Evaluating the performance of CE-QUAL-W2 version 4.5 sediment diagenesis model

Manuel Almeida[1], Pedro Coelho[1]

[1]MARE - Marine and Environmental Sciences Centre, ARNET - Aquatic Research Network Associate Laboratory, NOVA School of Science and Technology, NOVA University Lisbon, Caparica, 2896-516, Portugal

Correspondence to: Manuel Almeida (mcvta@fct.unl.pt)

Abstract. This research evaluates the performance of the CE-QUAL-W2 v4.5 sediment diagenesis model in simulating water temperature, dissolved oxygen, total phosphorus, total nitrogen, chlorophyll-a, and biochemical oxygen demand in a Portuguese reservoir over a six-year period (2016–2021). The model was calibrated using 35 observed profiles of temperature and dissolved oxygen, as well as six annual measurements of total nitrogen, total phosphorus, chlorophyll-a, and biochemical oxygen demand at multiple depths. To benchmark performance, three alternative sediment oxygen demand formulations—a Zero-order, First-order, and a Hybrid model combining both approaches—were also implemented and compared. All models achieved NSE and RMSE values within or near the ranges reported in the literature, effectively capturing the system's water quality dynamics. Among them, the Hybrid model yielded the best overall performance while maintaining a simpler structure (Water temperature - NSE: 0.96±0.18; RMSE: 1.09±0.23 ºC; Dissolved oxygen - NSE: 0.76±0.30; RMSE: 1.87±0.72 mg/L). The sediment diagenesis model exhibited similar performance metrics (Water temperature - NSE: 0.95 ± 0.18; RMSE: 1.13 ± 0.28 °C; Dissolved oxygen - NSE: 0.71 ± 0.14; RMSE: 2.01 ± 0.59 mg/L). Overall, the results suggest that the diagenesis model may be better suited for capturing detailed process-based dynamics over extended timeframes, whereas simpler models, such as the Hybrid model, are more appropriate for short- to medium-term applications or situations with limited data availability. Hopefully, the results of this study will help improve water management strategies by supporting more informed model selection tailored to the temporal scope and data constraints of reservoir monitoring programs.~~This study set out to assess the performance of the state-of-the-art CE-QUAL-W2 v4.5 sediment diagenesis model. The model was applied to a reservoir in Portugal using observed sediment particulate organic carbon values corresponding to a six-year period (2016-2021). The model was calibrated by comparing its results with 35 observed dissolved oxygen and water temperature profiles, as well as annual total nitrogen, total phosphorus, biochemical oxygen demand, and chlorophyll-a measurements corresponding to three different depths. In addition to model calibration, a sensitivity analysis was also conducted by varying the input particulate organic carbon values and applying a user-specified sediment oxygen model (zero-order model). The results demonstrated the overall effectiveness of the sediment diagenesis model, which accurately simulated dissolved oxygen profiles, nutrient concentrations, and organic matter levels (Dissolved oxygen profiles: NSE = 0.41 ± 0.67; RMSE = 1.73 mg/L ± 0.69), highlighting its potential as an effective tool for simulating lakes and reservoirs and supporting water management processes. The study further suggests that the zero-order model is~~

~~able to serve as an effective starting point for implementing the sediment diagenesis model, providing an initial estimate for mean reservoir sediment oxygen demand (SOD) values.~~

1 Introduction

Modeling water quality plays a crucial role in managing lakes and reservoirs, providing essential insights into the dynamics of nutrients, organic matter, and phytoplankton within aquatic systems (Abbaspour et al., 2015; Whitehead et al., 2009). These models simulate the physical, chemical, and biological processes that influence water quality, with examples including widely-used tools like CE-QUAL-W2 (Wells, 2021), MIKE21 (Chapman, 1996), and DYRESM (Hamilton and Schladow, 1997). The value of such modeling lies in its capacity to aid researchers and policymakers in understanding the complex interactions between various factors that impact the ecological health of water bodies (Varis et al., 1994; Loucks and Beek, 2017). However, the intricacy of these systems, combined with the substantial data requirements, often presents significant challenges for those developing and applying water quality models. Effective inflow data characterization (quantity and quality) is hard to obtain, both for major river branches and small tributaries, as is waterbody sediment characterization related to carbon and nutrients due to the significant cost associated with the sampling and laboratorial analysis process and the fact that water management stakeholders are still more focused on the classification of waterbody water quality rather than the collection of water quality forcing data. The absence of sediment initial particulate organic carbon (POC), particulate organic nitrogen (PON) and particulate organic phosphorus (POP) data can be decisive to the overall performance of a water quality model, in essence generating an imbalance right from the start of the simulation with regard to the sediment concentration of POC, PON and POP, which then has a considerable impact on the SOD and, consequently, the waterbody dissolved oxygen (DO). When calibrating the model, water quality modelers therefore need to plug this gap by evaluating the model performance considering: i) different initial sediment oxygen demand (SOD) where a zero-order model is applied, ii) different POC, PON and POP values where a predictive diagenesis model is considered.

The main challenge with these modeling approaches is that the sources of DO depletion—such as the inflow of organic matter or algal mortality—can significantly influence DO dynamics, and these sources must be well characterized to ensure accurate predictions. While the baseline model can reproduce observed DO profiles with reasonable accuracy, its predictive reliability may be compromised if key DO sinks and sources are not well defined.

For example, the model’s response to a reduction in external phosphorus loading is influenced by internal phosphorus release from sediments during anoxic periods. In CE-QUAL-W2, when a zero-order SOD model is used, the anoxic release of phosphate (PO_4) is modeled as a linear function of SOD: SOD [g O_2/m^2day] × PO_4 release rate [g P/g O_2]. Thus, any error in the estimation of SOD will directly affect the predicted internal phosphorus loading, and by extension, the overall phosphorus balance in the waterbody. In contrast, when using the predictive sediment diagenesis model, internal phosphorus loading depends on the organic and nutrient inputs from particulate matter in the water column and the sediment’s biogeochemical response, which is highly influenced by the initial value of particulate organic carbon (POC). As a result, this approach introduces additional uncertainty when key particulate components are not adequately measured or constrained in both the water column and sediments. Calibrating other constituents, such as orthophosphate (P-PO_4), can help reduce uncertainty. P-PO_4 is released from sediments under anaerobic conditions, and its calibration can enhance the accuracy of DO modeling. Still, this release is influenced by multiple factors, including the initial sediment P-PO_4 concentration and the release rate (in the zero-order model), or the mineralization of POP (in the diagenesis model). In both cases, significant uncertainty remains without observed data for POC, PON, and POP in both the water column and sediments. Of these, POC has the most significant influence on SOD, making access to sediment POC data essential for improving model accuracy, even when PON and POP measurements are lacking. ~~The main problem with these solutions is that the source of DO decay can also be driven by the inflow of organic matter or algae mortality. In other words, the final baseline model results are able to fit the observed DO profiles with reasonable accuracy but the model prediction capability may be unbalanced if the DO sinks/sources are not well defined. For example, the model's response to a reduction of inflow phosphorous is a function of the phosphorus released from the sediments during anoxia periods and, therefore, if the SOD is not accurately computed the waterbody phosphorous balance will, in turn, be incorrect. The calibration of other constituents, such as nutrients can help to minimize this uncertainty. For example, orthophosphates (P-PO₄) are released from the sediments under anaerobic conditions and therefore the calibration of this nutrient can help with the overall calibration of the water body DO concentration. The release of P-PO₄ from the sediments is, however, a function of several variables, for example, the initial P-PO₄ dissolved concentration and the P-PO₄ release rate (zero-order approach), or a diagenesis rate for POP (diagenesis model). In other words, the modeling uncertainty may diminish but will persist without observed POC, PON and POP. Of the three variables (POC, PON, and POP), POC has the most significant impact on SOD. Consequently, access to sediment POC values is crucial for ensuring accurate modeling even where PON and POP concentrations are unavailable.~~

The CE-QUAL-W2 model has been widely used to simulate various water bodies and water quality scenarios, including reservoir physical and biochemical dynamics in response to warming projections (Mi et al., 2020; Mi et al., 2023). This model has also been used to predict DO in a number of water bodies worldwide, although the SOD has always been modeled with a zero-order and/or 1-order model (e.g. Park et al., 2014; ~~Zouabi~~ Zouabi-Aloui et al., 2015; Terry et al., 2017; Sadeghian et al., 2018; Lindenschmidt et al., 2019). The bibliographic research conducted before and during this study suggests that the CE-QUAL-W2 sediment diagenesis model has not been applied to any waterbodies other than the Wahiawa Reservoir in central Oahu (Berger and Wells, 2014). Moreover, no scientific publications on the evaluation of this model in other contexts have been identified, further highlighting the importance of the primary motivation for this study, namely, to evaluate the performance of the CE-QUAL-W2 model with its new sediment diagenesis component. This study benefited from having access to observed reservoir sediment total organic carbon (TOC) values, which are rare. Although, in theory, these values are typically higher than particulate organic carbon (POC) values, they provided an excellent starting point for this study. The methodological approach was, therefore, defined to evaluate the performance of the CE-QUAL-W2 model considering the new state-of the art sediment diagenesis model in modeling a reservoir, DO, Total Phosphorus (Total P); Total Nitrogen (Total N), Biochemical Oxygen Demand (BOD_5), Chlorophyll~~-~~-a (Chl-a) and SOD.

To achieve this, the water quality of a highly productive reservoir was simulated over a six-year period (2016–2021) using the CE-QUAL-W2 v4.5 model. The simulation incorporated a Zero-order sediment model, a First-order model, a Hybrid model combining both approaches, and a sediment diagenesis model. The Zero-order, First-order, and Hybrid models were included to provide alternative representations of sediment oxygen demand, enabling comparative analysis and supporting the calibration and evaluation of the more complex sediment diagenesis model. ~~To achieve this, the water quality of a highly productive reservoir was simulated using the CE-QUAL-W2 v4.5 model, incorporating both a zero-order sediment model and the sediment diagenesis model over a six-year period, spanning 2016 to 2021. It should be noted that the zero-model was included in this study to back-calculate the DO uptake rate in the reservoir water column as a function of the reservoir boundary conditions. The zero-order model is not a predictive model, as it fails to include the accumulation of particulate organic matter and algae in the sediments and rates do not vary over time, except as a result of decay rate temperature dependence. Nevertheless, the consideration of this model proved useful for addressing potential issues related to the calibration quality of the sediment diagenesis model. If the predicted zero-order model DO profiles in the water column are exactly matched, then the values for SOD used in calibration are very close to the actual dissolved oxygen uptake rates in the water column (Wells, 2011). Therefore, if the W2 model with the sediment diagenesis function performs in the same way under the same boundary conditions as the zero-order model in terms of DO in the water column we can conclude that the performance of the sediment diagenesis function is accurate.~~ In the case of water temperature and DO, the modeling results were compared with 35 water column profiles observed near the dam. The remaining parameters were calibrated using time series datasets collected at multiple depths, with six annual values available for each parameter. ~~The remaining parameters were calibrated against time series data sets observed at different depths.~~ A sensitivity analysis was performed to evaluate the reservoir water quality response, namely DO, to the variation of POC, PON and POP concentration in the reservoir sediments. The results of this study will hopefully prove useful by helping to improve lake and reservoir water quality modeling and, therefore, the water management process from a practical perspective.

2 Methods

2.1 Site Location and Main Characteristics

Portugal experiences a temperate maritime climate characterized by a wet, cool season and a dry summer. Despite most of the precipitation occurring during the winter months, there is significant inter-annual variability. Precipitation patterns are spatially and temporally heterogeneous, with annual maxima exceeding 2500 mm in the rugged highlands of the northwest, while the low-lying plains of the southeast receive around 400 mm per year (Cardoso et al., 2013; Soares et al., 2015) (Fig.1). The Torrão dam, located in the northern region of mainland Portugal in the Tâmega River, is a significant hydraulic structure designed for multiple purposes, including water supply, irrigation, and hydroelectric power generation. The reservoir has a substantial storage capacity, contributing to regional water management and flood control. This infrastructure plays a crucial role in the socio-economic development of the region, balancing resource management and environmental preservation. However, it is also important to note that the reservoir was classified as eutrophic for all the simulated years, a condition that can lead to persistent water quality issues.

Table 1: Main features of Torrão dam and reservoir

Full supply volume (hm^3)	Mean inflow (m^3/s)	Average annual inflow (local basin)	Active storage volume (hm^3)	Surface area at FLV (km^2)	Structural height (m)	Max depth (m)	Turbine number/power	Hydraulic residence time (days)	Watershed area (km^2)	Trophic state (2016 - 2021)(1)
123.9	76.98	2 147	40.4	6.5	70	58	2 reversible pump-turbines/146 (MW)	13.59	3 252	Eutrophic

(1) Classification according to OECD Trophic State limits (OECD, 1982)

~~2.2 Modeling **approach**~~

~~The bathymetry of the Torrão reservoir was initially defined using a Digital Elevation Model (DEM) provided by Energies of Portugal, S.A. (EDP) and structured according to the methodology outlined in Wells (2021). The reservoir comprises one main branch (the Tâmega River), three tributaries and one distributed tributary (Fig. 1). Tributaries 1 and 2 are depicted in Fig 1. Tributary 3 represents the inflow from the Douro River into the pump-back system of the Torrão Reservoir. The bathymetric map includes 27 segments, each measuring 1000 meters in length, and a maximum number of 58 layers, each with a depth of 1 meter. Following this preliminary step, the reservoir boundary conditions (including water quality, hydrology, meteorology, and sediment characterization) were defined according to the methods described in Section 1.2.3. Due to the lack of available information, the model structure only includes a single algae group (Diatoms). Subsequently, two different CE-QUAL-W2 model versions (see section 2.3) were calibrated for the 2016-2021 period (vide section 1.2.4): a) incorporating the sediment diagenesis model (W2_SD), and b) considering a user-specified sediment oxygen model (zero-order model) that was not coupled with the water column (W2_zero-order). The zero-order model was selected for the analysis because it typically provides an accurate approximation of the sediment oxygen demand (SOD) in a reservoir without relying on sediment concentrations or requiring a separate sediment compartment and was thus a useful tool for~~

~~establishing a SOD benchmark for the reservoir. This is not, however, a predictive approach, as, other than variations resulting from the temperature dependence of the decay rate, the rates remain constant over time (Wells, 2021).~~

~~The models were calibrated by adjusting their parameters to improve the fit between the model output and observed data. Please refer to Wells (2021) for a detailed account of the model calibration parameters and default values. Water temperature was the first constituent to be calibrated. The wind sheltering coefficient (WSC) was manually adjusted to achieve the best fit between the modeled and observed water temperature profiles, resulting in a final value of 1. A value of 1 implies that the WSC has no effect over the wind velocity forcing the model. The zero-order model for sediment oxygen demand (SOD) was then manually adjusted to improve dissolved oxygen (DO) predictions based on 35 DO profiles. The optimal result was achieved with a constant SOD value of 2.5 g~~ ~~O_2/m^2/day~~. ~~Following this calibration, the phosphorus sediment release rate (PO4R) in the zero-order model was modified from its default value of 0.001 to 0.015. All other parameters were kept at their relevant default values and the default settings for the sediment diagenesis model were also maintained. The observed data included water temperature (WT), dissolved oxygen (DO), Total P, Total N, BOD5, and chlorophyll-a. These two runs were named: W2_SD_baseline and W2_zero-order_baseline. A sensitivity analysis was conducted after the calibration process to evaluate the model's response: a) to different POC, PON and POP values in the case of the W2_SD model; and b) to different SOD values in the case of the W2_zero-order model. Section 1.2.5 details the methodological approach employed for the sensitivity analysis. The evaluation of the model's performance, along with the results obtained from the sensitivity analysis, provided deeper insights into modeling SOD using the diagenesis model.~~

2.2~~3~~. CE-QUAL-W2 v4.5 model

This study employed the latest version of CE-QUAL-W2 (Version 4.5), a model originally developed in 1975 by the US Army Corps of Engineers and written in Fortran. Since its inception, the model has undergone regular updates and enhancements, primarily by researchers at Portland State University (Cole and Wells, 2006). CE-QUAL-W2 is a two-dimensional, laterally averaged hydrodynamic and water quality model capable of simulating free surface elevation, hydrostatic pressure, density, horizontal and vertical velocities, as well as constituent concentrations. The model uses the finite difference method to solve key equations, including mean transverse momentum in the x- and z-directions, the continuity equation, state equations, and water surface elevation equations (~~Adelena et al., 2015;~~ Tavera-Quiroz et al., 2024; Wells, 2021). This model represents SOD through four distinct approaches: (i) a user-defined zero-order formulation that is decoupled from the water column, (ii) a simple predictive first-order model, (iii) a hybrid approach combining the zero- and first-order methods, and (iv) a comprehensive sediment diagenesis model. The zero-order model is not a predictive approach, as, other than variations resulting from the temperature dependence of the decay rate, the rates remain constant over time (Wells, 2021). Additionally, under anoxic conditions in the water column, SOD is disabled in the model. The first-order sediment model does not function as a full sediment diagenesis model, as it lacks the capability to track the fate of organic nutrients delivered to the sediments, their breakdown, and the release of byproducts into the water column under low-oxygen conditions. However, it does represent the deposition of particulate organic matter and dead algal biomass, along with the resulting oxygen demand imposed on the water column. By including this first-order sediment process, the model becomes sensitive to increased organic loading to the sediment, which in turn influences sediment oxygen demand. A combination of the zero and first order model can be considered where organic materials accumulate and decay in the sediments under aerobic conditions and are released based on the SOD zero-order decay rate under anaerobic conditions. In contrast, the sediment diagenesis model simulates kinetic processes occurring within the sediment and at the sediment–water interface. This module originally developed for application ~~A sediment diagenesis module, originally tailored for~~ in oil sand pit lakes, has been adapted for application in other aquatic environments and integrated into version 4.0 (Vandenberg et al., 2015). The conceptual framework of the model has been elaborated in works by Prakash et al. (2014), Berger and Wells (2014), and Vandenberg et al. (2015). It is important to note that significant enhancements to the sediment diagenesis module were introduced in version 4.5 of the model, as detailed in the User Manual (Wells, 2021). These improvements mark a substantial advancement over the initial version 4, which was more limited in its capabilities. The CE-QUAL-W2 model has demonstrated its utility in simulating hydrodynamic and ecological processes—such as stratification, internal waves, ~~oxygen~~ DO dynamics, and phytoplankton blooms—in lakes and reservoirs worldwide (Zhang et al., 2015; Chuo et al., 2019; Kobler et al., 2018; Uhlmann, 2017; Terry et al., 2017; Mi et al., 2020). Additional details about the model's structure, algorithms, and historical applications can be found in the user manual (Wells, 2021).

2.2.1 Model Setup

The bathymetry of the Torrão reservoir was initially defined using a Digital Elevation Model (DEM) provided by Energies of Portugal, S.A. (EDP) and structured according to the methodology outlined in Wells (2021). The reservoir comprises one main branch (the Tâmega River), three tributaries and one distributed tributary (Fig. 1). Tributaries 1 and 2 are depicted in Fig 1. Tributary 3 represents the inflow from the Douro River into the pump-back system of the Torrão Reservoir. The bathymetric map includes 27 segments, each measuring 1000 meters in length, and a maximum number of 58 layers, each with a depth of 1 meter. Following this preliminary step, the reservoir boundary conditions (including water quality, hydrology, meteorology, and sediment characterization) were defined according to the methods described in Section ~~1.2.3.~~2.4. Due to the lack of available information, the model structure only includes a single algae group (Diatoms).

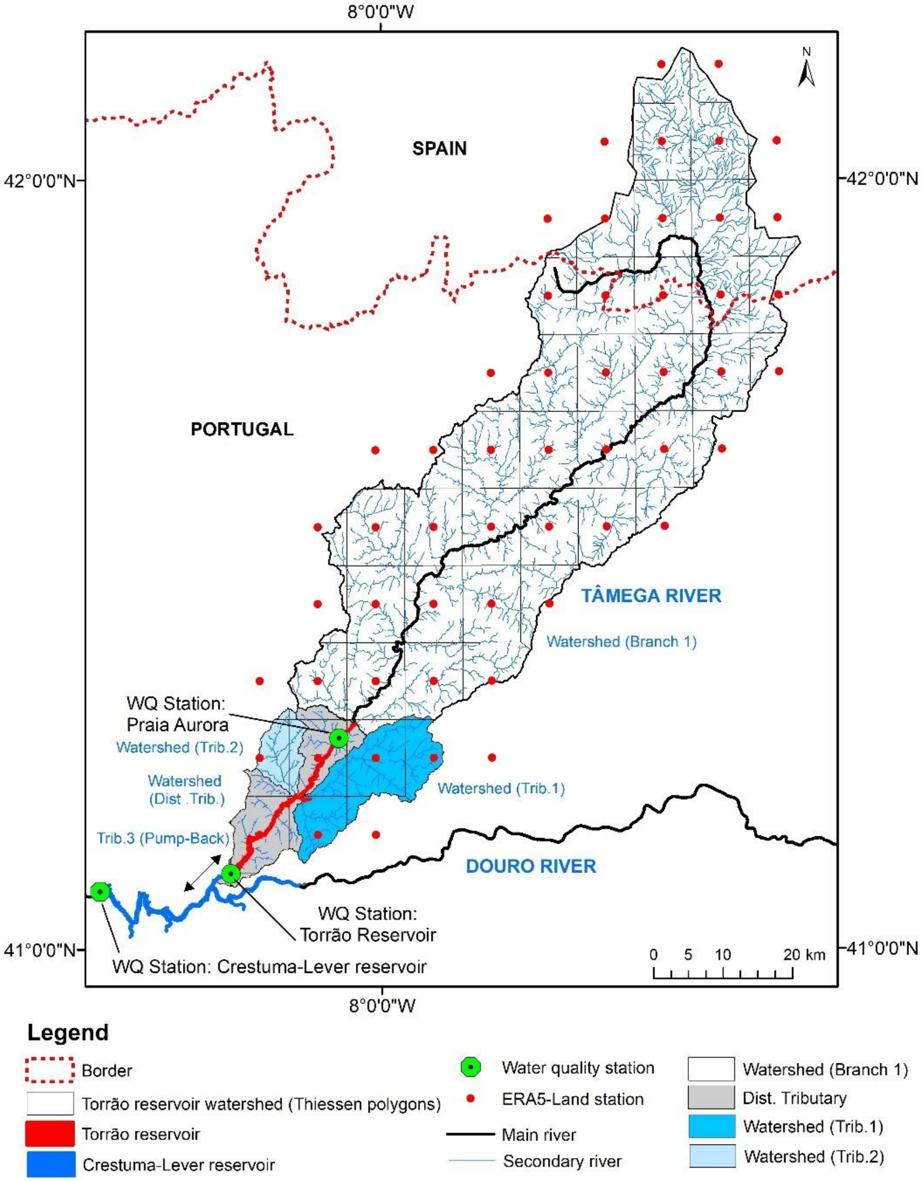


Figure 1~~1~~: Torrão reservoir watershed. Thiessen polygons. Water quality stations

2.3 Modeling approach

To thoroughly evaluate the capability of CE-QUAL-W2 in modeling dissolved oxygen using the sediment diagenesis module, the four available SOD modeling approaches were considered: Zero-order model; First-order model; Zero/First-order model (Hybrid model) and the sediment diagenesis model (SG model). The models were calibrated for the 2016–2021 period (see Section 2.5). During the results analysis, the performance metrics obtained during each model's calibration process were compared, along with the SOD values across the bottom layers of each model. A sensitivity analysis was conducted following calibration to evaluate each model's response: a) to varying POC, PON, and POP values in the case of the SG model; b) to different SOD values in the Zero-order and Hybrid models and c) to varying the initial first order sediment concentration in the case of the First-order model. Section 2.6 details the methodological approach used for the sensitivity analysis. To assess the sensitivity of each model to reductions in external organic matter (OM) and phosphorus (PO_4-P) inputs, two separate scenario analyses were conducted. The first scenario involved an 80% reduction in OM inflow load, while the second applied an 80% reduction in both OM and PO_4-P inflow loads. These reductions were implemented specifically in the main reservoir branch (Branch 1 – Tâmega River), where the majority of nutrient and organic inputs occur. Each sediment model—SD, Zero-order, First-order, and Hybrid—was run under baseline conditions and under both reduction scenarios. The impact on DO dynamics was evaluated using time series of depth- and segment-averaged DO concentrations. Each model—SD, Zero-order, First-order, and Hybrid—was run under baseline conditions and then under this reduced-loading scenario. The evaluation of model performance, along with the results of the sensitivity analysis, provided deeper insights into simulating SOD dynamics using the sediment diagenesis approach in comparison to the other SOD formulations.

