Systems & Control Transactions

Original Research Article
Peer Reviewed Conference Proceeding
ESCAPE 36 - European Symposium on Computer Aided Process Engineering
Sheffield, UK. 21-24 June 2026

PSE PRESS

Hybrid Modeling of Wastewater Treatment Dynamics Using Hammerstein-Wiener Structures

Arne Tirez[a], Niels Stevens[a], Dominik Bongartz[a], and José Matias Assumpcao[b]

[a] KU Leuven, Department of Chemical Engineering, Leuven, Flemish-Brabant, Belgium
[b] KU Leuven, Department of Chemical Engineering, Sint-Katelijne-Waver, Flemish-Brabant, Belgium

ABSTRACT

The zero-pollution ambition of the European Union requires improvements in wastewater treatment to meet increasingly stringent regulations at achievable cost. One promising approach consists in model-based optimal control. However, wastewater treatment plants involve highly nonlinear and time-varying processes, making existing mechanistic models such as the Benchmark Simulation Model no.1 (BSM1) challenging for direct use in online control. Therefore, this study explores a hybrid modeling approach using the Hammerstein-Wiener (HW) structure. The proposed model combines a mechanistic steady-state model, derived from BSM1, with a data-driven approximation of the system dynamics, incorporating low-order linear dynamic models. In this work, the HW model was used as a surrogate for BSM1. The HW surrogate model attained coefficients of determination (R^2) often exceeding 0.95 across key water quality indicators, such as total nitrogen and ammonium concentration. This accuracy was found to be substantially better than that of a First-Order-Plus Dead-Time model as conventional low-order linear model, demonstrating the usefulness of the HW structure in capturing key nonlinearities.

Keywords: Dynamic Modelling, Modelling, Modelling and Simulations, Wastewater, System Identification

INTRODUCTION

New regulation is imposing increasingly stringent requirements on wastewater treatment to tackle emerging sources of urban pollution and address societal challenges like the energy transition and climate neutrality [1]. One approach towards meeting these requirements consists in model-based methods, and in particular model-based optimal control. However, accurate and computationally tractable modeling of wastewater treatment dynamics remains a challenge.

In the 80s, the Activated Sludge Model no.1 (ASM1) became the first accepted mechanistic model across the industry, giving insights into the dynamic behavior of the carbon and nitrogen removal of activated sludge systems. Several versions have since been developed to target specific biological processes, such as phosphorus recovery and anaerobic digestion [2].

Due to the success of the ASM1 and its variants, mechanistic models are still used in wastewater treatment plants (WWTPs) because of the process insight and predictive capabilities they provide in plant design and operation optimization. The Benchmark Simulation Model No. 1 (BSM1) [3] is the industry standard for benchmarking and control application. Conversely, dynamic black-box models (a.k.a. data-driven models) are the most effective in the presence of a high volume of data. However, they have limited interpretability, large data requirements, and are unreliable in extrapolation [4].

A promising intermediate approach is hybrid modeling, which has already been successfully applied to certain aspects of wastewater treatment where some parameters, such as micropollutant concentration, colloid charge, and pathogen abundance, are hard or expensive to measure with sufficient precision [5]. By using data-driven surrogates for these complex variables, hybrid models offer a robust solution for real-time monitoring and process control. For instance, Gaublomme et al. [6] used hybrid models for solution-diffusion processes in membrane processes. Wang et al. [7] created an integrated framework to improve forecasting of salinity and dissolved oxygen with hybrid models. Nordstrand & Dutta [8] achieved a more robust and flexible control of capacitive-deionisation desalination processes by applying the hybrid linear-state-space dynamic Langmuir model. Tomperi et al. [9] used a series-connected ARMAX model to estimate effluent COD concentration in full-scale industrial WWTPs.

Despite the fact that current hybrid models offer significant predictive capacity, they often face a trade-off between computational complexity and integration into real-time control and optimization frameworks. A critical gap remains in identifying hybrid structures that are complex enough to capture wastewater dynamics but simple enough to be used efficiently in mathematical optimization [10–12].

In this work, we address this by utilizing Hammerstein–Wiener (HW) models. These models constitute a class of block-structured dynamic models that represent systems as a linear dynamic block embedded between two nonlinear static transformations [13]. Specifically, we aim at using a mechanistic steady-state model as nonlinear static block, combined with a data-driven low-order dynamic block. We hypothesize that this hybrid model structure is well-suited for WWTP for two reasons: First, many dynamics observed in WWTP pilot plants can be effectively described as first-order models [2]. Second, by including a mechanistic model as static block, we can capture nonlinearities in the gains and represents steady states accurately. In this work, we use BSM1 both to derive the static block, and as a high-fidelity representation of the true system (cf. Figure 1).

