FACULTY OF
BIOSCIENCE ENGINEERING

HYBRID MODELLING OF WASTEWATER TREATMENT PROCESSES

Loes Verhaeghe
Student ID: 01804432

Promotors: Dr. ir. Saba Daneshgar, Prof. dr. ir. Peter Vanrolleghem
Tutor: Prof. dr. ir. Elena Torfs

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of master in Bioscience Engineering: land, water and climate management .
Academic year: 2022 - 2023

GHENT
UNIVERSITY

Gent, June 9, 2023

The promotors,	The tutor,	The author,
Dr. Ir. Saba Daneshgar & Prof. dr. ir. Peter Vanrolleghem	Prof. dr. ir. Elena Torfs	Loes Verhaeghe

ACKNOWLEDGMENTS

Above all, I want to extend a heartfelt thank you to my tutor, Prof. dr. ir. Elena Torfs. Her invaluable guidance, insightful feedback, and unwavering support have played a crucial role in writing my master's thesis. She has been a constant source of motivation, especially during the more challenging moments. I am truly grateful for the opportunity she provided me to explore my own ideas, which has sparked a passion for further research.

I would also like to express my heartfelt thanks to promotor, Prof. dr. ir. Peter Vanrolleghem, for warmly welcoming me to Québec and to the modelEAU research group. His guidance has been immensely valuable in shaping the direction of my research. Additionally, I would like to express my gratitude to my promotor, Dr. ir. Saba Daneshgar, for his guidance in Python programming and his insightful feedback on my thesis. His expertise and advice have been invaluable in enhancing the quality of my work.

Several individuals who have contributed to the success of this research deserve a special mention. Firstly, I would like to thank Marcello for our discussions on hybrid models, the inspiration I gained from his research, and his overall support. I am also grateful to Gamze for taking the time to explain her model to me. Jean-David, thank you for your assistance in the machine learning aspects of my thesis, and Thomas, thank you for your help in with the mechanistic modelling. A special acknowledgement is reserved for the members of the modelEAU research group for their valuable feedback on my research, the enjoyable activities we shared and the much-needed coffee breaks. I am especially grateful to the pilEAUte team, Sanaz, Karen, Nathalia, Laleh, Camille, and Peter, whose work on the pilEAUte made this research possible.

Lastly, I would like to thank my friends and family for the support during my studies and during the writing of my thesis. A special mention goes to Elisabeth for the incredible memories we have created together and for all the ways she has supported me, decreasing my stress level on countless occasions.

CONTENTS

SAMENVATTING

Een hybride model (HM) combineert een mechanistisch model gebaseerd op fysische proceskennis met een data-gedreven component die leert uit online sensor data. In deze thesis wordt een parallel HM van een piloot-afvalwaterzuiveringsinstallatie ontwikkeld om een model te verkrijgen met een betere voorspellende kracht. Het HM werd opgesteld door een neuraal netwerk (NN) te trainen op de fout tussen het mechanistische model en metingen van het effluent van nitraat en zwevende stoffen. Het onderzoek toont aan dat het parallelle HM significant betere voorspellingen biedt in vergelijking met het gekalibreerde mechanistische model. Zowel Long-Short Term Memory (LSTM) als Convolutional Neural Networks (CNN's) worden getest, waarbij CNN betere resultaten levert. De thesis onderzoekt ook de balans tussen de kalibratie-inspanning voor het mechanistische model en de compensatie door de NN-component. Hiervoor worden verschillende versies van het parallelle HM gebouwd waarbij gebruik gemaakt wordt van een kalibratiedataset, validatiedataset of een ongekalibreerd mechanistisch model. De beste prestatie wordt bekomen voor het parallelle HM op basis van het ongekalibreerde mechanistisch model. Een parallel HM kan dus de kalibratie-inspanning van mechanistische modellen kan verminderen, vooral voor sterk onzekere of niet-identificeerbare parameters. Een serie HM integreert een data-gedreven model dat een minder goed-beschreven subproces modelleert in het mechanistische model. In deze thesis wordt een serie HM ontwikkeld voor het voorspellen van de zuurstof-overdrachtscoëfficiënt (K_La) in een beluchtingsbekken. Dit seriemodel heeft potentieel om in de toekomst real-time K_La te voorspellen in beluchtingsbekkens. Hoewel het seriemodel redelijke prestaties vertoont, heeft het moeite met extrapolatie naar nieuwe situaties. Verder onderzoek is nodig om de K_La-voorspelling in serie HM's te verbeteren.

SUMMARY

A hybrid model (HM) combines a mechanistic model based on physical process knowledge with a data-driven component that learns from online sensor data. This thesis develops a parallel HM of a pilot wastewater treatment plant to obtain a model with improved predictive power. The HM was constructed by training a neural network (NN) on the residual between the mechanistic model and measurements of effluent nitrate and total suspended solids. The research demonstrates that the parallel HM provides significantly better predictions compared to the calibrated mechanistic model. Both Long-Short Term Memory (LSTM) and Convolutional Neural Networks (CNNs) are tested, with CNNs yielding better results. The thesis also investigates the balance between the calibration effort for the mechanistic model and the compensation by the NN component. For this purpose, different versions of the parallel HM are built using a calibration dataset, validation dataset, or an uncalibrated mechanistic model. The best performance is achieved for the parallel HM based on the uncalibrated mechanistic model. A parallel HM can thus reduce the calibration effort of mechanistic models, particularly for highly uncertain or non-identifiable parameters. A serial HM integrates a data-driven model that models a less well-described subprocess within the mechanistic model. In this thesis, a series HM is developed to predict the oxygen transfer coefficient (K_{La}) in an aeration tank. This serial model has the potential to predict real-time K_{La} in aeration tanks in the future. Although the serial HM shows reasonable performance, it struggles with extrapolation to new situations. Further research is needed to improve K_{La} prediction in series HM models.

1. INTRODUCTION

Clean water supply has become increasingly important in recent decades, as water scarcity worsens due to factors such as rapid population growth, ongoing economic development and industrialization, heightened water pollution, expanding agricultural practices, and the impact of climate change. Next to the increased demand, also regulation on drinking water quality and environmental discharge of wastewater effluents has become more stringent. The European Union's Water Framework Directive (WFD) is an important directive aimed at protecting and enhancing the quality of water resources within its member states. Under the WFD, member states are obligated to attain and sustain favorable ecological and chemical conditions in their water bodies. The directive defines quality benchmarks for drinking water and outlines criteria for evaluating and managing contaminants in surface waters and groundwater. For the wastewater industry this means that more stringent quality requirements need to be met, putting increased pressure on the improvement of their operational performance. Moreover, the wastewater industry is currently facing dramatic changes, moving from energy-intensive wastewater treatment methods as end of pipe treatment to adopting low-energy, sustainable and more circular technologies capable of achieving energy-positive operation and resource recovery (Regmi et al., 2019). There is critical need for innovative approaches that address these challenges.

In this context, modelling is a powerful tool to support Water Resource Recovery Facility (WRRF) operators and engineers. Modelling biological, chemical, and physical processes is useful to acquire process understanding, simulate and test control strategies, and predict future behaviour under changing conditions (Gernaey et al., 2004). Mechanistic models in wastewater treatment have traditionally been favored by engineers due to their first-principles approach. However, these models often simplify complex processes and parameterizing, calibrating, and validating mechanistic models often require laborious experiments, which can be hindered by economic, time, or measurability constraints. On the other hand, data-driven methods are gaining interest with advancements in sensor technologies, enabling ubiquitous data collection and real-time monitoring. While data-driven approaches excel with large and high-quality datasets, data quality, extrapolation limitations and the lack of mechanistic interpretability pose chal-

lenges in the wastewater treatment sector. Hybrid modelling is a solution to bring forth the advantages of both mechanistic and data-driven models (Schneider et al., 2022). The application of hybrid modelling in the domain of WRRFs has the potential to foster automation (Rodriguez-Roda et al., 2002), increase efficiency, and increase the predictive power of models (Von Stosch et al., 2014). Whereas, online data is becoming more and more available from wastewater treatment plants, the application of hybrid models in this domain is still very limited. This thesis aims to investigate the potential of hybrid modelling on a unique, continuously monitored pilot-wastewater treatment plant.

2. LITERATURE REVIEW

2.1 Wastewater treatment

Wastewater engineering has come a long way from open dumping to collection and treatment prior to reuse in Water Resource Recovery Facilities (WRRFs), formerly called Waste Water Treatment Plants (WWTPs). Yet there is a need for more efficient WRRFs due to emerging concerns such as scarcity of freshwater sources due to increasing population, rising energy costs for the operation of treatment plants, more stringent discharge limits due to continued degradation of water bodies, etc. (Riffat and Husnain, 2013). The most important contaminants in wastewater effluent are Total Suspended Solids (TSS), organic matter (either expressed as Biological Oxygen Demand (BOD) or Chemical Oxygen Demand (COD)) and nutrients such as nitrogen and phosphorus. The EU Water Framework Directive (WFD) has set up a number of policy instruments for protecting inland waters. The Urban Waste Water Treatment Directive (UWWTD) supports the achievement of the WFD and aims at controlling emissions of nitrogen and phosphorus from WRRFs (Aloe et al., 2014). Nitrogen and phosphorus limits are applied to treatment facilities that discharge to sensitive areas. The UWWTD also contains other universal standards for the effluent quality in WRRFs. The standards are expressed as quantitative limits or removal rates for the parameters BOD, COD and TSS (Benedetti, 2006).

2.1.1 Wastewater treatment process

The wastewater treatment process usually starts with screening and the physical removal of constituents such as rags, sticks, grit, etc. (Gerba and Pepper, 2019). Then the primary treatment ensures the separation of particulate pollutants (e.g. debris, sand, grease, oils) in a primary settling tank due to gravitational separation (Hreiz et al., 2015). Approximately half the suspended organic solids settle to the bottom as primary sludge (Gerba and Pepper, 2019). Sometimes an advanced primary treatment is executed where solids and organic matter are removed by Chemically Enhanced Primary Treatment (CEPT). CEPT's main objective is to remove pollutants such as organic carbon and nutrients by adding chemicals and concentrating them in the sludge produced,

which can be considered valuable resources to increase energy recovery (Shewa and Dagnew, 2020). Another primary treatment concept currently under extensive development and testing is Rotating Belt Filtration (RBF), which is a technology designed for the removal of suspended solids and effluent organic matter as wastewater flows through the inclined section of a continuously rotating belt screen filter (Franchi and Santoro, 2015).

In a secondary treatment stage, the wastewater is biologically treated to remove organic carbon and soluble nitrogen and phosphates (Gerba and Pepper, 2019). As secondary treatment the activated sludge process is most commonly used given its efficiency, relatively simple operation and low cost (Hreiz et al., 2015). In this stage the wastewater is mixed with a quantity of microorganisms responsible for the degradation of organic matter and/or nutrients. An overview of the most important mechanisms in the tanks is listed in Table 2.1. The coupling of nitrification (converting ammonia to nitrate in the presence of oxygen: reactions 1 and 2 in Table 2.1) and denitrification, (converting nitrate to nitrogen gas in absence of oxygen: reaction 4 in Table 2.1) is the conventional method for N-removal. Organic matter and ammonium are removed in an aerobic tank where aeration provides oxygen for metabolising microorganisms that grow and remove pollutants and also provides mixing, allowing microorganisms to consume organic matter and nutrients. Aeration requires the pumping of air or pure oxygen into the tank, a process for which the energy used can be up to 40-65% of the total energy consumption in activated sludge systems including N-removal (Kirim, 2022). To achieve nitrate removal, an anoxic tank is needed for facilitating denitrification. Biological phosphorus removal necessitates the presence of both anaerobic and aerobic zones. The anaerobic zone enables bio-P bacteria (PAOs) to release phosphate as orthophosphate into the water. The aerobic zone allows PAOs to uptake the orthophosphate from the water. The removal of surplus sludge effectively eliminates stored phosphate from the treatment system (Janssen et al., 2002). In some cases phosphorus is chemically removed.

Table 2.1: Most important chemical reactions in the activated sludge process (Tchobanoglus et al., 2013).

Reaction	**Tank type**	**Organisms**
$NH_4^+ + \frac{3}{2}O_2 \longrightarrow Biomass + NO_2^- + H_2O + 2H^+$	Aerobic	Autotrophic
$NO_2^- + \frac{1}{2}O_2 \longrightarrow Biomass + NO_3^-$	Aerobic	Autotrophic
$O_2 + COD \longrightarrow Biomass + CO_2$	Aerobic	Heterotrophic
$NO_3^- + 6H^+ + COD \longrightarrow Biomass + \frac{1}{2}N_2 + 3H_2O$	Anoxic	Heterotrophic

As a final step of the secondary treatment stage the wastewater goes to the secondary settling tank (SST), where the denser solid phase (the activated sludge mass) is separated from the liquid phase (effluent) by gravitational settling. The SST combines the function of a thickener (producing a continuous underflow of thickened sludge for return to the biological reactor), a clarifier (producing a clarified final effluent) and a storage tank to store sludge during peak flows (Ekama et al., 1997). Water flows off the top of the tank and sludge is removed with a pump from the bottom (Gerba and Pepper, 2019). In the SST a part of the sludge is wasted and a part is recirculated to the secondary treatment (Torfs, 2015).

Some treatment plants also include a tertiary treatment step. In the tertiary treatment a series of additional steps further reduce organics, turbidity, nutrients, metals, and pathogens. Often some type of physicochemical treatment is involved such as coagulation, sand filtration, activated carbon adsorption of organics, reverse osmosis, and additional disinfection steps (Gerba and Pepper, 2019). Disinfection applied on technical scale include chemical (with the use of chlorine, ozone or peracetic acid) and physical methods (with the use of UV radiation or ultrafiltration) (Bray et al., 2021). Tertiary treatment becomes increasingly important, as stricter effluent quality regulations apply and the wastewater effluent is also reused, e.g. for irrigation (Illueca-Muñoz et al., 2008).

Primary, secondary and even tertiary sludges generated during wastewater treatment are usually subjected to a variety of treatments: screening, thickening, dewatering, conditioning and stabilization. Anaerobic digestion has gained popularity for energy recovery through biogas production (Section 2.1.3) (Gerba and Pepper, 2019). A schematic overview of the wastewater treatment process is given in Figure 2.1.

2.1.2 Challenges in the water sector

Energy is needed in all stages of wastewater treatment. In general, aeration is the largest energy consumer, followed by pumping, sludge treatment and mixing processes (Kirim, 2022). As the number of treatment plants worldwide increases and effluent quality requirements become more stringent, the issue of energy efficiency is receiving increasing attention from an environmental and economic point of view. A 130% increment in total energy consumption is estimated for the water treatment industry until 2040 (International Energy Agency, 2019). Thus, extensive research is currently being conducted to make WRRFs more energy-saving or even energy self-sufficient.

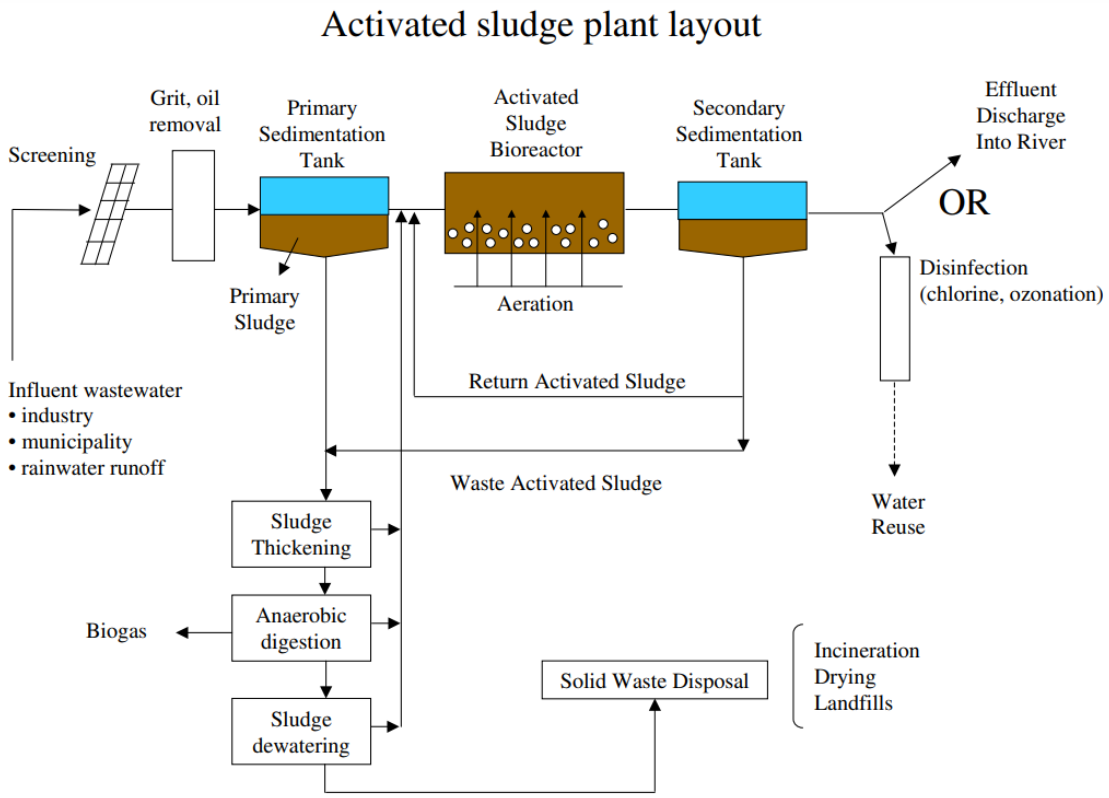

Figure 2.1: Overview of a biological wastewater treatment facility (Nopens, 2005).

Emerging contaminants (ECs) have recently attracted concern as they have been shown to significantly affect the natural environment but also to pose a major challenge to existing water treatment systems in terms of the effectiveness of their removal. The prominent classes of ECs include pharmaceuticals and personal care products, surfactants, plasticizers, pesticides, fire retardants, and nanomaterials (Rout et al., 2021). Conventional wastewater treatment processes are inefficient when it comes to ECs removal, hence innovative technologies are needed to tackle this problem.

Recent studies have identified WRRFs as potential sources of anthropogenic greenhouse gas emissions, contributing to climate change and air pollution. During biological wastewater treatment several greenhouses gasses are produced and emitted. CO_2 is mainly indirectly emitted due to the production of the energy required for the plant operation. CH_4 is mainly emitted trough the anaerobic digestion process. N_2O is intensively emitted during the N-removal process. This is problematic, because the global warming potention of N_2O is 298 times greater than that of CO_2, which it is why it is the main GHG of concern in WRRFs. Emissions of these gases need to be limited and research on mitigation strategies is ongoing (Campos et al., 2016).

2.1.3 Innovation in the water industry

The wastewater industry is currently shifting away from energy-intensive wastewater treatment towards low-energy, sustainable technologies. A recently developed energy-saving method to remove nitrogen is the anaerobic ammonia oxidation (anammox) process. In this process anammox bacteria oxidize the ammonium using nitrite as electron acceptor under anaerobic conditions without the need for an external carbon source of the conventional denitrification process (see Table 2.1). Nitrite can be produced through partial oxidation of ammonia by ammonia oxidizing bacteria known as partial nitrification or nitrate reduction by denitrifiers known as partial denitrification (Zhang et al., 2019). By using the anammox process instead of the conventional denitrification/nitrification process, oxygen demand is reduced by approximately 63%, allowing the WRRF to have a much lower energy consumption, and the sludge production can likewise be remarkably reduced by approximately 80%, while the addition of external COD is no longer required (Gerba and Pepper, 2019). Anammox is difficult to control and therefore the implementation of anammox technology on full-scale installations is still challenging and subject to research (Karthikeyan and Joseph, 2007). Next to energy saving technology also energy production from wastewater is gaining increased attention. Anaerobic digestion (AD) is a well developed and robust technology commonly used to recover energy from organic streams. AD is a biological process able to transform organic compounds into biogas, a mixture of CH_4, and CO_2. It is often implemented in the sewage sludge treatment to stabilize the sludge and recover energy in the form of biogas (Silvestre et al., 2015).

The water industry is currently trending towards a circular economy where resources in the wastewater are recovered. Many valuable resources can be recovered from WRRFs including nitrogen, phosphorus and heavy metals, but also water itself. Anaerobically digested sewage sludge is a valuable resource for recovery due to its rich organic matter and nutrient content, and recent advancements in recovery technology have been made. The methods can be broadly categorized based on their ability to recover a specific nutrient or multiple elements simultaneously, with examples including ammonia stripping, struvite precipitation, ion exchange, membrane filtration, and thermal treatments (Di Costanzo et al., 2021). Nutrient recycling recovers nutrients in the wastewater as soil amendments or fertilizers for beneficial uses. However, there are also risks associated with this technique. The main risks related to agricultural use of sewage sludge are the potential presence of pathogens and pollutant enrichment in soils, plants and an-

imal pastures and the subsequent entry into the food (Gianico et al., 2021). Controlled struvite crystallization is another way of recovering nutrients by precipitating struvite from sludge digester liquors. Struvite is a phosphate mineral and a slow release fertilizer rich in magnesium, ammonium and phosphate. It can also be used as a building material and adsorbent (Li et al., 2019a).

Also water itself can be recoverd from wastewater treatment plants through using wastewater effluent for different applications. To tackle water scarcity and drought, the Flemish Government created the Blue Deal. It is an ambitious program that addresses water scarcity and drought through numerous actions. An example of a project being implemented is the investment in five business processes that will use 33.3 million m^3 of treated sewage water as a source of water for process or drinking water in the future (VMM, 2022). Treated wastewater also has a high potential to be used in agricultural irrigation. Even though wastewater effluent is an interesting water source for irrigation, health risks related to pathogen exposure, accumulation of heavy metals, salts, antibiotics, growth hormones and other hazardous substances into the soil, etc. need to be considered and mitigated (Ungureanu et al., 2020). The reuse of effluent in other, less evident applications like cooling, flushing toilets, ecosystem restoration, etc., is also an emerging concept (Neczaj and Grosser, 2018).

2.2 Modelling wastewater treatment

Mathematical models are a powerful tool to address current challenges in wastewater treatment. Models provide a simplified representation of the physical system and can thus be used to simulate a system's behaviour under different operational or design scenarios and for optimization. Models can be used for process understanding, operation and control, design and diagnosis. Innovations in the water sector also create new challenges in the operation and control of a WRRF. The purpose of a model will significantly affect its structure and complexity (Olsson and Newell, 1999). Many different classifications have been produced for the different model types: mechanistic versus data-driven models, dynamic versus static models, deterministic versus stochastic models, etc. (Jeppsson, 1996b).

2.2.1 Difference between mechanistic and data-driven modelling

The most important classification of models for this work is the difference between mechanistic and data-driven models. The approach of the two models differs fundamentally. Mechanistic models, often referred to as white-box models, are based on fundamental engineering and scientific knowledge about the physical, chemical, and biological mechanisms that affect a system, with the relationships themselves defined by modellers and assumed to be known (Schneider et al., 2022). Data-driven models are often associated with black-box models, because they are developed using algorithms and other methods that do not reference fundamental mass, charge and energy balances. However, it is misleading to consider data-driven models as purely black-box models and mechanistic models as purely white-box models. For example, some mechanistic models used in wastewater treatment modelling include a variety of Monod-type switching functions which are mathematically tractable but are not supported with theory, while some data-driven methods can be transparant, e.g. linear regression or decision trees (Regmi et al., 2019; Schneider et al., 2022).

In recent decades, engineers have favoured mechanistic approaches for modelling wastewater treatment processes. These models usually simplify the complex processes in the WRRF, such as aeration, mixing or aggregation of particulates. These type of models require only limited data-input and have a high interpretability. Mechanistic models have a pre-defined structure and thus require extensive knowledge about the system (Schneider et al., 2022). They are capable of extrapolating the process performance to a wide variety of process operating conditions (Hvala and Kocijan, 2020). This type of modelling requires extensive and time-consuming parameterisation, calibration and validation. Parameter estimation based on numerical optimization algorithms may lead to non-unique parameter estimates and many local optima (Hvala and Kocijan, 2020).

A data-driven model is based on empirical relationships between the input and the output (Jeppsson, 1996b). In contrast to mechanistic models, the structure of data-driven models is directly derived from data and does not require extensive knowledge about the system. Data-driven models have a low interpretability and have high data requirements. They are also limited in extrapolation power (Schneider et al., 2022).

Mechanistic modellers use the term calibration when referring to the parameter identification process and the term validation to check if the calibrated model is successful in unseen data. In data-driven modelling, the training phase involves iteratively adjust-

ing the model's parameters to minimize the difference between predicted and observed outputs in the training dataset. The validation dataset is a separate dataset used to evaluate the model's performance during training. In neural networks, validation specifically focuses on optimizing the model's hyperparameters for improved performance. Once the a data-driven model is trained and validated, it can be used in a non-training mode to process new input of a test dataset, allowing evaluation of its performance on unseen data.

2.2.2 Mechanistic models

2.2.2.1 Biokinetic process modelling

A generally-accepted mechanistic model for the biokinetic processes in the biological reactor of wastewater treatment plants is the family of Activated Sludge Models (ASMs), of which Activated Sludge Model No.1 (ASM1) is the most commonly used. These models are suitable to model activated sludge processes, but they can also be integrated with other models for specific modelling purposes. For instance, they can be coupled with membrane filtration models to model membrane reactors. (Henze et al., 2000). The ASM1 was primarily developed to model the removal of organic carbon and nitrogen, but it also aims to accurately describe sludge production and oxygen consumption (Gernaey et al., 2004). The model uses a COD based modelling technique where all carbon materials have been expressed as equivalent amounts of COD, as it provides a link between electron equivalents in the organic substrate, the biomass and the oxygen used. Organic carbon compounds and nitrogenous compounds were classified into a limited number of fractions based on biodegradability and solubility leading to 13 state variables and 8 fundamental processes, resulting in 13 mass balance equations with 19 parameters (Henze et al., 2000). The biokinetic mass balances in the system are described using Ordinary Differential Equations (ODEs).

In summary, the following processes are considered: the *aerobic growth* of heterotrophic biomass uses biodegradable substrate and oxygen as an electron acceptor. In the absence of oxygen the heterotrophic organisms are capable of using biodegradable substrate with nitrate as the terminal electron acceptor. The process will lead to a production of heterotrophic biomass and nitrogen gas (*denitrification*). Ammonia is oxidized to nitrate via a single-step process (*nitrification*) resulting in production of autotrophic biomass. Ammonia is also used as the nitrogen source for synthesis and incorporated

into cell mass. Biodegradable soluble organic nitrogen is converted to ammonia (*ammonification*) in a first-order process by heterotrophs. *Hydrolysis* of entrapped organics and organic nitrogen result in the conversion of particulate organic compounds to readily available soluble compounds (Jeppsson, 1996a).

Finally, decay of heterotrophic and autotrophic organisms is considered. The ASM1 uses the death-regeneration hypothesis in an attempt to single out the different reactions that take place when organisms die. In this hypothesis, decayed cell material is released again through lysis. One fraction is non-biodegradable and remains as an inert residue while the remaining fraction is considered to be slowly biodegradable and used for cell growth (Jeppsson, 1996a). An overview of the ASM1 equations is presented in the Gujer matrix format.

The ASM1 has been extended by adding new components or processes to the original model. The Activated Sludge Model No. 2 (ASM2) is an extension of ASM1 and reuses the same concepts as ASM1. The most remarkable change in ASM2 is that biological and chemical processes for phosphorus removal are included. The ASM2 contains 19 processes, 19 state variables and 57 parameters. The Activated Sludge Model No. 3 (ASM3) has the same focus as ASM1, but it fixes some shortcomings of the original ASM1 model. For example, endogenous respiration instead of the death-regeneration concept is used and the storage in cell internal components is included as a process (Henze et al., 2000). Several extensions to the ASM model family have also been developed to simulate greenhouse gas emissions of WRRFs, among which the dominant nitrous oxide (N_2O) (Maktabifard et al., 2022; Guo and Vanrolleghem, 2014). The ASM1_AN model is an extension of ASM1 with anammox and two-step nitrification and denitrification. The model considers NO_2-N as an intermediate model variable and distinguishes the growth of ammonium oxidizers and nitrite oxidizers (Van Hulle, 2005).

2.2.2.2 Hydraulic modelling

Mixing is a critical process within wastewater treatment plants. The movement of the biomass caused by aeration and mixing improves the contact of the particles in the mixture. It helps prevent the formation of stagnant, untreated areas within the treatment system. However, in large reactors of WWTPs, non-ideal mixing frequently occurs leading to dead zones, shortcuts or mass-transfer limiting conditions. Investigating the mixing behavior of such tanks is critical in understanding and optimizing the transport of various

components involved in the process. Hydraulic modelling is used to describe mixing patterns in wastewater treatment tanks. Wastewater treatment processes can be modelled using simplified hydraulic assumptions such as the Tanks In Series (TIS) approach (Laurent et al., 2014). The TIS model consists of the serial connection of Completely Stirred Tank Reactors (CSTRs). By increasing the number of tanks, different mixing regimes can be covered, from perfectly mixed flow (one CSTR) to plug flow (infinite CSTRs) (De Clercq et al., 1999). Modelling using TIS has been useful to limit the model's computational complexity while still describing the mixing behaviour, yet it doesn't capture complex 3-dimensional transport-reaction interactions that occur in multi-phase, multi-scale WRRFs (Laurent et al., 2014).