2.~~4~~ 4 Model Forcing Datasets

The meteorological data used to drive the model, including hourly air temperature, dew point, solar radiation, cloud cover, and wind characteristics, were sourced from ERA5-Land, a high-resolution reanalysis dataset optimized for land applications. Although no on-site meteorological stations are available in the study area for direct validation, studies by Almeida and Coelho (2023b) and Barbosa et al. (2022) have demonstrated a strong correlation between ERA5-Land air temperature data and observed measurements at regional scales, supporting the reliability of this dataset for our modeling purposes. Furthermore, the accuracy of water temperature predictions in our simulations indicates that the meteorological forcing was well represented, confirming the suitability of ERA5-Land data for driving the model. Reservoir data, such as daily inflow/outflow, water levels, and water quality, covering the years 2016–2021, were provided by EDP. Water quality data specific to Branch 1 originated from the Praia Aurora Station, accessed via the Portuguese National Water Resources Information System (SNIRH, 2024). With only 21 recorded measurements for Branch 1 during this period, three modeling methods were employed to address the 99.04% of missing data. The variables include: water temperature; DO; Total P; Ammonium (N-NH_4); Nitrate+Nitrite (N-NO_X~~; –~~); BOD_5~~)~~; ~~Chlorophyll-a~~Chl-a; Alkalinity; Conductivity and Total

Suspended Solids (SST). The first method employed regression models implemented through the LOADEST package (Runkel et al., 2004) developed by the U.S. Geological Survey. The second method utilized the Extreme Gradient Boosting (XGBoost) machine learning algorithm, implemented using the Chen and Guestrin (2016) open-source library, a method proven effective in various environmental studies (Feigl et al., 2021; Adedeji et al., 2022; Xu et al., 2022). For additional details on the algorithm, refer to Almeida and Coelho (2023b). The third approach relied on Support Vector Regression (SVR), implemented via the scikit-learn library (Pedregosa et al., 2011), which has also demonstrated strong performance in environmental modeling applications (Adedeji et al., 2022; Ji and Lu, 2018). For machine learning approaches, datasets were split into training (80%) and testing (20%) sets. Hyperparameters for these models were optimized using the Tree-structured Parzen Estimators (TPE) algorithm, executed with the Hyperopt library (Bergstra et al., 2013) and 100 iterations. The Nash-Sutcliffe Efficiency (NSE) was used to determine the best model. Table A1 describes the input features of each model. Correlations derived from Branch 1 informed data extrapolation to other tributaries using flow as the predictor. Observed data for Tributary 3 was retrieved from the Crestuma-Lever reservoir monitoring station.

Water quality variables used for model inputs included water temperature, DO, orthophosphates (P-PO_4), N-NH_4, N-NOx, labile and refractory dissolved and particulate organic matter (LDOM, RDOM, LPOM, RPOM), alkalinity, inorganic suspended solids (ISS), total dissolved solids (TDS), total inorganic carbon (TIC), and algal biomass (diatoms). For non-monitored variables, estimations were made based on available data: i) P-PO_4: Derived from total phosphorus, assuming inorganic phosphorus represents 70% of the total; ii) Organic matter: BOD_5 was converted to organic matter using a stoichiometric ratio of 1.4 g O_2/1.0 g organic matter, with 60% assumed refractory and 40% labile; iii) ISS: Estimated as 97.4% of TSS; iv) TDS: Calculated from electrical conductivity (Eq.1); v) TIC: Estimated from alkalinity (Eq. 2); vi) Algae biomass: Chl-a~~Chlorophyll-a~~ was converted to biomass using the following ratio: Algal Biomass (mg/L)/ Chl-a ~~Chlorophyll-a~~ (µg/L) = 0.05

$$\text{TDS (mg/L)}=0.65\times\text{Electrical Conductivity}(\mu\text{S/cm}) \quad (1)$$

$$\text{TIC (mg/L)} = 0.2782\times\text{Alkalinity (mg/L)}0.9706 \quad (2)$$

This equation was derived from the relationship between TIC and alkalinity values observed in four reservoirs within the United States, utilizing a dataset comprising 55232 value pairs available in the CE-QUAL-W2 v4.5 model examples (Wells, 2021). The analysis achieved an R^2 value of 0.99.

Figure 2 illustrates the locations of the five sediment sampling sites used to define the SD model baseline run. The spatial distribution of the sediment samples depicted in the figure were linked to specific reservoir segments to characterize the initial sediment content of ~~particulate organic carbon (~~POC~~)~~, ~~particulate organic nitrogen (~~PON~~)~~, and ~~particulate organic phosphorus (~~POP~~)~~, as detailed in Table 2. Sediment values were assigned as follows: site A to segments 25–28, site B to segments 20–24, site C to segments 16–19, site D to segments 11–15, and site E to segments 2–10. Several assumptions were made to establish the sediment characterization: i) A sediment density of 960 kg/m³ (density of dried sediment with air in the pore space) (Minear, 2007) was applied to convert sample values from mg/kg to mg/L; ii) POP values were set at 25 mg/L, based on established literature benchmarks (Wells, 2021); iii) The T~~otal ~~N value observed at site B was used to characterize sites C, D, and E; iv) TOC and T~~otal ~~N were assumed to exist entirely in particulate form, represented as particulate organic carbon and nitrogen. This approach ensured a consistent and representative characterization of sediment properties across the reservoir segments.

Table 2: Torrão Reservoir sediment chemical characterization obtained for each sampling site.

Observed values			
Sampling site	**TOC, mg/kg**	**T~~otal ~~N, mg/kg**	**POP, mg/kg**
A	25 000	6020	-
B	20 900	5990	-
C	22 300	-	-
D	20 100	-	-
E	5 600	-	-
Final values included in the sediment diagenesis model			
	POC, mg/L	**PON, mg/L**	**POP, mg/L**
A	24 000	5779	25
B	20 064	5750	25
C	21 408	5750	25
D	19 296	5750	25
E	5376	5750	25

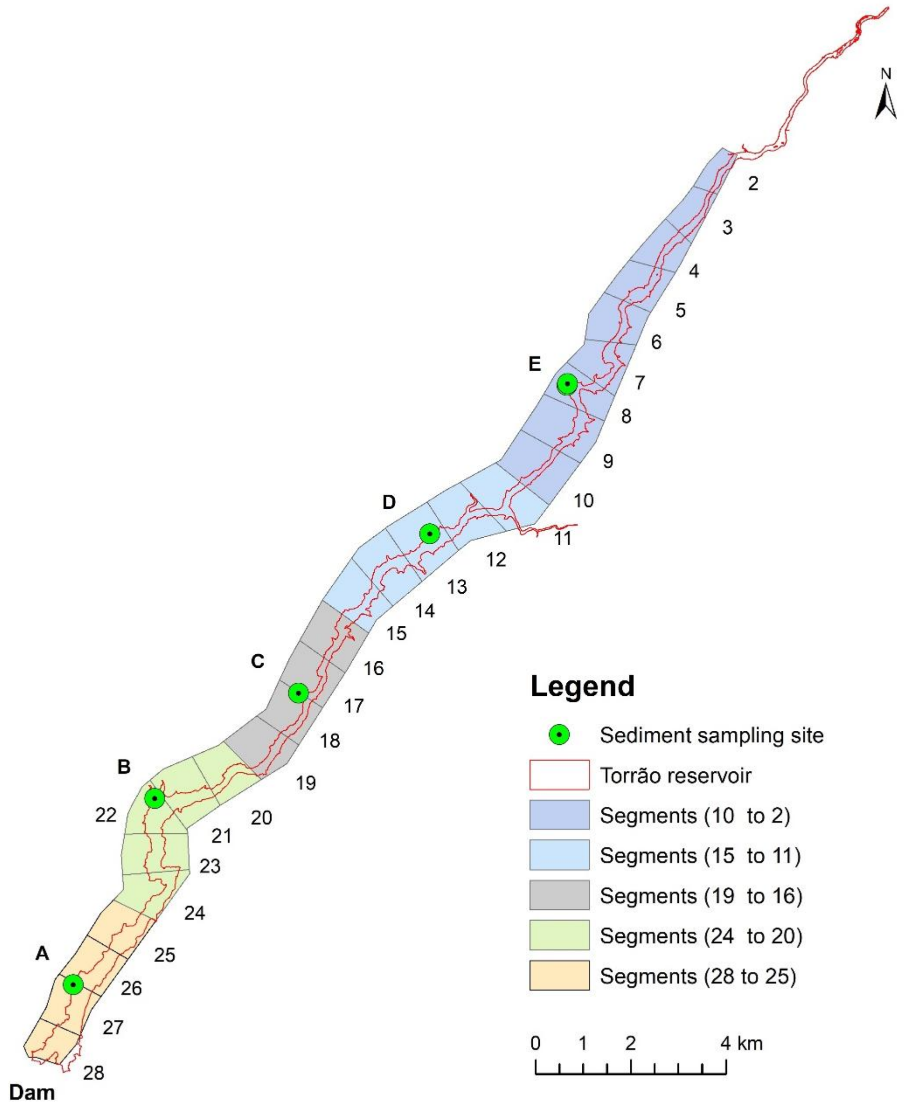


Figure 2: Sediment sampling sites. CE-QUAL-W2 model segments

2.~~5~~ 5 Water Quality Model (CE-QUAL-W2) Calibration

The simulation period considered for this study spanned 2016 to 2021. This period was selected due to the availability of flow and water quality data. The trial-and-error technique was applied to calibrate the model for the simulation period, considering the default calibration parameters described in Wells, 2021. The error between observed and predicted values of six state variables was evaluated with five different metrics (~~vide~~ see section ~~1.2.6~~2.7). The observed data included six annual values for water temperature~~ (WT)~~, ~~dissolved oxygen (~~DO~~)~~, T~~otal~~ P, T~~otal~~ N, BOD_5, and Chl-a~~chlorophyll-a~~. These time series were obtained from: (a) an integrated sample between the reservoir surface and a depth of 5.8 meters, (b) a depth of 23 meters, and (c) a depth of 43.7 meters. In addition, 35 water temperature and DO profiles—six per year from 2016 to 2021—were also included.~~ 35 water temperature and DO profiles, with six profiles per year from 2016 to 2021, were also considered.~~ These profiles were observed 300 meters upstream from the Torrão Dam. Details on the models' initial conditions, parameters, constants, and forcing datasets can be found in Almeida and Coelho (2025) and in Tables A2 to A8. ~~The models' initial conditions, parameters, constants, and the forcing datasets can be obtained from Almeida and Coelho, 2025.~~The models were calibrated by adjusting their parameters to improve the fit between the model output and observed data. Please refer to Wells (2021) for a detailed account of the model calibration parameters and default values. Water temperature was the first constituent to be calibrated. The wind sheltering coefficient (WSC) was manually adjusted to achieve the best fit between the modeled and observed water temperature profiles, resulting in a final value of 1. A value of 1 implies that the WSC has no effect over the wind velocity forcing the model. The zero-order model for ~~sediment oxygen demand (~~SOD~~)~~ was then manually adjusted to improve ~~dissolved oxygen (~~DO~~)~~ predictions based on 35 DO profiles. The optimal result was achieved with a zero-order SOD value of 2.5 g ~~$O_2/m^2/day$~~O_2/m^2day. Following this calibration, the phosphorus sediment release rate (PO4R) in the zero-order model was modified from its default value of 0.015 to 0.001. The same process was applied to the Hybrid model, where the best results were achieved using a PO4R of 0.001 and a zero-order SOD value of 1.0 g ~~$O_2/m^2/day$~~O_2/m^2day. In the first-order model, the PO_4R parameter was adjusted to 0.001, and the initial concentration of first-order sediment was set to 0.5 g/m². ~~For the first-order model, the PO4R was also modified to 0.001, while the initial first-order sediment concentration was maintained at its default value of 0.0 g/m².~~ All other parameters were kept at their relevant default values and the default settings for the sediment diagenesis model were also maintained. The observed data included water temperature~~ (WT)~~, ~~dissolved oxygen (~~DO~~O)~~,~~,~~ ~~Total~~ P, ~~Total~~ N, BOD$_5$, and Chl-a~~chlorophyll-a~~.

2.~~5~~ 6 Sensitivity analysis

A sensitivity analysis was conducted after the calibration process to evaluate the model's response:

i) Different initial sediment values for POC, PON, and POP were used in the SD model (Table 3). It is important to note that for each of the 24 runs described in Table 2, only the corresponding parameter was modified, while the other two parameters retained their default values shown in Table ~~1~~3. The number of runs varying the PON and POP values is higher than the number of runs considered for POC, with 6 versus 9 runs, respectively. This adjustment was necessary to achieve a minimal RMSE in the predictions of dissolved oxygen in the water column;~~.~~

ii) Different zero-order SOD values for the Zero-order model (0.5, 1.0, 1.5, 2.0, 2.5 and 3.0 g O_2/m^2.~~/~~day);

~~ii)~~iii) Different initial first order sediment concentration (ISC) for the First-order model (0.0, 0.5, 1.0, 1.5, 2.0, 2.5 and 3.0 g/m^2);

~~iii)~~iv) Different zero-order SOD values for the Hybrid model (0.5, 1.0, 1.5, 2.0, 2.5 and 3.0 g O_2/m^2.day.

In the results analysis for each run and for both scenarios (i) and (ii), the prediction error for ~~dissolved oxygen (~~DO~~)~~ was compared with the ~~sediment oxygen demand (~~SOD~~)~~ values derived from each model. Specifically, ~~Runs~~runs 5, 8, and 20 were forced with the ~~particulate organic carbon (~~POC~~)~~, ~~particulate organic nitrogen (~~PON~~)~~, and ~~particulate organic phosphorus (~~POP~~)~~ values defined in the SD model baseline run.

Table 3. TOC, PON, and POP initial sediment values for the SD model sensitivity analyses.

Parameter	Sampling site/Model segments					~~W2~~ SD model run
	A	B	C	D	E	
	28-25	24-20	19-16	15-11	10-2	
POC	4 800	4 013	4 282	3 859	1 075	1
	9 600	8 026	8 563	7 718	2 150	2
	14 400	12 038	12 845	11 578	3 226	3
	19 200	16 051	17 126	15 437	4 301	4
	24 000	20 064	21 408	19 296	5 376	5 (SD model-baseline)
	28 800	24 077	25 690	23 155	6 451	6
PON	8 669	8 625	8 625	8 625	8 625	7
	5 779	5 750	5 750	5 750	5 750	8 (SD model-baseline)
	2 890	2 875	2 875	2 875	2 875	9
	1 445	1 438	1 438	1 438	1 438	10

	722	719	719	719	719	11
	361	359	359	359	359	12
	181	180	180	180	180	13
	90	90	90	90	90	14
	45	45	45	45	45	15
POP	128	128	128	128	128	16
	85	85	85	85	85	17
	57	57	57	57	57	18
	38	38	38	38	38	19
	25	25	25	25	25	20 (SD model-baseline)
	13	13	13	13	13	21
	6	6	6	6	6	22
	3	3	3	3	3	23
	2	2	2	2	2	24

2.~~6~~ 7 Metrics

The evaluation of model calibration and the analysis of quantitative differences across simulation scenarios utilized various performance metrics. These included the root mean square error (RMSE), mean absolute error (MAE), Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970), percent bias (PBIAS), and the coefficient of determination (R^2). The calculations were carried out using equations where m_i and o_i represent the simulated and observed values, respectively, and $\bar{o_i}$ the observed values mean.

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(m_i - o_i)^2} \tag{3}$$

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|m_i - o_i| \tag{4}$$

$$NSE = 1 - \left[\frac{\sum_{i=1}^{N}(o_i - m_i)^2}{\sum_{i=1}^{N}(o_i - \bar{o_i})^2}\right] \tag{5}$$

$$PBIAS = \frac{\sum_{i=1}^{N}(o_i - m_i)}{\sum_{i=1}^{N}(o_i)} \times 100 \tag{6}$$

$$\text{R}^2 = \frac{\sum_{i=1}^{N}(m_i - \bar{o})^2}{\sum_{i=1}^{N}(o_i - \bar{o})^2} \times 100 \tag{7}$$

3 Results

3.1 Observed Inflow Water Quality Characterization

The SVR algorithm was more effective at predicting the inflow water temperature compared to the other models. The R^2 and PBIAS values achieved with the SVR were 0.87, and 3.77%, respectively, indicating that the water temperature trends and average magnitudes are well described (Table A1). Additionally, the RMSE and MAE values of 2.1ºC and 1.6ºC, respectively, demonstrate an accurate approximation of the observed datasets. The SVR algorithm was also the best model in predicting DO. The R^2, PBIAS, RMSE, and MAE values reached, 0.91, 0.92%, 0.40 mg/L and 0.26 mg/L, respectively, indicating that the model performed well. This was not the case for the remaining parameters. In fact, the Loadest regression outperformed the other models for the remaining water quality variables. This was primarily due to the limited number of training samples. Simpler models like regressions can have lower variance (i.e., be less susceptible to overfitting) compared to SVR and XBOOST algorithms. Overall, the PBIAS obtained for NH_4, N-NOx, and Chl-a ~~Chlorophyll-a~~ (10.88%, 43.64%, and 30.00%) suggests that the average magnitude was reasonably well represented.

3.2 CE-QUAL-W2 calibration

Tables A2 through A8 display the most significant CE-QUAL-W2 coefficients obtained after the calibration process. The results of the calibration process for all models, are presented in Table 4 and Table ~~A2~~ A9 and illustrated in figures 3 to 6 and figures 8 and 9.~~.~~ The performance metrics for water temperature across the different sediment models show consistent accuracy, with NSE and R^2 values ranging from 0.95 to 0.96 and minimal variation across models. The RMSE and MAE for temperature also remain low, indicating reliable thermal performance regardless of the sediment model applied. In contrast, DO predictions show more variability. The Hybrid model achieved the best overall DO performance, with the highest NSE (0.76 ± 0.30) and R^2 (0.76 ± 0.31), as well as the lowest RMSE (1.87 ± 0.72) and MAE (1.22 ± 0.55), while maintaining a near-zero PBIAS (-0.55 ± 11.14), indicating minimal systemic bias. The Zero-order model also performed reasonably well, with slightly lower error metrics than the SD model. The First-order model, however, showed the weakest DO performance, with a lower NSE (0.68 ± 0.22), higher RMSE (2.15 ± 0.82), and a significant negative PBIAS (-12.17 ± 15.44), suggesting an underestimation of oxygen concentrations. Overall, the results suggest that while temperature simulation is robust across all models, DO dynamics are better captured using the Hybrid or Zero-order models, with the Hybrid model offering the most balanced and accurate representation under the tested conditions. However, the differences in performance metrics for DO among the models are relatively small and often fall within overlapping standard deviations, with the exception of the First-order model, which consistently shows lower accuracy and higher bias, suggesting that while the Hybrid model offers slightly better overall performance, the improvements over the SD and Zero-order models are modest and should be interpreted with caution. In terms of nutrient dynamics, the Hybrid and Zero-order models improve TN and TP predictions relative to the SD and First-order models. The Hybrid model, for example, improves TN R^2 to 0.31 and TP to 0.27, although the associated biases remain significant (e.g., −18.75% for TN and +36.49% for TP). BOD_5 and Chl-a remain poorly simulated across all models, with R^2 values consistently low (≤0.06 for Chl-a and ≤0.03 for BOD_5), and large PBIAS values, particularly in the SD and First-order configurations. The Zero-order model slightly reduces bias in Chl-a and Total N compared to the SD model but performs poorly for TP due to a large overestimation (PBIAS = 103.43%) (Fig.4D). Notably, the SD and First-order models failed to reproduce observed phosphorus release events from sediments on 2018-09-18, 2020-09-08, and 2021-08-31 (Figures 3D and 5D). In contrast, the Hybrid model successfully captured these events by modeling phosphorus release as a linear function of SOD, providing a more realistic representation of sediment–water nutrient interactions (Fig.6D). Overall, while no model fully captures the complexity of all constituents, the Hybrid model consistently provides the most balanced and improved representation, particularly for DO and nutrient parameters.

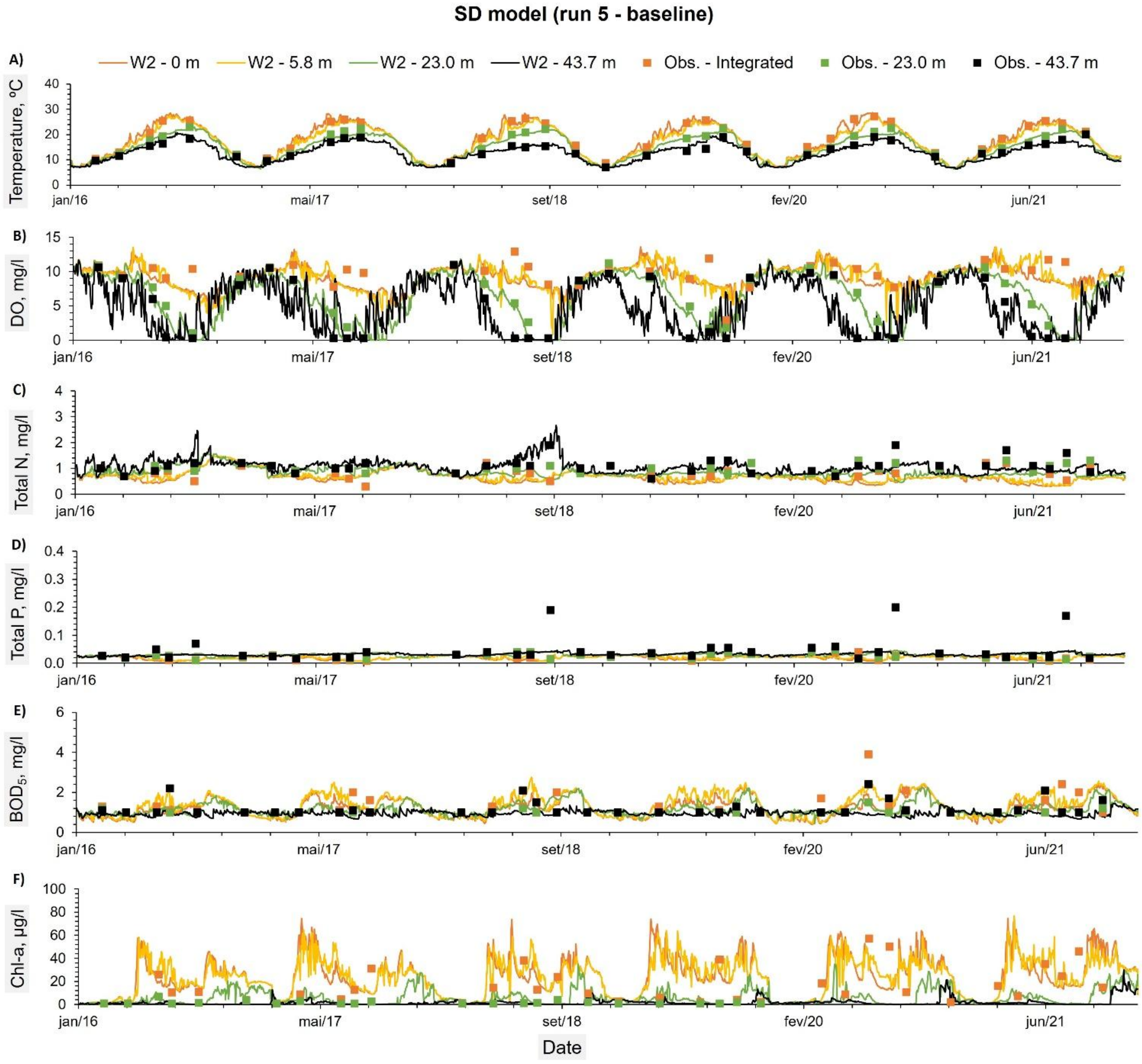
~~Fig. 3 shows that both models predicted the reservoir water temperature reasonably well, a conclusion confirmed by the metrics obtained for this constituent (Table 4). The water column DO was accurately predicted by both models, however, the W2_zero-order model performed slightly better according to all metrics, with the exception of PBIAS. According to Fig.3, the remaining parameters were also well simulated by both models. The only exception to this were the Total N and Total P concentration values predicted by the W2_zero-order model for the reservoir bottom, which were overestimated (Table A2).~~

~~Based on the analysis of Fig. 3 and the values obtained for RMSE, MAE, and PBIAS, it is reasonable to conclude that both models performed equally well in predicting reservoir water quality. However, the W2_zero-order model was slightly better at predicting reservoir DO. Despite this, the model overestimated Chlorophyll-a concentrations from the surface down to a depth of 5.8 meters, and the concentrations of Total P and Total N at the reservoir bottom were also overestimated. During anoxic periods, the model exaggerated the release of $P\text{-}PO_4$ and $N\text{-}NH_4$, an issue that was not mitigated by tuning the release rates of $P\text{-}PO_4$ and $N\text{-}NH_4$ from the sediments, which suggests that the problem was driven by excessive anoxia in the~~

~~bottom layers. Conversely, the W2_SD model provided a more balanced response, predicting the mean evolution of all constituents reasonably well.~~

Table 4: Metrics between observed and predicted values for ~~both~~ all models. Water temperature and DO metrics were obtained from 36 observed and predicted profiles.

