Science of the Total Environment 1038 (2026) 181900

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Modelling nutrient loads in data-scarce large catchments using spatially regularized ensemble calibration

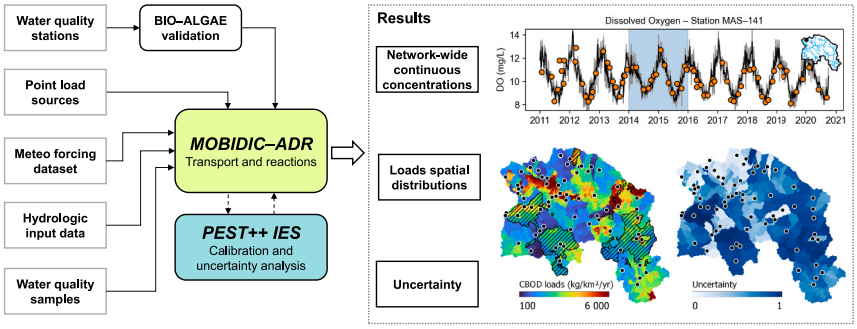
Matteo Masi [a,*], Maryam Barati Moghaddam [b], Fabio Castelli [a], Chiara Arrighi [a]

[a] Department of Civil and Environmental Engineering, University of Florence, Via di S. Marta 3, 50139, Firenze, Italy
[b] Department of Water Resources and Environmental Modelling, Faculty of Environmental Sciences, Czech University of Life Sciences Prague (CZU), Czechia

HIGHLIGHTS

- A water quality model is developed for data-poor large catchments.
- Inversion with Tikhonov regularization is used to estimate nutrient loads.
- Calibration is achieved efficiently using an iterative ensemble smoother method.
- Model parameter and predictive uncertainty are quantified.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Surface water quality
Reactive transport
Nutrient dynamics
Uncertainty quantification
Iterative ensemble smoother
Tikhonov regularization

ABSTRACT

Nutrient pollution in freshwater systems poses major ecological challenges, requiring robust modelling tools to support effective water management. However, catchment-scale water quality modelling is often constrained by sparse monitoring networks and high parameter uncertainty associated with complex biogeochemical systems. This study presents an integrated modelling framework combining the MOBIDIC–ADR hydrological model with a newly developed BIO–ALGAE reactive component to simulate nutrient dynamics across large catchments. The framework employs spatially regularized ensemble calibration using PEST++ iterative ensemble smoother to estimate distributed pollutant loads while quantifying parameter and predictive uncertainty. The model was applied to the Arno River catchment (7990 km^2) in Italy, simulating 8 water quality constituents including dissolved oxygen, nitrogen species, phosphorus compounds, and algal biomass over a ten-year period (2011–2020). Using 8151 observations from sparse sampling locations, the calibration demonstrated the model's ability to reproduce observed patterns across multiple constituents. The model proved effective in identifying pollution hotspots, highlighting strong associations between urban areas and elevated carbonaceous biochemical oxygen demand and ammonium loads, whereas phosphorus displayed a more heterogeneous spatial distribution indicative of multiple source contributions. Despite limitations under low dissolved oxygen conditions, the approach effectively captured first-order reactive processes and provided spatially explicit load estimates with uncertainty bounds. This framework offers a practical decision-support tool for targeted water quality management in data-scarce environments.

* Corresponding author.
E-mail address: matteo.masi@unifi.it (M. Masi).

https://doi.org/10.1016/j.scitotenv.2026.181900
Received 24 November 2025; Received in revised form 24 April 2026; Accepted 19 May 2026
Available online 22 May 2026

1. Introduction

Nutrient pollution in freshwater systems is a key ecological concern that has led to the development of numerous European water policies, including the Water Framework Directive (WFD), the Nitrates Directive (91/676/EEC), and the Urban Waste Water Treatment Directive (91/271/EEC) (Arheimer and Olsson, 2003). The complex interplay between flow regimes, temperature fluctuations, biogeochemical reactions, and nutrient availability, within the broader context of climate and land-use change, makes nutrient dynamics modelling essential for water management, albeit extremely challenging (Abbaspour et al., 2015; Cao et al., 2025; Grizzetti et al., 2008; Stutter et al., 2023; Vystavna et al., 2023; Whitehead et al., 2006). While the complexity is reasonably resolved when making simulation at the river reach scale (Bai et al., 2022; Costa et al., 2021; Lucas et al., 2025), modelling large catchments to support the implementation of policy is still challenging (Keller et al., 2023; Strokal et al., 2024) due limited and uneven observations, scale mismatches (process and data), parameter identifiability and uncertainty, and difficulties validating models at large spatial extents.