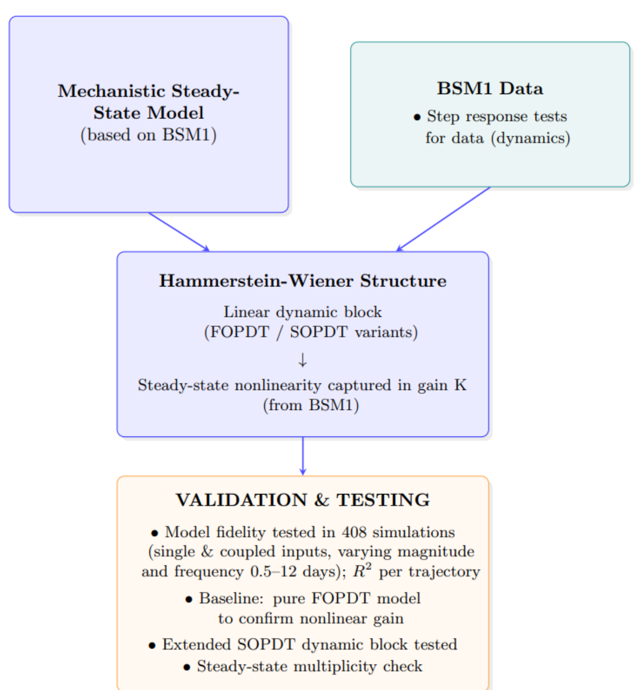

Figure 1. Simplified Flowsheet of Methodology

Our contribution lies in the development of a hybrid HW model that aims to strike a balance between accuracy and computational efficiency for real-time applications. In the following, we first describe the structure of the model as well as the experiments used for fitting the data-driven blocks. Then we evaluate the accuracy of the model in a number of validation experiments. Finally, we compare it to a first-order-plus-dead-time model as conventional, linear model to assess the benefit of retaining the nonlinear mechanistic part.

METHODOLOGY

BSM1 Model

The base model configuration of BSM1 consists of an activated sludge unit for biological wastewater treatment, followed by a circular secondary settler for separating treated water from sludge. The model consists of 205 differential equations and a minimum of 65 algebraic equations, depending on the specific implementation, and has 28 individual control handles [14]. The BSM1 reflects the standard design of many real-world secondary treatment stages in WWTPs. BSM1 provides standardized simulation environments to evaluate control strategies in SIMULINK.

In this work, we use the BSM1 in two ways: 1. We use a steady-state version of it within the HW structure as described below. 2. We use the full dynamic BSM1 as virtual plant, i.e., to generate data for the identification of the dynamic part of the HW structure, and as well as the ground truth for evaluating the accuracy of the HW model.

Hammerstein-Wiener Model

The Hammerstein-Wiener model is commonly used in nonlinear dynamic system identification. As illustrated in Figure 2, it is constructed from a linear block between two static nonlinear transformations.

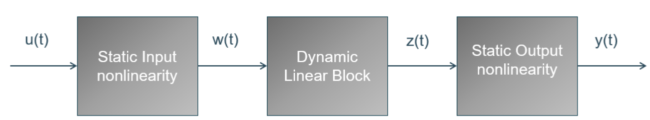

Figure 2. Hammerstein-Wiener Structure

The formulation for a Hammerstein-Wiener model is presented in Equation (1)-(4) provided by Wills et al. [13]:

$$w(t) = f_H(u(t)), \forall t \in [t_0, t_f] \quad (1)$$

$$\dot{x}(t) = Ax(t) + Bw(t), \forall t \in [t_0, t_f] \quad (2)$$

$$z(t) = Cx(t) + Dw(t), \forall t \in [t_0, t_f] \quad (3)$$

$$y(t) = f_W(z(t)), \quad \forall t \in [t_0, t_f] \quad (4)$$

Here, $f_H: \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_w}$ represents the Hammerstein function for the input nonlinearity and $f_W: R^{n_z} \rightarrow R^{n_y}$ the