A more accurate tool for modelling hydraulics is Computational Fluid Dynamics (CFD). CFD involves solving Partial Differential Equations (PDEs) of continuity, momentum, and energy related to fluid dynamics in 2 or 3 dimensions using numerical approximations. The main benefit of using CFD is its ability to predict and visualize the flow pattern, mixing behavior, and other important flow characteristics of various types of fluids in complex geometries with extremely high spatial resolution (e.g. in centimeters). The method can simulate fluid flow in both 2-dimensional and 3-dimensional domains (Jalilnejad et al., 2022). CFD is however computationally expensive and therefore is not widely applied for dynamic modelling of full-scale wastewater treatment processes. Recently, compartmental models are also being developed. These are models where several CSTRs are coupled in a 2-dimensional grid with exchange flows between them. Their structure can be derived from a CFD model. They have the advantage of being more detailed than TIS but much less computationally demanding than a full CFD model (Jourdan et al., 2019).

2.2.2.3 Modelling the secondary settling process

The secondary settling process is has a crucial role as it directly influences the effluent quality and the biomass concentration in the system. The development of a general settling process model is quite challenging due to the co-occurrence of distinct settling regimes. A sedimentation model includes both time and space dependence and thus describes the process using PDEs (Bürger et al., 2011). The SST model is combined with biokinetic mass balance models composed of ODEs and implemented using numerical solvers available in common WRRF modelling software. To facilitate this implementation, the PDE that describes the SST process is discretized into a set of ODEs. To achieve this, the settling tank is divided into horizontal layers with a uniform concentration within

each layer. Historically, the most widely adopted model for SST is the model by Takács et al. (1991). Since then several models have been developed aiming to incorporate more realistic functions for the different settling regimes or improving the numerical methods for the PDE (Ramin et al., 2014; Bürger et al., 2011). Among these is the model developed by Bürger et al. (2011) (BD model). The BD model ensures the solution of the governing PDE by reliable numerical methods for ODEs. The number of layers can be set by the user. The model accounts for sludge compression and inlet dispersion phenomena by adding a compression function (d_{comp}) and a dispersion function (d_{disp}) to the PDE (Equation 2.1). The BD model includes several phenomena in a modular way instead of trying to lump different phenomena in a single parameter or function.

$$\frac{\partial X}{\partial t} = -\frac{\partial}{\partial z}\Big(F(X,z,t)\Big) + \frac{\partial}{\partial z}\left(\Big(d_{comp}(X) + d_{disp}(z,Q_f(t))\Big)\frac{\partial X}{\partial z}\right) + \frac{Q_f(t)X_f(t)}{A}\partial(z) \qquad (2.1)$$

where X: =X(z,t) is the total solid concentration as function of depth z and time t, F(X, z, t) is the flux function comprised of both the advective and hindered settling flux. The feed source term contains the feed flow (Q_f), the feed concentration (C_f) and the Dirac function ($\partial(z)$).

Previous modelling efforts have mainly considered SSTs as nonreactive, but some reactive settler models have been developed (Bürger et al., 2016; Kirim et al., 2022). A significant amount of the overall sludge inventory of the WWTP can be stored at the bottom of the SST. This accumulated sludge mass can turn the bottom of the settling tank into an additional biological reactor. Within this sludge blanket biological reactions may occur. Especially denitrification is known to happen when sufficient sludge mass in present at the bottom of the SST. Reactive settler models combine a SST model with biological reactions (Kirim et al., 2022).

2.2.3 Data driven models

Data-driven models are particularly successful when dealing with problems involving large and high-quality data sets. They can be categorised based on their learning algorithms, i.e. supervised, unsupervised, and reinforcement learning. Supervised learning methods are applied when the data is in the form of input variables and output target values. The algorithm learns the mapping function from the input to the output. A variety of algorithms, including linear regression, Artificial Neural Network (ANN), Decision Tree (DT), Support Vector Machine (SVM), naive Bayes, K-Nearest Neighbor (KNN), Ran-

dom Forest (RF), etc. have been developed. Unsupervised learning is usually used to handle data without labels. It classifies the training data into different categories according to their different characteristics, mainly based on dimensionality reduction and clustering. Principal Component Analysis (PCA) and K-means are the commonly used unsupervised machine learning algorithms (Zhu et al., 2022). In reinforcement learning, an agent learns the optimal mapping of situations to actions (called policy) through a trial-and-error search guided by a scalar reward signal. During the learning process, the model receives either rewards or punishments for the actions it performs (Nian et al., 2020).

2.2.3.1 Artificial neural networks

Research of Bahramian et al. (2022) reports that neural networks are currently the most popular method for data-driven modelling of wastewater treatment. NNs have many advantages that make them a tool used in numerous applications. One of the main advantages of NNs their ability to model nonlinear and unknown dynamics. NNs can capture functional relationships among the water quality data. Even when the underlying relationships of obtained data are difficult to describe, ANN models still work. Moreover, ANNs require fewer prior assumptions and can achieve higher accuracy compared with traditional approaches (Chen et al., 2020). Other reported advantages of NNs are more fault tolerance and the ability to work with incomplete data and missing data (Mijwel, 2018). Numerous studies have demonstrated the usefulness of neural networks in the analysis of time series (Wu et al., 2020; Borovykh et al., 2017).

Several studies have successfully implemented neural networks for modelling wastewater treatment processes. Research on wastewater treatment using neural networks has primarily focused on simulation of treatment performance, monitoring and control, and classification. For instance, in the study by Ráduly et al. (2007), an ANN was trained on available input-output data from a WWTP to predict the concentrations of effluent ammonium, total nitrogen, BOD_5, COD, and TSS. In the study conducted by Hong et al. (2007), a neural network approach was developed to estimate nutrient concentrations in real-time, addressing the issue of delayed measurements. Furthermore, Onkal-Engin et al. (2005) used a convolutional NN to classify microbeads in urban wastewater.

The central idea of NNs is to extract linear combinations of the inputs as derived features, and then model the target as a nonlinear function of these features (Hastie et al.,

2009). A neural network is a structure of interconnected units or layers of large number of neurons. Each neuron within the network is capable of receiving input signals, processing them, and generating an output signal. The output signal is composed of weighted synapses that determine the strength of connections, an adder that combines the weighted input data, and an activation function that regulates the output amplitude of the neuron (Zhang et al., 2018). To account for nonlinearities within systems, it becomes necessary to use nonlinear activation functions. Various types of nonlinear activation functions exist, some frequently used functions are the sigmoidal (or logistic) activation function, the hyperbolic tangent function and the nonsigmoidal rectified linear unit (ReLU) (Pomerat et al., 2019). The first layer of an NN is the input layer and the last one forms the output layer. In between one or more neuron layers, also called hidden layers, can be located that act like feature detectors. The number of neurons in each layer is determined by the desired accuracy in the predictions. The accuracy and convergence of the different types of NN highly depend on hyperparameters (e.g. the numbers of hidden layers and neurons in each layer), which are often arbitrarily chosen Nikbakht et al. (2021).

Neural networks can either be supervised or unsupervised (Park and Lek, 2016). Considering different characteristics of input data, there exist many different types and structures of neural networks. Bahramian et al. (2022) reports that feedforward neural networks, radial basis function neural networks, recurrent neural networks, convolutional neural networks, long short-term memory neural networks and Kohonen self-organizing neural networks are the most popular ones used as stand-alone models in the wastewater treatment context. Feedforward NNs (FFNNs) were the first type of NN invented. The name originates from the fact that the information only travels forward in the network, there are no loops. In Figure 2.2 a simple example of an FFNN is given in which each circle is a neuron and each arrow is an interconnection between two neurons in which information is transferred. This type of neural network has a popular structure that has been used for many applications, and it gives a good performance in many cases. The Radial Basis Function (RBF) network uses radial basis functions as activation function. The output of the network is a linear combination of radial basis functions of the inputs and neuron parameters. RBF networks have the advantages of simple design, good generalization, and strong tolerance towards input noise. They can be used for various applications, including time series forecasting (Sharkawy, 2020).

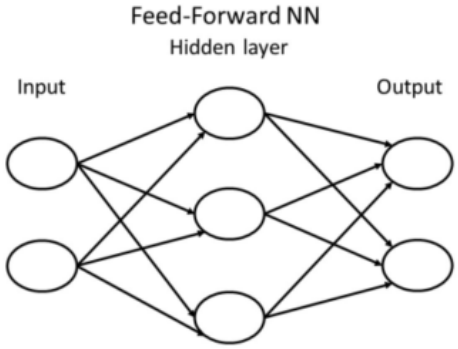

Figure 2.2: Schematic representation of a feedforward neural network with two neurons in the input layer, three hidden cells and two neurons in the output layer (Singh and Chauhan, 2009).

Recurrent Neural Networks and Long Short-Term Memory networks

Recurrent neural networks (RNNs) are models that have at least one feedback loop which allows them to take some context into account in their decision function. Hence RNNs are able to remember information through time, which make them a useful tool for time series forecasting (Sharkawy, 2020). They are limited to look back in time for approximately ten timesteps. This is because the feedback signal vanishes or explodes. Figure 2.3 shows a schematic representation of a simple RNN.

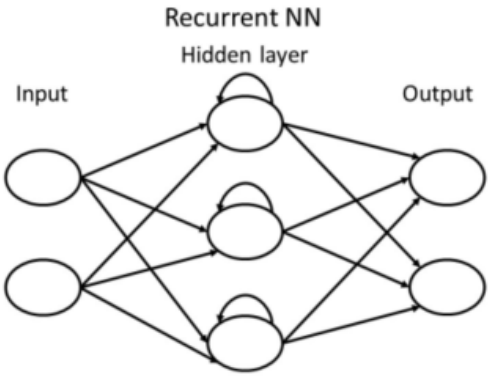

Figure 2.3: Schematic representation of a recurrent neural network (RNN) with two neurons in the input layer, three recurrent cells and two neurons in the output layer (Shukla and Iriondo, 2020).

Long Short-Term Memory Recurrent Neural Networks (LSTM-RNNs) address the issues concerning RNN as they are capable of learning long-term dependencies. LSTM networks have the ability to remember more than 1 000 time steps, depending on the complexity of the network's architecture (Staudemeyer and Morris, 2019). An LSTM network is the same as a standard RNN, except that the summation units in the hidden layer are replaced by memory blocks. Each block contains one or more self-connected memory cells and three multiplicative units, the input, output and forget gates (Graves, 2012). The input gate controls the input activation flow, the output gate controls the flow of the output

activation and the forget gate scales the internal state of the cell before it goes through the self-connection, thereby adapting the cell's memory. Some modern LSTM blocks also have peephole connections from the gates to the internal cell state (Sak et al., 2014). Figure 2.4 provides an illustration of an LSTM memory block with a single cell.

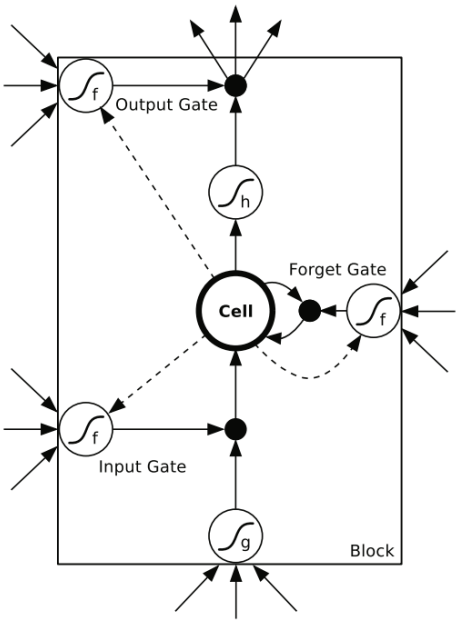

Figure 2.4: LSTM memory block with one cell, with 'f' the gate activation function 'f' and 'g' and 'h' the cell input and output activation functions. The weighted 'peephole' connections from the cell to the gates are shown with dashed lines (Graves, 2012).

Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are often used for processing multidimensional inputs. CNNs have shown impressive performance in classification tasks such as image recognition, speech recognition and natural language processing (Koprinska et al., 2018). However, in more recent times, CNNs have gained attention as a valuable tool for time series forecasting (Wang et al., 2017; Durairaj and Mohan, 2022). A typical architecture consists of repetitions of a stack of several convolution layers and a pooling layer, followed by one or more fully connected layers. A convolution layer performs feature extraction, which typically consists of a combination of a convolution operation and activation function. The convolution operation involves applying a small array of numbers, called a kernel, across the input. An element-wise product between each element of the kernel and the input tensor is calculated at each location of the input tensor and summed to obtain the output value in the corresponding position of the output tensor, called a feature map. Multiple kernels can be applied to form different feature maps. The size and number of kernels are important hyperparameters of the convolution operation. Feature maps are then passed through a nonlinear activation function. A pooling layer

provides a typical downsampling operation which reduces the in-plane dimensionality of the feature maps. The output feature maps of the final convolution or pooling layer is typically flattened, i.e., transformed into a one-dimensional array, and connected to one or more fully connected layers, also known as dense layers. The activation function applied to the last fully connected layer is usually different from the others and needs to be selected according to the goal of the CNN (Yamashita et al., 2018). CNNs have high potential to predict time series because they have the capability to autonomously learn and extract features from raw data without the need for prior knowledge or manual feature engineering. CNNs also learn filters that represent recurrent patterns in the series and use them to predict future values (Koprinska et al., 2018). They can also be effective for handling noisy time series by removing noise at each layer and extracting only the meaningful patterns (Borovykh et al., 2017).

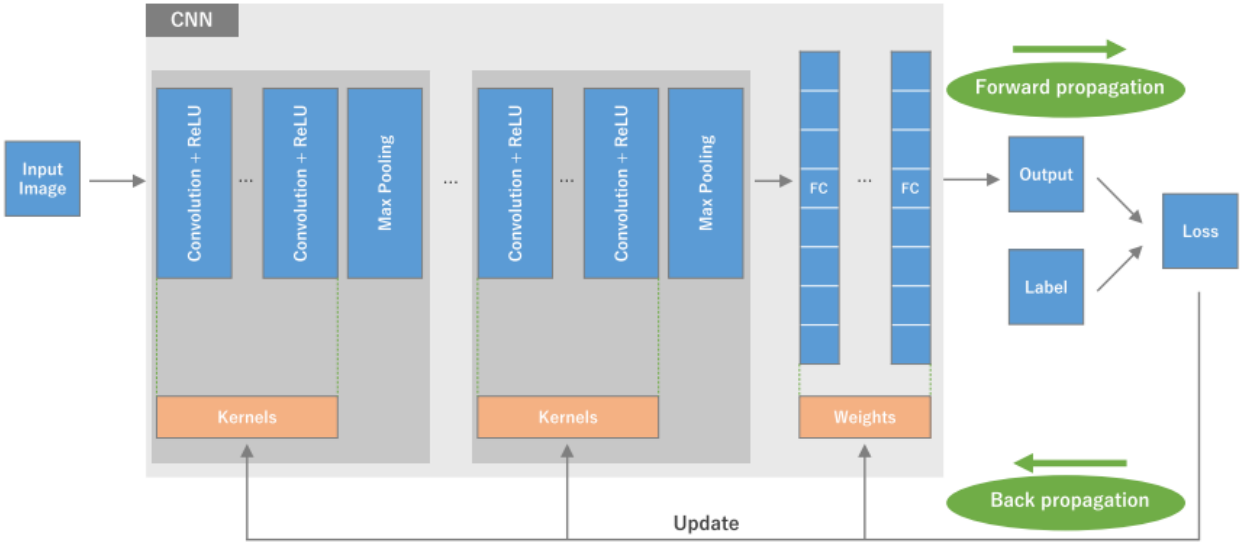

Figure 2.5: An overview of a convolutional neural network architecture and the training process (Koprinska et al., 2018).

2.2.4 Hybrid models

The application of mechanistic models can be limited due to knowledge gaps or the (over)simplification of complex processes. There are still some relevant processes which are not understood clearly enough to be put in a model for use in simulations, such as the formation of nitrous oxide and the conversions within the biological phosphorus removal process (due to the variability of phosphate accumulating organism metabolics) (Regmi et al., 2019). This results in only a partial description of the process or system, which can be valuable, but often is not robust or resistant to significant changes (Schneider et al., 2022). On the other hand, the lack of sufficient high-quality data is a limitation for data-

driven models. Typical wastewater process sensor data and laboratory data are noisy, contain inconsistent values or biases and are often not available for periods of time or at the required sampling frequency (Regmi et al., 2019). The advancements in water treatment technology and the shift towards a more circular economy have resulted in a transformation in the water treatment modelling objectives. Water reuse may require more case-specific quality targets that need more detailed modelling and the digitalisation of the water sector offers more input data for modelling. This creates an environment an environment where the use of hybrid models (HMs) becomes increasingly advantageous. Hybrid models combine mechanistic and data-driven modelling techniques and benefit in this way from the advantages of the two approaches. Mechanistic models have the capability to conserve critical process knowledge and maintain extrapolation capabilities, while data-driven approaches are capable of discovering hidden relationships and patterns that may not be detected by mechanistic models. Regmi et al. (2019) reports that in the future hybrid models will emerge as dominant methods to model wastewater treatment processes, which necessitates the collaboration of mathematicians, computer scientists, systems engineers, and software developers as well as chemical, environmental and civil engineers, biochemists and biologists.

2.2.4.1 Different architectures

Various architectures have been explored for merging mechanistic and data-driven components in hybrid models. Three different architectures can be distinguished: serial, parallel and surrogate models. In a serial hybrid model, the output of one model is used as input for the other model (Schneider et al., 2022). Generally, the output of the data-driven model is used as the input for the mechanistic model since the data-driven component is capable of predicting dynamics that are not accurately modelled by the mechanistic component, such as poorly defined reaction kinetics (Lee et al., 2002). The parallel architecture is a suitable option when assigning the missing knowledge to a specific (sub)process becomes challenging or when delineating this subprocess is difficult. In cases where the structure of the mechanistic model is inaccurate, the parallel arrangement can have a better performance than the serial arrangement, as the parallel data-driven model can partially compensate for any structural mismatch in the mechanistic model (Von Stosch et al., 2014). One way to make parallel hybrid models is to train the data-driven component to learn the residuals between the mechanistic model and historical data, after which the results are fused (Anderson et al., 2000). In another

type of parallel hybrid models, both mechanistic and data-driven models are trained to make predictions, after which the results are weighted and combined (Peres et al., 2001). Lastly, surrogate models are data-driven models that are trained on the output of a mechanistic model to create a computationally more efficient model, allowing rapid and/or large-scale simulations.

Once the model architecture has been defined, the parameter values for each component can be identified, either independently or in conjunction (Schneider et al., 2022).

2.2.4.2 Current use of hybrid modelling in wastewater treatment

Around 10-20 years ago, there was an increase in the number of papers on hybrid modelling, which coincided with the growing popularity of data-driven methods. Hybrid models have proven useful in a number of applications within the field of wastewater treatment. For example, dynamic influent data is a major bottleneck for applying ASMs to evaluate the design and operational scenarios for WRRFs. Hybrid models have emerged as effective solutions to address this problem, by developing influent generators focusing on the wastewater dynamics at the entrance of WRRFs, which can be coupled with ASMs. Different influent generator models using different data-driven methods have been developed to predict the incoming flow rate and pollutant loads to WRRFs (Zhu et al., 2015; Flores-Alsina et al., 2014). Hybrid models also demonstrate their usefulness in the development of soft-sensing models. Soft-sensing models, often data-driven, can predict variables that serve as input information to controllers, which are used in conjunction with ASM models (Wang et al., 2022). As a result, soft-sensing models are valuable in the effective operation of advanced control systems. The implementation and advancement of hybrid models also have the capacity to encourage and enhance the effective deployment of digital twin technology in WRRFs (Torfs et al., 2022). The study conducted by Lee et al. (2002) provided a proof of concept for the effectiveness of parallel hybrid modelling in improving the prediction of process variables. However, there is a scarcity of subsequent research that delves deeper into the application of parallel hybrid modelling, particularly in the context of full-scale wastewater treatment plants. Serial hybrid modelling, in which where a data-driven model is developed of a subprocess that is less well-described and subsequently integrated in the overall mechanistic model, has been relatively underexplored in the literature. Despite its potential advantages, there is a scarcity of studies investigating this approach in depth.

2.2.4.3 Challenges for hybrid models

The hybrid modelling of wastewater treatment processes encounters various critical challenges that need corresponding development efforts. The availability of Good Modelling Practice guidelines for the development of mechanistic models is well-established (Rieger et al., 2012). The integration of hybrid modelling paradigms into existing frameworks for WRRF modelling raises important considerations regarding the extension of existing protocols. A comprehensive protocol for developing hybrid models should address critical questions. For instance, it should outline how to effectively couple a data-driven model with a mechanistic model. Another crucial aspect is determining which data should be used to construct each model: should the same data be used to determine the parameters of both the mechanistic and data-driven methods?

A second challenge for hybrid models is the fact that uncertainty arises from both the mechanistic and data-driven components. However, previous studies often lack the quantification of uncertainty, limiting the assessment of trust in hybrid models. Dynamic quantification of uncertainty is essential to fully establish the potential of hybrid models.

A third challenge for using hybrid models is the task of balancing complexity between the mechanistic and data-driven components. One aspect to address regarding this challenge is the effort required for calibration of the mechanistic model compared to the data-driven model. Another important aspect is determining the acceptable level of error that can be compensated by the data-driven model.

A fourth challenge encountered when developing hybrid models is the possibility of obtaining inaccurate output from the data-driven component, necessitating the need to address data quality issues to ensure reliable results.

Selecting the suitable architecture for constructing hybrid models presents a fifth challenge in the development of hybrid model. It is important to ascertain the specific purposes for employing parallel models and sequential models. Identifying the most effective data-driven models for distinct purposes is also crucial in achieving desired outcomes.

Finally, for using hybrid models to their full potential as a standard practice there is a need for compatible platforms or Application Programming Interfaces (APIs) to be able to connect the mechanistic and data-driven parts of the model.

In this thesis, the first, third and fifth challenges are addressed (Section 3).

3. RESEARCH OBJECTIVES

This thesis aims to investigate the potential of different hybrid modelling architectures to improve the predictive power of wastewater treatment plant models. The study is performed based on the pilot-scale wastewater treatment plant. In previous research, a mechanistic model was developed for a pilot-scale wastewater treatment plant. However, accurate effluent quality prediction remains a challenge in this model. The objectives of this thesis are:

1. To explore the feasibility to develop a parallel hybrid model for enhancing effluent quality prediction in the wastewater treatment plant.
2. To explore the optimal model structure of the data driven component by comparing the performance of a parallel hybrid model with a Long Short-Term Memory Recurrent Neural Network (LSTM-RNN) as data-driven component and parallel hybrid model with a Convolutional Neural Network (CNN) as data-driven component.
3. To explore best practice in setting up parallel hybrid models with respect to calibration and training efforts between the mechanistic and the data-driven component.
4. To investigate the suitability of using a serial hybrid model to forecast the oxygen transfer coefficient (K_{La}), as well as the corresponding Dissolved Oxygen (DO) output to improve the predictions of oxygen needs in the system.

4. MATERIALS AND METHODS

4.1 PilEAUte WRRF

In this study, a hybrid model is developed for a pilot-scale WRRF at Université Laval, called pilEAUte. A scheme of the plant setup is shown in Figure 4.1. The pilEAUte receives domestic wastewater from the student residence and kindergarten, and rainwater from a parking lot. The water is transferred from the campus sewer system to the plant's inlet via a pumping station, which also shreds the large particles present in the raw wastewater. The water is pumped to a storage tank with a volume of 5 m^3 where the influent is homogenized. After equalization, the wastewater is directed to a primary settler with a flow rate of 1.1 m^3/h. The primary settler is designed to allow gravity settling of particles in the wastewater and has a volume of 2.8 m^3. A Y-strainer with a 7/8-inch mesh size is installed on the pipe between the store tank and the primary settler. Its purpose is to screen large solid particles from the influent wastewater, protecting the pump. All pumps in the pilEAUte plant are equipped with Variable Frequency Drives (VFDs) to adjust flow or pressure according to demand. The primary effluent is conveyed to two biological reactors called pilot and co-pilot which are operating in parallel with a flow rate of 0.5 m^3/h. If there is an excessive amount of primary effluent that cannot be processed by the biological reactors, it is discharged into the sewer system along with the waste primary sludge.

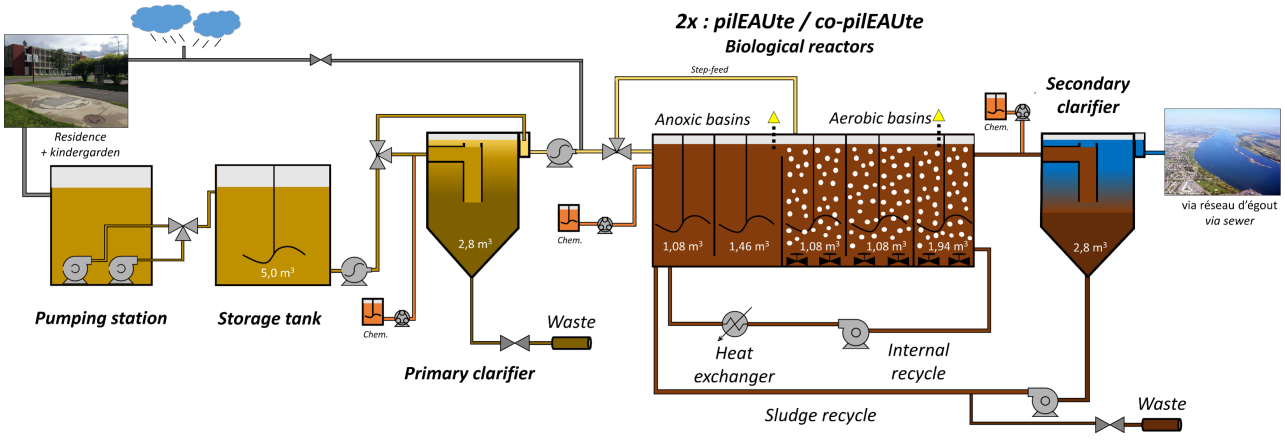

Figure 4.1: Scheme of the pilEAUte WRRF at Université Laval.

Each biological reactor has an identical design and is equipped with a pre-denitrification configuration, enabling them to effectively remove carbon and nitrogen. Within each biological reactor, there are two anoxic basins equipped with mechanical stirrers, three aerobic basins, and a secondary clarifier. The aerobic basins are aerated with compressed air via diffusers at the bottoms of the tanks. It is possible to control the aeration flowrate for each basin through mass flow controllers based on the DO-concentration in basin 4. The DO setpoint in basin 4 is set to 3 mg/L. Each basin is equipped with individual air flow lines, which can be adjusted using ratio controllers. The biological reactors both have an internal recycle system that circulates water from the last basin back to the first basin to sustain pre-denitrification, with an internal recycle flow rate of 1.5 m^3/h. Additionally, a sludge recycle loop transfers thickened sludge from the secondary clarifiers to the first basins to maintain an optimal mixed liquor concentration, with a flow rate of 0.5 m^3/h.

A waste flow rate from the bottom of the secondary clarifiers is controlled to ensure a sufficient sludge age for the growth of nitrifiers. The effluent and sludge wastage of the secondary clarifier are discharged back to the sewer system. Under standard operational conditions, the inflow from the biological reactors to the secondary clarifiers is 1.0 m^3/h. Each of the secondary clarifiers has a volume of 2.8 m^3. The clarifiers have a height of 2.5 m, with the feeding point situated 1.1 m above the bottom where the conical shape begins. To facilitate sludge transportation to the bottom, the conical part is equipped with a rotating chain that operates on an hourly basis.

The pilEAUte gathers monitoring data through two data acquisition systems: SCADA and monEAU. All data, including their metadata, from these sources is automatically archived in a comprehensive SQL database known as datEAUbase. The pilEAUte is monitored with sensors at the outlet of the primary clarifier, the biological reactors, the recycle streams and the outlet of the secondary clarifiers. All sensors have a sampling frequency of 1 minute, except for the RODTOX sensor, which can make a measurement every 5 seconds. An overview of the pilEAUte online monitoring system is given in Table 4.1.

Location of Sensor	Sensor name	Manufacturer	Monitored Parameters	Principle	Monitoring System
Primary effluent	Spectro::lyser	S::CAN	COD_{total}, $COD_{soluble}$, TSS	UV-VIS spectrophotometry	monEAU
	Ammo::lyser	S::CAN	NH_4-N, K, Temperature, pH	Ion selective electrode	monEAU
	Varion	WTW	NH_4-N, K, Temperature, pH	Ion selective electrode	monEAU
	Conductivity meter (Inductive)	Hach	Conductivity, Temperature	Potentiometric	SCADA
	RODTOX	Hach	BOD_{short_term} and Toxicity	Respirometry	monEAU
Biological reactors – Basin 2	Solitax	Hach	TSS	Infrared duo-scattered light technique	SCADA
Biological reactors – Basin 4	DO meter (LDO)	Hach	Dissolved oxygen, Temperature	Luminescent	SCADA
Biological reactors – Basins 2&5, Secondary effluent	Trescon/Purcon	Hach	NH_4-N, NO_2-N, NO_X-N (NO_2-N & NO_3-N total)	Potentiometric or Photometric	monEAU
Sludge recycle line	Solitax	Hach	TSS	Infrared duo-scattered light technique	SCADA
Secondary effluent	Turbidity meter	WTW	Turbidity	Nephelometric	monEAU
	Varion	WTW	NH_4-N, NO_3-N, K, Cl	Ion selective electrode	monEAU
	pH meter	Hach	pH, Temperature	Potentiometric	monEAU

Table 4.1: Online monitoring equipment of pilEAUte (Kirim, 2022).