Constituent	**~~W2_~~SD model (run 5 - baseline)**				
	NSE	**R^2**	**PBIAS**	**RMSE**	**MAE**
Water temperature	~~0.56~~0.95±~~0.19~~0.18	~~0.56~~0.96±0.07~~13~~	~~-0.11~~1.96±~~0.71~~3.08	~~3.51~~1.13±0.~~32~~28	~~1.56~~0.89±0.~~10~~26
DO	~~0.34~~0.71±0.~~69~~14	~~0.55~~0.73±0.~~31~~29	-~~8.26~~4.43±~~14.70~~15.06	~~1.79~~2.01±0.~~63~~59	~~1.34~~1.38±0.~~47~~46
Constituent	**Zero-order model (zero~~-~~order SOD = 2.5 g O_2/m^2day - baseline)~~W2_zero-order model (SOD = 2.5 g O_2/m^2/day - baseline)~~**				
	NSE	**R^2**	**PBIAS**	**RMSE**	**MAE**
Water temperature	0.95±0.19~~0.56±0.19~~	0.96±0.07~~0.56±0.13~~	1.91±3.09~~-0.12±0.80~~	1.13±0.28~~3.51±0.32~~	0.89±0.25~~1.57±0.10~~
DO	0.73±0.20~~0.50±0.59~~	0.74±0.30~~0.54±0.31~~	1.75±15.87~~-4.77±15.52~~	1.97±0.74~~1.62±0.73~~	1.29±0.57~~1.20±0.55~~
Constituent	**First-order model (ISC= 0.5 g/m² - baseline)**				
	NSE	**R^2**	**PBIAS**	**RMSE**	**MAE**
Water temperature	0.96±0.19	0.96±0.08	1.46±2.97	1.09±0.23	0.85±0.20
DO	0.68±0.22	0.73±0.27	-12.17±15.44	2.15±0.82	1.50±0.65
Constituent	**Hybrid model (zero~~-~~order SOD= 1.0 g O_2/m^2day - baseline)**				
	NSE	**R^2**	**PBIAS**	**RMSE**	**MAE**
Water temperature	0.96±0.18	0.96±0.08	1.45±2.97	1.09±0.23	0.85±0.20
DO	0.76±0.30	0.76±0.31	-0.55±11.14	1.87±0.72	1.22±0.55


SD model (run 5 - baseline)
A)
W2 - 0 m
W2 - 5.8 m
W2 - 23.0 m
W2 - 43.7 m
Obs. - Integrated
Obs. - 23.0 m
Obs. - 43.7 m
Temperature, ºC
0
10
20
30
40
jan/16
mai/17
set/18
fev/20
jun/21
B)
DO, mg/l
0
5
10
15
jan/16
mai/17
set/18
fev/20
jun/21
C)
Total N, mg/l
0
1
2
3
4
jan/16
mai/17
set/18
fev/20
jun/21
D)
Total P, mg/l
0.0
0.1
0.2
0.3
0.4
jan/16
mai/17
set/18
fev/20
jun/21
E)
BOD$_5$, mg/l
0
2
4
6
jan/16
mai/17
set/18
fev/20
jun/21
F)
Chl-a, µg/l
0
20
40
60
80
100
jan/16
mai/17
set/18
fev/20
jun/21
Date

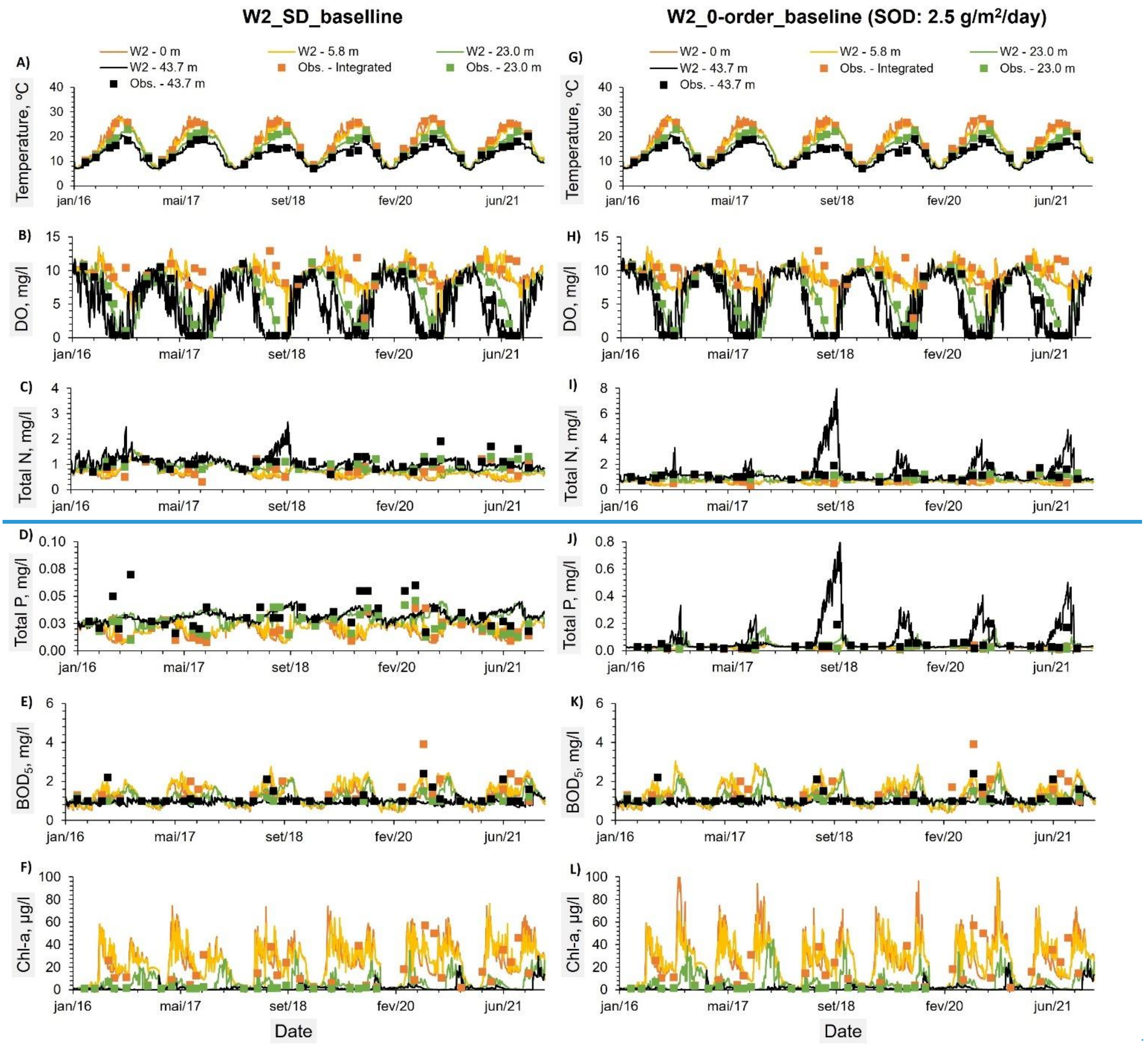
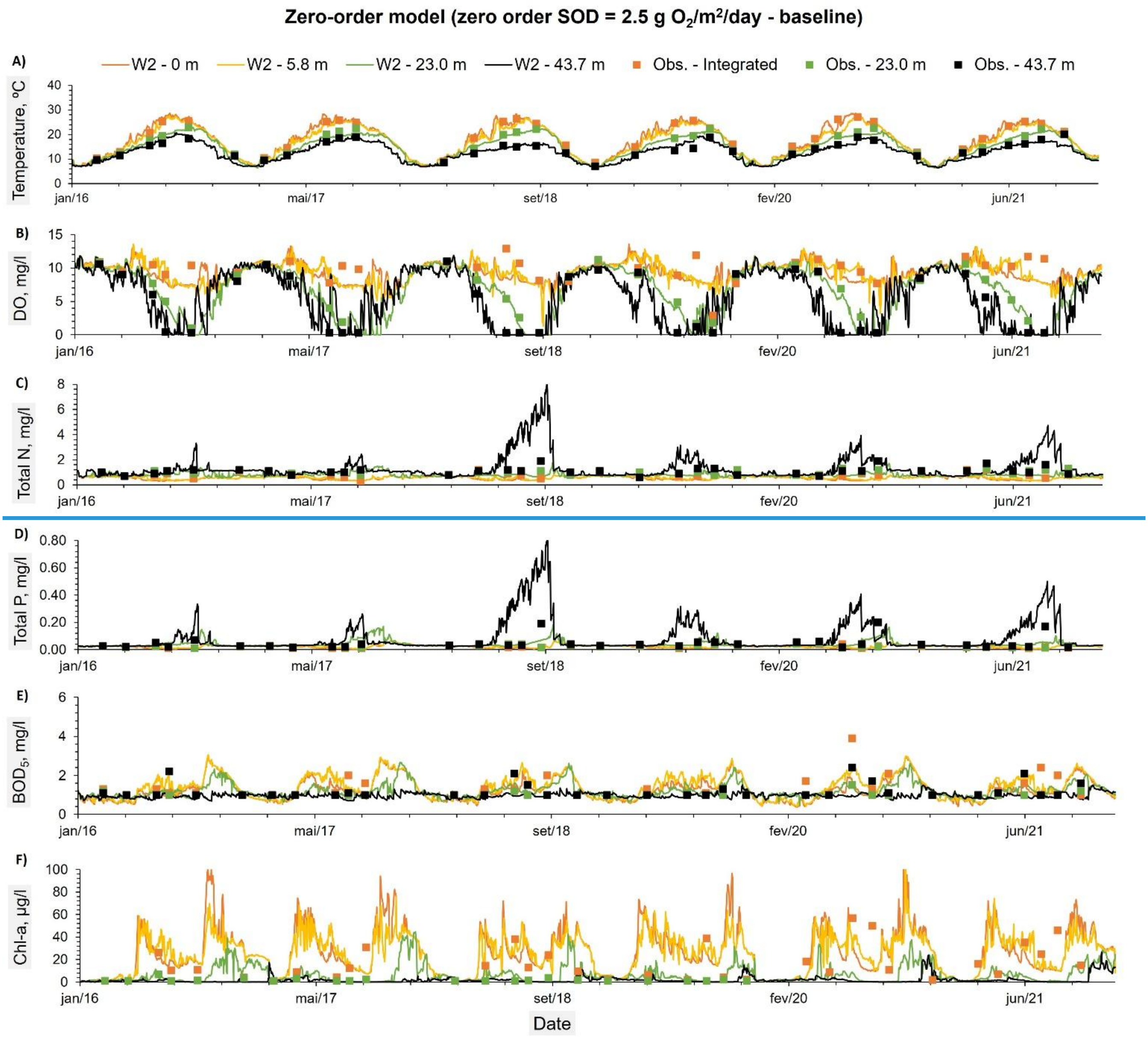


Figure 3: Constituents observed values at three different depths: (a) an integrated sample between the reservoir surface and an average depth of 5.8 meters, (b) an average depth of 23 meters, and (c) an average depth of 43.7 meters. These observed values were compared with the predicted time series from the ~~W2-~~_ SD model (run 5 - baseline)~~SD_baseline~~ (A to F) ~~and W2_zero-order_baseline (SOD: 2.5 g/m^2/day) (G to L)~~ for the same depths.



Zero-order model (zero order SOD = 2.5 g O_2/m^2/day - baseline)

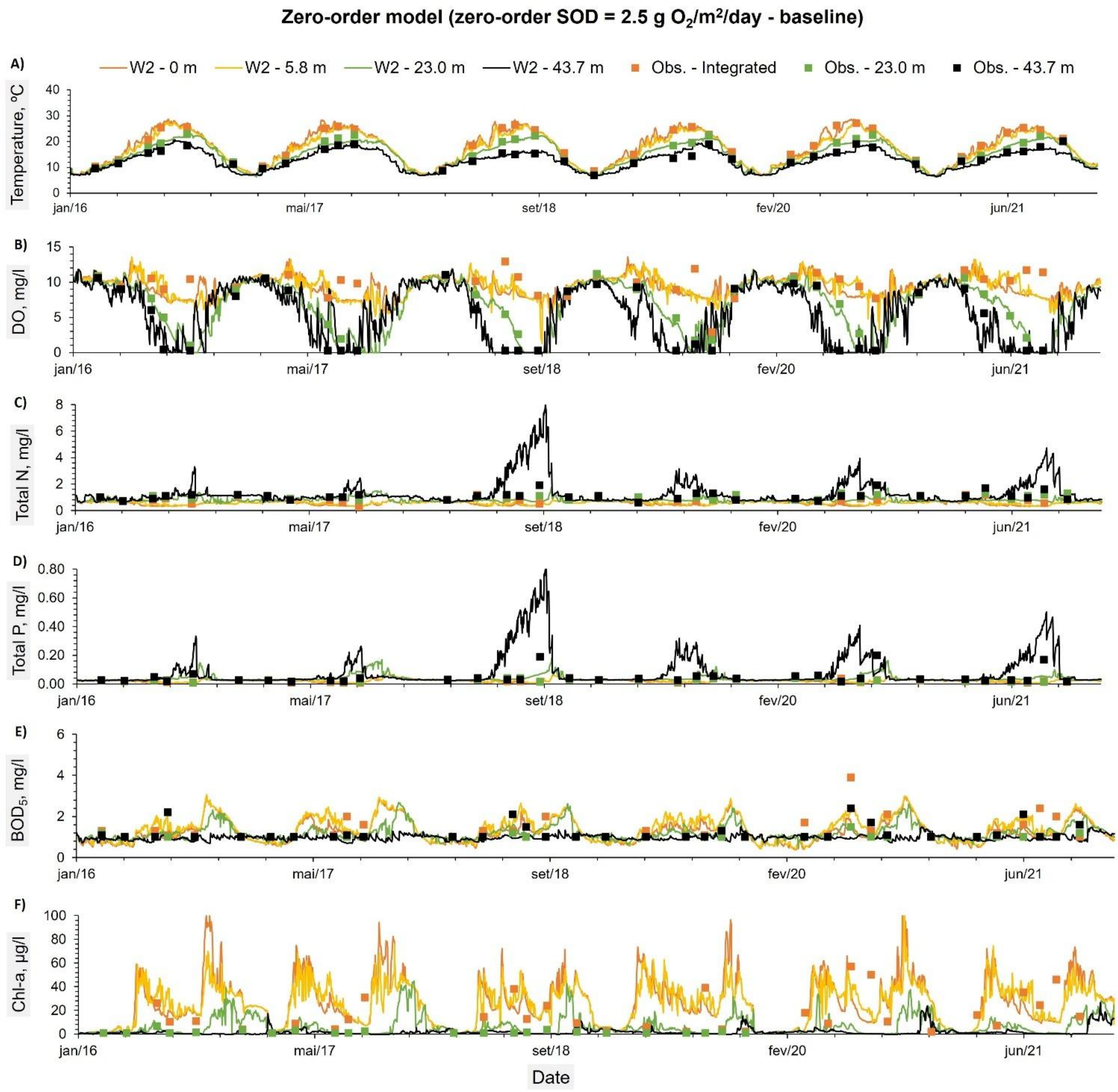
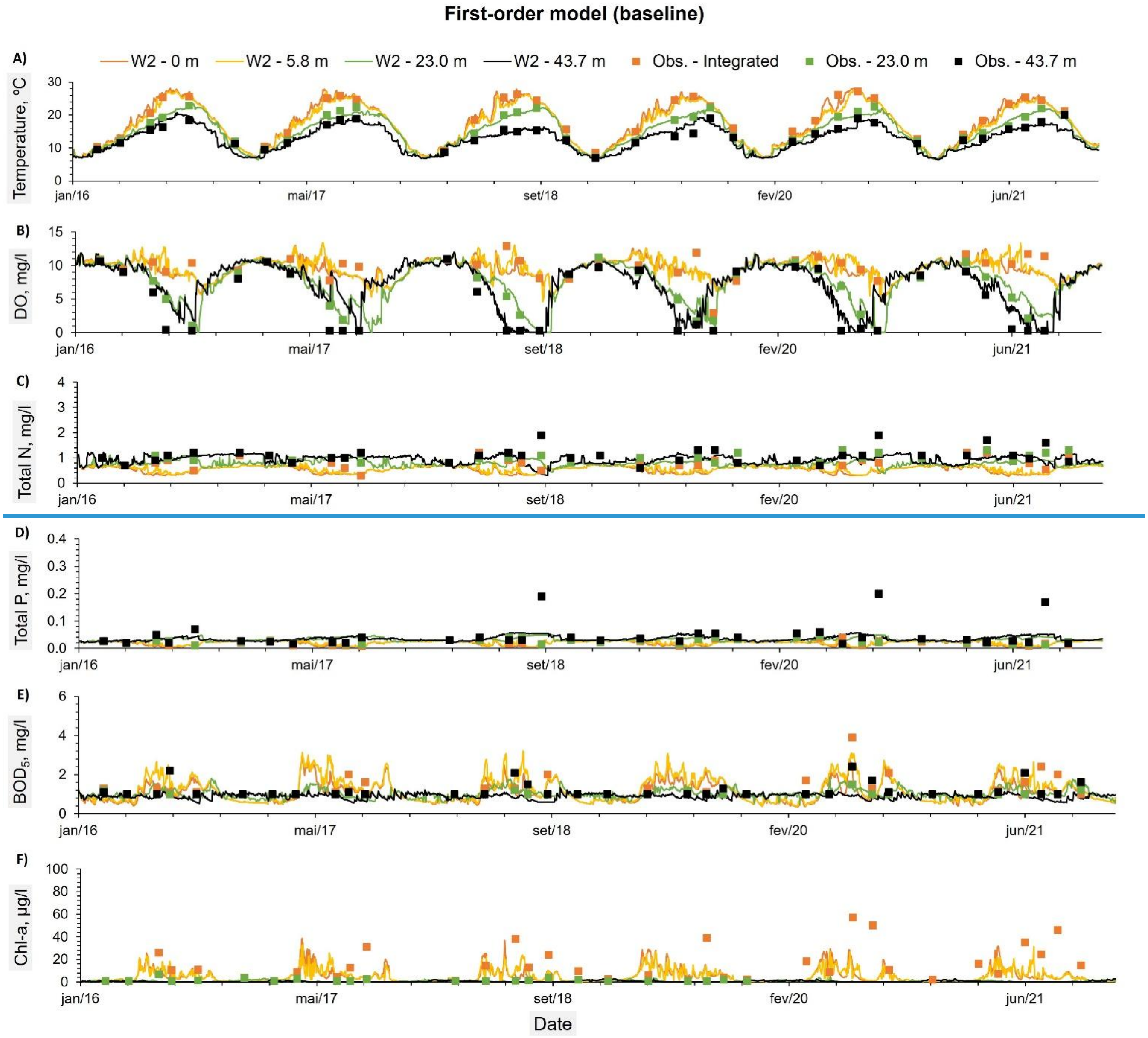


Figure 4: Constituents observed values at three different depths: (a) an integrated sample between the reservoir surface and an average depth of 5.8 meters, (b) an average depth of 23 meters, and (c) an average depth of 43.7 meters. These observed values were compared with the predicted time series from the Zero-order model (zero order SOD = 2.5 g O_2/m^2day - baseline) (A to F) for the same depths.


First-order model (baseline)
A)
W2 - 0 m
W2 - 5.8 m
W2 - 23.0 m
W2 - 43.7 m
Obs. - Integrated
Obs. - 23.0 m
Obs. - 43.7 m
Temperature, °C
0
10
20
30
jan/16
mai/17
set/18
fev/20
jun/21
B)
DO, mg/l
0
5
10
15
jan/16
mai/17
set/18
fev/20
jun/21
C)
Total N, mg/l
0
1
2
3
4
jan/16
mai/17
set/18
fev/20
jun/21
D)
Total P, mg/l
0.0
0.1
0.2
0.3
0.4
jan/16
mai/17
set/18
fev/20
jun/21
E)
BOD5, mg/l
0
2
4
6
jan/16
mai/17
set/18
fev/20
jun/21
F)
Chl-a, µg/l
0
20
40
60
80
100
jan/16
mai/17
set/18
fev/20
jun/21
Date

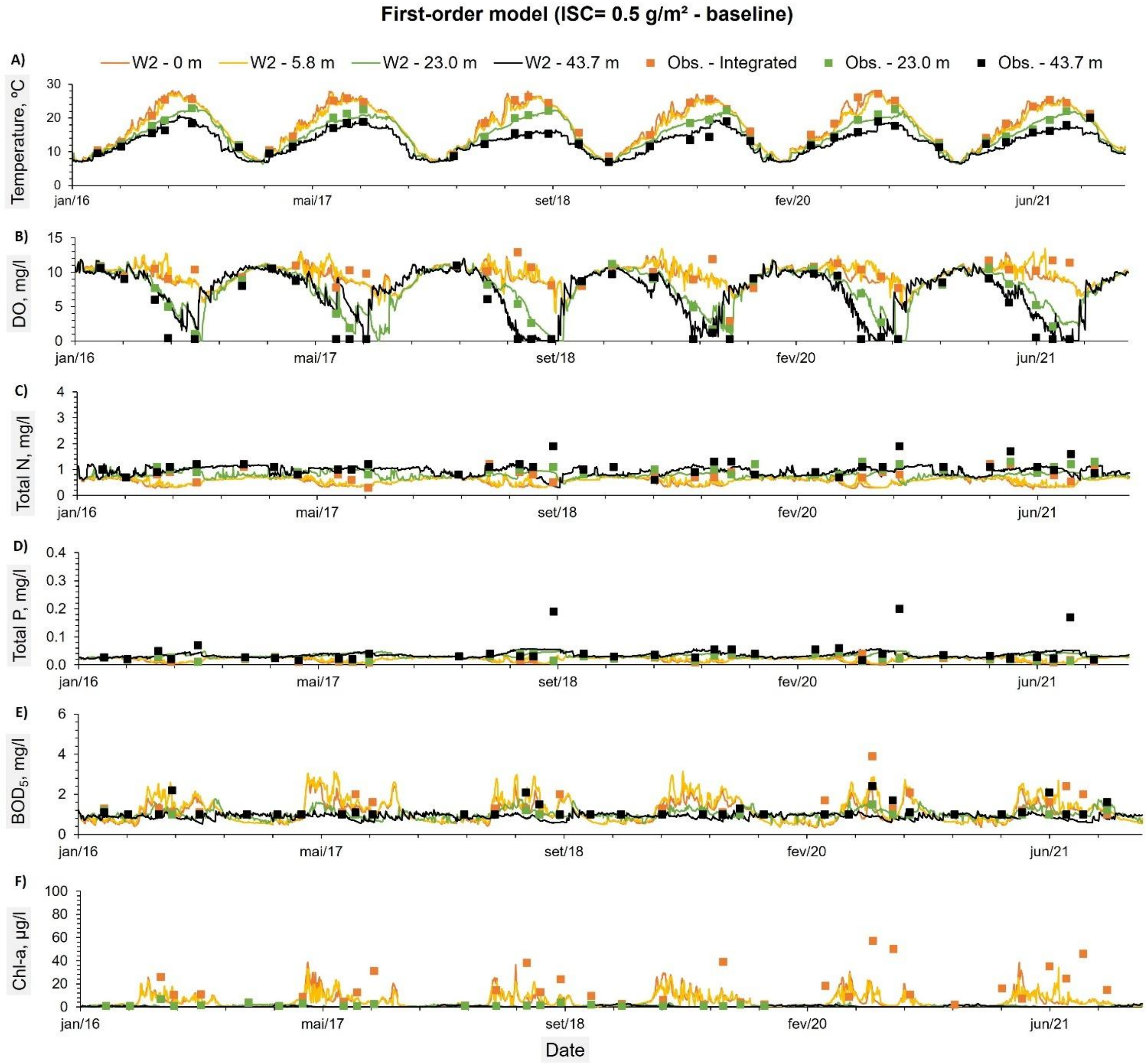
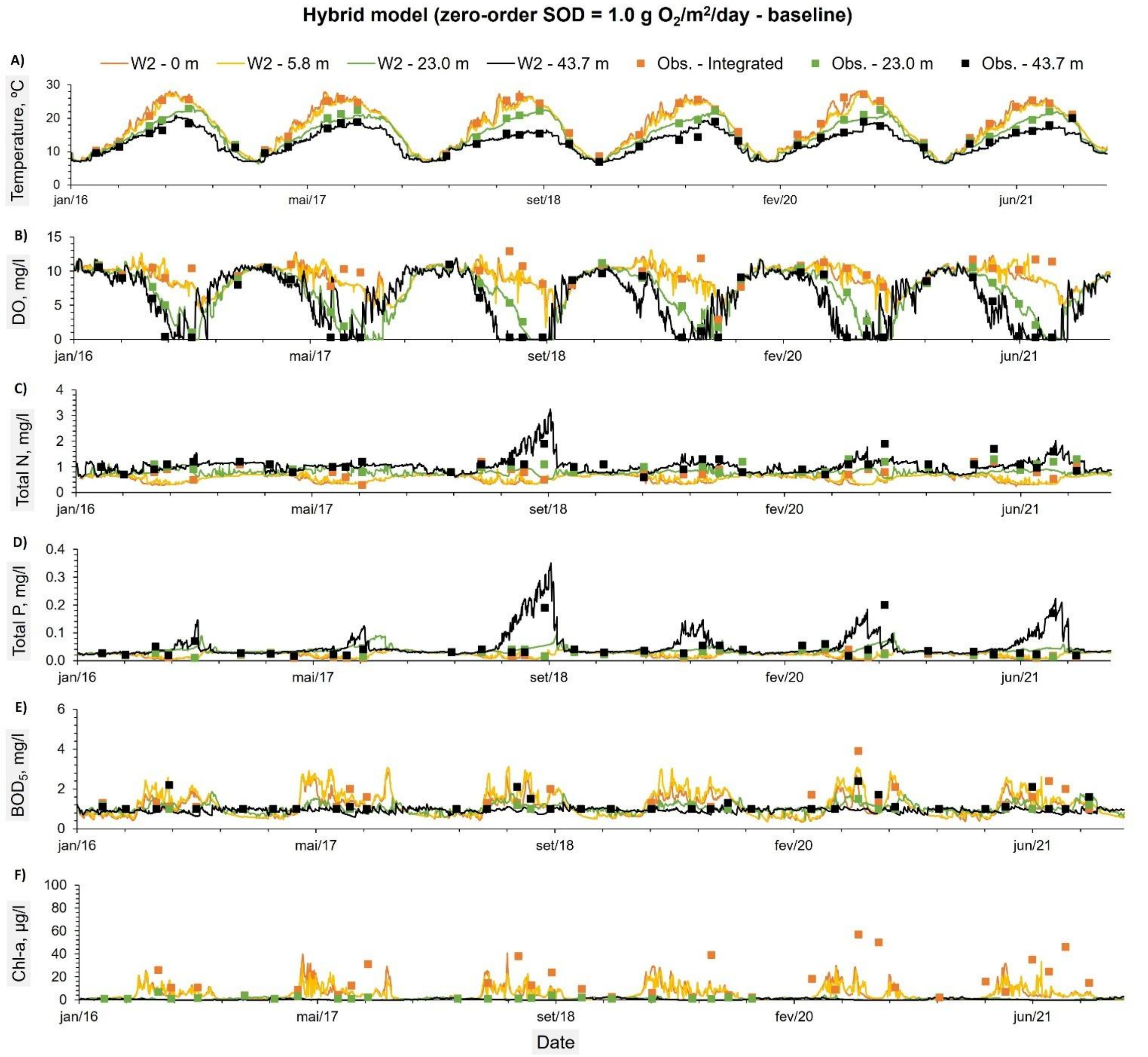