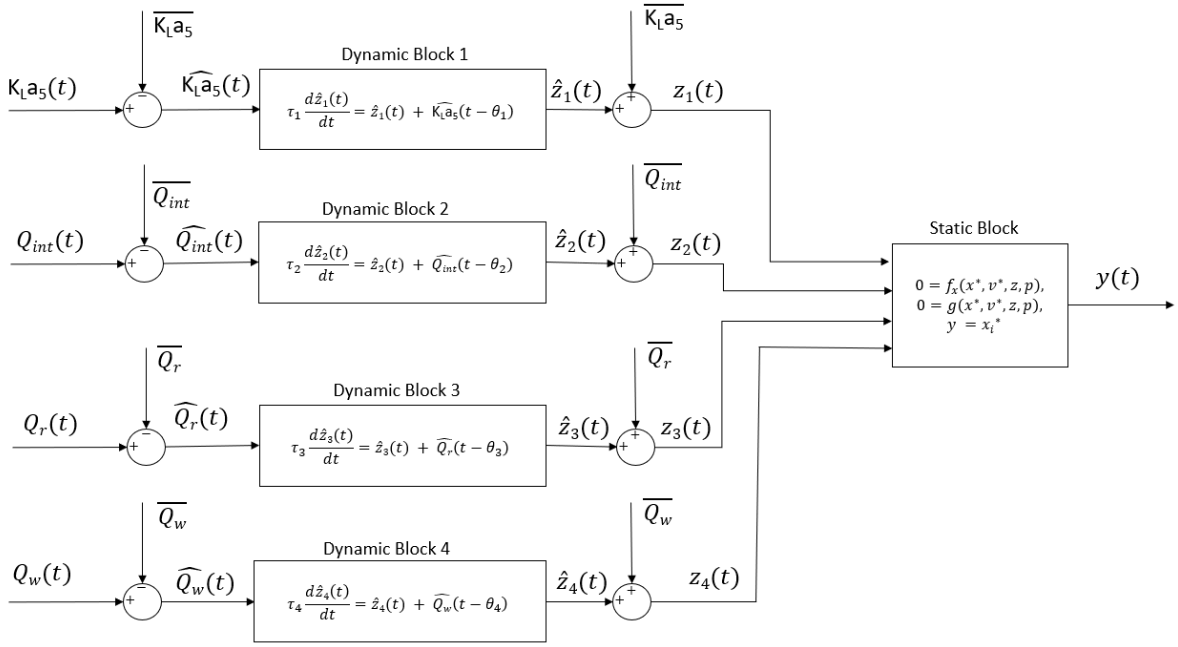

Figure 3. Hammerstein-Wiener surrogate model. $y(t)$ can be any of the 17 outputs in Table 1; the same model structure must be run separately for each output

Wiener function for the output nonlinearity. The inputs of the system are represented as $u: [t_0, t_f] \to \mathbb{R}^{n_u}$, while the inputs to the linear time-invariant (LTI) system are denoted by $w: [t_0, t_f] \to \mathbb{R}^{n_w}$. The system states are given by $x: [t_0, t_f] \to \mathbb{R}^{n_x}$, and the outputs of the LTI system are described by $z: [t_0, t_f] \to \mathbb{R}^{n_z}$. Finally, the overall system outputs are expressed as $y: [t_0, t_f] \to \mathbb{R}^{n_y}$.

Hammerstein-Wiener models can also be observed as two simpler variants: the Hammerstein model uses a nonlinear-then-linear (N-L) structure, and the Wiener model follows a linear-then-nonlinear (L-N) structure.

Hybrid modeling Approach

The HW model developed is inspired by the Wiener structure (linear-then-nonlinear) and takes 4 inputs: aeration intensity in the 5th aeration tank (K_La_5), internal recycle flow (Q_{int}), sludge recycle flow (Q_r), and waste sludge flow (Q_w). This selection was based on the most common control handles of a WWTP. Furthermore, aeration intensity and internal recycle flow have a built-in PID controller in BSM1, which provide a ground for control comparison. The modeled outputs are the effluent variables of secondary clarifier in BSM1, which can be found in Table 1. The resulting HW model architecture is shown in Figure 3.

Note that, to capture the distinct system dynamics of each variable, we developed a series of Multi-Input Single-Output (MISO) structures, where a dedicated HW model is trained for each specific effluent output. This is because the time constants τ_i and dead times θ_i of the FOPDT models in Equation (7) only captures the relation between a single input and output. Therefore, different outputs will require different parameter values. For example, ammonium concentrations and active heterotrophic biomass are affected differently by an increase in aeration intensity.

Parallel Linear Dynamic Blocks

The input entering each linear dynamic block i is defined as a deviation form by subtracting the nominal operating point from the actual input value:

$$\hat{u}_i(t) = u_i(t) - \overline{u}_i, (6)$$

where $\hat{u}_i(t)$ is the deviation input, the actual input $u_i(t)$, and nominal operating point $\overline{u}_i$. The linear block models only the time-dependence of the effect of input u_i on the output y, and consists of a FOPDT model with unit gain:

$$\tau_i \frac{d\hat{z}_i(t)}{dt} = -\hat{z}_i(t) + \hat{u}_i(t - \theta_i), \qquad (7)$$

where $\hat{z}_i(t)$ is the internal state (output) of the i^{th} block, $\hat{u}_i(t)$ is the input, τ_i is the time constant, and θ_i is the dead time. The output $\hat{z}_i$ thus represents the same physical quantity as $\hat{u}_i$, but delayed in a way that represents the response *dynamics* of the overall output of the HW model (i.e., the output of the nonlinear block, y).