Abbott et al. (1986) laid the conceptual foundation of catchment-scale distributed hydrological modelling by introducing the Système Hydrologique Européen (SHE), a physically based modelling framework that emphasized the importance of capturing spatial variability in flow paths, soil characteristics, and land use. Their work catalysed the development of integrated modelling frameworks designed to simulate water movement and nutrient transport across heterogeneous catchments. By the late 1990s, researchers began incorporating biogeochemical processes into hydrological frameworks, leading to development of models such as INCA (Integrated Nitrogen in Catchments) by the Aquatic Environments Research Centre at the University of Reading (Wade et al., 2002; Whitehead et al., 1998). INCA simulates nitrogen dynamics using mass balance and reaction kinetics, capturing key processes such as mineralisation, immobilisation, nitrification, and denitrification. Its semi-distributed structure allows for the representation of land use classes within sub-catchments, which are sequentially linked to a multi-reach river system to produce daily time series of nitrogen concentrations and fluxes. INCA has been widely applied to estimate nutrient loads across diverse catchment types (McIntyre et al., 2005; Rankinen et al., 2021; Whitehead et al., 2015).

Parallel to INCA's development, the Soil and Water Assessment Tool (SWAT) emerged as a widely adopted semi-distributed model for simulating hydrology, sediment transport, and nutrient cycling across agricultural landscapes. Developed by the USDA Agricultural Research Service, SWAT divides watersheds into hydrologic response units based on land use, soil type, and slope, enabling detailed assessments of non-point source pollution and land management impacts. Its flexibility and GIS integration have made it a key tool in evaluating best management practices, forecasting water quality, and supporting policy development related to nutrient control, watershed restoration, and agricultural sustainability (Arnold et al., 1998; Gassman et al., 2007). SWAT has been extensively applied to simulate nitrogen and phosphorus dynamics across a range of agricultural catchments (Salim Aoubid and Opp, 2023; Saravanan et al., 2023).

As integrated modelling frameworks expanded in scale and scope, the Hydrological Predictions for the Environment (HYPE) model emerged as a tool for simulating water flow and nutrient transport from source to sea. Developed by the Swedish Meteorological and Hydrological Institute (SMHI), HYPE employs conceptual hydrological algorithms and operates on a daily time step, enabling consistent simulation across multiple basins. By linking hydrological processes with nutrient fate across spatial scales, HYPE has become a key instrument for evaluating policy scenarios, identifying pollution hotspots, and informing strategic planning in large-scale catchments (Capell et al., 2021; Lindström et al., 2010; Namugize et al., 2017). More specialized models have also been developed to address fine-scale biogeochemical processes with greater physical detail. For instance, Flipo et al. (2007) applied the CAWAQS (Catchment Water Quality Simulator) model to the 1200 km^2 Grand Morin basin in France, coupling hydrological, hydrogeological, and biogeochemical modules to simulate nitrate fluxes.

Most of the modelling approaches mentioned earlier rely on simplified frameworks that typically simulate nitrogen or phosphorus export at the catchment scale. Given the need for more robust simulation tools, recent advances in integrated catchment-scale modelling frameworks have expanded the range of simulated water quality parameters and improved process integration. Johnson and Zhang (2016) introduced modular nutrient simulation tools (NSMs) for hydrologic models, allowing user-defined constituents such as DOC and algae to be flexibly integrated. Crossman et al. (2021) expanded the INCA framework to develop INCA-PEco, a model capable of simulating phytoplankton growth, biological oxygen demand (BOD), and dissolved oxygen (DO) alongside phosphorus dynamics, enabling scenario testing for eutrophication and hypoxia. Similarly, Zhi et al. (2022) applied BioRT-Flux-PIHM across 236 minimally disturbed U.S. catchments (5–25,791 km^2), coupling hydrological partitioning with reactive transport of carbon (DOC), nitrogen, and phosphorus, and capturing biogeochemical hotspots across diverse environmental gradients. Khorashadi Zadeh et al. (2022) developed a multi-parameter reactive transport model for the 704 km^2 River Dender, explicitly simulating DO, BOD, NO_3^-, and NH_3 concentrations. These models reflect a shift toward holistic water quality simulation, where internal cycling, ecological feedback, and constituent interactions are explicitly represented to support restoration planning and nutrient management under changing hydrological regimes.

