TREATMENT WETLANDS FOR SECONDARY DOMESTIC WASTEWATER: AN EXPERIMENTAL AND MODELLING APPROACH

by

Mario Alberto Salinas Toledano

B.Eng., Universidad Autónoma Metropolitana, 2021

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF APPLIED SCIENCE IN ENGINEERING

UNIVERSITY OF NORTHERN BRITISH COLUMBIA

December 2024

© Mario Alberto Salinas Toledano, 2024

ABSTRACT

This thesis explores the use of treatment wetlands (TWs) for domestic wastewater (WW), addressing the need for more experimental trials in secondary treatment scenarios and cold climates. Additionally, a data-driven model, the Sparse Identification of Nonlinear Dynamics (SINDy) algorithm, was employed for the first time to identify a nonlinear system of ordinary differential equations (ODEs) to describe pollutant removal rates in TWs. A 2x2x2 full factorial experiment was conducted to examine the effects of WW strength, temperature and vegetation on lab-scale horizontal subsurface flow TWs. High-resolution monitoring of biologically relevant parameters such as oxidation reduction potential (ORP), dissolved oxygen (DO), pH, and temperature was implemented. Experimental datasets were utilized for training and validating the chemical oxygen demand (COD), ammonia and phosphates reaction rate models. Results indicated no significant difference in COD removal between high and low WW strength, indicating TWs can effectively manage raw domestic WW.

Results indicated no significant difference between high and low organic loads, indicating TWs can effectively manage high-strength domestic WW. However, temperature significantly decreased the COD removal (86-96% in warm and 58-76% in cold conditions). Additionally, systems with the largest increase in ORP values exhibited the best COD removal efficiencies. However, not enough oxygenic levels were wealthy enough to enhance ammonia degradation efficiently. The pH also presented a slight increase in the systems, which could affect the phosphate removal by promoting their release in the water. Finally, ORP offered more reliable insights into oxygenic conditions than direct DO measurements. These results underscore the potential of TWs to be implemented for secondary treatment, as well as the importance of continuous monitoring of ORP and pH parameters as just-in-time indicators of TW performance.

ODEs were identified for modelling pollutant removal, with simplified equations employing 3 to 6 polynomial basis functions in low-noise scenarios and up to 11 polynomial basis functions in high-noise situations. Key parameters such as COD*ammonia, ORP, and pH were found to have the most significant impact on the pollutant removal rates. Validation results demonstrated robust predictive capabilities for COD models, achieving R² values between 0.81 and 0.98 with experimental data, while ammonia and phosphate models showed varying accuracy levels. These results suggest the need for effective application of the SNDy algorithm to identify the TWs dynamics and more datasets to be tested for model validation. They also highlight the importance of incorporating ORP and pH into TWs monitoring and modelling.

TABLE OF CONTENTS

ABSTRACT	ii
TABLE OF CON	ΓENTSiv
LIST OF TABLES	Svii
LIST OF FIGURE	ES viii
GLOSSARY	Х
ACKNOWLEDG	EMENTSxi
1 INTRODUCT	FION1
1.1	Study background1
1.2	Treatment wetlands for secondary domestic wastewater
1.2.1	Treatment wetlands classification
1.2.2	Design and operational criteria4
1.3	Data-driven models for treatment wetlands4
1.4	Hypotheses7
1.5	Research objectives7
1.6	Thesis outline7
2 TREATMEN	T WETLANDS FOR SECONDARY TREATMENT OF DOMESTIC
WASTEWATER:	FACTORIAL EXPERIMENT AND HIGH-RESOLUTION DATA
MONITORING	9
2.1	Abstract9
2.2	Introduction10
2.3	Materials and methods12
2.3.1	Treatment wetland systems
2.3.2	Influent preparation14

	2.3.3	Factorial design	
	2.3.4	Sample collection and analysis	16
	2.3.5	Calculations	
	2.4	Results and discussion	
	2.4.1	System performance	
	2.4.2	Evaluation of factors	19
	2.4	.2.1 Significant factors	19
	2.4	.2.2 Interactive effects	23
	2.4.3	Change in pollutants and physical-chemical parameters	
	2.4	.3.1 Chemical oxygen demand soluble (CODs)	
	2.4	.3.2 Ammonia	
	2.4	.3.3 Phosphates	
	2.4	.3.4 Oxidation reduction potential (ORP)	
	2.4	.3.5 Dissolved oxygen (DO)	
	2.4	.3.6 pH	
	25	Conclusions	36
			MODEL
3 D	ISCOVER	ING NONLINEAR DIFFERENTIAL EQUATIONS TO	MODEL
POLL	UTANT C	ONCENTRATION IN TREATMENT WETLANDS	
	3.1	Abstract	
	3.2	Introduction	
	3.3	Methodology	41
	3.3.1	SINDy algorithm	41
	3.3.2	Application of SINDy algorithm to treatment wetlands	
	3.3	.2.1 Data preparation	
	3.3	.2.2 Model training and validation	
	3.4	Results and discussion	46
	3.4.1	Chemical oxygen demand (COD) modelling	
	3.4.2	Ammonia modelling	53
	3.4.3	Phosphates modelling	
	25	Conclusions	(0
	3.3	CONCLUSIONS	

4.1	Research summary	70
4.2	Limitations and future work	71
REFERENCES		

LIST OF TABLES

Table 2.1. Influent characterization (n=8). 15
Table 2.2. Factors and levels of each experiment of the factorial design
Table 2.3. Sampling schedule per batch
Table 2.4. Water analyses, methods and instruments
Table 2.5. Statistical F- and P-value of factors on CODt, ammonia and phosphates removal.22
Table 3.1. Best polynomial degree for the COD trend
Table 3.2. Best hyperparameter value for the COD modelling using SINDy. 47
Table 3.3. Sparse coefficients estimated for CODs modelling (values x10 ⁻³)
Table 3.4. SINDy model validation with COD experimental data
Table 3.5. Best polynomial fit for the ammonia trend
Table 3.6. Best hyperparameter value for the ammonia modelling using SINDy. 54
Table 3.7. Sparse coefficient estimated for ammonia modelling (values x10 ⁻⁵)
Table 3.8. SINDy model validation with ammonia experimental data 60
Table 3.9. Best polynomial fit for the phosphates trend
Table 3.10. Best hyperparameter value for the phosphates modelling using SINDy61
Table 3.11. Sparse coefficient estimated for phosphates modelling (values x10 ⁻⁵)65
Table 3.12. SINDy model validation with phosphates experimental data 68

LIST OF FIGURES

Figure 1.1 Year-round publications between 2000 and 2023, Science Direct. Keywords search
as dark red: "treatment wetland," "constructed wetlands"; orange: "treatment wetland,"
"constructed wetlands," "domestic"; blue: "treatment wetland," "constructed wetlands,"
"domestic," "decentralized."
Figure 1.2 Year-round publications between 2000 and 2023, Science Direct. Keywords
research as blue: "treatment wetland," "constructed wetlands," "modelling"; purple:
"treatment wetland," "constructed wetlands," "data-driven models."
Figure 2.1. Experimental set-up for evaluating the effect of WW strength, temperature and
plants on pollutant removal efficiencies14
Figure 2.2. Average removal efficiencies of CODt, ammonia and phosphates in the different
TWs (N=8 per system)
Figure 2.3. Pareto chart related to the factors evaluated for the removal of a) CODt,
b) ammonia, and c) phosphates (A = WW strength, B = temperature, and C = plants)23
Figure 2.4. The interaction effects of WW strength, temperature and plants for a)CODt,
b)ammonia and c)phosphates25
Figure 2.5. Temperature*WW strength contour plots obtained by Box-Behnken design for the
removal of a) CODt, b) ammonia, and c) phosphates in i)planted and ii) unplanted TWs
Figure 2.6. CODs concentrations in TW systems
Figure 2.7. Ammonia concentrations in TW systems
Figure 2.8. Phosphate concentrations in TW systems
Figure 2.9. ORP values in TW systems
Figure 2.10. Dissolved oxygen values in TW systems

Figure 2.11. pH levels in TW systems throughout the HRT operation time
Figure 3.1 Schematic of SINDy algorithm45
Figure 3.2. SINDy model approximation to COD polynomial curve (training data) for
A)W20P, B)W20U, C)C20P, D)C20U, E)W5P, F)W5U, G)C5P and H)C5U51
Figure 3.3. SINDy model approximation to ammonia polynomial curve (training data) for
A)W20P, B)W20U, C)C20P, D)C20U, E)W5P, F)W5U, G)C5P and H)C5U58
Figure 3.4. SINDy model approximation to phosphates polynomial curve (training data) for
A)W20P, B)W20U, C)C20P, D)C20U, E)W5P, F)W5U, G)C5P and H)C5U66

GLOSSARY

ANOVA	Analysis of Variance
CODt	Chemical Oxygen Demand (total)
CODs	Chemical Oxygen Demand (soluble)
DO	Dissolved Oxygen
FWSF	Free Water Surface Flow
HRT	Hydraulic Retention Time
HSSFW	Horizontal Subsurface Flow Wetland
ORP	Oxidation reduction Potential
ODE	Ordinary Differential Equations
RMSE	Root Mean Squared Error
RK4	Runge-Kutta 4
SINDy	Sparse Identification of Nonlinear Dynamics
SSF	Subsurface Flow
TW	Treatment Wetland
VSSFW	Vertical Subsurface Flow Wetland
WW	Wastewater

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisors, Dr. June Garcia-Becerra and Dr. Ron Thring, for their invaluable guidance, support, and expertise throughout the completion of this thesis. Thanks for your motivation and encouragement.

I would also like to express my appreciation to my committee members, Dr. Deborah Roberts for their insightful feedback on the experiemntal approach and Andy Wang for their guidance on the modelling conceptualization and interpretation.

Additionally, I am grateful to Doug Thompson and Dr. Kenedy Boateng, curators at the Enhanced Forestry Lab, for providing the space and resources for setting up my experiments. I would also like to thank Fryderyk Packowski from the hydrology lab for his cooperation and assistance.

I gratefully acknowledge the financial support from the UNBC, through the Research Project Award, which provided essential funding for this research, and the Research Strategic Initiatives Grant, which enabled me to present my work at prestigious national and international conferences. I also appreciate the additional financial support provided by the Natural Sciences and Engineering Research Council (NSERC), which helped secure the necessary resources for this work.

I would like to thank my lab mates, Erik Groenenberg and Bridget Atta, for their companionship and support during the experimental tests, and the WASH-T research group for the insightful discussions at our biweekly meetings.

Finally, I express my heartfelt thanks to my father, mother, and brothers for their unwavering support throughout this journey.

xi

1 INTRODUCTION

1.1 Study background

Water is a precious resource for living beings. In 2010, the United Nations General Assembly explicitly recognized the human right to water and sanitation, acknowledging that clean drinking water and sanitation are essential to the realization of all human rights (UN-WHO, 2010). It includes ensuring safe, accessible, and affordable ways to manage wastewater (WW). Nevertheless, billions of people worldwide are reported to live without safely managed sanitation (UN-Water, 2021). Globally, 44 % of all household WW flows are not safely treated, i.e. treated by secondary or higher processes or with effluent discharges meeting relevant standards. Under this situation, the deterioration of rivers, lakes, and oceans, as well as public health concerns, have increased (Cosgrove and Loucks, 2015; Hoekstra et al., 2018).

Traditionally, centralized gray infrastructure has been used to supply water and WW services for storage, conveyance, treatment, and disposal. In these systems, WW is collected and transported to treatment facilities far from the generation point (de Anda et al., 2018). Consequently, it is capital and energy-intensive, with long construction cycles, and often incompatible with the environment (Li et al., 2017; Palmer et al., 2015). This prompts a review of alternative blue-green infrastructure technologies, such as treatment wetlands (TWs) for on-site WW treatment (WWAP and UN-Water, 2018).

TWs are nature-based solutions consisting of a waterbed with plants and a filter media layer as its foundation, where physical, chemical and biological processes combine to clean WW (Fu et al., 2018). They are described as a high-efficiency, low-cost, easy-to-operate, and low-maintenance technology, proposing TWs as a feasible solution to face water sanitation issues (Meng et al., 2014; Puigagut et al., 2008; Solano et al., 2004).

1.2 Treatment wetlands for secondary domestic wastewater

The interest in studying TWs, also called constructed wetlands (CW), has increased significantly in the past 10 years. A bibliometric analysis conducted in the Science Direct database indicates an exponential growth of scientific publications in the last seven years (orange in Figure 1.1). Similar trends were found for Google Scholar and Web of Science search engines (not shown). This is partly due to TWs versatility in treating a wide range of WW, such as domestic sewage, urban runoff and stormwater, industrial and agricultural WW, landfill leachate, and polluted river water (Kumar and Dutta, 2019; Saeed and Sun, 2012; Vohla et al., 2011; Wu et al., 2015). However, in the context of domestic WW, the research interest started only a few years ago (green in Figure 1.1), where most TW systems are implemented to perform the tertiary or polishing treatment as the main process occurring in a centralized WW treatment plant. It limits our understanding of their performance as a secondary treatment system. Experimental trials under high pollutant concentrations might help to evaluate the feasibility of TWs for decentralized (i.e. small-scale) domestic WW treatment since only a few studies have been reported in the literature (blue in Figure 1.1). This is especially important in developing countries where the majority of raw WW is discharged directly into rivers due to a lack of suitable and effective treatment technologies, operational failures of larger treatment plants, and higher costs involved in constructing new treatment units (Zhang et al., 2014).



Figure 1.1 Year-round publications between 2000 and 2023, Science Direct. Keywords search as dark red: "treatment wetland," "constructed wetlands"; orange: "treatment wetland," "constructed wetlands," "domestic"; blue: "treatment wetland," "constructed wetlands," "domestic," "decentralized."

1.2.1 Treatment wetlands classification

TWs can be classified according to their hydraulics into free water surface flow (FWSF) and subsurface flow (SSF) (Kadlec and Wallace, 2009). In the FWSF configuration, the water circulates over the saturated substrate through the stems of the plants. The water is directly exposed to the atmosphere, similar to the conventional lagoon systems. In the SSF wetlands, water circulation is through a granular medium and the plant rhizomes and roots. Depending on the direction of the flow, SSF systems are divided into horizontal subsurface flow wetlands (HSSFW) and vertical subsurface flow wetlands (VSSFW) wetlands. In the HSSFW, the water is fed at the top of one end; however, in VSSFWs, the system receives WW at the top along the wetland surface area (Delgadillo et al., 2010).

In particular, SSF systems are the most common types of TWs used around the world since they improve anaerobic and anoxic conditions and interactions between pollutants and soil substrate,

plant roots and microorganisms (García et al., 2010; Kumar and Dutta, 2019). HSSFW is preferred over VSSFW due to its less complex operation and maintenance (del Castillo et al., 2022).

1.2.2 Design and operational criteria

The design of TWs is often carried out using the black box concept. Although numerous manuals are available currently, most of the design and operational recommendations are still considered "rule of thumb" or empirically based on previous experiences (Kadlec and Wallace, 2009; UN-Habitat, 2008; US EPA, 2000). Key factors influencing TW efficiencies include influent quality, temperature, and plant species. However, the interconnections between these factors and their impact on pollutant removal rates are not fully understood (Fu et al., 2018).

In addition, oxidation reduction potential (ORP), dissolved oxygen (DO) and pH are reported to affect the physical, chemical, and biological processes occurring within WW treatment processes, thus they can aid in enhancing the understanding and prediction of removal rates for conventional pollutants (Grimalt-Alemany et al., 2021; Zhai et al., 2012). However, these parameters are not often reported in previous studies and if so, measurements are typically taken only at the influent and effluent of the treatment system, limiting the comprehension of their roles within the TWs.

1.3 Data-driven models for treatment wetlands

The pollutant dynamics, i.e. their transportation, transformation and fate, in TWs have not been elucidated yet, which limits proposing design and operational optimization strategies (Meyer et al., 2015; Yuan et al., 2020). Over time, mathematical modelling has been applied to overcome this limitation. It describes the conceptual understanding of pollutant degradation mechanisms into mathematical expressions (i.e., algebraic or differential equations) (Meyer et al., 2015). As

shown in Figure 1.2, publications that included a modelling approach for studying TWs have increased over time, with most of them involving mathematical models for expanding the understanding of TWs. Nevertheless, these models exhibit some limitations since they are generally based on assumptions; contain parameters that rely on real-time measurement; and often present high complexity due to typically large numbers of equations, kinetic constants and stoichiometric factors (Li et al., 2021; Meyer et al., 2015). Moreover, such models generally do not consider nonlinear interactions for pollutant degradation via biological activity which has been reported (Li et al., 2021).

Recently, data-driven models based on machine learning algorithms have gained attention in simulating complex engineering systems. These models consist of finding relationships between the system state input and output variables without exact knowledge about the physical behaviour of the system (Solomatine et al., 2008). Although data-driven models have been reported to be successfully applied in different fields such as mechanics, hydraulics, cell biology, among others (Brunton et al., 2020; Li et al., 2021; Zitnik et al., 2019), its implementation on TWs started increasing only three years ago with 661 publications in 2023 (purple in Figure 1.2). Hence, exploring data-driven models stands out as a research opportunity to obtain a simplified equation to better understand pollutant dynamics in TWs.



