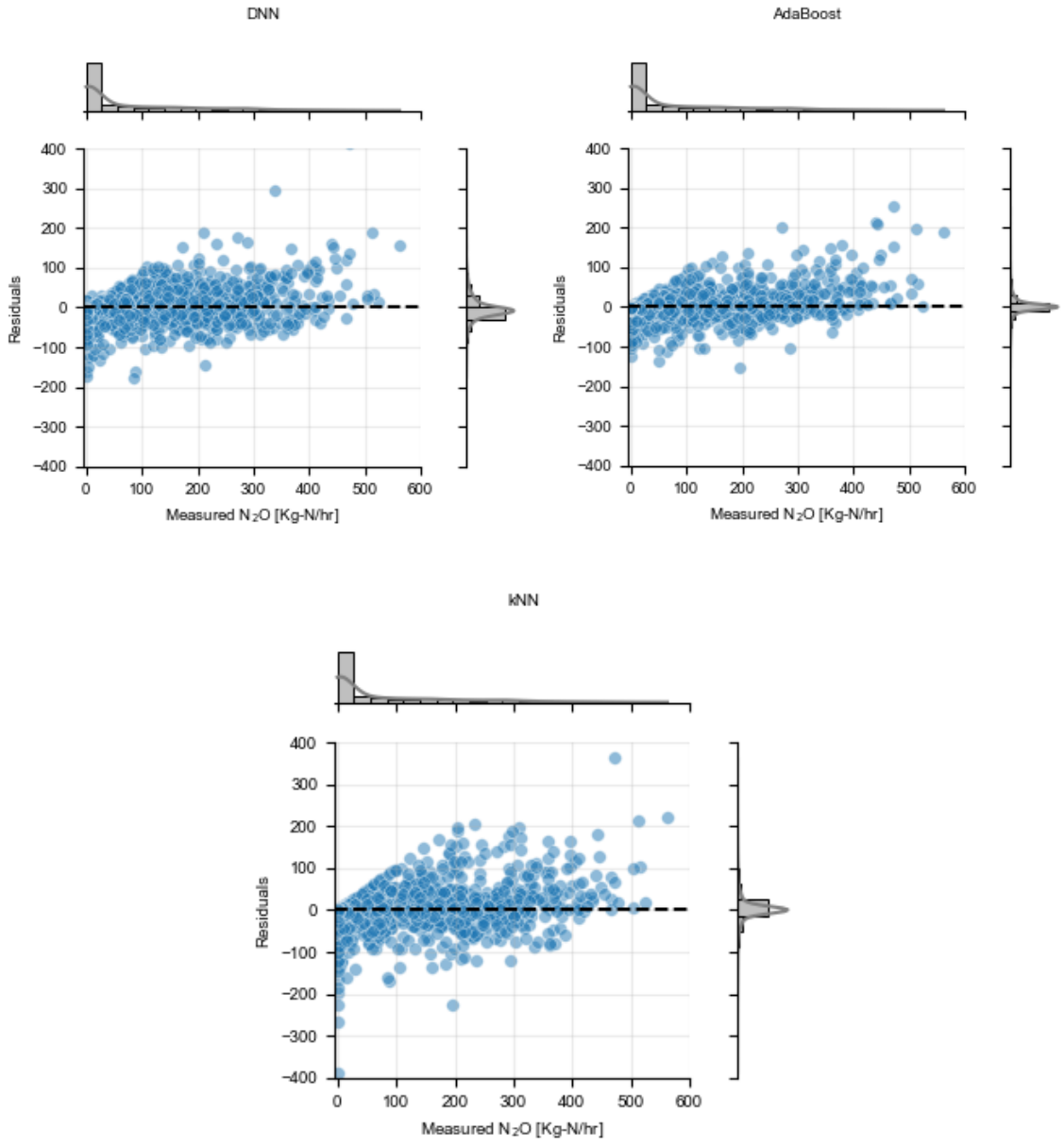


Figure 4-11 Residual analysis of kNN, DNN, and AdaBoost models

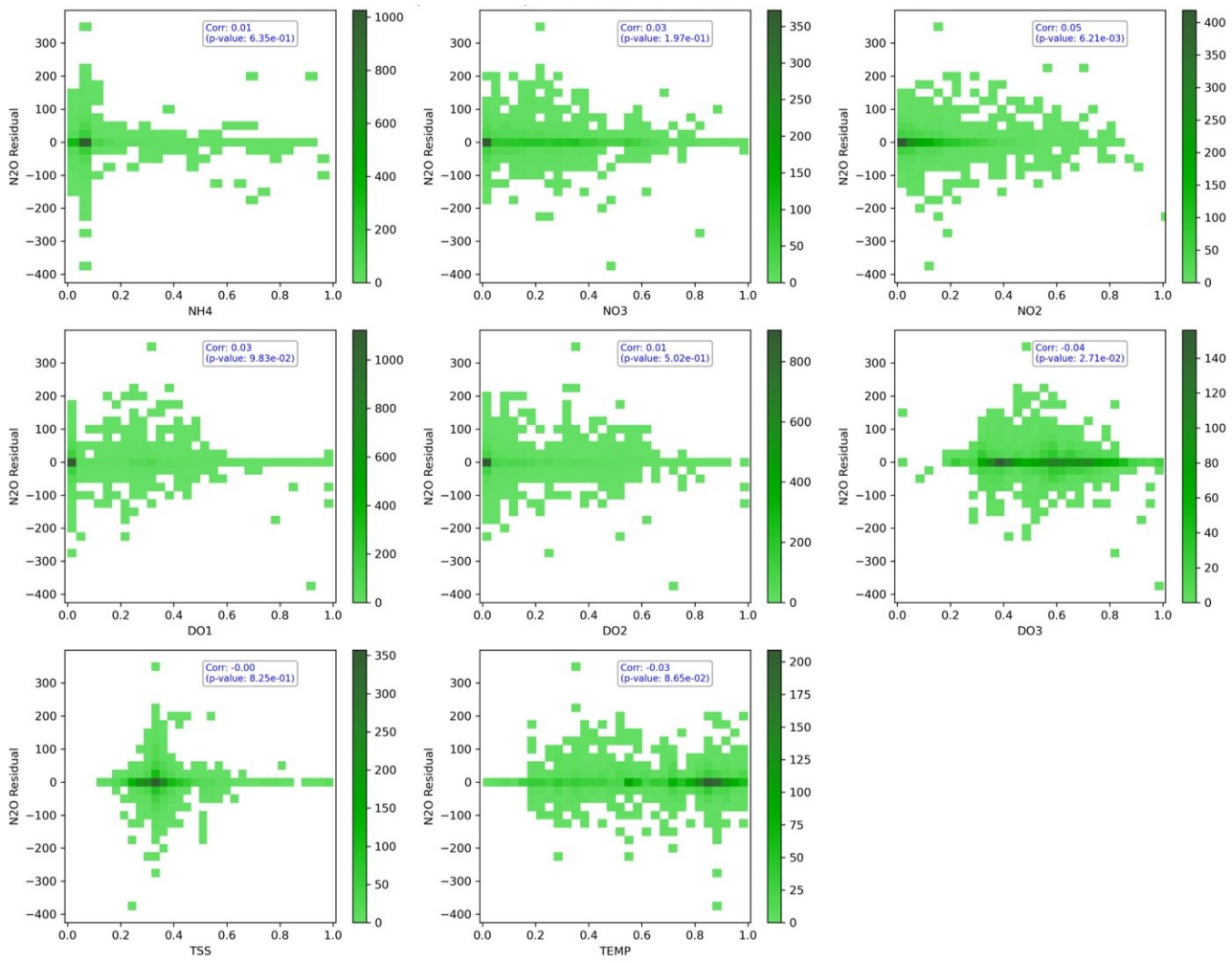


Figure 4-12 Residual plot of kNN model predictions against each input feature including the calculated Pearson correlation between each input feature and the model residual

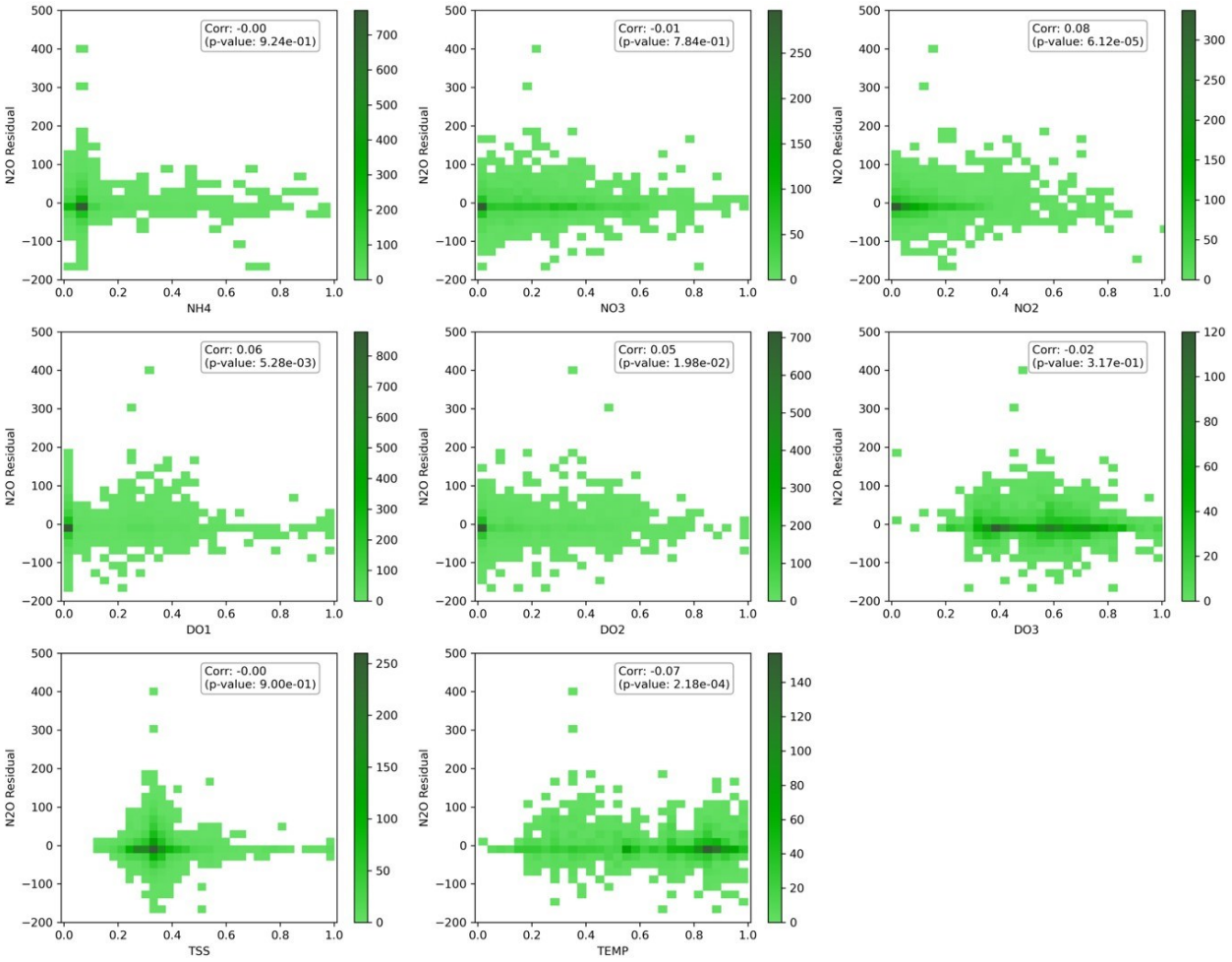


Figure 4-13 Residual plot of AdaBoost model predictions against each input feature including the calculated Pearson correlation between each input feature and the model residual

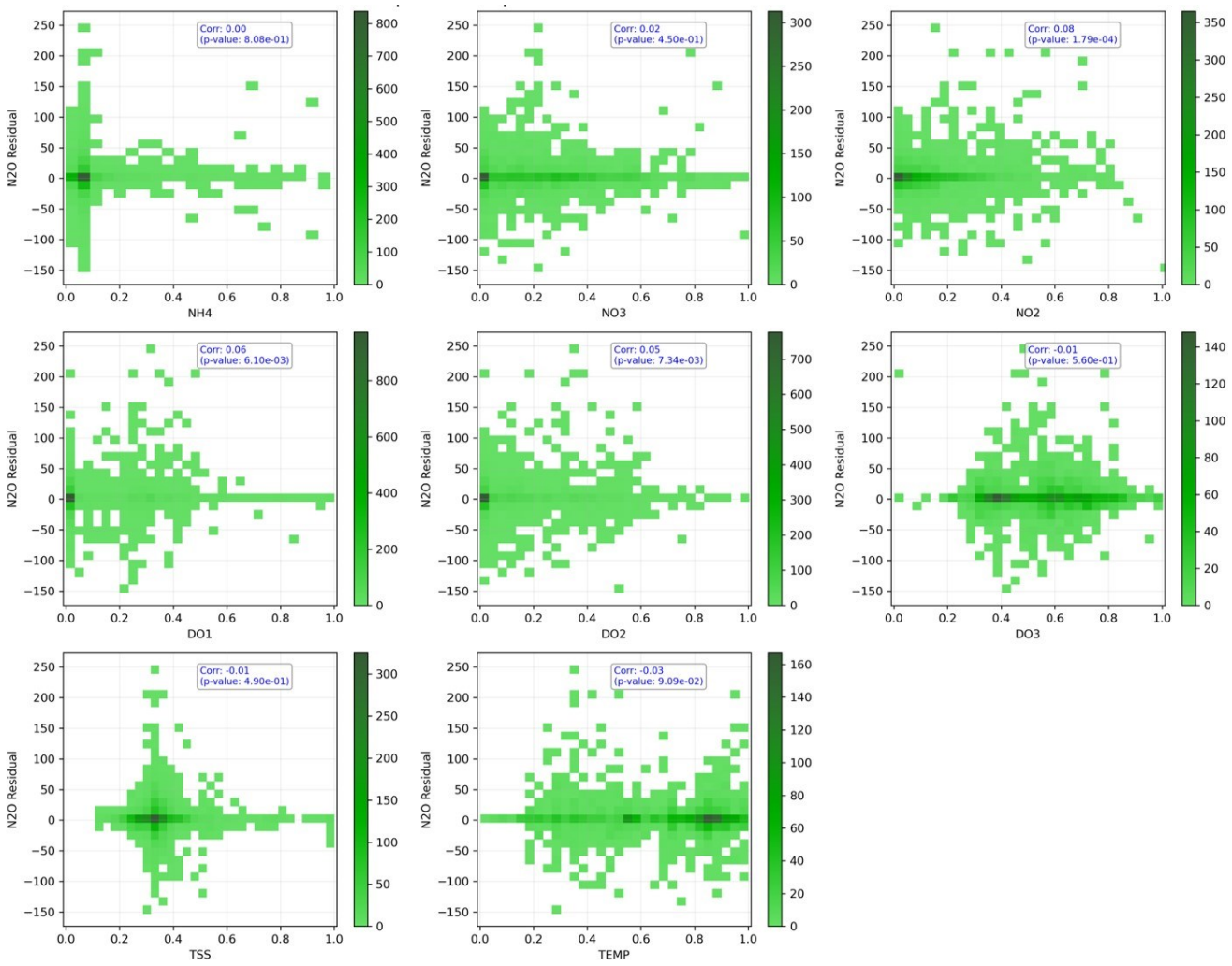


Figure 4-14 Residual plot of DNN model predictions against each input feature including the calculated Pearson correlation between each input feature and the model residual

4.3.6 Process understanding and feature importance analysis

Compliance between model feature importance and wastewater domain knowledge has significant impact on model interpretability, trustability, and decision-making process. Experts and stakeholders are more likely to accept the adoption of data-driven models when the model decisions are based on the same domain knowledge factors. This also validates that the model is learning from the correct resources, increasing confidence and trust in model outcomes. In the next steps, the feature importance ranking is investigated for the three selected models and their compliance with the process meaning.

4.3.6.1 Model-based feature importance

One of the advantages of ensemble tree-based models, such as AdaBoost, lies in their capacity to compute feature importance based on each feature's contribution to individual decision trees. In this study, the model revealed that NO_2^-, NO_3^-, and TEMP accounted for 50% of the importance when considering all available features, and these same features were ranked as the top three contributors, making up 60% of the importance following feature reduction (Figure 4.15). Out of the eight features selected by the mRMR method, seven were found to have the highest feature importance ranking when using all features as the model input, which underscores the built-in feature importance method's reliability in AdaBoost, provided the model performs well on training and validation data.

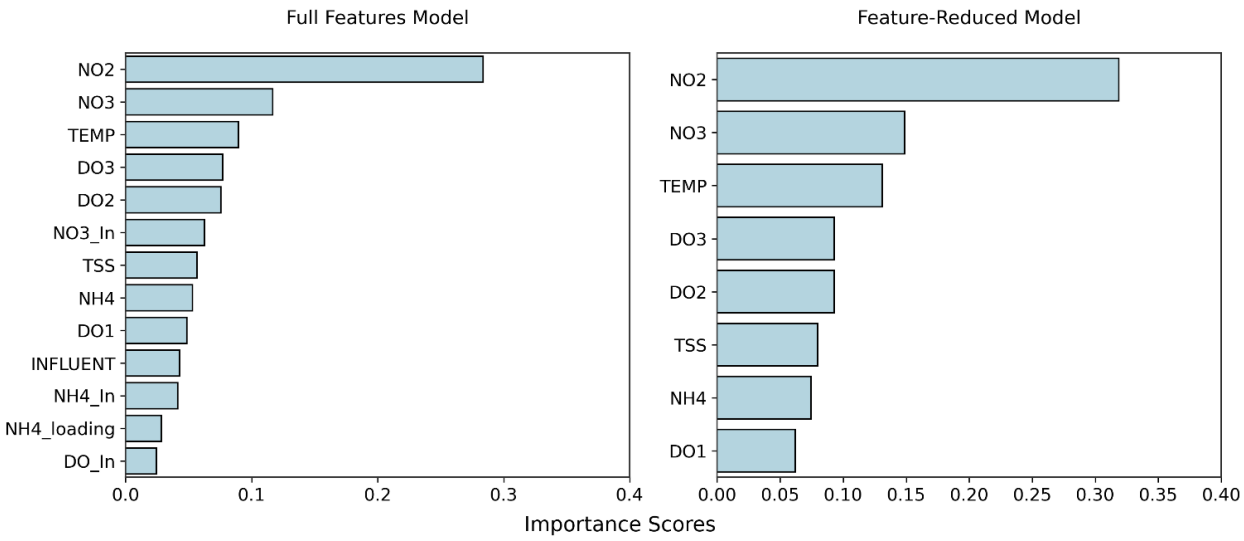


Figure 4-15 AdaBoost built-in feature performance

Nonetheless, the most crucial feature identified by this method differs from that indicated by mRMR used for feature selection. This discrepancy arises because data-driven models can sometimes exhibit bias, relying on features that contribute to overfitting the data, possibly due to a less complex relationship with the target variable. Although this does not negate the feature's potential as a strong predictor of the target variable, it should not be used for feature selection since it overlooks redundancy. This observation can be corroborated by examining Pearson's correlation matrix, as illustrated in Figure 4.16, it demonstrates that NO_2^- has the highest linear correlation with N_2O, which aligns with the selection as the highest important feature based on AdaBoost built-in feature importance method.