4.2 Mechanistic model

A pre-existing mechanistic model created by Kirim (2022) is used in this study to create a hybrid model (Figure 4.2). The model was developed in the software platform WEST 2017 (DHI, 2017). The pilEAUte operational data for the time period February 1st - March 31st 2018 was chosen to set-up the model since this period reflects the normal operational conditions in terms of flow patterns. The raw monitoring data was subjected

to the univariate data validation method developed by Alferes and Vanrolleghem (2016), and further processed using outlier detection and data smoothing filters before modelling (Alferes and Vanrolleghem, 2016; Philippe, 2018). The time periods used for the different mechanistic modelling purposes are presented in Figure 4.3. An input file is generated to be used as influent data in the mechanistic model that includes the influent flow rate, total COD (CODt), soluble COD (CODs), and ammonium (NH_4-N). Prior to incorporating these inputs, a full influent fractionation was performed. For the fractionation of COD, a backward calculation was done between the effluent and influent of the primary clarifier to be able to apply the influent COD fractionation study for the pilEAUte by Li et al. (2019b), resulting in the fractionation of COD into inert particulate COD (X_i), biodegradable COD (X_s), inert soluble COD (S_i), and readily biodegradable COD (S_s). Since NH_4-N is the only parameter that is measured at the primary effluent, an influent total nitrogen fractionation was performed to determine the ammonium nitrogen (S_{NH}), soluble organic nitrogen (S_{ND}) and the particulate organic nitrogen (X_{ND}). Since continuous alkalinity data is not available for the influent, the influent alkalinity is calculated based on the measurements from the influent ammonium sensor. In order to perform simulations, the mechanistic model also requires input data including the measured temperature in basin 4 and the sludge waste flowrate in the secondary settling tank.

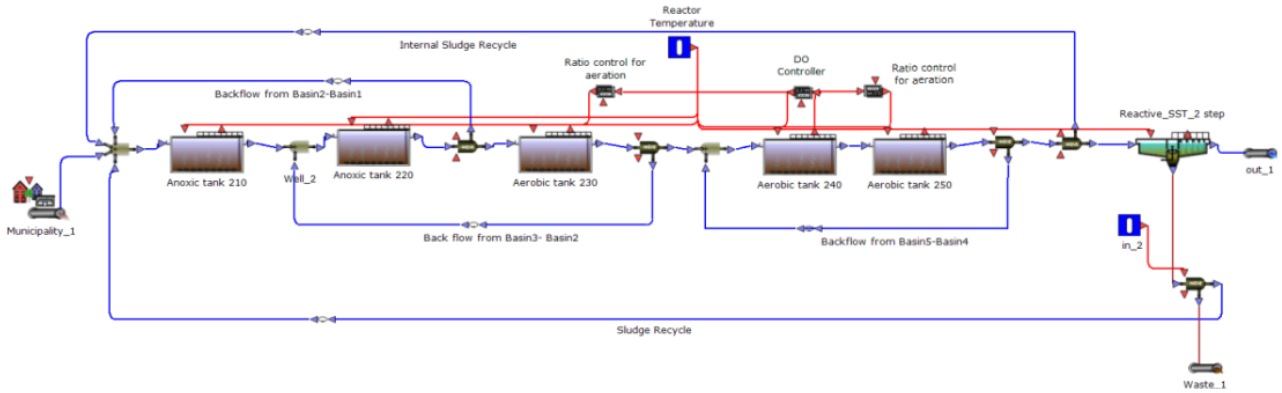

Figure 4.2: The mechanistic model layout of the pilEAUte WRRF developed in WEST by Kirim (2022).

Figure 4.3: Timeline with the time periods used for the different mechanistic modelling purposes.

The initial phase of model calibration involved performing a hydraulic characterization of the plant and calibrating the hydraulic model. This step was particularly crucial due

to significant backflows observed between the different basins of the biological reactors. The wastewater normally flows from basin 1 to basin 5 in the plant. However, the baffles between basins 4 & 5, basins 2 & 3 and also basins 1 & 2 were insufficient in effectively separating the reactors when aeration is on in the tanks, resulting in significant and undesirable backflows. Due to significant backflows occurring between basins 4 & 5, it was necessary to treat these basins as a single fully mixed reactor. Two tracer test were conducted for standard operational conditions (serving as a reference) and step-feed operational conditions (involving the feeding of influent and internal sludge recycle to basin 3). The results of the hydraulic model for both tracer tests can be found in Appendix A. The final hydraulic model incorporates a Tanks In Series (TIS) configuration with five Continuously Stirred Tank Reactors (CSTRs). The flow rates of the backflows between the basins were determined using data from tracer tests and integrated into the final model set-up.

The pilEAUte model implemented in WEST describes each basin as an ideally mixed, activated sludge tank with constant volume. The oxygen transfer coefficient is the key factor for characterizing aeration and the transfer of mass between gas and liquid phases. The air flowrate is the only measured variable in the pilEAUte's biological reactors that allows for quantifying aeration in the actual system. So, the air flow rate is used for calibrating the aeration model. This calibration involves adopting a correlation between the oxygen transfer coefficient (K_{La}) and the air flow rate (Qair) to predict aeration in the pilEAUte model:

$$K_{La} = \frac{\rho * Q_{air} * OTE}{(\beta * S_{O,sat} - S_O) * V} \Rightarrow \frac{K_{La}}{\frac{\rho * OTE}{(\beta * S_{O,sat} - S_O) * V}} = K_{La} * \gamma = Q_{air} \tag{4.1}$$

where ρ is the density of air, β denotes a correction factor for the oxygen saturation, $S_{O,sat}$ and S_O are the standard oxygen saturation and the dissolved oxygen concentration in the aerated tank and V is the volume of the tank. The "γ" conversion factor is dermined to be 1.9 by comparing the calculated air flowrate with the measured values. In the pilEAUte's actual system, dissolved oxygen is controlled with a setpoint of 3 mg/L in Basin 4 by adjusting the air flowrate, and the same air flowrate is applied to all basins. To incorporate this into the model, a proportional-integral controller is used to regulate the KLa in Basin 4, with a DO setpoint of 3 mg/L. The model uses the inverse relationship between reactor volume and K_{La}, and ratio controllers are employed to adjust the aeration in Basins 3 and 5 based on the K_{La} in Basin 4. For the secondary clarifier, a one-dimensional reactive settler model was used based on the Bürger-Diehl framework including hindered settling and compression processes (Kirim et al., 2022).

The ASM1_AN biokinetic model, an extension of ASM1, was used as the biokinetic model because in the research goal of Kirim (2022) was to investigate process optimization scenarios and applicability of short-cut N removal processes. For calibration of the biokinetic parameters, stepwise Monte Carlo-based calibration methodology inspired by Mannina et al. (2011) was followed and 1-month of online data with a 1min interval was used. The calibration of the biokinetic model used operational data from the biological reactors, including air flowrate, dissolved oxygen, and TSS concentrations. Additionally, TSS concentration in the sludge recycle stream was taken into account, after the data was adjusted by considering the TSS mass balance around the secondary clarifier. Also effluent NH_4-N and NO_3-N were used in the model calibration process. The ASM1_AN model has 51 model parameters and 27 model parameters including, all the kinetics, were selected after a pre-selection procedure based on engineering expertise and the available data. A local sensitivity analysis was applied to determine the influential model parameters, after which 17 model parameters were selected for calibration. Prior to calibration, three different parameter subsets were selected, each focusing on a different group of output variables in the calibration procedure. The next step involved the group calibration of the model parameters, where each subset of parameters was calibrated based on the model outputs and objective function by carrying out Monte Carlo simulations. Parameter sampling was performed using Latin Hypercube Sampling. The final calibrated parameter values can be found in Table 4.2.

The calibrated model results for nitrate and ammonium nitrogen for the calibration and validation period can be found in Appendix B. The validation period indicated that the model lacks predictive power for effluent nitrate. The order of magnitude is the same for most of the validation time period for effluent nitrate, but significant differences were observed for a few days. A more detailed description of the mechanistic model can be found in the PhD thesis by Kirim (2022).

4.3 Data-driven models

4.3.1 Data preprocessing

The data preprocessing and the data-driven model set-up was performed using Python (version 3.9.13, Van Rossum and Drake (2009)), Jupyter Notebook (version 6.4.12, Kluyver et al. (2016)) and the packages pandas (version 1.4.4, McKinney et al. (2010)) and

	Parameter		Default Value	Calibrated Value	Unit
Subset 1	b_H	Decay Coefficient for Heterotrophic Biomass	0.62	0.70	1/d
	f_XI	Fraction of Biomass Converted to Particulate Inert Matter	0.1	0.13	-
	k_h	Maximum Specific Hydrolysis Rate	3	3.20	gCOD/(gCOD*d)
Subset 2	b_NH	Decay Coefficient for NH4 Oxidizing Autotrophic Biomass	0.05	0.06	1/d
	b_NO	Decay Coefficient for NO Oxidizing Autotrophic Biomass	0.033	0.04	1/d
	K_HNO2_NO	Nitrous Acid Half-Saturation Coefficient for NO Oxidizing Autotrophic Biomass	0.000872	0.000061	$gCOD/m^3$
	K_NH3_NH	Ammonia Half-Saturation Coefficient for NH4 Oxidizing Autotrophic Biomass	0.75	0.0057	$gNH_3\text{-}N/m^3$
	K_SH	Substrate Half-Saturation Coefficient for Heterotrophic Biomass	20	8.74	$gCOD/m^3$
	mu_H	Maximum Specific Growth Rate for Heterotrophic Biomass	6	4.77	1/d
	mu_NH	Maximum Specific Growth Rate for NH4 Oxidizing Autotrophic Biomass	0.8	0.71	1/d
	mu_NO	Maximum Specific Growth Rate for NO Oxidizing Autotrophic Biomass	0.79	0.95	1/d
Subset 3	K_NO2_H	Nitrite Half-Saturation Coefficient for Denitrifying Heterotrophic Biomass	1	3.31	$gCOD/m^3$
	K_O_NH	Oxygen Half-Saturation Coefficient for NH4 Oxidizing Autotrophic Biomass	0.6	0.25	gO_2/m^3
	K_O_NO	Oxygen Half-Saturation Coefficient for NO Oxidizing Autotrophic Biomass	1.5	0.27	gO_2/m^3
	K_OH	Oxygen Half-Saturation Coefficient for Heterotrophic Biomass	0.2	0.10	gO_2/m^3
	n_NO2	Correction Factor for Anoxic Growth of Heterotrophs on Nitrite	0.6	0.92	-
	n_NO3	Correction Factor for Anoxic Growth of Heterotrophs on Nitrate	0.6	0.49	-

Table 4.2: Calibrated biokinetic model parameter values (Kirim, 2022).

NumPy (version 1.21.5, Oliphant et al. (2006)). The input variables used for the data-driven models are outlined in Section 5.1.1. The preprocessing of the input variables was already carried out in Kirim (2022) prior to developing the mechanistic model. Prior to conducting the training process, the input dataset undergoes standard scaling using the "StandardScaler" from the sklearn preprocessing module (Equation 4.2). To reduce the computational burden, the input data set is downsized by subsampling data points at 10-minute intervals from the original 1-minute interval. This downsampling process is achieved with the "interp" function from the NumPy package in Python, which performs a one-dimensional linear interpolation. The measurements used for comparison with the results are also downsampled using this method. The data splitting method is explained in Section 5.

$$z = \frac{x - \mu}{\sigma} \tag{4.2}$$

in which z represents the scaled value of the variable x, μ represents the mean of x and σ represents the standard deviation of x.

4.3.2 Neural network set-up

In this research, hybrid models are created by combining neural networks with the mechanistic model. For the development of the neural networks, the Keras library (version 2.11.0, Chollet et al. (2015)) from the package TensorFlow (version 2.11.0, Abadi et al. (2016)) was used. In order to connect the neural network models developed in python to the mechanistic model, the WEST Tornado kernel is called through the command line interface. This allows the python scripts to interact with an already compiled WEST model in an autonomous way for various tasks including simulation execution, model parameter adjustment and running scenario analysis. To do this, a windows batch file is created that includes the necessary commands to set the appropriate path to the Tornado kernel library and execute the XML file associated to the WEST model.

In this research, LSTM-RNN and CNN are used as data-driven components of the hybrid models due to their effectiveness in handling time series data. The input data is reshaped into a 3D matrix to align with the required format for LSTM-RNNs and one-dimensional CNNs. The first dimension of this matrix is the number of time series data points that are used for training, validating or testing the model, also called samples. The second dimension is the number of previous steps that should be taken into account, this number

of time steps is considered as a model parameter that needs to be optimized. The third dimension is the number of input variables. The neural networks are trained on a training dataset, the hyperparameters are determined using a validation dataset, and the performance is evaluated using a test dataset. The following hyperparameters of the LSTM-RNN and CNN are optimized:

- Time steps: The number of time steps taken into account is also considered as an hyperparameter that needs tuning, as mentioned before.
- Number of layers: Optimized between 1 to 5 layers.
- LSTM hidden units: Optimized for each LSTM layer.
- CNN kernel size and filters: Optimized to determine the dimensions of the convolutional filter and the number of filters applied to the input data.
- Activation function: Chosen per layer (ReLU or tanh) to introduce non-linear transformations.
- Epochs: Optimized for each neural network to determine the number of times the entire dataset is passed through during training.
- Learning rate: Optimized to control the step size at which the optimizer adjusts the model's parameters during training, with the Adam optimizer used in this case (Kingma and Ba, 2015).
- Loss function: Chosen between Mean Absolute Error (MAE) and Mean Square Error (MSE) to measure the discrepancy between predicted and actual values during training.
- Batch size: Selected for each constructed neural network, this defines the number of samples processed per iteration during training

To overcome the potential of overfitting a drop-out rate of 0.1 was chosen between each layer of the neural networks. This drop-out rate indicates what fraction of the neurons should be dropped (ignored) during the training of the model. In the initial trial, the hyperparameters of the neural networks were optimized using a random search tuner with the KerasTuner framework (O'Malley et al., 2019). This approach involves automatically exploring different combinations of hyperparameter values within predefined ranges to find the configuration that minimizes the validation loss. However, neural networks are

inherently stochastic in nature due to factors such as the random initialization of weights in neurons, which still led to large variations in model performance across different training runs, and no optimal parametersets were found using this procedure. It was decided to use a manual hyperparameter optimization approach. Manual optimization allows for a more thorough exploration of hyperparameter space by trying out a variety of random values for each hyperparameter and observing their impact on the validation loss. By systematically testing different values for each hyperparameter, the hyperparameterset that yields the lowest validation loss was determined for every developed neural network. For every neural network, the MAE was selected as optimal loss function and a batch size of 64 was selected. The final values of the other hyperparameters for every developed neural network are reported in Section 5. The Root Mean Square Error (RMSE) is used to evaluate the performance of the computed models, with

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{4.3}$$

where $\hat{y}_i$ represents the model prediction value, y_i represents the measurement value, and n represents the number of data points.

5. RESULTS AND DISCUSSION

5.1 Parallel hybrid modelling

The work of Kirim (2022) revealed that despite an extensive and thorough calibration procedure, the mechanistic model of the pilEAUte plant still exhibits limited predictive power for effluent nitrate. This limitation was the driving force behind investigation undertaken in this work, aiming to explore the potential and identify an optimal methodology for developing a parallel hybrid model. The purpose of this hybrid model is to compensate for the missing information in the mechanistic model by incorporating additional information and ultimately enhancing the overall predictive power of the model.

5.1.1 Model set-up

In the first part of this study, parallel hybrid models are developed to address inaccuracies in predictions made by the mechanistic model. To create the parallel hybrid models, both LSTM-RNNs and CNNs are trained to predict the residuals between the mechanistic model outputs and the measurement data for effluent nitrate and TSS. The predicted residuals are then added to the mechanistic model outputs. This approach is motivated by the potential presence of valuable dynamic information within the residuals, and the neural networks are expected to uncover this information by learning patterns in the residuals. The input data used to train the NNs consists of the mechanistic model output of effluent nitrate or TSS, the measured air flow rate in basin 4, and the measured, non-fractionated variables that are also fed to the mechanistic model: temperature in basin 4 and influent CODt, CODs, TSS and NH4-N. Additionally, the input data for the neural network includes measurements of pH in the influent. The influent flow rate and sludge waste flow were not included as inputs for the neural network. This is because the influent flow rate is kept constant at 12 m^3/d and thus does not provide useful information for the data-driven model. The sludge flow rate only varies in the long term, so its effect is expected to be primarily seen over a very long period, which is why it was also not included. On the other hand, the Mixed Liquor Suspended Solids (MLSS) in basin 2 was measured in the pilEAUte and shows significant fluctuation over the used time

period. This variable would be interesting to include in future studies as it could provide valuable information for the NNs. The graphs in Appendix C present the input data for the neural networks throughout the entire modelling period. A diagram of the parallel hybrid network for effluent nitrate can be found in Figure 5.1.

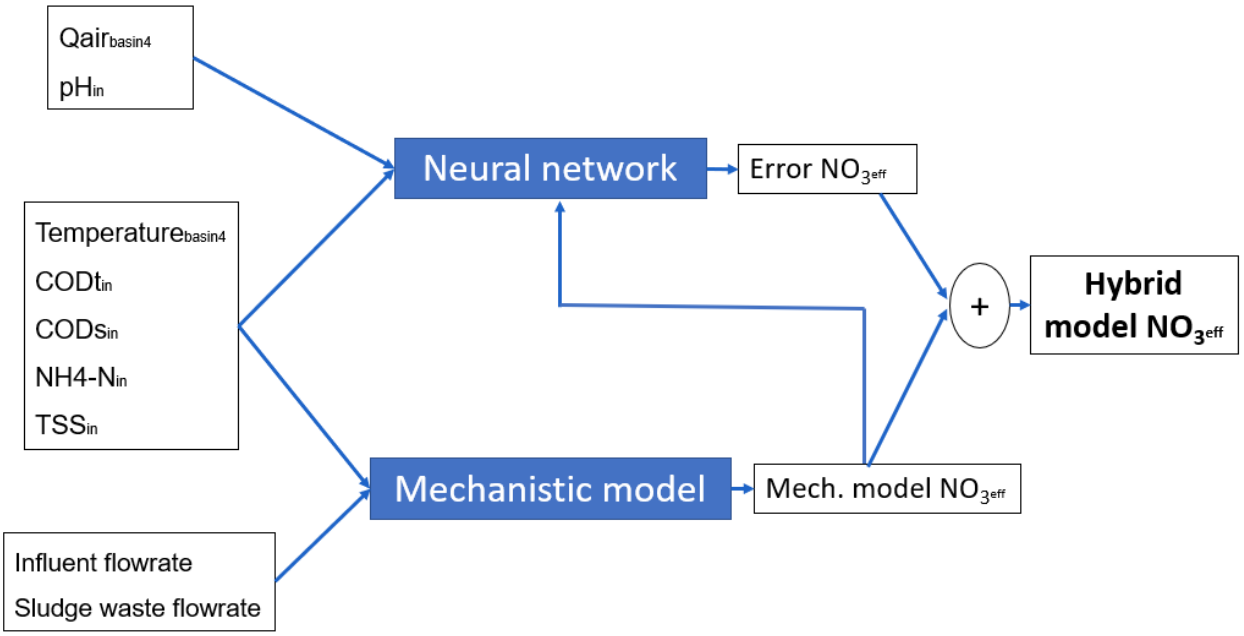

Figure 5.1: Overview of a parallel hybrid model developed to predict effluent nitrate.

5.1.2 Parallel hybrid modelling for improved effluent nitrate predictions

5.1.2.1 Parallel hybrid model using LSTM-RNN and CNN - a comparative analysis

Parallel hybrid models have been established using various approaches to address specific research objectives. The first objective is to assess the difference between using an LSTM-RNN and a CNN as a data-driven component of the parallel hybrid model. Both types of neural networks are known to perform well for time series prediction. However, it would be valuable to determine how effectively they work in combination with a mechanistic model and which type excels in terms of information extraction in this specific application. To examine this, both an LSTM-RNN and a CNN architecture were established to model the residuals between the effluent nitrate output generated by the mechanistic model and the measured effluent nitrate. Subsequently, the predicted residuals were added to the mechanistic effluent nitrate outputs. The NNs are trained and validated using the calibration dataset of the mechanistic model. The neural network training is

Table 5.1: Final hyperparameter sets for a LSTM-RNN hybrid model and a CNN hybrid model trained on the calibration dataset of the mech. model for effluent nitrate.

Prediction	Network	Time steps	Structure details	Learning rate Adam optimizer	Epochs
NO3	LSTM-RNN	30	Number of layers: 1 Units: 10 Activation: ReLu	0.0005	75
NO3	CNN	30	Number of layers: 1 Filter: 12 Kernel size: 9 Activation: ReLu	0.00005	200

conducted using the largest portion of the calibration dataset (78%). A smaller portion of this dataset (22%) is reserved for validating the neural network to determine the optimal neural network structure and parameter values. The NNs also generate outputs for a test period, for which the validation dataset of the mechanistic model is used. The data splitting approach can be found in Figure 5.2. Appendix C contains boxplots of the neural networks' input data for each variable during each period. The Table 5.1 contains the final set of hyperparameters determined for each neural network used in this experiment.

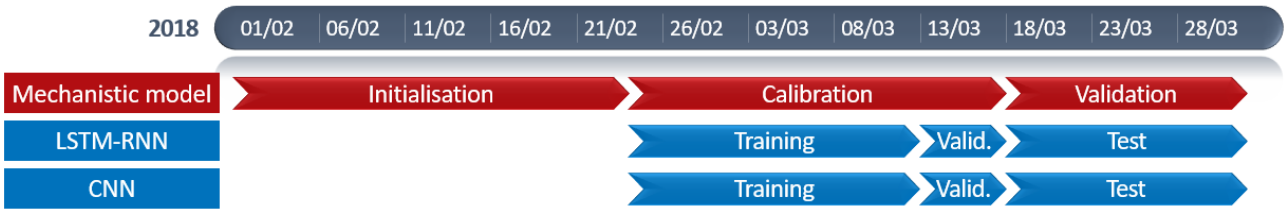

Figure 5.2: Timeline with the time periods used to develop a parallel LSTM-RNN hybrid model and a parallel CNN hybrid model for effluent nitrate.

Figure 5.3 shows the outputs of the hybrid model using LSTM-RNN and the hybrid model using CNN compared to the mechanistic model outputs and the measured values for effluent nitrate during the training, validation and test periods. The computed RMSE values for the training, validation and testing period for this experiment can be found on the timelines in Figure 5.4. The optimal parameter value for the number of past time steps to consider is determined to be 30 for both CNN and LSTM-RNN, with each step representing a duration of 10 minutes. This implies that the neural networks effectively analyze the preceding 5-hour period to predict a residual value. It is interesting to compare this time interval with the Hydraulic Retention Time (HRT), which represents the average residence time of wastewater within a biological reactor, thereby determining the contact

duration between pollutants and microorganisms (Von Sperling, 2007). The average HRT in the pilEAUte WRRF typically falls within the range of 10 to 15 hours. Additionally, it is worth considering the Solids Retention Time (SRT), which denotes the average duration that activated-sludge solids remain in the system. The SRT is longer than the HRT and can range from days to weeks, depending on various factors within a WRRF.

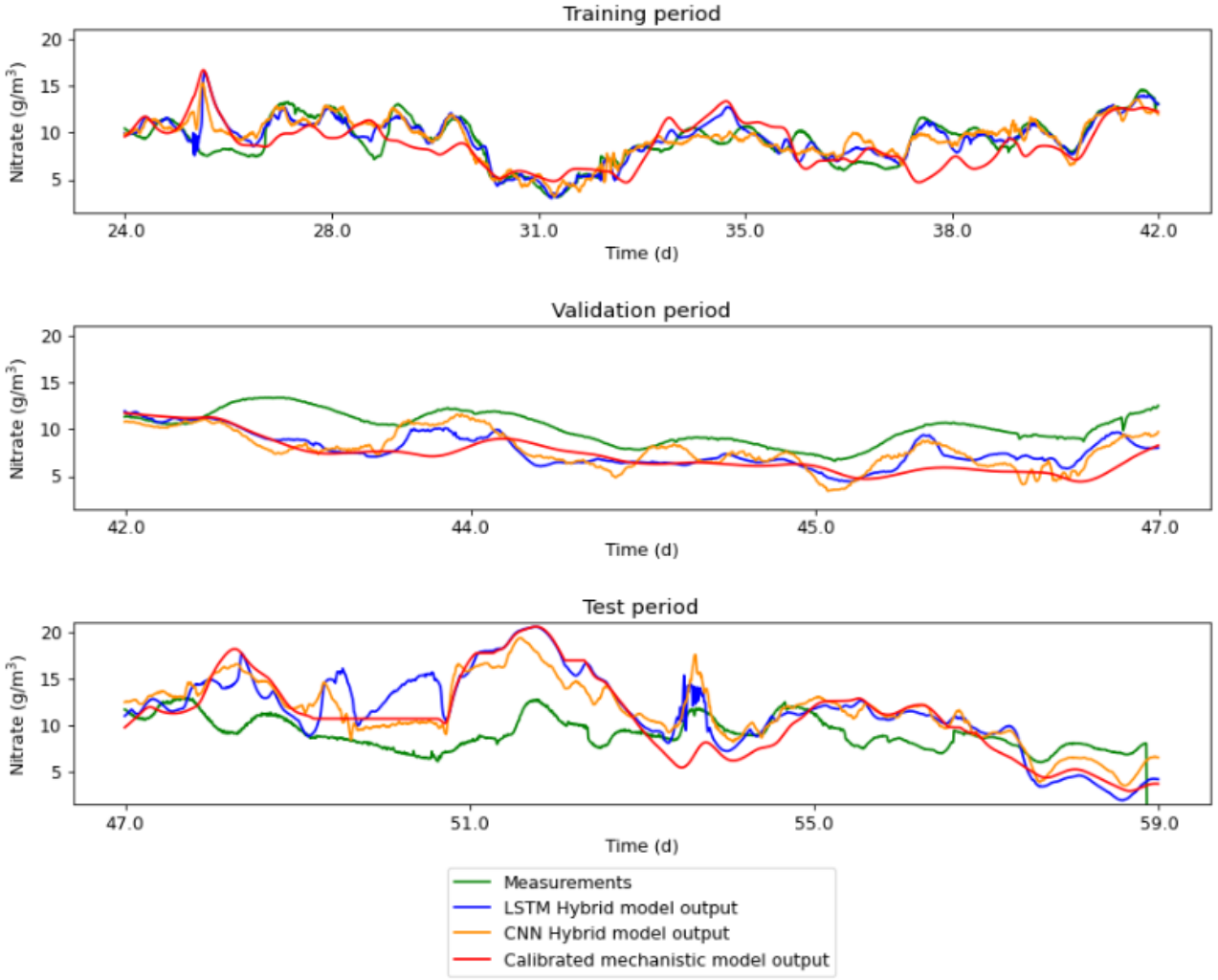

Figure 5.3: LSTM-RNN hybrid model output and CNN hybrid model output compared to mechanistic model output and measurements for effluent nitrate for the train, validation and test period. The NNs are trained on the calibration dataset of the mech. model.

Figure 5.4: Timeline with the RMSE per time period (Tr=train, V=validation, Te=test) used to develop a parallel LSTM-RNN hybrid model and a parallel CNN hybrid model for effluent nitrate, compared to the RMSE of the mechanistic model in these periods.

Based on visual assessment, the effluent nitrate outputs of the hybrid model incorporating LSTM as the data-driven component reveal the following findings. The output of the LSTM hybrid model during the training period demonstrates satisfactory performance, capturing the appropriate dynamics for the majority of the period. However, on day 25 and day 34 the HM output still shows peaks that didn't occur in the measurements. During the validation period, the hybrid model's output is less accurate, occasionally aligning closer to the measured values but it overall fails to reproduce the correct dynamics. The LSTM hybrid model's performance deteriorates during the test period. The model incorrectly predicts peaks that are not present in the measured data, for example from day 47 to 51, and fails to capture the underlying dynamics entirely during this period. These findings are supported by the results for the RMSE, with the RMSE of the LSTM hybrid model during the test period being higher than the RMSE of the mechanistic model. During the development of the LSTM HM, it was observed that the predictions generated by the LSTM exhibited substantial variability. Even when training the LSTM multiple times with identical hyperparameters, there was considerable variation in the results during each subsequent testing phase. There are several potential factors that can account for the observed phenomena. One possible explanation is that the LSTM's ability to extract meaningful dynamics from the residuals during the training phase may be insufficient. According to the findings of Lowe et al. (2022), training of LSTM requires large and diverse datasets making the training process difficult. This limitation could hinder its performance during subsequent validation and testing phases. Insufficient performance of the model could also be attributed to the fact that the model might not have been properly set-up, by which the LSTM fails to gather the necessary information required for making accurate predictions. Other possible factors of the insufficient performance of this hybrid model are discussed below.

The hybrid model using CNN as a data-driven component clearly outperforms the mechanistic model during the NN training period and successfully captures the underlying dynamics during the first half of the training period, except for the peak on day 25. However, the dynamics of the second half of the training period are captured to a somewhat lesser extent. The CNN outputs for the training period exhibit slightly lower performance compared to those of the LSTM, as reflected in the corresponding RMSE values. Throughout the validation period, the output of the hybrid model with CNN shows reduced accuracy. The CNN HM does not achieve the level of accuracy necessary to fully capture the precise dynamics. The CNN HM output is closer to the measurements compared to the

mechanistic model during this period, but the LSTM HM performs slightly better than the CNN HM, which is again reflected in the RMSE values. During the test period, the predictions of the hybrid model with CNN outperforms the hybrid model with LSTM-RNN and the mechanistic model, also reflected in the observed RMSE values. The hybrid model with CNN shows a capability to capture specific dynamics to some extent. An example of this is observed in the HM outputs during the period around day 51-53, where despite not accurately reproducing the correct order of magnitude, it manages to capture the underlying dynamic in the effluent nitrate. This suggests that the CNN HM has learned a structural pattern related to this specific behavior. However, it is crucial to acknowledge that the CNN HM still faces challenges in accurately representing the dynamics present in the test dataset, as indicated by the observed RMSE. These findings emphasize that while the CNN hybrid model shows improvement, it has not yet reached the desired level of accuracy required to fully capture the nuances of the system's dynamics.