Figure 1.2 Year-round publications between 2000 and 2023, Science Direct. Keywords research as blue: "treatment wetland," "constructed wetlands," "modelling"; purple: "treatment wetland," "constructed wetlands," "data-driven models."

The general objective of this thesis is to support the effective design and operation of TWs for secondary domestic WW treatment by evaluating their performance under different conditions and developing pollutant removal prediction models in the form of ordinary differential equations (ODEs) based on a data-driven model. This approach addresses the complexity of TWs by viewing them as adaptive-evolving systems that respond to operational and design changes such as organic loading, temperature and plant growth. Based on experimental observations, a data-driven model can learn from the ongoing variations in TW processes and identify nonlinear ODEs to describe pollutant dynamics.

This research expects to provide a better understanding of TW performance and propose operational strategies and design recommendations to improve the performance of small-scale TWs for secondary domestic WW. Also, it is expected to provide insights into using data-driven models to study TWs.

1.4 Hypotheses

The hypotheses to be tested in this work are:

- The performance of TWs, in terms of pollutant removal efficiencies, can vary due to factors such as WW strength, temperature, and plants.
- Physicochemical parameters such as ORP, DO and pH can serve as indicators of the pollutant removal rates in TWs.
- Data-driven modelling based on machine learning regression can produce governing nonlinear ODE to describe pollutant removal rates in TWs.

1.5 Research objectives

Based on the main research gaps, the objectives of this thesis are:

- To examine the influence of WW strength, temperature, and the presence of plants in HSSFW treatment efficiency (in terms of conventional pollutants concentration and physicochemical parameters).
- To implement a data-driven model that generates the non-linear governing ODE for pollutant dynamics, emphasizing the interactions between key parameters influencing the degradation of conventional pollutants in HSSFWs.

1.6 Thesis outline

The thesis proposal is composed of four chapters. Chapter one presents the general overview and objectives of the thesis work. Chapter two presents the experimental results, including the effect and the interactions between factors such as WW strength, temperature, and vegetation, as well as changes in physicochemical parameter values to explain pollutant dynamics over the reaction time. Chapter three presents the modelling results where a data-driven model algorithm was implemented using experimental data to elucidate the non-linear ODE for conventional pollutants removal in TWs. Finally, chapter four provides the conclusions and suggests directions for future research.

2 TREATMENT WETLANDS FOR SECONDARY TREATMENT OF DOMESTIC WASTEWATER: FACTORIAL EXPERIMENT AND HIGH-RESOLUTION DATA MONITORING

2.1 Abstract

TWs can be used as decentralized treatment technology by removing pollutants such as organics and nutrients from domestic WW; however, more research is needed under secondary treatment scenarios and cold climates while monitoring biologically relevant parameters. In this work, a 2x2x2 full factorial experiment was implemented to investigate the impact of WW strength (500 and 2,000 mgCODt/L), temperature (6 and 20°C), and vegetation (planted and unplanted) on lab-scale HSSFW, with the high-resolution monitoring of ORP, DO, pH and temperature as biologically relevant parameters. Results indicated no significant difference between high and low organic loads, indicating TWs can effectively manage high-strength domestic WW. However, temperature significantly decreased the COD removal (86-96% in warm and 58-76% in cold conditions). Additionally, systems with the largest increase in ORP values exhibited the best COD removal efficiencies. However, not enough oxygenic levels were wealthy enough to enhance ammonia degradation efficiently. The pH also presented a slight increase in the systems, which could affect the phosphate removal by promoting their release in the water. Finally, ORP offered more reliable insights into oxygenic conditions than direct DO measurements. These results underscore the potential of TWs to be implemented for secondary treatment, as well as the importance of continuous monitoring of ORP and pH parameters as just-in-time indicators of TW performance.

2.2 Introduction

TWs have become a viable and affordable natural option in recent years to treat municipal WW. They have proven effective in a variety of settings and sizes, working both alone and in tandem with conventional methods (Arias et al., 2021). HSSFW are reported to effectively remove organics with efficiencies ranging between 75% and 95% (Ilyas and Masih, 2018) and, to a lesser extent, nutrients (N and P) with 0-75% (Ilyas and Masih, 2018, 2017). Their efficiency can be attributed to their ability to enhance aerobic and anoxic conditions and, thus, promote complex biological and physicochemical processes (García et al., 2010; Kumar and Dutta, 2019). In this sense, HSSFWs show potential as a decentralized secondary WW treatment alternative, especially in scenarios where communities face challenges accessing traditional sanitation systems. However, further experimental tests are needed to explore this potential application fully (Ghimire et al., 2019), including the effect and interactions of operational and design parameters, as well as obtaining high-resolution data on conventional pollutants and more biologically relevant parameters like ORP, DO and pH to explain pollutant dynamics.

The performance of HSSFW can be improved by the proper selection of operational and design parameters such as WW strength, temperature and plant types (Kadlec and Wallace, 2009; Wu et al., 2015). While many studies have shown high efficiencies in treating low-strength WW (50-500 mg CODt/L), there is only a limited number that focuses on high-strength WW (1,000-2,000 mg CODt/L), which levels can also be seen in raw black domestic WW (Gómez-Borraz et al., 2022; Wang et al., 2016). Hence, to our knowledge, no optimal reaction time has been reported at high-strength WW. The impact of temperature on plant and microbial activity has been noted, with reports indicating decreased efficiencies in organic and nutrient removal in cold climates (min. and max. temperatures of -3°C and 10°C, respectively) (Wang et al., 2017). While recent studies show that HSSFWs may be feasible in cold climates, there still exists

uncertainty about how temperature can affect contaminant removal efficiencies and treatment processes (Varma et al., 2021). Plants have been shown to fulfill a crucial role by facilitating nutrient assimilation, oxygenation of the rhizosphere, and creating favourable conditions for biofilm development (Lai et al., 2012). However, this is contradicted by other studies that have reported no significant differences in organics removal compared to unplanted systems (Von Sperling and de Paoli, 2013). From the above-mentioned points, it is noted that TW systems are complex, and parameters such as WW strength, temperature, and plants may interact with each other. However, the interactions among influencing factors in the operation of TW systems remain unknown as they have been studied separately. In this sense, factorial analysis can help better explore and understand such internal interactive effects to propose design and operation strategies (Demir-Duz et al., 2020; Song et al., 2018).

Additional parameters such as ORP, DO, and pH may offer insights into processes occurring under these conditions, given their significant roles in driving both biological and physicochemical removal processes (Ilyas and Masih, 2018; Varma et al., 2021). ORP is a dynamic indicator that reflects the presence of oxidizers or reducers in a liquid and is reported to correlate with biological substances like enzymes, vitamins, and most metabolic processes, as well as concentrations of COD, DO, and NO^{3–}/NH⁴⁺ (Grimalt-Alemany et al., 2021; Li and Bishop, 2002). However, despite its importance in indicating organic degradation, nitrification, and denitrification, research on ORP changes within TWs remains limited. Recent review papers indicate that oxygen limitation is the primary factor limiting biological pollutant removal rates in TWs, whereas neutral pH levels promote nitrification and denitrification processes (Ilyas and Masih, 2018, 2017; Tang et al., 2020a). Nevertheless, studies typically only report values at the influent and effluent, creating a gap in understanding how concentrations change and affect other processes within the system when monitored at high resolution during hydraulic retention time (HRT). While phosphorus removal is reported to be primarily attributed to physical processes rather than biological mechanisms, ORP, DO, and pH are known to influence phosphorus speciation, affecting its potential for adsorption, precipitation, and microbial assimilation (Kadlec, 2016). Overall, ORP, DO and pH levels can be measured online and are directly involved in key pollutant removal processes. As such, they can be used as complementary to traditionally measured parameters such as COD and biological oxygen demand (BOD). This could be important since COD and BOD are measured ex-situ, take from a couple of hours to days to be processed, and do not offer just-in-time insights.

The present work's objective is thus to factorially analyze the performance of HSSFW in the removal of COD, ammonia and phosphates under various operating conditions. The probable interactions between WW strength, temperature and plants were investigated, and the relative impacts of each factor were evaluated using a 2x2x2 factorial design. Also, biologically relevant and just-in-time parameters were monitored to see if they helped follow the operations of the conventional pollutant measurements. The fluctuations within the TWs of ORP, DO, and pH were observed and analyzed in relation to the changes in the concentration of the targeted pollutants using a high-resolution data set. The results aim to better understand the application of TWs for secondary treatment and better monitor performance efficiency.

2.3 Materials and methods

2.3.1 Treatment wetland systems

The study consisted of operating eight identical lab-scale HSSFW prototypes located in a greenhouse and a cold room in the Enhanced Forestry Laboratory at the University of Northern British Columbia campus in Prince George, BC ($53^{\circ}53'8.63''$ N -122°48'29.63'' W). Four systems operated at room-controlled temperatures in a greenhouse at 20 ± 2.5 °C and 12 hours

average of sunlight per day, while the other four were operated in a cold room at 6 ± 1.0 °C and 12 hours of 3000K artificial warm light per day (Figure 2.1). Each TW comprised a rectangular dark-plastic tank (58 cm long, 38 cm wide, and 32 cm deep) filled with 27 cm deep gravel with a working volume of 13 L. The aspect ratio (length: width) of 1:1.5 was selected following the US EPA, 2000 design guidelines. Figure A1 shows the experimental set up in the cold room and greenhouse. The HSSFW systems were divided into three major sections: surface, treatment, and bottom. Their respective substrate further distinguished each section: the surface and bottom layers are associated with a 5 cm layer of course river rock (20 - 25 mm), while the treatment section consisted of a 17 cm layer of gravel substrate (4.75 - 9.5 mm) to increase surface area for biofilm formation. At the center of each TW, a 5 cm diameter PVC pipe with even perforations was placed vertically, functioning as the sampling and measurement port.

The vegetation in the planted systems was *Carex utriculata* acquired in May 2023 from the Western Acres Wetland, a full-scale surface TW that receives the effluent of a lagoon system located southwest of Prince George. These plants have been growing in the Western Acre Wetland since 2001, and it could be assumed that they are adapted to local domestic WW contaminants and weather (from -40 °C in the winter to 30 °C in the summer). After harvesting, the macrophytes were acclimatized to the experimental conditions by exposing the plants to the working temperatures in the greenhouse (20 ± 2.5 °C) and the cold room (6 ± 1.0 °C) fed with low-strength synthetic WW (SWW) (Table 2.1) for two weeks. Six individuals were planted per system.



Figure 2.1. Experimental set-up for evaluating the effect of WW strength, temperature and plants on pollutant removal efficiencies.

2.3.2 Influent preparation

Commercial pet food was used to mimic the complex composition of domestic WW (Picos-Benítez et al., 2017; Sun et al., 2022). In the present study, the Kargol et al., 2023 technique was adapted, wherein dog food was blended instead of soaked, and commercial fertilizer Scotts Turf Builder Lawn Food was added for nitrogen supplementation. The SWW was prepared with a CODt: Ammonia ratio of 10:1, as reported for raw black domestic WW (Gómez-Borraz et al., 2022). The SWW was held in a 60L holding tank situated before the TWs for three days to gain anoxic conditions. The pH of the feeding was adjusted to 6.75 ± 2.5 pH with 5M NaOH. Influent characterization is shown in Table 2.1. The TWs operated in batch mode with a pump delivering a mean volume of 15L of SWW every six days.

Parameter	High-strength WW	Low-strength WW
CODt (mg/L)	1980 ± 107	492 ± 25
CODs (mg/L)	601 ± 94	179 ± 22
Ammonia (mg/L)	193 ± 13.4	48.3 ± 3.3
Nitrates (mg/L)	39.4 ± 11.2	24.7 ± 12.2
Nitrites (mg/L)	< 0.01	< 0.01
Phosphates (mg/L)	54.0 ± 7.1	16.5 ± 3.3
ORP (mV)	-299 ± 63	-258 ± 44
DO (mg/L)	1.8 ± 0.7	1.9 ± 0.7
pH	6.7 ± 0.2	6.6 ± 0.3

Table 2.1. Influent characterization (n=8).

CODt: total chemical oxygen demand; CODs: soluble chemical oxygen demand; ORP: oxide-reduction potential; DO: dissolved oxygen

2.3.3 Factorial design

A full factorial design was adopted to determine the factors influencing the secondary treatment of domestic WW using TWs and to study their interactions. Factors considered in the design were WW strength, temperature, and plants, while the percent removal of COD, ammonia and phosphates were the response variables. The range of values was determined based on previous works (Gómez-Borraz et al., 2022; Kadlec and Reddy, 2001; Wang et al., 2017, 2016): WW strength 500–2000 mg CODt/L, temperature 6–20 °C, and plants as unplanted and planted. In a full factorial design, each factor has 2 levels, minimum and maximum. Eight TWs were designated with the names W20P, W20U, C20P, C20U, W5P, W5U, C5P, and C5U, based on the parameters under which they operated. The designations W or C indicate warm or cold temperature, the number 20 or 5 denotes high- or low-strength SWW intake and the letters P or U specify whether they were planted or unplanted, respectively. Table 2.2 shows the level of the factor for each experiment.

Experiment	WW strength	Temperature	Vegetation
	(mg CODt/L)	(°C)	
W20P	2,000	20	Planted
W20U	2,000	20	Unplanted
C20P	2,000	6	Planted
C20U	2,000	6	Unplanted
W5P	500	20	Planted
W5U	500	20	Unplanted
C5P	500	6	Planted
C5U	500	6	Unplanted

Table 2.2. Factors and levels of each experiment of the factorial design.

2.3.4 Sample collection and analysis

The HSSFW started operating in June 2023, and the monitoring consisted of everyday sampling from August to October 2023 (i.e. eight batches). Water loss due to evapotranspiration was replenished at night to maintain a constant volume during sampling. To increase the data resolution of the parameters' behaviour in the 6-day HRT, samples were collected in triplicate at various times for each batch (Table 2.3). This sampling strategy was designed to maintain an evenly distributed time interval between each data point (3 hours timestep) when staggering.

Batch	Sampling time
1	10 am
2	7 am
3	7 pm
4	10 pm
5	4 pm
6	1 am
7	1 pm
8	4 am

Table 2.3. Sampling schedule per batch

Note: The systems were fed and drained every six days at 4 pm during the study.

The following conventional water quality parameters and the respective standard method used to measure each one was determined from grab samples: total COD, CODt (closed reflux colorimetric: APHA 2017: 5220-D); soluble COD, CODs (UV-Vis absorption: Xu et al., 2023); phosphate (ascorbic acid: APHA 2017: 4500 P-E); ammonium (ammonia selective electrode: APHA 2017: 4500 D); nitrate (cadmium reduction method: APHA 2017: 4500 NO3-B); nitrite (colorimetric method: APHA 2017: 4500 NO2-B); DO (membrane electrode: APHA 2017: 4500 O-G); ORP (Pt electrode method: APHA 2017: 2580-B); pH (electrode: APHA 2017: 2580-B); temperature (probe method: APHA 2017: 2550-B). The CODs, DO, and water temperature were measured *in situ*, while the remaining parameters were measured in the lab the same day after sampling. The COD probe values were contrasted with those obtained through the digestion method to ensure the validation of the equipment (Figure A2). Table 2.4 shows detailed information regarding the equipment and reagents used.

Parameter	Method	Reference	Instrument/reagents
CODt	Closed reflux colorimetric	(APHA, 2017: 5220-D)	COD HACH Digestion vials
CODs	UV-Vis absorption	(Xu et al., 2023)	APURE
Phosphate	Ascorbic acid (adaptation)	(APHA, 2017: 4500 P- E)	HANNA HI 93706-01
Ammonium	Ammonia Selective Electrode	(APHA, 2017:4500 D)	YSI TruLine Ammonium ISE
Nitrate	Cadmium reduction	(APHA, 2017: 4500 NO3-B)	YSI TruLine Nitrate ISE
Nitrite	Colorimetric	(APHA, 2017:4500 NO2-B)	HANNA HI 93707-01
DO	Membrane electrode	(APHA, 2017: 4500 O-G)	Myron Ultrapen PT5
ORP	Pt electrode	(APHA, 2017: 2580-B)	Atlas Scientific ENV-30- ORP
pН	Electrode	(APHA, 2017: 2580-B)	Thermo Scientific STAR A2116
Temperature	Probe	(APHA, 2017: 2550-B)	Myron Ultrapen PT5

Table 2.4. Water analyses, methods and instruments

CODt = total chemical oxygen demand; CODs = soluble chemical oxygen demand; DO= dissolved oxygen; ORP = oxide-reduction potential

2.3.5 Calculations

The removal efficiencies (RE, %) for CODt, ammonia and phosphates were calculated according to Eq. (1):

$$RE = \frac{C_i - C_e}{C_i} x100\% \tag{1}$$

Where C_i and C_e denote the pollutant concentrations in the influent and day six within the HSSFWs, respectively (mg/L). The factorial experimental design and statistical analysis were conducted using MINITAB 21 statistical software.

2.4 Results and discussion

2.4.1 System performance

Varying removal efficiencies for CODt, ammonia, and phosphates, ranging from 63 to 95%, from -7 to 96% and from 46 to 87%, respectively, were found in the TWs. In general, prior studies have also reported elevated removal rates for CODt, 80-95% (Ilyas and Masih, 2017; Wang et al., 2017) and lower removal rates for ammonia, -19-91% (Fountoulakis et al., 2017; Trang et al., 2010; Yadav et al., 2018), and phosphates, 53-85% (Stefanakis and Tsihrintzis, 2009; Zhong et al., 2015).