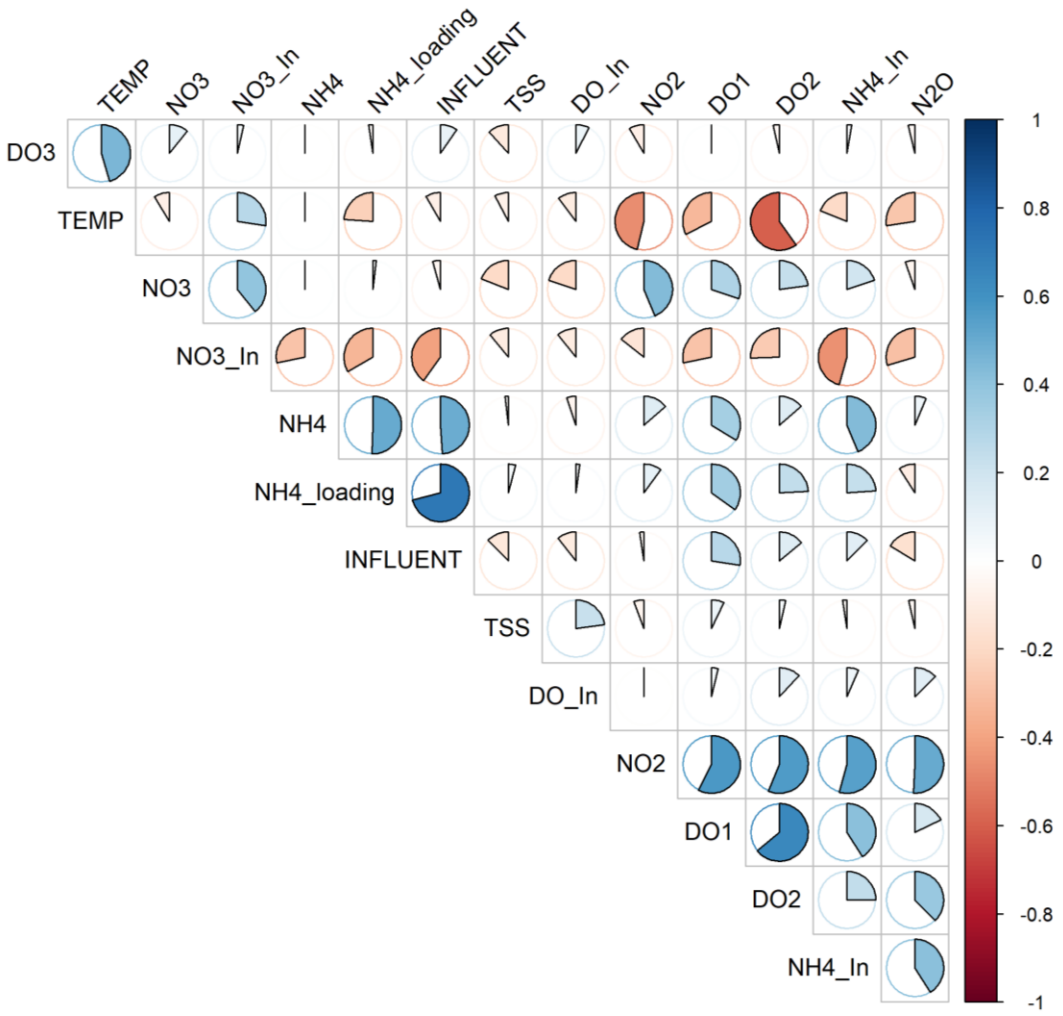


Figure 4-16 The Pearson correlation matrix of the entire dataset features

4.3.6.2 Model-agnostic feature importance and model interpretability

Permutation importance was used to estimate the importance of features in kNN, AdaBoost, and DNN models as shown in Figure 4.17. Common across all three models is the heightened importance assigned to NO_2^-, registering permutation scores at 65%, 45%, and 35% for kNN, AdaBoost, and DNN, respectively. This elevated impact on model performance can be attributed to its central role in the nitrifier denitrification (ND) pathway, in which AOB rely on NO_2^- as the electron acceptor to produce N_2O under low

DO conditions (M J Kampschreur et al., 2009). The contribution of the ND pathway to the total N_2O was reported to be the main pathway when there is NO_2^- accumulation, confirming the importance of the NO_2^- feature in the model (Wunderlin et al., 2013). Also, the three models agreed on the high importance of temperature, with permutation scores of 78%, 42%, and 45% for kNN, AdaBoost, and DNN, respectively. The effect of temperature on N_2O emissions can be explained by the increased enzymatic activities of both nitrifiers and denitrifiers at higher temperatures (Ahn et al., 2010). A negative correlation was demonstrated between temperature and NO_2^- in Figure 4.16, which can relate to higher N_2O emission by the ND pathway. This is in line with the study of Philips et al. (2002), where it was found that NO_2^- accumulates at low temperatures in WWTPs. Adouani et al. (2015) found that lower temperatures slow down NO and N_2O reductase activities, inducing N_2O emissions. Another essential factor to consider is the impact of temperature on N_2O solubility, influencing the rate of liquid-gas mass transfer. This, in turn, can have implications on the rate of N_2O emissions (Baeten et al., 2020). While these studies confirm the high impact of temperature on N_2O emissions, the exact mechanisms remain to be further investigated.

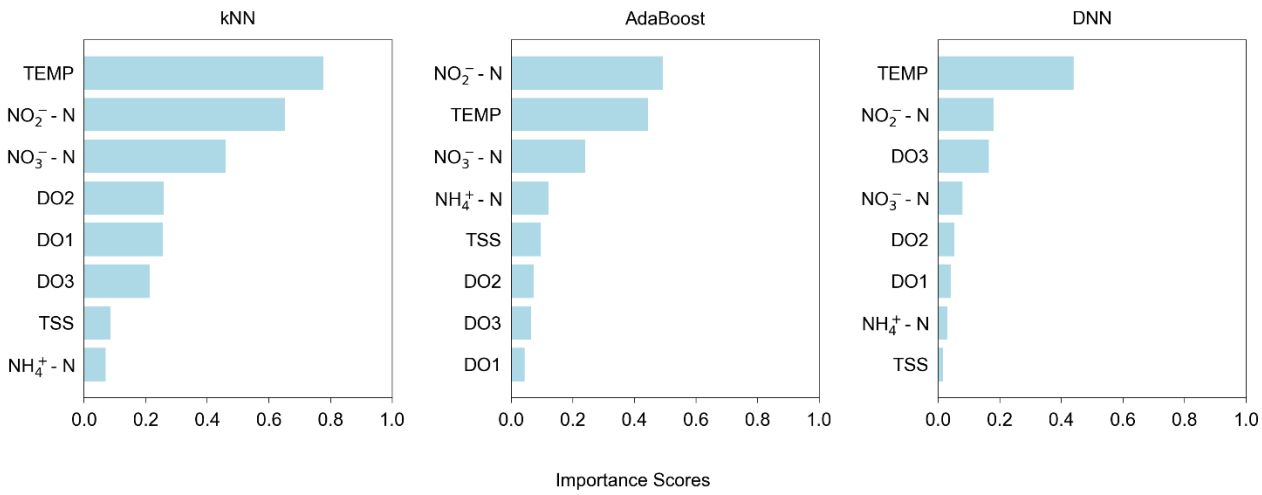


Figure 4-17 Permutation feature importance of feature reduced models: kNN, AdaBoost, and DNN models. The graph shows the relative change of model mean squared error as a result of feature imputation

Additionally, there was a strong agreement among the three models regarding the moderate permutation scores of NO_3^-, with values of 45%, 25%, and 10%, for kNN, AdaBoost, and DNN, respectively. The effect of NO_3^- can be explained by the fact that its accumulation can lead to incomplete denitrification in the anoxic part of the carrousel when there is a lack of carbon source, leading to poor N_2O consumption by heterotrophic biomass (Chen et al., 2019). This can result in N_2O generation via the heterotrophic denitrification pathway (Daelman et al., 2015). Furthermore, the permutation effect of NH_4 on N_2O emissions was significantly lower than both NO_2^- and NO_3^- in the three investigated models, which agrees with the conclusions of Daelman et al. (2015) that the ND pathway was more dominant in the present dataset. However, the insignificance of the NN pathway can still not be confirmed as the NH_4 permutation score was 12% in AdaBoost, and about 7% in kNN. In addition to nitrogen species, the permutation of all DO features was still relatively high for

the entire set of models. Despite the moderate correlation between DO2 and DO3 (r = 0.64), they had a similar impact on model performance. This can be attributed to the high heterogeneity in the DO concentrations due to aeration using surface aerators and the plug flow arrangement of the reactor, leading to a steep DO gradient along the carrousel. Even with a moderate degree of correlation between DO at different locations, this DO heterogeneity can explain data points representing complex interaction between N_2O pathways.

Besides the importance rankings, permutation scores varied in magnitude for the same feature across different models. Specifically, DNN and AdaBoost displayed comparatively lower magnitudes across all features when contrasted with kNN. This discrepancy underscores kNN's heightened sensitivity to features in comparison to the relatively more robust nature of AdaBoost and DNN. This observation points to the increased intricacy inherent to AdaBoost and DNN, enabling them to effectively capture complex interactions between features.

4.3.6.3 Understanding N_2O production pathways' contribution

Despite that NO_2^- is a key parameter influencing N_2O production through the ND pathway, it was not reported on previous datasets and not included in high accuracy models (Hwangbo et al., 2020, 2021). Whether this is due to insignificant contribution of ND pathway in the mentioned dataset or due to the existence of other variables that shares mutual information remains unclear. To investigate the pathway contribution in the current dataset, NO_2^- has been omitted from the input features to investigate the impact that this might have on feature importance ranking and model performance. Figure 4.18 shows that NO_2^- removal has changed the permutation feature importance ranking with a higher

weight to DO3 in all models compared to before NO_2^- removal. This might be perceived as evidence of the higher contribution of ND pathway as the models extracted the same information in NO_2^- from DO3 because the ND pathway only activates at both low DO and accumulated NO_2^-. NO_2^- has been identified as one of the major factors that influence the N_2O generation from WWTPs (Duan et al., 2021). This aligns with the pathway knowledge when NO_2^- is being used as the electron acceptor in the ND pathway and getting reduced sequentially to NO, and N_2O catalyzed by nitrite reductase (NIR) and nitric oxide reductase (NOR) (Yu et al., 2010). This pathway has been also considered the main contributor to N_2O production by AOBs (Yang et al., 2009). However, previous modelling of N_2O emissions using regression ML models does not include NO_2^- as an input feature (Hwangbo et al., 2020, 2021). The high performance achieved by the DNN model in the latter study cannot be used as evidence of the insignificance of ND pathway contribution to N_2O emissions in the reported case. Insights on pathway contributions should be better provided in the case of presence of NO_2^- to allow a comprehensive feature importance analysis that supports model interpretability. This aligns with the recommendation by the original study that collected and analyzed the same dataset (Chen et al., 2019). It is noteworthy that the purpose of this analysis is to improve interpretability by ensuring that the model is using the correct features, which is a different objective than maximizing model performance. The models in this study showed a drop in performance in response to NO_2 elimination in terms of R^2 values, registering at 0.83, 0.91, and 0.89 for kNN, AdaBoost, and DNN, respectively (a more comprehensive analysis was not performed). The drop may seem insignificant from the perspective of model performance, especially in the DNN. This might be attributed to the complexity of DNN that allows it to capture

complex relationships from features potentially sharing part of the same information; something simpler models are not able to do. Hence, the model and features choice are highly dependent on the objective and the available resources.

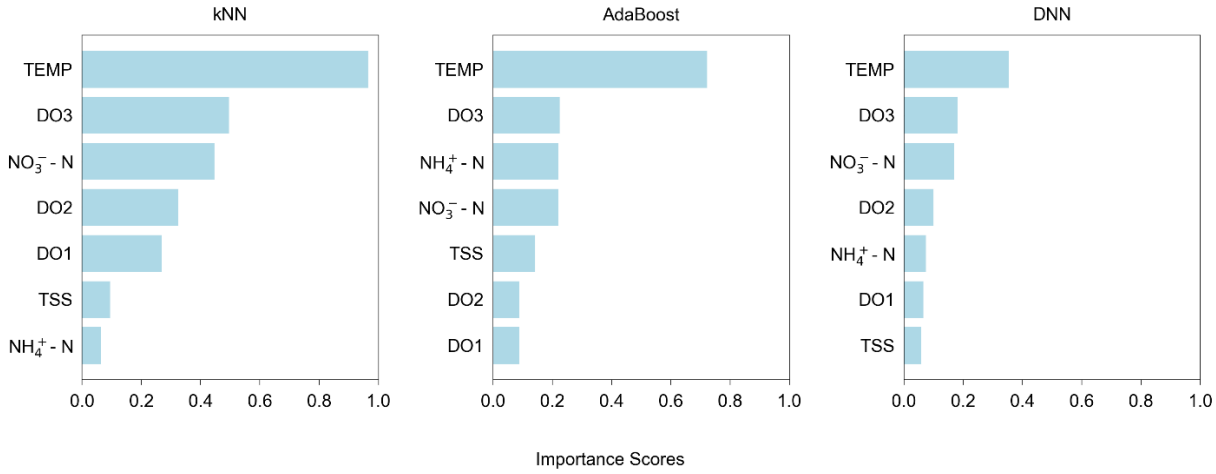


Figure 4-18 Permutation feature importance of feature reduced models: kNN, AdaBoost, and DNN models after elimination of Nitrite (NO_2) from the dataset

4.4 Process and environmental implications

There is a growing interest in incorporating GHG emissions into the evaluation of WWTP performance, which will probably lead to decisions supporting the mitigation of GHG emissions. One of the means to reduce GHG emissions is the selection of process control and operational strategies that account for N_2O emissions. By employing an ML framework to generate model predictions consistent with process knowledge, it becomes possible to evaluate emission-influencing factors on a case-by-case basis using instance-based model explainability tools. Empowered by this approach, operators and decision-makers can confidently make informed choices regarding control and operational strategies. These choices effectively integrate the reduction of N_2O emissions with other control objectives.

In this study, surface aerators in the reactor were solely controlled based on NH_4^+-N concentrations, representing the conventional control approach based on effluent quality. Nevertheless, this approach seems insufficient in terms of mitigation of N_2O emissions from the present WWTP. As a result, a more comprehensive control strategy founded on multiple variables should be considered, including NO_2^--N and DO levels at various points within the carrousel to eliminate peaks in NO_2^--N, reducing N_2O emissions accordingly. Nevertheless, the development of a comprehensive control philosophy requires further investigations. For instance, one can dynamically adjust the length and duration of anoxic and aerobic phases as needed to prevent unfavorable DO conditions. This concept aligns with the conclusions drawn by Daelman et al. (2015), who identified suboptimal DO conditions as a key factor contributing to elevated N_2O levels in the present WWTP. Given the intricate nature of this control objective, a ML-based model predictive control (MPC) strategy might also be applied. A similar approach was proposed by Hwangbo et al. (2021). In this context, a soft sensor relying on the developed ML model may play a crucial role, either by integration into the MPC system or by supplying accurate N_2O emission data to the MPC. This data will serve as valuable historical information for training the predictive model within the MPC framework. Employing this strategy emphasizes the sustainability aspects of WWTPs rather than solely focusing on performance indices.

4.5 Summary

This study confirmed the potential of ML models as viable alternatives to conventional N_2O quantification methods, and as a replacement for mechanistic models given their limitations in full-scale applications. It was demonstrated that such models hold the dual advantage of describing N_2O emissions, while also guiding decision-making processes

through a balanced consideration of model performance, processing time, complexity, and interpretability. Such advantages can be harnessed by adopting the comprehensive framework established in the present study incorporating 1) data preprocessing considering the specific characteristics of wastewater data, 2) input feature reduction that only considers the most relevant and least redundant feature set, 3) development of ML models with different degrees of complexity, 4) comprehensive evaluation of model performance, and 5) facilitating model interpretability using feature importance analysis.

The developed framework showed a potential to a fit-for-purpose soft-sensor application that models and provides understanding of N_2O emissions from WWTPs. The framework was successfully applied to a full-scale long-term dataset using models spanning three levels of complexity and interpretability. Among the ML-models tested, AdaBoost demonstrated a good balance between accuracy, processing time, complexity, and interpretability. While the DNN displayed a marginally lower accuracy, it also came with a higher processing time and larger complexity. On the other hand, kNN compensated its least optimal accuracy with the lowest complexity and processing time, which may still be useful for some applications. Finally, utilizing an interpretable model can provide necessary insights for a multi-variate control approach considering GHG emission mitigation in addition to effluent quality compliance.

Chapter 5 - An integrated feature selection and hyperparameter optimization algorithm for balanced machine learning models predicting nitrous oxide emissions from wastewater treatment plants*

5.1 Introduction

Various ML models were employed in the wastewater field to learn from input data in order to predict a continuous output variable, such as deep neural networks and tree-based models. Such learning process, known as model "training", aims at optimizing the value of model parameters to best fit the training data. The efficacy of model training dictates model's overall performance and its ability to generalize to new data. Significantly, model training is influenced by the tunning of its "hyperparameters". Hyperparameters are the external configuration settings of a machine learning model that are not learned from the data but are set prior to the training process. These settings control various aspects of how the model learns and operates.

Among the tree-based algorithms used for prediction models, Adaptive boosting (AdaBoost) showed high performance in prediction of N_2O emissions from WWTP in chapter 4. The AdaBoost algorithm works by combining multiple weak learners to form a strong regressor, in which the number of weak learners (such as decision trees) to be used is an example of a hyperparameter.

---

* A modified version of this chapter has been submitted in December, 2023 for publication in Journal of *Water Process Engineering*

This hyperparameter plays a crucial role in the performance of the AdaBoost model, in addition to other hyperparameters such as the maximum depth of each weak learner (tree). A model with a low number of estimators might underfit due to the low model complexity that fails to capture high non-linearity in the data. Conversely, complex models are adept at uncovering hidden relationships within the data and able to achieve higher performance thanks to their enhanced flexibility. However, the use of highly complex models comes with certain drawbacks. They demand significant computational resources, requiring more time and memory for both training and making predictions on new data instances. Additionally, complex models are more prone to overfitting, resulting in poor performance when exposed to new data instances (low generalizability) (Bishop, 2006). Furthermore, increased complexity can negatively impact the interpretability of the model (Molnar, 2022). Therefore, the attainment of a trade-off between model complexity on one side, and model accuracy and interpretability (denoted as "model performance" hereafter) on the other side is a crucial consideration in the application of ML models (Doshi-Velez and Kim, 2017).