Figure 5: Constituents observed values at three different depths: (a) an integrated sample between the reservoir surface and an average depth of 5.8 meters, (b) an average depth of 23 meters, and (c) an average depth of 43.7 meters. These observed values were compared with the predicted time series from the First-order model (ISC=0.5 g/m^2 - baseline) (A to F) for the same depths. 475


Hybrid model (zero-order SOD = 1.0 g O_2/m^2/day - baseline)

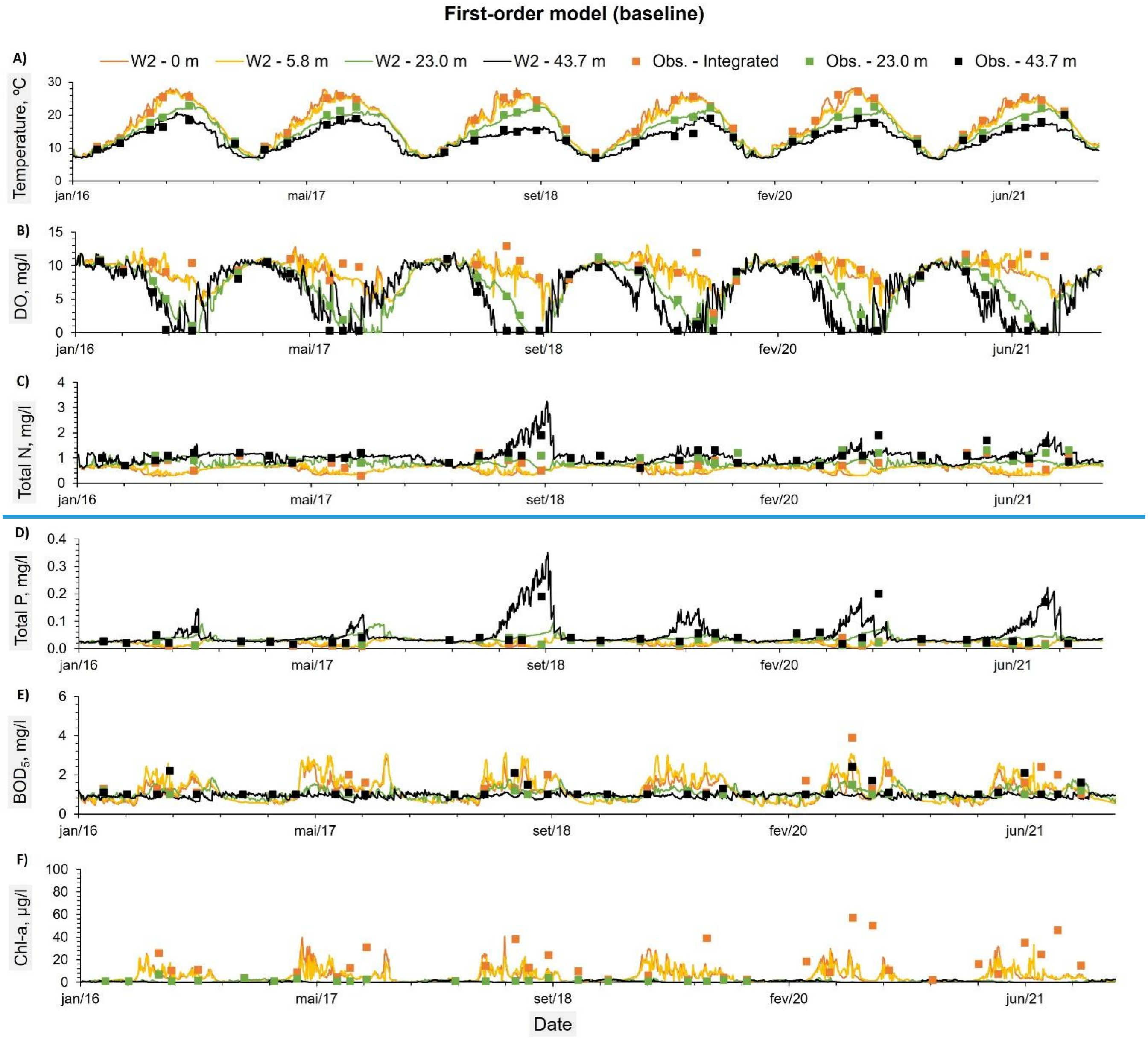


Figure 6: Constituents observed values at three different depths: (a) an integrated sample between the reservoir surface and an average depth of 5.8 meters, (b) an average depth of 23 meters, and (c) an average depth of 43.7 meters. These observed values were compared with the predicted time series from the Hybrid model (zero order SOD= 1.0 g O_2/m^{2}day - baseline) (A to F) for the same depths

3.3 Sensitivity analysis

The SOD values strongly influence the water column DO; therefore, this parameter was considered to support this analysis. Figure ~~4~~ 7 shows the SOD values from the reservoir bottom layer, predicted by the ~~W2-~~SD model for Runs 1 to 6, compared with the RMSE (~~Fig4A~~Fig7A) and the NSE (~~Fig4A~~Fig7B) values obtained between the predicted water column DO profiles and the mean initial POC values (across all sites values) for each run. These results suggest that Run ~~2~~ 4 was the best modeling solution. Considering the results obtained for Run 5 (baseline), Run 4 reduced the RMSE from 2.015 mg/L (Run 5) to 2.011 mg/L (Run 4) and increased the NSE from 0.714 (Run 5) to 0.716 (Run 4). The average SOD value in the bottom layer of the reservoir (across all model segments) decreased from 1.162 g O_2/m^{2}day (Run 5) to 1.071 g O_2/m^{2}day (Run 4). Although the reduction is modest and had only a minor effect on the DO profile predictions (Fig. 9), it suggests that the initial POC values used in Run 5 were likely overestimated. This outcome aligns with the assumption made in Run 5, where all observed TOC was considered to exist entirely as POC. In contrast, Run 4 was characterized using a lower average sediment concentration. Specifically, the mean value used in Run 4 (14170 mg/L) represents approximately 80% of the TOC value used in Run 5 (17712 mg/L), which was derived from observed TOC measurements (see Table 3). This comparison suggests that a more realistic estimate is that about 80% of the total organic carbon exists in particulate form, with the remainder composed of dissolved organic carbon. Run 4 and Run 5 show negligible differences in the predicted water temperature and DO profiles (Fig. 8 and 9). Table A10 presents the performance metrics for water temperature, DO, TN, TP, BOD$_5$, and Chl-a obtained for Run 4. While this run improved the DO simulation in the reservoir, results for the other constituents remained very similar to those of Run 5 (baseline). Overall, the water temperature profiles are very well captured by all models (Fig. 8), reflecting their robustness in simulating thermal dynamics. In contrast, DO profiles are more complex and challenging to model due to their sensitivity to multiple interacting processes. Nevertheless, the models were able to capture the main seasonal and vertical trends in DO concentrations, including stratification patterns and general oxygen depletion in bottom layers during warmer months (Fig.9). ~~Considering the results obtained for Run 5 (baseline), Run 2 reduced the RMSE from 1.79 mg/L to 1.73 mg/L and increased the NSE from 0.34 to 0.41. The average SOD value in the bottom layer of the reservoir (across all model segments) decreased from 1.2 g O_2/m$_2$/day (Run 5) to 0.81 g O_2/m^2/day (Run 2). Although not a significant reduction, it suggests an initial overestimation of POC values. This outcome was expected, given the initial assumption in Run 5 that all TOC existed in the form of POC. Based on the initial POC values of both runs, it can be concluded that the particulate fraction of organic carbon constitutes 40% of the TOC. Table A4 presents the metrics for water temperature, DO, Total N, Total P, BOD$_5$, and Chlorophyll-a obtained for Run 2. Although this run improved the reservoir DO prediction, the results for the other constituents are very similar to those obtained for Run 5 (baseline).~~

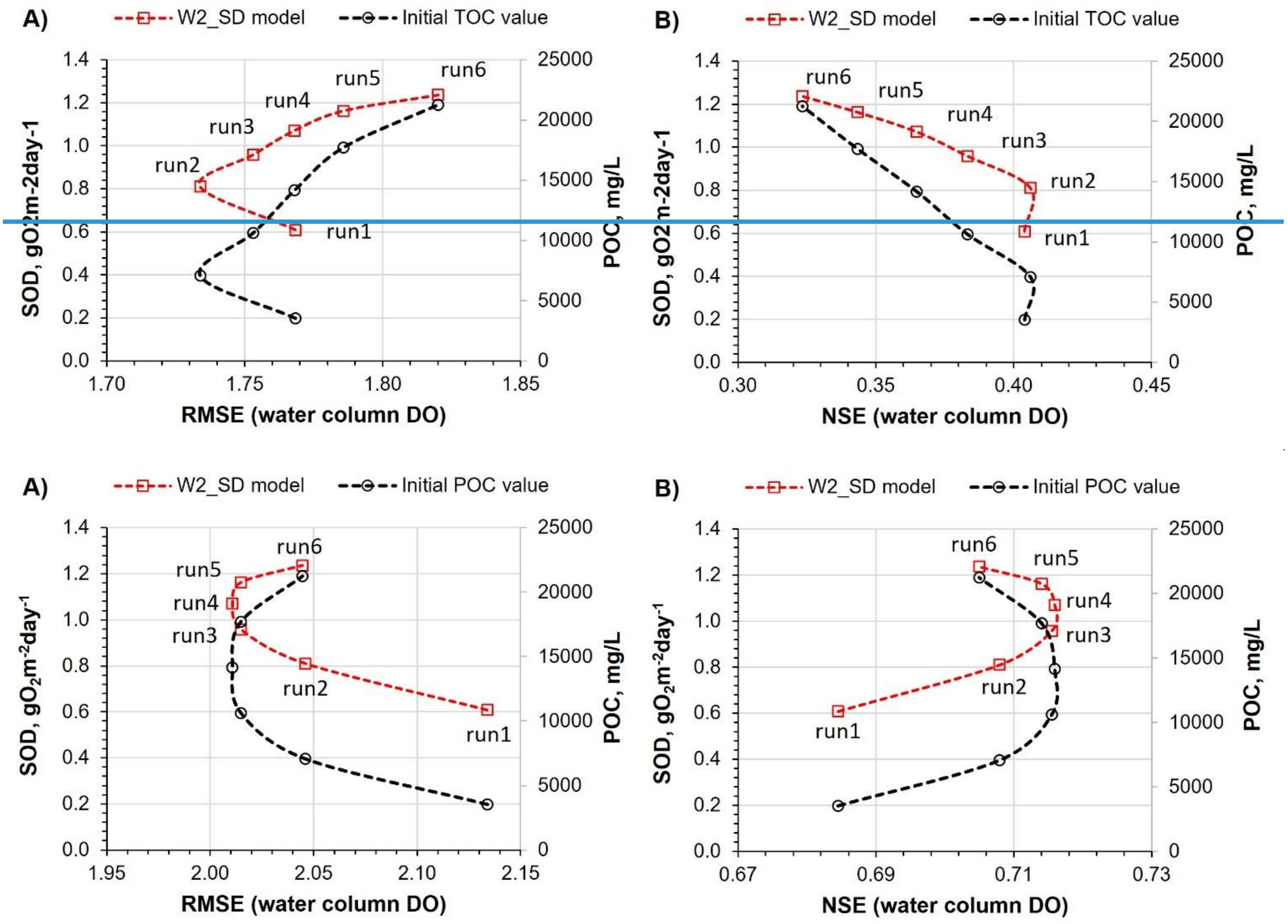
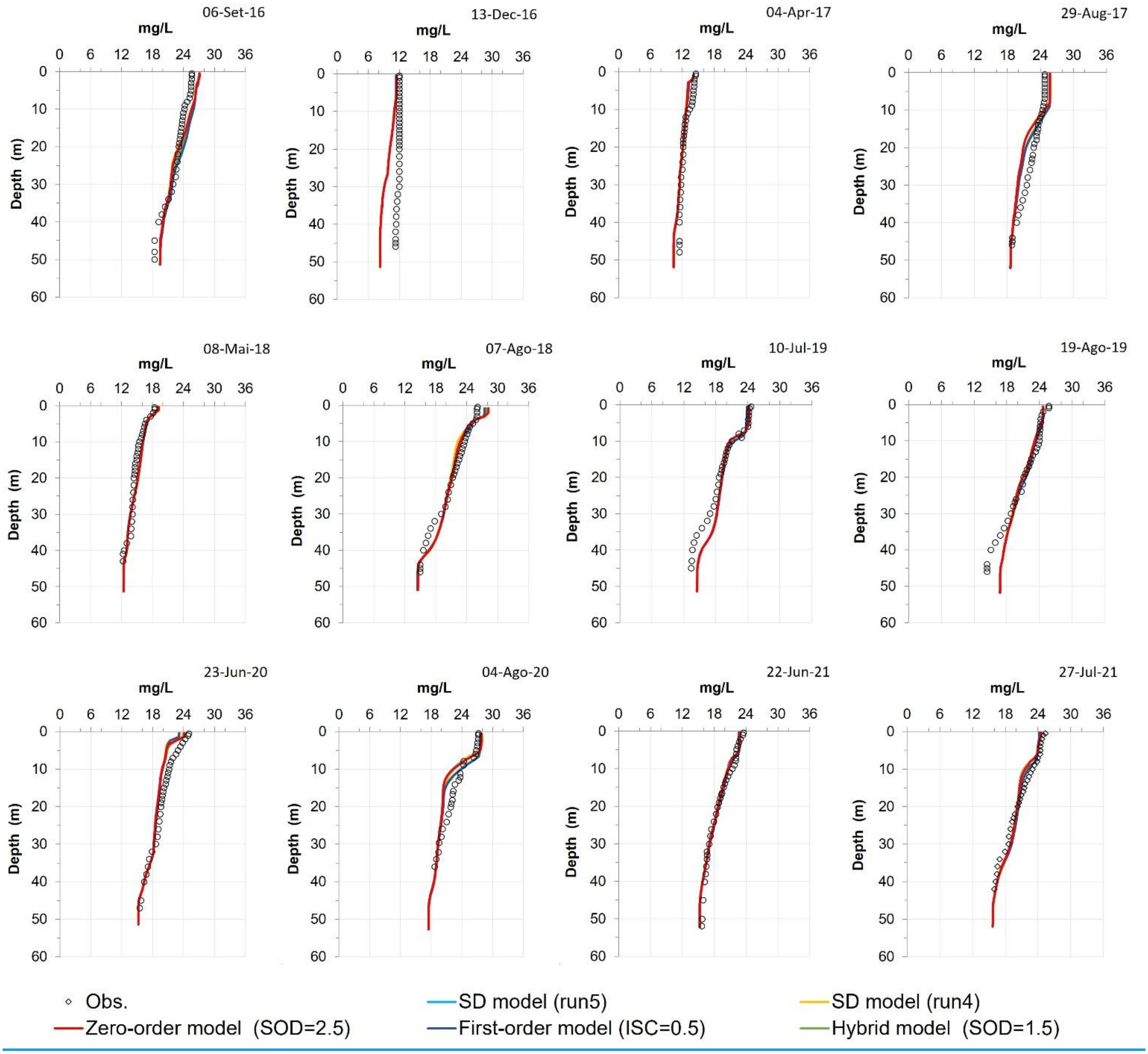


Figure ~~4~~7: A) SOD values from the reservoir bottom layer, predicted by the ~~W2_~~SD_model for Runs 1 to 6, compared with the RMSE obtained between the predicted water column DO profiles and the mean initial POC values (across all sites values) for each run of the ~~W2_~~SD_model. B) Similar to A but considering the NSE metric.


06-Set-16
13-Dec-16
04-Apr-17
29-Aug-17
08-Mai-18
07-Ago-18
10-Jul-19
19-Ago-19
23-Jun-20
04-Ago-20
22-Jun-21
27-Jul-21
mg/L
0 6 12 18 24 30 36
Depth (m)
0 10 20 30 40 50 60
Obs.
SD model (run5)
SD model (run4)
Zero-order model (SOD=2.5)
First-order model (ISC=0.5)
Hybrid model (SOD=1.5)

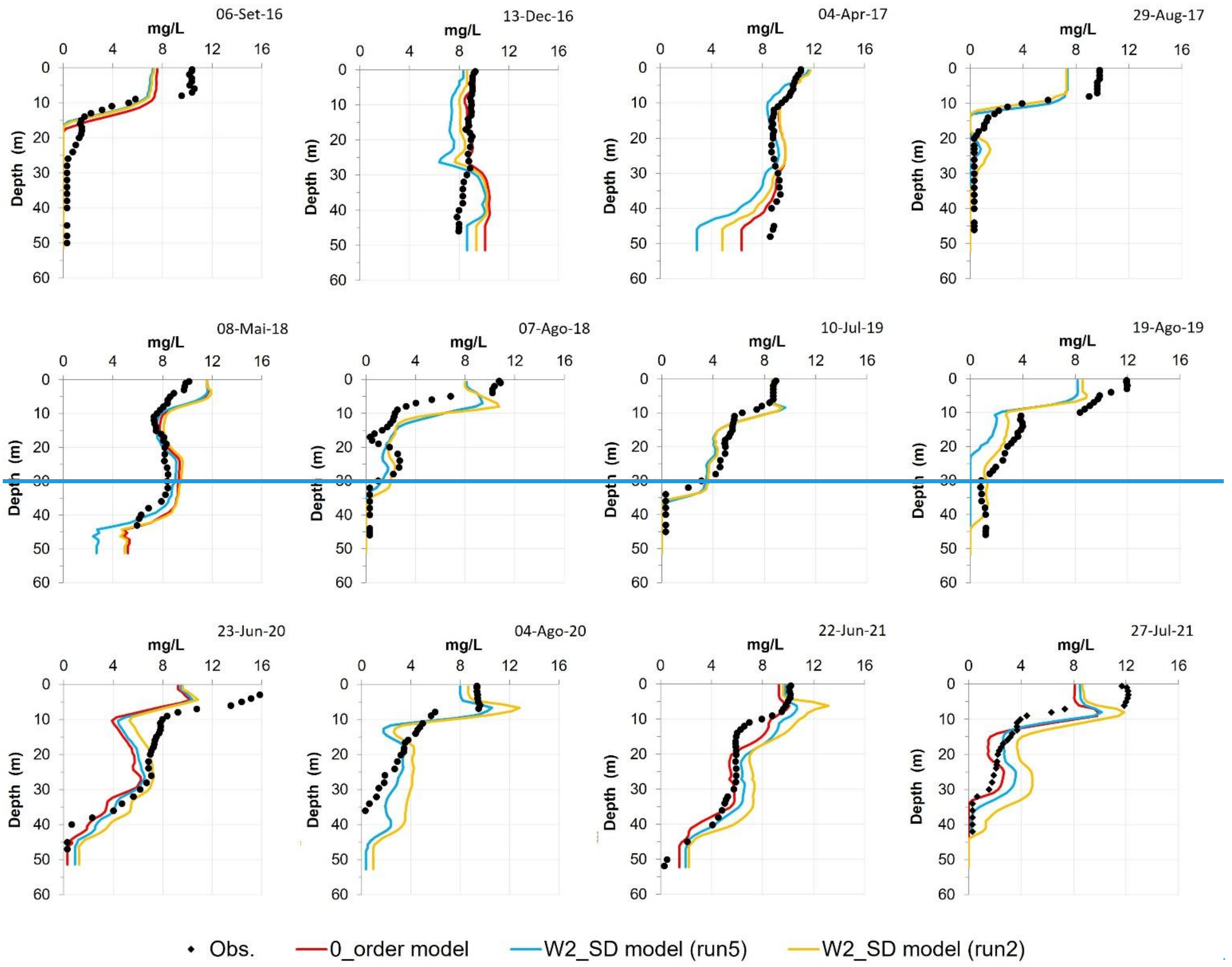


Figure ~~5~~8: Observed ~~DO~~ water temperature profiles (300 m from the dam) compared to predicted profiles using the SD model (Run 4) and (Run 5 - baseline), ~~W2~~ Zero-order model (zero-order SOD = 2.5 g O_2/m^2day - baseline); First-order model (ISC= 0.5 g/m² - baseline) and the Hybrid model (zero order SOD= 1.0 g O_2/m^2day - baseline).~~zero-order model (baseline), W2_SD model (Run 2) and (Run 5; baseline).~~