Figure 2.2 shows the removal of CODt, ammonia and phosphates under different conditions of the factorial design. It can be observed that the TWs exposed to a warmer temperature presented the highest CODt mean removal with 95% for the unplanted fed with 500 mg CODt/L, followed with 94% for the planted fed 2,000 mgCODt/L, then with 93% for the unplanted fed with 2,000 mgCODt/L, and 91% for the planted fed with 500 mgCODt/L. It can be explained by the fact that organic matter is mostly a result of microbial activity that is highly influenced by

temperature (Wang et al., 2017). On the other hand, the planted system fed 500 mgCODt/L at 20°C produced significantly higher mean removal of ammonia and phosphates, with 96 and 87%, respectively. Thus, plant activity plays an essential role in nutrient removal, as reported in previous studies (Wang et al., 2019). Nevertheless, under strong WW conditions (2,000 mg CODt/L and 200 mg ammonia/L), significantly lower removals of ammonia with mean values of 2.0% for the planted at 20°C; 1.0% for the unplanted at 20°C; -2.3% for the planted at 6°C; and -6.6% for the unplanted at 6°C. It might suggest that TWs cannot remove ammonia from strong WW at these conditions. Further discussion regarding the effect and interactions between WW strength, temperature and plants is discussed in section 2.4.2, while the CODs, phosphates, ammonia, ORP, DO, and pH trends over the HRT are presented in section 2.4.3.



Figure 2.2. Average removal efficiencies of CODt, ammonia and phosphates in the different TWs (N=8 per system).

2.4.2 Evaluation of factors

2.4.2.1 Significant factors

The analysis of variance (ANOVA) resulting from the factorial analysis of each pollutant is shown in Table 2.5. Significant interaction is presented at P<0.05. It can be observed that only

linear interactions are significant for CODt, while for ammonia and phosphates, 2-way interactions are presented. The residual plots presented in Figure A3 demonstrate a normal distribution as the points were approximately along a straight line (Song et al., 2018). The Pareto chart of effects (Figure 2.3) was also used to distinguish all components' main influences and interactions. In each Pareto chart, all standard effects (SE) were arranged in decreasing order, and the column length showed their significant levels.

The findings suggest that temperature (F=554, p=0, SE=24) as a factor exerts the greatest independent influence on CODt removal, whereas plants also contributed significantly at a lesser degree (F=5.8, p=0.019, SE=2.5). These two factors emerged as the sole significant contributors to the CODt removal. Higher temperatures and the absence of plants were favourable for organic degradation in TWs. This might be explained since the microbial activity and metabolic rate in the TW systems at the lower temperature were reduced, which significantly impeded the heterotrophic bacteria in decomposing organic pollutants (Ji et al., 2020). Plants are reported to enhance carbon oxidation by introducing oxygen into the rhizosphere (Březinová & Vymazal, 2015). However, certain plant species may not release enough oxygen to facilitate organic reduction notably. In addition, plants release exudates, predominantly organic in nature, which could contribute to the measured COD content (Kantawanichkul et al., 2009). This might be the situation with the C. utriculata species utilized in this study. Similar observations have been reported by (Von Sperling and de Paoli, 2013) where unplanted systems exhibited lower COD concentrations along the TW compared to planted systems with Typha latifolia. The fact that the WW strength was not a significant factor affecting CODt removal (F=1.09, p=0.302, SE=1.0) suggests that TWs could efficiently treat high-strength WW, thus being effective for secondary treatment.

In terms of ammonia removal, the three factors, including both independent and 2-way interactions, presented significant effects. WW strength (F=46.9, p=0, SE=6.92) and temperature (F=45.7, p=0, SE=6.88) presented the greatest influence on ammonia removal, followed by the interaction between these two factors, WW strength*temperature (F=30.8, p=0, SE=5.56). It indicates that stronger WW (i.e. higher ammonia concentrations) limited ammonia removal. Kantawanichkul et al., 2009 evaluated the performance of vertical SSFW under high ammonia concentrations ($290 \pm 16 \text{ mg NH}_3\text{-N/L}$). They also observed high concentrations in the effluent after the treatment using *Cyperus involucratus* ($163 \pm 16 \text{ mg NH}_3\text{-N/L}$), *Typha* angustifolia ($110 \pm 26 \text{ mg NH}_3\text{-N/L}$) and unplanted ($126 \pm 68 \text{ mg NH}_3\text{-N/L}$). Furthermore, it is widely recognized that higher temperatures promote ammonia oxidation, nitrification, and denitrification processes, which constitute the primary mechanisms for ammonia removal (Gersberg et al., 1986; Tang et al., 2020b). Finally, a lower but still significant effect was presented by plants (F=13.77, p=0, SE=3.8) and its interaction with temperature (F=11.8, p=0.001, SE=3.45) and WW strength (F=10.0, p=0.002, SE=3.12). It is documented that plants primarily enhance nitrogen removal through the adsorption of solutes, uptake of inorganic nitrogen compounds, and release of oxygen and root exudates to influence microbial activity and diversity (Kantawanichkul et al., 2009; Tang et al., 2020b)

The most important parameters affecting phosphate removal were WW strength (F=21.5, p=0, SE=4.55) and plants (F=21.0, p=0, SE=4.50). It indicates that higher phosphate concentrations limit its removal, and the plants significantly promote it. This finding aligns with previous research, such as that conducted by Körner & Vermaat, 1998, which indicated that plants and microbes could remove up to 75% of phosphorus from which macrophytes were found to uptake around 52%, with the remaining portion removed by associated organisms and microorganisms. Plants stood out as a highly important factor affecting phosphates since their interaction with

temperature (F=15.62, p=0, SE=3.9) and WW strength (F=4.71, p=0.034, SE=2.1) presented a significant effect, too. Also, it mainly indicates that the effect of WW strength and temperature on plant activity drives phosphate removal. This can be explained by considering that the strength of WW (i.e., the level of toxicity resulting from the presence of pollutants in high concentrations) and temperature influence plant physiology. Consequently, these two parameters may constrain root processes involved in phosphorous uptake and promote diversity among microorganisms, including phosphate-accumulating organisms (Ilyas and Masih, 2018; Lu et al., 2023).

Table 2.5. Statistical F- and P-value of factors on CODt, ammonia and phosphates removal

Response	CODt removal		Ammonia removal		Phosphate removal	
Source	F-Value	P-Value	F-Value	P-Value	F-Value	P-Value
WW strength	1.09	0.302	46.96	0	21.55	0
Temperature	554.9	0	45.75	0	0.19	0.663
Plants	5.8	0.019	13.77	0	21.03	0
WW strength*temperature	1.21	0.275	30.83	0	2.15	0.148
WW strength*plants	0.69	0.409	10.04	0.002	4.71	0.034
Temperature*plants	1.55	0.219	11.79	0.001	15.62	0

Note: bold numbers indicate significant factors.



Figure 2.3. Pareto chart related to the factors evaluated for the removal of a) CODt, b) ammonia, and c) phosphates (A = WW strength, B = temperature, and C = plants).

2.4.2.2 Interactive effects

Interaction and contour plots were obtained to assess the interactive effects of the factors. Figure 2.4 illustrates interaction plots depicting the combined effects of three factors (WW strength, temperature, and plants) on the removal of CODt, ammonia, and phosphate. These plots demonstrate how the behaviour of individual factors changes in the presence of others and highlight their level of independence. The slope of the line is proportional to the magnitude of the factor's importance. Figure 2.5 shows the contour plots where the effect of the three factors is presented. Distinctly coloured regions represent different levels of removal efficiency corresponding to the specified values for each parameter.

The WW strength*temperature interaction was only significant for ammonia removal, and the interaction plot (Figure 2.4.b) shows a negative slope in relation to these factors. It indicates higher removal rates at low-strength WW, where the effect of temperature is evident as significantly higher removal is presented at warmer temperatures. Although WW strength*temperature interaction was not significant for phosphates, similar trends are presented regarding WW strength; however, no significant difference (p>0.05) is presented for temperature. In terms of removing CODt, the interaction plots of WW strength and temperature reveal no significant difference between high- and low-strength WW, as evidenced by nearly identical slopes (slope \approx 0). Supporting the statement that TWs can effectively treat secondary raw WW. However, distinct variations emerge when different temperatures are applied, particularly with significantly higher removal rates observed at warmer temperatures.

The WW strength*plants interaction was significant for ammonia and phosphates, and similar trends were observed in their interaction plots (Figure 2.4. b and c). For both, plants significantly improve their removal when having low-strength WW. However, by increasing the concentration of pollutants, there is no significant difference in the removal either by having or not plants. In terms of removing CODt, the interaction plots of WW strength and plants reveal no significant difference between high- and low-strength WW, neither by having nor not having plants.

The temperature*plants interaction was significant for ammonia and phosphates, and similar trends are observed in their interaction plots (Figure 2.4. b and c). For both, there is no significant difference at low temperatures, either by using plants or not. However, at high temperatures, planted systems increase the removal significantly. In terms of removing CODt, the interaction plots of temperature and plants reveal no significant difference between planted and unplanted systems but significantly higher removal at warmer temperatures.



Figure 2.4. The interaction effects of WW strength, temperature and plants for a)CODt, b)ammonia and c)phosphates
Contour plots were obtained to show the areas with high and low removal at different WW strength levels, temperatures and plants. For CODt removal, linear plots are shown since there are no significant 2-way interactions for its removal (Figure 2.5.a). Similar plots are presented for planted and unplanted systems, and the removal increases mainly as the temperature does and not towards a change in the SWW strength. For ammonia removal, curves are observed since 2-way interactions are significant (Figure 2.5.b). Planted systems present higher removal efficiencies (up to >80%), while unplanted systems present up to >40%. Although different plots are shown for those planted and unplanted systems, the highest removal is presented when lower WW strength and higher temperatures are presented in both cases. The unplanted systems plot presents zones (shown in blue) where null removal is presented, indicating that temperatures below 8 °C and WW strength concentration above 1,800 mgCOD/L affect the performance of the systems for its removal. Finally, different plots are presented for phosphate removal for planted and unplanted systems (Figure 2.5.c). Planted systems presented higher phosphate removal, particularly under optimal conditions like warmer temperatures and reduced pollutant concentrations. Conversely, in unplanted systems, observations indicate that lower temperatures can lead to increased removal of phosphates due to the enhancement of adsorption and precipitation processes (He et al., 2023). Studies have also revealed that at warmer temperatures, the release of phosphorus in the water increases (Sarkar et al., 2017).



Figure 2.5. Temperature*WW strength contour plots obtained by Box-Behnken design for the removal of a) CODt, b) ammonia, and c) phosphates in i)planted and ii) unplanted TWs

2.4.3 Change in pollutants and physical-chemical parameters

2.4.3.1 Chemical oxygen demand soluble (CODs)

Figure 2.6 shows the concentration of CODs over the HRT in the TWs after staggering the data. A comparable decreasing trend was observed in the eight systems, albeit to a different magnitude. At the beginning of the batches, the CODs for the high- and low-strength WW systems were 580 ± 18 mg/L and 155 ± 8 mg/L, respectively. On the first day of operation, the COD values in the high-strength WW systems dropped sharply to 220 ± 20 mg/L at a similar rate in both the warm and cold systems, highlighting the performance of TW to achieve comparable performance at the beginning of the reaction time. In contrast, this significant decrease occurred within the first 18 hours in the low-strength WW systems, reaching a concentration of 35 ± 5 mg/L. At this stage, the DO in the system may be almost depleted, as it has been consumed by heterotrophic bacteria during the removal of CODs (Rehman et al., 2017). Thereafter, the CODs values experienced a slight decline before entering a plateau stage. After the third day, the high-strength WW systems in the cold room showed a clear difference in concentration compared to the warmer systems, achieving $200 \pm 5 \text{ mg/L}$ and $100 \pm 10 \text{ mg/L}$, respectively. For the low-strength WW systems, those at cold temperatures reached similar final concentrations of 55 ± 5 mg/L for unplanted fed 500 mgCODt/L and 52 ± 3 mg/L for the planted fed 500 mgCODt/L, indicating no significant role of the plants in the system. In warm systems, the lowest final concentration was achieved by unplanted fed 500 mgCODt/L with 35 ± 5 mg/L, followed by the planted fed 500 mgCODt/L with 43 ± 5 mg/L. At higher concentrations, the impact of temperature on COD depletion becomes more evident, suggesting the need to install an insulating system, such as mulch, over the wetlands. However, no major modifications are necessary for treating low-strength wastewater, as no significant variation was observed in the effluent concentration after 6 days of HRT. As the concentration stabilizes after four days in

high-strength WW systems, a HRT of four days could be proposed for secondary treatment purposes. For low-strength WW, an HRT of two days would be sufficient.



Figure 2.6. CODs concentrations in TW systems

2.4.3.2 Ammonia

The concentrations of ammonia within the TWs are shown in Figure 2.7. High-strength WW systems exhibited similar trends, with initial concentrations at 205 ± 5 mg/L. The concentrations increased slightly over the first three days, peaking at 242 ± 10 mg/L, before gradually decreasing over the remaining days. The initial increase in concentrations may result from the breakdown of organic matter, leading to the release of ammonia as a byproduct (Lamers et al., 2012). Subsequent declines could be attributed to microbial activity, where ammonia is utilized as a substrate for nitrification, converting it into nitrate (Lamers et al., 2012; Zhu et al., 2014). The final concentrations for high-strength WW was 185 ± 9 mg/L for planted at 20 °C; 179 ± 8 mg/L for unplanted at 20 °C; 172 ± 7 mg/L for planted at 6 °C; and 155 ± 7 mg/L for unplanted at 6 °C.

In contrast, low-strength WW systems showed fluctuating concentrations without a consistently decreasing trend, with a final value of $45 \pm 3 \text{ mg/L}$ for the unplanted system at 6 °C, $45 \pm 8 \text{ mg/L}$ for unplanted at 6 °C and $40 \pm 7 \text{ mg/L}$ for unplanted at 20 °C. Only the planted system fed with 500 mg/L at 20 °C exhibited a decrease similar to the CODs, with a rapid decline on the first day followed by a slower decrease, reaching a final concentration of 3 mg/L. This highlights the importance of plant activities, which promote nutrient assimilation and provide favourable conditions for microbial activity (Li et al., 2018). Nitrate concentrations were also measured, showing fluctuations, but a downward trend is observed across all systems (Figure A4). This suggests that anoxic conditions were predominant, allowing the denitrification process to occur at a higher rate than nitrification. Specifically, it will be after the third day of operation.



Figure 2.7. Ammonia concentrations in TW systems

2.4.3.3 Phosphates

Figure 2.8 illustrates the phosphate concentrations in the TWs. For the high-strength WW systems, the concentrations decreased gradually over the six-day period, except for the

unplanted system fed with 2,000 mg CODt/L at 20 °C, where the decrease halted by the second day, and then the concentrations stabilized. The final concentrations achieved by high-strength WW systems were 22 ± 3 mg/L for the planted at 20 °C; 30 ± 4 mg/L for the unplanted at 20 °C; 24 ± 6 mg/L for the planted at 6 °C; and 19 ± 4 mg/L for the unplanted 6 °C. Decreased phosphate concentrations might indicate effective plant uptake and microbial assimilation (Li et al., 2020). Stabilizing phosphate concentrations in unplanted fed with 2,000 mg/L at 20 °C could be attributed to factors such as limited microbial activity or substrate availability (Li et al., 2020). The low-strength WW systems exhibited a slighter continuous decrease throughout the operation period, with the lowest concentration observed in the planted at 20 °C with 1 ± 3 mg/L, followed by the planted at 6 °C with 5 ± 3 mg/L; the unplanted at 6 °C with 7 ± 2 mg/L; and the planted at 20 °C with 9 ± 3 mg/L, respectively. The lower final concentrations observed in low-strength WW and planted systems compared to unplanted underscore the significant role of vegetation in enhancing nutrient removal efficiency. The presence of plants promotes nutrient assimilation and provides favourable conditions for microbial activity, leading to more pronounced reductions in phosphate concentrations (Ilyas and Masih, 2018).



Figure 2.8. Phosphate concentrations in TW systems

2.4.3.4 Oxidation reduction potential (ORP)

ORP reflects the balance between oxidized and reduced substances within the TW environment, influencing the availability and transformation of nutrients and contaminants (Ilyas and Masih, 2018; Zhai et al., 2012). The value of ORP in the high and low-strength influent was -299 \pm 63 and -258 \pm 44 mV, respectively, both representing anaerobic conditions and in the range of raw domestic WW (Foladori et al., 2013, 2012; Ong et al., 2009). In the warm systems, ORP levels remained relatively steady initially, registering -300 \pm 20 mV for high-strength WW systems and -160 \pm 20 mV for low-strength WW systems (Figure 2.9). Subsequently, ORP levels exhibited a gradual increase, reaching -30 \pm 20 mV in the planted system fed 2,000 mg CODt/L at 20 °C; -90 \pm 18 mV in the unplanted system fed 2,000 mg CODt/L at 20 °C; 30 \pm 8 mV in the planted system fed 5,000 mg CODt/L at 20 °C; and -28 \pm 11 mV in the unplanted system fed 5,000 mg CODt/L at 20 °C. These levels were sustained on days 5 and 6. The progressive rise in ORP within the TWs indicated the development of favourable conditions for completely biodegradable COD oxidation, either by using O₂ or by exploiting NO₃ (Zhai et al., 2012; Zhu et al., 2014). Studies have noted that the ascending ORP profile promotes enzymatic activities such as heightened dehydrogenase and catalase activities, along with intensified carbon degradation (cellulose and β -glucosidase activity), nitrogen cycling (urease activity), and phosphorus cycling (phosphatase activity). This leads to more efficient removal of various pollutants (Truu et al., 2009; Wießner et al., 2005; Zhou et al., 2005).