N_2O modelling is no exception that a trade-off between model performance and complexity is crucial for successful model development. In fact, it can be argued that such a trade-off is more significant in the case of N_2O modelling given the limited capabilities of its mechanistic models to quantitively describe N_2O emissions (Wan and Volcke, 2022). An optimal balance between performance and complexity can be achieved by (i) input feature selection to optimize the number of features induced complexity (feature complexity), and/or (ii) hyperparameter optimization (HPO) to reduce the model architecture induced complexity (architecture complexity).

Different methods for input feature selection can be used, generally divided into model-based and filter-based methods (Hvala and Kocijan, 2021). Filter-based methods are independent of the model and focus on the statistical relationships between features. They utilize statistical metrics like Pearson correlation, Spearman correlation, F-test, or information theory metrics such as Mutual Information. These methods assess features based on their individual statistical properties, disregarding the model's structure, and have been frequently used in ML models in wastewater applications to reduce dimensionality and the resulting model complexity (Deepnarain et al., 2019; Zaghloul and Achari, 2022). Model-based methods, in contrast, rely on the model itself to determine the predictive power of the input features. This is done by using validation data, where the model's performance with various subsets of features is evaluated to identify those that contribute most to its predictive accuracy. Another method for effective feature selection is the minimum redundancy maximum relevance (mRMR) method which selects the highly relevant features while minimizing multicollinearity. While it is considered a filter-based method, a recent study utilized mRMR as a wrapper-method where the number of features is determined based on model performance. As a result, it effectively streamlined the complexity of the examined models, ensuring they are both efficient and less complex. Moreover, it was demonstrated that the selected features using mRMR are well-aligned with the mechanistic understanding, which underscores the method's ability to identify pertinent features based on their relevance to the predictive modeling of N_2O emissions. One should be aware, however, that the inverse is not true: the selection of features should not be interpreted as proof for causality between the feature and the model output.

Despite the significance of feature selection in reducing the overall model complexity and easing data acquisition efforts, the role of HPO in managing model complexity is often overlooked, particularly in wastewater research studies utilizing ML tools (Zhu et al., 2023). To the best of our knowledge, the common practice for HPO, in prior ML wastewater studies, is employing grid search or random search algorithms. For example, grid search was used to select the optimal hyperparameters of Support Vector Machine predicting N_2O emissions (Vasilaki et al., 2020b, 2020a). These algorithms focus on exhaustively or randomly testing a pre-defined space of hyperparameters to identify the set that yields the best performance on validation or cross-validation sets (Bergstra et al., 2011). However, the grid/random search approaches are highly dependent on the pre-defined hyperparameter space and do not consider the trade-off between model architectural complexity and performance. To address such a challenge, Hwangbo et al. (2020) utilized metrics like AIC (Akaike information criterion) and BIC (Bayesian information criterion) to identify the optimal structure of a deep neural network from a pre-defined set of model architectures. However, this approach still did not investigate a wide grid-space of solutions and is expected to miss a possible optimal model structure. Moreover, more automated/generalizable approach is still needed. A more advanced HPO approach can be conducted using metaheuristic search algorithms. An example is the optimization of the architecture of a deep neural network predicting wastewater was implemented by (Zhu et al., 2022) using genetic algorithm (GA), particle swarm optimization (PSO), and grey wolf optimization (GWO).

Nevertheless, there are several limitations to the approaches commonly used in literature to HPO and feature selection. The HPO focuses solely on model performance, overlooking

an explicit representation of model complexity, potentially leading to unnecessary complexity. Furthermore, both input feature selection and HPO methods are typically performed in two isolated steps. While this practice can lead to high accuracies, the present article argues that it leads to suboptimal models in terms of complexity-trade off resulting in unnecessarily complex models which keeps such models prone to overfitting, limited generalizability, and limited interpretability.

The primary objective of this chapter is to do develop an approach aimed at optimizing model complexity alongside performance, addressing a critical gap in current N_2O emission prediction methodologies. By investigating the synchronization between input feature selection and HPO, their combined effect on model complexity and performance is elucidated. The performed comprehensive analysis includes an exhaustive grid search across a wide spectrum of input features and model hyperparameters, highlighting the significance of the developed algorithm that integrates and automates input feature selection coupled with HPO. This methodology leverages GA for identifying the optimal model configuration, which is then compared against a simpler search algorithm to demonstrate its superiority. It is hypothesized that this integrated strategy will not only bolster model efficiency but also enhance generalizability, thereby mitigating the risk of overfitting. Furthermore, this study seeks to underscore the often-overlooked importance of model complexity in environmental engineering applications of ML tools, with a particular emphasis on modeling N_2O emissions from WWTPs. This research offers invaluable insights and tools for engineers and researchers striving to navigate the complexities of ML in environmental applications.

5.2 Approach and Methodology

5.2.1 Model training and evaluation

5.2.1.1 Adaptive boosting model

The Adaptive Boosting (AdaBoost) regression model was employed in this study, an ensemble technique known for its ability to improve the accuracy of weak learning algorithms (Freund and Schapire, 1997). AdaBoost was designed by sequentially applying a decision tree regressor, to repeatedly modified versions of the data. Each subsequent model in the sequence focuses on the instances that were incorrectly predicted by the previous models, thereby adaptively boosting their importance in the dataset. The final prediction is made through a weighted average of the predictions from all learners in the ensemble, thus reducing both bias and variance, and leading to improved prediction accuracy. The model training and evaluation was performed using Sklearn open-source library in python. Below is a simplified mathematical overview of our AdaBoost implementation. Initially, the weights (w_i) for all training instances (N) are set equally:

$$w_i = \frac{1}{N} \qquad 5.1$$

A weak learner is then trained to make predictions (y_i) for each data point, and the loss (L_i) for each prediction is calculated using the square loss function, normalized by the maximum absolute error:

$$L_i = \left(\frac{y_i - \hat{y_i}}{\max |y_i - \hat{y_i}|} \right)^2 \qquad 5.2$$

The average weighted loss (L) for the weak learner is computed as:

$$\bar{L} = \sum L_i w_i \qquad 5.3$$

This loss informs the performance weight (β) of the weak learner:

$$\beta = \frac{\bar{L}}{1 - \bar{L}} \qquad 5.4$$

Subsequently, the weights for the next learner are updated, reflecting each instance's loss and the previous learner's performance:

$$w_i \rightarrow w_i \beta^{1 - L_i} \qquad 5.5$$

The final prediction model is determined by the value corresponding to the weighted median of all individual learner predictions, where the weight of each learner is inversely proportional to its error.

5.2.1.2 Model evaluation and comparison

The comparison among different models was based on the root mean squared error (RMSE) and the mean absolute error (MAE) calculated using the following formulas:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y_i})^2}{n}} \qquad 5.6$$

$$MAE = \frac{\sum_{i=1}^{n}\left|y_i - \hat{y_i}\right|}{n} \qquad 5.7$$

where y_i is the model predictions, $\hat{y_i}$ is the true value, and n is the number of data points. Moreover, a more comprehensive evaluation strategy was used when comparing the outcomes of the developed algorithm integrating feature selection and HPO against the uncoupled method used in Chapter 4 to allow for a nuanced understanding of the trade-offs between predictive performance and model complexity. Thus, the mean error (ME) and coefficient of determination were also calculated, as follows:

$$ME = \frac{\sum_{i=1}^{n}(y_i - \hat{y_i})}{n} \quad 5.8$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y_i})^2}{\sum_{i=1}^{n}(y_i - y_i)^2} \quad 5.9$$

5.2.2 Investigation and data analysis of the hyperparameters-input features grid space

To investigate the combined effect of hyperparameters and the number input features on model accuracy, the exhaustive grid search method coupled with cross-validation was employed. For each n number of input features selected by mRMR method, the hyperparameters of the AdaBoost regressor were optimized using the GridSearchCV available in Sklearn. This exhaustive method systematically explores a wide range of hyperparameter values, providing an extensive understanding of the hyperparameter space (Bergstra et al., 2011). The grid of potential values for key hyperparameters, including the number of estimators, learning rate, decision tree maximum depth, and the used loss function, were defined with the ranges presented in Table 5.1. For each combination in this grid, the AdaBoost regressor is trained and evaluated, based on the average model performance across cross-validation sets. K-fold cross-validation was employed, which involves dividing the dataset into K distinct subsets, iteratively training the model on K-1 subsets and validating it on the remaining subset. The exhaustive exploration of the hyperparameter space through grid search with cross-validation enables a thorough exploration and data analysis of the combined features and hyperparameter space.

5.2.3 An integrated algorithm for feature selection and hyperparameter optimization

An enhanced and automated algorithm that integrates mRMR feature selection with a multi-objective hyperparameter optimization has been developed, as illustrated in Figure 5.1. This algorithm emphasizes the tuning of hyperparameters for the model, utilizing every

possible set of n input features to ensure that the optimal combination of features and hyperparameters is identified. Moreover, the hyperparameters are optimized through a multi-objective optimization algorithm that select the set of hyperparameters that maximizes the model performance and minimizes the ML model's complexity. This enables an optimized hyperparameter optimization process that results in a less complex yet accurate model, which is necessary for online applications of N_2O emission prediction. Thus, the following objective functions were utilized to represent the fitness evaluation and selection:

$$Mean\ squared\ error = \frac{1}{n}\sum(Y_s - Y_o)^2 \qquad 5.10$$

$$Number\ of\ nodes = 2^{Tree\ maximum\ depth} \times Number\ of\ Estimators \qquad 5.11$$

where Y_s represents the simulated data, Y_o the actual testing data, n the number of observations, $Number\ of\ Estimators$ the number of weak learners in the model, max_depth is the maximum depth of each decision tree. For the multi-objective optimization problem, two algorithms were compared: Nelder-Mead and GA.

5.2.3.1 Nelder-Mead optimization algorithm

The Nelder-Mead algorithm is a numerical minimization method that is used for optimization of non-linear problems, and is often called "Simplex" algorithm. The algorithm starts with N+1 arrays of hyperparameter values, where N is the number of hyperparameters that need to be optimized, representing a polytope with N+1 vertices in an N-dimensional space. The fundamental operations include reflection, expansion, contraction, and shrinkage, which iteratively adapt the polytope in the parameter space

based on the objective function's response. These operations enable the algorithm to navigate through the optimization landscape without the need for derivative calculations, a major advantage. The first step is reflection, where the algorithm mirrors the worst performing point across the centroid of the other points in the polytope, creating a new point. This process is akin to reflecting an image across a mirror line, seeking a better position in the search space. Then, expansion extends the simplex further in the direction of the reflected point, reaching towards areas of potential lower function values, essentially capitalizing on the successful direction found during reflection. If reflection fails to yield an improvement, the algorithm tries contraction. This step pulls the simplex inwards, towards the centroid, aiming to probe the nearby space more carefully for a better solution. Lastly, when both reflection and contraction fail to find a better solution, the algorithm employs shrinkage, aiming to reduce the overall size of the simplex, bringing all its points closer to the best point.

These operations allow the Nelder-Mead algorithm to adaptively search the parameter space. The combination of these steps enables the algorithm to efficiently explore and exploit the landscape of the objective function, moving towards the optimum solution even in the absence of gradient information. The Nelder-Mead algorithm is efficient in optimization of low-dimensional problems.

5.2.3.2 Genetic algorithm

Genetic algorithms are popular metaheuristic optimization algorithms that are inspired by the process of natural selection (Holland, 1992). The nondominated sorting genetic algorithm II (NSGA-II) is a widely used algorithm in solving multi-objective optimization problems (Deb et al., 2002), and has been used in machine learning applications in the

water engineering field (Li and Vanrolleghem, 2022). The algorithm starts with initializing a population of individual solutions (chromosomes) that are evaluated using the predefined fitness evaluation function. Chromosomes with the highest fitness are selected as the parent chromosomes for the subsequent processes and other chromosomes are eliminated. The non-dominated sorting allows determining a set of optimal solutions on the Pareto front thar are ranked by the evaluation of their fitness. Moreover, a two-point crossover was used to produce new offspring from the parent chromosomes until the new offspring count reach the total population count. This operator randomly selects two points within the individuals' sequence of hyperparameters and exchanges the segments between these points, creating offspring with mixed traits from both parents. Lastly, mutation was also utilized to randomly modify chromosomes with a chosen probability of 20%. The implementation of NSGA-II for optimizing the AdaBoost model parameters utilized the DEAP (Distributed Evolutionary Algorithms in Python) framework. This optimization was designed to target four critical hyperparameters of the AdaBoost model: the number of estimators, the maximum depth of the decision trees, the learning rate, and the loss function. NSGA-II's non-dominated sorting employed allows determining a set of optimal solutions on the Pareto front thar are ranked by the evaluation of their fitness.

5.2.4 Model interpretability

Feature importance was leveraged to interpret the AdaBoost model results by quantitatively measuring the impact of each feature on the predictive power of the model. In the context of AdaBoost, the feature importance scores are computed based on the contribution of each feature to the reduction of variance in the predictions across all the trees in the ensemble (Louppe, 2014). This method aggregates the weighted importance scores assigned to each

feature within the individual decision trees constituting the AdaBoost ensemble. The weights are proportional to the reduction in prediction error brought about by each feature, reflecting their relative importance in the model's decision-making process. By analyzing these importance scores, we can interpret the model in terms of which features are most influential in predicting the outcome, thereby providing insights into the underlying model outcome-generating process.

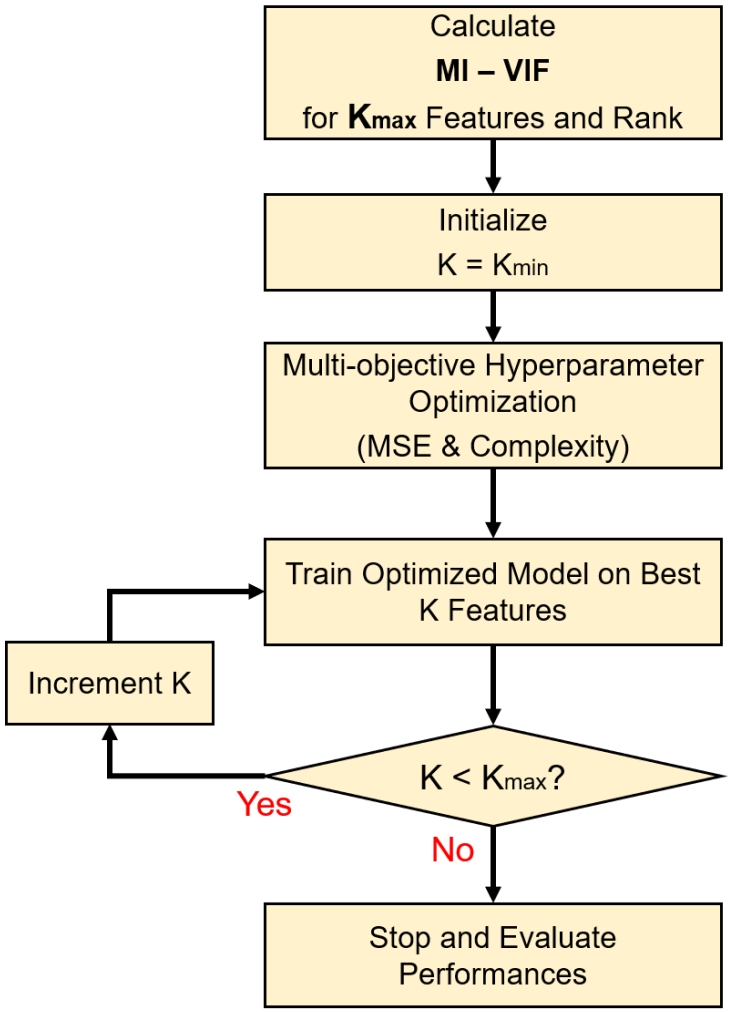


Figure 5-1 General scheme of the suggested automated procedure for optimal model selection using both mRMR for feature selection by calculation of MI and VIF, and HPO.

This was implemented using the NSGA-II and Nelder-Mead algorithm. Kmin and Kmax are the boundary of the number of features that are to be investigated.

5.3 Results and Discussion

5.3.1 Exhaustive search grid search output

For HPO in developing ML tools for wastewater applications, grid search optimization has been predominantly used. For example, it was employed to optimize ML models for biogas production from anaerobic digestion (Sappl et al., 2023), bio-hydrogen from dark fermentation (Hosseinzadeh et al., 2022), and activated sludge effluent quality (Wang et al., 2021; Park et al., 2022; Xu et al., 2023). Despite being computationally expensive, grid search optimization has always been adopted thanks to its success in achieving high accuracy (Ly et al., 2022). Yet, grid search optimization does not consider the possible associated increase in model complexity. Such an increase may negatively affect the model's computational efficiency, interpretability, and susceptibility to overfitting. In this section, we pre-defined a large range of values for Adaboost's key hyperparameters, see Table 5.1. Adaboost model was selected given its high performance in predicting N_2O emissions from WWTPs and its high ability to magnify its performance as a result of HPO.

Table 5-1 Definition of the investigated hyperparameter grid-space

Hyperparameter	Minimum	Maximum
Number of estimators	100	1200
Maximum depth of tree	2	30
Learning rate	0.001	1.25
Loss function	Linear, square, exponential	

Model complexity stems from the employed number of features (features-driven complexity) and/or the selected hyperparamters (HP-driven complexity). Figure 5.2 shows the averaged effect of various combinations of number of estimators and the number of input features on the AdaBoost model performance. Each cell represents the average RMSE of the model aggregated over various combinations of other hyperparameters using k-fold cross validation data. The heatmap values exhibit highly irregular pattern indicating there is no clear trend of the combined effect of the number of model estimators. However, it was shown that the lower number of features (i.e., four input features) scored higher RMSE values within the range of plotted number of estimators. However, it was shown that the model utilizing only four input features—namely NH_4, NO_2, NO_3, and TEMP—scored higher RMSE values within the range of the plotted number of estimators.

Figure 5.3 represents the results of conducting the same analysis and aggregation for the maximum depth of decision tree and the number of input features. Unlike Figure 5.2, this heatmap exhibits a clear pattern of the effect of tree depth on the model performance. Models with maximum tree depth higher than 10 performed significantly better than models with shallower decision trees. This behavior is aligned with the mathematical representation of the ensemble models, where the number of nodes in the model grows exponentially with increasing the tree depth. This increases the model overall complexity and ability to capture more non-linear relationship.