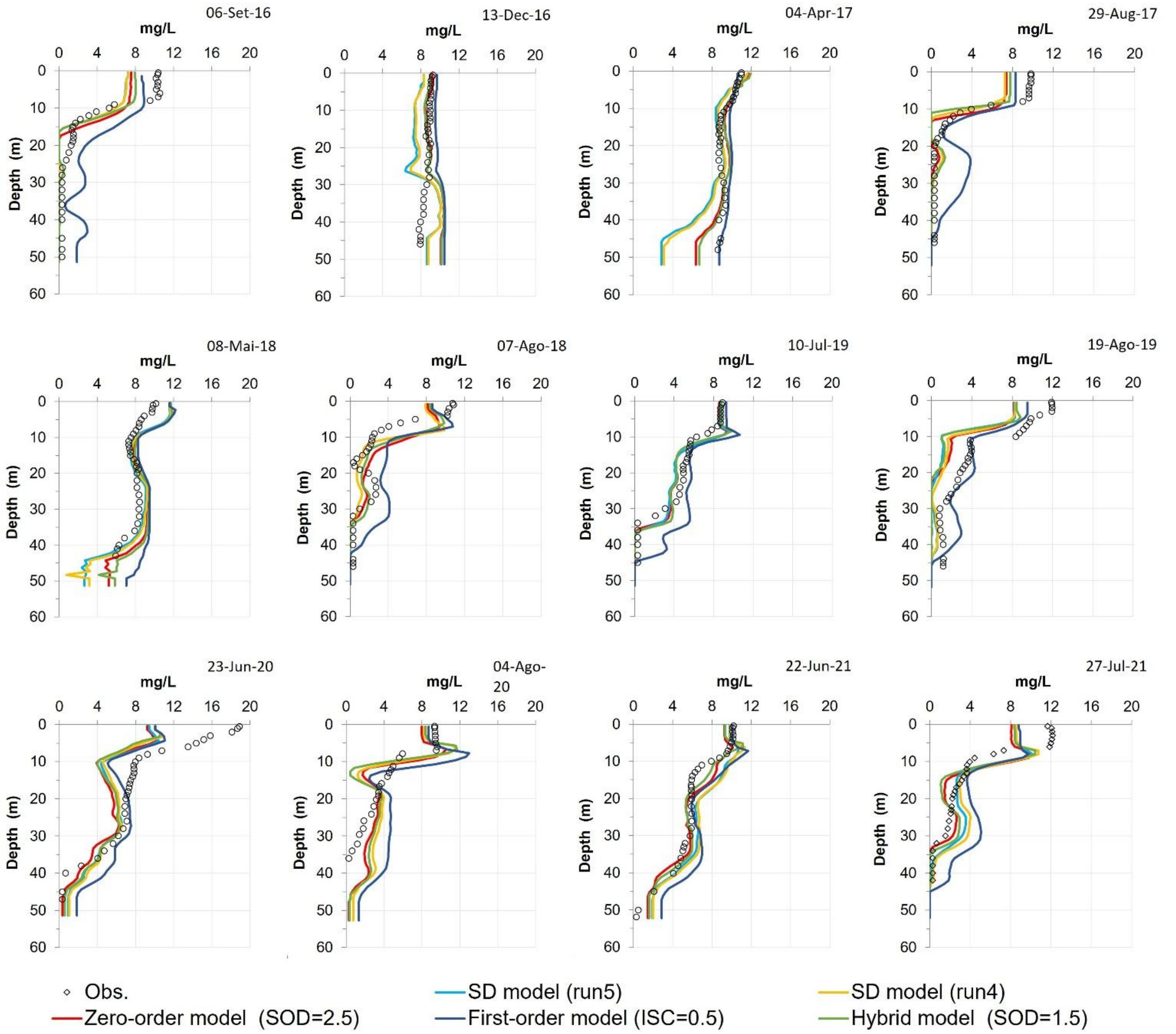
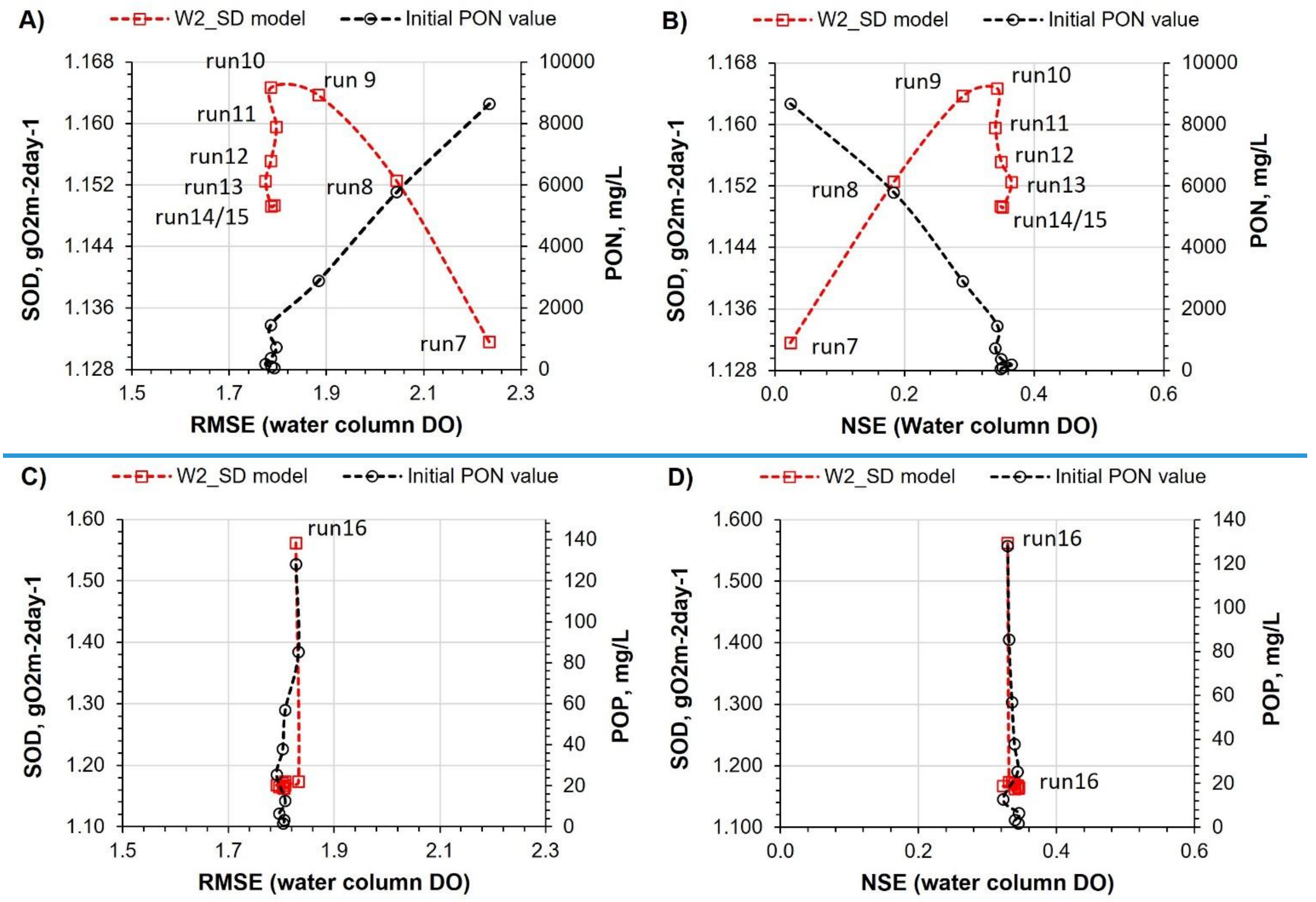


Figure 9: Observed DO profiles (300 m from the dam) compared to predicted profiles using the SD model (Run 4) and (Run 5 - baseline), Zero-order model (zero-order SOD = 2.5 g O_2/m^2day - baseline); First-order model (ISC= 0.5 g/m² - baseline) and the Hybrid model (zero order SOD= 1.0 g O_2/m^2day - baseline).

The sensitivity analysis also involved varying the initial values of PON and POP for each run. The results indicate that mean reservoir SOD values remained nearly constant, as depicted in Fig. ~~6~~10, suggesting that the ~~W2~~ SD model was not significantly affected by variations in the initial PON and POP values in the sediments. However, in Runs 7, 8, and 9, where

PON values were higher, there was a significant increase in the release of $N\text{-}NH_4$ and $N\text{-}NO_X$ from the reservoir sediments, leading to an impact on water column DO. This is evidenced by the notable increase in RMSE and the reduction of NSE values, as shown in Fig. 10~~6~~A and 10~~6~~B.


A)
W2_SD model
Initial PON value
SOD, gO2m-2day-1
PON, mg/L
RMSE (water column DO)
1.168
1.160
1.152
1.144
1.136
1.128
10000
8000
6000
4000
2000
0
1.5
1.7
1.9
2.1
2.3
run10
run 9
run11
run12
run13
run8
run14/15
run7
B)
W2_SD model
Initial PON value
SOD, gO2m-2day-1
PON, mg/L
NSE (Water column DO)
1.168
1.160
1.152
1.144
1.136
1.128
10000
8000
6000
4000
2000
0
0.0
0.2
0.4
0.6
run9
run10
run11
run12
run8
run13
run14/15
run7
C)
W2_SD model
Initial PON value
SOD, gO2m-2day-1
POP, mg/L
RMSE (water column DO)
1.60
1.50
1.40
1.30
1.20
1.10
140
120
100
80
60
40
20
0
1.5
1.7
1.9
2.1
2.3
run16
D)
W2_SD model
Initial PON value
SOD, gO2m-2day-1
POP, mg/L
NSE (water column DO)
1.600
1.500
1.400
1.300
1.200
1.100
140
120
100
80
60
40
20
0
0.0
0.2
0.4
0.6
run16
run16

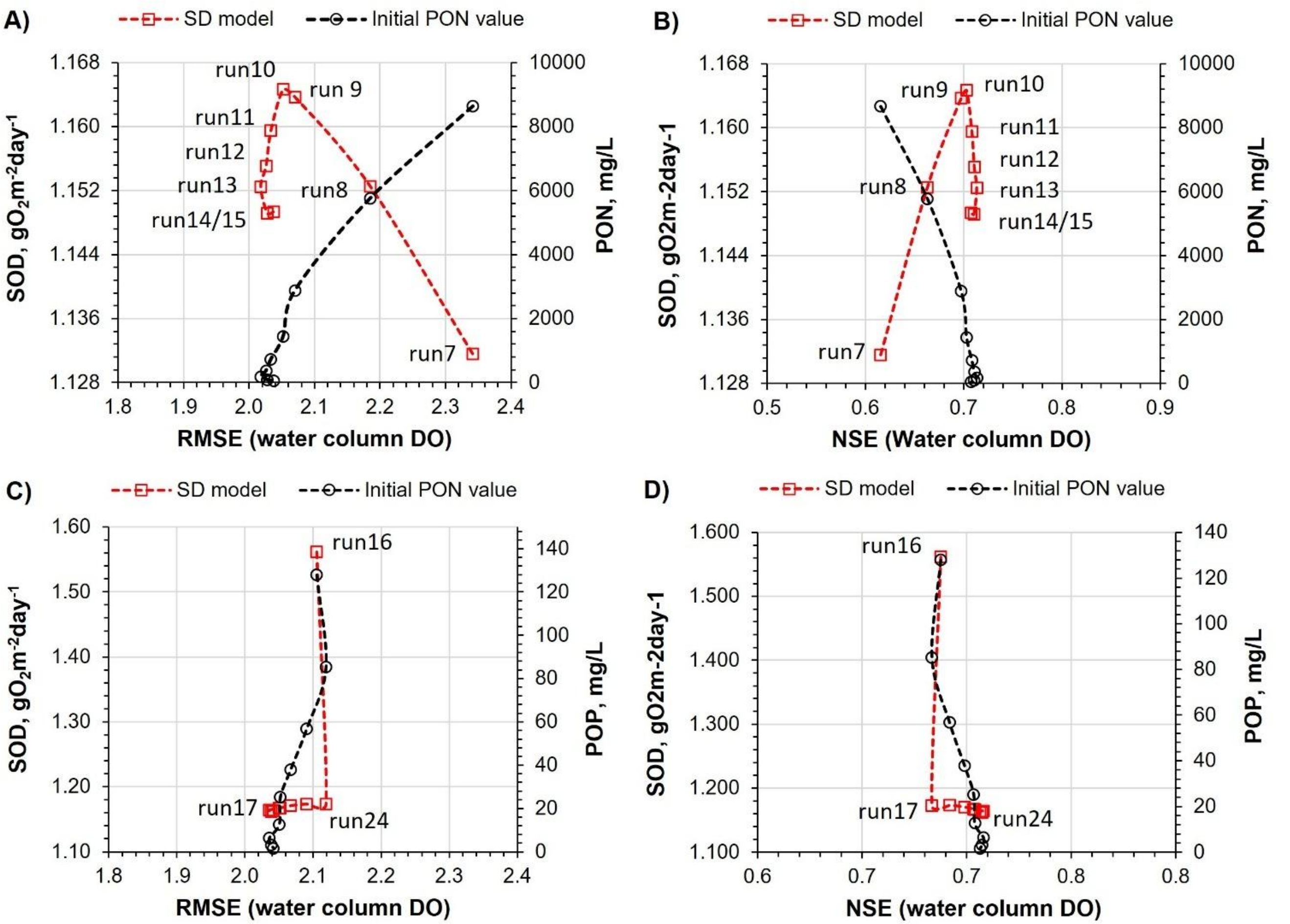
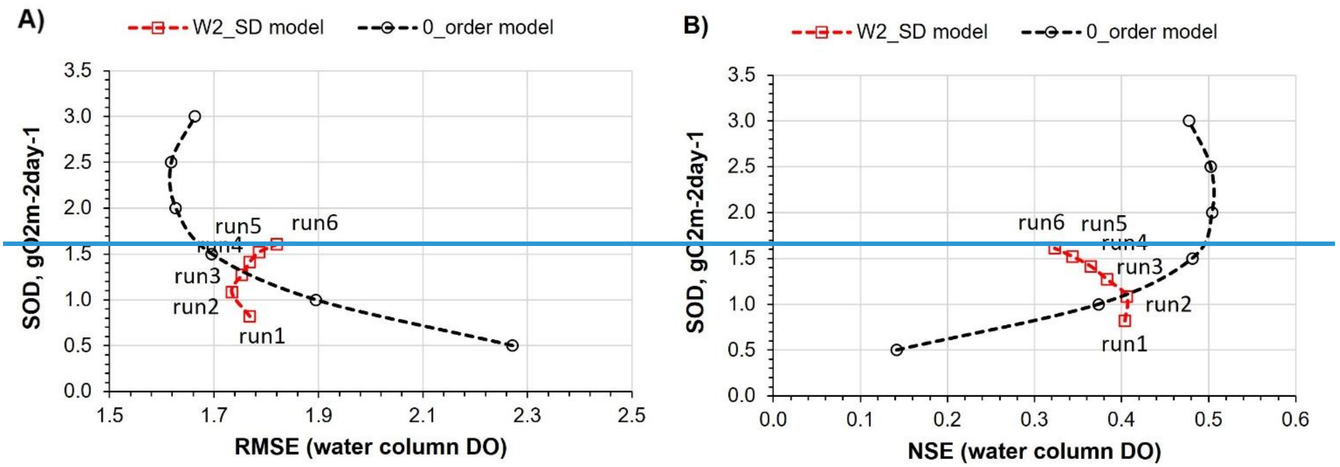


Figure 10~~6~~: A) SOD values from the reservoir bottom layer, predicted by the ~~W2~~ SD model for Runs 7 to 15, compared with the RMSE obtained between the predicted water column DO profiles and the mean initial PON values (across all sites) for each run B) Similar to A) but considering the NSE metric. C) SOD values from the reservoir bottom layer, predicted by the ~~W2~~ SD model for Runs 16 to 24, compared with the RMSE obtained between the predicted water column DO profiles and the mean initial POP values (across all sites) for each run. D) Similar to C) but considering the NSE metric.

Figure ~~7~~ 11 shows the RMSE (Fig. ~~7A~~11A) and the NSE (Fig. ~~7B~~11B) values between observed and predicted water column DO profiles for ~~both~~ all models: ~~W2~~ SD model (Runs 1 to 6), ~~and the~~ Z~~W2 z~~ero-order model and Hybrid model, each with six different SOD values ranging from 0.5 to 3.0 g/m^2~~/~~day, along with the corresponding reservoir SOD values. Additionally, this figure illustrates how the First-order model varies with the initial sediment concentration. Among the four models evaluated, the Hybrid model demonstrated the best overall performance in predicting DO concentrations in the reservoir. With an average SOD of 1.49 g O_2/m^{2}day, the hybrid model achieved the lowest RMSE (1.87 mg/L) and highest NSE (0.76), demonstrating superior predictive accuracy. The Zero-order model followed closely, reaching optimal performance at an average zero-order SOD of 1.43 g O_2/m²day, with an RMSE of 1.965 mg/L and an NSE of 0.732. The SD model also performed well, attaining its best accuracy at an average SOD of 1.07 g O_2/m²day, where the RMSE decreased to 2.011 mg/L and the NSE peaked at 0.716; however, further improvements plateaued beyond this point. In contrast, the First-order model consistently exhibited higher RMSE values (ranging from 2.15 mg/L to 2.22 mg/L) and lower NSE values (between 0.66 and 0.68), regardless of the initial sediment concentration. Moreover, its SOD at the bottom layer remained relatively stable, indicating limited sensitivity to input variations. Overall, these results underscore the hybrid model's robustness and accuracy, followed by the Zero-order and SD models, while the First-order model demonstrated the weakest performance in this context.~~Overall, these results show that the zero-order model with a constant SOD of 2.5 g O_2/m²/day was better than the W2_SD model at predicting reservoir DO. The analysis of this figure also indicates that varying the SOD in the zero-order model from 0.5 to 3.0 g O_2/m²/day resulted in changes in DO prediction, with RMSE values ranging from 2.27 to 1.62 g O_2/m²/day and NSE values ranging from 0.14 to 0.50. The reservoir SOD obtained with the W2_SD model was less sensitive to changes in the initial sediment POC value, suggesting that the initial POC value must be significantly higher than 20 000 mg/L to achieve SOD values of more than 2.0 g O_2/m²/day.~~



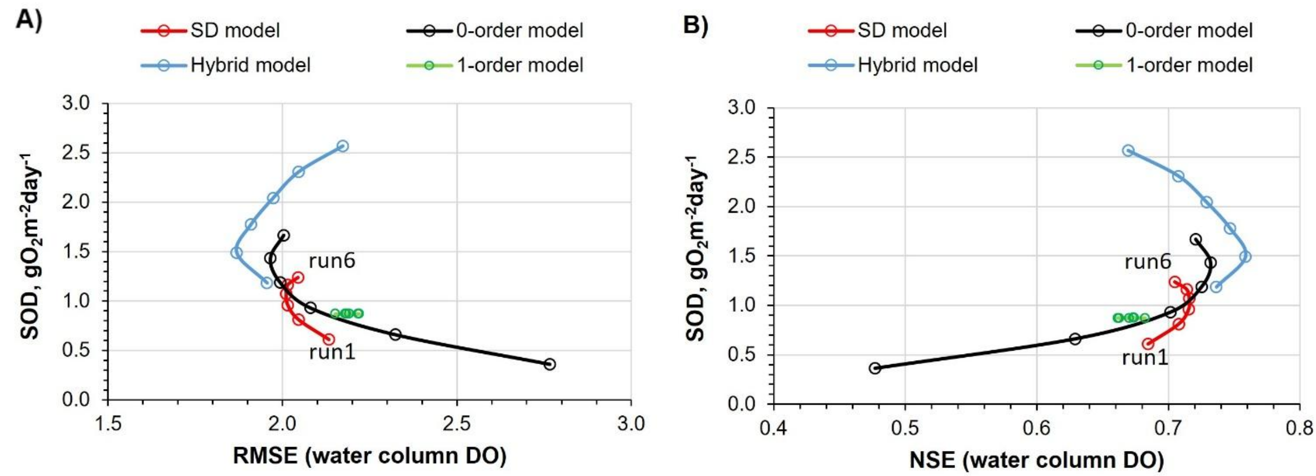


Figure 11~~7~~. **(A) RMSE between observed and simulated DO profiles in the water column for all models: the SD model (Runs 1–6), the Zero-order model, the Hybrid model with six SOD values ranging from 0.5 to 3.0 g O_2/m^{2}day, and the First-order model with initial sediment organic matter concentrations from 0.0 to 3.0 g m^{-2}. (B) Same as (A), but using the Nash–Sutcliffe Efficiency (NSE) as the performance metric.~~A) RMSE values between observed and predicted water column DO profiles predicted by both models (W2_SD model Runs 1 to 6 and 0_order model with six different SOD values ranging from 0.5 to 3.0 g/m^2/day) compared with the reservoir SOD values. B) Similar to A) but considering the NSE metric.~~**

3.4 Inflow Organic Matter and Phosphorus Load Reduction Scenarios

The results reveal clear differences in model sensitivity to inflow load reductions, with the First-order and Hybrid models exhibiting a stronger response compared to the SD and Zero-order models (Figures 12 and 13). The SD model showed minimal change, indicating limited sensitivity to external loading (Figures 12a and 13a), likely due to strong internal loading feedback from legacy phosphorus and organic matter stored in sediments. The Zero-order model demonstrated limited utility for management scenarios because it is decoupled from the water column, reducing its responsiveness to external changes. The First-order model may overestimate sensitivity as it tends to underestimate internal loading contributions. The Hybrid model, which combines both approaches, is less reactive than the First-order model due to the influence of the Zero-order component, offering a more balanced response. However, the Zero-order SOD component in the Hybrid model depends solely on temperature and remains decoupled from water column conditions; this limitation may gradually reduce the model's accuracy in long-term simulations. These differences in model sensitivity are further reflected in the evolution of average SOD across scenarios (Table 5). While the Zero-order and SD models show virtually no change in bottom-layer SOD under reduced loading conditions, the First-order and Hybrid models register clear declines. The First-order model's SOD drops from 0.87 g O_2/m^2day to 0.42 g O_2/m^2day (80% OM reduction) and 0.29 g O_2/m^2day (80% OM and P reduction) and the Hybrid model from 1.49 g O_2/m^2day to 1.07 g O_2/m^2.day (80% OM reduction) and 0.94 g O_2/m^2day (80% OM and P reduction).

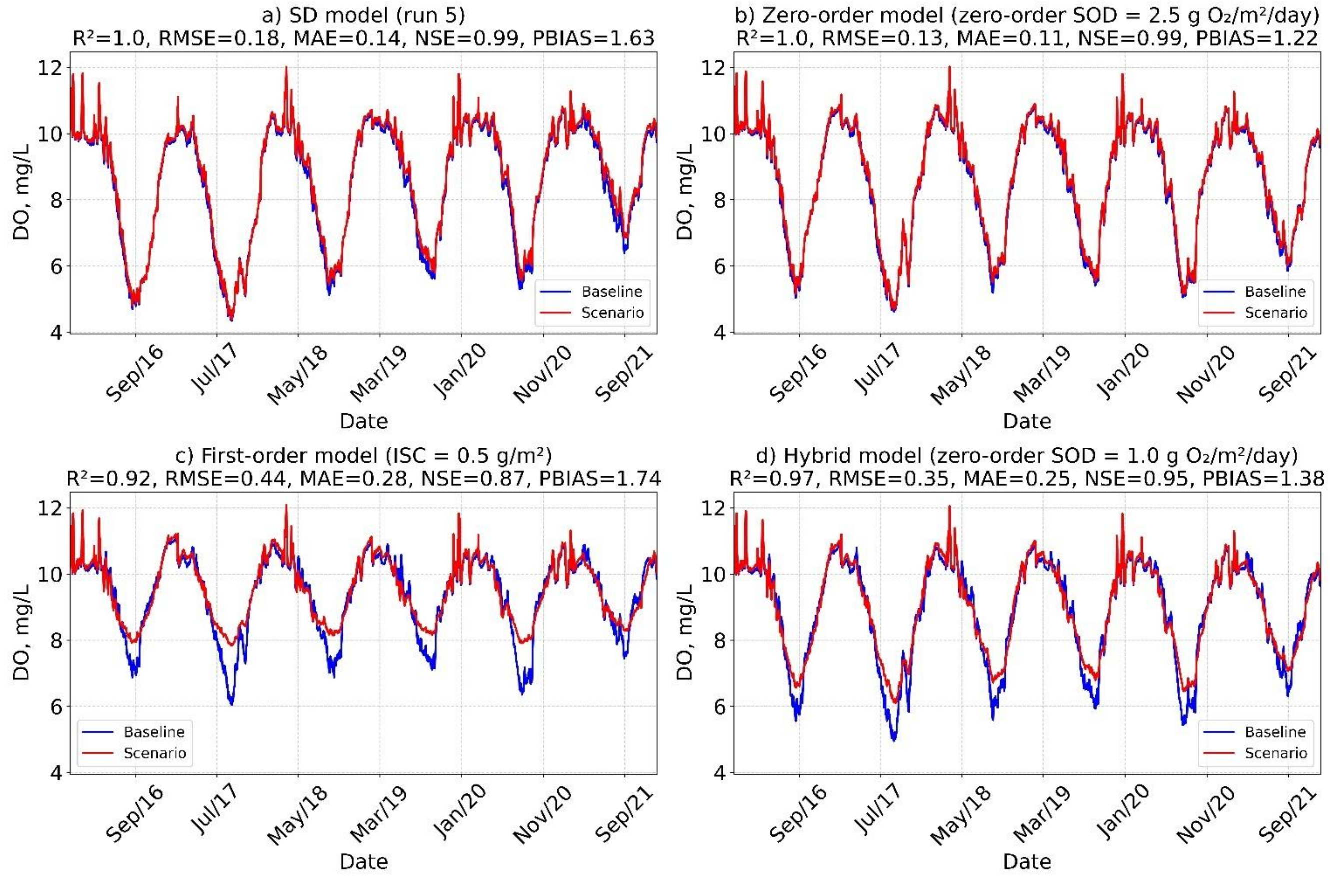


Figure 12. Time series of DO, averaged across all model layers and segments, for each baseline model scenario: SD model (Run 5), Zero-order model (SOD = 2.5 g O_2/m^2day), First-order model (initial sediment concentration = 0.5 g/m²), and Hybrid model (Zero-order SOD = 1.0 g O_2/m^2day). The figure compares baseline conditions with an 80% reduction in organic matter inflow load in the main reservoir branch (Branch 1 – Tâmega River). Performance metrics (R², RMSE, MAE, NSE, and PBIAS) are also shown for each case. 610

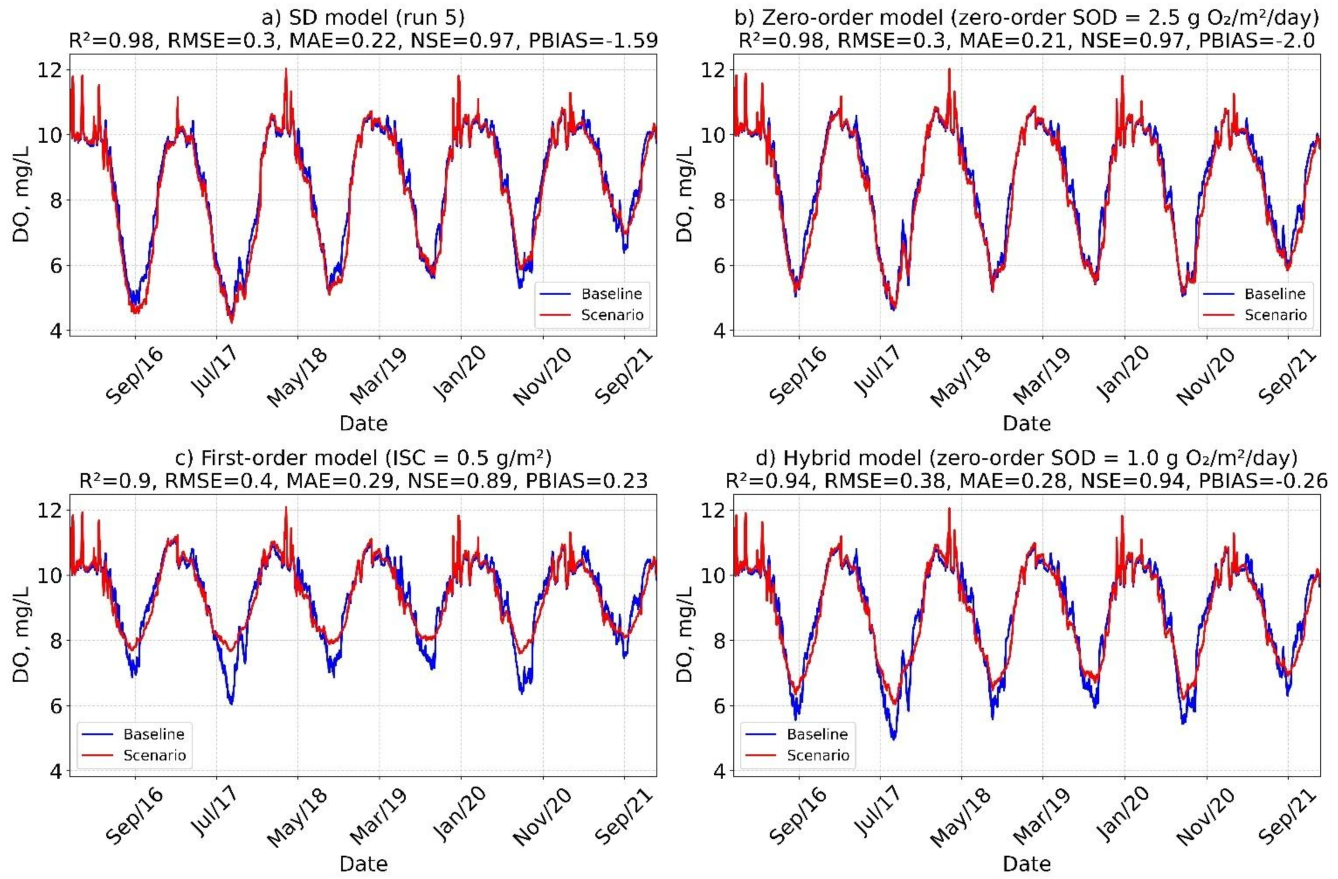


Figure 13. Time series of DO, averaged across all model layers and segments, for each baseline model scenario: SD model (Run 5), Zero-order model (SOD = 2.5 g O_2/m^2day), First-order model (initial sediment concentration = 0.5 g/m²), and Hybrid model (Zero-order SOD = 1.0 g O_2/m^2day). The figure compares baseline conditions with an 80% reduction in organic matter and P-PO_4 inflow loads in the main reservoir branch (Branch 1 – Tâmega River). Performance metrics (R^2, RMSE, MAE, NSE, and PBIAS) are also shown for each case.