Conversely, systems operating at colder temperatures exhibited minimal fluctuation in ORP levels throughout the treatment process, maintaining values within the range of -280 to -220 mV. This sustained low ORP, indicating low oxygen conditions, can account for the notably lower removals of COD, ammonia, and phosphates observed in cold temperatures compared to warm systems. In environments with low ORP, microorganisms utilize oxygen from oxidized pollutants, leading to a slower breakdown and subsequent decrease in COD (Wießner et al., 2005).



Figure 2.9. ORP values in TW systems

2.4.3.5 Dissolved oxygen (DO)

DO is vital for supporting aerobic microbial activity, which plays a significant role in the breakdown and biodegradation of organic pollutants and nitrogen compounds present in TWs (Zhai et al., 2012). Figure 2.10 illustrates the DO concentrations in the TWs. The TWs exposed to warmer temperatures initially exhibited dissolved oxygen (DO) concentrations of 1.3 ± 0.3 mg/L, while those kept in colder conditions recorded 2.8 ± 0.6 mg/L. Systems operating at warmer temperatures showed a gradual increase, reaching an average of 2.0 ± 0.3 mg/L across all treatment processes, with a notable higher increment observed in planted systems.

Conversely, systems under cold temperatures experienced a slight rise in DO levels during the first two days, levelling off thereafter at 3.5 ± 0.4 mg/L. Overall, the observed increase in DO across all systems could be influenced by several factors that may affect its accuracy. In warmer systems, the presence of emergent vegetation, such *as Cyperus sp.*, likely contributed to enhanced oxygen transfer through root arenchyma (Rehman et al., 2017), while in colder systems, the increase in DO can be attributed to their ability to retain more oxygen due to lower temperatures and reduced evaporation rates (Varma et al., 2021). However, it has been reported that the direct measurement of DO in the matrix, particularly in deeper soil layers where concentrations tend to be lower, can be challenging and subject to variability (Li et al., 2016a). In comparison, ORP has been shown to correlate more strongly with microbial activity indicators, such as population density and oxygen uptake rate, making it a useful alternative for assessing oxidative-reductive processes where oxygen consumption is involved (Kadam et al., 2009; Li et al., 2016a; Wang et al., 2007). ORP measurements were, therefore, used in place of direct DO measurements to provide additional insights into the system's oxygen dynamics.



Figure 2.10. Dissolved oxygen values in TW systems

2.4.3.6 pH

The pH values within the TWs over the 6 days HRT are shown in Figure 2.11. Influent pH values ranged between 6.6-7.0. For the eight systems, the pH increased gradually over the operation time. By the end of the batch, the unplanted system fed with 2,000 mgCODt/L at 20 °C reached the highest pH at 7.6, followed by both the unplanted system fed with 2,000 mgCODt/L and the planted system fed with 500 mgCODt/L at 20 °C with 7.3; the planted and unplanted fed with 500 mgCODt/L at 6 °C reached 6.9; the unplanted fed with 2,000 mgCODt/L at 20 °C and the planted 500 mgCODt/L at 6 °C reached 6.6; and finally the planted system fed with 500 mgCODt/L at 20 °C and the planted fed with 2,000 mgCODt at 6 °C reached 6.4. It indicates that the TWs worked in a neutral environment. This range of values is reported to favour nitrification processes (Grinberga et al., 2022). W20U presented the highest final pH and phosphates concentration (Figure 2.8). This suggests that an increase in pH affects the phosphate adsorption capacities in TWs, as reported in previous studies (Bai et al., 2017). Fluctuations in the values for the warm systems might be due to the fluctuations in temperature (Figure A5).



Figure 2.11. pH levels in TW systems throughout the HRT operation time.

2.5 Conclusions

The performance of TWs in removing COD, ammonia, and phosphates was evaluated, focusing on the effects of WW strength, temperature, and vegetation. High CODt removal and no significant difference was found for WW strength in CODt removal, highlighting the technology's potential for implementing as secondary treatment technology receiving black domestic WW characteristics. Higher temperatures were most influential for CODt removal, promoting organic degradation, while unplanted systems showed higher removal rates than planted ones. ORP levels are suggested as an indicator of COD level fluctuations within the system. In warm systems, the increase in ORP levels corresponds to a decrease in COD concentration, which is due to a more oxidative environment. Ammonia removal was generally low across most systems, except those with low-strength WW, warm temperatures, and vegetation. Further research is necessary to improve ammonia removal under conditions of high-strength WW and low temperatures, with an emphasis on promoting plant activity, as indicated by this study's findings highlighting the significant importance of plants. Improving aerobic conditions to enhance nitrification rates is suggested. Phosphate removal was mainly influenced by WW strength and plant presence, with their interactions with temperature enhancing efficiency. The results imply that elevated temperatures and rising pH levels reduce the potential for phosphate adsorption. Ultimately, ORP proved to offer more reliable insights into oxygenic conditions than direct DO measurements, particularly in deeper layers or sampling locations.

This study offers valuable insights into key parameters for measurement (such as ORP, DO, and pH) that reflect the system's performance and the potential of TWs to treat high-strength domestic WW. Further research is needed to gain deeper mathematical insights into the role of these parameters and their interactions in affecting pollutant concentrations.

3 DISCOVERING NONLINEAR DIFFERENTIAL EQUATIONS TO MODEL POLLUTANT CONCENTRATION IN TREATMENT WETLANDS

3.1 Abstract

Treatment wetlands (TWs) harbour complex physical, chemical and biological processes that may interact with each other to efficiently remove pollutants from wastewater (WW). Up to date, there are limitations on the design and operation of TWs due to a lack of accurate and interpretable models. To the best of our knowledge, this study is the first to apply a data-driven approach, specifically utilizing the Sparse Identification of Nonlinear Dynamics (SINDy) algorithm (Brunton et al., 2016), to derive model equations corresponding to TWs under varying WW strengths, temperatures, and plants. Using experimental data, the models reveal the interactions between biologically relevant parameters (ORP, DO, pH) and conventional pollutants (COD, ammonia, and phosphates) to predict pollutant concentrations over the reaction time. Nonlinear systems of differential equations were identified for modelling pollutant removal, with simplified equations employing 3 to 6 polynomial basis functions in low-noise scenarios and up to 11 polynomial basis functions in high-noise situations. Key parameters such as ORP and the interaction between COD and ammonia were identified as critical for estimating COD and ammonia concentrations, while pH and ORP were found to have the most significant impact on phosphates. Validation results demonstrated robust predictive capabilities for COD models, achieving R² values between 0.81 and 0.98, while ammonia and phosphate models showed varying accuracy levels. The results suggest the need for improved data quality and highlight the importance of incorporating ORP and pH into pollutant modelling.

3.2 Introduction

TWs are engineering systems able to remove pollutants as they provide the environment for diverse removal processes, including sedimentation, precipitation, adsorption, microbial degradation and plant uptake (Kadlec and Wallace, 2009). These processes rely on the water's physicochemical properties, plant growth, microbial composition, WW composition, and nature of the TW substrate (media) (Garcia et al., 2010). Such complexity makes it challenging to model pollutant patterns using stand-alone physical or chemical equations, leaving the conserved quantities that govern the system's dynamics unknown. First-order chemical reactions, commonly used for designing TWs, do not consider interactions between multiple pollutants and the physicochemical environment, failing to match experimental data (Nguyen et al., 2018; Ventura et al., 2022). This leads to a limited understanding of pollutant behaviour and impedes the optimization of TWs designs. Instead, alternative modelling tools, such as data-driven models, have been introduced and might applied to describe TWs performance.

Data-driven models have been shown to effectively establish relationships between pollutant removal rates and the TW characteristics, requiring minimal understanding of the underlying removal processes (Gupta et al., 2021; Jafarzadeh et al., 2022). For example, (Nguyen et al., 2021) applied data-driven modelling, based on machine learning algorithms, such as Random Forest, Cubist, Support Vector Machines, and K-nearest Neighbors to predict ammonia concentrations in subsurface TWs. While these models achieved high correlation with the training data (R² values ranging from 0.80 to 0.91 and RMSE values between 0.9 and 1.9 mg/L), none of them offered a simple output as system identification, i.e., the model's structure or equation that describes the pollutant behaviour. Alternatively, regression analysis can be used as a simpler data-driven approach to achieve this. Singh et al., 2022 employed multiple linear regression (MLR) to develop a model predicting biological oxygen demand (BOD), chemical

oxygen demand (COD), Ammonium-Nitrogen, Total Nitrogen, and Total Phosphorus in the effluent, using independent variables such as hydraulic retention time (HRT), temperature, flow rate, depth, area, and influent concentrations. This resulted in equations with 11 variables, yielding from low to high accuracy (R² ranging from 0.47 to 0.90). Limitations of MLR include the assumption of linearity and the risk of overfitting, as it requires assigning a coefficient to each input variable. Additionally, the need to assign a coefficient to each input variable increases the risk of overfitting, especially when there are many variables, as the model may become too tailored to the training data and perform poorly on unseen datasets. When modelling TWs, no previous studies have incorporated physicochemical parameters such as ORP, DO and pH within TWs to derive a reaction equation for pollutants in TWs.

In 2016, Brunton et al., 2016 reported the Sparse Identification of Nonlinear Dynamics (SINDy) algorithm to identify the governing equations from time series data. SINDy is based on sparse regression (i.e. assigning zero coefficients to less important variables), aiming to promote model interpretability and avoid overfitting. It consists of selecting a reduced set of nonlinear functions from a library of basis functions that represent potentially relevant variables and determining their coefficients using the least absolute shrinkage and selection operator (LASSO) regression as the sparsity-inducing regularization. Recently, SINDy has gained popularity in different fields such as economy, physics, chemistry, and biology, showing promising results, but its use has primarily focused on generic data, limiting its application to real-world data (Abdullah and Christofides, 2023; Hoffmann et al., 2019; Loiseau, 2020; Mangan et al., 2016; Prokop and Gelens, 2024; Sandoz et al., 2023). To the best of our knowledge, SINDy has not been tested on WW systems, but previous studies suggest the algorithm's potential to identify pollutant reaction rates in TWs.

This study aims to implement the SINDy algorithm to derive a nonlinear system of ODEs that describe the dynamics of COD, ammonia, and phosphates in TWs. Also, from these model structures, it will be able to identify key physicochemical parameters and their interactions with pollutants that promote the removal of the pollutants. The data-driven models were trained and validated using experimental time series data, as described in Chapter Two, from TWs under various operational conditions (WW strength, temperature, and plants). The resulting eight models (one for each TW) per pollutant were used to find relationships between the system state input and output variables, considering that the exact knowledge about the system's physical behaviour is unknown. These models were developed using the SINDy algorithm to estimate the mathematical equations of the physical system using the post-processed input data based on polynomial fitting. With comprehensive training and validation, the models were able to generate a set of non-linear ODEs representing the system dynamics. The equations were then integrated over time using the Runge-Kutta 4 (RK4) numerical method and the initial values.

3.3 Methodology

In this section, the data-based model identification method is briefly reviewed. The governing process dynamical system under study is represented as

$$\frac{d}{dt}\boldsymbol{x}(t) = \boldsymbol{f}(\boldsymbol{x}(t)) \tag{1}$$

where the state of the vector $\mathbf{x}(t)$ denotes the states of the system at time t and the nonlinear function $\mathbf{f}(\mathbf{x}(t))$ represents the governing equations of motion of the system. In this study, the SINDY method is applied to identify a candidate function \mathbf{f} .

3.3.1 SINDy algorithm

In this section, a brief overview of the SINDy method is presented. For more information, the readers can refer to the original paper by (Brunton et al., 2016). Given a time-series data of the state $x(t) = [x_1(t) x_2(t) \dots x_n(t)]^T$ where the time difference between each data point is the same, they can then be arranged into one large matrix, shown below.

$$X = \begin{pmatrix} x_1(t_1) & x_2(t_1) & \dots & x_n(t_1) \\ x_1(t_2) & x_2(t_2) & \dots & x_n(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ x_1(t_m) & x_2(t_m) & \dots & x_n(t_m) \end{pmatrix}$$
(2)

Next, the derivatives of the state is numerically computed, producing a matrix shown as follows:

$$\dot{\mathbf{X}} = \begin{pmatrix} \dot{\mathbf{x}}_{1}(t_{1}) & \dot{\mathbf{x}}_{2}(t_{1}) & \dots & \dot{\mathbf{x}}_{n}(t_{1}) \\ \dot{\mathbf{x}}_{1}(t_{2}) & \dot{\mathbf{x}}_{2}(t_{2}) & \dots & \dot{\mathbf{x}}_{n}(t_{2}) \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \dot{\mathbf{x}}_{1}(t_{m}) & \dot{\mathbf{x}}_{2}(t_{m}) & \dots & \dot{\mathbf{x}}_{n}(t_{m}) \end{pmatrix}$$
(3)

After obtaining the derivatives of the state variables, the collected time-series data of X are utilized to build a candidate function library containing all possible non-linear functions as:

$$\theta(X) = \begin{bmatrix} 1 & X & X^2 & X^3 & \dots & \sin(X) & \cos(X) & \dots \end{bmatrix}$$
 (4)

The choice of selecting the potential functions can be based on the knowledge of physics and prior information about the system. There is significant flexibility in constructing the entries in this matrix since only the regression will reveal those that represent the process dynamics of the system. The regression problem can be expressed as

$$\dot{\mathbf{X}} = \boldsymbol{\Theta}(\boldsymbol{X})\boldsymbol{\Xi} \tag{5}$$

where \dot{X} denotes the time-series of derivatives of state variables, $\boldsymbol{\theta}(X)$ is the library of potential functions representing the system dynamics, and $\boldsymbol{\Xi}$ is the vector containing the function coefficients. In this sense, $\boldsymbol{\Xi}$ can be expressed as $\boldsymbol{\Xi} = [\xi_1 \ \xi_1 \ \dots \ \xi_n]$ where each column $\boldsymbol{\xi}_k$ represents a sparse vector of coefficients obtained by a convex l₁-regularized (LASSO) regression. This optimization problem seeks to minimize the sum of the squared residuals and an l₁ regularization term to produce a sparse solution for the coefficient vector Ξ . By combining the squared error term $\|\Theta \Xi - \dot{X}\|_2^2$ with the l₁ norm penalty $\alpha \|\Xi\|_1$, it balances accurate data fit with sparsity, encouraging certain coefficients to be exactly zero. It is expressed as follow

$$\underset{\Xi_{\gamma}\in\mathbb{R}^{n}}{\operatorname{Min}}\|\boldsymbol{\Theta}\boldsymbol{\Xi}-\dot{\mathbf{X}}\|_{2}^{2}+\boldsymbol{\alpha}\|\boldsymbol{\Xi}\|_{1}$$
(6)

where α is the hyperparameter that establishes the "strength" of the regularization. In other words, if alpha is zero, all features are considered, and it is equivalent to linear regression, where only the residual sum of squares is computed to build a predictive model. On the other hand, as alpha closes to infinity, it eliminates more and more features. Thus, it is important to set an appropriate value for α . A common way to determine the best α is by hyperparameter tuning, which is based on picking the α value that produced the lowest error between the model and the training data. In order to evaluate the accuracy of the obtained model, the root mean squared error (RMSE) of the predicted and true values is calculated. RMSE is shown as follows

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$
(7)

where n is the number of observations, y_i is the actual value and \hat{y}_i is the predicted value.

3.3.2 Application of SINDy algorithm to treatment wetlands

3.3.2.1 Data preparation

Data was obtained from eight experimental TW systems studied for domestic WW treatment under different WW strengths (500 and 2,000 mg CODt/L), room temperature (6 and 20 °C)

and vegetation (unplanted and planted). Chapter 2 describes the experimental setup, sampling strategy and analytical methods in more detail. The sampling yielded 48 observations for 8 parameters on each experimental TW. Parameters included COD, ammonia, nitrates, phosphates, ORP, DO and pH. Each data point has a time difference of 3 hours between each other. The objective of this dataset is to infer the model on the removal rate of COD, ammonia and phosphates during the working HRT of 6 days based on the other parameter values. The experimental data was split into training and validation sets in the following steps. A random 20% of the data was set aside for model validation, while the remaining 80% was used to train the model.

3.3.2.2 Model training and validation

Using the training data, a polynomial regression was employed to model the trends of each parameter. The best polynomial fit was determined through a tuning process in Python. The obtained polynomial regression was used to generate simulation time-series data (n=48) to identify the governing dynamics of the process. This approach was taken to reduce inherent noise in the raw data, as directly modelling the raw data could risk overfitting and compromise the trend analysis.