Furthermore, it can be deduced from Figure 5.2 and 5.3 that the combined approach of searching the hyperparameter grid space in association with the number of input features can lead to finding two closely performing models with varying degrees of complexities. For example, based on Figure 5.3, the average performance of models with 13 input features with tree depths higher than 15 achieved performances similar to models with 9 features with tree depths higher than 20. This indicates that the expected deterioration in performance resulting from feature reduction can be mitigated by further optimizing the model hyperparameters; something that cannot be fully achieved without coupling the feature selection and hyperparameter optimization processes. This synergetic effect is not typically considered in ML models in environmental engineering applications that performed HPO. For example in (Xiao et al., 2018), grid-search were used to select the optimal hyperparameters for tree-based extreme gradient boosting (XGBoost) to predict $PM_{2.5}$ concentrations. Another example is Hwangbo et al. (2021) where the DNN model hyperparameters were optimized using grid-search without considering the effect of number of features on the model complexity and the resulting hyperparameters. However,

the HPO implemented in these studies overlooked the potential increase in model complexity that might be associated with minor performance increase, which may adversely affect the model's computational efficiency, interpretability, and vulnerability to overfitting.

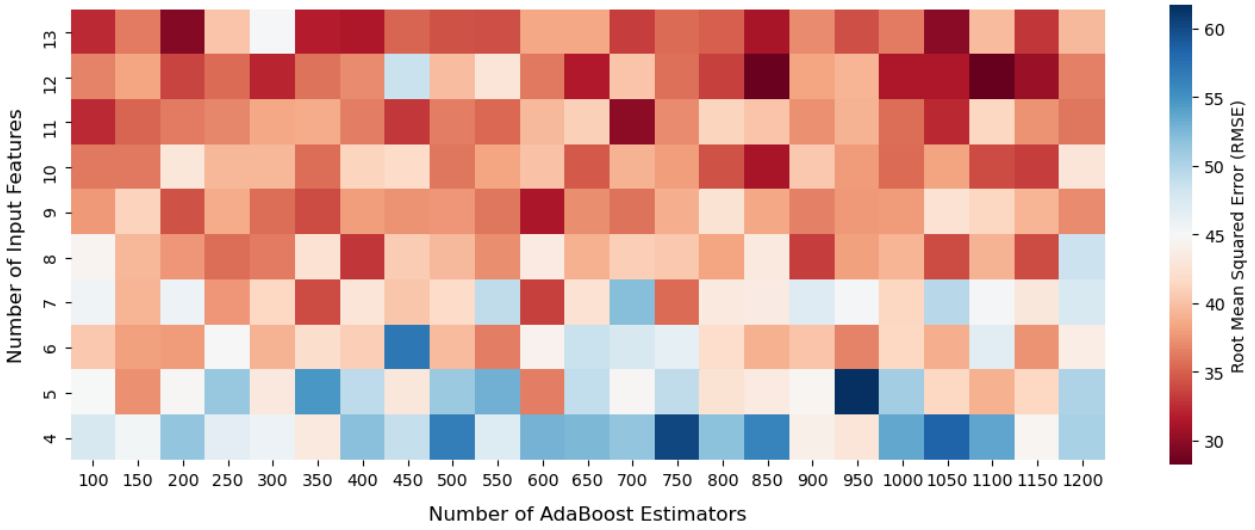


Figure 5-2 Heatmap of RMSE calculated using cross-validation based on every combination of number of AdaBoost decision tree estimators and the number of input features. A lower value (red) stands for a higher model performance, and higher values (blue) stands for lower model performance.

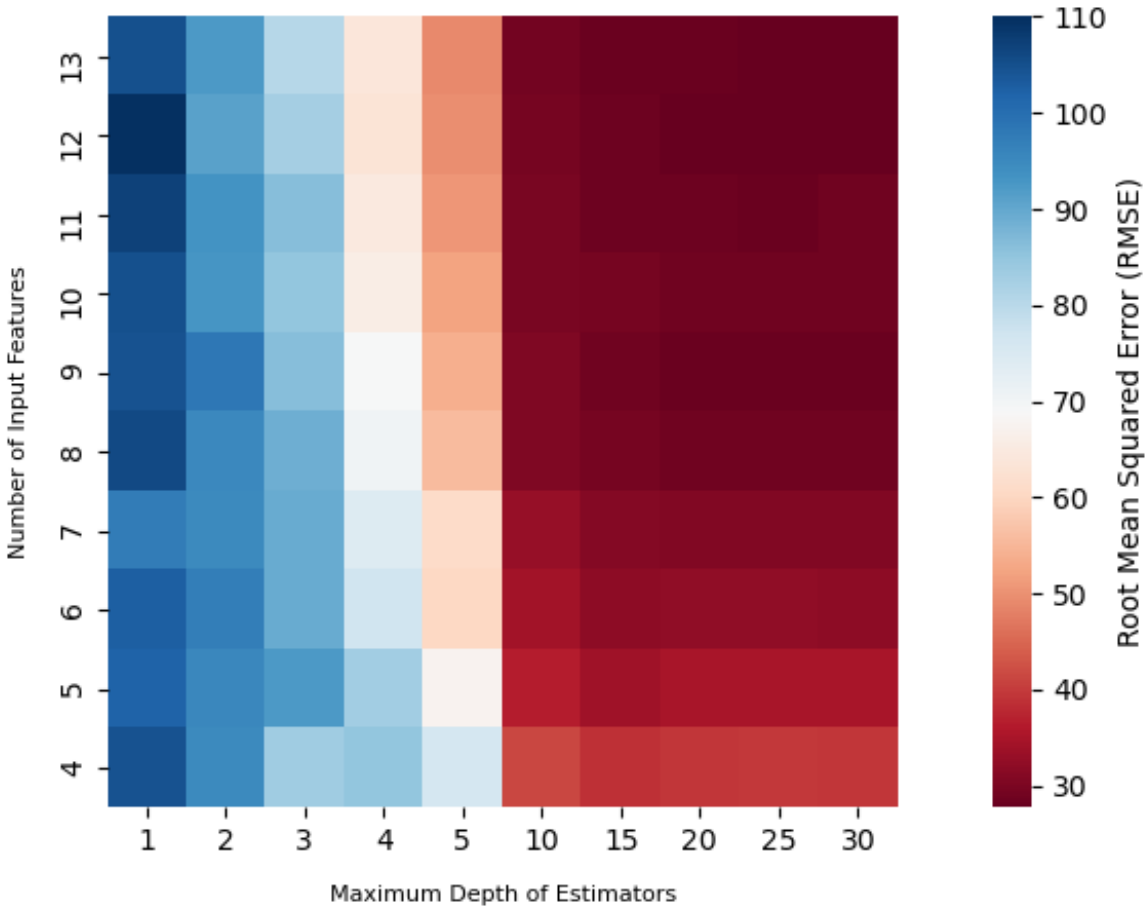


Figure 5-3 Heatmap of RMSE calculated using cross-validation based on every combination of number of explored AdaBoost based decision tree maximum depth and the number of input features. A lower value (red) stands for a higher model performance, and higher values (blue) stands for lower model performance.

5.3.2 Investigating model complexity-performance relationship

To further investigate the synergetic effect of model hyperparameters and number of input features on model performance, a comparative analysis between four extreme models with adverse degrees of complexity and input features was performed as shown in Figure 5.4. The four selected models represented high and low complexity based on hyperparameters matched with the minimum viable and maximum available number of input features. The

hyperparameter values and the number of input features for each model (Models 1 – 4) are shown in Table 5.2, as well as the same information for the model that was ultimately selected with the proposed approach (Model 5).

Model 1 displayed the lowest prediction error values across training, testing, and cross-validation datasets, in comparison with the other models. RMSE and MAE values of 27 and 12.5 were obtained for the testing dataset. The significantly higher RMSE values is a result of the skewed data distribution which typically inflates RMSE values, which was discussed in detail in Chapter 4. Interestingly, model training error, either RMSE or MAE, showed significantly lower values which were near zero. Such discrepancy between training and testing is indicative of overfitting where the model's complexity leads to a memorization of the training data, including noise, thereby impairing its ability to generalize to new, unseen data. In contrast, Model 2, which maintains the maximum number of input features but with extremely much lower complexity compared to Model 1, showed smaller gap between training and test error metrics, aligning with the hypothesis that the reduction in complexity prevented model overfitting and thus resulted in more consistent performance across various datasets. Meanwhile, it achieved similar RMSE, with a test value near 27, and slightly higher MAE, around 16 for the test dataset.

Similar to Models 1 and 2, Models 3 and 4 were trained using only four input features, with high complexity for Model 3 and low complexity for Model 4. Both models exhibited very similar performance on testing. The RMSE and MAE values were higher than those of Model 1 and 2, with an RMSE of 39 and 38, and an MAE of 19 and 20 for the test dataset for Model 3 and 4, respectively. This confirms that using a small number of features will have a performance penalty, despite the increased complexity. Model 3 showed the highest

gap between training and testing datasets, indicating that the increased complexity with fewer features tends to increase overfitting. Whilst Model 4 exhibited a slightly higher error than Model 3 due to its lower complexity, its performance across training, testing, and cross-validation is indicative of less overfitting.

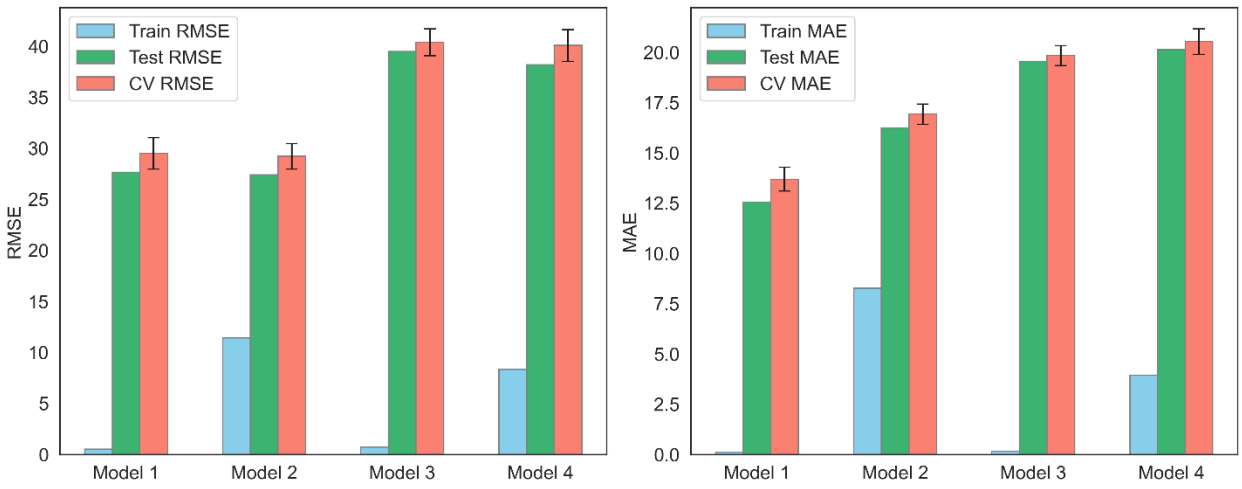


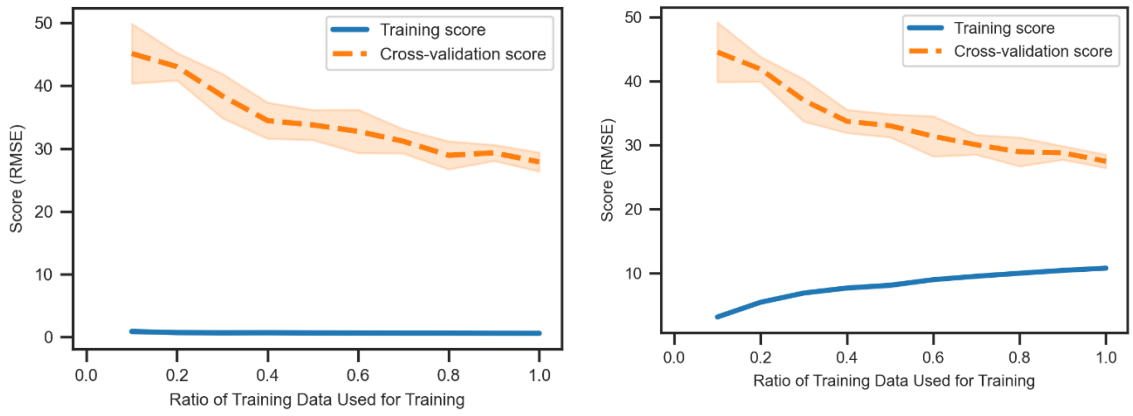
Figure 5-4 Performance comparison of the four selected models based on RMSE and MAE for the training, cross-validation, and testing sets.

Table 5-2 Definition of all the AdaBoost model candidates

Model	Number of Input Features	Number of Estimators	Decision Tree Depth	Selection Method
Model 1	13	850	30	High features high complexity - Manual selection from grid-search results
Model 2	13	250	10	High features high complexity - Manual selection from grid-search results
Model 3	4	900	30	High features high complexity - Manual selection from grid-search results
Model 4	4	150	15	High features high complexity - Manual selection from grid-search results
Model 5	8	231	11	Based on coupled optimization of complexity and performance using genetic algorithm and mRMR
Model 6	8	800	18	Based on uncoupled optimization of performance and feature selection based on mRMR

The overfitting and generalization ability of the four models were further investigated by plotting the learning curves for each model, offering insights into diagnosing their bias and variance with varying volumes of data volumes used for model training. Figure 5.5 shows that the models with high complexity (Model 1 and Model 3) have a persistent gap between training and cross-validation RMSE at all data volumes, pointing to the model's ongoing struggle with overfitting due to its high complexity. For the less complex models (Model 2 and Model 4), showed a converging behaviour between the model training and cross-validation RMSE, suggesting that the models captured less variance from the data, and that the models benefit from more data.

Combined together, it can be deduced that HP-driven complexity makes the model very prone to overfitting. Moreover, the model's potential to overfit increases as complexity increases and the number of features is limited. This is in spite of the fact that a slightly higher performance can always be achieved at higher HP-driven complexity. This is more observed at higher number of features.



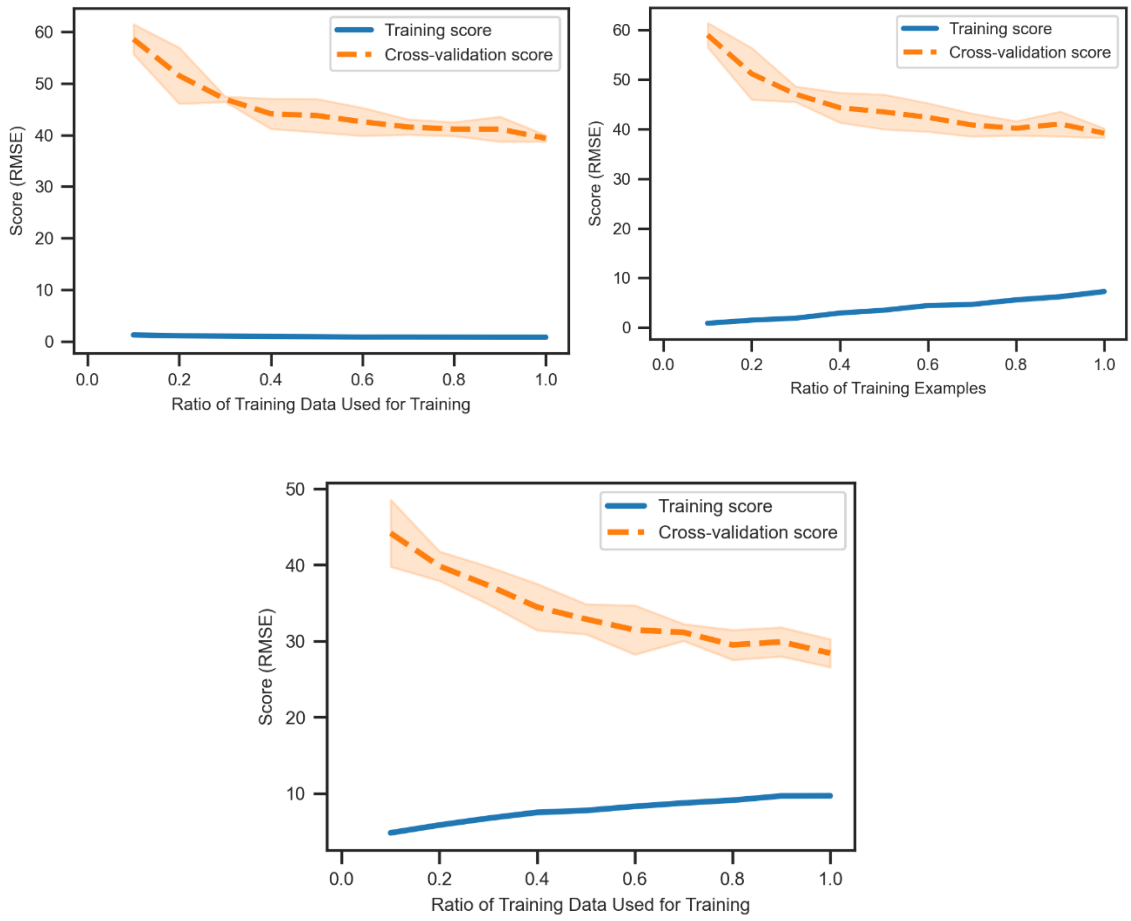


Figure 5-5 Learning curves comparison of five models based calculated as the RMSE of the model prediction for both the training and cross-validation datasets when the model is being trained on varying sizes of data as a percentage of the full training dataset.

Prior to deployment of ML models for N_2O applications, trust in the model predictions is needed by operators and decision-makers. A model will be trusted if the model predictions are aligned with domain expert knowledge, which requires the model to be interpretable in the first place. There is a relationship between model complexity and interpretability; models with less complexity are easier to understand in terms of the underlying mechanisms, but are possibly not able to capture the highly nonlinear patterns in the data.

On the other hand, highly complex models can yield predictions with significantly higher accuracy without understanding how predictions are being made, often called “black-box” models. However, for the same model, a wide range of complexities can be found, based on number of features and values of hyperparameters, indicating that the same relationship between complexity and interpretability is valid. It is necessary to investigate the implications of the level of complexity of the selected model on the interpretability of model results.

Figure 5.6 shows the feature importance as derived from AdaBoost. In models trained with 13 input features, the model with greater complexity demonstrates a propensity to attribute increased importance to the most influential feature (NO_2). This contrasts with the model of lower complexity, where importance values are slightly more evenly distributed among the features. This observation suggests that high complexity models may be more sensitive to the most predictive features, potentially capitalizing on complex interactions or non-linear relationships that are less pronounced in simpler models. On the other hard, models trained with only four input features showed an increase in the overall feature importance values. This compensatory mechanism suggests that when the model is constrained by fewer features, each feature's contribution is inherently amplified to maintain predictive performance. This is consistent with the principle that in a model with fewer inputs, each input assumes a greater relative significance. The observed patterns in feature importance underscore the complex relationship between model complexity, interpretability, and predictive performance. Increased model complexity resulted in an elevated dependence on single, yet highly important, feature (NO_2). However, the increased dependence on a single feature without significant increase in model performance on the testing data is

indicative of overfitting. Furthermore, given the different pathways affecting N_2O net production such as nitrifier nitrification, nitrifier denitrification, heterotrophic denitrification, and abiotic pathways, it is less likely that a single feature can contribute most to N_2O production data.