The observed outcomes of the two model could be ascribed to several potential factors. A possible explanation to consider is that the training data may not be entirely representative of the test dataset. For example, in the input data for the neural network, there is an influent stop around day 48-50 during the test period, which also affects other measured variables such as Qair, CODt and CODs (Appendix C). The boxplots of the input data for every time period show that the values for CODs, CODt and NH4-N for the test period are overall slightly lower than those in the training and validation periods. Also the Qair is slightly lower and has wider range of variation. However, considering the overall lack of significant differences in the boxplots across the three periods, it can still be concluded that the training period is representative for predicting the test period. Another potential explanation for the observed phenomena could be a lack of a structural relationship between the neural network's inputs and the residuals. The absence of a clear and consistent relationship could inhibit the neural networks of generating reliable predictions during periods on which they are not trained. The lack in prediction power during the test period could also potentially be attributed to overfitting of the NN to the training period, despite the inclusion of regularization parameters in the neural network structure. The HM outputs exhibit some noise, particularly during the training phase. However, considering that the latter half of the calibration period also presents suboptimal performance, the likelihood of overfitting is relatively low in this scenario. It is less likely that the prediction method alone is the primary cause of the issue. In the subse-

quent discussion, supporting evidence and explanations will be provided to elaborate on the reasons behind the aforementioned statement.

To verify the validity of the employed methodology to develop parallel hybrid models, two additional experiments were conducted. The discussed methodology of this thesis was applied to a confidential dataset, for which the results cannot be presented due to confidentiality constraints. This confidential dataset was already used in an independent study conducted by an external researcher that successfully developed a parallel hybrid model to predict residuals. The external research showed an increased accuracy of the parallel hybrid model outputs for effluent nitrate relative to the outputs of the mechanistic model. Thus, the methodology employed to create a parallel hybrid model in this thesis (outlined in section 4.3) is tested on the confidential dataset to evaluate its effectiveness in achieving more accurate outputs for effluent nitrate. The results for the training and validation periods exhibit significant improvements compared to the mechanistic model outputs. However, the most significant observation is that in this experiment, the test period is accurately predicted, demonstrating the ability of the method to capture the underlying dynamics effectively.

In the aforementioned external research, the MATLAB toolbox 'Regression Learner App' was employed to construct the data-driven component of the parallel HM (version 2021b, (MathWorks, 2022)). The Regression Learner App trains regression models to predict data. The application can be used to perform automated training to search for the best regression model type, including linear regression models, regression trees, Gaussian process regression models, support vector machines, kernel approximation models, ensembles of regression trees, and neural network regression models. As a second experiment to verify the validity of the employed methodology, the Regression Learner App was used to develop a parallel hybrid model aimed at predicting the residual of effluent nitrate, using the pilEAUte plant data from the current study. This predicted residual was then added to the outputs of the mechanistic model. The results of this experiment indicate that even with this approach, the prediction of effluent nitrate in the test period remained poor. This strongly suggests that the cause of the inadequate prediction in the test period lies not in the methodology of this study itself.

LSTM-RNNs are typically used to predict time series due to their ability to capture long-term dependencies in sequential data. However, in this study, the model fails to effectively learn meaningful patterns from the input data during the training phase. The LSTM-RNN struggles to capture and represent the underlying patterns and dependencies

present in the time series data. As a result, when faced with the test data, the model lacks the necessary knowledge and context to accurately recognize and predict patterns in the residuals. This observation is noteworthy because the existing literature suggests that LSTM-RNNs perform well as data-driven components in hybrid models, as for example demonstrated in study of Dong et al. (2023). It is possible that the dataset is not large and diverse enough for the LSTM-RNN to learn something useful in this study. On the contrary, the hybrid model with CNN as data-driven component is able to detect some dynamics in the residuals during the test period, however the predictive performance of the model still falls short of achieving high levels of accuracy. The CNN has the ability to concentrate on important time serie features and can extract these features through its own learning process. The convolution filter in CNN can be useful for identifying patterns at different time scales or detecting specific events in the time series. This was also confirmed by Wang et al. (2019). The training process of the LSTM-RNN model is also observed to be slower compared to that of the CNN. As CNN outperformed LSTM-RNN as data driven component of the hybrid model, the decision was made to employ the CNN architecture for the next part of the study.

The results of this experiment could indicate that training a data-driven component of a hybrid model on the calibration period of the mechanistic model may not be a good idea, since both LSTM and CNN became tuned to the specific characteristics of the error of the calibration period, resulting in reduced ability to make accurate predictions for other periods. To set-up the mechanistic model, an extensive and thorough calibration procedure was executed. Hence, it raises the question of whether there is still enough relevant structural information left in the residual of this calibration period for the data-driven models to capture. Additionally, the question arises to what extent the missing dynamics and knowledge in the output of the mechanistic model are already compensated for by the calibration itself. If indeed the residuals of the calibration period in the mechanistic model output contain minimal or no relevant structural information, and the missing dynamics and knowledge during this period are already compensated for by the calibration process, the data-driven component faces a challenging task to perform effectively. It must successfully tackle two objectives in order to accurately predict in unseen time periods: 1) recognize and compensate for the overfitting of the mechanistic model, and 2) incorporate the missing dynamics. These requirements are demanding, and there is a realistic possibility of failure of the data-driven component to predict accurately in this case.

5.1.2.2 Balancing calibration efforts

To investigates the balance between calibration effort for the mechanistic model and acceptable compensation by the neural network component, two additional experiments are executed. In a first experiment, the goal is to examine the difference between training the neural network component of the parallel hybrid model on the calibration dataset of the mechanistic model and on an independent validation dataset. The focus is to understand the difference in information contained within the residuals of both datasets and to determine what the best approach is to create parallel hybrid models. So, a CNN is trained and validated on the residual of the calibration dataset of the mechanistic model and tested on the validation dataset of the mechanistic model, using the same data splitting approach as the previous experiment. Another CNN is trained on the residual of an independent validation dataset. It should be noted that the training period in the second approach is relatively shorter than in the first approach (11 days compared to 18 days) due to the limited length of the validation dataset of the mechanistic model. The validation for the CNN in the second approach is performed on the calibration dataset of the mechanistic model, and the testing period corresponds to the initialization period of the mechanistic model. The details regarding the dataset splitting in this experiment can be found in Figure 5.5. Table 5.2 contains the final set of hyperparameters determined for each neural network used in this experiment.

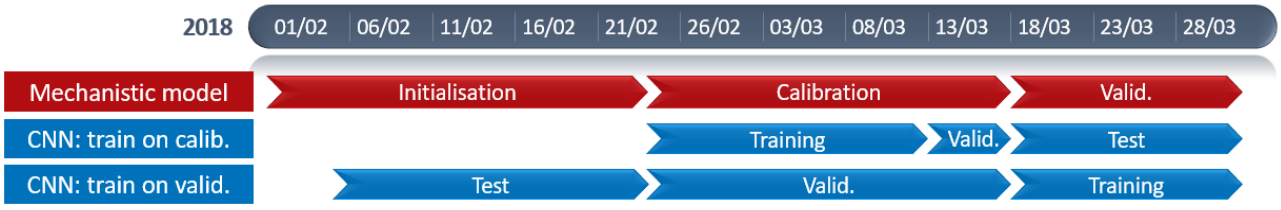

Figure 5.5: Timeline with the time periods used to develop a CNN HM trained on the calibration dataset and a CNN HM trained on a validation dataset.

Figure 5.7 shows the output of the mechanistic model, the measured values, and the output of the CNN hybrid models trained on the calibration dataset and trained on an independent validation dataset of the mechanistic model during the training, validation and test periods. The corresponding RMSE during each of the time period for both approaches are shown in Figure 5.6. The results for the CNN hybrid model output using the first data splitting approach are extensively discussed in section 5.1.2.1. The CNN hybrid model output demonstrated improved accuracy in nitrate levels compared to the mechanistic model output during the training period, as well as improved modelling of

Table 5.2: Final hyperparameter sets for a CNN hybrid model trained on the calibration dataset of the mech. model and a CNN hybrid model trained on the validation dataset of the mech. model for effluent nitrate.

Prediction	Network	Time steps	Structure details	Learning rate Adam optimizer	Epochs
NO3	CNN trained on calib. data	30	Number of layers: 1 Filters: 12 Kernel size: 9 Activation: ReLu	0.00005	200
NO3	CNN trained on valid. data	50	Number of layers: 5 Filter: 25, 12, 9, 6, 3 Kernel size: 9, 5, 3, 3, 3 Activation: ReLu, ReLu, ReLu, ReLu, tanh	0.0001	30

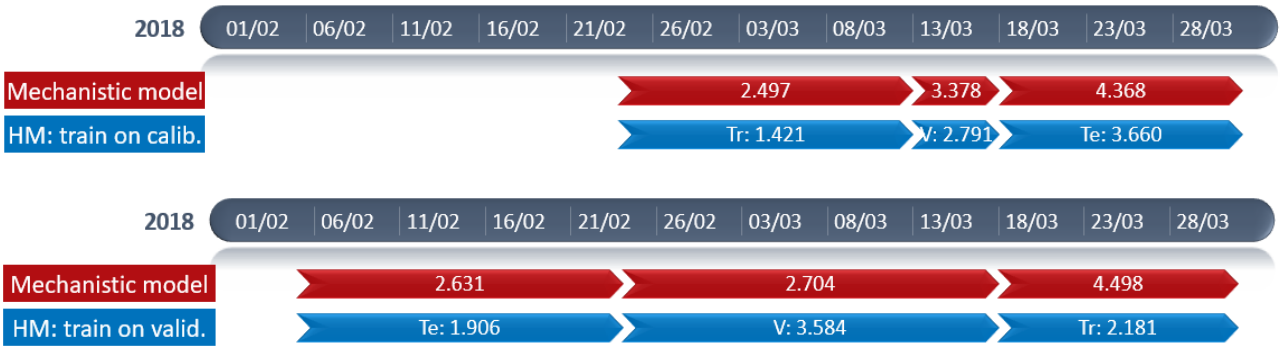

Figure 5.6: Timeline with RMSE per time period (Tr=train, V=validation, Te=test) used to develop a CNN hybrid model trained on the calibration dataset and a CNN hybrid model trained on a validation dataset for effluent nitrate, compared to the RMSE of the mechanistic model in these periods.

the dynamics, mainly in the first half of the training period. However, this performance could not be sufficiently extrapolated to the test period.

In the second approach the CNN hybrid model is trained on the remaining error of the mechanistic model with respect to an independent validation dataset. The hyperparameter tuning process proved to be challenging in this approach as no hyperparameter set was found that resulted in improved predictions during the validation period. Despite attempting various parameter configurations, the validation loss did not show any decrease. Consequently, a parameter set was selected that minimally increased the validation loss while managing to achieve a reasonable reduction in training loss. The CNN hybrid model output shows some improvement in its performance during the training period compared to the mechanistic model output. The CNN hybrid model's output demonstrates a closer alignment in terms of the order of magnitude with the actual measurements, and this improvement is also reflected in the reduction of the RMSE compared to the mechanistic model. During the validation period, the CNN hybrid model

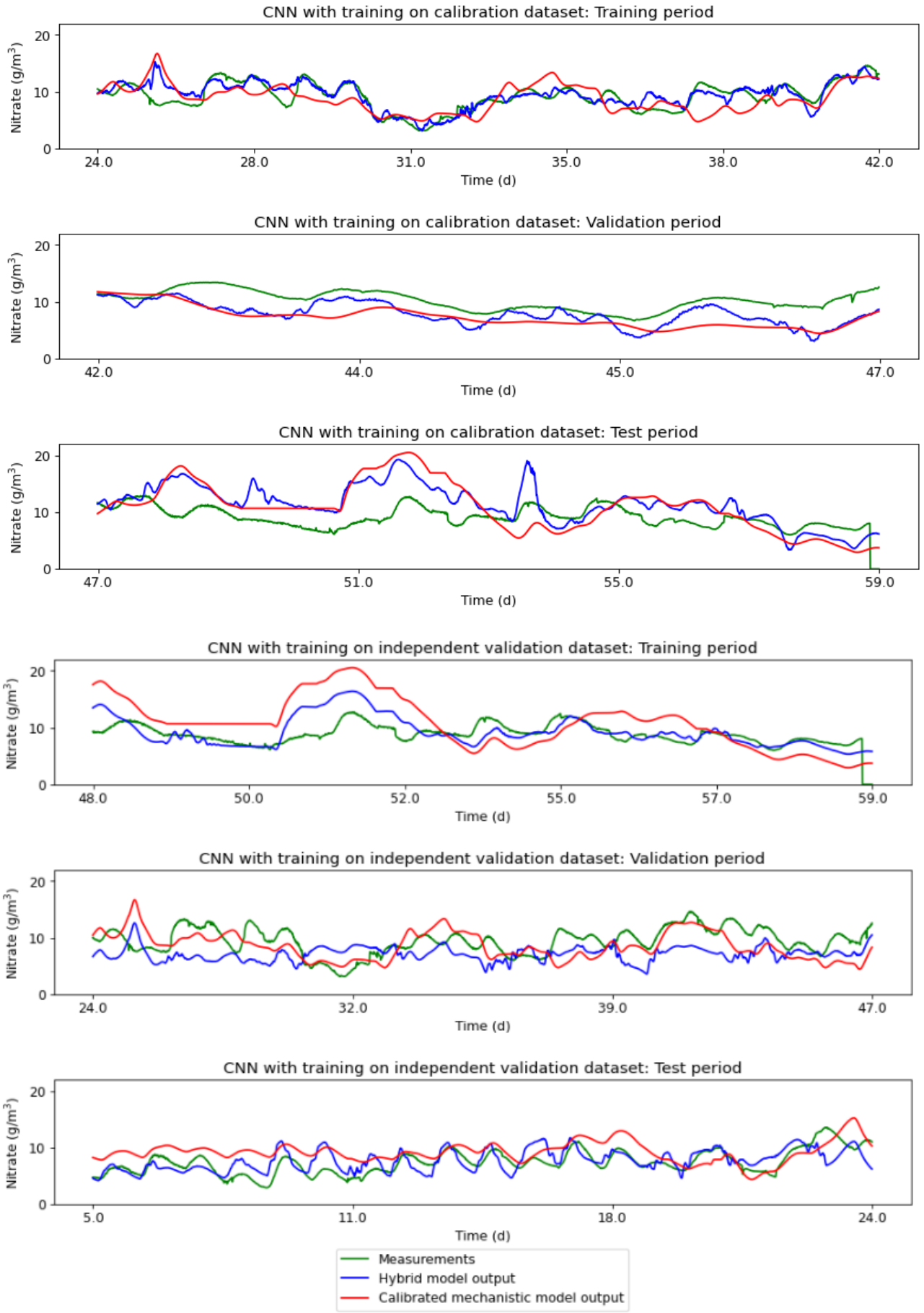

Figure 5.7: CNN hybrid model output trained on a calibration dataset of a mech. model for the train, validation and test period (top) and CNN hybrid model output trained on a validation dataset of a mech. model for the train, validation and test period (bottom) compared to mechanistic model output and measurements for effluent nitrate.

significantly fails to accurately capture any dynamics. Its outputs are notably worse compared to those of the mechanistic model. This failure is evident in an increase in the RMSE in comparison to the RMSE of the mechanistic model. The validation of the CNN is conducted using the calibration period of the mechanistic model. The poor predictions observed in this phase could be attributed to the issue mentioned in the previous section. It is possible that the mechanistic model parameters are overfit, resulting in limited useful information to predict in the residuals during this period. Additionally, some dynamics may already be compensated for through the calibration process. The CNN hybrid model was able to predict effluent nitrate during the test period with a more accurate order of magnitude compared to the mechanistic model output. While the HM shows an improvement in capturing the dynamics compared to the mechanistic model, the predicted dynamics still slightly deviate from those of the measurements during some days, for example day 20-24 of the test period. However, the CNN HM trained on the validation dataset performs significantly better during the test period than the CNN HM trained on the calibration dataset of the mechanistic model.

It is noteworthy that despite the suboptimal optimization of the CNN hyperparameters in the second approach, the test period is still predicted well. Also, despite using a shorter period for training the CNN (11 days in the second approach compared to 18 days in the first approach), the test period is predicted more accurately compared to the previous approach. This improved performance is further supported by comparing the RMSE values obtained in the two approaches during the test periods. By examining the boxplots in Appendix C, it can also be observed that the training dataset is representative of the test dataset in the second approach. Despite occasional lower values and wider ranges of variables in the training data, there is no significant difference between the two datasets. However, it is possible that results may have been influenced by the nature of the data in the second approach. The test data in this approach, i.e. the initialisation period of the mechanistic model, exhibits a daily pattern. The presence of such patterns may lead to relatively better predictability and potentially favorable results during the testing phase.

The findings of this experiment suggest that training a neural network in a parallel hybrid model configuration using a calibration dataset may not yield effective results, as the remaining error during this period contains limited information. Additionally, some dynamics in the calibration dataset may already be compensated for through the calibration process. So, training the data-driven component of a HM solely on the calibration

dataset may not be effective in enabling the network to acquire the essential knowledge required for accurate predictions on unseen datasets. It may be more beneficial to explore alternative approaches or datasets to improve the performance of the neural network in predicting the desired outputs.

To further investigate the balance between calibration effort for the mechanistic model and acceptable compensation by the neural network component a second experiment is executed. In this experiment, the goal is to examine the difference between training the neural network component of the parallel hybrid model on residual of the calibrated mechanistic model and on the residual of the uncalibrated mechanistic model. To accomplish this, the mechanistic model is set to default biokinetic parameters and the backflows included in the hydraulic model are removed. A CNN is then trained and validated on the error between the uncalibrated model and the calibration dataset. Thus in this experiment, the neural network is trained to predict the residual between the output of the mechanistic model with default parameters for effluent nitrate, and the corresponding measurements. The predicted residuals are then added to the output of the uncalibrated mechanistic model. The results are then compared with the output of a hybrid model using a CNN trained on the residual of a calibrated mechanistic model and the calibration dataset. The data partitioning in this experiment is visualised in Figure 5.8. Table 5.3 contains the final set of hyperparameters determined for each neural network used in this experiment.

Figure 5.8: Timeline with the time periods used to develop a CNN hybrid model trained on the calibrated mechanistic model residual and a CNN hybrid model trained on the uncalibrated mechanistic model residual.

As mentioned in the previous sections, it is likely that the mechanistic model is exhibiting signs of overfitting to the data. Potential calibration issues can be observed in two different submodels for the work of Kirim (2022): the hydraulic and the biokinetic model. The hydraulic model was established using the data obtained from two tracer tests, which led to the determination of various backflows between the basins that were incorporated into the final model configuration (Section 4.2). The results of the hydraulic model can be found in Appendix A. The results demonstrate that the hydraulic model prediction is

Table 5.3: Final hyperparameter sets for CNN hybrid model trained on the calibrated mechanistic model residual and a CNN hybrid model trained on the uncalibrated mechanistic model residual for effluent nitrate.

Prediction	**Network**	**Time steps**	**Structure details**	**Learning rate Adam optimizer**	**Epochs**
NO3	CNN trained on calib. model	30	Number of layers: 1 Filters: 12 Kernel size: 9 Activation: ReLu	0.00005	200
NO3	CNN trained on uncalib.model	30	Number of layers: 4 Filter: 10, 5, 5, 2 Kernel size: 5, 3, 3, 3 Activation: ReLu, tanh, ReLu, ReLu	0.0002	70

still suboptimal. In the first experiment, the model consistently overestimates the observed peak in each tank. In the second experiment, the shape of the peak deviates from the actual peak in each tank. A suboptimal hydraulic will influence the subsequent calibration of biokinetic model parameters as they may be used to compensate for missing dynamics in the hydraulic model. The calibrated parameters of the biokinetic model can be found in Table 4.2. It can be observed that the calibrated values for the model parameters *K_HNO2_NO* and *K_NH3_NH* related to the biokinetics of the 2-step nitrification process are significantly lower than their default values (factor 10-100 difference). The same situation holds for the oxygen halfsaturation concentrations (*K_O_NH* and *K_O_NO*) (factor 3-5 difference). This could be an indication that the biokinetic model parameters are calibrated to compensate for other missing phenomena. However, lumping different phenomena into biokinetic parameters will negatively impact the model's predictive power.

To reduce the influence of the potential overfit parameters, all the parameter values are reset to the default values in Table 4.2. Also the backflows are removed from the hydraulic model structure. The model is re-executed, generating new mechanistic model predictions for effluent nitrate (Figure 5.9). The removal of the backflows had no impact on the model predictions, however, resetting the parameters to their default values did result in a significant difference in the output of the mechanistic model. The uncalibrated model clearly performs worse than the calibrated one as expected.

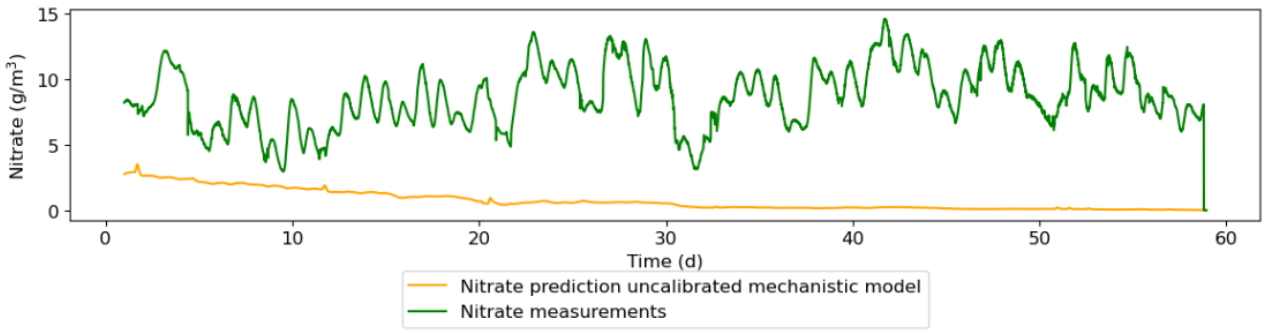

Figure 5.9: Uncalibrated mechanistic model output for effluent nitrate compared to effluent nitrate measurements for the entire modelling period.

The results for this experiment are shown in Figure 5.11 and the corresponding RMSE values can be found in Figure 5.10. The results for the CNN hybrid model output trained on the residual of the calibrated mechanistic model output are extensively discussed in section 5.1.2.1. The results obtained for the CNN hybrid model trained on the residual of the uncalibrated mechanistic model and the calibration dataset, demonstrate improved accuracy and prediction of dynamics during the training period compared to the mechanistic model. While occasionally slightly deviating from the desired dynamics, the overall performance remains satisfactory. During the validation period, the outputs of the HM exhibit closer agreement with the measurements compared to those of the mechanistic model and the HM in the first approach. However, the true dynamics are not fully captured by the HM in this period. The output of the CNN hybrid model during the test period demonstrates superior performance compared to the mechanistic model. The HM's predictions are also significantly better than those of the HM trained on the output of the calibrated mechanistic model. Particularly during three notable peaks in the predictions of the mechanistic model on days 48, 51-53, and 55-57, the HM trained on the uncalibrated mechanistic model successfully improves the accuracy of the predictions, whereas the HM trained on the calibrated model fails to compensate for the mistakes of the mechanistic model. These observations are further supported by the RMSE values, which are the lowest for the three period for the CNN hybrid model trained on the uncalibrated mechanistic model output. However, the difference in RMSE values becomes particularly pronounced during the test period, where the HM trained on the uncalibrated mechanistic model output demonstrates a significantly lower RMSE value compared to mechanistic model and the HM trained on the calibrated mechanistic model output.

To investigate if the information that the hybrid model gains out of the uncalibrated mechanistic model is still significant, a standalone data-driven model is set-up to pre-

Figure 5.10: Timeline with RMSE per time period (Tr=train, V=validation, Te=test) used to develop a CNN hybrid model trained on the calibrated mechanistic model residual and a CNN hybrid model trained on the uncalibrated mechanistic model residual for effluent nitrate, compared to the RMSE of the mechanistic model in these periods.

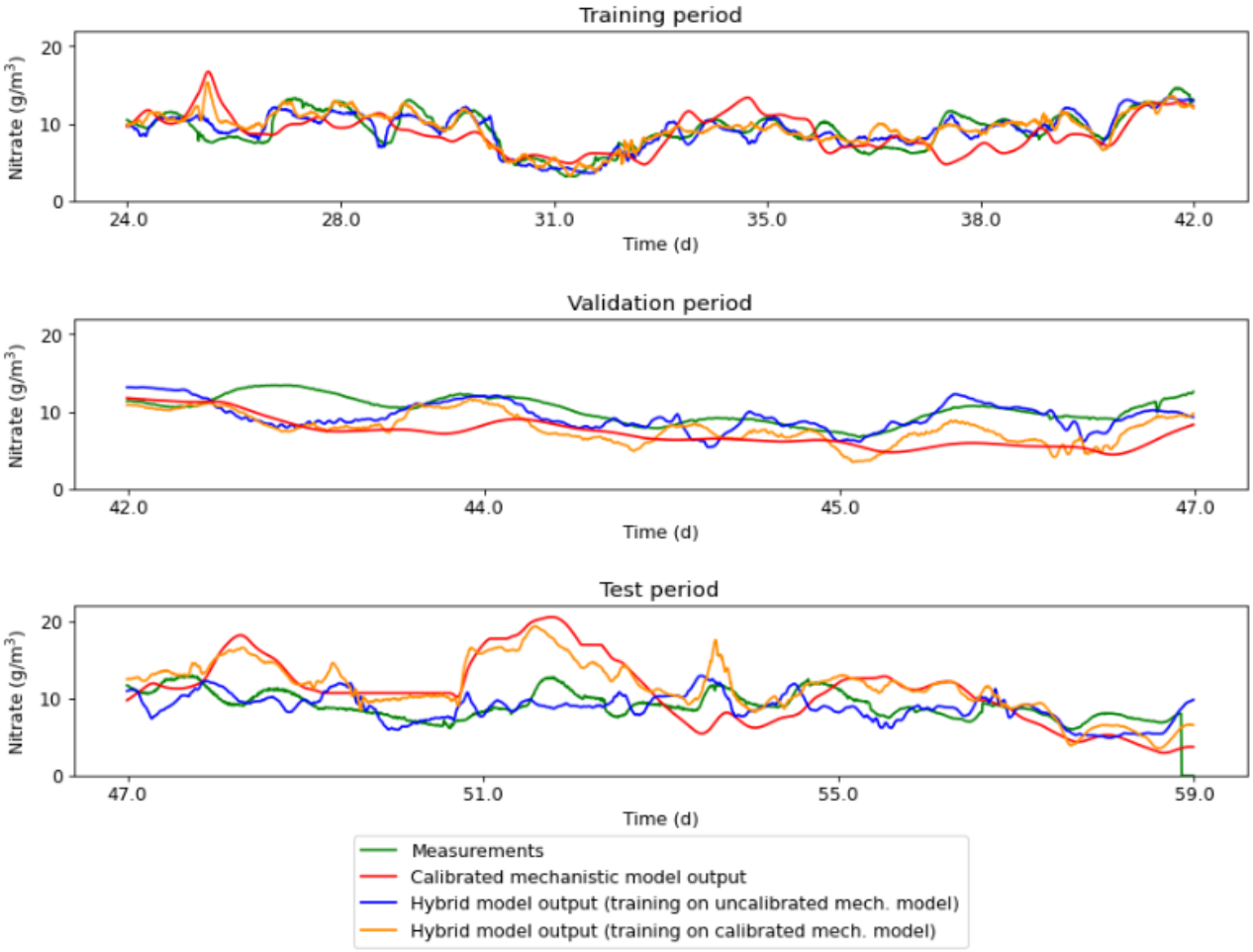

Figure 5.11: CNN hybrid model output trained on the residual of a calibrated mechanistic model and CNN hybrid model output trained on the residual of an uncalibrated mechanistic model for the train, validation and test period compared to mechanistic model output and measurements for effluent nitrate.

dict the effluent nitrate in the pilEAUte. Therefore, a CNN is trained using the effluent nitrate measurements as the target variable and produces predictions of effluent nitrate as its output. The CNN has the same input data as the data-driven component in the HM, except for the mechanistic model output for nitrate. The data splitting, the final hyperparameter set and the results of the data-driven model can be found in appendix D. The results of the data-driven model show improvements compared to the mechanistic model in the training, validation, and testing periods. However, the dynamics are rarely accurately predicted in the three periods. The hybrid model trained on the uncalibrated mechanistic model performs better than the data-driven model, particularly in the training period. However, the difference in performance between this HM and the data-driven model is not substantial, as expected. These observations are further supported by the lower RMSE values of the hybrid model in all three periods, with notably superior performance during the training period. So in this study, a data-driven model can make more accurate predictions than a calibrated mechanistic model. However, the combination of a data-driven model with an uncalibrated mechanistic model, which itself provides poor nitrate predictions, yields even better results. This suggests that the mechanistic model still offers valuable information to the hybrid model, although to a lesser extent than the data-driven model. The advantage obtained by the parallel hybrid model is thus mainly derived from the data-driven component, but the contribution of the uncalibrated mechanistic model should not be disregarded.