Figure 3.1 summarizes the SINDy algorithm from a given pollutant concentration (COD as C, ammonia as N and phosphates as P) time-series data. The regression coefficients are calculated for those possible interactions that describe the pollutant dynamics.



Figure 3.1 Schematic of SINDy algorithm

The derivative of the time series data was computed using the finite difference method RK4, a numerical approach that approximates the derivative by applying discrete difference quotients between successive data points. For creating the *O* matrix contained up to second-degree polynomial. It includes the concentration of pollutants (C, N, NO: nitrates, and P) and the physicochemical parameters (pH, DO, and ORP) based on their relevance to explaining TW performance, identified in Chapter 2.

To get the $\boldsymbol{\xi}$ sparce vector for each parameter, the LASSO regression was computed using the convex optimization problems (CVXPY) package in Python. The variable coefficients in $\boldsymbol{\xi}$ with values below the thresholding parameter are set to zero, and the regression problem is solved iteratively until the coefficients converge. The hyperparameter α , which is crucial for setting an optimal thresholding value, and, thus, eliminating unwanted functions, was optimized through a tuning process in Python by iterating over values in the range [10⁻¹, 10⁵].

The model validation was performed by calculating the RMSE between the predicted values derived from the obtained differential equations and the observed data. This was done using the validation dataset, which was excluded from the model's training process. The predicted and observed values were then plotted against each other to evaluate the R² coefficient. A strong linear correlation between these values indicates a close approximation to the actual values.

3.4 Results and discussion

This section presents the results obtained for model identification of the TW dynamics using experimental data. The identified model is evaluated based on the key parameters identified in each model and the prediction accuracy obtained from experimental data.

3.4.1 Chemical oxygen demand (COD) modelling

Table 3.1 shows the best polynomial fit for COD, indicating the degree of the polynomial where the RMSE between the modelled and training data was at its minimum. For all systems, the best polynomial fit falls between the 4th and 6th order, supporting the hypothesis that pollutant dynamics in TWs are nonlinear. The high RMSE value for the unplanted system fed 2,000 mgCODt/L at 6 °C (136 mg/L) shows significant noise in the data for this dataset. The remaining systems present lower RMSE values, ranging from 6 to 33 mg/L.

System	Best polynomial degree	RMSE (mg/L)
W20P	6	14
W20U	5	33
C20P	5	23
C20U	4	136
W5P	6	11
W5U	5	11
C5P	5	6
C5U	5	6

Table 3.1. Best polynomial degree for the COD trend.

The hyperparameter α influences the degree of sparsity observed (Goharoodi et al., 2018), meaning that different values of α lead to different resulting models. Table 3.2 presents the best α for each system and their corresponding RMSE values resulting from the tuning process. The α values vary significantly across different systems, reflecting the influence of data complexity (e.g., polynomial degree) and noise (data points distribution). Higher α values, like 150 for the planted system fed 2,000 mgCODt/L at 6 °C, might be linked to more complex or noisy datasets, requiring stronger regularization to avoid overfitting, whereas smaller values, such as 0.5 for the unplanted system fed 500 mgCODt/L at 6 °C, might be associated with cleaner data where minimal regularization suffices. The planted systems generally require higher α values than the unplanted ones, suggesting that plants may introduce additional complexity to the data sets. This may be explained by the roles of plants in TWs such as providing surface area for microbial growth on their roots, oxygenating the rhizosphere, and releasing exudates that promote organic breakdown and nutrient uptake, leading to more complex dynamics in COD concentrations (Rahi et al., 2020).

System	Best-hyperparameter value	RMSE (mg/L)	
W20P	80	47	
W20U	40	38	
C20P	150	25	
C20U	1	36	
W5P	11	13	
W5U	51	7	
C5P	1	11	
C5U	0.5	10	

Table 3.2. Best hyperparameter value for the COD modelling using SINDy.

Table 3.3 presents the resulting model coefficients obtained by solving equation 6 using each system's optimal hyperparameter, α . A matrix can be observed where the majority of the potential candidates' coefficients are zero. The most simplified model was found for the planted

system fed 500 mgCODt/L at 20 °C, which included only 5 out of 27 potential candidates. In contrast, planted system fed 2,000 mgCODt/L at 6 °C had the highest number of possible candidates, with 11 variables used to describe the model. From all the potential interactions, ORP stands out as a significant parameter in all the systems models. This underscores the significance of ORP in estimating COD concentrations and aligns with previous research that highlighted the correlation between the gradual rise in ORP values and microbial activity and abundance in subsurface treatment systems (Li et al., 2016a), as well as conventional wastewater treatment technologies (Zhao et al., 2023). The strong interaction between COD and ammonia (C*N) in all the warm systems and the planted system fed 2,000 mgCODt/L at 6 °C suggests that ammonia availability directly supports the microbial processes responsible for organics removal, especially at warm temperatures when higher microbial activity is presented. This dependence may be attributed to the role of organic carbon as an essential energy source for nitrifying bacteria, which enhances their capacity to conduct nitrification (Yang et al., 2018). The C*ORP interaction is significant in all planted systems, supporting that microbial metabolic processes and redox conditions govern organic matter breakdown. In planted systems, where root exudates provide a source of organic carbon to stimulate microbial activity and stabilize redox conditions, the relationship between COD and ORP becomes increasingly important. (Truu et al., 2015). Finally, pH is observed only in the cold temperature systems, interacting with COD in the unplanted system fed 500 mgCODt/L at 6 °C, with ammonia in the planted system fed 2,000 mgCODt/L at 6 °C, and with ORP in both the planted and unplanted system fed 2,000 mgCODt/L at 6 °C. pH may primarily affect cold temperature systems, where variations impact microbial activity and enzymatic processes. In contrast, warm systems typically have more adaptable microbial communities, making them less sensitive to pH fluctuations and allowing for stable performance (Varma et al., 2021).

It can be noted that the unplanted system fed 2,000 mgCODt/L at 6 °C got the coefficients with the highest values (up to $-3,857.9 \times 10^{-3}$). This dataset exhibited the highest RMSE during the polynomial regression fit, suggesting a high level of noise. This could result in overfitting, even with the application of LASSO regularization.

• Planted system fed with 2,000 mgCODt/L at 20 °C

$$\frac{dC}{dt} = 23.7C + 56.40RP - 0.2C * N + 5.2C * DO + 4.7P * ORP + 3.60RP^2 [x10^{-3}]$$

• Unplanted system fed with 2,000 mgCODt/L at 20 °C

$$\frac{dC}{dt} = 2.27C + 22.40RP - 0.1C * N + 0.1C * ORP + 14.1C * DO - 0.2N * DO [x10-3]$$

• Planted system fed with 2,000 mgCODt/L at 6 °C

$$\frac{dC}{dt} = 22.10RP - 2C^2 + 0.8C * N + 1.1C * P - 0.1C * ORP + 1.3C * DO - 31.6N * DO + 4.4N * pH + 7.8P * ORP + 4.80RP * pH + 5.20RP^2 [x10^{-3}]$$

• Unplanted system fed with 2,000 mgCODt/L at 6 °C

$$\frac{dc}{dt} = -25.7C - 241N - 3040RP - 0.7N^2 + 176.3N * pH + 39.80RP * pH + 3.20RP^2 - 3,858DO^2 [x10^{-3}]$$

• Planted system fed with 500 mgCODt/L at 20 °C

$$\frac{dC}{dt} = 2.30RP - 0.2C^2 - 0.4C * N - 0.5C * ORP + 0.5N * ORP \quad [x10^{-3}]$$

• Unplanted system fed with 500 mgCODt/L at 20 °C

$$\frac{dC}{dt} = 4.6N - 1.20RP + 0.4C^2 + 1.6C * N + 0.8N^2 - 0.1N * ORP - 0.50RP * DO [x10^{-3}]$$

• Planted system fed with 500 mgCODt/L at 6 $^{\circ}$ C

$$\frac{dc}{dt} = 0.50RP + 0.4C^2 + 9.2C * P + 6.3C * 0RP - 17.4N * P - 4.4N * 0RP + 6.1P^2 + 22.2P * 0RP + 0.20RP^2 [x10^{-3}]$$

• Unplanted system fed with 500 mgCODt/L at 6 $^{\circ}$ C

$$\frac{dC}{dt} = 40.6C + 5.40RP + 0.1C * ORP + 7.4C * DO + 1.1C * pH + 23.8N * DO + 5.70RP * DO [x10-3]$$

Table 3.3. Sparse coefficients estimated for CODs modelling (values x10⁻³)

System	W20P	W20U	C20P	C20U	W5P	W5U	C5P	C5U
С	23.7	2.27		-25.7				40.6
Ν				-241.4		4.6		
Р								
ORP	56.4	22.4	22.1	-303.7	2.3	-1.2	0.5	5.4
DO								
pН								
C^2			-2.0		-0.2	0.4	0.4	
C*N	-0.2	-0.1	0.8		-0.4	1.6		
C*P			1.1				9.2	
C*ORP		0.1	-0.1		-0.5		6.3	0.1
C*DO	5.2	14.1	1.3					7.4
C*pH								1.1
N^2				-0.7		0.8		
N*P							-17.4	
N*ORP					0.5	-0.1	-4.4	
N*DO		-0.2	-31.6	176.3				23.8
N*pH			4.4					
\mathbf{P}^2							6.1	
P*ORP	4.7		7.8				22.2	
P*DO								
P*pH								
ORP*DO						-0.5		5.7
ORP*pH			4.8	39.8				
ORP ²	3.6		5.2	3.2			0.2	
DO*pH								
DO^2				-3,857.9				

Note: grey areas represent coefficients with a value of zero.

Figure 3.2 presents the model fit obtained using the SINDy algorithm compared to the polynomial curve that served as the source of the training data. The model successfully captures the exponential decay trends of the COD throughout the 6-day HRT across various systems. The model presents a strong correlation for all the TWs with R² values of 0.98 or 0.98 for all the cases. The RMSE ranges from 11.30 to 16.29 mg/L for high-strength WW and from 3.92 to 5.54 mg/L for low-strength WW.



Figure 3.2. SINDy model approximation to COD polynomial curve (training data) for A)W20P, B)W20U, C)C20P, D)C20U, E)W5P, F)W5U, G)C5P and H)C5U



Figure 3.2. (continuation) SINDy model approximation to COD polynomial curve (training data) for A)W20P, B)W20U, C)C20P, D)C20U, E)W5P, F)W5U, G)C5P and H)C5U

Table 3.4 displays the R2 and RMSE to evaluate the model validation with experimental data. R² values range from 0.81 to 0.98, indicating a strong correlation between the observed and predicted COD concentrations. The RMSE ranges from 32.3 to 45.8 mg/L for high-strength WW and from 5.2 to 14.0 mg/L for low-strength WW.

	1	
System	\mathbb{R}^2	RMSE (mg/L)
W20P	0.93	40.7
W20U	0.92	39.8
C20P	0.95	32.3
C20U	0.94	45.8
W5P	0.81	14.0
W5U	0.98	5.2
C5P	0.93	8.7
C5U	0.87	12.4

Table 3.4. SINDy model validation with COD experimental data

3.4.2 Ammonia modelling

Table 3.5 shows the best polynomial fit for ammonia values. For all systems, the optimal polynomial fit falls between the 2nd and 6th order, indicating that ammonia dynamics in TWs are nonlinear. The high RMSE values for the planted system fed 2,000 mgCODt/L at 20 °C (106 mg/L) and the unplanted system fed 2,000 mgCODt/L at 20 °C (386 mg/L) suggest significant noise in the data for these two datasets. The remaining systems present lower RMSE values, ranging from 11 to 37 mg/L.

System	Best-polynomial fit	RMSE (mg/L)
W20P	3	106
W20U	3	386
C20P	4	28
C20U	4	37
W5P	6	32
W5U	3	11
C5P	2	20
C5U	2	18

Table 3.5. Best polynomial fit for the ammonia trend.

Table 3.6 presents the best α for each system after the tuning process. Ammonia hyperparameter values are considerably higher, ranging from 320 to 40,000, whereas COD values are much lower, spanning from 0.5 to 150. This suggests that ammonia dynamics are more complex than

those of COD, as ammonia data sets require stronger regularization. The planted system fed 2,000 mgCODt/L at 20 °C, which has a hyperparameter value of 1,000, results in a higher RMSE of 38 mg/L, indicating challenges in accurately predicting ammonia levels. In contrast, the planted system fed 500 mgCODt/L at 6 °C system achieves the lowest RMSE of 6 mg/L.

System	Best-hyperparameter value	RMSE (mg/L)
W20P	1,000	38
W20U	577	30
C20P	400	26
C20U	40,000	17
W5P	9,450	14
W5U	670	8
C5P	455	6
C5U	320	7

Table 3.6. Best hyperparameter value for the ammonia modelling using SINDy.

Table 3.7 presents the resulting model structure obtained by solving equation 6 using each system's optimal hyperparameter, α . The most simplified model was found for W5U, which included only 3 out of 35 potential candidates. In contrast, planted system fed 2,000 mgCODt/L at 6 °C had the highest number of possible candidates, with 11 variables used to describe the model. The COD², ORP, and pH² functions affect ammonia concentration in all systems, indicating a strong dependence of ammonia on these parameters. This may suggest a potential cometabolism with COD or a dependency on the rate of COD consumption, considering that COD is oxidized before bacteria can utilize oxygen for the nitrification process (Nivala et al., 2013). Additionally, ORP and pH reflect the environmental conditions under which nitrification rates occur. Although ORP has been reported as an indicator of potential nitrification processes in conventional WW treatment technologies and TWs, the optimal ranges for this parameter vary. (Fuerhacker et al., n.d.) identified an optimal range between -33 and +290 mV, while (Weißbach et al., 2018) found the optimal ORP for nitrification to be between 200-260 mV,

ensuring sufficient oxygen availability. The interaction between ammonia and ORP (N*ORP) stands out in six out of eight systems, highlighting the relationship between the gradual rise in ORP and the microbial activity and abundance of nitrifiers and denitrifiers in TWs (Li et al., 2016a). Regarding the high dependence on pH, (Zhao et al., 2023) reported that nitrification-denitrification rates are highly sensitive to pH changes, with optimal activity occurring within a narrow range of 7.20–7.50. A significant decline in activity was observed when the pH rose above 8.0 or fell below 7.00. In all systems receiving high-strength WW, ammonia is consistently present as N² function throughout. It indicates a nonlinear relationship, where higher ammonia concentrations significantly influence system dynamics, such as microbial activity and nutrient removal efficiency. This suggests that small changes in operational conditions may significantly influence ammonia removal at high ammonia concentrations.

It is observed that the pH coefficients have the highest values, ranging from 29.0 to 2,993.1. This can be explained by the fact that pH values are between 6.5 and 8.0 in all cases, whereas other parameters, such as COD, ammonia, and ORP, vary in the hundreds. Therefore, to include the impact of pH within the model, its coefficient should reflect a comparable influence to that of other parameters. Additional data preparation, such as the standardization process, might help avoid this issue. It is important to note that standardizing data sets does not alter the parameters displayed by the model; it only changes the coefficient values.

• Planted system fed with 2,000 mgCODt/L at 20 °C

$$\frac{dN}{dt} = C^2 - C * N + 2N^2 + N * ORP + ORP^2 - 2,993pH^2 \qquad [x10^{-5}]$$

• Unplanted system fed with 2,000 mgCODt/L at 20 °C

$$\frac{dN}{dt} = 0.08C^2 - 0.8C * N - 0.02C * NO + 0.63C * ORP + 0.52N^2 + 0.02N * NO + 0.28N * ORP - 0.23NO * ORP + 0.68ORP^2 - 2,016pH^2 [x10^{-5}]$$

• Planted system fed with 2,000 mgCODt/L at 6 °C

$$\frac{dN}{dt} = -0.7C^2 + 3.8C * N - 4.2C * NO + 0.5C * ORP - 1.3N^2 - 3.7N * NO - 0.6N * ORP - 6.8NO * ORP - 11.6P * ORP + ORP^2 - 3,161pH^2 [x10^{-5}]$$

• Unplanted system fed with 2,000 mgCODt/L at 6 °C

$$\frac{dN}{dt} = C^2 + 3.8C * N + 1.3N^2 + 0.10RP^2 - 4,171pH^2 \quad [x10^{-5}]$$

• Planted system fed with 500 mgCODt/L at 20 °C

$$\frac{dN}{dt} = -2.9C^2 + 2.7N * ORP - 0.90RP^2 + 29pH^2 \qquad [x10^{-5}]$$

• Unplanted system fed with 500 mgCODt/L at 20 °C

$$\frac{dN}{dt} = 0.5C^2 + 4.6N^2 + 0.30RP^2 - 561pH^2 \qquad [x10^{-5}]$$

• Planted system fed with 500 mgCODt/L at 6 °C

$$\frac{dN}{dt} = 4.5C^2 + 2.4C * ORP - 17.4N * P - 0.4N * ORP - 0.8NO * ORP + 0.2NO * DO + 0.20RP^2 - 368pH^2 [x10^{-5}]$$

• Unplanted system fed with 500 mgCODt/L at 6 $^{\circ}$ C

$$\frac{dN}{dt} = 1.3C^2 + 0.6C * ORP - 0.4N * ORP + 0.8NO * ORP - 2NO * DO + 0.10RP^2 - 229pH^2 [x10^{-5}]$$

System	W20P	W20U	C20P	C20U	W5P	W5U	C5P	C5U
С		-				-	-	
Ν								
NO								
Р								
ORP								
DO								
pН								
C^2	1.0	0.08	-0.7	1.0	-2.9	0.5	4.5	1.3
C*N	-1.0	0.8	3.8					
C*NO		-0.02	-4.2					
C*P								
C*ORP		0.63	0.5				2.4	0.6
C*DO								
C*pH								
N^2	2.0	0.52	-1.3	1.3		4.6		
N*NO		0.02	-3.7					
N*P							-17.4	
N*ORP	1.0	0.28	-0.6		2.7		-0.4	-0.4
N*DO								
N*pH								
NO^2								
NO*P								
NO*ORP		-0.23	-6.8				-0.8	0.8
NO*DO							0.2	-2.0
NO*pH								
\mathbf{P}^2			_					
P*ORP			-11.6					
P*DO								
P*pH								
ORP ²	1.0	0.68	1.0	0.1	-0.9	0.3	0.2	0.1
ORP*DO								
ORP*pH								
DO^2								
DO*pH								
pH ²	-2993.1	-2016.2	-3161.3	-4170.6	29.0	-561.2	-368.0	-228.5

Table 3.7. Sparse coefficient estimated for ammonia modelling (values x10⁻⁵)

Note: grey areas represent coefficients with a value of zero.