It is essential to acknowledge that while the feature importance scores provide valuable insights into the predictive capabilities of various parameters within the model, they do not directly imply causation. For instance, the importance attributed to features like NO_2 reflects their statistical significance in the context of the predictive framework, rather than a definitive causal relationship with N_2O emissions. Future research could benefit from incorporating causal discovery methods alongside predictive modeling to further elucidate the underlying mechanisms that drive N_2O emissions based on the model.

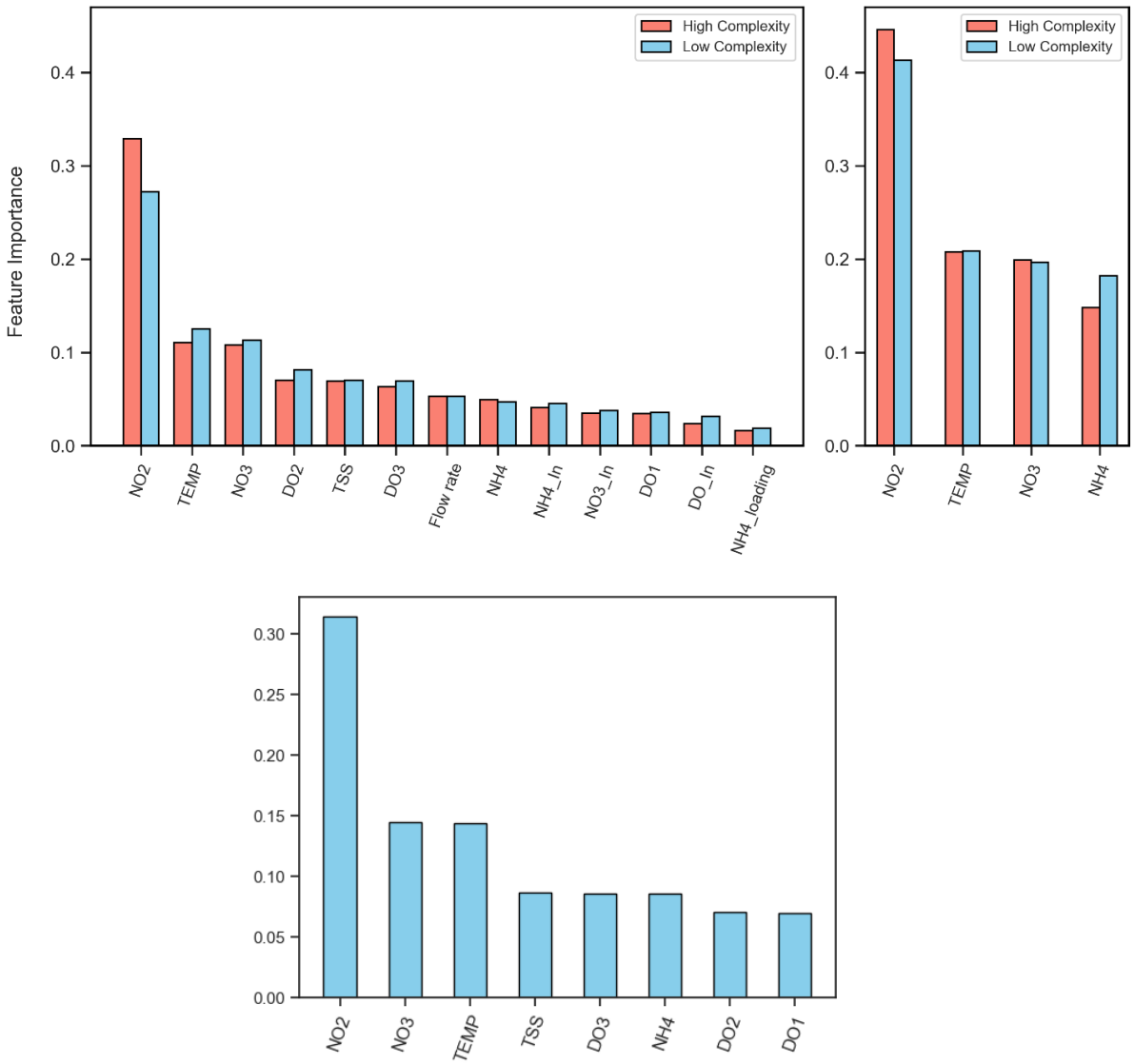


Figure 5-6 A comparison of feature importance score calculated based on the AdaBoost algorithm for (a) Models with 13 input features with varying complexity. (b) Models with four input features with varying complexity. (c) Model with number of input features and hyperparmeters selected based on the developed algorithm combining mRMR and NSGA-II

5.3.3 Selection of optimization algorithm

The previous analysis demonstrated that a multi-objective optimization algorithm is needed to strike the desirable balance between model performance on one side, and HP-driven complexity and feature-driven complexity on the other side. Hence, in this section, an extended search of optimal combination of input features and model hyperparameters were implemented using both the Nelder-Mead and the NSGA-II algorithm to select the model with the best performance – complexity trade-off based on cross-validation. In contrast to the Nelder-Mead algorithm where the optimal solution is given, the NSGA-II provides Pareto front set of solutions. Figure 5.7a and b show a comparison in terms of both RMSE and complexity between the NSGA-II Pareto front solutions plotted as boxplots and the Nelder-Mead's optimal solution showed as points. Based on Figure 5.7a, it is clear that the solution provided by the Nelder-Mead's algorithm showed higher RMSE than most of the pareto front solutions provided by NSGA-II for all number of input features. The RMSE provided by Nelder-Mead exceeded the 75^{th} quantile for all input features and the maximum RMSE for some input features of the solutions provided by NSGA-II. Conversely, Figure 5.7b shows that the complexity functions of the models selected as optimal by Nelder-Mead was lower than the 25^{th} quantile for all input features and the minimum complexity for some input features of the solutions provided by NSGA-II. Figure 5.8 shows a detailed visual comparison of model performance across a spectrum of complexity levels and input feature counts, as determined through the multi-objective optimization processes using both the Nelder-Mead algorithm and NSGA-II's most effective solutions. This three-dimensional scatter plot highlights the relationship between the number of input features, the model's complexity, and the resulting average RMSE

obtained from cross-validation datasets. It was demonstrated that the increase in complexity does not always yield higher model accuracy. For example, Figure 7 shows different levels of complexity for the NSGA-II points with a very similar RMSE at the same number of features. Such an unnecessary complexity may be chosen by an algorithm that only focuses on adding accuracy – even if minor – such as grid search. The results also confirms that the NSGA-II was able to yield more balanced models in their Pareto Front solutions, with achieving overall higher RMSE than the Nelder-Mead algorithm, which failed to provide reasonable trade-off between performance and complexity. The superiority of NSGA-II over Nelder-Mead can be attributed to the fact that the optimization objective is highly non-linear and irregular, as shown in Figure 5.2. This non-linearity was also shown in previous research optimizing deep neural network to predict wastewater treatment plant variables in the hyperparameter space of the deep neural network (Zhu et al., 2022). Developing an automated approach that performs HPO while considering model complexity and preventing overfitting is of great importance to environmental engineering applications, particularly modelling N_2O emissions. However, according to the literature analysis of the highly cited environmental engineering research articles performed by (Zhu et al., 2023), about 80% of these studies did not perform an automated HPO such as grid or random search. Moreover, using the best provided hyperparameters calculated by grid-search automated tools provided by open-source packages such as *Python Sklearn* without deeper exploration of the explored grid might lead to selection of a highly complex model that only adds minor extra-performance to other solutions. An approach to overcome this issue was implemented by (Xiao et al., 2018), where the possible parameter set were explored and parameters that provide an improvement in accuracy larger than a certain

threshold were selected. However, grid and random search algorithms are unidirectional compared to metaheuristic algorithms such as NSGA-II, where the optimal solutions get updated based on the feedback from the cross-validation fitness function evaluation (Zhu et al., 2023). This makes genetic algorithm such as NSGA-II superior to other optimization methods, given the complexity of the problem shown in the nonconvex grid-space shown in Figure 5.2 and Figure 5.3. The limitation of such algorithms is the high requirement of computational resources. However, this can be mitigated by utilization of parallel computing, where the algorithm is implemented in parallel parts with each part using one of the available logical cores of the computer processor achieving a faster overall implementation of the algorithm.

It is obvious from the optimization results that the optimal number of input features to the model are ranging from 8 – 10 features to yield reasonable performance and complexity. The choice of the number of features to be used in the deployed model within this range can be highly dependent on decision made by the responsible engineers and researchers. This decision might consider the complexities and cost associated with data acquisition. For example, the current dataset has three DO sensors, a privilege that might not be available for other facilities. Nevertheless, the present automated approach can help providing the range that satisfies the performance and model complexity requirement. The selected set of model hyperparameters and number of input features as the optimal model - denoted in this study as "Model 5" with its hyperparameters defined in Table 3 – followed the same approach. Model 5 showed performance similar to the models with a large number of features (Model 1 and 2) in terms of RMSE. This further shows the efficiency of the

algorithm to find the optimal set of input features, and model hyperparameters while keeping model performance high.

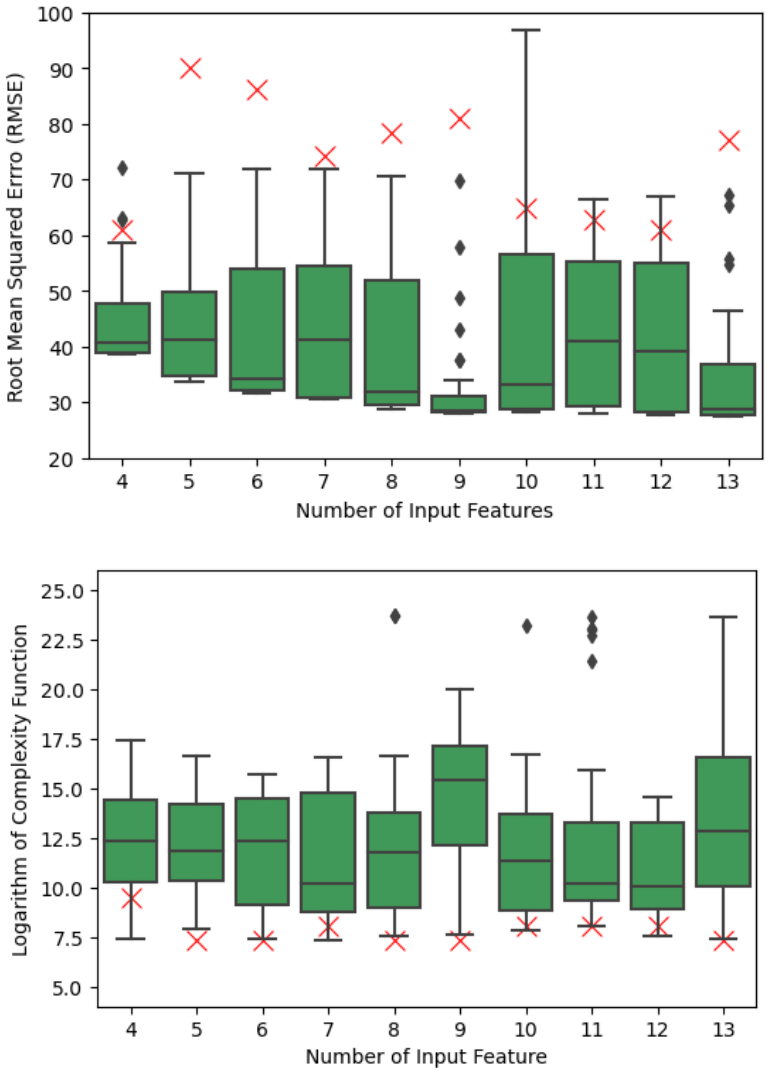


*Figure 5-7 Comparison of the performance of the Pareto-front solutions (box-plots) provided by the NSGA-II and the optimal solution by Nelder-Mead algorithm (red cross points) for each number of input features based on (a) RMSE of the cross-validated results. (b) Complexity represented by the logarithm of the function representing number of maximum nodes in the AdaBoost algorithm. The black diamonds represent outlier (values > 75th percentile + 1.5 * Interquartile range (IQR)).*

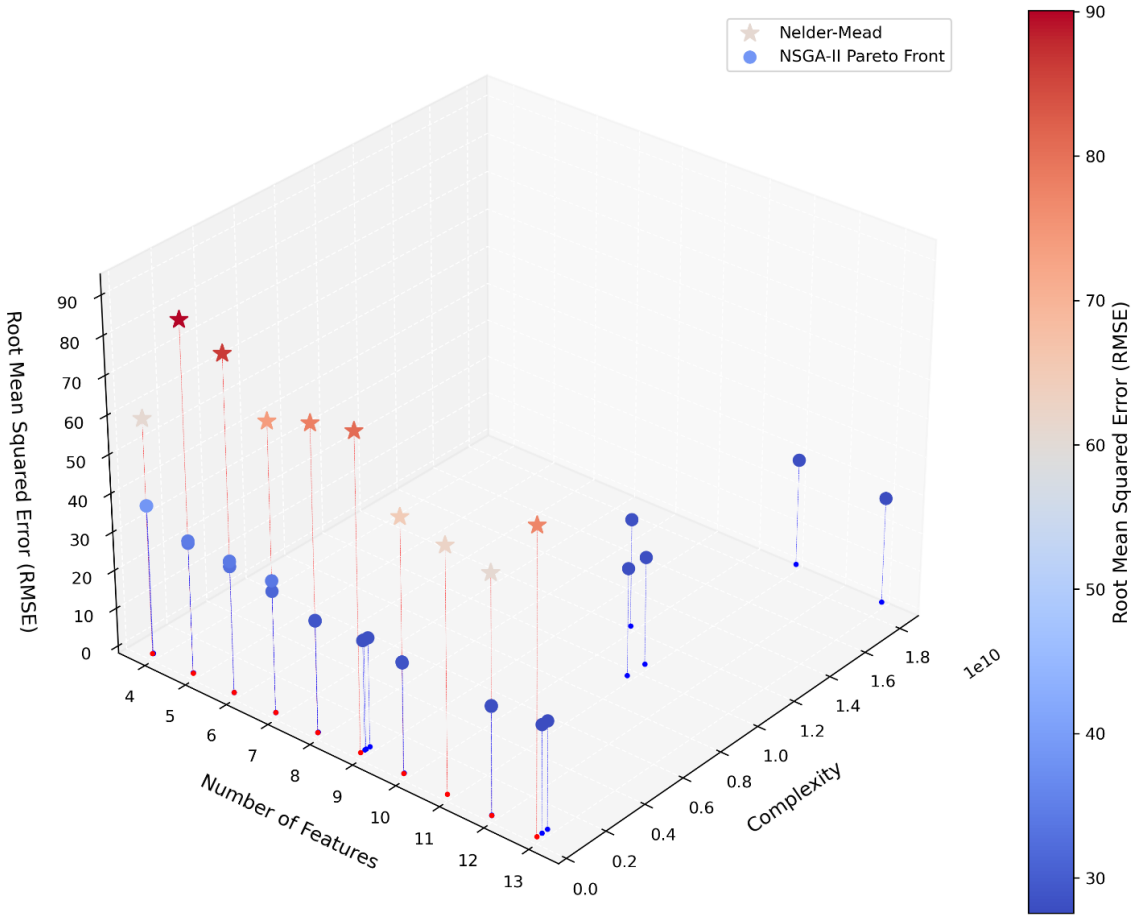


Figure 5-8 Three-dimensional scatter plot representing a comparison between the best selected solutions from the NSGA-II Pareto-front and the optimal solutions by Nelder-Mead algorithm

5.3.4 Model performance comparison and evaluation

Table 5.4 contrasts the performance of the optimized model using the coupled GA and mRMR methods (Model 5) with that of the model utilizing uncoupled feature selection and grid search-based HPO (Model 6), as implemented in Chapter 4, across both training and testing datasets. Performance metrics, including R^2, RMSE, and MAE, reveal a minimal performance decrement as a consequence of the complexity reduction achieved by the developed approach in this study. Notably, the discrepancies between the metrics for

training and testing datasets are more pronounced for Model 6 than for Model 5. This discrepancy underscores the effectiveness of the developed algorithm in minimizing model complexity, thereby mitigating overfitting and enhancing the model's generalization capabilities across unseen data. The investigation of ME revealed further insights into the performance of both models. For Model 5, the ME during training is observed at 1.71 while being reduced to 0.76 during the testing phase, suggesting that despite the slight bias, the model is able to generalize well to unseen data with a more balanced prediction error. Model 6 presents a different pattern, with a negligible ME of -0.06 during training, which shifts to a more pronounced negative bias (ME of -2.33) in the testing phase. This shift implies that while Model 6 achieves nearly unbiased predictions on the training set, it tends to underestimate the target variable when exposed to new data. The negative ME in the testing phase for Model 6 could indicate a model that, despite being well-calibrated to the training data, struggles to maintain this accuracy when predicting on unseen data, potentially due to overfitting or insufficient generalization. The contrasting ME values between the training and testing phases for both models underscore the importance of balancing model complexity and generalization capability. Model 5 demonstrates a more consistent performance, suggesting a better trade-off between model accuracy and complexity thanks to the integrated algorithm used for input feature selection and HPO. In contrast, Model 6, despite its minimal bias in training, reveals vulnerabilities in its predictive stability across different data sets. Furthermore, Model 5 did not suffer from the potential overfitting demonstrated in Model 6, as shown in the learning curves presented in Figure S2, where the gap between model performance on the training and cross-validation sets is decreased when increasing the data volume used for model training. These

results suggest the efficiency of the developed algorithm to yield a more efficient and accurate model compared to conventional methods reported in literature.

Direct comparisons with existing literature on ML models for N_2O emission prediction may be hindered by variations in datasets, monitoring campaign lengths, and the range or lack of detailed performance metrics. N_2O To our knowledge, the study by [44], which reports the highest performance using machine learning on long-term datasets, achieved an R^2 of over 0.90 with a complex DNN. Notably, our previous research (Model 6) implemented in Chapter 4 reached an R^2 of 0.94, showcasing high accuracy. The significance of the current study, however, lies in the developed algorithm's ability to attain comparable accuracy with Model 5—a more streamlined and less complex model. This advancement underscores our approach's efficacy in simplifying the model without sacrificing performance, marking a significant contribution to the field of environmental ML applications.

Table 5-3 Comparison of performance of the optimized model using the developed algorithm coupling GA with mRMR and a model with the uncoupled approach (HPO and mRMR)

Metrics		Model 5	Model 6
R^2	Training	0.98	1
	Testing	0.94	0.94
RMSE	Training	10	1.9
	Testing	27.25	26.27
MAE	Training	7.3	0.7
	Testing	16.6	13.12
ME	Training	1.71	-0.06
	Testing	0.76	-2.33

5.4 Summary

The present study highlights the critical balance between model complexity and performance, underlining the importance of integrating input feature selection and HPO for effective ML model development. The followed approach synchronized feature selection and HPO successfully addressed the challenge of model overfitting to enhance the model's generalizability and interpretability. The use of mRMR for feature selection, coupled with advanced HPO methods like the NSGA-II algorithm, provided an innovative way to optimize both the model's complexity and its performance. The results underscore the superiority of the GA in navigating the complex, non-linear optimization landscape, as evident from the comparative analysis with the Nelder-Mead algorithm.