In this section the balance between calibration effort for the mechanistic model and acceptable compensation by the neural network component is investigated by conducting two experiments. The first experiment involved training a neural network component of a hybrid model using both a calibration dataset and an independent validation dataset from the mechanistic model. A comparison of the results revealed that the HM performed better when trained on the independent validation dataset. This could be attributed to the fact that the residual information in the calibration dataset is limited due to the calibration process, and that certain dynamics or knowledge may already be accounted for during the calibration process itself. When a neural network is trained solely on the calibration data, it may not learn the necessary patterns and relationships to make accurate predictions outside the calibration period. In the second experiment, a neural network is trained on the residual of both a calibrated and an uncalibrated mechanistic model. The results of this experiment demonstrate that training the NN on the residual of the uncalibrated mechanistic model yields better performance, particularly during the test

period. This suggests that the residual of the uncalibrated mechanistic model contains valuable information and patterns that are not present in the residual of the calibrated model. By incorporating this residual information into the training process, the NN is able to improve its predictive capabilities, resulting in superior performance during the test period. The results of these experiment thus support the hypothesis that the mechanistic model may have been slightly overfitted. This is always a risk when some missing dynamics are present in a mechanistic model. Detailed calibration of the mechanistic model with several unidentifiable parameters may lead to overfitting which makes it more difficult for a parallel neural network component to learn the missing model dynamics. The present work shows that a parallel hybrid model approach can compensate for non calibrated phenomena. This could potentially save time in the calibration effort of mechanistic models as well.

To reaffirm the findings of Section 5.1.2.1, where the performance of a CNN hybrid model is compared with an LSTM-RNN hybrid model, an additional hybrid model is developed with an LSTM-RNN as the data-driven component. In this experiment, the LSTM-RNN is trained on the residual of the uncalibrated mechanistic model. This experiment aims to confirm that the inferior performance of the LSTM HM, as compared to the CNN HM, can be attributed to the nature of the data-driven model component in the hybrid model, rather than the specific characteristics of the calibrated mechanistic model. The data splitting, the final hyperparameter set and the results of this experiment can be found in Appendix E. During the training period, the output of the LSTM hybrid model is significantly more accurate than the output of the mechanistic model. The LSTM hybrid model consistently captures the correct dynamics during this period. However, it is noteworthy that the correct dynamics are absent in the outputs of the LSTM hybrid model during the validation and test periods. The LSTM hybrid model predicts several peaks that do not occur in the measurement data during these periods, for example on day 46-47 in the validation period and day 53 and 59 in the test period. This reaffirms the findings from Section 5.1.2.1, where the LSTM hybrid model also exhibited poorer performance compared to the CNN hybrid model.

5.1.3 Parallel hybrid modelling for improved effluent TSS predictions

Parallel hybrid models are also developed for effluent TSS prediction, using the same model set up as described in Section 5.1.1. Given that the effluent TSS was not utilized in the previous work conducted by Kirim (2022), the preprocessing procedure outlined

Table 5.4: Final hyperparameter set for a CNN hybrid model for effluent TSS trained on the calibration dataset of the mech. model.

Prediction	Network	Time steps	Structure details	Learning rate Adam optimizer	Epochs
TSS	CNN	30	Number of layers: 4 Filter: 30, 9, 6, 3 Kernel size: 9, 3, 3, 3 Activation: ReLu, ReLu, ReLu, tanh	0.0001	30

in Section 4.2 was implemented on the effluent TSS data in the present study. Both an LSTM-RNN and a CNN architecture were established to predict the residuals between the effluent TSS output generated by the mechanistic model and the measured effluent TSS. Subsequently, the predicted residuals were added to the mechanistic effluent TSS outputs. Both the hybrid models are trained and validated on the calibration dataset of the mechanistic model and tested on the validation dataset of the mechanistic model. The data splitting technique is visualised in Figure 5.12. Since both hybrid models produced very similar outputs, the results corresponding to the hybrid model with LSTM RNN can be found in Appendix F. Table 5.4 contains the final set of hyperparameters determined for the CNN.

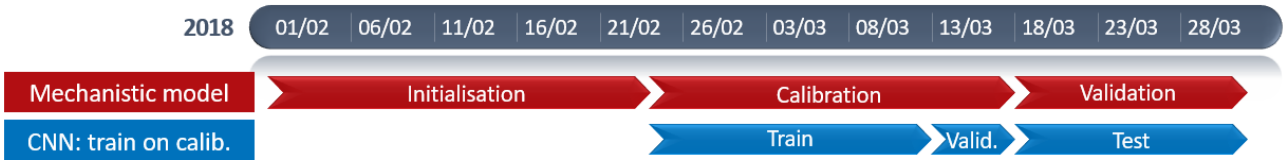

Figure 5.12: Timeline with the time periods used to develop a CNN hybrid model for effluent TSS.

In Figure 5.13 result of the CNN hybrid model for effluent TSS are shown for the training, validation and testing period. As TSS measurements are not directly available in the effluent, they are derived through a correlation with the turbidity data in the effluent. The TSS output of the mechanistic model significantly deviates from the measurements in all three periods. Only the global trend is predicted reasonably well; however, there is a significant discrepancy in the magnitude of the predictions. A possible explanation for this deviation is that the mechanistic model was not specifically calibrated on the effluent TSS data (Kirim, 2022). The CNN hybrid model demonstrates consistent performance across the training, validation, and test datasets. The HM improves the prediction compared to the mechanistic model by capturing an order of magnitude closer to the measurements; however, the dynamics are still not correctly captured. Due to the closer alignment between the magnitude of the HM outputs for effluent TSS and the measure-

ments, the RMSE has significantly improved in both the training, validation, and testing periods (Figure 5.14). The same results are observed for the LSTM hybrid model outputs (Appendix F). The primary factor contributing to the absence of accurate dynamics in the CNN and LSTM hybrid model output is most likely attributed to the TSS measurements. It is known that the effluent TSS measurements of the pilEAUte may not be highly accurate. For instance, the drop to zero during days 30-32 in the training period does not reflect behaviour that could be expected in reality. There are also sometimes a rapid decline or increase in the data, such as on day 36 of the train period, which would not occur in reality. The presence of unreliable data in the train period poses a significant challenge in effectively training a neural network to accurately learn and generalize from the training data. This leads to difficulties in accurately predicting in the test period.

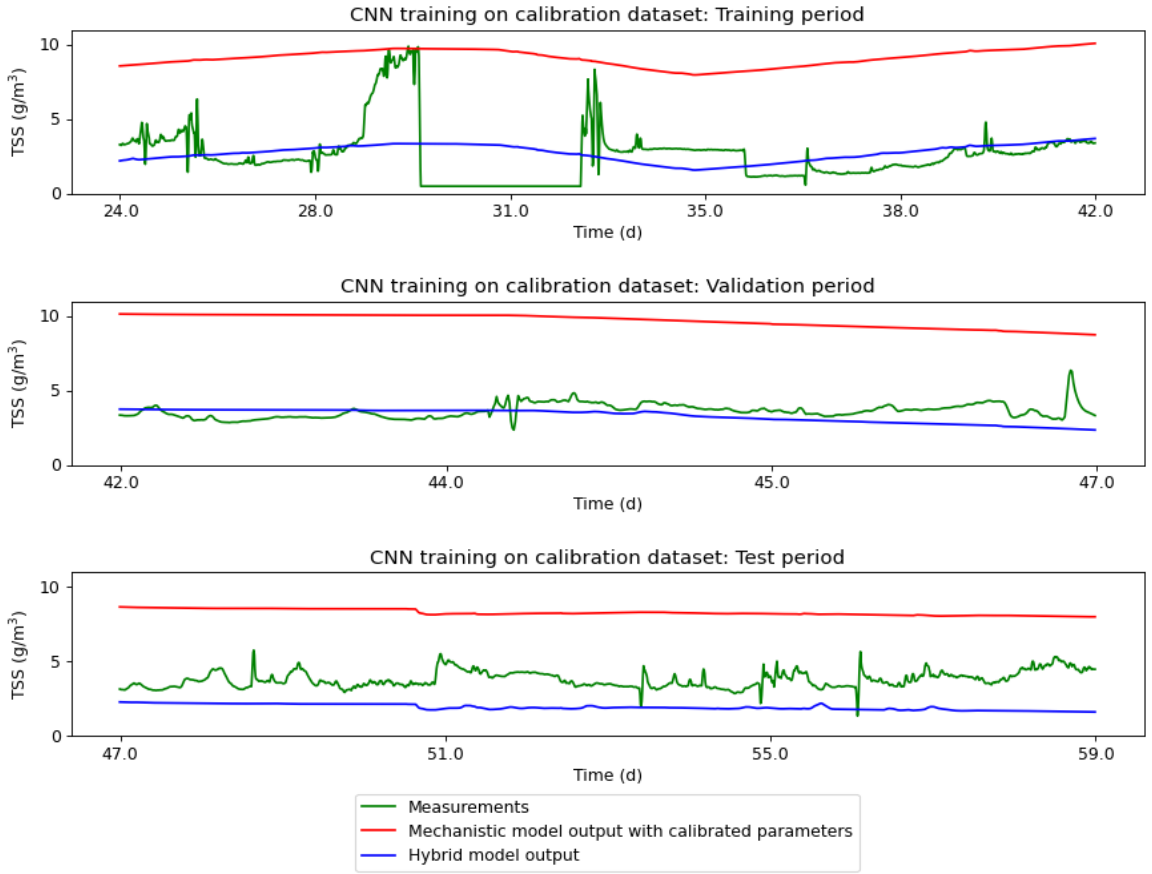

Figure 5.13: CNN hybrid model output for the train, validation and test period compared to mechanistic model output and measurements for effluent TSS. The CNN is trained on the calibration dataset of the mech. model.

Table 5.5: Final hyperparameter set for a CNN hybrid model for effluent TSS trained on the validation dataset of the mech. model.

Prediction	Network	Time steps	Structure details	Learning rate Adam optimizer	Epochs
TSS	CNN trained on valid. data	30	Number of layers: 4 Filter: 12, 9, 6, 3 Kernel size: 6, 5, 3, 3 Activation: ReLu, ReLu, ReLu, ReLu	0.0002	70

Figure 5.14: Timeline with RMSE per time period (Tr=train, V=validation, Te=test) used to develop a CNN hybrid model for effluent TSS, compared to the RMSE of the mechanistic model in these periods.

To investigate whether a parallel hybrid model can predict effluent TSS dynamics effectively when there are no significant errors in the measurements during the training period, a CNN hybrid model is developed in which the CNN component is trained on the residual of the validation dataset of the mechanistic model. The CNN component is validated using the calibration dataset and tested on the initialisation dataset of the mechanistic model. The time periods used in this experiment are visualized in Figure 5.15. The final set of hyperparameters for the CNN is outlined in Table 5.5.

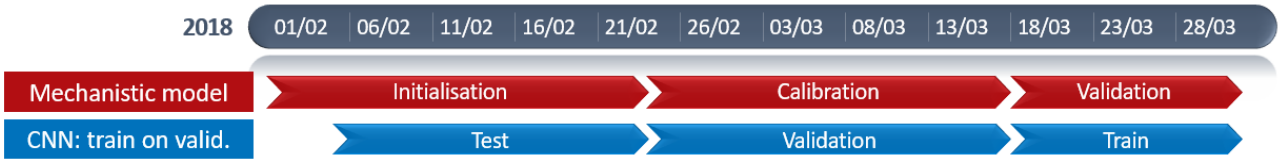

Figure 5.15: Timeline with the time periods used to develop a CNN hybrid model for effluent TSS.

In this experiment, during the training period, the CNN hybrid model demonstrates improved accuracy in capturing the correct magnitude compared to the mechanistic model. Moreover, the output of the hybrid model exhibits increased dynamics compared to the previous experiment where the CNN was trained on a dataset with sensor errors. In specific segments of the training period, such as during day 50-51 and 57-59, the dynamics are even captured with remarkable precision. In the validation period, the hybrid model also shows more dynamic behavior in the data compared to the previous experiment. However, the dynamics of the data are not accurately predicted. This is expected as this

period contains too many sensor errors to make reliable predictions. During the testing period, the predictions of the hybrid model align better with the order of magnitude of the measurements, compared to the mechanistic model. Again, the hybrid model's output displays more dynamics, although they are not captured accurately.

Overall, it can be concluded that neither the mechanistic model nor the hybrid models can accurately predict effluent TSS, even when trained on a dataset with limited sensor errors. The hybrid models can estimate the approximate magnitude better, but fail to capture the correct dynamics. This could be attributed to several factors. A possible reason could be that the neural network fails to detect patterns in the input variables that could help increase the accuracy of the TSS predictions. It is possible that the hybrid model predictions could be improved by providing additional input variables to the neural network component. These could include supplementary measurements at the end of the aeration zone, for instance. It is also possible that the duration of the training period is insufficient for the model to acquire an adequate level of knowledge to make accurate predictions. A longer training period could potentially provide the neural network with the opportunity to learn sufficient patterns in the input data and make more accurate predictions. However, due to the relatively stable nature of the data over time, extending the training period may not lead to any significant improvements.

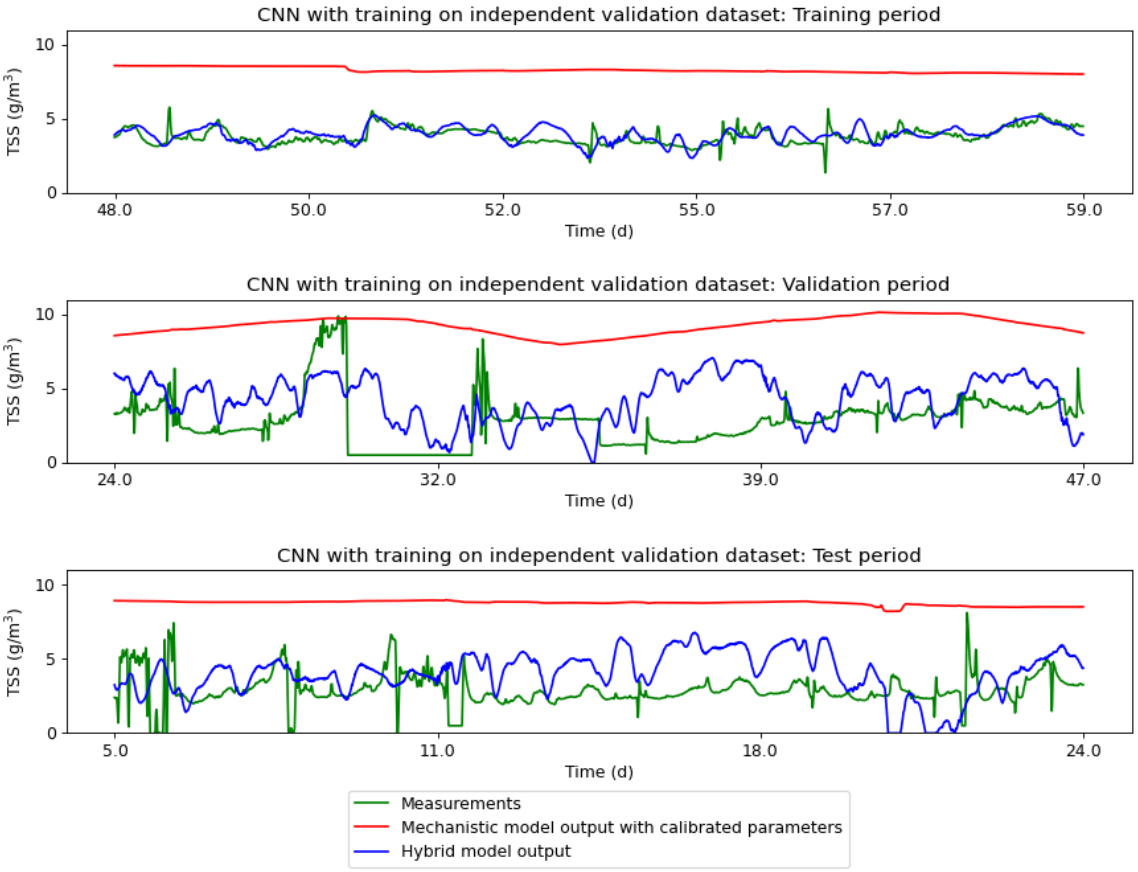

Figure 5.16: CNN hybrid model output for the train, validation and test period compared to mechanistic model output and measurements for effluent TSS. The CNN is trained on the validation dataset of the mech. model.

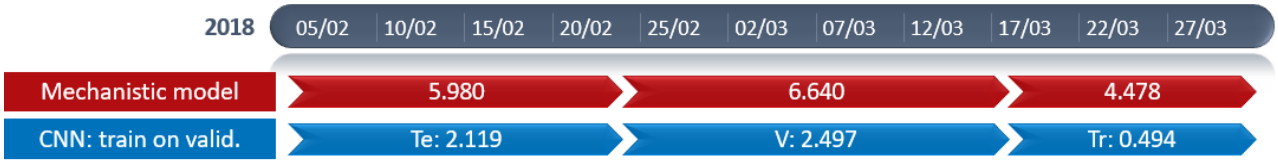

Figure 5.17: Timeline with RMSE per time period (Tr=train, V=validation, Te=test) used to develop a CNN hybrid model for effluent TSS, compared to the RMSE of the mechanistic model in these periods.

An additional experiment was conducted where a CNN was trained on the residual of an uncalibrated model and the calibration period. The data splitting technique, the final hyperparameter set, the output of this hybrid model, and the calculated RMSE values, as well as the output of the uncalibrated model for effluent TSS, can be found in Appendix F. The output of this hybrid model exhibits similar behavior to the experiment where a CNN hybrid model was trained on the calibrated model and the calibration period. This provides further evidence that the poor hybrid model performance is caused by sensor errors in the training period.

It is important to acknowledge that parallel hybrid models may not be applicable for predicting all variables in wastewater treatment processes. The potential improvement that a parallel hybrid model can offer depends on several factors. Firstly, it relies on the variability present in the data. The effectiveness of a parallel hybrid model heavily depends on the availability of diverse and informative data patterns. When the data lacks significant variability or exhibits homogeneity, the data-driven component of the hybrid model may face challenges in capturing relevant information from the residuals. Consequently, extracting meaningful insights and improving predictions beyond what the mechanistic model already provides can become difficult. The improvement brought by a parallel hybrid model are also influenced by the mechanistic model structure and its ability to accurately represent the physical phenomena. If the mechanistic model wrongly captures the underlying processes, the potential advantages of a parallel hybrid model may be constrained. The quality of the data used for modelling is also a crucial factor in assessing the potential improvement that a parallel hybrid model can offer. The reliability, accuracy, and representativeness of the data directly impact the model's performance and its ability to capture meaningful patterns and relationships. These considerations highlight the importance of carefully assessing the suitability and potential advantages of parallel hybrid models for specific variables within wastewater treatment processes.

5.2 Serial hybrid modelling

In this second part of the research, the potential of serial hybrid modelling for wastewater treatment systems is explored. In a serial hybrid model a data-driven model is developed of a subprocess that is less well-described and subsequently integrated in the overall mechanistic model. Aeration is one of the highest costs in operating a treatment plant and aeration models are known to have high uncertainty. Therefore, in this section a serial model for aeration in wastewater treatment is developed.

In this respect, the determination of the oxygen transfer coefficient (K_{La}) is important in wastewater treatment because it quantifies the efficiency of oxygen transfer from the gas phase to the liquid phase. It is a critical parameter in aeration systems. The K_{La} is not a constant but varies over time, and its value depends on several variables. These variables include operational parameters such as airflow rate, mixed liquor suspended solids (MLSS), sludge age, and viscosity and design parameters such as the type of aeration system and the configuration of tanks (Rosso et al., 2005). The K_{La} can be used

to calculate how much air flow rate is needed to obtain a certain DO level in the system or inversely what the DO concentration will be for a certain fixed air flow rate. Accurate knowledge of K_{La} allows optimisation of aeration system design, operational conditions, and process control strategies to achieve desired treatment goals. Traditionally, the off-gas method has been employed to determine oxygen transfer coefficient, as described by Redmon et al. (1983). Although entirely accurate, this method is expensive to adapt during operation as it requires additional sensors to measure the volumetric fraction of the outlet gas fraction. Besides, the K_{La} calculated with this method may not be representative for all regions in the aeration tanks (Pan and Dagnew, 2022). In this context, the prediction of K_{La} using a data-driven model emerges as a promising approach, offering valuable insights for informed decision-making and comprehensive analysis across various applications.

5.2.1 Model set-up

In this part of the research, a CNN is developed to predict the oxygen transfer coefficient. However, in order to train, validate, and evaluate the performance of the prediction model, it is essential to have prior knowledge of the K_{La} values. These known K_{La} values serve as the reference or ground truth against which the predictions made by the CNN model are compared. This allows for the assessment of the accuracy and reliability of the model's predictions in estimating K_{La} values. Direct measurements of K_{LA} could be obtained through off-gas analysis as explained above. However, in the mechanistic model of the pilEAUte, no such off-gas measurements were available. In the pilEAUte plant, the DO concentration is controlled at a fixed setpoint of 3 mg/L. Therefore, the real K_{LA} can be estimated in the model by using a controller that regulates the K_{La} in Basin 4, to obtain a DO setpoint of 3 mg/L. Subsequently, ratio controllers are used to adjust the aeration in Basins 3 and 5 based on the K_{La} in Basin 4 (Section 4.2). The K_{La} can also be calculated from the oxygen mass balance in the tank:

$$\frac{dS_O}{dt} = Q_{in}S_{O,in} - Q_{out}S_O + K_{La}(S^*_O - S_O) - OUR \tag{5.1}$$

where S_O is the DO concentration in the tank, S^*_O is the saturation DO concentration, $S_{O,in}$ is the DO concentration entering the tank, Q_{in} is the flow rate of the liquid entering the tank, Q_{out} is the flow rate of the liquid leaving the tank and OUR is the oxygen uptake

rate in the tank. Equation 5.1 can be transformed to calculate K_{La} in a tank (Equation 5.2).

$$K_{La} = \frac{\frac{dS_O}{dt} + OUR - Q_{in}S_{O,in} + Q_{out}S_O}{S^*_O - S_O} \tag{5.2}$$

The oxygen transfer coefficients in basin 4 of the pilEAUte, obtained using the two described methods, can be found in Figure 5.18. The components of the mass balance were extracted from the output of the mechanistic model. The derivative of the DO concentration is calculated by applying Equation 5.3. The extracted components for the mass balance underwent a resampling process using a one-dimensional linear interpolation function, implemented through the utilization of the "interp" function from the NumPy package in Python. The resampling transformed the data from a 1-minute interval to a 10-minute interval. Additionally, the K_{La} output directly derived from the mechanistic model was downsampled using the same interpolation technique.

$$\frac{dS_O}{dt} \approx \frac{S_O(t) - S_O(t-1)}{t - (t-1)} \tag{5.3}$$

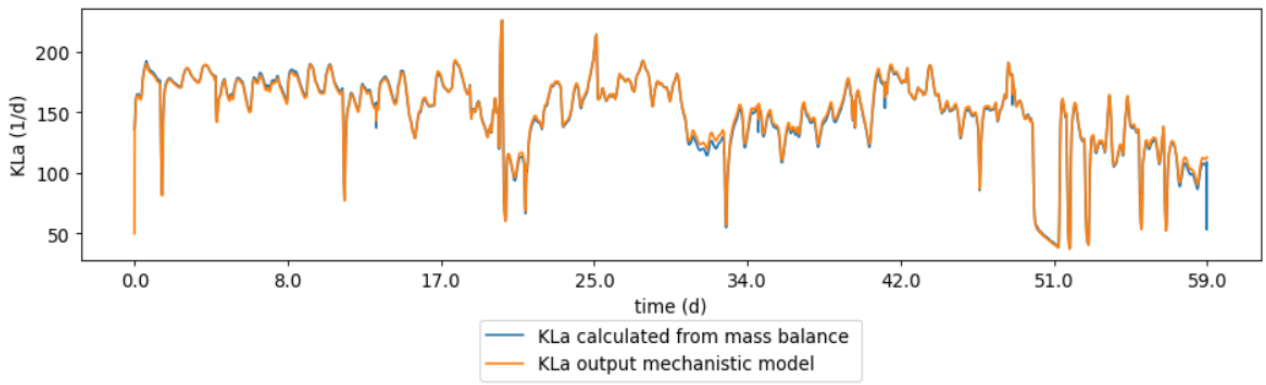

Figure 5.18: Oxygen transfer coefficient calculated from the DO mass balance (Equation 5.2) and oxygen transfer coefficient predicted by mechanistic model based on DO control during the data time period used to set-up the mechanistic model.

Given the high degree of agreement between the K_{La} calculated from the DO mass balance and the K_{La} output obtained from the mechanistic model, the choice between them as the "ground truth" for training, validation, and evaluation of the CNN becomes insignificant. So, in this context, the ground truth is represented by the K_{La} output estimated through the controller in the mechanistic model. It should be noted that both approaches to calculate the true K_{La} are imperfect since both are based on the DO values of the

mechanistic model that assumes an (almost) perfectly controlled DO at a setpoint of 3 mg/L. In reality, the DO concentration is not always perfectly controlled as can be seen in Figure 5.19.

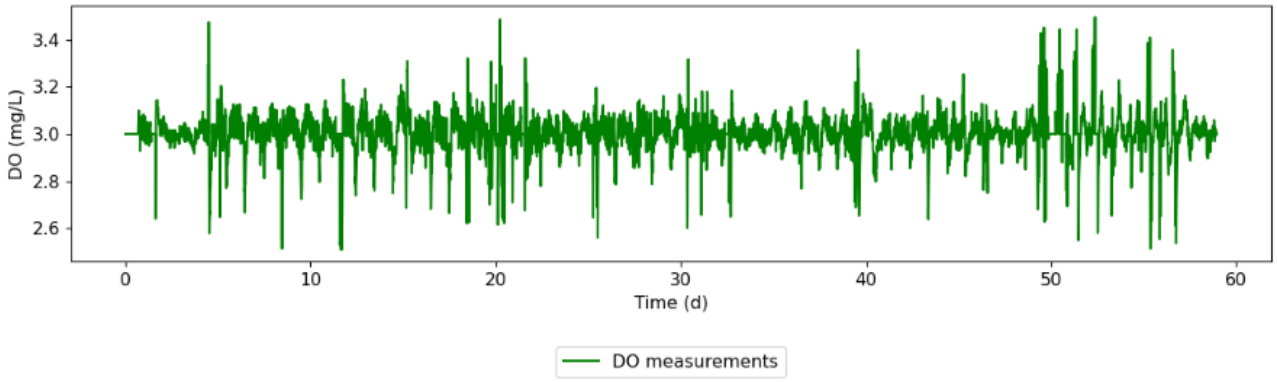

Figure 5.19: DO measurements in the pilEAUte plant during the whole modelling period.

The training data for the CNN includes the measured air flow rate in basin 4, as well as the measured, non-fractionated variables that are also fed to the mechanistic model: temperature in basin 4, influent CODt, CODs, and NH4-N. The input data also incorporates measurements of TSS and pH in the influent. The Oxygen Uptake Rate (OUR) in basin 4, as calculated by the mechanistic model in the WEST, is incorporated as an additional input variable to train the neural network. Note that influent flow rate was not used as an input in the CNN since it is constant over the data period. The CNN is thus used to predict the K_{La} based on the provided input data. The DO controller for basin 4 in the mechanistic model is then replaced with an input file containing the newly predicted K_{La} values. The predicted K_{La} values are also used as inputs for the ratio controllers that govern the K_{La} in basins 3 and 5. A diagram of the serial hybrid model can be found in Figure 5.20.

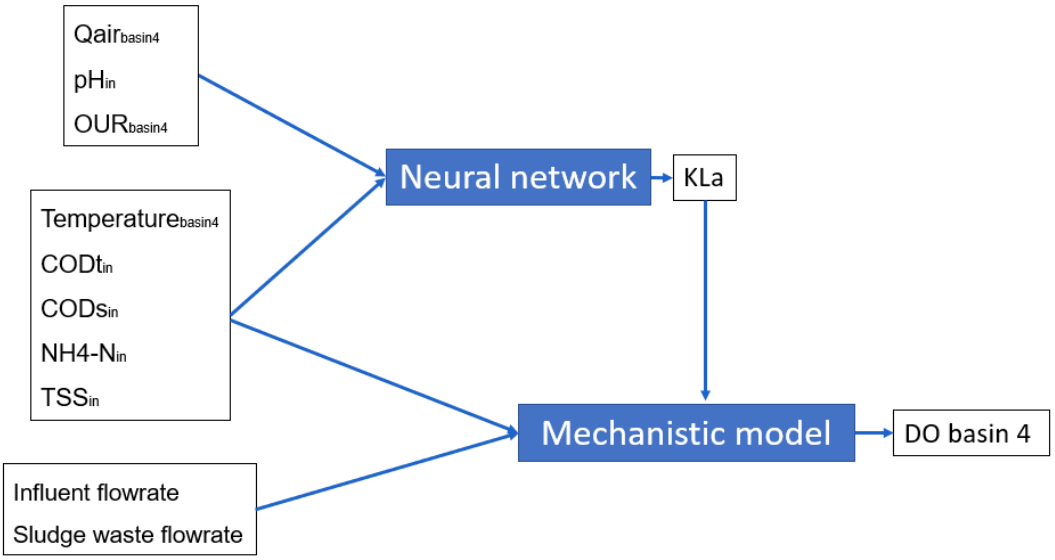

Figure 5.20: Overview of the serial hybrid model.

In this part of the study, again two different data splitting techniques are used, and their prediction performances are compared. In the first approach, the CNN is trained and validated using input data during the calibration period of the mechanistic model. Subsequently, the CNN is tested on the validation dataset of the mechanistic model. The data splitting periods for this approach are visualized in Figure 5.21. In the second approach, the CNN is trained using input data from a time period used to validate the mechanistic model. The CNN is then validated (i.e., hyperparameter-tuned) using the calibration dataset and tested on the initialisation period of the mechanistic model. The data splitting periods for this experiment are visualized in Figure 5.22. The final hyperparameter sets used in the two approached can be found in Table 5.6.