Table 5. Average sediment oxygen demand (SOD) in the bottom layers of the reservoir, calculated across all segments, for each model under three scenarios: Reference (baseline conditions), 80% reduction in organic matter inflow (OM 80%), and combined 80% reduction in organic matter and phosphorus inflow (OM and P%) in the in the main reservoir branch (Branch 1 – Tâmega River)

Scenario	**SD model (Run 5)**	**Zero-order model (SOD = 2.5 g O_2/m^2day)**	**First-order model (ISC = 0.5 g/m²)**	**Hybrid model (Zero-order SOD = 1.0 g O_2/m.day)**		
				Aggregate	**Zero-order term**	**First-order term**
Baseline	1.16±0.82	1.43±2.12	0.87±1.19	1.49±2.02	0.59± 0.85	0.90± 0.75
OM 80% reduction	1.13±0.83	1.44±2.12	0.42±0.61	1.07±1.43	0.61± 0.85	0.46± 0.75
OM and P% 80% reduction	1.13±0.82	1.46±2.11	0.29±0.37	0.94±1.19	0.64± 0.84	0.30± 0.42

4 Discussion

Overall, the temperature and DO predictions for the reservoir boundary conditions (Tâmega river) were quite good: PBIAS: 0.76% and 0.92%, respectively. When a significant number of samples and forcing variables are available the accuracy of machine learning algorithms can be greatly enhanced. This was demonstrated in the studies by Lu et al. (2020), Rajesh and Rehana (2021), and Feigl et al. (2021), where the RMSE for river water temperature prediction reached 1.04ºC, 1.03ºC, and 0.58ºC, respectively. The results obtained for alkalinity, conductivity and TSS were also good: Alkalinity-PBIAS: 17.44%; Conductivity - PBIAS: 8.23%; TSS - ~~0.73;~~ PBIAS: 11.86%. However, as expected, the PBIAS values obtained for the remaining constituents were not as favorable (Total P- PBIAS: 7.11%; N-NO_X- PBIAS: 3.92%; BOD_5- PBIAS: 6.93%; Chla-PBIAS: 30%). The modeling of these constituents involves complex biological, chemical, and physical processes that are harder to model accurately. However, except for Chla, the PBIAS values were generally less than 10%, reflecting acceptable levels of bias. Ammonium (N-NH_4) was the only parameter for which performance was significantly lower, generating a PBIAS of 28.27%. Moriasi et al. (2015) suggest that ±10≤PBIAS≤ ± 25 is indicative of a satisfactory model performance.

Based on the RMSE, the overall reservoir calibration results obtained for all constituents with ~~both~~ all models for the 2016-2021 period were consistent with the results seen in other studies (see Table A11~~3~~). The mean RMSE values for Chl-a ~~Chlorophyll-a~~ obtained with ~~both~~ all models (~~W2~~ SD model (run 5 - baseline)~~SD model~~: ~~13.43~~17.72 µg/L; ~~W2 z~~ Zero-order model (zero-order SOD = 2.5 g O_2/m^2 day - baseline)~~ero-order model~~: ~~15.84~~17.78 µg/L; First-order model (ISC= 0.5 g/m² - baseline): 14.88 µg/L and the Hybrid model (zero-order SOD=1.0 g O_2/m^2 day - baseline): 14.88 µg/L) are aligned with the results of other modeling studies (Brito et al., 2018: 62.9 µg/L; Kim et al., 2019: 6.7 to 13.2 µg/L; Tasnim et al., 2021: 0.6 to 27.6 µg/L; Almeida et al., 2023: 19.36 to 25.57 µg/L). For TP, the mean RMSE values were 0.03 mg/L for both the SD model (Run 5 – baseline) and the First-order model (ISC = 0.5 g/m² – baseline), while the Hybrid model (zero-order SOD = 1.0 g O_2/m^2day – baseline) showed a slightly higher value of 0.04 mg/L. These results fall within the range reported in previous studies, including Brett et al. (2016) at 0.012 mg/L, Kim et al. (2019) between 0.014 and 0.068 mg/L, Tasnim et al. (2021) from 0.005 to 0.036 mg/L, and Almeida et al. (2023) ranging from 0.07 to 0.09 mg/L. The only exception was the Zero-order model (SOD = 2.5 g O_2/m^{2}day – baseline), which overestimated phosphorus export from sediments during the summer months (July to September) of 2018 to 2021, resulting in a notably higher RMSE of 0.1 mg/L.~~This can also be said of Total P, with the mean RMSE values obtained with both models (W2_SD model: 0.02 mg/L; W2_zero-order model: 0.06 mg/L) corresponding to those of other studies (Brett et al., 2016: 0.012 mg/L; Kim et al., 2019: 0.014 to 0.068 mg/L; Tasnim et al., 2021: 0.005 to 0.036 mg/L; Almeida et al., 2023: 0.07 to 0.09 mg/L).~~ Even with a very low phosphorus release rate from the sediments—representing a fraction of the SOD (0.001)—the Zero-order model still overestimated phosphorus concentrations, particularly during periods of elevated sediment oxygen demand. This suggests that the model may lack the sensitivity needed to accurately simulate low-level sediment-phosphorus interactions under such conditions. The mean RMSE values obtained for TN were lower than the only reference value available in the literature—0.77 mg/L reported by Deliman et al. (2002). Specifically, the SD model (Run 5 – baseline) yielded an RMSE of 0.33 mg/L, the First-order model (ISC = 0.5 g/m^2 – baseline) produced 0.36 mg/L, and the Hybrid model (zero-order SOD = 1.0 g O_2/m^{2}day – baseline) resulted in 0.35 mg/L. The only exception was the Zero-order model (SOD = 2.5 g O_2/m^{2}day – baseline), which had a significantly higher RMSE of 0.79 mg/L—slightly exceeding the value reported by Deliman et al., yet still within a comparable range. ~~The mean RMSE values obtained with both models (W2_SD model: 0.37 mg/L; W2_zero-order model: 0.64 mg/L) for Total N were, however, lower than the only reference found for this constituent, which was recorded as 0.77 mg/L (Deliman et al., 2002).~~

The RMSE obtained with the SD model (Run 5 - baseline), Zero-order model (zero-order SOD = 2.5 g O_2/m^2/day - baseline); First-order model (ISC= 0.5 g/m^2 - baseline) and the Hybrid model (zero order SOD= 1.0 g O_2/m^2/day - baseline)~~W2_SD model and the zero-order model~~ for DO, ~~(1.79~~2.01 mg/L, ~~and 1.62~~1.97 mg/L, 2.15 mg/L and 1.87 mg/L respectively) are also in line with the results obtained in other studies (e.g., Deliman et al., 2002: 1.34 mg/L; Brett et al., 2016: 1.2 mg/L; Brito et al., 2018: 7.6 mg/L; Luo et al., 2018: 1.78 mg/L; Tasnim et al., 2021: 2.33 mg/L). In the SD model (Run 5 – baseline), bottom-layer SOD values ranged from 0.015 to 5.152 g O_2/m^{2}day (μ = 1.162; σ = 0.823), reflecting moderate variability driven by seasonal biogeochemical processes. In comparison, the Zero-order model (SOD = 2.5 g O_2/m^{2}day - baseline) showed a broader but more temperature-driven range, from 0.000 to 15.640 g O_2/m^{2}day (μ = 1.432; σ = 2.122). The First-order model (ISC = 0.5 g/m^2 - baseline) yielded values between 0.000 and 20.000 g O_2/m^{2}day, with a much lower mean (μ = 0.870) and relatively high variability (σ = 1.920), consistent with its sensitivity to organic matter loading. The Hybrid model (zero-order SOD = 1.0 g O_2/m^{2}day - baseline) incorporated both zero- and first-order processes and produced the widest overall range, from 0.000 to 21.938 g O_2/m^{2}day (μ = 1.491; σ = 2.024), highlighting its enhanced responsiveness to both physical (e.g., temperature) and biogeochemical (e.g., organic matter) drivers. The monthly variation in SOD across the four models reveals distinct seasonal patterns influenced by their underlying formulations (Fig. A2). All models show notable peaks in May and October, corresponding to periods of elevated organic matter inflow, while a consistent decline is observed during the summer months (June to August), when external organic inputs are comparatively low. The Zero-order model (baseline SOD = 2.5 g $O_2/m^2 \cdot day$) exhibits a sharp rise from winter to a peak of 1.919 g $O_2/m^2 \cdot day$ in May, then gradually declines over the summer, before increasing again in October (1.910 g $O_2/m^2 day$). A similar double-peak pattern is observed in the Hybrid model (zero-order SOD = 1.0 g $O_2/m^2 \cdot day$, baseline), with SOD reaching 1.715 g $O_2/m^2 \cdot day$ in May and a more pronounced maximum of 2.338 g $O_2/m^2 \cdot day$ in October, reflecting the combined effects of temperature and organic matter availability. The SD model (Run 5 – baseline) shows more moderate seasonal variation, with values dipping to 0.679 g $O_2/m^2 \cdot day$ in August, then rising to 1.501 g $O_2/m^2 \cdot day$ in November, consistent with internal sediment dynamics. The First-order model (ISC = 0.5 g/m^2, baseline), which is most sensitive to organic matter loading, also mirrors this seasonal structure, peaking in October (1.235 g $O_2/m^2 day$) after a gradual summer decline. Collectively, these patterns underscore the importance of organic matter availability—particularly in spring and autumn—as a key driver of SOD across the different modeling approaches. This pattern indicates the model's responsiveness to both organic matter inputs and temperature, leading to a more nuanced representation of seasonal variation compared to the other models. ~~In general terms, the metrics obtained for both models and visual analysis of the predicted time series suggest that the CE-QUAL-W2 diagenesis model performed well and produced a more balanced response than the W2_zero-order model. It is also important to evaluate the magnitude of the SOD value associated with each model. Both models performed similarly well with regard to modeling the water column dissolved oxygen (DO). The SOD considered for the W2_zero-order model (2.50 gO_2/m^2/day) was significantly higher than the mean SOD computed with the best W2_SD model (Run 2) (0.810 gO_2/m^2/day). This can be explained by the fact that the W2_zero-order model SOD represents all of the reservoir's DO uptake rate in the water column and not just the sediment uptake. Overall, despite their limitations and assumptions, both models are capable of adequately reproducing the hydrodynamic and water quality dynamics of the reservoir. While not perfect, they effectively capture the main water quality trends given the current boundary conditions. It is important to stress the fact that this study did not set out to identify the best model overall but rather to evaluate the performance of the sediment diagenesis model. Specifically, we sought to determine whether this model could replicate dissolved oxygen (DO) profiles and perform comparably to the zero-order model.~~

~~The zero-order model employs a constant SOD value that only varies with water temperature and does not account for organic matter decay or its impact on SOD values. As a result, it cannot predict water quality trends since it is not integrated with the water column. Nevertheless, it can and was used to back-calculate the oxygen uptake rate in the water column. The fact that the reservoir's dissolved oxygen profiles matched the zero-order model SOD values used in calibration indicated that these values were close to the actual uptake rates of dissolved oxygen (Wells, 2011). Therefore, since the W2 model with the sediment diagenesis function performs similarly to the zero-order model in reproducing dissolved oxygen profiles under the same boundary conditions, it can be concluded that the sediment diagenesis function is performing accurately. This approach proved useful when it came to compensating for the uncertainty related to boundary condition limitations and all the assumptions made for the model's definition.~~

~~The W2_SD model includes a variable SOD computed from the sediment diagenesis model. In the best W2_SD model (Run 2), the SOD ranged from 0.013 to 4.110 gO_2/m^2/day (μ = 0.810; σ = 0.659). The SOD remained below 1 g/m^2/day from~~

~~December to April, increasing during the remaining months and peaking in September each year.~~ These values are consistent with the SOD values obtained in other studies, such as those of Schnoor and Fruh (1979), which concluded that the SOD values of Lake Lydon B. Johnson (located in the U.S.) ranged from 1.7 to 5.8 g ~~$O_2/m^2/day$~~O_2/m^2day, and of Beutel (2015), which measured SOD values in different locations around Lake Hodges (located in the U.S.) ranging from 0.6 to 2.3 g ~~$O_2/m^2/day$~~O_2/m^2day. It would~~, however,~~ be useful to be able to compare these results with SOD values measured at different sites within the Torrão reservoir.~~ Additional SOD monitoring studies need to be conducted in lakes and reservoirs and be extended to other regions, also focusing on the chemical characterization of sediments and the definition of sediment burial rates.~~

It is important to emphasize that this study was primarily designed to evaluate the performance of the sediment diagenesis model. However, by incorporating alternative SOD modeling approaches, it inevitably allowed for a comparative ranking of model performance, highlighting the relative strengths and limitations of each formulation. The performance limitations of the Zero-order and First-order models can be attributed to their structural simplifications. Specifically, the Zero-order model's strong temperature dependence, coupled with its disregard for the dynamics of organic matter loading, reduces its ability to capture temporal variability driven by external inputs. Similarly, the lower accuracy of the First-order model likely stems from its exclusion of anaerobic decay processes and limited representation of sediment biogeochemistry, which becomes especially relevant under low-oxygen conditions. The Hybrid model outperformed all other approaches. Considering the principle of parsimony (Occam's razor) (Burnham and Henderson, 2002), the simpler Hybrid model proved more effective than the complex SD model, making it the preferred choice for simulating SOD dynamics in the reservoir. These findings underscore the importance of selecting models that align with the specific characteristics of the system being studied. Simpler models, such as the Hybrid model, may be adequate for steady-state conditions, short- to medium-term forecasts, or scenarios with limited data. The zero-order SOD component of the Hybrid model relies solely on temperature and is decoupled from the water column; therefore, in long-term simulations, this limitation can gradually undermine the model's accuracy. In contrast, the SD model may be more appropriate when the goal is to explore system-wide feedbacks and temporal dynamics over extended periods—especially those involving sediment accumulation and nutrient cycling—where it may provide valuable insight into underlying processes, provided that sufficient observational data become available to support its additional state variables. Moreover, a model's effectiveness heavily depends on the user's familiarity with its structure and their skill in calibration. Yet, it is unrealistic to expect researchers to master the implementation of every available modeling approach. As such, comparisons between models should be interpreted carefully, acknowledging the influence of user expertise on performance outcomes (Piccolroaz et al. 2024). Overall, to strengthen the analysis, it is recommended that users apply all available SOD modeling approaches in the case of the CE-QUAL-W2 model and assess the model's behavior. This comprehensive evaluation provides a solid foundation for further modeling efforts and helps ensure that the chosen approach is well-suited to the system's specific conditions and objectives.

The results also revealed that the particulate fraction of organic carbon in the reservoir sediments corresponded to ~~40~~80% of the ~~total organic carbon (~~TOC~~)~~. This value is small compared to the results obtained for Taihu Lake by Yu et al. (2022), where the ratio of ~~particulate organic phosphorus (~~POP~~)~~ to TOC varied from 97.85% to 89.53%. However, this value (~~40~~80%) was obtained indirectly through the analysis of the reservoir's predicted SOD values as a function of different initial POC values and may, therefore, reflect other sources of uncertainty, such as inflow organic matter characterization. Given the fact that the magnitude of TOC in the sediment can be affected by numerous factors, including water column productivity, terrestrial inputs of organic materials, sediment properties, and microbial activity rates (Gireeshkumar et al., 2013), and that, partly due to differences in reservoir productivity and morphology, the spatial distribution and sources of organic carbon vary greatly across regions (Anderson et al., 2009), it is reasonable to assume that the only way to accurately assess the POC prediction is by monitoring the reservoir POC content.

Furthermore, this study has highlighted the need to expand research to additional waterbodies across diverse regions to improve our understanding of the CE-QUAL-W2 diagenesis model's performance under varying environmental conditions. This includes evaluating its applicability in long-term scenarios, which are essential for capturing cumulative sediment dynamics and climate-driven trends. Additional SOD monitoring studies need to be conducted in lakes and reservoirs and extended to other latitudes, with particular focus on the chemical characterization of sediments and the definition of sediment burial rates.

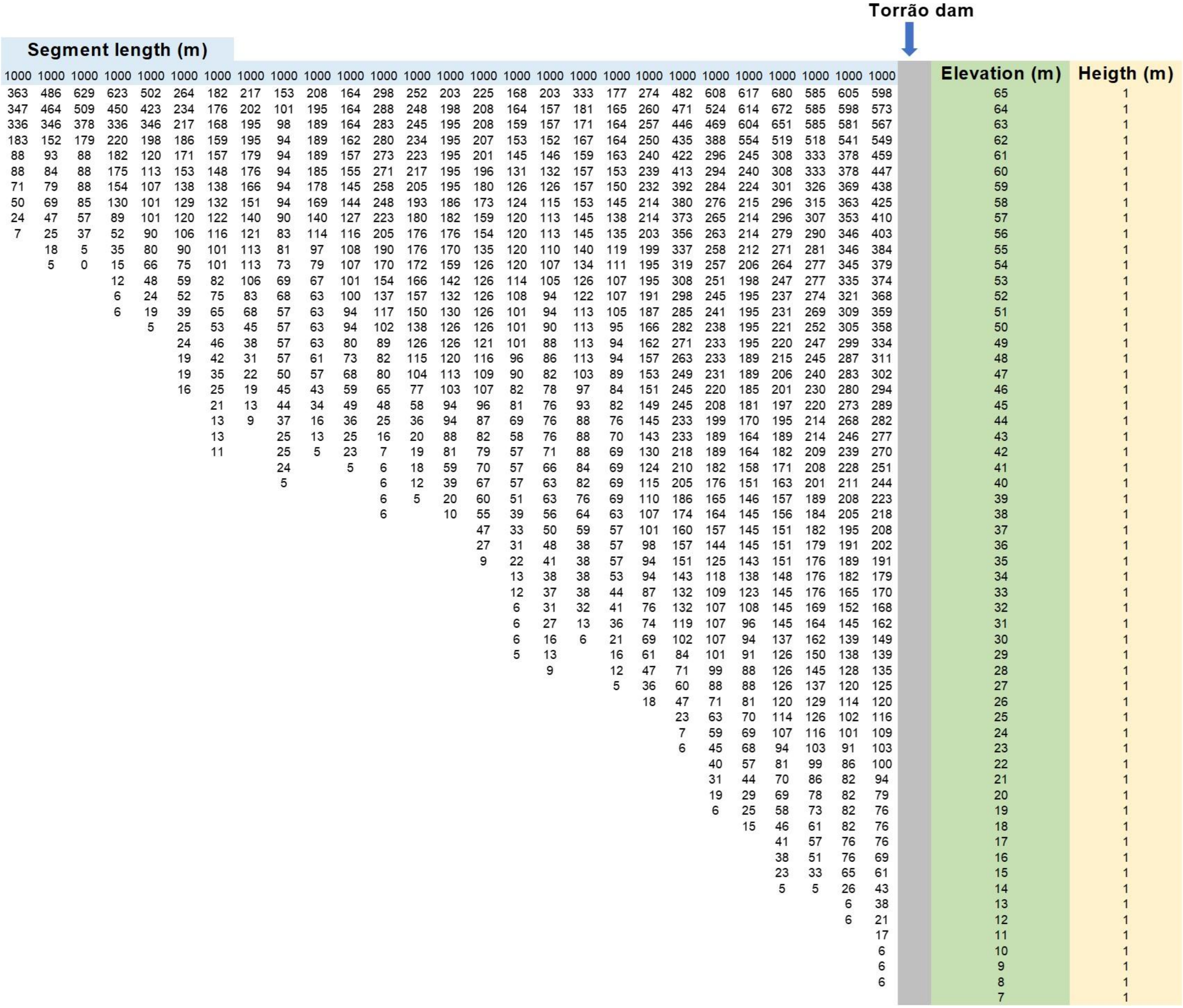
~~Furthermore, this project has highlighted a need to further expand the study to include other water bodies in different regions to enhance our understanding of the performance of the CE-QUAL-W2 diagenesis model and the significant heterogeneity of water body characteristics, as well as underlining the vital importance of establishing performance measures and evaluation criteria for daily water quality simulations in reservoirs and lakes.~~

5 Conclusions

This research evaluates the performance of the CE-QUAL-W2 v4.5 sediment diagenesis model in simulating water temperature, dissolved oxygen, total phosphorus, total nitrogen, chlorophyll-a, and biochemical oxygen demand in a Portuguese reservoir over the period from 2016 to 2021. Calibration was based on 35 sets of observed temperature and dissolved oxygen profiles, supplemented by six annual measurements of total nitrogen, total phosphorus, chlorophyll-a, and biochemical oxygen demand collected at various depths. To evaluate model accuracy, three alternative sediment oxygen demand formulations — a Zero-order model, a First-order model, and a Hybrid approach combining features of both — were also applied and compared. The Hybrid model consistently outperformed the other formulations, striking an effective balance between accuracy and simplicity. It therefore represents the most suitable choice for modeling the reservoir. In contrast, the Zero- and First-order models exhibited limitations related to temperature dependence and inadequate sediment process representation, respectively. Simpler models, such as the Hybrid model, may be adequate for steady-state conditions, short- to medium-term forecasts, or scenarios with limited data. In contrast, the SD model — despite its good performance

— may be more appropriate when the goal is to explore system-wide feedbacks and temporal dynamics over extended periods, especially in cases involving sediment accumulation and nutrient cycling. In such contexts, it may offer valuable insights, provided that sufficient observational data are available to support its additional state variables. Overall, the study reinforces the importance of choosing models based on site characteristics, available data, and simulation goals. Future work should broaden the evaluation of these models across various waterbodies and extended timeframes, while highlighting the need for enhanced sediment monitoring to support detailed process-based modelling.~~The primary objective of this study was to evaluate the performance of the state-of-the-art CE-QUAL-W2 sediment diagenesis model. This evaluation was conducted by applying two different approaches to conceptualize the sediment compartment in the CE-QUAL-W2 model: a) incorporating a sediment diagenesis model, and b) using a zero-order model. The models were applied to the Torrão reservoir located in northern Portugal. Overall, the study's results clearly demonstrate that the CE-QUAL-W2 water quality model performed very well when coupled with the advanced sediment diagenesis model despite the poor characterization of the reservoir inflow water quality. The mean NSE and RMSE obtained for the reservoir water column DO were 0.41±0.67 and 1.73 mg/L±0.69, respectively. The results also suggest that the zero-order model can serve as a good starting point for applying the sediment diagenesis model, as it helps to define an initial threshold for the mean reservoir SOD value.~~