The presented methodological framework is crucial for environmental engineers and researchers utilizing ML tools, particularly for modelling N_2O emission. Future research should focus on further refining this integrated approach, exploring its applicability to other environmental parameters, and testing its effectiveness in different WWTP contexts. Additionally, considering the computational demands of such advanced modelling techniques, there is a need for developing more efficient computational strategies or harnessing cloud computing resources to make these methods more accessible to researchers and practitioners in the field.

Chapter 6: Synthesis, Conclusions, and Future Directions

6.1 Introduction

This thesis is the outcome of an exploratory journey through two different modelling approaches for N_2O emissions from WWTPs: mechanistic (i.e., knowledge-based), and machine learning (i.e., data-driven) models. Each chapter has contributed distinct insights and findings, reflecting the multifaceted nature of this research. This final chapter aims to weave together the diverse strands of research presented in the preceding chapters, drawing concluding remarks and mapping out potential avenues for future research. By synthesizing these elements, this chapter seeks to provide a cohesive summary and a forward-looking perspective.

6.2 Comparative overview of modelling approaches

Process models could be primarily used as either a prognostic, diagnostic, or educational tool (Hug et al., 2009). Prognostic models are used to predict outcomes based on a specific input, such as using a model to predict the N_2O emissions from a biological reactor based on a specific influent characteristics and operational conditions. On the other hand, diagnostic models can be used to analyze and understand existing conditions or past events. They help in identifying causes of observed phenomena. An example of diagnostic usages of models in the context of the current research is the determination of the pathways and operational conditions causing N_2O emissions from a biological treatment reactor. These models are crucial for problem-solving and improving understanding of studied systems. Lastly, educational models are designed to teach and explain complex systems and processes to water professionals, students, policymakers, or the general public, but this is out of the context of the current research. A deeper look on the status of advancements in

the WWTP's N_2O emissions field would lead to a realization that there is a need for N_2O models to be both prognostic and diagnostic. Using the model for a model-informed quantification is a clear and indispensable goal of the models; however, quantification and reporting of N_2O is merely useful beyond the level of data collection for generating country-wide EFs. Rather, utilities would be mainly looking forward to getting insights associated with the predictions that could help to mitigate N_2O emissions. In this context, the desired model should both be accurate, practical, and interpretable to serve the intended purposes.

Mechanistic models are inherently interpretable thanks to the first-principal equations embedded into the models, and can be referred to as "white-box" models, where the user could see how predictions are made and understand the root-causes of N_2O emissions. Nonetheless, as evident from the results of the mechanistic model calibrated in Chapter 3, they struggle with accuracy, especially when validated with data generated from operational conditions different from those when collecting the calibration data. The accuracy remained limited despite the fact that the model was calibrated and validated using data collected from intense and well-conducted laboratory-scale experiments, which are a more controlled environment than full-scale applications. The model shortcoming in extrapolation to another operational scheme in the validation data might be attributed to the need for different values of kinetic parameters. Given the time span between the two operational phases in the experiment, the microbial species could have been acclimated to the operational and environmental conditions requiring an update in their values when operating on different conditions. The acclimation to operational conditions was previously

investigated in literature (Sabba et al., 2024); however, its effect on N_2O emissions should be further investigated.

Moreover, the calibration process of N_2O mechanistic models was shown to be impractical. Therefore, in Chapter 3, a multi-step calibration protocol was suggested, including sensitivity and identifiability analyses. While this can be understood in the context of academic research where this investigation can lead to better calibration schemes, this complex protocol limits the practical applications of mechanistic models. This was also evident through several calibration protocols that have been suggested for N_2O mechanistic models' calibration using full-scale datasets (Blomberg et al., 2018; Zaborowska et al., 2019; Solís et al., 2022). Moreover, there is no consensus yet on a calibration protocol that has been successfully tested on several case studies, and using long-term data. The attempted calibration efforts are, however, intricate, hindering their potential to be automated. The problem of impractical calibration is because the N_2O mechanistic models, including the two-pathway model used in the current research, suffer from the overparameterization problem. This means that the model has a large number of parameters and there is insufficient informative data (Dochain and Vanrolleghem, P.A, 2005), making the calibration process complicated. For instance, the model calibrated in Chapter 3 has a total of 38 kinetic and stoichiometric parameters, not including parameters related to mass-transfer, diffusivity, and temperature effects. Given this case of overparameterization, either re-estimation of all parameters or studying which parameters to be better estimated would be a time-consuming process. In the case of the current research, sensitivity and identifiability analyses were used to reduce the number of parameters to be estimated, resulting in eventually calibrating the model by only estimating 10 parameters. Although

this number of parameters was remarkably lower than the total number of parameters, the number of identifiable parameters cannot be generalized given that there is no full agreement in literature on the set of parameters to be estimated. However, there is an agreement on some parameters, such as μ_{AOB}, and the oxygen affinity constants for AOB. It is also worth noting that the selection of parameters to be estimated depends on the cut-off specified to select the most influential parameters from sensitivity analysis, whether it is a fraction of the certain parameters starting from the ones with the highest influence or taking the (n) most influencing parameters (Dochain and Vanrolleghem, P.A, 2005).

Another factor that affects the final subset of estimated parameters is the practical identifiability analysis (Dochain and Vanrolleghem, P.A, 2005). In the mechanistic model in Chapter 3, investigation of the sensitivity functions of the parameters with respect to the same state variable was used to detect collinearity among model parameters, helping to reduce the number of parameters that can practically be estimated during calibration. This approach effectively highlights how variations in parameters influence the model outputs, thereby pinpointing parameters that exhibit a lack of independent identifiability due to their interdependencies. However, other methods for parameter selection were also used in literature, with the analysis of confidence intervals generated from parameter estimations as the most prevalent (Solís et al., 2022). Although this also offers insights into poor identifiability, it does so from a broader perspective. Large confidence intervals are indicative of parameters that cannot be estimated with precision, a phenomenon that may not solely stem from identifiability issues but could also be attributed to other factors such as model overfitting, the inherent complexity of the model in relation to the volume of

available data, or both. Thus, fundamentally, the results obtained from this method should not be generalized given that the utilized data may affect the final outcome.

On the other hand, ML are data-driven models that are generally considered opaque, and often referred to as "black-box" due to the high complexity of the model hindering interpretability. This complexity makes the modelling approach powerful and capable of capturing hidden and non-linear relationships in the data, allowing for a high accuracy that could be leveraged for N_2O emissions predictive modelling. In Chapter 4, many models showed acceptable to high accuracy and captured most of the dynamics in N_2O emissions from a full-scale WWTP. For example, an AdaBoost model demonstrated high accuracy across various statistical, and dynamic metrics, enabling the model to be utilized for prediction of N_2O emissions. However, the optimized model had a complex architecture that might reduce interpretability. Interpretability is needed not only to understand the factors affecting N_2O emissions and guide decision-making, but also for justifying that this is the right model for the use case and that the model does not rely on spurious relationships. Here, the interpretability was investigated using post-hoc methods, such as permutation feature importance, in which a global interpretation of how the model makes decisions was investigated and successfully aligned with the domain knowledge to ensure the model overall validity. As the increased complexity generally reduces interpretability, a reduction in complexity was another model development goal. The type of model is not the only factor affecting complexity; in particular, factors such as number of input features and model hyperparameters also have a great effect on the resulting model's complexity. This was achieved using a combined feature selection-hyperparameter multi-objective optimization method using GA in Chapter 5. These efforts were effective in developing a

framework that leads to high accuracy models with relatively low complexity in terms of number of features and model architecture, and reduced overfitting. Although reducing overfitting is indicative of a model's ability to generalize, a distinction should be made between overfitting and generalization. Generalization is the model's ability to apply its learned patterns to new, unseen data effectively. This ability is crucial for the model to be useful after deployment in full-scale WWTPs, where data conditions and characteristics can change over time. Therefore, further demonstration of ML models is still needed to confirm their generalization ability in N_2O emissions applications.

6.3 Data requirements and utilization

Developing an effective model, irrespective of its type or purpose, fundamentally requires a thorough data collection phase. Emphasizing the importance of data gathering is crucial given its profound impact on modelling accuracy. Particularly, the number and type of monitored variables, which represent the different attributes or characteristics measured, bring unique challenges and considerations in data collection, particularly when contrasting mechanistic models with ML models.

Contrary to ML models, which are inherently data-driven, mechanistic models do not primarily rely on extensive datasets for their construction. However, data variables are essential for providing input to these models, as well as for their calibration and validation. This process involves utilizing the monitored variables that represent the model's state variables. For example, the mechanistic model in Chapter 3 was calibrated using NH_4^+, NO_2^-, and N_2O data. Although the primary purpose of the model is to describe N_2O concentrations, calibration using NH_4^+ and NO_2^- data was essential given their interaction with the autotroph kinetics and stoichiometry that would eventually affect N_2O

concentrations, and the required mass balance of nitrogen species. However, the model could not be calibrated using NO and NH_2OH data despite being critical intermediates in N_2O production pathways. This is a primary challenge in creating accurate N_2O mechanistic models because not all potentially relevant species are commonly measured in full-scale WWTPs. Thus, the provided mass balance of nitrogen species by the model cannot be fully calibrated or validated by experimental data. These intermediates are not available in most of the current N_2O calibration attempts using full-scale datasets (Blomberg et al., 2018; Zaborowska et al., 2019; Solís et al., 2022). Other variables were used as input to the model, such as influent characteristics, DO concentrations in the bulk liquid, and the dosing rate and concentration of the supplemented carbon source. These variables, while not explicitly used in calibration, are essential to ensure that the model operates under the same conditions as the experiment.

On the other hand, unlike mechanistic models, ML models have a higher flexibility in terms of the variables needed as input to the model. An ML model could be adapted and trained using the available features. Nevertheless, a small number or irrelevant input variables might not be enough to train a reliable model that is not biased due to missing an important feature. However, the model performance might not be affected if another feature is available that is highly correlated to the missing feature. Moreover, if the collected data has an extremely large number of features, the data are considered highly dimensional. A model trained on high dimensional data has a high risk of overfitting. In Chapter 4 and 5, the mRMR feature selection method was leveraged to select the most relevant data to the target variable, yet include the least possible redundancy. Aside from reducing data dimensionality, feature selection was important to reduce the number of features used in

the model to reduce the complexity of future monitoring campaigns in WWTPs. A high number of monitored parameters means higher cost and increased complexities resulting from sensor installation or lab measurements.

Another challenge is when a variable is absent from the dataset used for ML model training, while the variable is known to have a mechanistic effect on N_2O emissions. An example is the ML model trained in this research, where features such as pH, and chemical oxygen demand (COD) were missing, despite the potential effect they might have on the N_2O emissions. Although the model may still benefit from further improvements, the absence of these variables did not demonstrate a significant negative effect on the model performance. This can be attributed to the specifics of the reactor and process under study, as pH will likely not have a significant effect on N_2O emissions as long as the pH variability remains within the neutral pH range (Su et al., 2019). Likewise, the impact of COD on N_2O emissions will potentially affect the rate of N_2O emissions under low COD/N ratio conditions (Pan et al., 2013b), which was not the case in the experimental data used in Chapter 4 and Chapter 5. This emphasizes the need for a process understanding of the experimental data before using ML tools and also the need for continuous monitoring of model performance after deployment. If monitoring of the model performance is not performed, there is a potential risk of inaccurate model results in case the operational and/or influent conditions changed in the future. Thus, the model is likely to perform better if the operational conditions are not too far from those present during collection of the training data.

6.4 A comprehensive methodological framework for ML models

In Chapter 4 and 5, a comprehensive methodological framework was introduced and refined that considers the development of an ML model to predict N_2O emissions. This framework focused on the aspects that help ML models be more suitable particularly for online monitoring applications that can support decision-making processes. By adopting these goals, the objective of the framework focused not only on high accuracy, but also complexity reduction and interpretability improvement. Moreover, the lack of familiarity with ML in the environmental engineering field and the special challenges of wastewater applications, especially N_2O emissions, urged the need for providing tailored ML methodologies that tackle such challenges. The following is the final refined framework for ML models that has been developed in thesis:

- **Data pre-processing:** in addition to regular pre-processing activities such as missing data handling, data synchronisation, feature scaling, the developed framework focuses on making the outlier detection method more adaptive and interpretable. An adaptive method applicable to the special characteristics of wastewater data was needed: wastewater data include high interactivity between features and are characterized by skewed distributions. The developed method using Mahalanobis distance identifies multivariate outliers using a parametric method by fitting the best distribution to the calculated distances, and selection of outliers based on a threshold. Unlike complex non-parametric methods in which the adopted algorithm is less transparent, the developed method is understandable. Moreover, the method does not rely on a specific distribution just as the interquartile (IQR) boxplot method.

- **Preliminary model selection:** to help proceed towards the optimal models, this step ensures having a variety of models with different levels of interpretability and complexity. The models are trained, optimized, and compared using AIC.
- **Feature selection and hyperparameter optimization:** in Chapter 4, feature selection was efficiently conducted using the mRMR method to reduce the number of features with the minimum effect on model performance, and hyperparameters were optimized using an exhaustive grid search method. The overall model showed high performance on the testing dataset. However, this overlooked the dimension of complexity as it was addressed only from the perspective of number of input features. This gap was considered in Chapter 5 by integration of feature selection and hyperparameter optimization and using a using multi-objective GA optimization to select hyperparameters that both achieve high performance and low complexity. This modification further improved the model to be more efficient by reducing complexity.
- **Comprehensive performance evaluation:** performance metrics act as an alarm to poor model choice, misconducts in data pre-processing, or sub-optimal hyperparameters. As it is a way to judge the model performance, the evaluation metrics should be comprehensive to represent the model performance and shortcomings under different scenarios. The developed framework focused on a holistic performance evaluation by using different statistical metrics such as RMSE, MAE, and R^2, dynamic evaluation such as time series plots, and cross-correlation, in addition to error analysis.

- ❖ **Feature importance analysis:** the decision-making process of the models was investigated globally by feature importance methods such as permutation feature importance. The decision-making process of the model was aligned with domain knowledge to ensure that the model does not rely on spurious relationships. This step is not stand-alone from previous steps, and all steps affect each other. For example, the feature selection that reduces redundancy using mRMR method helps the model rely on more realistic relationships.

6.5 Conclusions

This thesis has explored the topics outlined in the preceding sections, leading to the following conclusions:

- The calibration of a mechanistic model against experimental data revealed the selected model's potential for accurate predictions within the same operational scheme used for calibration. However, its generalization across varying operational schemes was limited, underscoring the challenges in adapting such models to diverse treatment settings.
- The calibration process for the mechanistic model was found to be both time-consuming and complex, highlighting a critical gap in the field: the absence of a standardized approach for calibrating models to N_2O emissions data in wastewater treatment scenarios.
- The concentration of N_2O in the reactor headspace increased with increasing the DO during the aerated periods in an IFAS-SBR reactor with alternating aeration operation, an indicating increase in N_2O emissions.

- Through scenario analyses of the mechanistic model, temperature was found to influence setting a DO setpoint that satisfies ammonia removal and N_2O emission mitigation in an IFAS-SBR reactor with alternating aeration operation.
- Mechanistic models can still be a valuable tool for exploring "what if" scenarios, aiding in the design and planning phases by leveraging the best available knowledge to anticipate N_2O emissions.
- ML models, including k-nearest neighbours, adaptive boosting, and deep neural networks, demonstrated high accuracy and reliability in predicting N_2O emissions from a full-scale, long-term dataset.
- Deep learning is not essentially needed to predict N_2O emissions as assumed in previous studies, as simpler models, such as kNN, achieved comparable, if not superior, performance.
- Comprehensive evaluation of the performance of ML models revealed comparative insights across different models' performance to aid model selection.
- The develop multivariate outlier detection using the Mahalanobis distance provided a simpler and more interpretable methodology than non-parametric methods, and more adaptable than traditional statistical approaches.
- Feature reduction through the mRMR method successfully reduced the data acquisition efforts without significantly compromising model accuracy, thus facilitating more efficient monitoring and measurement strategies in WWTPs.
- The developed method to integrate feature selection with hyperparameter optimization through a multi-objective genetic algorithm (using NSGA-II) successfully enhanced the efficiency of model tuning.

- Using the NSGA-II algorithm was superior for the above multi-objective optimization problem than a more traditional optimization algorithm such as the Nelder-Mead method.

6.6 Recommendations for future research

Based on the research, experience, and insights gained through the work done in this thesis, it is recommended that further investigation of ML approaches, or a combination of ML and mechanistic approaches termed "hybrid modelling" can be pursued, through study of the following research gaps:

- ❖ The mechanistic model in this study explored the potential combined effect of temperature and influent NH_4 loads on N_2O emissions and the required DO for optimal N_2O mitigation and NH_4 removal. The findings warrant further experimental investigation to validate these results, including the effect of temperature and pH on N_2O emissions.
- ❖ Comparing the efficiency of model calibration with and without data on the intermediates such as NH_2OH and NO is needed to assess the necessity of their availability in calibration and validation data.
- ❖ Developing good modelling practices and a unified calibration protocol for N_2O emissions will make mechanistic modelling by practitioners feasible.
- ❖ In this research, the interpretability of ML models was approached by reducing complexity and applying post-hoc global interpretability methods, such as permutation feature importance. For future work, it is recommended to explore the use of instance-based interpretability methods, like Shapley or SHAP values (Molnar, 2022), to gain deeper insights into specific instances where the model

faces difficulties. Additionally, investigating the impact of designing features that embody a mechanistic understanding on the interpretability of ML models would be valuable. This could help in further elucidating how feature engineering influences model transparency and decision-making processes.

- The input features to the developed ML models for N_2O prediction did not include certain variables that could potentially affect N_2O emissions under other conditions, such as pH and COD load, due to their unavailability in the online monitored data. Future work should explore the feasibility of incorporating these variables, derived from low-frequency laboratory measurements, into the input variables used to train the ML model. This investigation could provide insights into improving model accuracy, predictive capability, and mitigating potential bias that may arise in case of change in the operational conditions.
- The outstanding performance of ML models demonstrated in this research, coupled with the knowledge-based approach of mechanistic models, suggests that hybrid modelling warrants further exploration. For robust development and evaluation, it is advisable to utilize long-term datasets. Additionally, it is crucial to investigate the optimal structure and methodology for the hybrid model. This includes determining the precise functions and integration strategies of both the mechanistic and data-driven components within the model, to maximize the effectiveness and accuracy of the hybrid approach. A detailed illustration of hybrid modelling, its types, and anticipated opportunities is available in (Schneider et al., 2022).