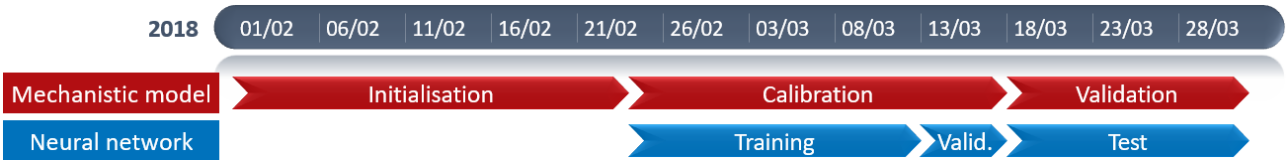

Figure 5.21: Timeline with the time periods used to develop a serial hybrid model in the first approach.

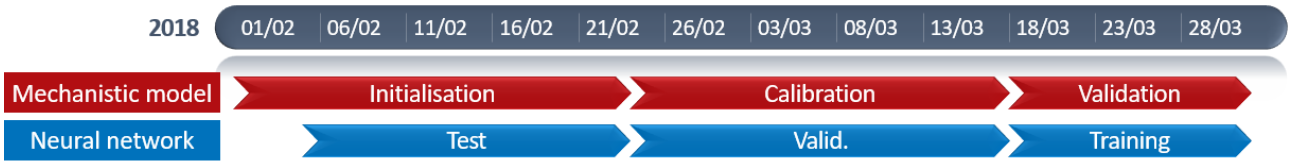

Figure 5.22: Timeline with the time periods used to develop a serial hybrid model in the second approach.

Table 5.6: Final hyperparameter sets determined for the CNN component of the serial hybrid models.

Prediction	Network	Time steps	Structure details	Learning Rate Adam optimizer	Epochs
KLa	CNN trained on calib. dataset	30	Number of layers: 2 Filters: 10, 5 Kernel size: 5, 3 Activation: ReLu, tanh	0.0005	300
KLa	CNN trained on valid. dataset	15	Number of layers: 1 Filter: 10 Kernel size: 5 Activation: ReLu	0.0005	550

5.2.2 KLa predictions

In this approach the CNN is trained on the calibration dataset of the mechanistic model. The K_{La} predictions and corresponding DO output of the mechanistic model for the training can be found in Figure 5.23. The CNN performed remarkably well in computing the K_{La} during the training phase. The predicted values closely aligned with the K_{La} output of the mechanistic model, indicating that the CNN has successfully learned the patterns and relationships within the training data. However, the CNN underestimates the K_{La} peak around days 32-33, which clearly leads to a corresponding peak in the DO. When the oxygen transfer coefficient is high, it can result in an elevated DO level. This is because a higher K_{La} facilitates a more efficient transfer of oxygen from the gas phase to the liquid phase, leading to a greater influx of oxygen into the system. As a result, the DO concentration increases. The accuracy of DO predictions is throughout the entire period strongly influenced by variations in the K_{La}, highlighting its direct impact on the efficiency of oxygen transfer in the aerated basin. As the oxygen transfer process is crucial for maintaining appropriate DO levels, even small deviations or inaccuracies in estimated K_{La} values can have a substantial effect on the DO predictions generated by the mechanistic model. This emphasizes the need for precise and reliable estimation methods for determining K_{La}, as any uncertainties or errors in its assessment directly influence the overall accuracy of the DO modelling process.

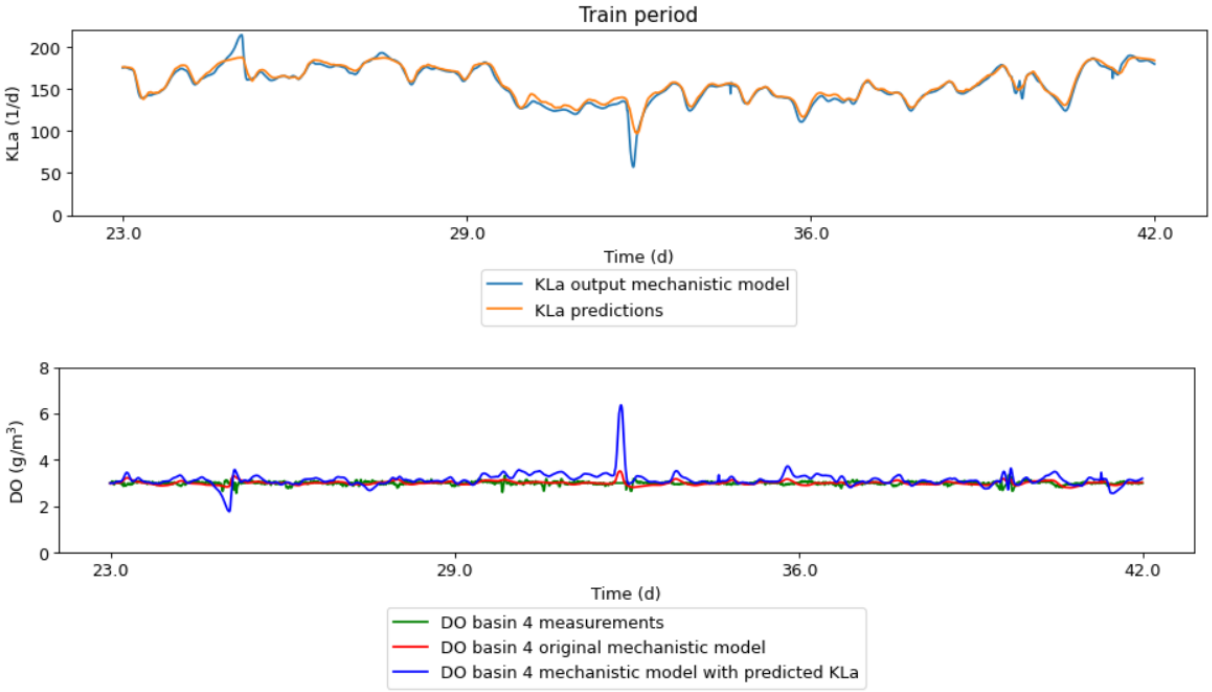

Figure 5.23: Serial hybrid model results during the training period for the first approach where the CNN component is trained on the calibration dataset of the mech. model. The K_{La} output of the CNN compared to the K_{La} output of the mech. model (top) and DO output of the mech. model with the predicted K_{La} compared to the DO output of the original mech. model and DO measurements (bottom).

The hyperparameters of the CNN are fine-tuned by considering the validation period. The results of the K_{La} prediction and the corresponding DO output from the mechanistic model during the validation period can be observed in Figure 5.24. The K_{La} predictions again closely align with the K_{La} output from the mechanistic model, but during the peak on day 46, the model again fails to accurately predict it. Considering the limitations in accurately predicting the peak during the training period, it is reasonable to expect comparable difficulties in forecasting it during following periods. While the DO output demonstrates a moderate level of correspondence with the measurements, it does exhibit noticeable deviations, particularly on day 46, where the CNN underestimates the K_{La} peak.

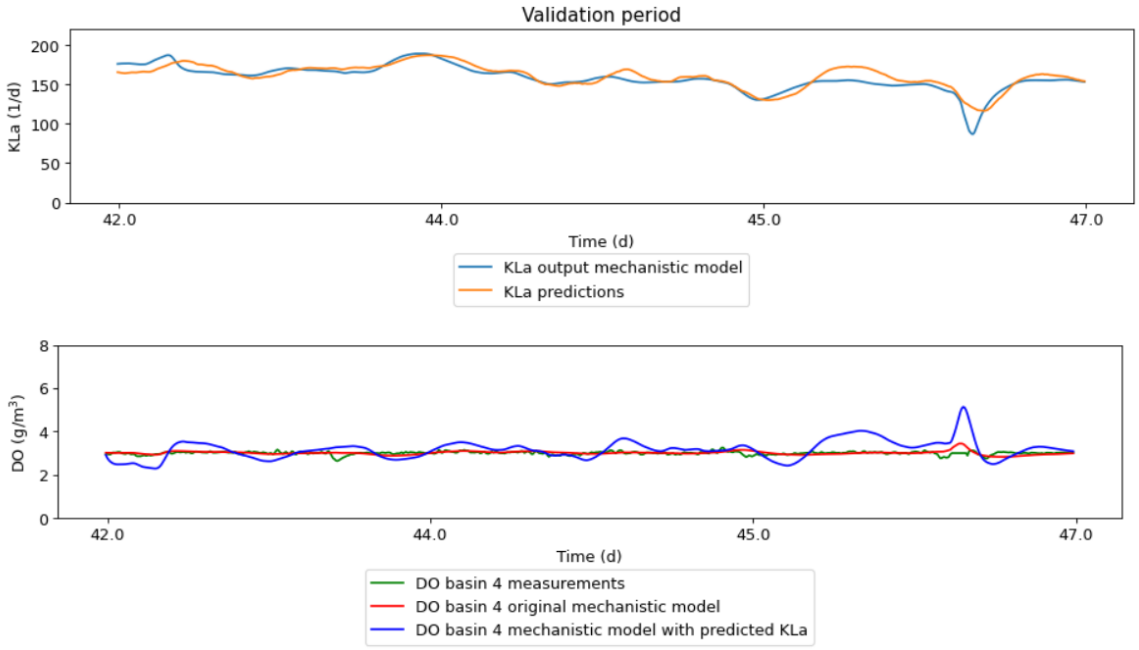

Figure 5.24: Serial hybrid model results during the validation period for the first approach where the CNN component is trained on the calibration dataset of the mech. model. The K_{La} output of the CNN compared to the K_{La} output of the mech. model (top) and DO output of the mech. model with the predicted K_{La} compared to the DO output of the original mech. model and DO measurements (bottom).

The K_{La} prediction made by the CNN and the corresponding DO output from the mechanistic model during the test period are shown in Figure 5.25. The K_{La} prediction during the test period exhibits a significant decline in accuracy. A potential factor contributing to this deviation is the reduction in K_{La} observed from days 48-50. During this two-day period, there was no influent flow in the pilEAUte, resulting in a decreased demand for DO, leading to a decrease in the K_{La}. These dynamics were not present during the training and validation period, which explains why the model encounters challenges in accurately predicting these patterns in the testing phase. The following K_{La} predictions may also be affected by this, as throughout the entire test period, the model encounters difficulties in predicting the correct order of magnitude. The CNN model used in this study did not include the influent flow rate as an input variable. This may have an adverse effect on the model's performance when attempting to predict situations characterized by variations in the influent flow rate. This is a shortcoming of the serial hybrid model employed in this study. However, some of the underlying dynamics are still captured by the CNN. A similar result is observed in the DO prediction, the magnitude is inaccurate, but when a DO peak occurs in the measurements, a disproportionately large peak is also present in the predictions by the CNN. It is important to note that the DO predictions generated by the original mechanistic model also exhibit inaccuracies in this time period.

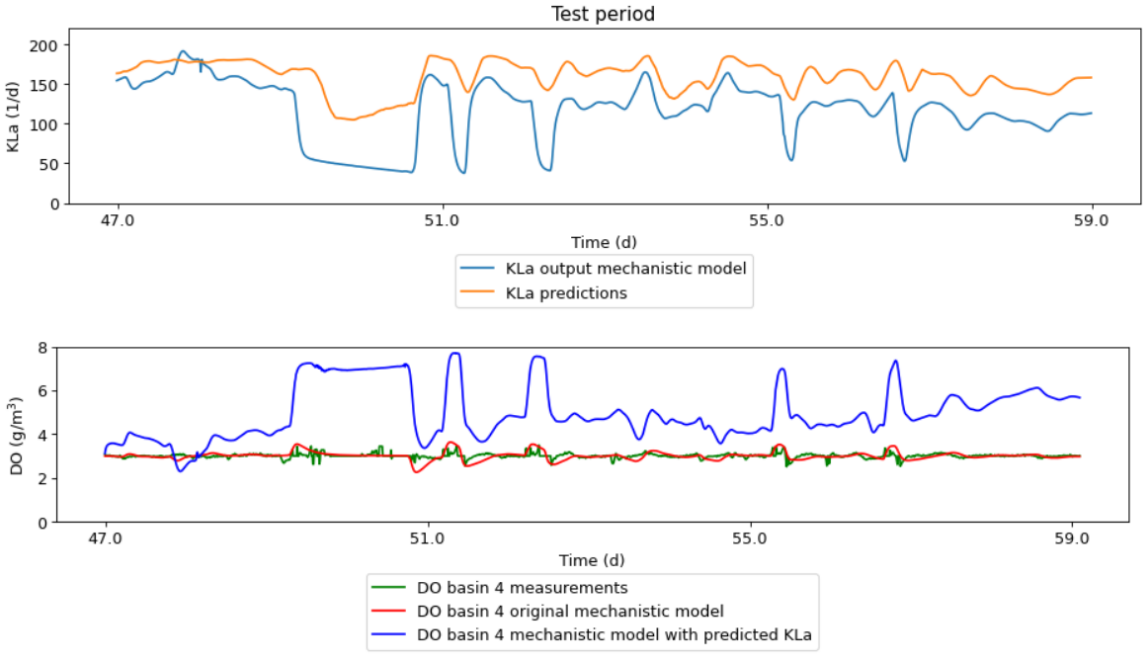

Figure 5.25: Serial hybrid model results during the test period for the first approach where the CNN component is trained on the calibration dataset of the mech. model. The K_{La} output of the CNN compared to the K_{La} output of the mech. model (top) and DO output of the mech. model with the predicted K_{La} compared to the DO output of the original mech. model and DO measurements (bottom).

So, the K_{La} can be accurately predicted by the CNN during the training and validation periods, but not during the test period. This inconsistency could potentially be attributed to the absence of influent stop in the training and validation periods. The DO predictions are highly sensitive to changes in K_{La}, which is justifiable considering the inherent relationship between the two variables.

Since the validation dataset of the mechanistic model clearly shows some specific features and additional dynamics that were not present in the calibration dataset of the mechanistic model (Figure 5.25), we also explore the serial model's predictive power when training on the mechanistic model's validation dataset. Using the period with influent stop as part of the training data could potentially be valuable in capturing and predicting this specific phenomenon. Note that the fact that the model was trained on the calibration dataset of the mechanistic model is less relevant here since the mechanistic model was not calibrated specifically for DO. The Qair measurements were only used to establish the relationship between K_{La} and Qair.

The K_{La} predictions made by the CNN and the corresponding DO output from the mechanistic model during the training period are shown in Figure 5.26. The K_{La} is accurately predicted here, even during the K_{La} decrease caused by the influent stop during day 48-

50. The small deviations in the predictions have a significant impact on DO prediction by the mechanistic model. A potential issue with the training on the validation dataset of the mechanistic model lies in the fact that the mechanistic model does not perform well to predict DO in this period either. The K_{La} in the mechanistic model is determined based on a controller aiming to maintain DO at 3 mg/L. However, in the validation dataset, several disturbances occur, making it challenging for the controller in the mechanistic model to keep DO consistently at 3 mg/L.

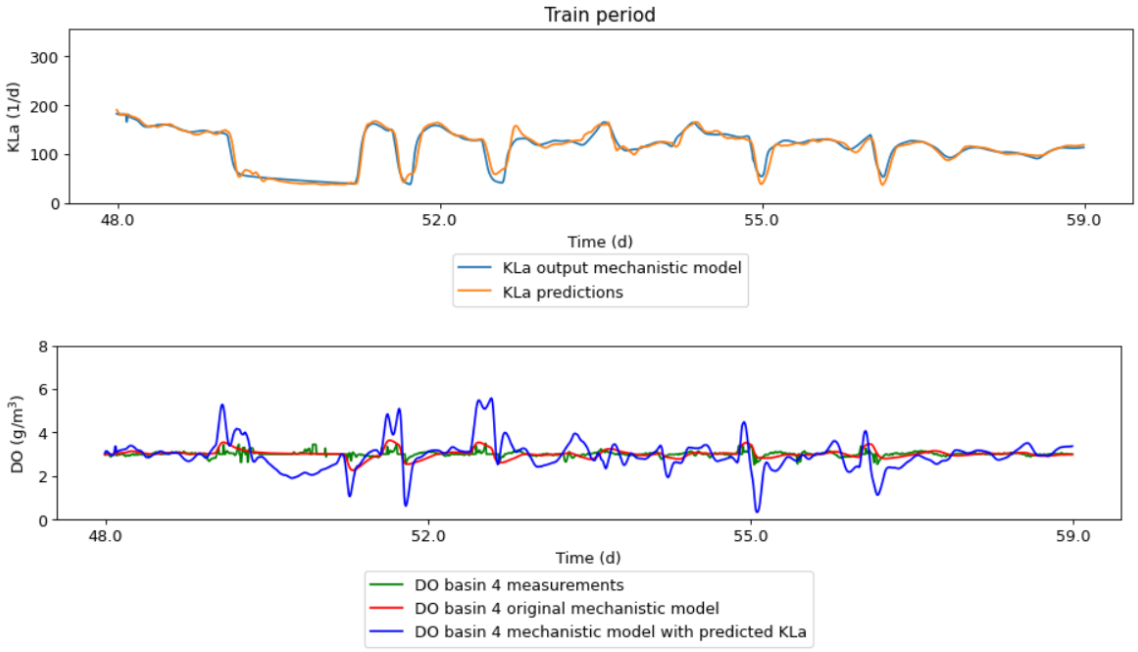

Figure 5.26: Serial hybrid model results during the training period for the second approach where the CNN component is trained on the validation dataset of the mech. model. The K_{La} output of the CNN compared to the K_{La} output of the mech. model (top) and DO output of the mech. model with the predicted K_{La} compared to the DO output of the original mech. model and DO measurements (bottom).

The results of the K_{La} prediction and the corresponding DO output from the mechanistic model during the validation period can be found in Figure 5.27. The K_{La} is no longer accurately predicted during this period, with only partial retrieval of some of the dynamics. Overall, the prediction is not satisfactory. This is also reflected in the DO prediction, which shows almost constant values close to 0 with occasional peaks above 3. The inaccurate prediction can be attributed to the consistently underestimated K_{La} values throughout the time period.

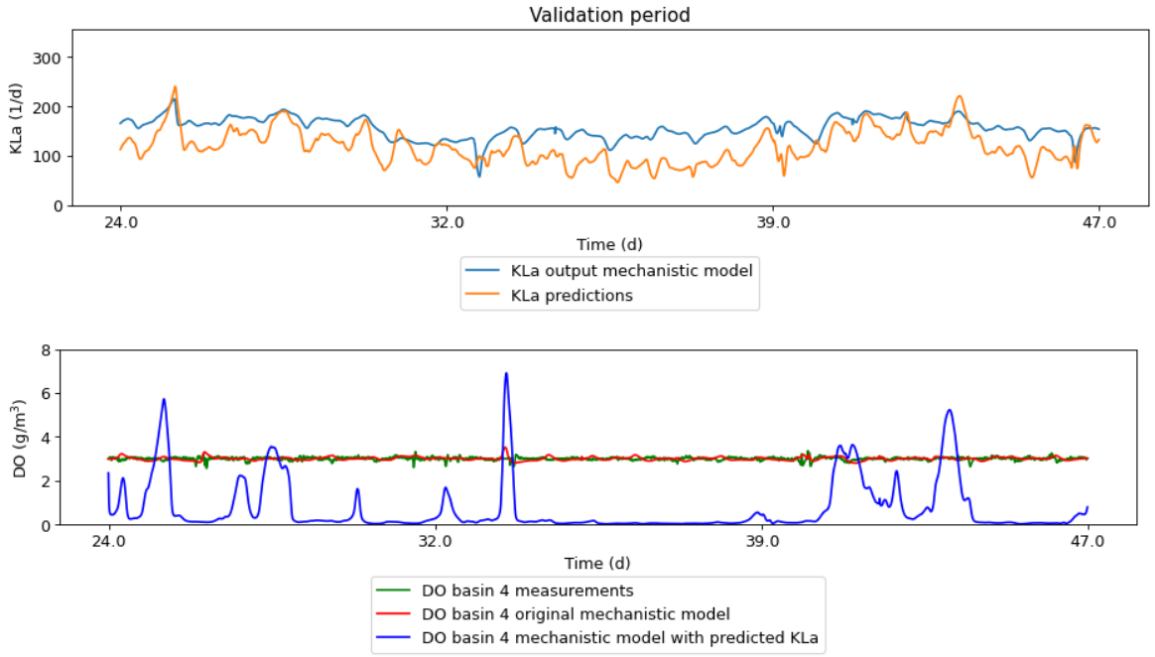

Figure 5.27: Serial hybrid model results during the validation period for the second approach where the CNN component is trained on the validation dataset of the mech. model. The K_{La} output of the CNN compared to the K_{La} output of the mech. model (top) and DO output of the mech. model with the predicted K_{La} compared to the DO output of the original mech. model and DO measurements (bottom).

The results of the K_{La} prediction and the corresponding DO output from the mechanistic model during the test period can be found in Figure 5.28. During the test period, similar findings are observed compared to the validation period. The K_{La} is underpredicted, however, the CNN manages to capture certain aspects of the correct dynamics. For instance, the peak on day 11 and the peak on day 21 are predicted by the CNN, but with incorrect magnitudes. This can again be observed in the corresponding DO output associated with these K_{La} predictions.

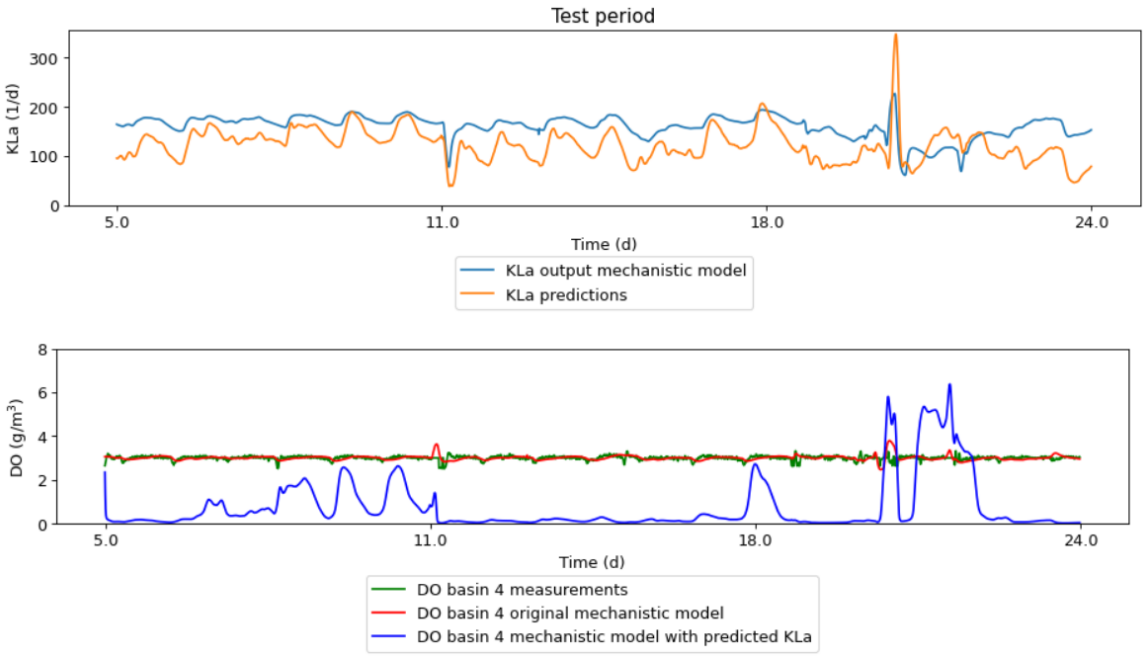

Figure 5.28: Serial hybrid model results during the test period for the second approach where the CNN component is trained on the validation dataset of the mech. model. The K_{La} output of the CNN compared to the K_{La} output of the mech. model (top) and DO output of the mech. model with the predicted K_{La} compared to the DO output of the original mech. model and DO measurements (bottom).

The challenge with training on the validation dataset is that the K_{La} output from the mechanistic model likely does not perform well in this case during the validation period. As mentioned, this K_{La} cannot accurately maintain the DO at 3 mg/L. By training a CNN on a K_{La} that may not represent the ground truth, the model struggles to predict K_{La} in other periods.

5.2.3 Serial hybrid model for enhanced KLa prediction: advantages and opportunities for improvement

The ability to predict the oxygen transfer coefficient using data-driven models can offer valuable insights for informed decision-making and comprehensive analysis across various applications. The prediction of K_{La} by the mechanistic model of the pilEAUte relies on the DO controller in basin 4. While the CNN demonstrates reasonable performance in predicting K_{La}, it still struggles to extrapolate to new situations. Moreover, the uncertainty of using the K_{La} output of the mechanistic model as the ground truth adds to the doubt regarding the effectiveness of this approach.

In Figure C.1, it can be observed that Qair and NH_4-N exhibit higher values during the calibration dataset of the mechanistic model (24/02-17/03) compared to the validation

dataset of the mechanistic model (17/03-31/03). Additionally, Qair shows a wider range during the validation period. Furthermore, it is worth noting that the MLSS concentration in tank 2 exhibits significantly lower values during the validation period in comparison to the calibration period. This observation is important because variations in the MLSS concentration influences the K_{La}. The two periods thus exhibit distinct characteristics. It is highly likely that the two CNN models struggle to extrapolate well to different scenarios as they were not trained on data with diverse characteristics. This implies that when constructing this serial model, extra caution is required to ensure sufficient variability in the training data. Consequently, a longer dataset may be necessary to capture and learn all significant behaviors effectively.

In this research, the influent flow rate and MLSS were not incorporated as inputs for the CNN component of the serial hybrid model. However, these variables exhibit significant variations during the validation dataset of the mechanistic model, which could potentially contribute to predicting patterns in the K_{La}. The inclusion of the influent flow rate and MLSS as input variables for the CNN component could potentially lead to significant improvements in the prediction accuracy. Unfortunately, due to time constraints, it was not possible to include them and recreate the serial models to assess their impact on the prediction performance within the scope of this thesis.

To address these challenges, the integration of reinforcement learning techniques could prove beneficial to predict K_{La} in wastewater treatment processes. Reinforcement learning might be used to adjust neural network parameters based on the deviation between predicted and actual DO measurements thus removing the dependency of intermediate K_{La} estimates. The reward system would thus be based on the accuracy of DO predictions. The neural network learns iteratively by receiving rewards for accurate DO predictions and adjusting its parameters (weights and biases) to maximize rewards and predict a more accurate K_{La}. The main benefit for developing an accurate serial hybrid model for K_{La} is that it can be used to predict in real-time the K_{La} value. As such it can be used for control of air flow rate to ensure that the air flow is adapted in a feedforward setting to the conditions in the biological reactor. It can also serve as an early warning system to alert operators of process conditions where oxygen transfer may become limiting.

6. CONCLUSION AND PERSPECTIVES

6.1 Conclusion

Hybrid models can combine the advantages of mechanistic models and data-driven models. However, there is currently no widely accepted or established methodology for developing hybrid models for wastewater treatment plants. In previous research, a mechanistic model was developed for a pilot-scale wastewater treatment plant, but it exhibited limitations in accurately predicting effluent quality. In the first part of this thesis, parallel hybrid models are developed, in which a data-driven component of the HM is trained to forecast the residual between the mechanistic model output and the measurements for effluent nitrate. The performances of a LSTM-RNN hybrid model and a CNN hybrid model were compared. The LSTM hybrid model struggled to learn meaningful patterns from the input data to predict correct dynamics, whereas the hybrid model with CNN was able to detect dynamics in the residuals and make more accurate predictions for effluent nitrate than the mechanistic model and the LSTM HM. This suggests that CNN is a more suitable data-driven component for parallel hybrid models of wastewater treatment system time series than LSTM-RNN. The final HM was also compared to a stand-alone CNN with the HM clearly outperforming the pure data-driven model.

Next, the balance between calibration effort for the mechanistic model and acceptable compensation by the neural network component is investigated. A first experiment involved training the CNN component of a parallel hybrid model using both a calibration dataset and an independent validation dataset from the mechanistic model. A comparison of these models revealed that the HM performed better when trained on the independent validation dataset. In a subsequent experiment, training the CNN component of a parallel hybrid model on the residual of an uncalibrated mechanistic model produced superior performance compared to training on the residual of a calibrated mechanistic model. The results of these experiment indicate that hybrid modes are more difficult to construct when the mechanistic component is overly calibrated and overfit for uncertain or unidentifiable parameters. Overfitting is always a risk in mechanistic models

of wastewater treatment plants as some missing dynamics are often present in typical mechanistic models for biological wastewater treatment. Detailed calibration of the mechanistic model with several unidentifiable parameters makes it more difficult for a parallel neural network component to learn the missing true model dynamics. The present work shows that a parallel hybrid model approach can compensate for missing dynamics and non calibrated phenomena. This could potentially save time in the calibration effort of mechanistic models.

Developing a parallel hybrid model to improve the effluent TSS prediction of the mechanistic model showed that parallel hybrid models may not be applicable for predicting all variables in wastewater treatment processes. The effectiveness of a parallel hybrid models highly depends on factors such as data variability, structure and accuracy of the mechanistic model, and quality of the data used for modelling. These considerations underscore the need for careful assessment of the suitability and potential advantages of parallel hybrid models for specific variables in wastewater treatment processes.

In a second part of this thesis also serial hybrid model was developed using a CNN to forecast the oxygen transfer coefficient (K_{La}) in an aerated basin of a pilot plant. Although the CNN showed reasonable performance in predicting K_{La}, it faced challenges in extrapolating to new situations. To accurately predict K_{La}, a long dataset with sufficient variability is required to capture all possible behaviors. Further research is needed to improve the prediction of K_{La} by serial hybrid models, however, the serial hybrid model shows promise and could potentially be used in the future for real-time prediction of K_{La} in WRRFs.

6.2 Perspectives

Further research is required to reach the full potential of hybrid modelling. Currently, there is limited research available on determining the effective data-driven components in parallel hybrid model structures. This study revealed the potential of CNN as data-driven component, but further exploration of alternative options may lead to even better methods. In recent years, CNN-LSTM models have emerged as a valuable data-driven approach for time series prediction. This model combines the strengths of CNN by leveraging the feature extraction capabilities and LSTM by capturing long-term dependencies in input time series data (Xie et al., 2020; Jin et al., 2020).