Appendix A



Torrão dam

Segment length (m)																											Elevation (m)	Heigth (m)
1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000		
363	486	629	623	502	264	182	217	153	208	164	298	252	203	225	168	203	333	177	274	482	608	617	680	585	605	598	65	1
347	464	509	450	423	234	176	202	101	195	164	288	248	198	208	164	157	181	165	260	471	524	614	672	585	598	573	64	1
336	346	378	336	346	217	168	195	98	189	164	283	245	195	208	159	157	171	164	257	446	469	604	651	585	581	567	63	1
183	152	179	220	198	186	159	195	94	189	162	280	234	195	207	153	152	167	164	250	435	388	554	519	518	541	549	62	1
88	93	88	182	120	171	157	179	94	189	157	273	223	195	201	145	146	159	163	240	422	296	245	308	333	378	459	61	1
88	84	88	175	113	153	148	176	94	185	155	271	217	195	196	131	132	157	153	239	413	294	240	308	333	378	447	60	1
71	79	88	154	107	138	138	166	94	178	145	258	205	195	180	126	126	157	150	232	392	284	224	301	326	369	438	59	1
50	69	85	130	101	129	132	151	94	169	144	248	193	186	173	124	115	153	145	214	380	276	215	296	315	363	425	58	1
24	47	57	89	101	120	122	140	90	140	127	223	180	182	159	120	113	145	138	214	373	265	214	296	307	353	410	57	1
7	25	37	52	90	106	116	121	83	114	116	205	176	176	154	120	113	145	135	203	356	263	214	279	290	346	403	56	1
	18	5	35	80	90	101	113	81	97	108	190	176	170	135	120	110	140	119	199	337	258	212	271	281	346	384	55	1
	5	0	15	66	75	101	113	73	79	107	170	172	159	126	120	107	134	111	195	319	257	206	264	277	345	379	54	1
			12	48	59	82	106	69	67	101	154	166	142	126	114	105	126	107	195	308	251	198	247	277	335	374	53	1
			6	24	52	75	83	68	63	100	137	157	132	126	108	94	122	107	191	298	245	195	237	274	321	368	52	1
			6	19	39	65	68	57	63	94	117	150	130	126	101	94	113	105	187	285	241	195	231	269	309	359	51	1
				5	25	53	45	57	63	94	102	138	126	126	101	90	113	95	166	282	238	195	221	252	305	358	50	1
					24	46	38	57	63	80	89	126	126	121	101	88	113	94	162	271	233	195	220	247	299	334	49	1
					19	42	31	57	61	73	82	115	120	116	96	86	113	94	157	263	233	189	215	245	287	311	48	1
					19	35	22	50	57	68	80	104	113	109	90	82	103	89	153	249	231	189	206	240	283	302	47	1
					16	25	19	45	43	59	65	77	103	107	82	78	97	84	151	245	220	185	201	230	280	294	46	1
						21	13	44	34	49	48	58	94	96	81	76	93	82	149	245	208	181	197	220	273	289	45	1
						13	9	37	16	36	25	36	94	87	69	76	88	76	145	233	199	170	195	214	268	282	44	1
						13		25	13	25	16	20	88	82	58	76	88	70	143	233	189	164	189	214	246	277	43	1
						11		25	5	23	7	19	81	79	57	71	88	69	130	218	189	164	182	209	239	270	42	1
								24		5	6	18	59	70	57	66	84	69	124	210	182	158	171	208	228	251	41	1
								5			6	12	39	67	57	63	82	69	115	205	176	151	163	201	211	244	40	1
											6	5	20	60	51	63	76	69	110	186	165	146	157	189	208	223	39	1
											6		10	55	39	56	64	63	107	174	164	145	156	184	205	218	38	1
														47	33	50	59	57	101	160	157	145	151	182	195	208	37	1
														27	31	48	38	57	98	157	144	145	151	179	191	202	36	1
														9	22	41	38	57	94	151	125	143	151	176	189	191	35	1
															13	38	38	53	94	143	118	138	148	176	182	179	34	1
															12	37	38	44	87	132	109	123	145	176	165	170	33	1
															6	31	32	41	76	132	107	108	145	169	152	168	32	1
															6	27	13	36	74	119	107	96	145	164	145	162	31	1
															6	16	6	21	69	102	107	94	137	162	139	149	30	1
															5	13		16	61	84	101	91	126	150	138	139	29	1
																9		12	47	71	99	88	126	145	128	135	28	1
																		5	36	60	88	88	126	137	120	125	27	1
																			18	47	71	81	120	129	114	120	26	1
																				23	63	70	114	126	102	116	25	1
																				7	59	69	107	116	101	109	24	1
																				6	45	68	94	103	91	103	23	1
																					40	57	81	99	86	100	22	1
																					31	44	70	86	82	94	21	1
																					19	29	69	78	82	79	20	1
																					6	25	58	73	82	76	19	1
																						15	46	61	82	76	18	1
																							41	57	76	76	17	1
																							38	51	76	69	16	1
																							23	33	65	61	15	1
																							5	5	26	43	14	1
																									6	38	13	1
																									6	21	12	1
																										17	11	1
																										6	10	1
																										6	9	1
																										6	8	1
																											7	1

Figure A1: CE-QUAL-W2 bathymetry - Cross section of the Tâmega River with the average segment width

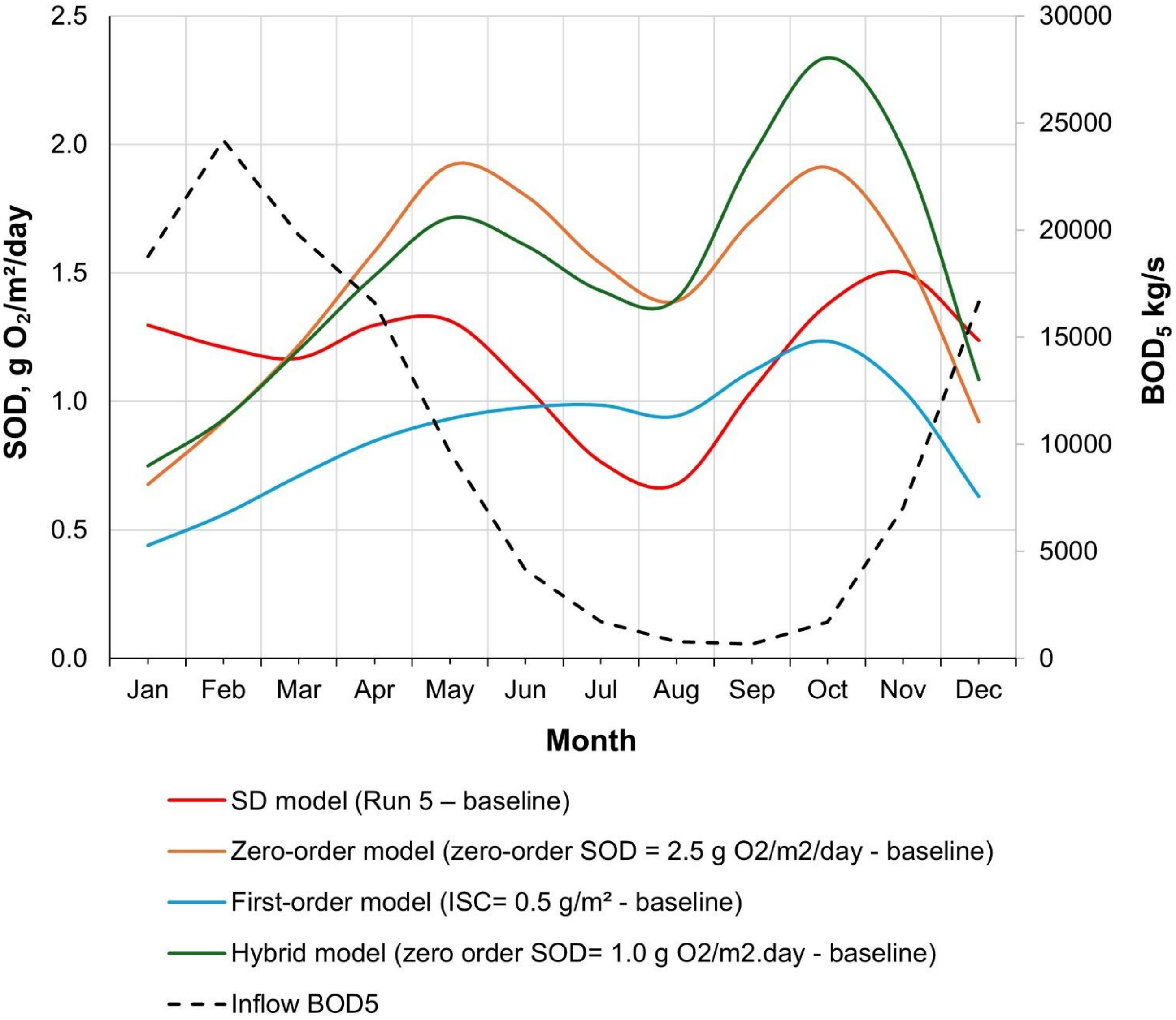


Figure A2: Average monthly sediment oxygen demand (SOD) at the reservoir bottom layer predicted by the SD model (Run 5 - baseline), Zero-order model (baseline: SOD = 2.5 g O_2/m^2day), First-order model (baseline: ISC = 0.5 g/m^2), and Hybrid model (baseline: zero-order SOD = 1.0 g O_2/m^2day). Also shown is the inflow BOD_5 load from the reservoir's main branch.

Table A1. Model metrics. Characterization of Tâmega river inflow

Constituent	Model	R^2	PBIAS	RMSE	MAE	Input features	Input sample number	Sample Time period
Water temperature	LOADEST	0.81	5.14	2.61	1.86	Flow; AirT; WT	21	2016-2021
	XGBOOST	0.85	0.76	2.19	1.76	Air T; WT		
	SVR	0.87	3.77	2.07	1.60	Air T; WT		
DO	LOADEST	0.11	5.02	1.32	1.03	Flow, DO	78	2010-2021
	XGBOOST	0.83	1.50	0.56	0.33	WT; DO; Month		
	SVR	0.91	0.92	0.40	0.26	WT; DO; Month		
Total P	LOADEST	0.21	7.11	0.02	0.01	Flow, Total P	47	2010-2021
	XGBOOST	0.12	-	45.72	32.73	WT; Total P; Month		
	SVR	-	-	38.74	38.74	WT; Total P; Month		
N-NH$_4$	LOADEST	0.11	43.64	0.02	0.02	Flow; N-NH$_4$	46	2010-2021
	XGBOOST	0.26	32.20	0.02	0.02	WT; N-NH$_4$; Month		
	SVR	0.03	28.27	0.02	0.02	WT; N-NH$_4$; Month		
N-NO$_X$	LOADEST	0.11	10.88	0.18	0.14	Flow; N-NO$_X$	63	2010-2021
	XGBOOST	0.15	3.92	0.16	0.12	WT; N-NO$_X$;Month		
	SVR	0.26	7.48	0.16	0.13	WT; N-NO$_X$;Month		
BOD$_5$	LOADEST	0.19	6.93	0.98	0.83	Flow; BOD$_5$	77	2010-2021
	XGBOOST	-	782.16	16.31	16.28	WT; BOD$_5$; Month		
	SVR	0.14	1438.34	30.64	29.93	WT; BOD$_5$; Month		
Chlorophyll-a	LOADEST	0.65	30.00	6.27	4.96	Flow; Chl-a	49	2010-2021
	XGBOOST	0.12	774.02	39.12	29.36	WT; Chl-a; Month		
	SVR	0.02	345.72	15.05	13.25	WT; Chl-a; Month		
Alkalinity	LOADEST	0.26	17.44	6.59	6.30	WT; Chl-a; Month	31	2013-2021
	XGBOOST	0.26	471.67	88.57	57.64	WT; Alkalinity;		
	SVR	0.20	242.35	37.97	30.52	WT; Alkalinity;		
Conductivity	LOADEST	0.80	8.23	8.81	7.69	Flow; Conductivity	77	2010-2021
	XGBOOST	0.32	184.79	133.16	120.16	WT; Conductivity;		
	SVR	0.02	17.67	21.96	17.32	WT; Conductivity;		
TSS	LOADEST	0.00	11.86	1.83	1.45	Flow; SST	78	2010-2021
	XGBOOST	0.19	7.06	1.96	1.38	WT; SST; Month		
	SVR	0.24	8.44	1.62	1.24	WT; SST; Month		

Table A2. Rates and constants: Hydraulic coefficients

Rates and constants	Value
Transport solution scheme	Ultimate
Time-weighting for vertical advection scheme	0.55
Longitudinal eddy viscosity (m^2/s)	1.0
Longitudinal eddy diffusivity (m^2/s)	1.0
Coefficient of bottom heat exchange ($W/m^2/s$)	0.3
Sediment temperature (°C)	14.0
Interfacial friction factor	0.0
Heat lost to sediments that is added back to water column	0.0
Vertical eddy viscosity	W2
Maximum value of vertical eddy speed (m^2/s)	1.0
Bottom friction solution	CHEZY
Wind roughnessheight (m)	0.001

Table A3. Rates and constants: Extinction coefficients

Rates and constants	Value
Extinction for pure water, m^{-1}	0.25
Suspended solids light extinction, 1/m	0.01
EXOM - extinction organic matter, 1/(m mg/l)	0.2
Fraction of incident solar radiation absorbed at the water surface (long-wave components of short-wave solar)	0.45

Table A4. Rates and constants: phytoplankton (Diatoms)

Rates and constants	Value
Maximum algal growth rate (day^{-1})	3
Maximum algal respiration rate (day^{-1})	0.04
Maximum algal excretion rate (day^{-1})	0.04
Maximum algal mortality rate (day^{-1})	0.1
Algal settling rate (day^{-1})	0.1
Algal half-saturation for phosphorus-limited growth, g m^{-3}	0.003
Algal half-saturation for nitrogen limited growth, g m^{-3}	0.014
Fraction of biomass going to POM at death	0.8
AT1 lower temperature for algal growth, °C	10
AT2 lower temperature for maximum algal growth, °C	30
AT3 upper temperature for maximum algal growth, °C	35
AT4 upper temperature for algal growth, °C	40
Fraction of growth rate at AT1 temperature	0.1
Fraction of maximum growth rate at AT2	0.99
Fraction of maximum growth rate at AT3	0.99
Fraction of growth rate at AT4	0.1
Chlorophyll a to algae biomass ratio	0.065

Table A5. Rates and constants: organic matter

Rates and constants	Value
Labile DOM decay rate, day^{-1}	0.3
Refractory DOM decay rate, day^{-1}	0.001
Labile to refractory DOM decay rate, day^{-1}	0.01
Labile POM decay rate, day^{-1}	0.08
Refractory POM decay rate, day^{-1}	0.01
Labile to refractory POM decay rate, day^{-1}	0.001
POM settling rate, m day^{-1}	0.5
Lower temperature for organic matter decay, °C, OMT1	4
Upper temperature for organic matter decay, °C, OMT2	25
Fraction of organic matter decay rate at OMT1	0.1
Fraction of organic matter decay rate at OMT2	0.99

Table A6. Rates and constants: nutrients

Rates and constants	Value
PO4R Sediment release rate of phosphorus, fraction of SOD	0.015
PARTP Phosphorus partitioning coefficient for suspended solids	0.0
NH4R Sediment release rate of ammonium, fraction of SOD	0.15
NH4DK Ammonium decay rate, day^{-1}	0.05
NH4T1 Lower temperature for ammonia decay, ºC	5
NH4T2 Lower temperature for maximum ammonia decay, ºC	25
NH4K1 Fraction of nitrification rate at NH4T1	0.1
NH4K2 Fraction of nitrification rate at NH4T2	0.99
O2NH4 Oxygen stoichiometry for nitrification (mg O_2/mg N)	4.57
NO3DK Water column denitrification rate or nitrate decay rate, day$_{-1}$	0.05
NO3S Nitrate loss velocity to the sediments because of sediment denitrification, m day$_{-1}$	0
FNO3SED Fraction of NO3-N diffused into the sediments that becomes part of organic N in the sediments (The rest is denitrified.)	0.37
NO3T1 Lower temperature for nitrate decay, ºC	5
NO3T2 Lower temperature for maximum nitrate decay, ºC	25
NO3K1 Fraction of denitrification rate at NO3T1	0.1
NO3K2 Fraction of denitrification rate at NO3T2	0.99

Table A7. Rates and constants: SOD rates

Rates and constants	Value
SEDCI Initial first order sediment concentration, g m^{-2}	0.5
SEDK First order sediment decay rate, day^{-1}	0.1
SEDS First order sediment settling or focusing rate, m day^{-1}	0.1
SEDBR First order sediment burial rate, day^{-1}	0.01
DYNSEDK Turns ON/OFF dynamic calculation of the first order sediment model decay rate	ON
SODT1 Lower temperature for zero-order SOD or first-order sediment decay, ºC	4
SODT2 Upper temperature for zero-order SOD or first-order sediment decay, ºC	30
SODK1 Fraction of SOD or sediment decay rate at lower temperature	0.1
SODK2 Fraction of SOD or sediment decay rate at upper temperature	0.99

Table A8. Rates and constants: Sediment diagenesis model (considering 5 regions)

```
"Initial sediment temperature ºC",10,10,10,10,10,
"Initial sediment pH (only used if include modeling dynamic pH)",7,7,7,7,7
"Initial POC (total) concentration
mgC/L",4301,15437,17126,16051,19200,12902
"Initial PON (total) concentration mgN/L",1000,1000,1000,1000,1000
"Initial POP (total) concentration mgP/L",0.001,0.001,0.001,0.001,0.001
"Initial SO4 concentration mgS/L",0.1,0.1,0.1,0.1,0.1
"Initial dissolved NH4 concentration mgN/L",0.01,0.01,0.01,0.01,0.01
"Initial dissolved NO3 concentration mgN/L",0.0,0.0,0.0,0.0,0.0
"Initial total PO4 concentration mgP/L",0.00,0.00,0.00,0.00,0.00
"Initial dissolved H2S concentration mgS/L",0.1,0.1,0.1,0.1,0.1
"Initial CH4 concentration mgC/L",0.092,0.092,0.092,0.092,0.092
"Initial TIC concentration mgC/L",4.011,4.011,4.011,4.011,4.011
"Initial Alkalinity concentration for each region mg/l as CaCO3",50,50,50,50,50
"Initial Ferrous Iron concentration for each region mgFe/l",0.1,0.1,0.1,0.1,0.1
"Initial Iron Oxyhydroxide concentration for each region
mgFe/l,",0.1,0.1,0.1,0.1,0.1
"Initial Mn(II) concentration for each region mgMn/l,",0.1,0.1,0.1,0.1,0.1
"Initial manganese dioxide concentration for each region
mgMn/l,",0.1,0.1,0.1,0.1,0.1
"Number of regions for different diagenesis related rates",1
,Region1,Region2,Region3,Region4,Region5
"Starting segment for regions",2
"Ending segment for regions",21
"Fraction of labile POC",0.1
"Fraction of refractory POC",0.89
"Fraction of labile PON",0.1
"Fraction of refractory PON",0.89
"Fraction of labile POP",0.1
"Fraction of refractory POP",0.89
"Pore water diffusion coefficient m2/d",0.0025
"Particle mixing velocity between aerobic and anaerobic layers m2/d",0.00006
"Burial velocity m/d",0.00000685
"CH4 production calculation method (0: Analytical  1: Numerical)",1,1
"DO threshold for aerobic layer oxidation rates mgO2/L",2,2
"Nitrification rate in aerobic layer at DO below threshold m/d",0.3
"Nitrification rate in aerobic layer at DO above threshold m/d",0.3
"Denitrification rate in aerobic layer at DO below threshold m/d",0.1
"Denitrification rate in aerobic layer at DO above threshold m/d",0.1
"Denitrification rate in anerobic layer m/d",0.1,0.1
"CH4 oxidation rate in aerobic layer m/d",0.7,0.7
"Half-saturation oxygen constant for CH4 oxidation mg-O2/L",0,0
"Nitrification half-saturation constant for NH4 in aerobic layer mgN/L",0.728
"Nitrification half-saturation constant for O2 in aerobic layer mgO2/L",0.37
"Temperature coefficient for pore water diffusion between layers",1.08
"Temperature coefficient for particle mixing diffusion coefficient",1
"Temperature coefficient for nitrification",1.123
"Temperature coefficient for denitrification",1.08
"Temperature coefficient for methane oxidation",1.079
"SO4 concentration above which sulfide over methane is produced mgS/L",20
"H2S oxidation rate in aerobic layer m/d",0.2
"Temperature coefficient for H2S oxidation",1.079
"H2S oxidation normalization constant for O2 mgO2/L",4
"Diagenesis rate for labile POC (G1) 1/d",0.035
"Diagenesis rate for refractory POC (G2) 1/d",0.0018
"Diagenesis rate for inert/slow refractory POC (G3) 1/d",0.0001
"Diagenesis rate for labile PON (G1) 1/d",0.035
"Diagenesis rate for refractory PON (G2) 1/d",0.0018
"Diagenesis rate for inert/slow refractory PON (G3) 1/d",0.0001
"Diagenesis rate for labile POP (G1) 1/d",0.035
"Diagenesis rate for refractory POP (G2) 1/d",0.0018
```

```
"Diagenesis rate for inert/slow refractory POP (G3) 1/d",0.0001
"Temperature coefficient for labile POC",1.1
"Temperature coefficient for refractory POC",1.15
"Temperature coefficient for inert/slow refractory POC",1
"Temperature coefficient for labile PON",1.1
"Temperature coefficient for refractory PON",1.15
"Temperature coefficient for inert/slow refractory PON",1,1
"Temperature coefficient for labile POP",1.1,1.1
"Temperature coefficient for refractory POP",1.15,1.15
"Temperature coefficient for inert/slow refractory POP",1
"PO4 sorption coefficient in anaerobic layer m3/kg",0.02
"Incremental PO4 partition coefficient",0,0
"Critical oxygen concentration for incremental sorption mg-O2/L",0.01
"NH4 sorption coefficient in aerobic layer m3/kg",0.001
"NH4 sorption coefficient in anaerobic layer m3/kg",0.001
"H2S sorption coefficient in aerobic layer m3/kg",0.1
"H2S sorption coefficient in anaerobic layer m3/kg",0.1
"Algorithm for POM resuspension (0: Wind induced resuspension, 1: Bottom scour resuspension) Only used if Include POM resuspension is TRUE.",1,0
"Fe(II) sorption coefficient in aerobic layer m3/g",0.00005
"Fe(II) sorption coefficient in anaerobic layer m3 /g",0.01
"Mn(II) sorption coefficient in aerobic layer m3 /g",0.00005
"Mn(II) sorption coefficient in anaerobic layer  m3 /g",0.01
"Write sediment fluxes", TRUE.
"Frequency of output days",1
```

Table ~~A2~~A9. Metrics between observed and predicted values for both models. The predicted values were compared with observed values at three different depths: (a) an integrated sample between the reservoir surface and an average depth of 5.8 meters, (b) an average depth of 23 meters, and (c) an average depth of 43.7 meters. The values shown in this table represent the mean value of the metrics obtained for each date and the corresponding standard deviation.