Bibliography

Adouani, N., Limousy, L., Lendormi, T., Sire, O., 2015. N_2O and NO emissions during wastewater denitrification step: Influence of temperature on the biological process. Comptes Rendus Chim. 18, 15–22. https://doi.org/https://doi.org/10.1016/j.crci.2014.11.005

Ahn, J.H., Kim, S., Park, H., Katehis, D., Pagilla, K., Chandran, K., 2010. Spatial and temporal variability in atmospheric nitrous oxide generation and emission from full-scale biological nitrogen removal and non-BNR processes. Water Environ. Res. 82, 2362–2372.

Altman, N.S., 1992. An introduction to kernel and nearest-neighbor nonparametric regression. Am. Stat. 46, 175–185. https://doi.org/10.2307/2685209

Baeten, J.E., van Loosdrecht, M.C.M., Volcke, E.I.P., 2020. When and why do gradients of the gas phase composition and pressure affect liquid-gas transfer? Water Res. 178. https://doi.org/10.1016/j.watres.2020.115844

Bartram, D., Short, M.D., Ebie, Y., Farkaš, J., Gueguen, C., Peters, G.M., Zanzottera, N.M., Karthik, M., 2019. Refinement to the 2006 IPCC Guidelines for national greenhouse gas inventories: Volume 5 Waste - Chapter 6 Wastewater Treatment and Discharge. IPCC 6.1-6.28.

Bellandi, G., Weijers, S., Gori, R., Nopens, I., 2020. Towards an online mitigation strategy for N_2O emissions through principal components analysis and clustering techniques. J. Environ. Manage. 261. https://doi.org/10.1016/j.jenvman.2020.110219

Bergstra, J., Bardenet, R., Bengio, Y., Kégl, B., 2011. Algorithms for hyper-parameter optimization, in: Shawe-Taylor, J., Zemel, R., Bartlett, P., Pereira, F., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems. Curran Associates, Inc.

Bishop, C.M., 2006. Pattern Recognition and Machine Learning. New York : Springer.

Blomberg, K., Kosse, P., Mikola, A., Kuokkanen, A., Fred, T., Heinonen, M., Mulas, M., Lübken, M., Wichern, M., Vahala, R., 2018. Development of an extended ASM3 model for predicting the nitrous oxide emissions in a full-scale wastewater treatment plant. Environ. Sci. Technol. 52, 5803–5811. https://doi.org/10.1021/acs.est.8b00386

Borisov, V., Leemann, T., Seßler, K., Haug, J., Pawelczyk, M., Kasneci, G., 2022. Deep neural networks and tabular data: A survey. IEEE Trans. Neural Networks Learn. Syst. 1–21. https://doi.org/10.1109/TNNLS.2022.3229161

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L., 1984. Classification and Regression Trees. Wadsworth International Group, New York.

Brotto, A.C., Lake, A., 2022. Reporting Guidelines, in: Ye, L., Porro, J., Nopens, I. (Eds.), Quantification and Modelling of Fugitive Greenhouse Gas Emissions from Urban Water Systems. IWA Publishing, London, UK.

Caranto, J.D., Lancaster, K.M., 2017. Nitric oxide is an obligate bacterial nitrification

intermediate produced by hydroxylamine oxidoreductase. Proc. Natl. Acad. Sci. U. S. A. 114, 8217–8222. https://doi.org/10.1073/pnas.1704504114

Chen, X., Mielczarek, A.T., Habicht, K., Andersen, M.H., Thornberg, D., Sin, G., 2019. Assessment of full-scale N_2O emission characteristics and testing of control concepts in an activated sudge wastewater treatment plant with alternating aerobic and anoxic phases. Environ. Sci. Technol. 53, 12485–12494. https://doi.org/10.1021/acs.est.9b04889

Ching, P.M.L., So, R.H.Y., Morck, T., 2021. Advances in soft sensors for wastewater treatment plants: A systematic review. J. Water Process Eng. 44, 102367. https://doi.org/https://doi.org/10.1016/j.jwpe.2021.102367

Corominas, L., Flores-Alsina, X., Snip, L., Vanrolleghem, P.A., 2012. Comparison of different modeling approaches to better evaluate greenhouse gas emissions from whole wastewater treatment plants. Biotechnol. Bioeng. 109, 2854–2863. https://doi.org/10.1002/bit.24544

Corominas, L., Garrido-Baserba, M., Villez, K., Olsson, G., Cortés, U., Poch, M., 2018. Transforming data into knowledge for improved wastewater treatment operation: A critical review of techniques. Environ. Model. Softw. 106, 89–103. https://doi.org/10.1016/j.envsoft.2017.11.023

Corominas, L., Sin, G., Puig, S., Balaguer, M.D., Vanrolleghem, P.A., Colprim, J., 2011. Modified calibration protocol evaluated in a model-based testing of SBR flexibility. Bioprocess Biosyst. Eng. 34, 205–214. https://doi.org/10.1007/s00449-010-0462-2

Daelman, M.R.J., van Voorthuizen, E.M., van Dongen, L.G.J.M., Volcke, E.I.P., van

Loosdrecht, M.C.M., 2013. Methane and nitrous oxide emissions from municipal wastewater treatment – results from a long-term study. Water Sci. Technol. 67, 2350–2355. https://doi.org/10.2166/wst.2013.109

Daelman, M.R.J., van Voorthuizen, E.M., van Dongen, U.G.J.M., Volcke, E.I.P., van Loosdrecht, M.C.M., 2015. Seasonal and diurnal variability of N_2O emissions from a full-scale municipal wastewater treatment plant. Sci. Total Environ. 536, 1–11. https://doi.org/10.1016/j.scitotenv.2015.06.122

Daigger, G.T., 2014. Oxygen and carbon requirements for biological nitrogen removal processes accomplishing. Water Environ. Res. 86, 3–8. https://doi.org/10.2175/106143013X13807328849459

Daims, H., Lebedeva, E. V., Pjevac, P., Han, P., Herbold, C., Albertsen, M., Jehmlich, N., Palatinszky, M., Vierheilig, J., Bulaev, A., Kirkegaard, R.H., Von Bergen, M., Rattei, T., Bendinger, B., Nielsen, P.H., Wagner, M., 2015. Complete nitrification by Nitrospira bacteria. Nature 528, 504–509. https://doi.org/10.1038/nature16461

Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans. Evol. Comput. 6, 182–197. https://doi.org/10.1109/4235.996017

Deepnarain, N., Nasr, M., Kumari, S., Stenström, T.A., Reddy, P., Pillay, K., Bux, F., 2019. Decision tree for identification and prediction of filamentous bulking at full-scale activated sludge wastewater treatment plant. Process Saf. Environ. Prot. 126, 25–34. https://doi.org/10.1016/j.psep.2019.02.023

Dietterich, T.G., 2000. Ensemble Methods in Machine Learning BT - Multiple Classifier

Systems. Springer, Heidelberg, Germany, pp. 1–15.

Ding, C., Peng, H., 2003. Minimum redundancy feature selection from microarray gene expression data. Proc. 2003 IEEE Bioinforma. Conf. CSB 2003 523–528. https://doi.org/10.1109/CSB.2003.1227396

Dochain, D., Vanrolleghem, P.A, 2005. Dynamical Modelling & Estimation in Wastewater Treatment Processes. https://doi.org/10.2166/9781780403045

Domingo-Félez, C., Smets, B.F., 2016. A consilience model to describe N_2O production during biological N removal. Environ. Sci. Water Res. Technol. 2, 923–930. https://doi.org/10.1039/c6ew00179c

Doshi-Velez, F., Kim, B., 2017. Towards a rigorous science of interpretable machine learning. arXiv 1–13.

Duan, H., van den Akker, B., Thwaites, B.J., Peng, L., Herman, C., Pan, Y., Ni, B.J., Watt, S., Yuan, Z., Ye, L., 2020. Mitigating nitrous oxide emissions at a full-scale wastewater treatment plant. Water Res. 185, 116196. https://doi.org/10.1016/j.watres.2020.116196

Duan, H., Zhao, Y., Koch, K., Wells, G.F., Zheng, M., Yuan, Z., Ye, L., 2021. Insights into nitrous oxide mitigation strategies in wastewater treatment and challenges for wider implementation. Environ. Sci. Technol. 55, 7208–7224. https://doi.org/10.1021/acs.est.1c00840

Fiat, J., Filali, A., Fayolle, Y., Bernier, J., Rocher, V., Spérandio, M., Gillot, S., 2019. Considering the plug-flow behavior of the gas phase in nitrifying BAF models

significantly improves the prediction of N_2O emissions. Water Res. 156, 337–346. https://doi.org/10.1016/j.watres.2019.03.047

Freund, Y., Schapire, R.E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. Syst. Sci. 55, 119–139. https://doi.org/https://doi.org/10.1006/jcss.1997.1504

Fux, C., Velten, S., Carozzi, V., Solley, D., Keller, J., 2006. Efficient and stable nitritation and denitritation of ammonium-rich sludge dewatering liquor using an SBR with continuous loading. Water Res. 40, 2765–2775. https://doi.org/https://doi.org/10.1016/j.watres.2006.05.003

Géron, A., 2019. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. O'Reilly Media: Sebastopol, CA.

Gillot, S., Ohtsuki, T., Rieger, L., Shaw, A., Takacs, I., Winkler, S., 2009. Development of a unified protocol for good modeling practice in activated sludge modeling. Influents 4, 70–72.

Gruber, W., Niederdorfer, R., Ringwald, J., Morgenroth, E., Bürgmann, H., Joss, A., 2021. Linking seasonal N_2O emissions and nitrification failures to microbial dynamics in a SBR wastewater treatment plant. Water Res. X 11, 100098. https://doi.org/10.1016/j.wroa.2021.100098

Gujer, W., 2008. Systems Analysis for Water Technology, Systems Analysis for Water Technology. Springer, Berlin Heidelberg. https://doi.org/10.1007/978-3-540-77278-1

Guo, L., Vanrolleghem, P.A., 2014. Calibration and validation of an activated sludge model for greenhouse gases no. 1 (ASMG1): Prediction of temperature-dependent N_2O emission dynamics. Bioprocess Biosyst. Eng. 37, 151–163. https://doi.org/10.1007/s00449-013-0978-3

Harper, W.F., Takeuchi, Y., Riya, S., Hosomi, M., Terada, A., 2015. Novel abiotic reactions increase nitrous oxide production during partial nitrification: Modeling and experiments. Chem. Eng. J. 281, 1017–1023. https://doi.org/10.1016/j.cej.2015.06.109

Henze, M., Grady, C.P.L., Gujer, W., Marais, G.V.R., Matsuo, T., 1987. Activated Sludge Model No. 1, Scientific and Technical Report No. 1.

Henze, M., Gujer, W., Mino, T., van Loosdrecht, M.C.M., 2000. Activated sludge models ASM1, ASM2, ASM2d and ASM3. IWA Publishing, London, UK. https://doi.org/10.1007/s13398-014-0173-7.2

Hiatt, W.C., Grady, C.P.L., 2008. An updated process model for carbon oxidation, nitrification, and denitrification. Water Environ. Res. 80, 2145–2156. https://doi.org/10.2175/106143008x304776

Höhne, N., Elzen, M. den, Rogelj, J., Metz, B., Fransen, T., Kuramochi, T., Olhoff, A., Alcamo, J., Winkler, H., Fu, S., Schaeffer, M., Schaeffer, R., Peters, G.P., Maxwell, S., Dubash, N.K., 2020. Emissions: world has four times the work or one-third of the time. Nature 579, 25–28.

Holland, J.H., 1992. Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence. The MIT

Press. https://doi.org/10.7551/mitpress/1090.001.0001

Hosseinzadeh, A., Zhou, J.L., Altaee, A., Li, D., 2022. Machine learning modeling and analysis of biohydrogen production from wastewater by dark fermentation process. Bioresour. Technol. 343, 126111. https://doi.org/https://doi.org/10.1016/j.biortech.2021.126111

Hubert, M., Vandervieren, E., 2008. An adjusted boxplot for skewed distributions. Comput. Stat. Data Anal. 52, 5186–5201. https://doi.org/https://doi.org/10.1016/j.csda.2007.11.008

Hug, T., Benedetti, L., Hall, E.R., Johnson, B.R., Morgenroth, E., Nopens, I., Rieger, L., Shaw, A., Vanrolleghem, P.A., 2009. Wastewater treatment models in teaching and training: The mismatch between education and requirements for jobs. Water Sci. Technol. 59, 745–753. https://doi.org/10.2166/wst.2009.595

Hvala, N., Kocijan, J., 2021. Input variable selection using machine learning and global sensitivity methods for the control of sludge bulking in a wastewater treatment plant. Comput. Chem. Eng. 154. https://doi.org/10.1016/j.compchemeng.2021.107493

Hwangbo, S., Al, R., Chen, X., Sin, G., 2021. Integrated model for understanding N_2O emissions from wastewater treatment plants: A deep learning approach. Environ. Sci. Technol. 55, 2143–2151. https://doi.org/10.1021/acs.est.0c05231

Hwangbo, S., Al, R., Sin, G., 2020. An integrated framework for plant data-driven process modeling using deep-learning with Monte-Carlo simulations. Comput. Chem. Eng. 143. https://doi.org/10.1016/j.compchemeng.2020.107071

Kampschreur, M J, Poldermans, R., Kleerebezem, R., van der Star, W.R.L., Haarhuis, R., Abma, W.R., Jetten, M.S.M., van Loosdrecht, M.C.M., 2009. Emission of nitrous oxide and nitric oxide from a full-scale single-stage nitritation-anammox reactor. Water Sci. Technol. 60, 3211–3217. https://doi.org/10.2166/wst.2009.608

Kampschreur, Marlies J., Temmink, H., Kleerebezem, R., Jetten, M.S.M., van Loosdrecht, M.C.M., 2009. Nitrous oxide emission during wastewater treatment. Water Res. 43, 4093–4103. https://doi.org/10.1016/j.watres.2009.03.001

Kirim, G., McCullough, K., Bressani-Ribeiro, T., Domingo-Félez, C., Duan, H., Al-Omari, A., De Clippeleir, H., Jimenez, J., Klaus, S., Ladipo-Obasa, M., Mehrani, M.J., Regmi, P., Torfs, E., Volcke, E.I.P., Vanrolleghem, P.A., 2022. Mainstream short-cut N removal modelling: current status and perspectives. Water Sci. Technol. 85, 2539–2564. https://doi.org/10.2166/wst.2022.131

Lackner, S., Terada, A., Smets, B.F., 2008. Heterotrophic activity compromises autotrophic nitrogen removal in membrane-aerated biofilms: Results of a modeling study. Water Res. 42, 1102–1112. https://doi.org/10.1016/j.watres.2007.08.025

Lang, L., Piveteau, S., Azimi, S., Rocher, V., Spérandio, M., 2019. Modelling N_2O emission from PNA process under oxygen limitation: model calibration and prospects, in: Proceeding of Watermatex 2019 Conference. pp. 1–4.

Lang, L., Pocquet, M., Ni, B.J., Yuan, Z., Spérandio, M., 2017. Comparison of different two-pathway models for describing the combined effect of DO and nitrite on the nitrous oxide production by ammonia-oxidizing bacteria. Water Sci. Technol. 75, 491–500. https://doi.org/10.2166/wst.2016.389

Law, Y., Ni, B.J., Lant, P., Yuan, Z., 2012. N_2O production rate of an enriched ammonia-oxidising bacteria culture exponentially correlates to its ammonia oxidation rate. Water Res. 46, 3409–3419. https://doi.org/10.1016/j.watres.2012.03.043

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539

Li, F., Vanrolleghem, P.A., 2022. An influent generator for WRRF design and operation based on a recurrent neural network with multi-objective optimization using a genetic algorithm. Water Sci. Technol. 85, 1444–1453. https://doi.org/10.2166/wst.2022.048

Lipton, Z.C., 2018. The mythos of model interpretability. Commun. ACM 61, 35–43. https://doi.org/10.1145/3233231

Liu, Yingrui, Zhu, T., Ren, S., Zhao, T., Chai, H., Xu, Y., Peng, L., Liu, Yiwen, 2022. Contribution of nitrification and denitrification to nitrous oxide turnovers in membrane-aerated biofilm reactors (MABR): A model-based evaluation. Sci. Total Environ. 806, 1–9. https://doi.org/10.1016/j.scitotenv.2021.151321

Louppe, G., 2014. Understanding random forests: From theory to practice. arXiv Prepr.

Ly, Q.V., Truong, V.H., Ji, B., Nguyen, X.C., Cho, K.H., Ngo, H.H., Zhang, Z., 2022. Exploring potential machine learning application based on big data for prediction of wastewater quality from different full-scale wastewater treatment plants. Sci. Total Environ. 832, 154930. https://doi.org/https://doi.org/10.1016/j.scitotenv.2022.154930

Maclin, R., Opitz, D.W., 2011. Popular ensemble methods: An empirical study. arXiv abs/1106.0.

Mampaey, K.E., Beuckels, B., Kampschreur, M.J., Kleerebezem, R., Van Loosdrecht, M.C.M., Volcke, E.I.P., 2013. Modelling nitrous and nitric oxide emissions by autotrophic ammonia-oxidizing bacteria. Environ. Technol. 34, 1555–1566. https://doi.org/10.1080/09593330.2012.758666

Mannina, G., Cosenza, A., Vanrolleghem, P.A., Viviani, G., 2011. A practical protocol for calibration of nutrient removal wastewater treatment models. J. Hydroinformatics 13, 575–595. https://doi.org/10.2166/hydro.2011.041

Mitchell, T., 1997. Machine Learning. McGraw-Hill, Blacklick, Ohio, USA. https://doi.org/10.1002/(sici)1099-1689(199909)9:3<191::aid-stvr184>3.0.co;2-e

Molnar, C., 2022. Interpretable Machine Learning: A Guide for Making Black Box Models Explainable, 2nd ed.

Montavon, G., Lapuschkin, S., Binder, A., Samek, W., Müller, K.R., 2017. Explaining nonlinear classification decisions with deep Taylor decomposition. Pattern Recognit. 65, 211–222. https://doi.org/10.1016/j.patcog.2016.11.008

Morgenroth, E., 2008. Modelling Biofilms, in: Henze, M., Loosdrecht, M.C.M. van, Ekama, G.A., Brdjanovic, D. (Eds.), Biological Wastewater Treatment: Principles, Modelling and Design. IWA Publishing, London, UK, pp. 457–492.