The parameters of each dynamic block i are only the time constant τ_i and time delay θ_i, which are determined from multiple step responses and validated against response signals from the ground truth system, where in each case the best fitting parameters are chosen based

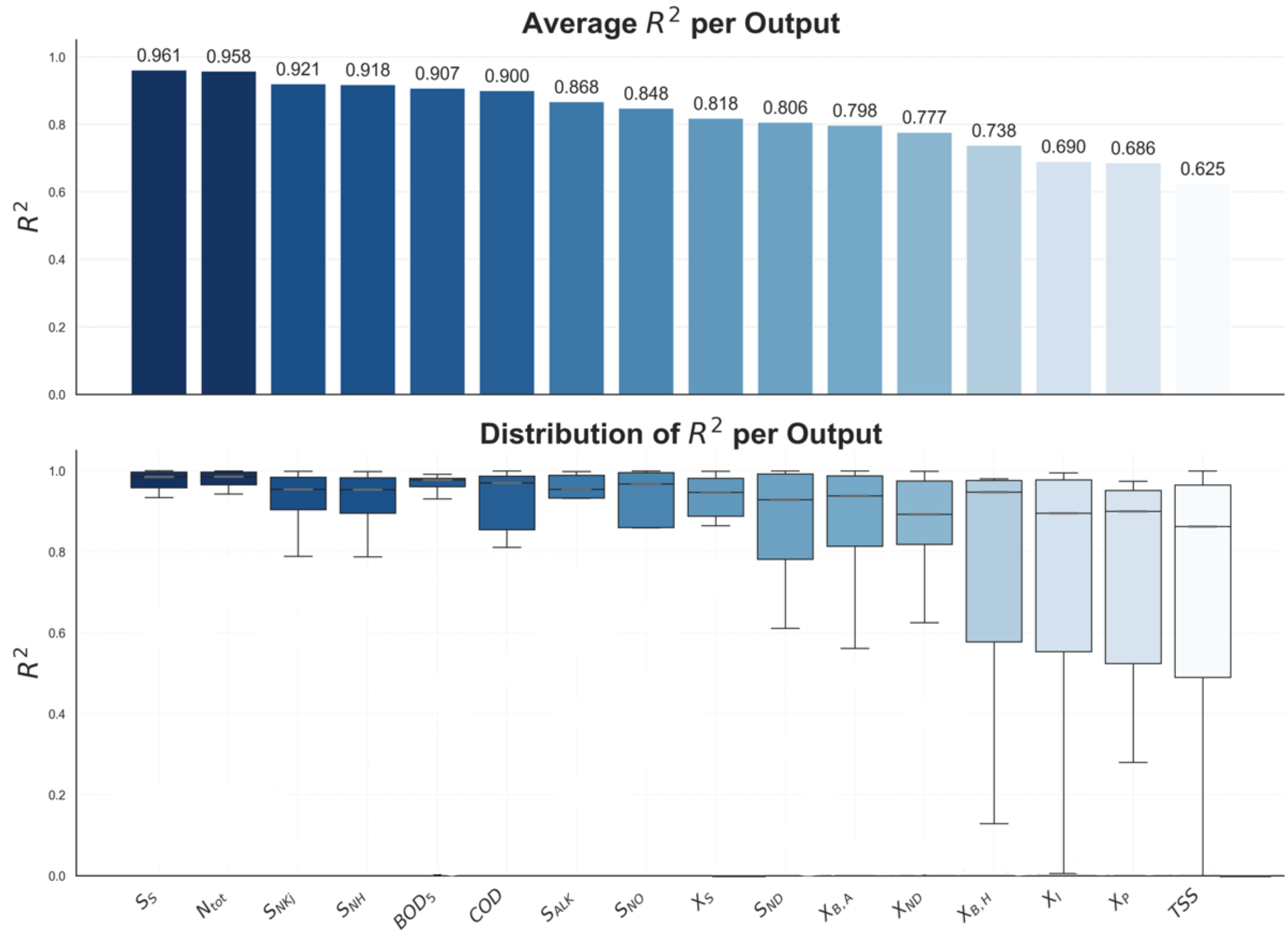

Figure 4: Predictive performance of the HW model across water quality outputs. Averages and distributions are computed over all time points of the 408 validation simulations.

on the coefficient of determination (R^2). The underlying assumption is that τ_i and θ_i stay rather constant across a wide operating range (i.e., that the main nonlinearity is in the gains). Moreover, it is assumed that the interaction between inputs in terms of dynamics can be ignored (superposition principle).

Finally, the operation point value is added to the dynamic block output to finalize the signal that flows into the nonlinear block:

$$z_i(t) = \hat{z}_i(t) + \bar{z}_i \quad (8)$$

where, $z_i(t)$ is the nominal term, $\hat{z}_i(t)$ is the deviation, and $\bar{z}_i = \bar{u}_i$ is the nominal operating point.

This modeling approach was used to enable dynamic prediction without introducing excessive model complexity, and much smaller data requirements than full black-box models. These smaller data requirements are enabled by the superposition principle of linear differential equations.