This research demonstrated the potential of serial hybrid models for predicting oxygen transfer coefficients in wastewater treatment processes, but also highlights the need for significant improvements in this approach. For instance, using different input data and exploring alternative data-driven methods could greatly enhance the accuracy of the predictions. Also the applicability of serial hybrid models to other subprocesses in wastewater treatment models can be explored. A specific challenge with respect to serial HMs is that true measurements of the variable to be predicted (for example K_{La}) are often lacking. Other methods to train a serial model on a secondary output (for example DO) should be explored.

This study aims to answer a few questions related to identifying the most effective methods for constructing hybrid models. However, a general framework for integrating mechanistic and data-driven models still needs to be developed. It should address key considerations such as data compatibility, model selection, parameter estimation, model validation, and interpretation of the hybrid model results. By developing this framework, researchers and practitioners can ensure a systematic and reliable integration of data-driven and mechanistic models, leading to improved accuracy and understanding of complex systems.

BIBLIOGRAPHY

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., et al. (2016). Tensorflow: A system for large-scale machine learning. In *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, pages 265–283.

Alferes, J. and Vanrolleghem, P. A. (2016). Efficient automated quality assessment: Dealing with faulty on-line water quality sensors. *AI Communications*, 29(6):701–709.

Aloe, A. K., Bouraoui, F., Grizzetti, B., Bidoglio, G., and Pistocchi, A. (2014). Managing nitrogen and phosphorus loads to water bodies: characterisation and solutions. *Towards macro-regional integrated nutrient management. Joint Research Centre, JRC-Ispra*.

Anderson, J., McAvoy, T., and Hao, O. (2000). Use of hybrid models in wastewater systems. *Industrial & Engineering Chemistry Research*, 39(6):1694–1704.

Bahramian, M., Dereli, R. K., Zhao, W., Giberti, M., and Casey, E. (2022). Data to intelligence: The role of data-driven models in wastewater treatment. *Expert Systems with Applications*, page 119453.

Benedetti, L. (2006). *Probabilistic design and upgrade of wastewater treatment plants in the EU Water Framework Directive context*. PhD thesis, Ghent University.

Borovykh, A., Bohte, S., and Oosterlee, C. W. (2017). Conditional time series forecasting with convolutional neural networks. *arXiv preprint arXiv:1703.04691*.

Bray, R. T., Jankowska, K., Kulbat, E., Łuczkiewicz, A., and Sokołowska, A. (2021). Ultrafiltration process in disinfection and advanced treatment of tertiary treated wastewater. *Membranes*, 11(3):221.

Bürger, R., Careaga, J., Diehl, S., Mejías, C., Nopens, I., Torfs, E., and Vanrolleghem, P. A. (2016). Simulations of reactive settling of activated sludge with a reduced biokinetic model. *Computers & Chemical Engineering*, 92:216–229.

Bürger, R., Diehl, S., and Nopens, I. (2011). A consistent modelling methodology for secondary settling tanks in wastewater treatment. *Water research*, 45(6):2247–2260.

Campos, J. L., Valenzuela-Heredia, D., Pedrouso, A., Val del Río, A., Belmonte, M., and Mosquera-Corral, A. (2016). Greenhouse gases emissions from wastewater treatment plants: minimization, treatment, and prevention. *Journal of Chemistry*, 2016.

Chen, Y., Song, L., Liu, Y., Yang, L., and Li, D. (2020). A review of the artificial neural network models for water quality prediction. *Applied Sciences*, 10(17):5776.

Chollet, F. et al. (2015). Keras. `https://github.com/fchollet/keras`.

De Clercq, B., Coen, F., Vanderhaegen, B., and Vanrolleghem, P. A. (1999). Calibrating simple models for mixing and flow propagation in waste water treatment plants. *Water Science and Technology*, 39(4):61–69.

DHI (2017). MIKE Powered by DHI. `https://www.mikepoweredbydhi.com/products/west`.

Di Costanzo, N., Cesaro, A., Di Capua, F., and Esposito, G. (2021). Exploiting the nutrient potential of anaerobically digested sewage sludge: a review. *Energies*, 14(23):8149.

Dong, S., Zhang, Y., and Zhou, X. (2023). Intelligent hybrid modeling of complex leaching system based on LSTM neural network. *Systems*, 11(2):78.

Durairaj, D. M. and Mohan, B. K. (2022). A convolutional neural network based approach to financial time series prediction. *Neural Computing and Applications*, 34(16):13319–13337.

Ekama, G., Barnard, G., Gunthert, F., Krebs, P., McCorquodale, J., Parker, D., and Wahlberg, E. (1997). Secondary settling tanks. *London: International Association on Water Quality*.

Flores-Alsina, X., Saagi, R., Lindblom, E., Thirsing, C., Thornberg, D., Gernaey, K. V., and Jeppsson, U. (2014). Calibration and validation of a phenomenological influent pollutant disturbance scenario generator using full-scale data. *Water Research*, 51:172–185.

Franchi, A. and Santoro, D. (2015). Current status of the rotating belt filtration (RBF) technology for municipal wastewater treatment. *Water Practice and Technology*, 10(2):319–327.

Gerba, C. P. and Pepper, I. L. (2019). Municipal wastewater treatment. In *Environmental and pollution science*, pages 393–418. Elsevier.

Gernaey, K. V., Van Loosdrecht, M. C., Henze, M., Lind, M., and Jørgensen, S. B. (2004). Activated sludge wastewater treatment plant modelling and simulation: state of the art. *Environmental modelling & software*, 19(9):763–783.

Gianico, A., Braguglia, C. M., Gallipoli, A., Montecchio, D., and Mininni, G. (2021). Land application of biosolids in Europe: possibilities, con-straints and future perspectives. *Water*, 13(1):103.

Graves, A. (2012). Long short-term memory. *Supervised sequence labelling with recurrent neural networks*, pages 37–45.

Guo, L. and 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 and biosystems engineering*, 37:151–163.

Hastie, T., Tibshirani, R., and Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer.

Henze, M., Gujer, W., Mino, T., and van Loosdrecht, M. C. (2000). *Activated sludge models ASM1, ASM2, ASM2d and ASM3*. IWA publishing.

Hong, S. H., Lee, M. W., Lee, D. S., and Park, J. M. (2007). Monitoring of sequencing batch reactor for nitrogen and phosphorus removal using neural networks. *Biochemical Engineering Journal*, 35(3):365–370.

Hreiz, R., Latifi, M., and Roche, N. (2015). Optimal design and operation of activated sludge processes: State-of-the-art. *Chemical Engineering Journal*, 281:900–920.

Hvala, N. and Kocijan, J. (2020). Design of a hybrid mechanistic/Gaussian process model to predict full-scale wastewater treatment plant effluent. *Computers & Chemical Engineering*, 140:106934.

Illueca-Muñoz, J., Mendoza-Roca, J., Iborra-Clar, A., Bes-Piá, A., Fajardo-Montañana, V., Martínez-Francisco, F., and Bernácer-Bonora, I. (2008). Study of different alternatives of tertiary treatments for wastewater reclamation to optimize the water quality for irrigation reuse. *Desalination*, 222(1-3):222–229.

International Energy Agency (2019). 2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories, Chapter 6: Wastewater Treatment and Discharge. *World Energy Outlook 2016*.

Jalilnejad, E., Jabbari, B., and Ghasemzadeh, K. (2022). Application of computational fluid dynamics technique in membrane bioreactor systems. In *Current Trends and Future Developments on (Bio-) Membranes*, pages 345–375. Elsevier.

Janssen, P., Meinema, K., and Van der Roest, H. (2002). *Biological phosphorus removal*. IWA publishing.

Jeppsson, U. (1996a). A general description of the activated sludge model no. 1 (ASM1). *Examensarb. Lund University*.

Jeppsson, U. (1996b). *Modelling aspects of wastewater treatment processes*. Lund Institute of Technology Lund.

Jin, X., Yu, X., Wang, X., Bai, Y., Su, T., and Kong, J. (2020). Prediction for Time Series with CNN and LSTM. In *Proceedings of the 11th International Conference on Modelling, Identification and Control (ICMIC2019)*, pages 631–641. Springer.

Jourdan, N., Neveux, T., Potier, O., Kanniche, M., Wicks, J., Nopens, I., Rehman, U., and Le Moullec, Y. (2019). Compartmental modelling in chemical engineering: A critical review. *Chemical Engineering Science*, 210:115196.

Karthikeyan, O. P. and Joseph, K. (2007). Anaerobic Ammonium Oxidation (ANAMMOX) process for nitrogen removal–A review. In *Biological Methods of Waste Treatment and Management in South India Symposium, Chennai, India*. Citeseer.

Kingma, D. P. and Ba, J. (2015). Adam: A Method for Stochastic Optimization. In *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pages 1–15.

Kirim, G. (2022). *Modelling and Model-Based Optimization of N-Removal WRRFs: Reactive Settling, Conventional & Short-Cut N-Removal Processes*. PhD thesis, Université Laval, Québec, QC, Canada.

Kirim, G., Torfs, E., and Vanrolleghem, P. A. (2022). An improved 1-D reactive Bürger-Diehl settler model for secondary settling tank denitrification. *Water Environment Research*, page e10825.

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J., Grout, J., Corlay, S., Ivanov, P., Avila, D., Abdalla, S., and Willing, C. (2016). Jupyter notebooks – a publishing format for reproducible computational workflows. In

Loizides, F. and Schmidt, B., editors, *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, pages 87–90. IOS Press.

Koprinska, I., Wu, D., and Wang, Z. (2018). Convolutional neural networks for energy time series forecasting. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE.

Laurent, J., Samstag, R., Ducoste, J., Griborio, A., Nopens, I., Batstone, D. J., Wicks, J., Saunders, S., and Potier, O. (2014). A protocol for the use of computational fluid dynamics as a supportive tool for wastewater treatment plant modelling. *Water Science and Technology*, 70(10):1575–1584.

Lee, D. S., Jeon, C. O., Park, J. M., and Chang, K. S. (2002). Hybrid neural network modeling of a full-scale industrial wastewater treatment process. *Biotechnology and bioengineering*, 78(6):670–682.

Li, B., Huang, H. M., Boiarkina, I., Yu, W., Huang, Y. F., Wang, G. Q., and Young, B. R. (2019a). Phosphorus recovery through struvite crystallisation: Recent developments in the understanding of operational factors. *Journal of environmental management*, 248:109254.

Li, F., Kirim, G., and Vanrolleghem, P. A. (2019b). Characterizing the dynamics of pollutant concentration and biodegradability in raw domestic wastewaters. Poster presented at the 2e Journée québécoise des étudiants CentrEau, Québec, QC, Canada.

Lowe, M., Qin, R., and Mao, X. (2022). A review on machine learning, artificial intelligence, and smart technology in water treatment and monitoring. *Water*, 14(9):1384.

Maktabifard, M., Blomberg, K., Zaborowska, E., Mikola, A., and Mąkinia, J. (2022). Model-based identification of the dominant N_2O emission pathway in a full-scale activated sludge system. *Journal of Cleaner Production*, 336:130347.

MathWorks (2022). Regression learner app. `https://nl.mathworks.com/help/stats/regressionlearner-app.html`. Accessed on 05/06/2023.

McKinney, W. et al. (2010). Data Structures for Statistical Computing in Python. In *Proceedings of the 9th Python in Science Conference*, volume 445, pages 51–56, Austin, TX.

Mijwel, M. M. (2018). Artificial neural networks advantages and disadvantages. *Retrieved from LinkedIn https//www. linkedin. com/pulse/artificial-neuralnet Work*.

Neczaj, E. and Grosser, A. (2018). Circular economy in wastewater treatment plant–challenges and barriers. *Multidisciplinary digital publishing institute proceedings*, 2(11):614.

Nian, R., Liu, J., and Huang, B. (2020). A review on reinforcement learning: Introduction and applications in industrial process control. *Computers & Chemical Engineering*, 139:106886.

Nikbakht, S., Anitescu, C., and Rabczuk, T. (2021). Optimizing the neural network hyperparameters utilizing genetic algorithm. *Journal of Zhejiang University-Science A*, 22(6):407–426.

Nopens, I. (2005). *Modelling the activated sludge flocculation process: a population balance approach*. PhD thesis, University Gent.

Oliphant, T. E. et al. (2006). *A guide to NumPy*, volume 1. Trelgol Publishing USA.

Olsson, G. and Newell, B. (1999). *Wastewater treatment systems*. IWA publishing.

O'Malley, T., Bursztein, E., Long, J., Chollet, F., Jin, H., Invernizzi, L., et al. (2019). Keras-tuner. `https://github.com/keras-team/keras-tuner`.

Onkal-Engin, G., Demir, I., and Engin, S. N. (2005). Determination of the relationship between sewage odour and BOD by neural networks. *Environmental Modelling & Software*, 20(7):843–850.

Pan, Y. and Dagnew, M. (2022). A new approach to estimating oxygen off-gas fraction and dynamic alpha factor in aeration systems using hybrid machine learning and mechanistic models. *Journal of Water Process Engineering*, 48:102924.

Park, Y.-S. and Lek, S. (2016). Artificial neural networks: Multilayer perceptron for ecological modeling. In *Developments in environmental modelling*, volume 28, pages 123–140. Elsevier.

Peres, J., Oliveira, R., and De Azevedo, S. F. (2001). Knowledge based modular networks for process modelling and control. *Computers & Chemical Engineering*, 25(4-6):783–791.

Philippe, R. (2018). Automatic data quality assessment tools for continuous monitoring of wastewater quality. Master's thesis, Université Laval, Québec, QC, Canada.

Pomerat, J., Segev, A., and Datta, R. (2019). On neural network activation functions and optimizers in relation to polynomial regression. In *2019 IEEE International Conference on Big Data (Big Data)*, pages 6183–6185. IEEE.

Ráduly, B., Gernaey, K. V., Capodaglio, A. G., Mikkelsen, P. S., and Henze, M. (2007). Artificial neural networks for rapid WWTP performance evaluation: Methodology and case study. *Environmental modelling & software*, 22(8):1208–1216.

Ramin, E., Wágner, D. S., Yde, L., Binning, P. J., Rasmussen, M. R., Mikkelsen, P. S., and Plósz, B. G. (2014). A new settling velocity model to describe secondary sedimentation. *Water Research*, 66:447–458.

Redmon, D., Boyle, W. C., and Ewing, L. (1983). Oxygen transfer efficiency measurements in mixed liquor using off-gas techniques. *Journal (Water Pollution Control Federation)*, pages 1338–1347.

Regmi, P., Stewart, H., Amerlinck, Y., Arnell, M., García, P. J., Johnson, B., Maere, T., Miletić, I., Miller, M., Rieger, L., et al. (2019). The future of WRRF modelling–outlook and challenges. *Water Science and Technology*, 79(1):3–14.

Rieger, L., Gillot, S., Langergraber, G., Ohtsuki, T., Shaw, A., Takács, I., and Winkler, S. (2012). *Guidelines for Using Activated Sludge Models*. IWA, London, UK.

Riffat, R. and Husnain, T. (2013). *Fundamentals of wastewater treatment and engineering*. Crc Press.

Rodriguez-Roda, I., Sànchez-Marrè, M., Comas, J., Baeza, J., Colprim, J., Lafuente, J., Cortés, U., and Poch, M. (2002). A hybrid supervisory system to support WWTP operation: implementation and validation. *Water science and technology*, 45(4-5):289–297.

Rosso, D., Iranpour, R., and Stenstrom, M. (2005). Fifteen years of offgas transfer efficiency measurements on fine-pore aerators: Key role of sludge age and normalized air flux. *Water environment research*, 77(3):266–273.

Rout, P. R., Zhang, T. C., Bhunia, P., and Surampalli, R. Y. (2021). Treatment technologies for emerging contaminants in wastewater treatment plants: A review. *Science of the Total Environment*, 753:141990.

Sak, H., Senior, A. W., and Beaufays, F. (2014). Long short-term memory recurrent neural network architectures for large scale acoustic modeling.

Schneider, M. Y., Quaghebeur, W., Borzooei, S., Froemelt, A., Li, F., Saagi, R., Wade, M. J., Zhu, J.-J., and Torfs, E. (2022). Hybrid modelling of water resource recovery facilities: status and opportunities. *Water Science and Technology*, 85(9):2503–2524.

Sharkawy, A.-N. (2020). Principle of neural network and its main types. *Journal of Advances in Applied & Computational Mathematics*, 7:8–19.

Shewa, W. A. and Dagnew, M. (2020). Revisiting chemically enhanced primary treatment of wastewater: A review. *Sustainability*, 12(15):5928.

Shukla, P. and Iriondo, R. (2020). Main types of neural networks and its applications. Medium article. Accessed: 08/06/2023.

Silvestre, G., Fernández, B., and Bonmatí, A. (2015). Significance of anaerobic digestion as a source of clean energy in wastewater treatment plants. *Energy Conversion and Management*, 101:255–262.

Singh, Y. and Chauhan, A. S. (2009). Neural networks in data mining. *Journal of Theoretical & Applied Information Technology*, 5(1).

Staudemeyer, R. C. and Morris, E. R. (2019). Understanding LSTM–a tutorial into long short-term memory recurrent neural networks. *arXiv preprint arXiv:1909.09586*.

Takács, I., Patry, G. G., and Nolasco, D. (1991). A dynamic model of the clarification-thickening process. *Water research*, 25(10):1263–1271.

Tchobanoglus, G., Burton, F., Tsuchihashi, F. L., and Stensel, H. D. (2013). *Water and Wastewater Engineering: treatment and resource recovery*. Metcalf & Eddy.

Torfs, E. (2015). *Different settling regimes in secondary settling tanks: experimental process analysis, model development and calibration*. PhD thesis, Ghent University.

Torfs, E., Nicolaï, N., Daneshgar, S., Copp, J. B., Haimi, H., Ikumi, D., Johnson, B., Plosz, B. B., Snowling, S., Townley, L. R., et al. (2022). The transition of WRRF models to digital twin applications. *Water Science and Technology*, 85(10):2840–2853.

Ungureanu, N., Vlăduț, V., and Voicu, G. (2020). Water scarcity and wastewater reuse in crop irrigation. *Sustainability*, 12(21):9055.

Van Hulle, S. (2005). *Modelling, simulation and optimization of autotrophic nitrogen removal processes*. PhD thesis, Ghent University.

Van Rossum, G. and Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA.

VMM (2022). Blue deal geeft 33 miljoen m^3 gezuiverd afvalwater nieuwe bestemming. `https://www.vmm.be/nieuws/archief/blue-deal-geeft-33-miljoen-m3-gezuiverd-afvalwater-nieuwe-bestemming`. Accessed on 31-10-2022.

Von Sperling, M. (2007). *Basic principles of wastewater treatment*. IWA publishing.

Von Stosch, M., Oliveira, R., Peres, J., and de Azevedo, S. F. (2014). Hybrid semi-parametric modeling in process systems engineering: Past, present and future. *Computers & Chemical Engineering*, 60:86–101.

Wang, G., Jia, Q.-S., Zhou, M., Bi, J., Qiao, J., and Abusorrah, A. (2022). Artificial neural networks for water quality soft-sensing in wastewater treatment: a review. *Artificial Intelligence Review*, 55(1):565–587.

Wang, K., Li, K., Zhou, L., Hu, Y., Cheng, Z., Liu, J., and Chen, C. (2019). Multiple convolutional neural networks for multivariate time series prediction. *Neurocomputing*, 360:107–119.

Wang, Z., Yan, W., and Oates, T. (2017). Time series classification from scratch with deep neural networks: A strong baseline. In *2017 International joint conference on neural networks (IJCNN)*, pages 1578–1585. IEEE.

Wu, Z., Pan, S., Long, G., Jiang, J., Chang, X., and Zhang, C. (2020). Connecting the dots: Multivariate time series forecasting with graph neural networks. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 753–763.

Xie, H., Zhang, L., and Lim, C. P. (2020). Evolving CNN-LSTM models for time series prediction using enhanced grey wolf optimizer. *IEEE access*, 8:161519–161541.

Yamashita, R., Nishio, M., Do, R. K. G., and Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights into imaging*, 9:611–629.

Zhang, M., Wang, S., Ji, B., and Liu, Y. (2019). Towards mainstream deammonification of municipal wastewater: Partial nitrification-anammox versus partial denitrification-anammox. *Science of the Total Environment*, 692:393–401.

Zhang, Z., Zhang, K., and Khelifi, A. (2018). *Multivariate time series analysis in climate and environmental research*. Springer.

Zhu, J.-J., Segovia, J., and Anderson, P. R. (2015). Defining influent scenarios: application of cluster analysis to a water reclamation plant. *Journal of Environmental Engineering*, 141(7):04015005.

Zhu, M., Wang, J., Yang, X., Zhang, Y., Zhang, L., Ren, H., Wu, B., and Ye, L. (2022). A review of the application of machine learning in water quality evaluation. *Eco-Environment & Health*.

APPENDIX A

HYDRAULIC MODEL RESULTS

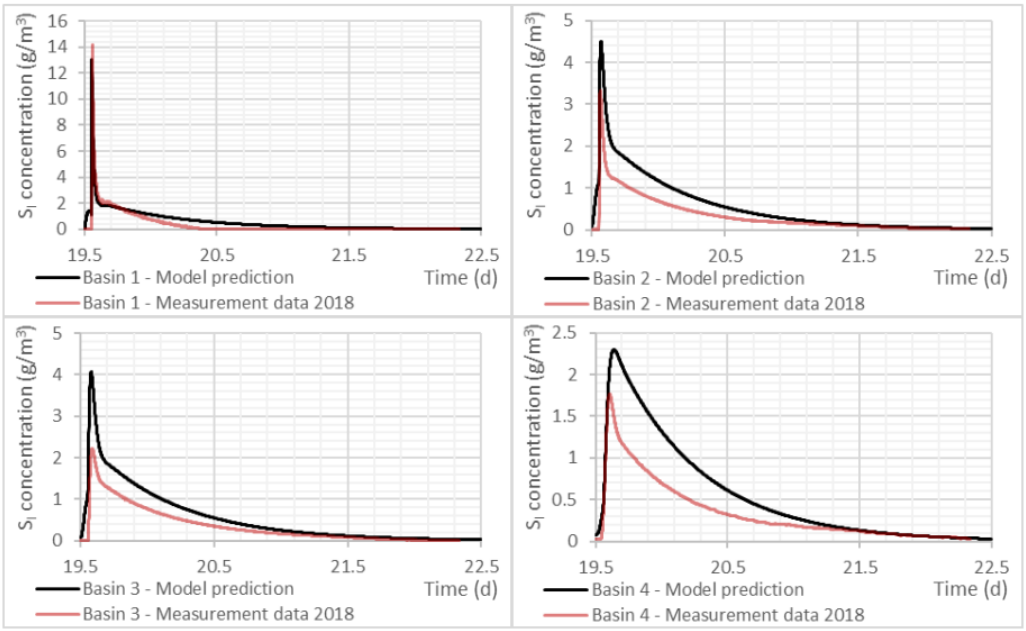

Figure A.1: Final hydraulic model results for reference case tracer experiment (Kirim, 2022).

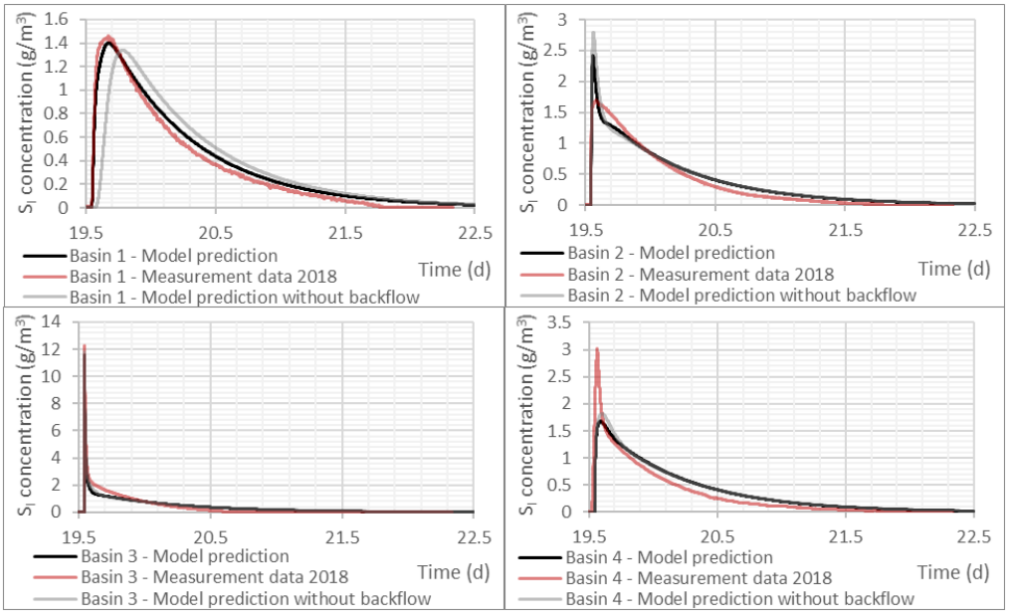

Figure A.2: Final hydraulic model results for step-feed case (Kirim, 2022).

APPENDIX B

MECHANISTIC MODEL RESULTS

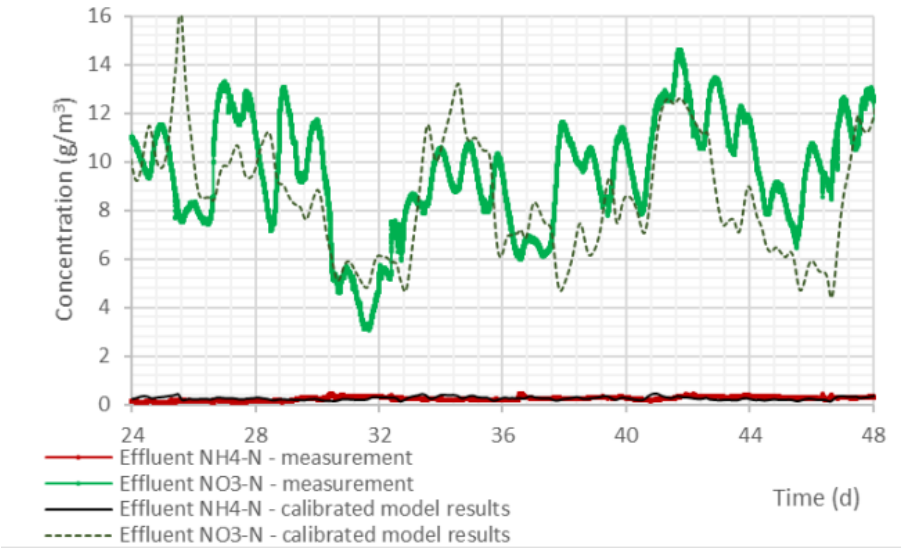

Figure B.1: Calibrated mechanistic model results and measurements for nitrate and ammonium nitrogen during the calibration period (Kirim, 2022).

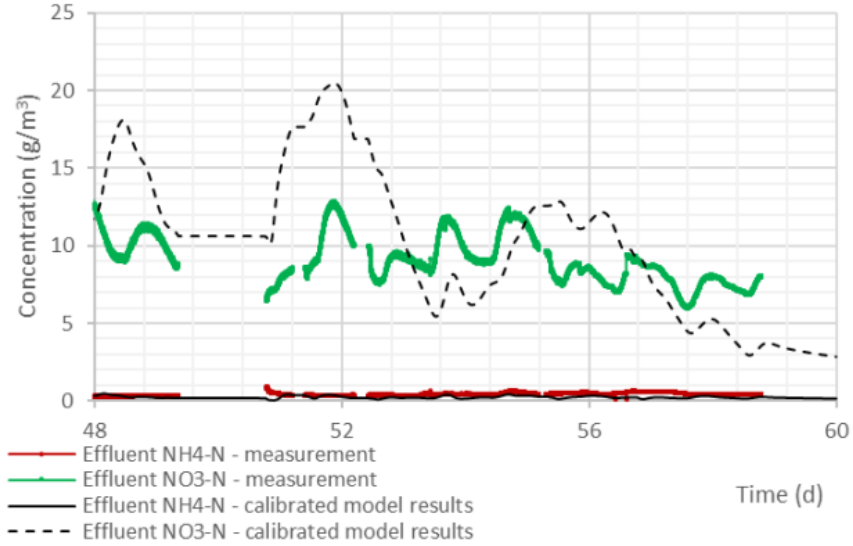

Figure B.2: Calibrated mechanistic model results and measurements for nitrate and ammonium nitrogen during the validation period (Kirim, 2022).