Ni, B.-J., Yu, H.-Q., 2010. Mathematical modeling of aerobic granular sludge: A review. Biotechnol. Adv. 28, 895–909.

https://doi.org/https://doi.org/10.1016/j.biotechadv.2010.08.004

Ni, B.J., Peng, L., Law, Y., Guo, J., Yuan, Z., 2014. Modeling of nitrous oxide production by autotrophic ammonia-oxidizing bacteria with multiple production pathways. Environ. Sci. Technol. 48, 3916–3924. https://doi.org/10.1021/es405592h

Ni, B.J., Ruscalleda, M., Pellicer-Nàcher, C., Smets, B.F., 2011. Modeling nitrous oxide production during biological nitrogen removal via nitrification and denitrification: Extensions to the general ASM models. Environ. Sci. Technol. 45, 7768–7776. https://doi.org/10.1021/es201489n

Ni, B.J., Ye, L., Law, Y., Byers, C., Yuan, Z., 2013a. Mathematical modeling of nitrous oxide (N_2O) emissions from full-scale wastewater treatment plants. Environ. Sci. Technol. 47, 7795–7803. https://doi.org/10.1021/es4005398

Ni, B.J., Yuan, Z., Chandran, K., Vanrolleghem, P.A., Murthy, S., 2013b. Evaluating four mathematical models for nitrous oxide production by autotrophic ammonia-oxidizing bacteria. Biotechnol. Bioeng 110, 153–163. https://doi.org/10.1002/bit.24620/abstract

Nopens, I., Porro, J., Ye, L., 2022. Perspectives on fugitive GHGs reduction from urban wastewater systems, in: Ye, L., Porro, J., Nopens, I. (Eds.), Quantification and Modelling of Fugitive Greenhouse Gas Emissions from Urban Water Systems. London, UK, pp. 245–257.

Pan, Y., Ni, B.J., Bond, P.L., Ye, L., Yuan, Z., 2013a. Electron competition among nitrogen oxides reduction during methanol-utilizing denitrification in wastewater treatment. Water Res. 47, 3273–3281. https://doi.org/10.1016/j.watres.2013.02.054

Pan, Y., Ni, B.J., Yuan, Z., 2013b. Modeling electron competition among nitrogen oxides reduction and N 2O accumulation in denitrification. Environ. Sci. Technol. 47, 11083–11091. https://doi.org/10.1021/es402348n

Park, J., Lee, W.H., Kim, K.T., Park, C.Y., Lee, S., Heo, T.-Y., 2022. Interpretation of ensemble learning to predict water quality using explainable artificial intelligence. Sci. Total Environ. 832, 155070. https://doi.org/https://doi.org/10.1016/j.scitotenv.2022.155070

Peng, L., Ni, B.J., Erler, D., Ye, L., Yuan, Z., 2014. The effect of dissolved oxygen on N_2O production by ammonia-oxidizing bacteria in an enriched nitrifying sludge. Water Res. 66, 12–21. https://doi.org/10.1016/j.watres.2014.08.009

Peng, L., Ni, B.J., Law, Y., Yuan, Z., 2015. Modeling of N_2O production by ammonia oxidizing bacteria: integration of catabolism and anabolism, in: The 9th IWA Symposium on Systems Analysis and Integrated Assessment (Watermatex 2015). Gold Coast, Australia.

Philips, S., Laanbroek, H.J., Verstraete, W., 2002. Origin, causes and effects of increased nitrite concentrations in aquatic environments. Rev. Environ. Sci. Biotechnol. 1, 115–141. https://doi.org/10.1023/A:1020892826575

Pijuan, M., Zhao, Y., 2022. Full-scale source, mechanisms and factors affecting nitrous oxide emissions, in: Ye, L., Porro, J., Nopens, I. (Eds.), Quantification and Modelling of Fugitive Greenhouse Gas Emissions from Urban Water Systems. IWA Publishing.

Pocquet, M., Wu, Z., Queinnec, I., Spérandio, M., 2016. A two pathway model for N_2O

emissions by ammonium oxidizing bacteria supported by the NO/N_2O variation. Water Res. 88, 948–959. https://doi.org/10.1016/j.watres.2015.11.029

Porro, J., Vasilaki, V., Bellandi, G., Katsou, E., 2022. Knowledge-based and data-driven approaches for assessing greenhouse gas emissions from wastewater systems, in: Ye, L., Porro, J., Nopens, I. (Eds.), Quantification and Modelling of Fugitive Greenhouse Gas Emissions from Urban Water Systems. UK, London, pp. 229–244.

Poughon, L., Dussap, C.-G., Gros, J.-B., 2001. Energy model and metabolic flux analysis for autotrophic nitrifiers. Biotechnol. Bioeng. 72, 416–433. https://doi.org/https://doi.org/10.1002/1097-0290(20000220)72:4<416::AID-BIT1004>3.0.CO;2-D

Quinlan, J.R., 1986. Induction of decision trees. Mach. Learn. 1, 81–106.

Ravishankara, A.R., Daniel, J.S., Portmann, R.W., 2009. Nitrous oxide (N_2O): The dominant ozone-depleting substance emitted in the 21st century. Science (80-.). 326, 123–125. https://doi.org/10.1126/science.1176985

Reichert, P., 1998. Computer Program for the Identi cation and Simulation of Aquatic Systems, Science And Technology.

Reichert, P., Ruchti, J., Simon, W., 1998. AQUASIM 2.0.

Reichert, P., Wanner, O., 1997. Movement of solids in biofilms: Significance of liquid phase transport. Water Sci. Technol. 36, 321–328.

Ribeiro, M.T., Singh, S., Guestrin, C., 2016. "Why Should I Trust You?" Explaining the Predictions of Any Classifier. NAACL-HLT 2016 - 2016 Conf. North Am. Chapter

Assoc. Comput. Linguist. Hum. Lang. Technol. Proc. Demonstr. Sess. 97–101. https://doi.org/10.18653/v1/n16-3020

Richardson, D., Felgate, H., Watmough, N., Thomson, A., Baggs, E., 2009. Mitigating release of the potent greenhouse gas N_2O from the nitrogen cycle - could enzymic regulation hold the key? Trends Biotechnol. 27, 388–397. https://doi.org/10.1016/j.tibtech.2009.03.009

Rogelj, J., Shindell, D., Jiang, K., Fifita, S., Forster, P., Ginzburg, V., Handa, C., Kheshgi, H., Kobayashi, S., Kriegler, Mundaca L., E., R., S., V., V.M., 2018. Mitigation Pathways Compatible with 1.5°C in the Context of Sustainable Development. An IPCC Spec. Rep. impacts Glob. Warm. 1.5°C above pre-industrial levels Relat. Glob. Greenh. gas Emiss. pathways, Context Strength. Glob. response to Threat Clim. Chang. Sustain. Dev. 93–174. https://doi.org/10.1017/9781009157940.004

Rokach, L., Maimon, O.Z., 2008. Data Mining with Decision Trees: Theory and Applications, Series in machine perception and artificial intelligence. World Scientific.

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. Nature 323, 533–536. https://doi.org/10.1038/323533a0

Sabba, F., Picioreanu, C., Pérez, J., Nerenberg, R., 2015. Hydroxylamine diffusion can enhance N_2O emissions in nitrifying biofilms: A modeling study. Environ. Sci. Technol. 49, 1486–1494. https://doi.org/10.1021/es5046919

Sabba, F., Redmond, E., Ruff, C., Ramirez, M., Young, M., Song, M.J., Yoon, S.,

Downing, L., 2024. Exploring community and kinetic shifts in nitrifying microbial communities in low dissolved oxygen activated sludge facilities for energy-efficient biological nitrogen removal. https://doi.org/10.1021/acsestwater.3c00715

Sappl, J., Harders, M., Rauch, W., 2023. Machine learning for quantile regression of biogas production rates in anaerobic digesters. Sci. Total Environ. 872. https://doi.org/10.1016/j.scitotenv.2023.161923

Schneider, M.Y., Quaghebeur, W., Borzooei, S., Froemelt, A., Li, F., Saagi, R., Wade, M.J., Zhu, J.J., Torfs, E., 2022. Hybrid modelling of water resource recovery facilities: status and opportunities. Water Sci. Technol. 85, 2503–2524. https://doi.org/10.2166/wst.2022.115

Schreiber, F., Loeffler, B., Polerecky, L., Kuypers, M.M.M., de Beer, D., 2009. Mechanisms of transient nitric oxide and nitrous oxide production in a complex biofilm. ISME J. 3, 1301–1313. https://doi.org/10.1038/ismej.2009.55

Seshan, S., Poinapen, J., Zandvoort, M.H., Lier, J.B. Van, 2024. Limitations of a biokinetic model to predict the seasonal variations of nitrous oxide emissions from a full-scale wastewater treatment plant. Sci. Total Environ. 917, 170370. https://doi.org/10.1016/j.scitotenv.2024.170370

Shao, Y., Florentino, A.P., Buchanan, I., Mohammed, A., Liu, Y., 2019. Microbial population dynamics in a partial nitrification reactor treating high ammonia strength supernatant from anaerobically digested sludge: Role of the feed water characteristics. Int. Biodeterior. Biodegrad. 137, 109–117. https://doi.org/10.1016/j.ibiod.2018.12.006

Sin, G., Kaelin, D., Kampschreur, M.J., Takács, I., Wett, B., Gernaey, K. V., Rieger, L., Siegrist, H., Van Loosdrecht, M.C.M., 2008. Modelling nitrite in wastewater treatment systems: A discussion of different modelling concepts. Water Sci. Technol. 58, 1155–1171. https://doi.org/10.2166/wst.2008.485

Smith, V.H., 2003. Eutrophication of freshwater and coastal marine ecosystems: A global problem. Environ. Sci. Pollut. Res. 10, 126–139. https://doi.org/10.1065/espr2002.12.142

Solís, B., Guisasola, A., Pijuan, M., Corominas, L., Baeza, J.A., 2022. Systematic calibration of N_2O emissions from a full-scale WWTP including a tracer test and a global sensitivity approach. Chem. Eng. J. 435, 134733. https://doi.org/10.1016/j.cej.2022.134733

Song, M.J., Choi, S., Bae, W. Bin, Lee, J., Han, H., Kim, D.D., Kwon, M., Myung, J., Kim, Y.M., Yoon, S., 2020. Identification of primary effecters of N_2O emissions from full-scale biological nitrogen removal systems using random forest approach. Water Res. 184, 116144. https://doi.org/10.1016/j.watres.2020.116144

Spérandio, M., Lang, L., Sabba, F., Nerenberg, R., Vanrolleghem, P., Domingo-Félez, C., Smets, B.F., Duan, H., Ni, B.-J., Yuan, Z., 2022. Modelling N_2O production and emissions, in: Ye, L., Porro, J., Nopens, I. (Eds.), Quantification and Modelling of Fugitive Greenhouse Gas Emissions from Urban Water Systems. London, UK, pp. 167–196.

Stein, L.Y., 2011. Surveying N_2O-producing pathways in bacteria. Methods Enzymol. 486, 131–152. https://doi.org/10.1016/B978-0-12-381294-0.00006-7

Su, Q., Domingo-Félez, C., Jensen, M.M., Smets, B.F., 2019. Abiotic nitrous oxide (N_2O) production is strongly pH dependent, but contributes little to overall N_2O emissions in biological nitrogen removal systems. Environ. Sci. Technol. 53, 3508–3516. https://doi.org/10.1021/acs.est.8b06193

Szeląg, B., Zaborowska, E., Mąkinia, J., 2023. An algorithm for selecting a machine learning method for predicting nitrous oxide emissions in municipal wastewater treatment plants. J. Water Process Eng. 54. https://doi.org/10.1016/j.jwpe.2023.103939

Tallec, G., Garnier, J., Billen, G., Gousailles, M., 2008. Nitrous oxide emissions from denitrifying activated sludge of urban wastewater treatment plants, under anoxia and low oxygenation. Bioresour. Technol. 99, 2200–2209. https://doi.org/10.1016/j.biortech.2007.05.025

Torfs, E., Nicolaï, N., Daneshgar, S., Copp, J.B., Haimi, H., Ikumi, D., Johnson, B., Plosz, B.B., Snowling, S., Townley, L.R., Valverde-Pérez, B., Vanrolleghem, P.A., Vezzaro, L., Nopens, I., 2022. The transition of WRRF models to digital twin applications. Water Sci. Technol. 85, 2840–2853. https://doi.org/10.2166/wst.2022.107

van Loosdrecht, M.C.M., Jetten, M.S.M., 1998. Microbiological conversions in nitrogen removal. Water Sci. Technol. 38, 1–7.

Vannecke, T.P.W., Volcke, E.I.P., 2015. Modelling microbial competition in nitrifying biofilm reactors. Biotechnol. Bioeng. 112, 2550–2561. https://doi.org/10.1002/bit.25680

Vasilaki, V., Conca, V., Frison, N., Eusebi, A.L., Fatone, F., Katsou, E., 2020a. A knowledge discovery framework to predict the N_2O emissions in the wastewater sector. Water Res. 178. https://doi.org/10.1016/j.watres.2020.115799

Vasilaki, V., Danishvar, S., Mousavi, A., Katsou, E., 2020b. Data-driven versus conventional N_2O EF quantification methods in wastewater; how can we quantify reliable annual EFs? Comput. Chem. Eng. 141, 106997. https://doi.org/10.1016/j.compchemeng.2020.106997

Vasilaki, V., Massara, T.M., Stanchev, P., Fatone, F., Katsou, E., 2019. A decade of nitrous oxide (N_2O) monitoring in full-scale wastewater treatment processes: A critical review. Water Res. 161, 392–412. https://doi.org/10.1016/j.watres.2019.04.022

Vasilaki, V., Volcke, E.I.P., Nandi, A.K., van Loosdrecht, M.C.M., Katsou, E., 2018. Relating N_2O emissions during biological nitrogen removal with operating conditions using multivariate statistical techniques. Water Res. 140, 387–402. https://doi.org/10.1016/j.watres.2018.04.052

Venard, J.K., Street, R.L., 1975. Elementary Fluid Mechanics, 5th ed. ed. Wiley, New York.

Wan, X., Volcke, E.I.P., 2022. Dynamic modelling of N_2O emissions from a full-scale granular sludge partial nitritation-anammox reactor. Biotechnol. Bioeng. 119, 1426–1438. https://doi.org/10.1002/bit.28054

Wang, X., Wang, K., Ding, J., Chen, X., Li, Y., Zhang, W., 2021. Predicting water quality during urbanization based on a causality-based input variable selection

method modified back-propagation neural network. Environ. Sci. Pollut. Res. Int. 28, 960–973. https://doi.org/10.1007/s11356-020-10514-8

Wanner, O., Gujer, W., 1986. A multispecies biofilm model. Biotechnol. Bioeng. 28, 314–328.

Wanner, O., Gujer, W., 1984. Competition in biofilms. Water Sci. Technol. 17, 27–44. https://doi.org/10.2166/wst.1994.0769

Wanner, O., Reichert, P., 1996. Mathematical modeling of mixed-culture biofilms. Biotechnol. Bioeng. 49, 172–184.

Wunderlin, P., Lehmann, M.F., Siegrist, H., Tuzson, B., Joss, A., Emmenegger, L., Mohn, J., 2013. Isotope signatures of N_2O in a mixed microbial population system: constraints on N_2O producing pathways in wastewater treatment. Environ. Sci. Technol. 47, 1339–1348. https://doi.org/10.1021/es303174x

Xiao, Q., Chang, H.H., Geng, G., Liu, Y., 2018. An ensemble machine-learning model to predict historical PM2.5 concentrations in China from satellite data. Environ. Sci. Technol. 52, 13260–13269. https://doi.org/10.1021/acs.est.8b02917

Xu, Y., Wang, Z., Nairat, S., Zhou, J., He, Z., 2023. Artificial intelligence-assisted prediction of effluent phosphorus in a full-scale wastewater treatment plant with missing phosphorus input and removal data. ACS ES&T Water. https://doi.org/10.1021/acsestwater.2c00517

Yang, Q., Liu, X., Peng, C., Wang, S., Sun, H., Peng, Y., 2009. N_2O production during nitrogen removal via nitrite from domestic wastewater: Main sources and control

method. Environ. Sci. Technol. 43, 9400–9406. https://doi.org/10.1021/es9019113

Yang, S., Xu, S., Zhou, Y., Mohammed, A., Ashbolt, N.J., Liu, Y., 2020. The importance of integrated fixed film activated sludge reactor and intermittent aeration in nitritation-anammox systems: Understanding reactor optimization for lagoon supernatant treatment. Int. Biodeterior. Biodegrad. 149, 104938. https://doi.org/10.1016/j.ibiod.2020.104938

Yu, R., Kampschreur, M.J., Loosdrecht, M.C.M. van, Chandran, K., 2010. Mechanisms and specific directionality of autotrophic nitrous oxide and nitric oxide generation during transient anoxia. Environ. Sci. Technol. 44, 1313–1319. https://doi.org/10.1021/es902794a

Zaborowska, E., Lu, X., Makinia, J., 2019. Strategies for mitigating nitrous oxide production and decreasing the carbon footprint of a full-scale combined nitrogen and phosphorus removal activated sludge system. Water Res. 162, 53–63. https://doi.org/10.1016/j.watres.2019.06.057

Zaghloul, M.S., Achari, G., 2022. Application of machine learning techniques to model a full-scale wastewater treatment plant with biological nutrient removal. J. Environ. Chem. Eng. 10, 107430. https://doi.org/10.1016/j.jece.2022.107430

Zhang, C., Zhang, Y., Shi, X., Almpanidis, G., Fan, G., Shen, X., 2019. On incremental learning for gradient boosting decision trees. Neural Process. Lett. 50, 957–987. https://doi.org/10.1007/s11063-019-09999-3

Zhu, J., Bernier, J., Patry, B., Azimi, S., Pauss, A., Rocher, V., Vanrolleghem, P.., 2019. Comprehensive modelling of full-scale nitrifying and post-denitrifying biofilters, in:

Proceedings WEF Nutrient Removal and Recovery Symposium 2019 – 21st Century Vision. Minneapolis, MN, USA.

Zhu, J.J., Borzooei, S., Sun, J., Ren, Z.J., 2022. Deep learning optimization for soft sensing of hard-to-measure wastewater key variables. ACS ES T Eng. 2, 1341–1355. https://doi.org/10.1021/acsestengg.1c00469

Zhu, J.J., Yang, M., Ren, Z.J., 2023. Machine learning in environmental research: Common pitfalls and best practices. Environ. Sci. Technol. https://doi.org/10.1021/acs.est.3c00026

Zou, X., Zhou, Y., Gao, M., Yang, S., Mohammed, A., Liu, Y., 2022. Effective N_2O emission control during the nitritation/denitritation treatment of ammonia rich wastewater. J. Environ. Chem. Eng. 10, 107234. https://doi.org/10.1016/j.jece.2022.107234

Zou, X., Zhou, Y., Guo, B., Shao, Y., Yang, S., Mohammed, A., Liu, Y., 2020. Single reactor nitritation-denitritation for high strength digested biosolid thickening lagoon supernatant treatment. Biochem. Eng. J. 160, 1–8. https://doi.org/10.1016/j.bej.2020.107630

Zumft, W.G., 1997. Cell biology and molecular basis of denitrification. Microbiol. Mol. Biol. Rev. 61, 533–616. https://doi.org/10.1128/mmbr.61.4.533-616.1997