However, the application of these models is often constrained by data availability; particularly in regions lacking permanent monitoring infrastructure. Most of these modelling approaches rely on high-frequency measurements, which are costly and logistically demanding, making them impractical for many case studies and remote catchments (Blaen et al., 2016; Jones et al., 2012; Bin Mamoon et al., 2025; Minaudo et al., 2017; Pellerin et al., 2016). In many catchments, nutrient data are sparse, discontinuous, or reliant on surrogate variables such as turbidity and conductivity, which introduce additional uncertainty in parameter estimation (Ator, 2019; Leigh et al., 2019). As a result, high-quality data availability remains a key limitation in both load estimation and water quality modelling (Kirchner et al., 2004; Pellerin et al., 2014), and many models suffer from sparse concentration datasets that hinder calibration and reduce predictive reliability (Neilson and Chapra, 2003). More recently, Lannergård et al. (2024) explored the use of high-frequency (HF) turbidity data as a proxy for calibrating a catchment-scale phosphorus model. Their study compared four calibration datasets: traditional monthly grab samples of total phosphorus (TP) and total suspended solids (TSS), HF turbidity proxies, and combinations thereof. Using the INCA-PEco framework, they demonstrated that HF data captured greater temporal variability and produced more dynamic model outputs, particularly under episodic flow conditions. Phillips et al. (2024) investigated whether comprehensive event-based sampling is necessary to effectively constrain phosphorus simulations in the HYPE model. Their study compared high-frequency (HF) and low-frequency (LF) sampling programs across multiple Canadian catchments, focusing on total phosphorus (TP) dynamics. They found that while LF data (e.g. monthly grab samples) provided reasonable long-term averages, it failed to capture episodic spikes and hysteresis effects during storm events. In contrast, HF sampling revealed short-term concentration peaks and improved model calibration, especially in headwater catchments with rapid runoff responses.

Beyond the challenges posed by data limitations, parameter uncertainty and identifiability represent critical challenges in water quality modelling. Even when sufficient input data are available, many models are affected by parameters that are either poorly constrained or non-identifiable. This issue is particularly pronounced when simulating complex processes such as nutrient cycling, oxygen dynamics, or sediment interactions. Parameters such as reaction rates, uptake coefficients, and settling velocities frequently display the problem of non-