Figure 3.3 illustrates the fit of the model derived from the SINDy algorithm alongside the polynomial curve. The model captures the initial rise in ammonia, followed by a continuous decline in all the systems receiving high-strength WW, as well as the unplanted system fed 500 mgCODt/L at 20 °C. In the planted system fed 500 mgCODt/L at 20 °C, an exponential decay trend similar to those observed for COD is identified. As the systems that received low-strength WW at 6 °C showed no significant decrease in concentration, the model only reflects a slight decline over the HRT. The model presents a strong correlation for all the TWs with R² values ranging from 0.75 to 1.00. The RMSE is variable and ranges from 3.72 to 17.53 mg/L for high-strength WW and from 0.98 to 4.33 mg/L for low-strength WW.



Figure 3.3. SINDy model approximation to ammonia polynomial curve (training data) for A)W20P, B)W20U, C)C20P, D)C20U, E)W5P, F)W5U, G)C5P and H)C5U



Figure 3.3. (Continuation) SINDy model approximation to ammonia polynomial curve (training data) for A)W20P, B)W20U, C)C20P, D)C20U, E)W5P, F)W5U, G)C5P and H)C5U

Table 3.8 presents the fit between the experimental data and the predicted values using the models generated by the SINDy algorithm. R² values range from 0.27 to 0.86, and the RMSE ranges from 15.1 to 20.8 mg/L for high-strength WW and from 3.8 to 6.6 mg/L for low-strength WW. The lowest performance accuracy was observed for planted system fed 500 mgCODt/L at 6 °C (R² = 0.27, RMSE = 3.8 mg/L) and unplanted system fed 2,000 mgCODt/L at 6 °C (R² = 0.32, RMSE = 16.8 mg/L). In contrast, the planted system fed 500 mgCODt/L at 20 °C

demonstrated strong prediction accuracy, with an R² of 0.6 and an RMSE of 6.55 mg/L. This is likely due to the arranging of the data, as the planted system fed 500 mgCODt/L at 20 °C shows a clear exponential decay trend in the raw data. In contrast, the other systems exhibit fluctuations throughout the 6-day HRT. In common practice, planted system fed 500 mgCODt/L at 20 °C conditions are typically represented, making this model more suitable for testing and application in evaluating the performance of HSSFW under similar conditions.

System	R ²	RMSE (mg/L)
W20P	0.57	15.1
W20U	0.60	20.8
C20P	0.72	20.1
C20U	0.32	16.8
W5P	0.86	6.6
W5U	0.59	4.7
C5P	0.27	3.8
C5U	0.35	5.3

Table 3.8. SINDy model validation with ammonia experimental data

3.4.3 Phosphates modelling

Table 3.9 shows the best polynomial fit for phosphates. For all systems, the optimal polynomial fit falls between the 2nd and 5th order, indicating that phosphate dynamics in TWs are nonlinear. The planted system fed 2,000 mgCODt/L at 20 °C employs a 4th-degree polynomial fit, resulting in an RMSE of 8.51 mg/L, suggesting a more complex relationship but with relatively moderate accuracy. In contrast, the planted system fed 500 mgCODt/L at 20 °C system utilizes a 3th-degree polynomial fit, achieving a low RMSE of 0.68 mg/L, indicating a strong fit for the data. Similar trends are observed in unplanted system fed 500 mgCODt/L at 20 °C.

System	Best-polynomial fit	MSE (mg/L)
W20P	4	8.51
W20U	2	6.10
C20P	2	4.35
C20U	3	3.15
W5P	3	0.68
W5U	3	1.18
C5P	3	0.65
C5U	3	1.26

Table 3.9. Best polynomial fit for the phosphates trend.

Table 3.10 outlines the best hyperparameter, α , values for phosphate modelling and their corresponding RMSE values in mg/L for various systems. The α values exhibit considerable variation across the different systems. The planted system fed 500 mgCODt/L at 20 °C system, with a high hyperparameter value of 9,450, achieves a low RMSE of 1.05 mg/L, indicating that this particular setting effectively captures the underlying dynamics despite the strong regularization process. Conversely, the planted system fed 2,000 mgCODt/L at 20 °C system, which has a lower hyperparameter value of 1,671, results in a higher RMSE of 12.32 mg/L, suggesting a less accurate model.

		8 8 1
System	Best-hyperparameter value	RMSE (mg/L)
W20P	1,671	12.32
W20U	324	2.68
C20P	125	4.39
C20U	341	2.36
W5P	9,450	1.05
W5U	50	0.45
C5P	70	0.39
C5U	69	1.33

Table 3.10. Best hyperparameter value for the phosphates modelling using SINDy.

Table 3.11 presents the resulting model structure obtained by solving equation 6 using each system's optimal hyperparameter, α . The most simplified model was found for the planted
system fed 500 mgCODt/L at 6 °C, which included only 5 out of 27 potential candidates. In contrast, the planted system fed 500 mgCODt/L at 20 °C had the highest number of possible candidates, with 12 variables used to describe the model. The high pollutant removal efficiencies in the planted system fed 500 mgCODt/L at 20 °C indicate greater microbial activity due to ideal conditions (20°C) and healthy plant growth, resulting in overlapping processes that may need extra parameters to describe phosphate dynamics accurately. The functions COD*ammonia (C*N) and pH² stand out in all the systems models. The COD*N interaction in TWs might affect phosphate removal by altering microbial activity and redox conditions. COD consumption lowers oxygen levels, favouring ammonia denitrification but limiting aerobic phosphate removal by phosphate-accumulating organisms (PAOs) (Ding et al., 2022; Tarayre et al., 2016). Additionally, it has been reported that anoxic conditions can lead to the release of bound phosphates from sediments up to 5 times higher than in aerobic conditions (1 mg P/kg in aerobic; 5 mg P/kg in anaerobic), significantly reducing phosphate removal efficiency (Wu et al., 2014). Similarly, phosphates release by PAOs is reported in anaerobic conditions (Tarayre et al., 2016). Regarding the high sensitivity to pH, phosphorus, and thus phosphates released from sediments, has been found to be significantly higher under alkaline conditions (pH 7.5-8.5) compared to acidic conditions (pH 6.5-7.0) (Wu et al., 2014). This suggests that even a slight increase in pH can have a notable impact on phosphorus concentrations, potentially influencing nutrient cycling. In all treatment wetland (TW) systems, ORP is included in the model structure, either by interacting with phosphates (P*ORP) or as squared functions (ORP²). This might be explained as ORP influences the solubility and bioavailability of phosphates. Lower ORP conditions can lead to increased phosphate release from sediments, often due to the reduction of iron oxides that typically bind phosphorus to the substrate in aerobic environments (Wu et al., 2014). The ORP² function also reflects the nonlinear impacts of ORP conditions on microbial communities, particularly PAOs, which have demonstrated increased activity at higher ORP values (from +25 to +225 mV) (Li et al., 2016b; Tarayre et al., 2016).

Similar to ammonia, the pH coefficients are among the highest, as pH values range narrowly between 6.5 and 8.0 in all cases, while other parameters, such as COD, ammonia, and ORP, vary in the hundreds.

• Planted system fed with 2,000 mgCODt/L at 20 °C

$$\frac{dP}{dt} = -0.6C^2 + 0.2C * N - 0.5C * ORP - 0.1N^2 + 0.6P * ORP - 0.3ORP^2 + 24.4pH^2$$
[x10⁻⁵]

• Unplanted system fed with 2,000 mgCODt/L at 20 °C

$$\frac{dP}{dt} = 0.2C^2 + 0.2C * N + 0.9C * ORP - 3.8N * P - 0.9N * ORP + 0.7P * ORP - 0.10RP^2 + 68pH^2 [x10^{-5}]$$

• Planted system fed with 2,000 mgCODt/L at 6 °C

$$\frac{dP}{dt} = 0.1C^2 - 0.3C * N + 0.1N * ORP + 2P * ORP + 0.10RP^2 + 171pH^2 \qquad [x10^{-5}]$$

• Unplanted system fed with 2,000 mgCODt/L at 6 °C

$$\frac{dP}{dt} = -0.25C * N + 0.02C * ORP - 0.18N^2 - 0.07ORP^2 - 303pH^2 \qquad [x10^5]$$

• Planted system fed with 500 mgCODt/L at 20 °C

$$\frac{dP}{dt} = 0.1C * N + 0.8C * P - 0.3C * DO + 0.3C * pH - 0.7N^{2} + 0.3N * P - 0.1N * ORP + 0.3N * pH + 0.1P^{2} - 0.2P * ORP + 0.1P * pH - 51pH^{2} [x10^{-5}]$$

• Unplanted system fed with 500 mgCODt/L at 20 °C

$$\frac{dP}{dt} = -0.10RP - 0.1C^2 + 0.3C * N + 0.1C * P + 0.7N^2 + 0.1N * P + 0.1N * pH - 0.6P * ORP - 0.10RP * DO - 0.60RP * pH - 73pH^2 [x10^{-5}]$$

• Planted system fed with 500 mgCODt/L at 6 $^{\circ}$ C

$$\frac{dP}{dt} = 0.1C * N + 0.12C * ORP + 0.25N * ORP + 0.04ORP^2 + 18.7pH^2 \qquad [x10^{-5}]$$

• Unplanted system fed with 500 mgCODt/L at 6 °C

$$\frac{dP}{dt} = 0.2C^2 - 0.8C * N - 0.1C * P + 0.1C * ORP - 0.2N * P + 0.7P * ORP + 31.8pH^2$$

$$[r10^{-5}]$$

$$[x10^{-5}]$$

System	W20P	W20U	C20P	C20U	W5P	W5U	C5P	C5U
С								
Ν								
Р								
ORP						-0.1		
DO								
pН								
C^2	-0.6	0.2	0.1			-0.1		0.2
C*N	0.2	0.2	-0.3	-0.25	0.1	0.3	0.1	-0.8
C*P					0.8	0.1		-0.1
C*ORP	-0.5	0.9		0.02			0.12	0.1
C*DO					-0.3			
С*рН					0.3			
N^2	-0.1			-0.18	-0.7	0.7		
N*P		-3.8			0.3	0.1		-0.2
N*ORP		-0.9	0.1		-0.1		0.25	
N*DO								
N*pH					0.3	0.1		
\mathbf{P}^2					0.1			
P*ORP	0.6	0.7	2.0		-0.2	-0.6		0.7
P*DO								
P*pH					0.1			
ORP ²	-0.3	0.1	0.1	-0.07			0.04	
ORP*DO						-0.1		
ORP*pH						-0.6		
DO^2								
DO*pH								
pH ²	24.4	68.0	170.8	303.1	-51.2	-73.4	18.69	31.8

Table 3.11. Sparse coefficient estimated for phosphates modelling (values $x10^{-5}$)

Note: grey areas represent coefficients with a value of zero.

Figure 3.4 presents the model fit obtained using the SINDy algorithm compared to the polynomial curve that served as the source of the training data. R² values indicate a strong correlation with values ranging from 0.87 to 0.99 for all the systems. The RMSE is variable and ranges from 4.22 to 9.76 mg/L for high-strength WW and from 0.21 to 1.84 mg/L for low-strength WW. In most systems, the trend indicates a gradual decrease in phosphate concentrations, which aligns with the experimental trend. In the case of unplanted fed 500 mgCOD/L at 20 °C, the model describes an initial increase in phosphates, followed by a

continuous decline, as seen in the polynomial curve. The planted fed 2,000 mgCOD/L at 20°C model shows signs of overfitting due to the oscillations in the phosphate values produced by the model's solution.



Figure 3.4. SINDy model approximation to phosphates polynomial curve (training data) for A)W20P, B)W20U, C)C20P, D)C20U, E)W5P, F)W5U, G)C5P and H)C5U



Figure 3.4. (Continuation) SINDy model approximation to phosphates polynomial curve (training data) for A)W20P, B)W20U, C)C20P, D)C20U, E)W5P, F)W5U, G)C5P and H)C5U

Table 3.12 presents the fit parameters between the experimental data and the predicted values using the models generated by the SINDy algorithm. R² values range from 0.11 to 0.85, and the RMSE ranges from 4.1 to 5.7 mg/L for high-strength WW and from 0.9 to 2.1 mg/L for low-strength WW. The lowest performance accuracy was observed for unplanted fed with 500 mgCOD/L at 20°C (R² = 0.11, RMSE = 1.8 mg/L). In contrast, unplanted fed with 2,000 mgCOD/L at 20°C demonstrated strong prediction accuracy, with an R² of 0.85 and an RMSE of 5.7 mg/L.

-		
System	R ²	RMSE (mg/L)
W20P	0.80	5.4
W20U	0.76	4.1
C20P	0.79	5.4
C20U	0.85	5.7
W5P	0.60	0.9
W5U	0.11	1.8
C5P	0.68	1.0
C5U	0.62	2.1

Table 3.12. SINDy model validation with phosphates experimental data

3.5 Conclusions

A data-driven model using the SINDy algorithm was implemented to identify ODEs that characterize the pollutant dynamics of TWs under different operational and design conditions, highlighting the dependency and interaction between physicochemical parameters and pollutants. For the three parameters, COD, ammonia and phosphates, polynomial regression degrees ranged from the 2nd to the 7th, confirming the nonlinear nature of pollutant dynamics. Moreover, the differences in hyperparameter, α , values among systems emphasize the significance of choosing the appropriate value based on the complexity and noise of each dataset. Planted systems typically demonstrate higher α values due to their complex dynamics and might be associated with the effects of root activity on microbial diversity and activity, as well as nutrient uptake.

Differential equations for modelling the removal of COD, ammonia, and phosphates were determined. In certain instances, simplified equations involving 3 to 6 functions were found, whereas in cases with high noise in the raw data, up to 11 functions were required to capture the dynamics accurately. ORP was identified as a crucial parameter across all systems, underscoring its importance in estimating COD concentrations and its correlation with microbial activity. For ammonia dynamics, the analysis highlights the COD², ORP², and pH² functions, illustrating a nonlinear relationship that indicates minor operational adjustments can

greatly influence ammonia removal efficiency, particularly in high-strength WW scenarios. In terms of phosphates, the functions COD*ammonia, pH² and phosphates*ORP underscores the intricate relationships that influence phosphate removal. These factors may play an essential role in determining the solubility and bioavailability of phosphates.

The COD models strongly correlated with training (R^2 from 0.98 to 0.99) and validation data (R^2 from 0.81 to 0.98). In contrast, ammonia models exhibited a more diverse performance, showing R^2 values from 0.27 to 1.0. Similarly, for phosphates, validation results indicated varying prediction accuracies, with R^2 ranging from 0.87 to 0.99 with training data and from 0.11 to 0.85 with validation data. These findings suggest the necessity for improved data quality and stability to enhance model robustness, particularly in system models with low prediction accuracy.

Overall, this study demonstrates the effectiveness of the SINDy algorithm in developing models to predict pollutant concentrations in TWs. Future work will focus on training the model with online measurements to enhance the data's quality and quantity, as well as testing the reported equations in TWs across various scales and conditions. Applying Bayesian inference to the identified parameters can quantify the model's uncertainty. Additionally, this study suggests incorporating ORP and pH into the performance analysis and modeling of TWs.