Nonlinear Static Block

This block consists of a steady-state model derived from the BSM1 equation set. Since the original BSM1 is implemented in Simulink, which limits the efficient computation of steady-states, we developed a standalone version in CasADi using the equations from [14]. This new model version allows for the calculation of steady states at a significantly lower computational cost. Note that, because the inputs are transformed based on their dynamic relationship with the output (captured via time constants and dead times), the steady-state block's solution represents a quasi-steady state, meaning the approximate response of the output to time-dependent input changes. The HW model is considered to be at a true steady state only when $u(t) = z(t)$, indicating the system has fully settled following input perturbations.

To compute the steady states, the BSM1, which is a nonlinear DAE, was converted into an algebraic system as follows. Let $x(t) \in \mathbb{R}^{n_x}$ denote the differential states, $v(t) \in \mathbb{R}^{n_v}$ the algebraic states, $z(t) \in \mathbb{R}^{n_z}$ the inputs, for which we here use the inputs transformed by the linear blocks rather than the actual external inputs to the system, and $p \in \mathbb{R}^{n_p}$ the model parameters. The BSM1 equations are

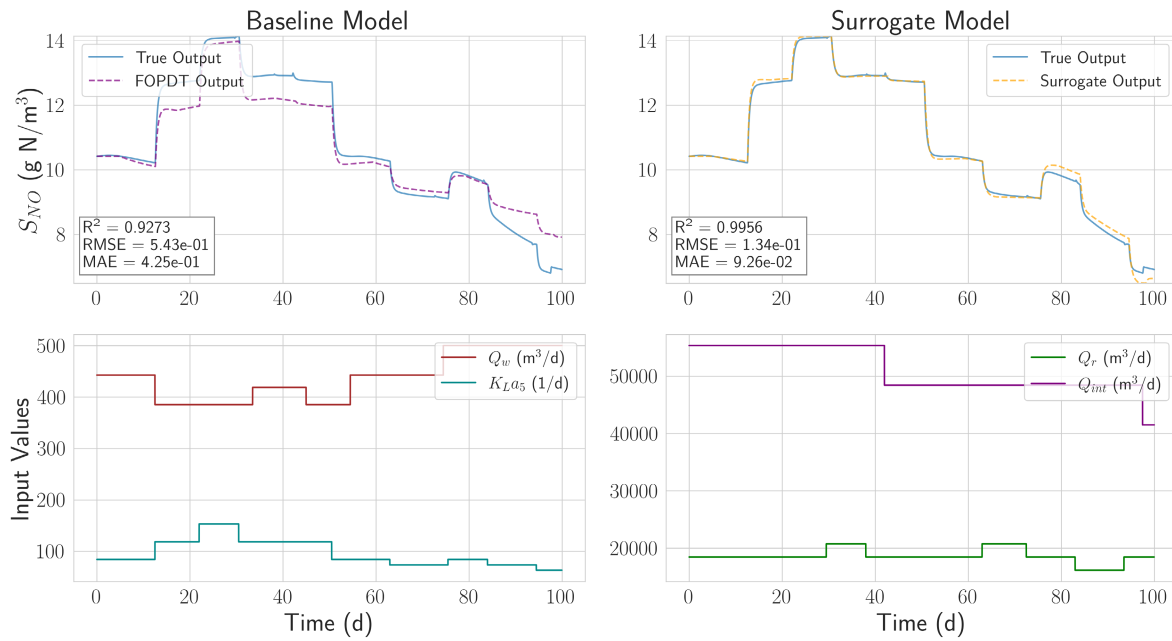

Figure 5. Comparison of a FOPDT baseline model (top left, dashed purple) against the proposed HW smodel (right, dashed orange) for nitrate and nitrite concentration, both in comparison to the BSM1 as ground truth (solid blue). The inputs are shown in the bottom left and right plots. Despite being shown in different figures, they were applied simultaneously to the models but were plotted on different graph because of the order of magnitude for the inputs.

$$\dot{x}(t) = f_x(x(t), v(t), z(t), p), (9)$$

$$0 = g(x(t), v(t), z(t), p), (10)$$

where f_x defines the differential equations (bioreactor, clarifier, and sensors), and g defines the algebraic equations (clarifier mappings and output connections). Computing the steady states reduces to solving the system

$$0 = f_x(x^*, v^*, z, p), (11)$$

$$0 = g(x^*, v^*, z, p), (12)$$

where (x^*, v^*) represent the steady-state values for the differential and algebraic variables, respectively. The effluent concentration i can be extracted from x^*:

$$y = x_i{}^* \ (15)$$

The fact that only one output can be extracted for solving the steady state model (as described above) is a limitation of the proposed structure.

Model Training and Validation

All data was sampled at 15 minute intervals. This was an upper limit of sampling time for the regression to capture the fastest dynamic in the system. For the training data, step responses were applied to the ground truth BSM1, starting from different operating points. Given the linear dynamic models used, all inputs were varied separately to fit the respective blocks. For each input, 20 points evenly spaced across the feasible range of input values were used as initial operating points as well as final points of the input steps. From these 20 step responses per input, a FOPDT model was fitted to every input-output combination through least-squares regression, resulting in 20 parameters sets (K_i, τ_i, θ_i) per output (i.e., one per input for every output). The gains K_i were then discarded, since this they are not used in the HW model.