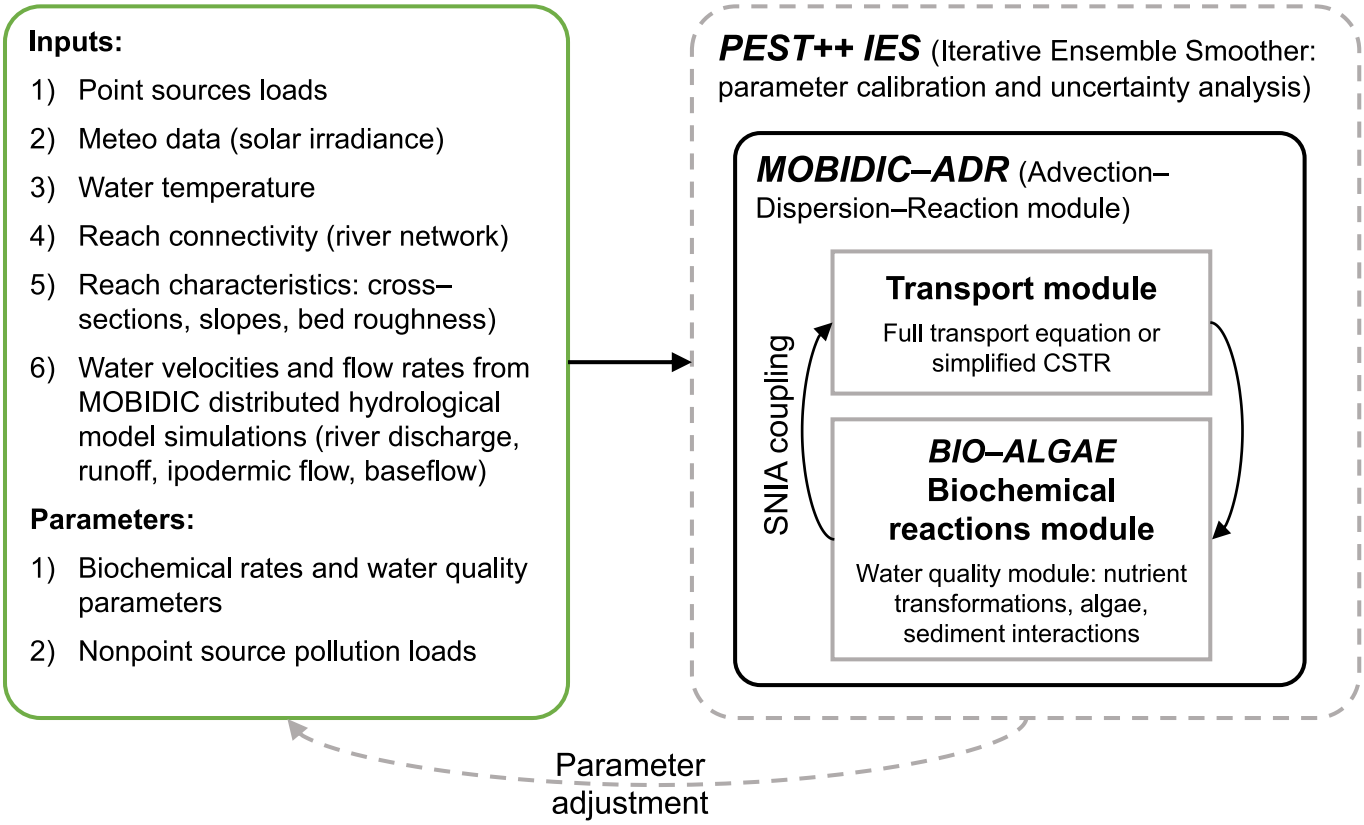


Fig. 1. Structure of the modelling framework.

uniqueness, where multiple combinations yield similar outputs, complicating calibration and reducing interpretability. To address the challenges of parameter uncertainty and identifiability, researchers have adopted a range of methodological innovations, including advanced probabilistic calibration such as DREAM(ZS) (Vrugt et al., 2009), GLUE (Beven and Binley, 1992; Spill and Gassmann, 2025), sensitivity screening such as LH-OAT (Khorashadi Zadeh et al., 2022), high-frequency surrogate data such as turbidity (Lannergård et al., 2024), multi-site and multi-criteria optimization (Udías et al., 2014), and strategic sampling design. These approaches enhance calibration robustness and improve model interpretability, particularly in data-scarce catchments. Fonseca et al. (2024) used multi-parametric sensitivity analysis and 10,000-run Monte Carlo simulations, which helped characterize constituent-specific uncertainty under limited data conditions. Ghaffar et al. (2024) advanced this further by calibrating the fully distributed mHM-Nitrate model in the Bode basin using a 300,000-iteration sequential optimization, extracting 95% uncertainty envelopes from top-performing simulations, and demonstrating that multi-site calibration improves nitrate predictions by capturing spatial heterogeneity.

Despite these advances, four persistent gaps continue to limit the effectiveness of catchment-scale nutrient modelling: (i) the scarcity, unevenness, and discontinuity of concentration data that limit calibration and validation; (ii) parameter non-identifiability and uncertainty arising in spatially heterogeneous systems; (iii) the difficulty of applying complex biogeochemical models at large spatial extents; and (iv) the lack of calibration procedures designed to perform under data-poor conditions. This study aims to address these gaps through three main objectives. First, it aims to develop a parsimonious yet process-consistent water-quality modelling framework capable of simulating nutrient cycling, oxygen dynamics, and algal processes across large catchments with limited and uneven monitoring data. Second, it seeks to find a calibration strategy for improving parameter identifiability and deriving spatially distributed nutrient loads under data-limited conditions. Third, it intends to quantify the predictive uncertainty associated with simulated nutrient concentrations and fluxes, providing a measure of model reliability that can inform water management decisions. To achieve these objectives we couple a watershed-scale water-quality model with a spatially regularized iterative ensemble smoother calibration technique implemented through the PEST++ software suite. The water quality model integrates the MOBIDIC–ADR hydrological and transport model (Masi et al., 2025) with the newly developed BIO–ALGAE reactive component, enabling the simulation of nutrient cycling, oxygen dynamics, and algal processes with a parsimonious number of parameters suitable for both reach-scale and watershed-scale applications. The calibration procedure employs spatial regularization to stabilize parameter estimation and to derive spatially distributed nutrient loads, improving identifiability even when observations are sparse or unevenly distributed. The ensemble-based framework enables efficient calibration in high-dimensional parameter spaces and provides robust uncertainty bounds without the computational burden typical of traditional global methods. Together, these elements address the mismatch between complex water-quality process representations and the practical limitations of sparse monitoring networks, enabling spatially distributed nutrient-load assessment at the catchment scale while simultaneously quantifying predictive reliability. The framework is applied to a 7990 km^2 portion of the Arno River catchment (total area 8230 km^2) to demonstrate its capability to generate spatially distributed nutrient-load estimates and associated predictive uncertainty under realistic data availability limitations.