APPENDIX C

NEURAL NETWORK INPUT DATA

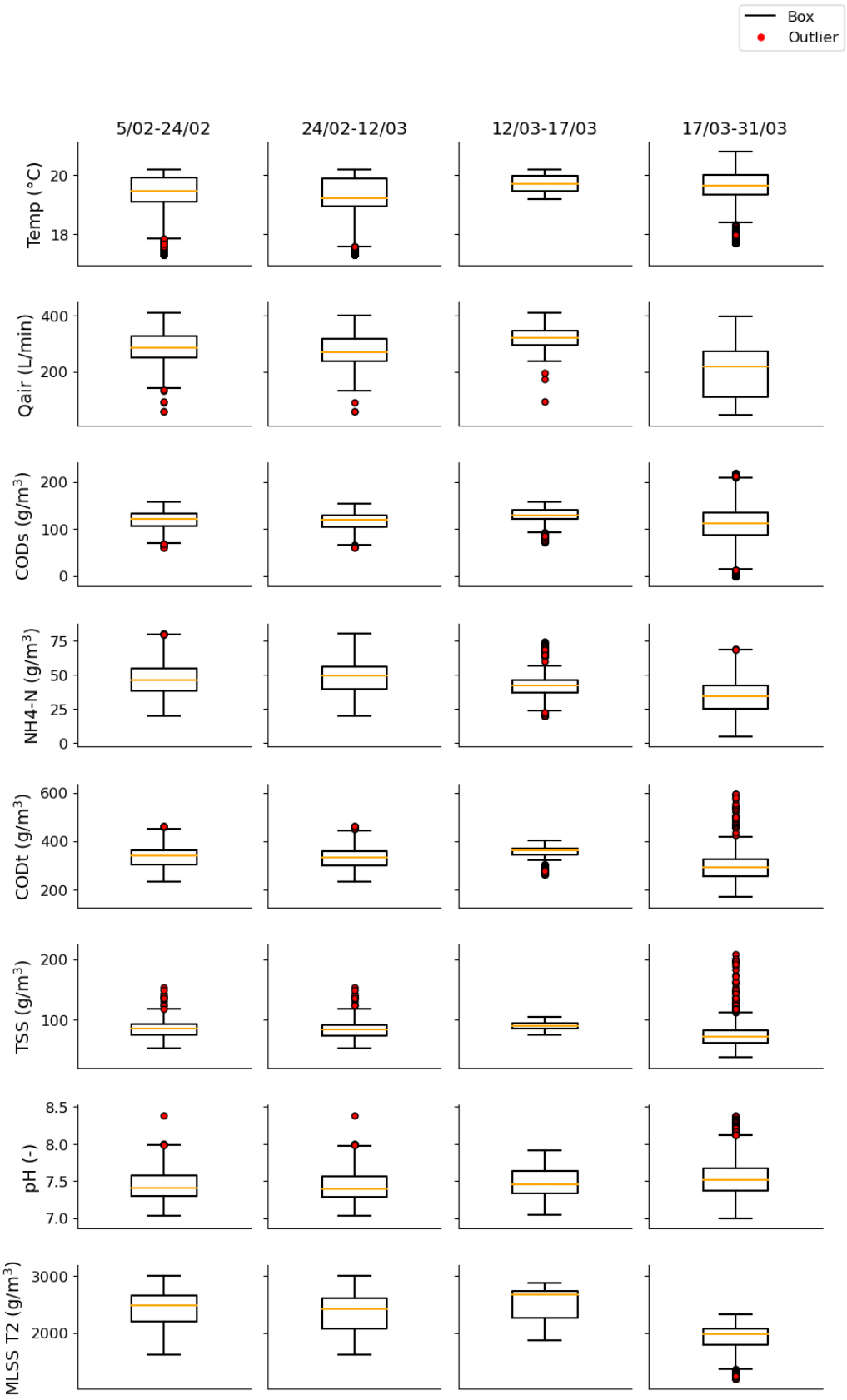

Figure C.1: Boxplots of the neural network input data for every variable for every period used to establish the neural networks. MLSS in tank 2 is not an input for the NNs, but is added for completeness.

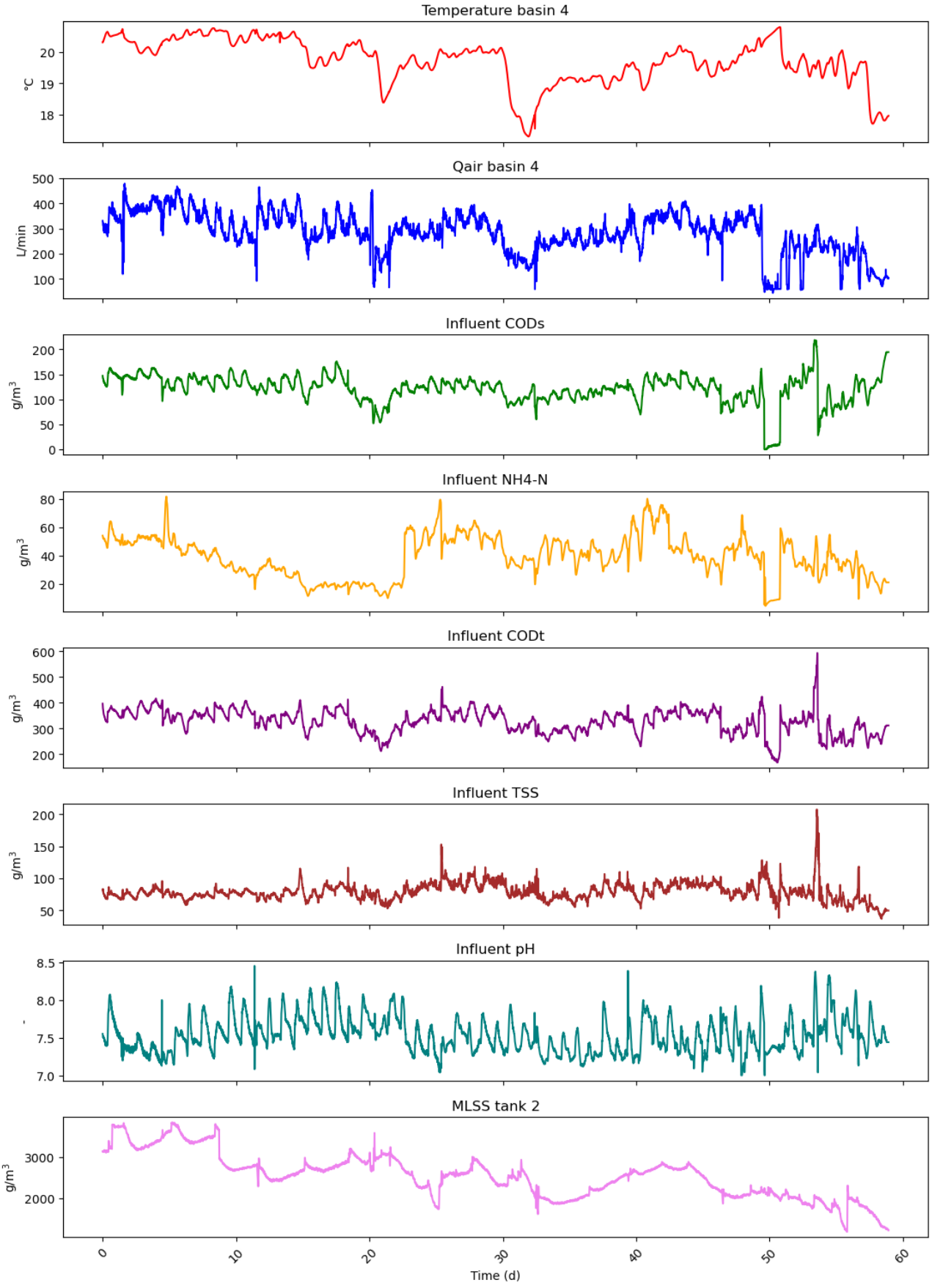

Figure C.2: Neural network input data during the whole time period. MLSS in tank 2 is not an input for the NNs, but is added for completeness.

APPENDIX D

DATA-DRIVEN MODEL RESULTS

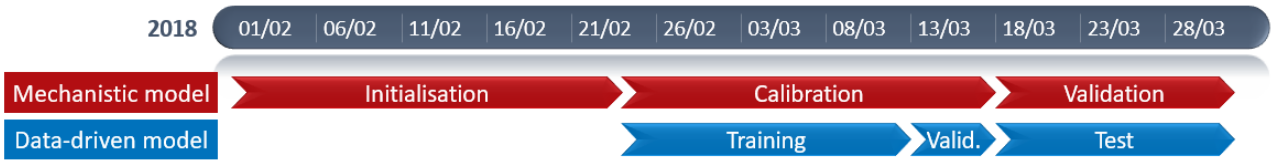

Figure D.1: Timeline with the time periods used to develop a data-driven model for effluent nitrate.

Table D.1: Final set of hyperparameters determined for data-driven model (CNN) for effluent nitrate.

Prediction	**Network**	**Time steps**	**Structure details**	**Learning rate Adam optimizer**	**Epochs**
NO3	CNN	30	Number of layers: 2	0.0001	25
			Filter: 10, 5		
			Kernel size: 3, 3		
			Activation: ReLu, tanh		

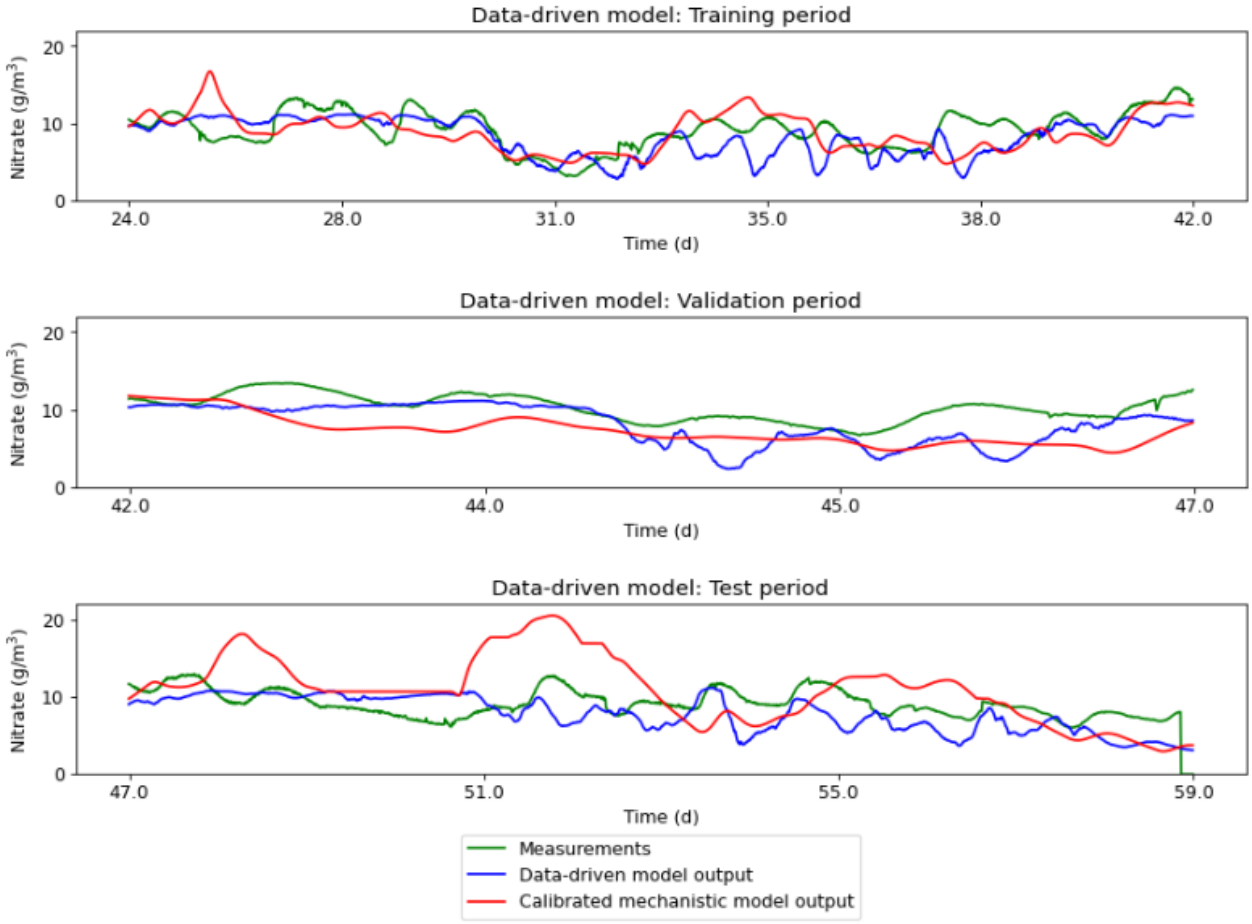

Figure D.2: Data-driven model output, mechanistic model output and measurements for effluent nitrate for the train, validation and test period.

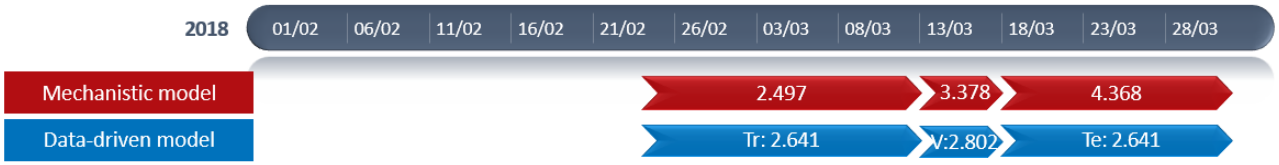

Figure D.3: Timeline with the RMSE per time period (Tr=train, V=validation, Te=test) used to develop a data-driven model for effluent nitrate, compared to the RMSE of the mechanistic model in these periods.

APPENDIX E

LSTM-RNN HYBRID MODEL TRAINED ON UNCALIBRATED MECH. MODEL RESULTS

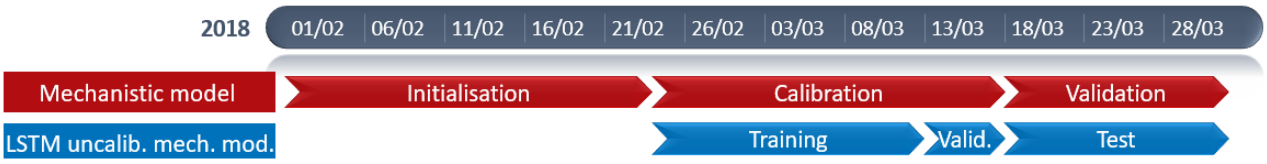

Figure E.1: Timeline with the time periods used to develop a LSTM hybrid model trained on the uncalibrated mechanistic model residual for effluent nitrate.

Table E.1: Final set of hyperparameters determined for LSTM hybrid model trained on the uncalibrated mechanistic model residual for effluent nitrate.

Prediction	**Network**	**Time steps**	**Structure details**	**Learning Rate Adam optimizer**	**Epochs**
NO3	LSTM-RNN	30	Number of layers: 1 Units: 10 Activation: ReLu	0.0005	15

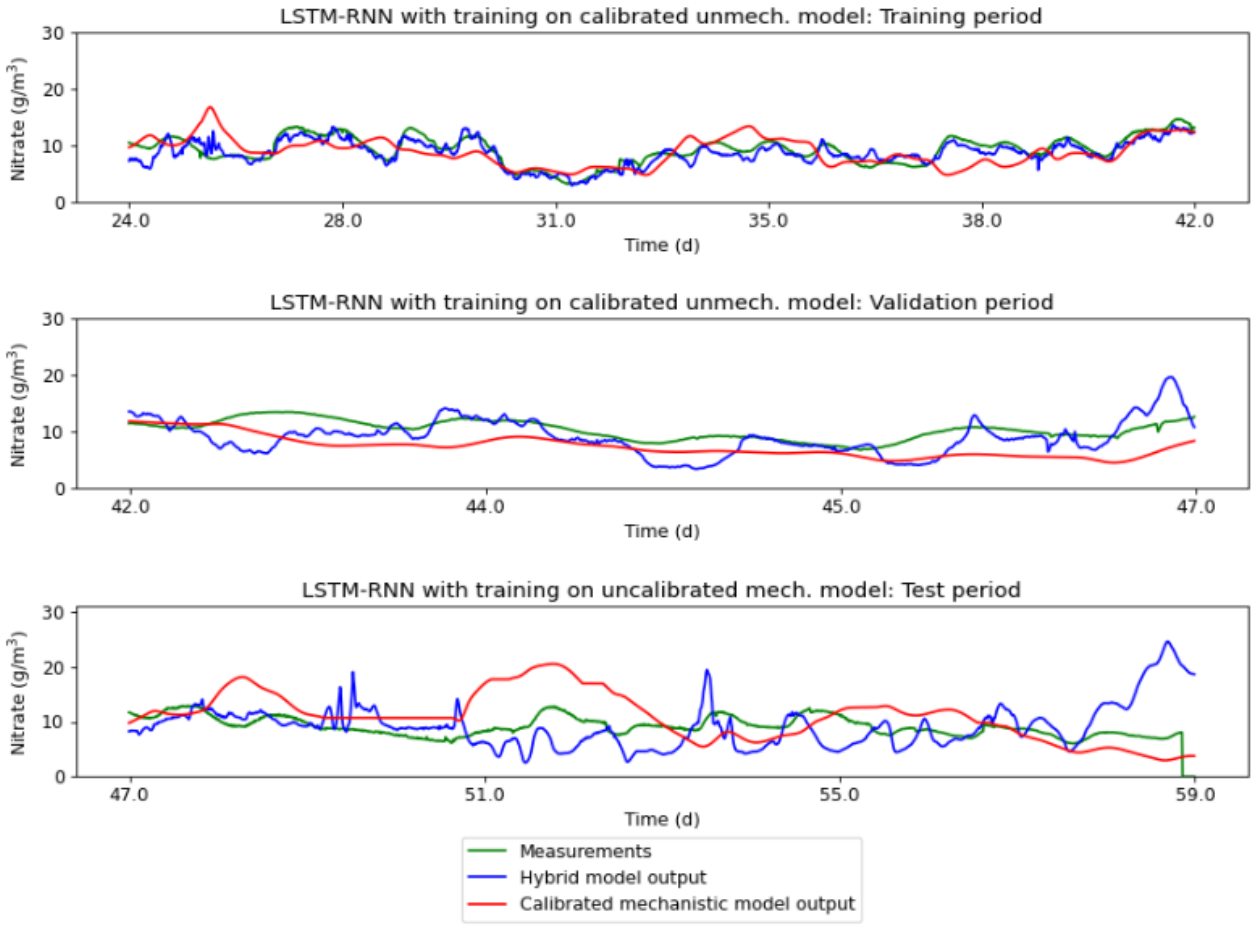

Figure E.2: Hybrid model output with LSTM trained on the uncalibrated mechanistic model residual, mechanistic model output and measurements for effluent nitrate for the train, validation and test period.

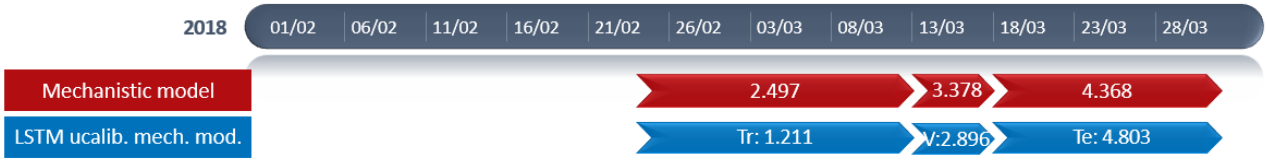

Figure E.3: Timeline with the RMSE per time period (Tr=train, V=validation, Te=test) used to develop a LSTM hybrid model trained on the uncalibrated mechanistic model residual for effluent nitrate, compared to the RMSE of the mechanistic model in these periods.

APPENDIX F

ADDITIONAL HYBRID MODEL OUTPUTS FOR EFFLUENT TSS

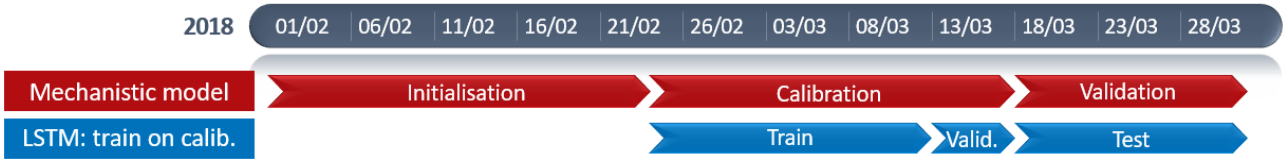

Figure F.1: Timeline with the time periods used to develop an LSTM hybrid model for effluent TSS.

Table F.1: Final set of hyperparameters determined for an LSTM hybrid model for effluent TSS trained on the calibration dataset of the mechanistic model.

Prediction	**Network**	**Time steps**	**Structure details**	**Learning rate Adam optimizer**	**Epochs**
TSS	LSTM-RNN	30	Number of layers: 3 Units: 36, 18, 10 Activation: ReLu, ReLu, ReLu	0.0005	15

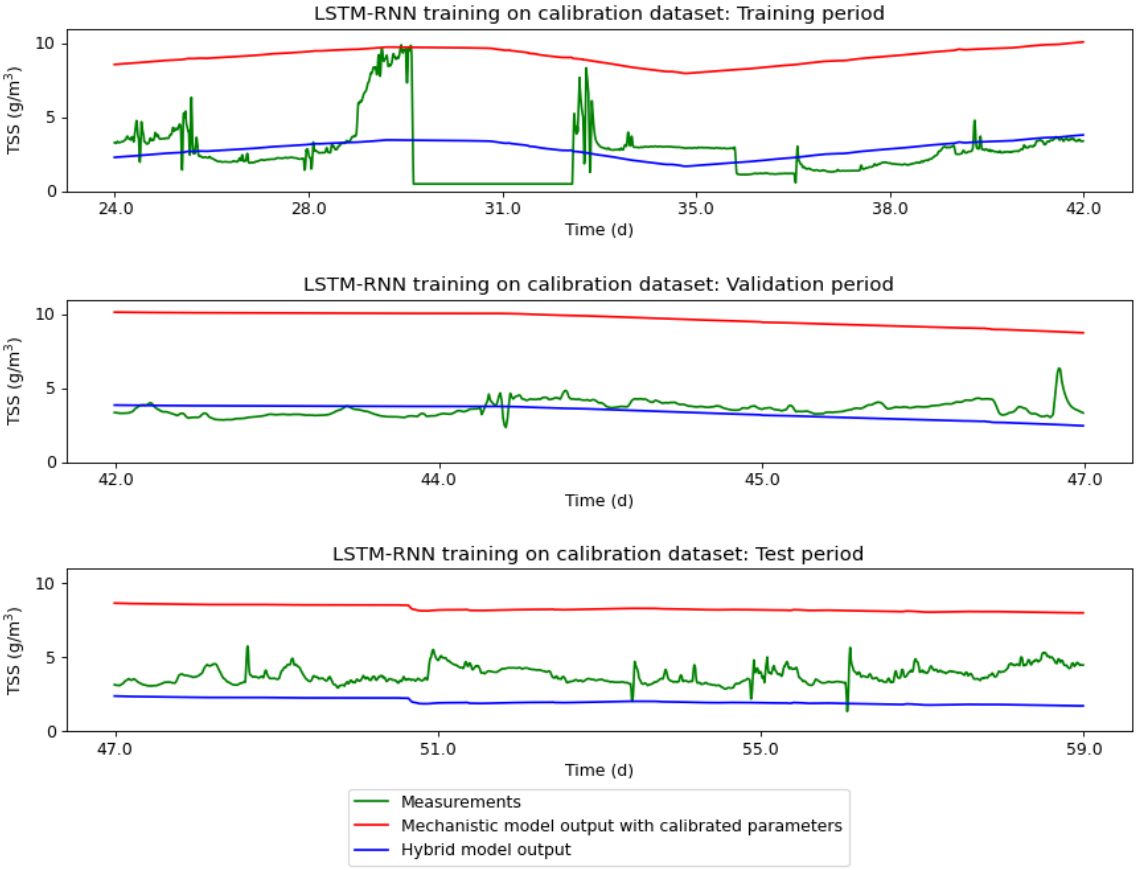

Figure F.2: Hybrid model output with LSTM trained on the calibration dataset of the mechanistic model, mechanistic model output and measurements for effluent TSS for the train, validation and test period.

Figure F.3: Timeline with the RMSE per time period (Tr=train, V=validation, Te=test) used to develop a LSTM hybrid model for effluent TSS, compared to the RMSE of the mechanistic model in these periods.

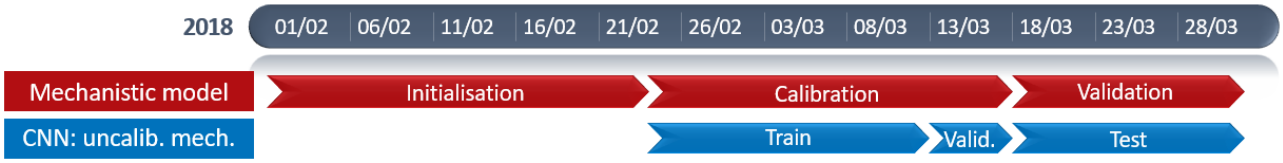

Figure F.4: Timeline with the time periods used to develop a CNN hybrid model trained on the uncalibrated mechanistic model residual for effluent TSS.

Table F.2: Final set of hyperparameters determined for CNN hybrid model trained on the uncalibrated mechanistic model residual for effluent TSS.

Prediction	**Network**	**Time steps**	**Structure details**	**Learning rate Adam optimizer**	**Epochs**
TSS	CNN trained on uncalib. model	30	Number of layers: 4	0.0001	30
			Filter: 30, 9, 6, 3		
			Kernel size: 9, 3, 3, 3		
			Activation: ReLu, ReLu, ReLu, tanh		

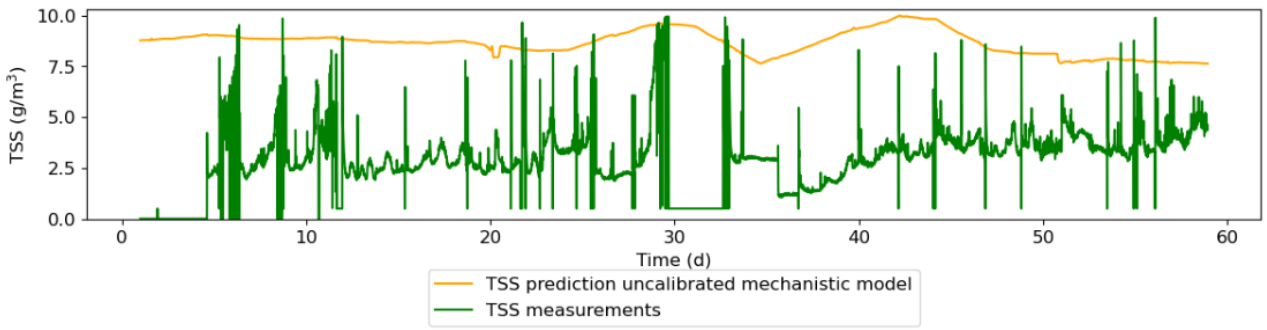

Figure F.5: Uncalibrated mechanistic model output for effluent TSS compared to effluent TSS measurements for the entire modelling period

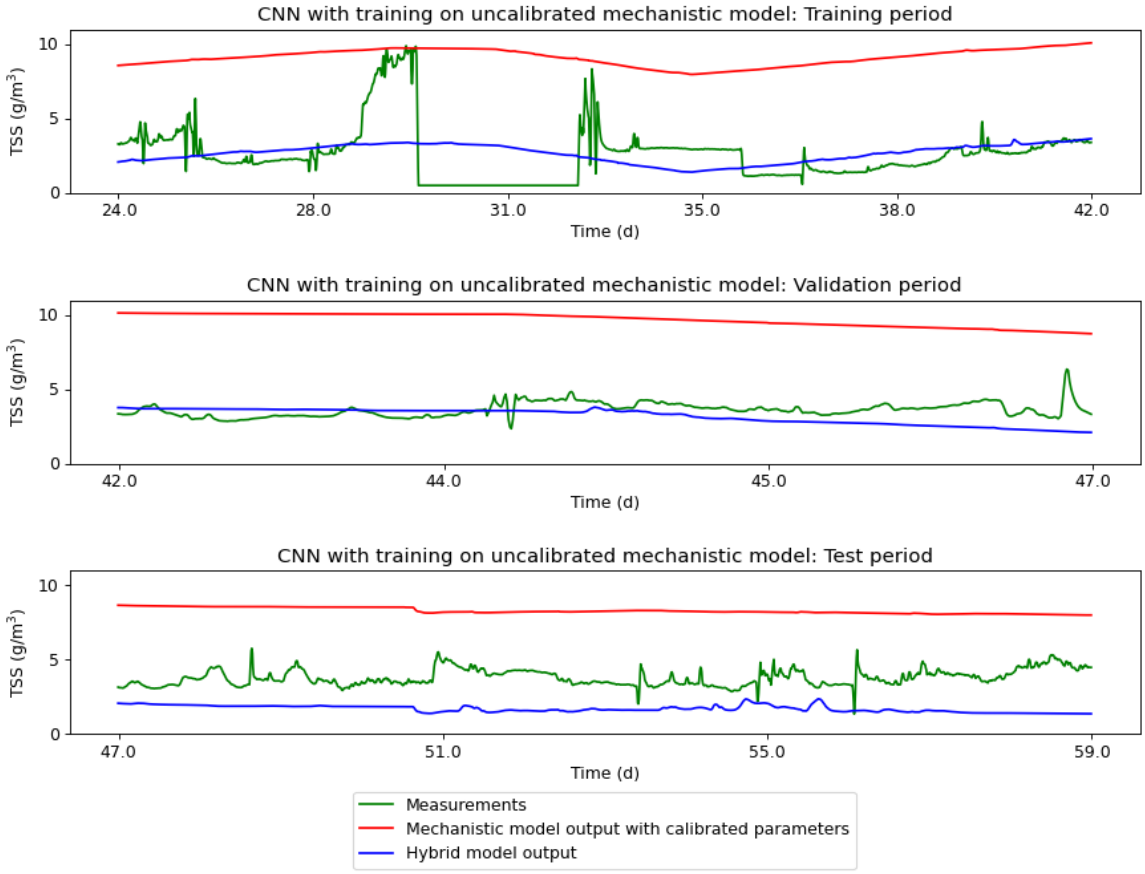

Figure F.6: Hybrid model output with CNN trained on the uncalibrated mechanistic model residual, mechanistic model output and measurements for effluent TSS for the train, validation and test period

Figure F.7: Timeline with the RMSE per time period (Tr=train, V=validation, Te=test) used to develop a CNN hybrid model trained on the uncalibrated mechanistic model residual for effluent TSS, compared to the RMSE of the mechanistic model in these periods.