The validation dataset was generated utilizing a Pseudorandom Multilevel Sequence (PRMS) with Pseudorandom Frequency Modulation (PRFM) as input signal. The multilevel amplitudes varied across different experiments: some only covered the training space, while other extrapolated beyond the training space while remaining within the physical operating bounds of the plant. . The frequency modulation was varied from a $1/(0.5\ \text{days})$ to $1/(12\ \text{days})$ across experiments. This was designed to validate all aspects of the model, including extrapolation capabilities, accuracy under fast and slow dynamics, and ability to predict steady states. The data was generated with the same sampling rate as the training data. In total, 408 output trajectories were generated by applying the described inputs signals.

Model validation was performed in two steps; first the best among the 20 best parameter sets per input-output combination was selected based on the coefficient of determination (R^2), which was computed along

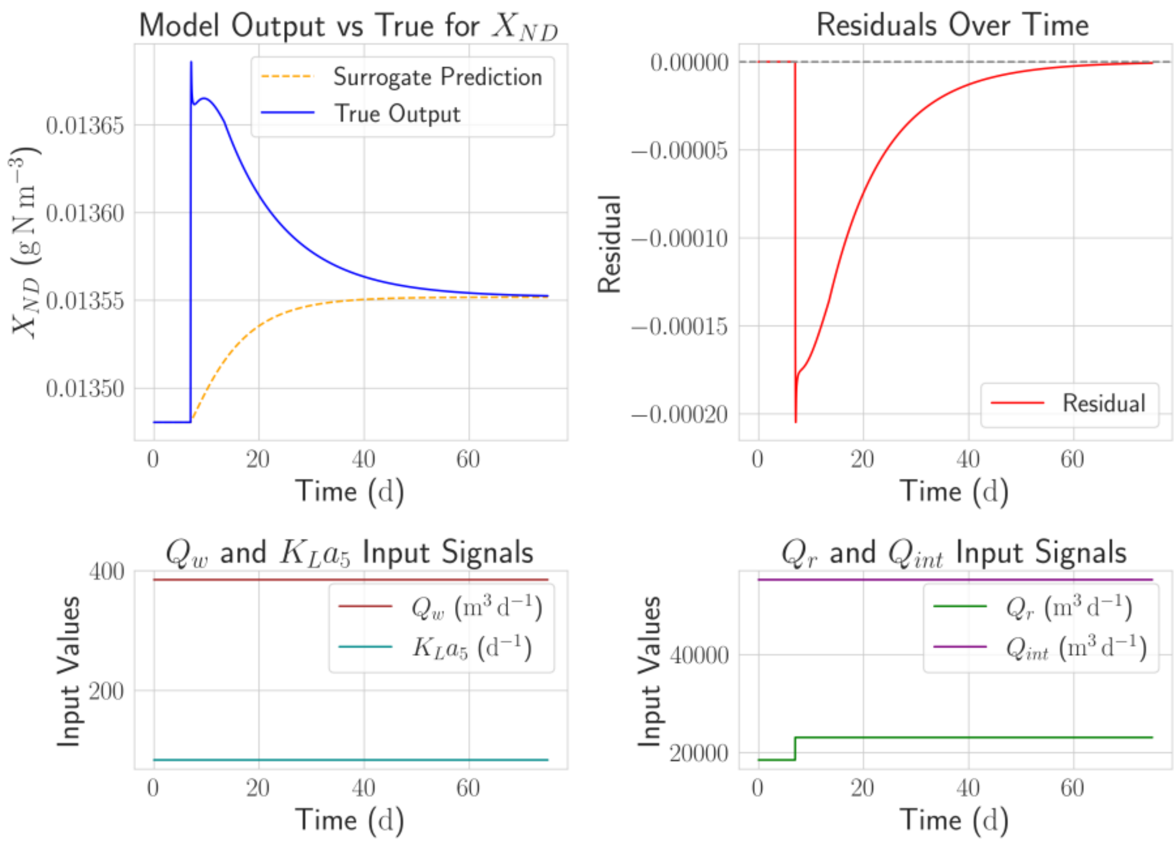

Figure 6. The particulate biodegradable organic nitrogen concentration cannot be captured by the proposed HW model with first-order dynamics because of its complex transient behavior.

the output trajectory. Afterwards, these best parameters were specified in the model.

In the second step of model validation (considered evaluation), the HW model's fidelity was tested. For each simulation the coefficient of determination (R^2) was captured for the whole trajectory of the simulation.