4 SIGNIFICANCE OF THE RESEARCH AND CONCLUSIONS

4.1 Research summary

In this study, the performance of TWs for secondary domestic WW treatment was evaluated under different conditions, and a pollutant removal model was obtained based on the SINDy algorithm. A full factorial experiment (2x2x2) was conducted to examine the effects of WW strength (500 and 2,000 mg CODt/L), temperature (6 and 20°C), and vegetation (planted and unplanted) on lab-scale HSSFW, incorporating high-resolution monitoring of biologically relevant parameters such as ORP, DO, pH, and temperature. The experimental data was then used to train and validate a derivative model based on the SINDy algorithm, which identified simplified differential reaction equations for COD, phosphates, and ammonia, as well as the relevance and interactions between physicochemical parameters and pollutants. The main findings are as follows:

- (1) High COD removal and no significant difference between high and low WW strength was presented, highlighting the TWs potential for implementing as secondary treatment technology receiving black domestic WW characteristics.
- (2) Ammonia removal was generally low across most systems, except those with lowstrength WW, warm temperatures, and vegetation.
- (3) Phosphate removal was mainly influenced by WW strength and plant presence, with their interactions with temperature enhancing efficiency. The indicates that elevated temperatures and rising pH levels reduce the potential for phosphate adsorption and biological assimilation.
- (4) ORP offered more reliable insights into oxygenic conditions than direct DO measurements, particularly in deeper layers or sampling locations.

- (5) The SINDy algorithm identified simplified differential equations involving 3 to 6 functions in some cases, while in instances with high noise in the raw data, as many as 11 functions were needed to capture the dynamics.
- (6) Key factors influencing COD removal rate include ORP and COD*ammonia, as critical functions for estimating COD concentrations. The COD models strongly correlated with training (R² of 0.98 or 0.99) and validation data (R² between 0.81 and 0.98).
- (7) The key factors for ammonia removal rate include COD², ORP², and pH². These interactions suggest that minor operational adjustments can significantly affect ammonia removal efficiency. The ammonia models displayed varied performance, with R² values between 0.75 and 1.0 with training and from 0.26 to 0.86 with validation data.
- (8) Key factors for phosphate removal include COD*ammonia, pH² and phosphates*ORP, which might affect their solubility and bioavailability. The phosphate models exhibited a strong correlation with training data (R² from 0.87 to 0.99), and varying correlation with validation data (R² from 0.11 to 0.85), indicating diverse prediction accuracies.

Overall, this study demonstrates the potential of TWs for domestic WW treatment; suggests the use of the SINDy algorithm in developing simplified models to predict pollutant concentrations in TWs; and offer insights into key parameters for measurement (ORP, DO, and pH) that reflect the system's performance.

4.2 Limitations and future work

In this study, the performance evaluation and development of reaction equations in TWs for domestic WW was conducted. While the results showed efficient removal of COD and, to a lesser extent, ammonia and phosphates, the observations and model validation in field applications has yet to be confirmed. Several additional steps can be taken to improve the applicability of these findings for decentralized domestic WW treatment. Recommendations for follow-up optimization and potential future studies are outlined as follows:

- (1) The trend observations and resulting equations can be polished by using a larger quantity of cleaner measurements, as this study only used 48 data points. Incorporating data from online COD, ORP, DO, and pH measurements could further improve the current models. Machine learning techniques such as data augmentation and denoising algorithms could also be applied for data preparation.
- (2) The models developed in this study need to be validated with TWs exposed to the elements and receiving real domestic wastewater.
- (3) Measuring additional parameters, such as biomass growth in the substrate (soil), plant biomass growth, nitrites, total nitrogen, and organic nitrogen, could enhance the accuracy of the models.
- (4) Quantify the uncertainty of the model by applying Bayesian inference to the identified parameters.

REFERENCES

- Abdullah, F., Christofides, P.D., 2023. Data-based modeling and control of nonlinear process systems using sparse identification: An overview of recent results, in: AIChE Annual Meeting, Conference Proceedings. American Institute of Chemical Engineers. https://doi.org/10.1016/j.compchemeng.2023.108247
- Arias, C.A., Amin, L., Ananthatmula, R., Andrews, L., Baxpehler, H., Behrends, L.L., Bresciani, R., Brodnik, U., Buttiglier, G., Byvägen, N., Castañares, L., Chazarenc, F., Comas, J., Cross, K., Erjavec, T., Esterlich, M., Fedler, C.B., Ganeshan, G., Gattringer, H., Gearheart, R., Groot, I., Huang, G., Istenič, D., Jacob, A., Junge, R., Kaggwa, R., Kangabam, R., Langergraber, G., Lechner, M., Liu, W., Lombard-Latune, R., Masi, G., Mendoza, E., Molle, P., Morvannou, A., Wallace, S., Waara, S., Zalaznik, A.M., Zapater-Pereyra, M., Zhai, J., 2021. Nature based solutions for sustainable sanitation. IWA Publishing.
- Bai, J., Ye, X., Jia, J., Zhang, G., Zhao, Q., Cui, B., Liu, X., 2017. Phosphorus sorptiondesorption and effects of temperature, pH and salinity on phosphorus sorption in marsh soils from coastal wetlands with different flooding conditions. Chemosphere 188, 677– 688. https://doi.org/10.1016/j.chemosphere.2017.08.117
- Březinová, T., Vymazal, J., 2015. Seasonal growth pattern of Phalaris arundinacea in constructed wetlands with horizontal subsurface flow. Ecol Eng 80, 62–68. https://doi.org/10.1016/j.ecoleng.2014.07.057
- Brunton, S.L., Hemati, M.S., Taira, K., 2020. Special issue on machine learning and datadriven methods in fluid dynamics. Theor Comput Fluid Dyn. https://doi.org/10.1007/s00162-020-00542-y
- Brunton, S.L., Proctor, J.L., Kutz, J.N., Bialek, W., 2016. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. Proc Natl Acad Sci U S A 113, 3932–3937. https://doi.org/10.1073/pnas.1517384113
- Cosgrove, W.J., Loucks, D.P., 2015. Water management: Current and future challenges and research directions. Water Resour Res. https://doi.org/10.1002/2014WR016869

- de Anda, J., López-López, A., Villegas-García, E., Valdivia-Aviña, K., 2018. High-strength domestic wastewater treatment and reuse with onsite passive methods. Water (Switzerland) 10. https://doi.org/10.3390/w10020099
- del Castillo, A.F., Garibay, M.V., Senés-Guerrero, C., Orozco-Nunnelly, D.A., de Anda, J., Gradilla-Hernández, M.S., 2022. A review of the sustainability of anaerobic reactors combined with constructed wetlands for decentralized wastewater treatment. J Clean Prod. https://doi.org/10.1016/j.jclepro.2022.133428
- Delgadillo, O., Camacho, A., Pérez, L., Andrade, M., 2010. Depuración de aguas residuales por medio de humedales artificiales. Cochabamba, Bolivia.
- Demir-Duz, H., Aktürk, A.S., Ayyildiz, O., Álvarez, M.G., Contreras, S., 2020. Reuse and recycle solutions in refineries by ozone-based advanced oxidation processes: A statistical approach. J Environ Manage 263. https://doi.org/10.1016/j.jenvman.2020.110346
- Ding, Y., Wang, H., Zhang, Q., Chai, B., Lei, X., Ye, M., Chen, B., 2022. Effects of dissolved oxygen on phosphorus transformation in reservoir sediments: novel insights on bacterial community and functional genes. J Soils Sediments 22, 2094–2104. https://doi.org/10.1007/s11368-022-03233-9
- Foladori, P., Ortigara, A.R.C., Ruaben, J., Andreottola, G., 2012. Influence of high organic loads during the summer period on the performance of hybrid constructed wetlands (VSSF + HSSF) treating domestic wastewater in the Alps region. Water Science and Technology 65, 890–897. https://doi.org/10.2166/wst.2012.932
- Foladori, P., Ruaben, J., Ortigara, A.R.C., 2013. Recirculation or artificial aeration in vertical flow constructed wetlands: A comparative study for treating high load wastewater. Bioresour Technol 149, 398–405. https://doi.org/10.1016/j.biortech.2013.09.099
- Fountoulakis, M.S., Daskalakis, G., Papadaki, A., Kalogerakis, N., Manios, T., 2017. Use of halophytes in pilot-scale horizontal flow constructed wetland treating domestic wastewater. Environmental Science and Pollution Research 24, 16682–16689. https://doi.org/10.1007/s11356-017-9295-8

- Fu, X., Wu, X., Zhou, S., Chen, Y., Chen, M., Chen, R., 2018. A constructed wetland system for rural household sewage treatment in subtropical regions. Water (Switzerland) 10. https://doi.org/10.3390/w10060716
- Fuerhacker, M., Bauer, H., Ellinger, R., Sree, U., Schmid, H., Zibuschka, F., Puxbaum, H., n.d. APPROACH FOR A NOVEL CONTROL STRATEGY FOR SIMULTANEOUS NITRIFICATION/DENITRIFICATION IN ACTIVATED SLUDGE REACTORS.
- García, J., Rousseau, D.P.L., Morató, J., Lesage, E., Matamoros, V., Bayona, J.M., 2010. Contaminant removal processes in subsurface-flow constructed wetlands: A review. Crit Rev Environ Sci Technol. https://doi.org/10.1080/10643380802471076
- Garcia, J., Rousseau, D.P.L., Morato, J., Lesage, E.L.S., Matamoros, V., Bayona, J.M., 2010. Contaminant Removal Processes in Subsurface-Flow Constructed Wetlands: A Review. Crit Rev Environ Sci Technol 40, 561–661. https://doi.org/10.1080/10643380802471076
- Gersberg, R.M., Elkins, B.V., Lyon, S.R., Goldman, C.R., 1986. Role of aquatic plants in wastewater treatment by artificial wetlands. Water Res 20, 363–368. https://doi.org/10.1016/0043-1354(86)90085-0
- Ghimire, U., Nandimandalam, H., Martinez-Guerra, E., Gude, V.G., 2019. Wetlands for wastewater treatment. Water Environment Research. https://doi.org/10.1002/wer.1232
- Goharoodi, S.K., Dekemele, K., Dupre, L., Loccufier, M., Crevecoeur, G., 2018. Sparse Identification of Nonlinear Duffing Oscillator From Measurement Data, in: IFAC-PapersOnLine. Elsevier B.V., pp. 162–167. https://doi.org/10.1016/j.ifacol.2018.12.111
- Gómez-Borraz, T.G., Salinas-Toledano, M.A., Garcia-Becerra, F.Y., 2022. A viable household-scale prototype based on treatment wetlands for on-site wastewater treatment and reuse in Mexico City, in: IWA 17 Th International Conference on Wetland Systems for Water Pollution Control. IWA, Lyon.
- Grimalt-Alemany, A., Etler, C., Asimakopoulos, K., Skiadas, I. V., Gavala, H.N., 2021. ORP control for boosting ethanol productivity in gas fermentation systems and dynamics of redox cofactor NADH/NAD+ under oxidative stress. Journal of CO2 Utilization 50, 101589. https://doi.org/10.1016/j.jcou.2021.101589

- Grinberga, L., Celms, A., Dzięcioł, J., Głuchowski, A., Grabuža, D., Grīnfelde, I., Lauva, D., Sas, W., 2022. Analysis of the removal of BOD5, COD and suspended solids in subsurface flow constructed wetland in Latvia. ACTA SCIENTIARUM POLONORUM
 Architectura Budownictwo 20, 21–28. https://doi.org/10.22630/aspa.2021.20.4.31
- Gupta, P.K., Yadav, B., Kumar, A., Himanshu, S.K., 2021. Machine learning and artificial intelligence application in constructed wetlands for industrial effluent treatment: advances and challenges in assessment and bioremediation modeling, in: Bioremediation for Environmental Sustainability. Elsevier, pp. 403–414. https://doi.org/10.1016/B978-0-12-820524-2.00016-X
- He, J., Feng, H., Diao, Z., Su, D., 2023. Effect of temperature variation on phosphorus flux at the sediment-water interface of the steppe wetlands. Environmental Science and Pollution Research 30, 12441–12452. https://doi.org/10.1007/s11356-022-23015-7
- Hoekstra, A.Y., Buurman, J., Van Ginkel, K.C.H., 2018. Urban water security: A review. Environmental Research Letters. https://doi.org/10.1088/1748-9326/aaba52
- Hoffmann, M., Fröhner, C., Noé, F., 2019. Reactive SINDy: Discovering governing reactions from concentration data. Journal of Chemical Physics 150. https://doi.org/10.1063/1.5066099
- Ilyas, H., Masih, I., 2018. The effects of different aeration strategies on the performance of constructed wetlands for phosphorus removal. Environmental Science and Pollution Research. https://doi.org/10.1007/s11356-017-1071-2
- Ilyas, H., Masih, I., 2017. The performance of the intensified constructed wetlands for organic matter and nitrogen removal: A review. J Environ Manage. https://doi.org/10.1016/j.jenvman.2017.04.098
- Jafarzadeh, H., Mahdianpari, M., Gill, E.W., Brisco, B., Mohammadimanesh, F., 2022. Remote Sensing and Machine Learning Tools to Support Wetland Monitoring: A Meta-Analysis of Three Decades of Research. Remote Sens (Basel). https://doi.org/10.3390/rs14236104

- Ji, M., Hu, Z., Hou, C., Liu, H., Ngo, H.H., Guo, W., Lu, S., Zhang, J., 2020. New insights for enhancing the performance of constructed wetlands at low temperatures. Bioresour Technol. https://doi.org/10.1016/j.biortech.2019.122722
- Kadam, A.M., Nemade, P.D., Oza, G.H., Shankar, H.S., 2009. Treatment of municipal wastewater using laterite-based constructed soil filter. Ecol Eng 35, 1051–1061. https://doi.org/10.1016/j.ecoleng.2009.03.008
- Kadlec, R.H., 2016. Large constructed wetlands for phosphorus control: A review. Water (Switzerland). https://doi.org/10.3390/w8060243
- Kadlec, R.H., Reddy, K.R., 2001. Temperature Effects in Treatment Wetlands. Water Environment Research 73, 543–557. https://doi.org/10.2175/106143001X139614
- Kadlec, R.H., Wallace, S.D., 2009. Treatment wetlands, Second. ed. CRC Press, Boca Raton, FL.
- Kantawanichkul, S., Kladprasert, S., Brix, H., 2009. Treatment of high-strength wastewater in tropical vertical flow constructed wetlands planted with Typha angustifolia and Cyperus involucratus. Ecol Eng 35, 238–247. https://doi.org/10.1016/j.ecoleng.2008.06.002
- Kargol, A.K., Burrell, S.R., Chakraborty, I., Gough, H.L., 2023. Synthetic wastewater prepared from readily available materials: Characteristics and economics. PLOS Water 2, e0000178. https://doi.org/10.1371/journal.pwat.0000178
- Körner, S., Vermaat, J.E., 1998. The relative importance of Lemna gibba L., bacteria and algae for the nitrogen and phosphorus removal in duckweed-covered domestic wastewater. Water Res 32, 3651–3661. https://doi.org/10.1016/S0043-1354(98)00166-3
- Kumar, S., Dutta, V., 2019. Constructed wetland microcosms as sustainable technology for domestic wastewater treatment: an overview. Environmental Science and Pollution Research. https://doi.org/10.1007/s11356-019-04816-9
- Lai, W.L., Zhang, Y., Chen, Z.H., 2012. Radial oxygen loss, photosynthesis, and nutrient removal of 35 wetland plants. Ecol Eng 39, 24–30. https://doi.org/10.1016/j.ecoleng.2011.11.010

- Lamers, L.P.M., van Diggelen, J.M.H., Op Den Camp, H.J.M., Visser, E.J.W., Lucassen, E.C.H.E.T., Vile, M.A., Jetten, M.S.M., Smolders, A.J.P., Roelofs, J.G.M., 2012. Microbial transformations of nitrogen, sulfur, and iron dictate vegetation composition in wetlands: A review. Front Microbiol 3. https://doi.org/10.3389/fmicb.2012.00156
- Li, B., Bishop, P., 2002. Oxidation-reduction potential (ORP) regulation of nutrient removal in activated sludge wastewater treatment plants. https://doi.org/https://doi.org/10.2166/wst.2002.0452
- Li, F., Liu, X., Zhang, X., Zhao, D., Liu, H., Zhou, C., Wang, R., 2017. Urban ecological infrastructure: an integrated network for ecosystem services and sustainable urban systems. J Clean Prod 163, S12–S18. https://doi.org/10.1016/j.jclepro.2016.02.079
- Li, L., Rong, S., Wang, R., Yu, S., 2021. Recent advances in artificial intelligence and machine learning for nonlinear relationship analysis and process control in drinking water treatment: A review. Chemical Engineering Journal 405. https://doi.org/10.1016/j.cej.2020.126673
- Li, M., Liang, Z., Callier, M.D., d'Orbcastel, E.R., Ma, X., Sun, L., Li, X., Wang, S., Song, X., Liu, Y., 2018. Nitrogen and organic matter removal and enzyme activities in constructed wetlands operated under different hydraulic operating regimes. Aquaculture 496, 247–254. https://doi.org/10.1016/j.aquaculture.2018.06.016
- Li, X., Guo, Q., Wang, Y., Xu, J., Wei, Q., Chen, L., Liao, L., 2020. Enhancing nitrogen and phosphorus removal by applying effective microorganisms to constructed wetlands. Water (Switzerland) 12. https://doi.org/10.3390/w12092443
- Li, Y.H., Li, H.B., Xu, X.Y., Zhou, Y.C., Gong, X., 2016a. Correlations between the oxidation-reduction potential characteristics and microorganism activities in the subsurface wastewater infiltration system. Desalination Water Treat 57, 5350–5357. https://doi.org/10.1080/19443994.2014.1003606
- Li, Y.H., Li, H.B., Xu, X.Y., Zhou, Y.C., Gong, X., 2016b. Correlations between the oxidation-reduction potential characteristics and microorganism activities in the subsurface wastewater infiltration system. Desalination Water Treat 57, 5350–5357. https://doi.org/10.1080/19443994.2014.1003606