Constituent	**SD model (run 5 - baseline)~~W2_SD model (run 5 - baseline)~~**			
	R^2	**PBIAS**	**RMSE**	**MAE**
Water temperature	0.95~~0.56±0.13~~	-3.34~~-0.11±0.71~~	1.31~~3.51±0.32~~	1.05~~1.56±0.10~~
DO	0.79~~0.55±0.31~~	-9.68~~-8.26±14.70~~	2.01~~1.79±0.63~~	1.40~~1.34±0.47~~
Total N	0.19~~0.03±0.03~~	-15.85~~-13.46±14.10~~	0.33~~0.37±0.03~~	0.27~~0.29±0.03~~
Total P	0.09~~0.16±0.08~~	-11.44~~-3.98±34.11~~	0.03~~0.02±0.02~~	0.01~~0.01±0.01~~
BOD$_5$	0.02~~0.03±0.01~~	-50.92~~-1.46±20.17~~	3.24~~0.52±0.17~~	1.58~~19.58±0.14~~
Chl-a	0.14~~0.01±0.01~~	51.62~~87.44±43.14~~	17.12~~13.43±12.81~~	11.51~~11.16±9.69~~
Constituent	**Zero-order model (zero order SOD = 2.5 g O_2/m^2/day - baseline)~~W2_zero-order model (SOD = 2.5 g O2/m2/day - baseline)~~**			
	R^2	**PBIAS**	**RMSE**	**MAE**
Water temperature	0.95~~0.56±0.13~~	-3.34~~-0.12±0.80~~	1.32~~3.51±0.32~~	1.06~~1.57±0.10~~
DO	0.84~~0.54±0.31~~	-5.91~~-4.77±15.52~~	1.69~~1.62±0.73~~	1.12~~1.20±0.55~~
Total N	0.22~~0.30±0.07~~	-1.07~~-4.19±44.37~~	0.79~~0.64±0.45~~	0.46~~0.45±0.22~~
Total P	0.26~~0.18±0.07~~	103.43~~74.68±89.24~~	0.10~~0.06±0.07~~	0.04~~0.04±0.04~~
BOD$_5$	0.03~~0.01±0.01~~	-50.93~~-0.73±18.73~~	3.24~~0.55±0.22~~	1.58~~19.73±0.17~~
Chl-a	0.09~~0.00±0.00~~	48.89~~190.46±183.89~~	17.78~~15.84±12.80~~	12.17~~12.11±10.90~~
Constituent	**First-order model (ISC= 0.5 g/m² - baseline)**			
	R^2	**PBIAS**	**RMSE**	**MAE**
Water temperature	0.95	-3.18	1.31	1.05
DO	0.83	9.21	1.84	1.28
Total N	0.14	-24.73	0.36	0.28
Total P	0.10	-3.71	0.03	0.02
BOD$_5$	0.00	-53.34	3.24	1.59
Chl-a	0.06	-60.39	14.88	8.49
Constituent	**Hybrid model (zero order SOD= 1.0 g O_2/m^2/day - baseline)**			
	R^2	**PBIAS**	**RMSE**	**MAE**
Water temperature	0.95	-3.20	1.31	1.05
DO	0.87	-2.34	1.52	1.03

Total N	0.31	-18.75	0.35	0.28
Total P	0.27	36.49	0.04	0.02
BOD_5	0.01	-51.93	3.25	1.61
Chl-a	0.06	-59.55	14.95	8.52

Table A10. Metrics between observed and predicted values for SD model (run 4). Water temperature and DO metrics were obtained from 36 observed and predicted profiles. The predicted values for the remaining constituents were compared with observed values at three different depths: (a) an integrated sample between the reservoir surface and an average depth of 5.8 meters, (b) an average depth of 23 meters and (c) an average depth of 43.7 meters. The values in this table represent the mean value of the metrics obtained at each date and the corresponding standard deviation or, in the case of water temperature and DO, the mean value of the metrics obtained for each profile and the standard deviation.

Constituent	**SD model (Run 4)**			
	R^2	**PBIAS**	**RMSE**	**MAE**
Water temperature	0.95	-3.45	1.34	1.07
DO	0.79	-8.20	1.98	1.37
Total N	0.18	-15.52	0.33	0.27
Total P	0.09	-11.18	0.03	0.01
BOD5	0.02	-50.93	3.24	1.58
Chl-a	0.15	52.95	16.97	11.48

Table A~~A3~~11. Root mean square error values obtained with different models and across different time frames

Constituent	Model	Simulation	RMSE	Author
Water temperature	CE-QUAL-W2	1984-1991	2.95 ºC	Deliman et al., 2002
	CE-QUAL-W2	2005-2010	1.93 ºC	Kim and Kim, 2006
	CE-QUAL-W2	1991-2000	0.56 ºC	Zhang et al., 2015
	CE-QUAL-W2	2011-2013	< 2.0 ºC	Lindenschmidt et al., 2019
	Delft3D-FLOW	2015-2017	0.96 to 1.0ºC	Piccioni, et al. 2020
	CE-QUAL-W2	2001-2011	1.80 ºC	Brito et al., 2018
	CE-QUAL-W2	2010	2.36 ºC	Liu et al., 2019
	MINLAKE2020	~~2-4 years~~2007-2009	1.51 ºC	Tasnim et al., 2021
	CE-QUAL-W2	2000-2019	3.01-3.17 ºC	Almeida et al., 2023
DO	CE-QUAL-W2	1984-1991	1.34 mg/l	Deliman et al., 2002
	CE-QUAL-W2	1991-2000	0.61 mg/l	Zhang et al., 2015
	CE-QUAL-W2	2001	1.2 mg/l	Brett et al., 2016
	CE-QUAL-W2	2001-2011	7.6 mg/l	Brito et al., 2018
	DYRESM 4.0	2009-2011	1.78 mg/l	Luo et al., 2018
	MINLAKE2020	2007-2009~~2-4 years~~	2.33 mg/l	Tasnim et al., 2021
	CE-QUAL-W2	2000-2019	2.22-3.46 mg/l	Almeida et al., 2023
Total P	CE-QUAL-W2	2001	0.012 mg/l	Brett et al., 2016
	CE-QUAL-W2	2005-2010	0.014 to 0.068 mg/l	Kim et al., 2019
	MINLAKE2020	2007-2009~~2-4 years~~	0.005 to 0.036 mg/l	Tasnim et al., 2021
	CE-QUAL-W2	2000-2019	0.07-0.09 mg/l	Almeida et al., 2023
Total N	CE-QUAL-W2	1984-1991	0.77 mg/L	Deliman et al., 2002
BOD_5	CE-QUAL-W2	2000-2019	3.06-4.83 mg/l	Almeida et al., 2023
	CE-QUAL-W2	1991-2000	1.08 µg/l	Zhang et al., 2015
Chl-a	CE-QUAL-W2	2001	4.6 µg/l	Brett et al., 2016
	CE-QUAL-W2	2005-2010	6.7 to 13.2 µg/l	Kim et al., 2019
	CE-QUAL-W2	2001-2011	62.9 µg/l	Brito et al., 2018
	MINLAKE2020	2007-2009~~2-4 years~~	0.6 - 27.6 µg/l	Tasnim et al., 2021
	CE-QUAL-W2	2000-2019	19.36-25.57 µg/l	Almeida et al., 2023

~~**Table A4. Metrics between observed and predicted values for W2_SD model (Run 2). Water temperature and DO metrics were obtained from 36 observed and predicted profiles. The predicted values for the remaining constituents were compared with observed values at three different depths: (a) an integrated sample between the reservoir surface and an average depth of 5.8 meters, (b) an average depth of 23 meters and (c) an average depth of 43.7 meters. The values in this table represent the mean value of the metrics obtained at each date and the corresponding standard deviation or, in the case of water temperature and DO, the mean value of the metrics obtained for each profile and the standard deviation.**~~

~~**Constituent**~~	~~**W2_SD model (Run 2)**~~			
	~~**R^2**~~	~~**PBIAS**~~	~~**RMSE**~~	~~**MAE**~~
~~Water temperature~~	~~0.56±0.13~~	~~-0.22±0.87~~	~~3.51±0.32~~	~~1.58±0.08~~
~~DO~~	~~0.53±0.30~~	~~0.93±16.44~~	~~1.73±0.69~~	~~1.33±0.56~~
~~Total N~~	~~0.03±0.03~~	~~-13.64±15.67~~	~~0.37±0.03~~	~~0.29±0.04~~
~~Total P~~	~~0.17±0.08~~	~~-3.98±34.46~~	~~0.02±0.02~~	~~0.01±0.01~~
~~BOD5~~	~~0.03±0.01~~	~~-3.61±19.35~~	~~0.51±0.18~~	~~19.70±0.15~~
~~Chl-a~~	~~0.01±0.01~~	~~85.06±42.15~~	~~13.26±12.65~~	~~11.11±9.77~~

Code availability

The exact version of the models' source code is archived on Zenodo at https://doi.org/10.5281/zenodo.14606105 (Almeida and Coelho 2025). The current version of the open-source CEQUAL-W2 model (version 4.5) used in this study is also available from the project website (http://www.ce.pdx.edu/w2/, last access 24 January 2024).

Data availability

Input files needed to run the models' and the hydrometric water quality and meteorological datasets used to force and validate each model are freely available and are archive on Zenodo at https://doi.org/10.5281/zenodo.14606105 (Almeida and Coelho 2025).

Competing interests

The contact author has declared that none of the authors has any competing interests.

Acknowledgements

The authors acknowledge the funding received from Fundação para a Ciência e a Tecnologia (FCT, Portugal), through the UIDB/04292/2020 and UIDP/04292/2020 strategic projects granted to the Marine and Environmental Sciences Centre (MARE) and the LA/P/0069/2020 project granted to the Aquatic Research Network Associate Laboratory (ARNET).

Author contribution: MA conceptualized the study, developed the methodology, and handled software and data curation, as well as writing the original draft. PC administered the project and contributed to reviewing and editing the manuscript

References

Abbaspour, K. C., Rouholahnejad, E., Vaghefi, S., Srinivasan, R., Yang, H. and Kløve, B.: A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model, Journal of Hydrology 524, 733–752, doi.org/10.1016/j.jhydrol.2015.03.027, 2015.

Adedeji, I.C., Ahmadisharaf, E., and Sun, Y.: Predicting in-stream water quality constituents at the watershed scale using machine learning. J. Contam. Hydrol., doi.org/10.1016/j.jconhyd.2022.104078, 2022.

Almeida, M., and Coelho, P.:Evaluating the performance of CE-QUAL-W2 version 4.5 sediment diagenesis model (Manuscript related material: input data and source code) (1.0.0) [Data set]. Zenodo https://doi.org/10.5281/zenodo.15775127~~. https://doi.org/10.5281/zenodo.14606105~~, 2025.

Almeida, M. and Coelho, P.S.: An integrated approach based on the correction of imbalanced small datasets and the application of machine learning algorithms to predict total phosphorus concentration in rivers, Eco. Inform., 76, 102138 doi.org/10.1016/j.ecoinf.2023.102138, 2023a.

Almeida, M., Coelho, P.: A first assessment of ERA5 and ERA5-Land reanalysis air temperature in Portugal, International Journal of Climatology, 43(14), 6643–6663, doi.org/10.1002/JOC.8225, 2023b.

Almeida, M., Rebelo, R., Costa, S. Rodrigues, A. C., and Coelho, P. S.: Long-Term Water Quality Modeling of a Shallow Eutrophic Lagoon with Limited Forcing Data. Environ Model Assess, doi.org/10.1007/s10666-022-09844-3, 2023.

Anderson, N. J., D'Andrea, W., and Fritz, S. C.: Holocene carbon burial by lakes in SW Greenland, Glob. Change Bio., 15, 2590–2598, doi.org/10.1111/j.1365-2486.2009.01942.x, 2009.

Barbosa, S., Scotto, M.G.: Extreme heat events in the Iberia Peninsula from extreme value mixture modeling of ERA5-Land air temperature, Weather. Clim. Extrem. 36, 100448, doi.org/10.1016/j.wace.2022.100448, 2022.

Berger, C. and Wells, S.: Updating the CEMA Oil Sands Pit Lake Model. Prepared for CEMA. August 2014, 2014.

Bergstra, J., Yamins, D., and Cox, D.D.: Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In: To Appear in Proc. of the 30th International Conference on Machine Learning (ICML 2013), 2013.

Beutel, M.: Lake Hodges Reservoir Oxygen Demand Study. Final Assessment Report. City of San Diego Public Utilities Department., 2015.

Brett, M. T., Ahopelto, S. K., Brown, H. K., Brynestad, B. E., Butcher, T.W., Coba, E. E., Curtis, C. A., Dara, J. T., Doeden, K. B., Evans, K. R., Fan L., Finley, J. D., Garguilo, N. J., Gebreeyesus, S. M., Goodman, M. K., Gray, K. W., Grinnell, C., Gross, K. L., Hite, B. R. E., Jones, A. J., Kenyon, P. T., Klock, A. M., Koshy, R. E., Lawler, A. M., Lu, M., Martinkosky, L., Miller-Schulze, J. R., Nguyen, Q. T. N., Runde, E. R., Stultz, J. M., Wang, S., White, F. P., Wilson, C. H., Wong, A. S., Wu, S. Y., Wurden, P. G., Young, T. R., and Arhonditsis, G. B.: The modeled and observed response of Lake Spokane hypolimnetic dissolved oxygen concentrations to phosphorus inputs, Lake and Reservoir Management, 32–3, 246-258, doi: 10.1080/10402381.2016.1170079, 2016.

Brito, D., Ramos, T.B., Gonçalves, M.C., Morais, M., and Neves, R.:Integrated modelling for water quality management in a eutrophic reservoir in south - eastern Portugal, Environ. Earth Sci. 77, doi.org/10.1007/s12665-017-7221-5., 2018.

Burnham, K. P., and Anderson, D. R.: Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach (2nd ed.). Springer, 2002.

Cardoso, R. M., Soares, P. M. M., Miranda, P. M. A., and Belo-Pereira, M.: WRF High resolution simulation of Iberian mean and extreme precipitation climate, Int. Journal Climat., 33, 2591–2608, doi.org/10.1002/joc.3616., 2013.

Chapman, D. V.: Water quality assessments: A guide to the use of biota, sediments, and water in environmental monitoring, second ed. CRC Press. ISBN 0 419 21590 5 (HB) 0 419 21600 6 (PB), 1996.

Chen, T. and Guestrin, C.: XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794), New York, NY, USA: ACM, doi.org/10.1145/2939672.2939785, 2016.

Chuo, M., Ma, J., Liu, D., and Yang, Z.: Effects of the impounding process during the flood season on algal blooms in Xiangxi Bay in the Three Gorges Reservoir, China. Ecol. Model, 392.236–249. https://doi.org/10.1016/j.ecolmodel.2018.11.017, 2019.

Cole, T.M. and Wells, S.A.: CE-QUAL-W2: A Two-dimensional, Laterally Averaged, Hydrodynamic and Water Quality Model, Version 3.5., 2006.

Deliman, P.N. and Gerald, J. A.: Application of the Two- Dimensional Hydrothermal and Water Quality Model, CE-QUAL-W2, to the Chesapeake Bay – Conowingo Reservoir, Lake and Reservoir Management. 18:1, 10-19. doi.org/10.1080/07438140209353925, 2002.

Feigl, M., Lebiedzinski, K., Herrnegger, M., and Schulz, K.: Machine-learning methods for stream water temperature prediction, Hydrol. Earth Syst. Sci. 25, 2951–2977, doi.org/10.5194/hess-25-2951-2021, 2021.

Galloway, J. M.: Determination of Organic and Inorganic Percentages and Mass of Suspended Material at Four Sites in the Illinois River in Northwestern Arkansas and Northeastern Oklahoma. Scientific Investigations Report. Geological Survey (U.S.), 2008.

Gireeshkumar,. R., Deepulal, P. M., and Chandramohanakumar, N.: Distribution and sources of sedimentary organic matter in a tropical estuary, southwest coast of India (Cochin estuary): A baseline study, Mar. Pollut. Bull., 66, 239–245, doi.org/10.1016/j.marpolbul.2012.10.002, 2013.

Hamilton, D.P. and Schladow, S.G.: Prediction of water quality in lakes and reservoirs. Part I — model description. Ecol. Model., 96, 91–110. doi.org/10.1016/ S0304-3800(96)00062-2, 1997.

Ji, X. and Lu, J.: Forecasting riverine total nitrogen loads using wavelet analysis and support vector regression combination model in an agricultural watershed. Environ. Sci. Pollut. Res. Int. 25 (26), 26405–26422, doi.org/10.1007/s11356-018- 2698-3, 2018.

Kim, D., Kim, Y., and Kim, B.: Simulation of eutrophication in a reservoir by CE-QUAL-W2 for the evaluation of the importance of point sources and summer monsoon. Lake Reservoir Manag. 35, 64–76, doi.org/10.1080/10402381.2018.1530318, 20019.

Kobler, U.G., Wüest, A., and Schmid, M.: Effects of lake–reservoir pumped-storage operations on temperature and water quality. Sustainability. 10, 1968. https://doi.org/ 10.3390/su10061968, 2018.Lewis, W. M. and Saunders, J. F.: Concentration and Transport of Dissolved and Suspended Substances in the Orinoco River. Biogeochemistry. 7(3), 203–240. http://www.jstor.org/stable/1468846, 1989.

Lindenschmidt, KE., Carr, M.K., Sadeghian, A. et al.: CE-QUAL-W2 model of dam outflow elevation impact on temperature, dissolved oxygen and nutrients in a reservoir. Sci Data 6, 312 (2019), doi.org/10.1038/s41597-019-0316-y., 2019.

Liu, M., Chen, X., Chen, Y., Gao, L., and Deng, H.: Nitrogen retention effects under reservoir regulation at multiple time scales in a subtropical river basin. Water (Switzerland) 11, doi.org/10.3390/w11081685, 2019.

Loucks, D.P. and van Beek, E.: Water Quality Modeling and Prediction. In: Water Resource Systems Planning and Management. Springer, Cham, doi.org/10.1007/978-3-319-44234-1_10, 2017.

Lu, H. and Ma, X.: Hybrid decision tree-based machine learning models for short-term water quality prediction, Chemosphere, 249, 126169, doi.org/10.1016/j.chemosphere.126169, 2020.

Luo, L., Hamilton, D., Lan, J., Mcbride, C., and Trolle, D.: Autocalibration of a one-dimensional hydrodynamic-ecological model (DYRESM 4.0-CAEDYM 3.1) using a Monte Carlo approach: simulations of hypoxic events in a polymictic lake, Geosci. Model Dev., (GMD) 11, 903–913, 2018.

Mi, C., Shatwell, T., Ma, J., Xu, Y., Su, F., and Rinke, K.: Ensemble warming projections in Germany's largest drinking water reservoir and potential adaptation strategies. Science of the Total Environment 748: 141366, doi.org/10.1016/j.scitotenv.2020.141366, 2020.

Mi, C., Shatwell, T., Kong, X., and Rinke, K.: Cascading climate effects in deep reservoirs: Full assessment of physical and biogeochemical dynamics under ensemble climate projections and ways towards adaptation, Ambio, doi.org/10.1007/s13280-023-01950-0, 2023.

Minear, J.T., and Kondolf, G.M.: Estimating reservoir sedimentation rates at large spatial and temporal scales: A case study of California, Water Resources Research, vol. 45, W12502, doi.org/10.1029/2007WR006703, 2009.

Moriasi, D.N., Gitau, M.W., Pai, N., and Daggupati, P.: Hydrologic and water quality models: performance measures and evaluation criteria, Trans. ASABE. 58, 1763–1785, doi.org/10.13031/trans.58.10715, 2015.

Muñoz-Sabater, J.: ERA5-Land hourly data from 1981 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), accessed 14 December 2022), doi.org/10.24381/cds.e2161bac, 2019.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models: Part 1. A discussion of principles. J. Hydrology, 10(3), 282-290, doi.org/10.1016/0022-1694(70)90255-6., 1970.

OECD: Organisation for the Economic Cooperation and Development. Eutrophication of Waters – Monitoring Assessment and Control. OECD (ed.), Paris, 125 pp., 1982.

Park, Y., Cho, K. H., Kang, J.-H., Lee, S. W. and Kim, J. H.: Developing a flow control strategy to reduce nutrient load in a reclaimed multi-reservoir system using a 2D hydrodynamic and water quality model. Sci. Total. Environ. 466–467, 871–880, 2014.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: machine learning in Python, JMLR, 12, 2825–2830, doi.org/10.48550/ arXiv.1201.0490, 2011.

Piccioni, F., Casenave, C., Lemaire, B.J., Le Moigne, P., Dubois, P., and Vinçon-Leite, B.: The response of small and shallow lakes to climate change: new insights from hindcast modelling. Earth Syst. Dyn. Discuss. 1–22. doi.org/10.5194/esd- 2020-51, 2020.

Piccolroaz, S., Zhu, S., Ladwig, R.,Carrea, L., Oliver, S., Piotrowski, A. P., Ptak. M., Shinohara, R., Sojka, M., Woolway, R.I., and Zhu, D.Z.: Lake water temperature modeling in an Era of climate change: Data sources, models, and future prospects. Reviews of Geophysics. 62. e2023RG000816. doi.org/10.1029/2023RG000816, 2024.

Prakash, S., Vandenberg, J.A. and Buchak, E.M.: Sediment Diagenesis Module for CE-QUAL-W2 Part 2: Numerical Formulation. Environ Model Assess 20, 249–258. https://doi.org/10.1007/s10666-015-9459-1, 2015.

Rajesh, M. and Rehana, S.: Prediction of river water temperature using machine learning algorithms: a tropical river system of India. J. Hydroinf. 23 (3), 605–626, doi.org/10.2166/hydro.2021.121., 2021.

Runkel, R. L., Crawford, C. G. and Cohn, T. A.: Load estimator (LOADEST): a FORTRAN program for estimating constituent loads in streams and rivers. USGS Techniques and Methods Book 4, Chapter A5. U.S. Geological Survey, Reston, Virginia, USA, 2004.Sadeghian, A., Chapra, S. C., Hudson, J., Wheater, H. and Lindenschmidt, K. E.: Improving in-lake water quality modeling using variable chlorophyll a/algal biomass ratios. Environ. Modell. softw. 101, 73–85, 2018.

Schnoor, J.L. and Fruh, E.G.: Dissolved Oxygen Model of a Short Detention Time Reservoir with Anaerobic Hypolimnion", Water Resources Bulletin, Vol 15, No. 2, pp 506-518, 1979.

SNIRH: National information system of water resources. Portuguese Environmental Agency [Sistema Nacional de Informacao de Recursos Hidricos], (Available at: http:// www. https:// snirh.apambiente.pt), 2024.

Soares, P. M. M, Cardoso, R. M., Ferreira, J.J., and Miranda, P. M. A.: Climate change and the Portuguese precipitation: ENSEMBLES regional climate models results, Clim Dyn. 45, 1771–1787, doi.org/10.1007/s00382-014-2432-x., 2015.

Tasnim, B., Fang, X., Hayworth, J. S., and Tian, D.: Simulating Nutrients and Phytoplankton Dynamics in Lakes: Model Development and Applications, Water 2021, 13, 2088, doi.org/10.3390/w13152088, 2021.

Tavera-Quiroz, H., Rosso-Pinto, M., Hernández, G., Pinto, S., Canales, F.A.: Water Quality Analysis of a Tropical Reservoir Based on Temperature and Dissolved Oxygen Modeling by CE-QUAL-W2. Water, 15, 1013. https://doi.org/10.3390/w15061013, 2023.Terry JA, Sadeghian A, and Lindenschmidt K.E. Modelling Dissolved Oxygen/Sediment Oxygen Demand under Ice in a Shallow Eutrophic Prairie Reservoir Water, 9(2):131. doi.org/10.3390/w9020131, 2017.

Uhlmann, W.: A Model-based Study on the Discharge of Iron-rich Groundwater Into the Lusatian Post-mining Lake Lohsa. Mine Water and Circular Economy, Germany, 2017.

Vandenberg, J.A., Prakash, S. and Buchak, E.M.: Sediment Diagenesis Module for CE-QUAL-W2. Part 1: Conceptual Formulation. Environ Model Assess 20, 239–247. https://doi.org/10.1007/s10666-014-9428-0, 2015.

Varis, O., Kuikka, S., and Taskinen, A.: Modeling for water quality decisions: Uncertainty and subjectivity in information, in objectives, and in model structure. Ecological Modelling, 74 (1-2) 91-101, doi.org/10.1016/0304-3800(94)90113-9, 1994.

Walker, W.W.: FLUX stream load computations. Version 4.4. Prep. for U.S. Army Corps Eng. Waterways Exp. Sta. Vicksburg, MS., 1990.

Wells, S.A.: CE-QUAL-W2: A Two-Dimensional, Laterally Averaged, Hydrodynamic and Water Quality Model, Version 4.5, Department of Civil and Environmental Engineering Portland State University, https://www.ce.pdx.edu/w2/, 2021.

Whitehead, P., Wilby, R., Battarbee, R., Kernan, M., and Wade, A. J.: A review of the potential impacts of climate change on surface water quality. Hydrological Sciences Journal, 54(1), 101–123, doi.org/10.1623/hysj.54.1.101, 2009.

Xu, G., Fan, H., Oliver, D.M., Dai, Y., Li, H., Shi, Y., Long, H., Xiong, K., and Zhao, Z.: Decoding river pollution trends and their landscape determinants in an ecologically fragile karst basin using a machine learning model. Environ. Res. 2022 (214), 113843, doi.org/10.1016/j.envres.2022.113843, 2022.

Yu, K., Zhang, Y., He, X., Zhao, Z., Zhang, M., Chen, Y., Lang, X., and Wang, Y.: Characteristics and environmental significance of organic carbon in sediments from taihu Lake, China. Ecol. Indic. 138, 108796, doi.org/10.1016/j. ecolind.2022.108796, 2022.

Zhang, Z., Sun, B., and Johnson, B.E.: Integration of a benthic sediment diagenesis module into the two dimensional hydrodynamic and water quality model—CE-QUAL-W2. Ecol. Model., 297, 213–231, doi.org/10.1016/j.ecolmodel.2014.10.025, 2015.

Zouabi-Aloui, B., Adelana, S. M. and Gueddari, M.: Effects of selective withdrawal on hydrodynamics and water quality of a thermally stratified reservoir in the southern side of the Mediterranean Sea: a simulation approach. Environ. Monit. Assess. 187, 292–311, 2015.