To evaluate the benefit of capturing the nonlinear gain through the mechanistic model in the HW structure, a pure FOPDT model was also developed as conventional low-order linear model. It was fitted to a step response where the final value was the upper bound of the training input range. As an extension, a HW model with a Second Order Plus Dead Time (SOPDT) function in its dynamic blocks was designed to consult test if an extra parameter in the structure leads to better predictive capabilities.

Although the BSM1 exhibits a unique steady state, the nonlinear blocks of the HW structure could introduce artificial steady-state multiplicity. To verify the reliability of the proposed model, we performed a multiplicity analysis using Latin Hypercube Sampling (LHS) to generate different initial guesses to the steady-state solver, ensuring the HW model maintains a unique, physically consistent mapping across the operational space.

RESULTS & DISCUSSION

Across all 408 validation simulations, the average R^2 for the outputs, as seen in Figure 4, achieved consistently high values, often exceeding 0.9, indicating accurate replication of both steady-state and dynamic behavior. The most important results from these experiments were the good predictive performance ($R^2 > 0.9$) across key water quality indicators such as readily biodegradable substrate (S_s), total nitrogen (N_{tot}), Kjeldahl nitrogen (S_{NKj}), and ammonium (S_{NH}), as well as Biochemical Oxygen Demand (BOD_5) and Chemical Oxygen Demand (COD). Lower-performing outputs were linked to specific limitations, particularly input-output relationships exhibiting complex nonlinear dynamics beyond FOPDT approximation. An example for such a case is shown in Figure 6.

The HW model achieved substantially better accuracy than the pure FOPDT model in the validation simulation: the HW model consistently attained higher coefficients of determination (R^2) for the whole trajectory of a simulation, often exceeding 0.90 whereas the FOPDT model yielded noticeably lower R^2 values. An example is shown in Figure 5. Preliminary experiments indicated that extending the HW model's linear blocks to second-order (SOPDT) dynamics yielded negligible improvements in predictive accuracy over the first-order structure.

Steady-state multiplicity analysis confirmed the presence of multiple equilibria in roughly 10% of tested conditions. These multiplicities were concentrated in specific input regions, most notably when the wastage sludge flow Q_w dropped below $300\ \mathrm{m}^3 d^{-1}$. This confirms that in operating regimes where multiplicity is present the HW model is not reliable.

Comparing the HW model to traditional linear hybrid structures, such as ARMAX, highlights the potential advantages of the our approach. For instance, while an ARMAX model achieved an $R^2 \approx 0.67$ for effluent COD in a specific industrial WWTP [9], our results suggest that augmenting low-order dynamics with nonlinear static map can further enhance predictive capabilities by capturing the system's inherent nonlinearities more effectively.

CONCLUSION

This work demonstrates that a hybrid Hammerstein–Wiener model can effectively approximate the dynamics of key outputs in a wastewater treatment process by combining mechanistic steady-state modeling with data-driven linear dynamic blocks. This separation enables reasonably accurate dynamic prediction without introducing excessive model complexity, and much smaller data requirements than full black-box models. Moreover, the HW model retains the gain of the original rigorous model which is useful for control and optimization applications. The linear dynamics of the HW model can be an advantage also because they can be exploited algorithmically in optimization algorithms [15]. These properties make the Hammerstein–Wiener framework particularly well suited for integration into DRTO or NMPC and allow operation with online optimization crucial for WWTPs that operate under increasingly strict regulatory constraints.

Table 1: Inputs and outputs of Hammerstein-Wiener model

Symbol	Description	Unit
K_{La5}	Oxygen transfer coefficient in reactor	d^{-1}
Q_{int}	Internal recycle flow rate	$\mathrm{m}^3\,\mathrm{d}^{-1}$
Q_r	Return sludge flow rate	$\mathrm{m}^3\,\mathrm{d}^{-1}$
Q_w	Waste sludge flow rate	$\mathrm{m}^3\,\mathrm{d}^{-1}$
N_{tot}	Total nitrogen in effluent	$\mathrm{g\,m}^{-3}$
$\mathrm{COD}_{\mathrm{tot}}$	Total chemical oxygen demand in effluent	$\mathrm{g\,m}^{-3}$
S_{NH}	Ammonium nitrogen in effluent	$\mathrm{g\,m}^{-3}$
TSS	Total suspended solids	$\mathrm{g\,m}^{-3}$
BOD_5	5-day biochemical oxygen demand in effluent	$\mathrm{g\,m}^{-3}$
S_I	Soluble inert organic matter	$\mathrm{g\,m}^{-3}$
S_S	Readily biodegradable substrate	$\mathrm{g\,m}^{-3}$
X_I	Particulate inert organic matter	$\mathrm{g\,m}^{-3}$
X_S	Slowly biodegradable substrate	$\mathrm{g\,m}^{-3}$
$X_{B,H}$	Active heterotrophic biomass	$\mathrm{g\,m}^{-3}$
$X_{B,A}$	Active autotrophic biomass	$\mathrm{g\,m}^{-3}$
X_P	Particulate products from biomass decay	$\mathrm{g\,m}^{-3}$
S_O	Dissolved oxygen	$\mathrm{g\,m}^{-3}$
S_{NO}	Nitrate and nitrite nitrogen	$\mathrm{g\,m}^{-3}$
S_{ND}	Soluble biodegradable organic nitrogen	$\mathrm{g\,m}^{-3}$
X_{ND}	Particulate biodegradable organic nitrogen	$\mathrm{g\,m}^{-3}$
S_{ALK}	Alkalinity	$\mathrm{mmol\,L}^{-1}$