- Loiseau, J.C., 2020. Data-driven modeling of the chaotic thermal convection in an annular thermosyphon. Theor Comput Fluid Dyn 34, 339–365. https://doi.org/10.1007/s00162-020-00536-w
- Lu, J., Dong, L., Guo, Z., Hu, Z., Dai, P., Zhang, J., Wu, H., 2023. Highly efficient phosphorous removal in constructed wetland with iron scrap: Insights into the microbial removal mechanism. J Environ Manage 347. https://doi.org/10.1016/j.jenvman.2023.119076
- Mangan, N.M., Brunton, S.L., Proctor, J.L., Kutz, J.N., 2016. Inferring Biological Networks by Sparse Identification of Nonlinear Dynamics. IEEE Trans Mol Biol Multiscale Commun 2, 52–63. https://doi.org/10.1109/TMBMC.2016.2633265
- Meng, P., Pei, H., Hu, W., Shao, Y., Li, Z., 2014. How to increase microbial degradation in constructed wetlands: Influencing factors and improvement measures. Bioresour Technol. https://doi.org/10.1016/j.biortech.2014.01.095
- Meyer, D., Chazarenc, F., Claveau-Mallet, D., Dittmer, U., Forquet, N., Molle, P., Morvannou,
 A., Pálfy, T., Petitjean, A., Rizzo, A., Samsó Campà, R., Scholz, M., Soric, A.,
 Langergraber, G., 2015. Modelling constructed wetlands: Scopes and aims a
 comparative review. Ecol Eng 80, 205–213.
 https://doi.org/10.1016/j.ecoleng.2014.10.031
- Nguyen, X.C., Chang, S.W., Nguyen, T.L., Ngo, H.H., Kumar, G., Banu, J.R., Vu, M.C., Le, H.S., Nguyen, D.D., 2018. A hybrid constructed wetland for organic-material and nutrient removal from sewage: Process performance and multi-kinetic models. J Environ Manage 222, 378–384. https://doi.org/10.1016/j.jenvman.2018.05.085
- Nguyen, X.C., Ly, Q.V., Li, J., Bae, H., Bui, X.T., Nguyen, T.T.H., Tran, Q.B., Vo, T.D.H., Nghiem, L.D., 2021. Nitrogen removal in subsurface constructed wetland: Assessment of the influence and prediction by data mining and machine learning. Environ Technol Innov 23. https://doi.org/10.1016/j.eti.2021.101712
- Nivala, J., Wallace, S., Headley, T., Kassa, K., Brix, H., van Afferden, M., Müller, R., 2013. Oxygen transfer and consumption in subsurface flow treatment wetlands. Ecol Eng 61, 544–554. https://doi.org/10.1016/j.ecoleng.2012.08.028

- Ong, S.A., Uchiyama, K., Inadama, D., Yamagiwa, K., 2009. Simultaneous removal of color, organic compounds and nutrients in azo dye-containing wastewater using up-flow constructed wetland. J Hazard Mater 165, 696–703. https://doi.org/10.1016/j.jhazmat.2008.10.071
- Palmer, M.A., Liu, J., Matthews, J.H., Mumba, M., D'Odorico, P., 2015. Manage water in a green way. Science (1979). https://doi.org/10.1126/science.aac7778
- Picos-Benítez, A.R., López-Hincapié, J.D., Chávez-Ramírez, A.U., Rodríguez-García, A., 2017. Artificial intelligence based model for optimization of COD removal efficiency of an up-flow anaerobic sludge blanket reactor in the saline wastewater treatment. Water Science and Technology 75, 1351–1361. https://doi.org/10.2166/wst.2017.005
- Prokop, B., Gelens, L., 2024. From biological data to oscillator models using SINDy. iScience 27. https://doi.org/10.1016/j.isci.2024.109316
- Puigagut, J., Caselles-Osorio, A., Vaello, N., García, J., 2008. Fractionation, biodegradability and particle-size distribution of organic matter in horizontal subsurface-flow constructed wetlands, in: Wastewater Treatment, Plant Dynamics and Management in Constructed and Natural Wetlands. Springer Netherlands, pp. 289–297. https://doi.org/10.1007/978-1-4020-8235-1_25
- Rahi, M.A., Faisal, A.A.H., Naji, L.A., Almuktar, S.A., Abed, S.N., Scholz, M., 2020. Biochemical performance modelling of non-vegetated and vegetated vertical subsurfaceflow constructed wetlands treating municipal wastewater in hot and dry climate. Journal of Water Process Engineering 33. https://doi.org/10.1016/j.jwpe.2019.101003
- Rehman, F., Pervez, A., Mahmood, Q., Nawab, B., 2017. Wastewater remediation by optimum dissolve oxygen enhanced by macrophytes in constructed wetlands. Ecol Eng 102, 112– 126. https://doi.org/10.1016/j.ecoleng.2017.01.030
- Saeed, T., Sun, G., 2012. A review on nitrogen and organics removal mechanisms in subsurface flow constructed wetlands: Dependency on environmental parameters, operating conditions and supporting media. J Environ Manage. https://doi.org/10.1016/j.jenvman.2012.08.011

- Sandoz, A., Ducret, V., Gottwald, G.A., Vilmart, G., Perron, K., 2023. SINDy for delaydifferential equations: application to model bacterial zinc response. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 479. https://doi.org/10.1098/rspa.2022.0556
- Sarkar, A., Biswas, D.R., Datta, S.C., Manjaiah, K.M., Roy, T., 2017. Release of Phosphorus from Laboratory Made Coated Phosphatic Fertilizers in Soil Under Different Temperature and Moisture Regimes. Proceedings of the National Academy of Sciences India Section B - Biological Sciences 87, 1299–1308. https://doi.org/10.1007/s40011-015-0693-8
- Singh, S., Kulshreshtha, N.M., Goyal, S., Brighu, U., Bezbaruah, A.N., Gupta, A.B., 2022. Performance prediction of horizontal flow constructed wetlands by employing machine learning. Journal of Water Process Engineering 50. https://doi.org/10.1016/j.jwpe.2022.103264
- Solano, M.L., Soriano, P., Ciria, M.P., 2004. Constructed Wetlands as a Sustainable Solution for Wastewater Treatment in Small Villages. Biosyst Eng 87, 109–118. https://doi.org/10.1016/j.biosystemseng.2003.10.005
- Solomatine, D., See, L.M., Abrahart, R.J., 2008. Chapter 2 Data-Driven Modelling: Concepts, Approaches and Experiences. Springer-Verlag.
- Song, P., Huang, G., An, C., Shen, Ju, Zhang, P., Chen, X., Shen, Jian, Yao, Y., Zheng, R., Sun, C., 2018. Treatment of rural domestic wastewater using multi-soil-layering systems: Performance evaluation, factorial analysis and numerical modeling. Science of the Total Environment 644, 536–546. https://doi.org/10.1016/j.scitotenv.2018.06.331
- Stefanakis, A.I., Tsihrintzis, V.A., 2009. Effect of outlet water level raising and effluent recirculation on removal efficiency of pilot-scale, horizontal subsurface flow constructed wetlands. Desalination 248, 961–976. https://doi.org/10.1016/j.desal.2008.08.008
- Sun, H., Yu, N., Liu, Y., 2022. Importance of low-abundance microbial species in response to disturbances in wastewater bioreactors. Process Safety and Environmental Protection 162, 663–671. https://doi.org/10.1016/j.psep.2022.04.047

- Tang, S., Liao, Y., Xu, Y., Dang, Z., Zhu, X., Ji, G., 2020a. Microbial coupling mechanisms of nitrogen removal in constructed wetlands: A review. Bioresour Technol. https://doi.org/10.1016/j.biortech.2020.123759
- Tang, S., Liao, Y., Xu, Y., Dang, Z., Zhu, X., Ji, G., 2020b. Microbial coupling mechanisms of nitrogen removal in constructed wetlands: A review. Bioresour Technol. https://doi.org/10.1016/j.biortech.2020.123759
- Tarayre, C., Nguyen, H.T., Brognaux, A., Delepierre, A., De Clercq, L., Charlier, R., Michels, E., Meers, E., Delvigne, F., 2016. Characterisation of phosphate accumulatingorganisms and techniques for polyphosphatedetection: A review. Sensors (Switzerland). https://doi.org/10.3390/s16060797
- Trang, N.T.D., Konnerup, D., Schierup, H.H., Chiem, N.H., Tuan, L.A., Brix, H., 2010. Kinetics of pollutant removal from domestic wastewater in a tropical horizontal subsurface flow constructed wetland system: Effects of hydraulic loading rate. Ecol Eng 36, 527–535. https://doi.org/10.1016/j.ecoleng.2009.11.022
- Truu, J., Truu, M., Espenberg, M., Nõlvak, H., Juhanson, J., 2015. Phytoremediation and Plant-Assisted Bioremediation in Soil and Treatment Wetlands: A Review. The Open of Biotechnology Journal 9, 85–92. https://doi.org/10.2174/1874070701509010085
- Truu, M., Juhanson, J., Truu, J., 2009. Microbial biomass, activity and community composition in constructed wetlands. Science of the Total Environment 407, 3958–3971. https://doi.org/10.1016/j.scitotenv.2008.11.036
- UN-Habitat, 2008. Constructed Wetlands Manual. UN-Habitat Water for Asian Cities Program, Nepal, Kathmandu.
- UN-Water, 2021. Summary Progress Update 2021 SDG 6 water and sanitation for all. Geneva, Switzerland.
- UN-WHO, 2010. Resolution A/RES/64/292. United Nations General Assembly, July 2010 General Comment No. 15. The right to water. UN Committee on Economic, Social and Cultural Rights.
- US EPA, 2000. Manual Constructed Wetlands Treatment of Municipal Wastewaters. Washington, DC, USA.

- Varma, M., Gupta, A.K., Ghosal, P.S., Majumder, A., 2021. A review on performance of constructed wetlands in tropical and cold climate: Insights of mechanism, role of influencing factors, and system modification in low temperature. Science of the Total Environment. https://doi.org/10.1016/j.scitotenv.2020.142540
- Ventura, D., Rapisarda, R., Sciuto, L., Milani, M., Consoli, S., Cirelli, G.L., Licciardello, F., 2022. Application of first-order kinetic removal models on constructed wetlands under Mediterranean climatic conditions. Ecol Eng 175. https://doi.org/10.1016/j.ecoleng.2021.106500
- Vohla, C., Kõiv, M., Bavor, H.J., Chazarenc, F., Mander, Ü., 2011. Filter materials for phosphorus removal from wastewater in treatment wetlands-A review. Ecol Eng 37, 70– 89. https://doi.org/10.1016/j.ecoleng.2009.08.003
- Von Sperling, M., de Paoli, A.C., 2013. First-order COD decay coefficients associated with different hydraulic models applied to planted and unplanted horizontal subsurface-flow constructed wetlands. Ecol Eng 57, 205–209. https://doi.org/10.1016/j.ecoleng.2013.04.036
- Wang, M., Zhang, D.Q., Dong, J.W., Tan, S.K., 2017. Constructed wetlands for wastewater treatment in cold climate — A review. Journal of Environmental Sciences 57, 293–311. https://doi.org/10.1016/j.jes.2016.12.019
- Wang, Q., Cao, Z., Liu, Q., Zhang, Jinyong, Hu, Y., Zhang, Ji, Xu, W., Kong, Q., Yuan, X., Chen, Q.F., 2019. Enhancement of COD removal in constructed wetlands treating saline wastewater: Intertidal wetland sediment as a novel inoculation. J Environ Manage 249. https://doi.org/10.1016/j.jenvman.2019.109398
- Wang, W., Ding, Y., Ullman, J.L., Ambrose, R.F., Wang, Y., Song, X., Zhao, Z., 2016. Nitrogen removal performance in planted and unplanted horizontal subsurface flow constructed wetlands treating different influent COD/N ratios. Environmental Science and Pollution Research 23, 9012–9018. https://doi.org/10.1007/s11356-016-6115-5
- Wang, Y., Kmiya, Y., Okuhara, T., 2007. Removal of low-concentration ammonia in water by ion-exchange using Na-mordenite. Water Res 41, 269–276. https://doi.org/10.1016/j.watres.2006.10.035

- Weißbach, M., Drewes, J.E., Koch, K., 2018. Application of the oxidation reduction potential (ORP) for process control and monitoring nitrite in a Coupled Aerobic-anoxic Nitrous Decomposition Operation (CANDO). Chemical Engineering Journal 343, 484–491. https://doi.org/10.1016/j.cej.2018.03.038
- Wießner, A., Kappelmeyer, U., Kuschk, P., Kästner, M., 2005. Influence of the redox condition dynamics on the removal efficiency of a laboratory-scale constructed wetland. Water Res 39, 248–256. https://doi.org/10.1016/j.watres.2004.08.032
- Wu, H., Zhang, J., Ngo, H.H., Guo, W., Hu, Z., Liang, S., Fan, J., Liu, H., 2015. A review on the sustainability of constructed wetlands for wastewater treatment: Design and operation. Bioresour Technol. https://doi.org/10.1016/j.biortech.2014.10.068
- Wu, Yunhai, Wen, Y., Zhou, J., Wu, Yunying, 2014. Phosphorus release from lake sediments: Effects of pH, temperature and dissolved oxygen. KSCE Journal of Civil Engineering 18, 323–329. https://doi.org/10.1007/s12205-014-0192-0
- WWAP, UN-Water, 2018. Nature-based solutions for water. UNESCO, Paris.
- Yadav, A., Chazarenc, F., Mutnuri, S., 2018. Development of the "French system" vertical flow constructed wetland to treat raw domestic wastewater in India. Ecol Eng 113, 88–93. https://doi.org/10.1016/j.ecoleng.2018.01.001
- Yang, Z., Yang, L., Wei, C., Wu, W., Zhao, X., Lu, T., 2018. Enhanced nitrogen removal using solid carbon source in constructed wetland with limited aeration. Bioresour Technol 248, 98–103. https://doi.org/10.1016/j.biortech.2017.07.188
- Yuan, C., Huang, T., Zhao, X., Zhao, Y., 2020. Numerical models of subsurface flow constructed wetlands: Review and future development. Sustainability (Switzerland). https://doi.org/10.3390/SU12083498
- Zhai, J., Zou, J., He, Q., Ning, K., Xiao, H., 2012. Variation of dissolved oxygen and redox potential and their correlation with microbial population along a novel horizontal subsurface flow wetland. Environ Technol 33, 1999–2006. https://doi.org/10.1080/09593330.2012.655320
- Zhang, D.Q., Jinadasa, K.B.S.N., Gersberg, R.M., Liu, Y., Ng, W.J., Tan, S.K., 2014. Application of constructed wetlands for wastewater treatment in developing countries -

A review of recent developments (2000-2013). J Environ Manage. https://doi.org/10.1016/j.jenvman.2014.03.015

- Zhao, L., Fu, G., Zeng, A., Cheng, B., Song, Z., Hu, Z., 2023. Effects of different aeration strategies and ammonia-nitrogen loads on nitrification performance and microbial community succession of mangrove constructed wetlands for saline wastewater treatment. Chemosphere 339, 139685. https://doi.org/10.1016/j.chemosphere.2023.139685
- Zhong, F., Wu, J., Dai, Y., Yang, L., Zhang, Z., Cheng, S., Zhang, Q., 2015. Bacterial community analysis by PCR-DGGE and 454-pyrosequencing of horizontal subsurface flow constructed wetlands with front aeration. Appl Microbiol Biotechnol 99, 1499– 1512. https://doi.org/10.1007/s00253-014-6063-2
- Zhou, Q.H., Wu, Z.B., Cheng, S.P., He, F., Fu, G.P., 2005. Enzymatic activities in constructed wetlands and di-n-butyl phthalate (DBP) biodegradation. Soil Biol Biochem 37, 1454– 1459. https://doi.org/10.1016/j.soilbio.2005.01.003
- Zhu, H., Yan, B., Xu, Y., Guan, J., Liu, S., 2014. Removal of nitrogen and COD in horizontal subsurface flow constructed wetlands under different influent C/N ratios. Ecol Eng 63, 58–63. https://doi.org/10.1016/j.ecoleng.2013.12.018
- Zitnik, M., Nguyen, F., Wang, B., Leskovec, J., Goldenberg, A., Hoffman, M.M., 2019.
 Machine learning for integrating data in biology and medicine: Principles, practice, and opportunities. Information Fusion 50, 71–91.
 https://doi.org/10.1016/j.inffus.2018.09.012

APPENDIX A

a)



b)



Figure A1. Experimental set ups in the a) greenhouse and b) cold room. W=warm; C=cold; 5= low strength WW; 20=high strength WW; P=plantes; U=unplanted



Figure A2. Correlation between CODs from digestion and CODs measurements from the probe at a) the beginning and b) the end of the experiment using SWW samples. The samples for the digestion method were previously filtered using a 45 mm membrane



Figure A3. Residual plots for a) CODt, b) ammonia and c) phosphates.



Figure A4. Nitrates concentration within the system in TWs.



Figure A5. Water temperature within the system in TWs.