ACKNOWLEDGEMENTS

DB gratefully acknowledges funding through Internal Funds KU Leuven (STG/22/060). JM gratefully acknowledges funding through Internal Funds KU Leuven (STG/23/057).

AUTHOR IDENTIFIERS

Author ORCIDs:
Bongartz D: 0000-0003-1790-0235
Matias Assumpcao J: 0000-0002-1094-6145

REFERENCES

1. European Parliament and Council of the European Union. Directive (EU) 2024/3019 concerning urban wastewater treatment (recast). Dec. 2024. url: https://eurlex.europa.eu/eli/dir/2024/3019/oj.
2. Olsson G. ICA and me – a subjective review. Water Research 46:1585-1624 (2012). https://doi.org/10.1016/j.watres.2011.12.054
3. Petersen B, Gernaey K, Devisscher M, Dochain D, Vanrolleghem PA. A simplified method to assess structurally identifiable parameters in monod-based activated sludge models. Water Research 37:2893-2904 (2003). https://doi.org/10.1016/s0043-1354(03)00114-3
4. Alizadeh R, Allen JK, Mistree F. Managing computational complexity using surrogate models: a critical review. Res Eng Design 31:275-298 (2020). https://doi.org/10.1007/s00163-020-00336-7
5. Schneider MY, Quaghebeur W, Borzooei S, Froemelt A, Li F, Saagi R, Wade MJ, Zhu JJ, Torfs E. Hybrid modelling of water resource recovery facilities: status and opportunities. Water Science and Technology 85:2503-2524 (2022). https://doi.org/10.2166/wst.2022.115
6. Gaublomme D, Strubbe L, Vanoppen M, Torfs E, Mortier S, Cornelissen E, De Gusseme B, Verliefde A, Nopens I. A generic reverse osmosis model for full-scale operation. Desalination 490:114509 (2020). https://doi.org/10.1016/j.desal.2020.114509
7. Wang X, Zhang J, Babovic V. Improving real-time forecasting of water quality indicators with combination of process-based models and data assimilation technique. Ecological Indicators 66:428-439 (2016). https://doi.org/10.1016/j.ecolind.2016.02.016
8. Nordstrand J, Dutta J. Flexible modeling and control of capacitive-deionization processes through a linear-state-space dynamic langmuir model. npj Clean Water 4: (2021). https://doi.org/10.1038/s41545-020-00094-y
9. Tomperi J, Sorsa A, Ruuska J, Ruusunen M. Series-connected data-based model to estimate effluent chemical oxygen demand in industrial wastewater treatment process. Journal of Environmental Management 373:123680 (2025). https://doi.org/10.1016/j.jenvman.2024.123680
10. Iratni A, Chang NB. Advances in control technologies for wastewater treatment processes: status, challenges, and perspectives. IEEE/CAA J. Autom. Sinica 6:337-363 (2019). https://doi.org/10.1109/jas.2019.1911372
11. Han HG, Zhang Y, Sun HY, Liu Z, Qiao J. Data-knowledge-driven multiobjective adaptive optimal control for wastewater treatment processes under multiple operating conditions. IEEE Trans. Cybern. 55:1056-1069 (2025). https://doi.org/10.1109/tcyb.2025.3531514
12. Wu XL, Han WH, Yang HY, Li X, Han HG. Robust soft constrained model predictive control and its application in wastewater treatment processes. IEEE Trans. Automat. Sci. Eng. 22:13198-13211 (2025). https://doi.org/10.1109/tase.2025.3552421
13. Wills A, Schön TB, Ljung L, Ninness B. Identification of hammerstein–wiener models. Automatica 49:70-81 (2013). https://doi.org/10.1016/j.automatica.2012.09.018
14. J. Alex, L. Benedetti, J. Copp, K.V. Gernaey, Ulf Jeppsson, I. Nopens, M.-N. Pons, J.-P. Steyer, and P.A. Vanrolleghem. Benchmark Simulation Model no. 1 (BSM1). Vol. TEIE-7229. Lund University, 2008.
15. Kappatou CD, Bongartz D, Najman J, Sass S, Mitsos A. Global dynamic optimization with hammerstein–wiener models embedded. J Glob Optim 84:321-347 (2022). https://doi.org/10.1007/s10898-022-01145-z