2. Materials and methods

2.1. Modelling framework

A new module for water quality simulation, named BIO–ALGAE, was

developed in this study to simulate nutrient transformations, algae and sediment interactions. This module was integrated into an already existing modelling framework, MOBIDIC–ADR, as illustrated in Fig. 1. MOBIDIC–ADR is a module of MOBIDIC hydrological model with an Advection–Dispersion–Reaction (ADR) package to simulate surface water quality processes (Masi et al., 2025). MOBIDIC (MOdello di Bilancio Idrologico DIstribuito e Continuo) is a raster-based distributed hydrological model that simulates the water and energy balance of the hydrological cycle at the cell level (Castelli et al., 2009; Castillo et al., 2015; Yang et al., 2014a). Originally developed for catchment-scale flood forecasting and water resources management applications (Ercolani and Castelli, 2017; Yang et al., 2014b), MOBIDIC employs horizontal discretization of the basin using a regular mesh grid and resolves vertical processes across five layers: vegetation, surface storage, gravity-dominated soil compartment, capillarity-dominated soil compartment, and aquifers. The water balance is coupled with the energy balance of vegetation and soil layers through evapotranspiration, with optional snow cover calculations. The hydrographic network is modelled as a system of prismatic channels, while major reservoirs and lakes are assigned specific regulation rules and interact with the river system. Soil moisture is represented as both plant-available water and percolation to surface and groundwater bodies. Groundwater flow is computed on a separate grid, using Darcy-based formulations that vary according to aquifer confinement conditions. Bidirectional exchanges between rivers and aquifers are computed based on relative water levels.

The ADR package utilizes these hydrological states to solve the full advection–dispersion equation and includes a basic nutrient reaction module, utilizing hydrological outputs from MOBIDIC. While the original implementation of MOBIDIC–ADR focused on short-term event-based simulations in small catchments, in the present study additional modifications were introduced to enable the simulation of water quality processes at larger spatial and temporal scales. Additionally, the model interface was adapted for integration with the PEST++ software suite to enable parameter estimation, sensitivity analysis, and uncertainty quantification in decision-support contexts. Fig. 1 outlines the model architecture and required inputs and parameters. It highlights the link between the distributed hydrological model MOBIDIC, which provides key flow-related inputs (e.g., discharge, runoff, baseflow), and the MOBIDIC–ADR module. The latter includes a transport module and the newly developed BIO–ALGAE module, which simulates the biochemical reactions. These components are coupled through a sequential non-iterative (SNIA) operator-splitting procedure (Carrayrou et al., 2004) and are designed for integration with the PEST++ IES iterative ensemble smoother for robust parameter estimation and uncertainty analysis. The BIO–ALGAE module and its implementation are described in detail in the following sections.

2.2. Transport model

The transport model (MOBIDIC–ADR) simulates the transport of dissolved species (water quality constituents) along river networks, including their reaction kinetics. It can be theoretically coupled with any hydrodynamic model a posteriori, as it utilizes outputs from hydrological and hydraulic simulations, such as flow rates and velocities, as inputs for water quality modelling. In this study, MOBIDIC was used to simulate river discharges, runoff rates, subsurface flow, and baseflow for each reach of the river network.

The transport of dissolved constituents along a river is described by the advection–dispersion–reaction equation, which in one dimension may be written in the following form:

$$\frac{\partial c}{\partial t} = -v\frac{\partial c}{\partial x} + D_L\frac{\partial^2 c}{\partial x^2} + R \tag{1}$$

where c (g m^{-3}) is the concentration of a dissolved constituent, t (s) the time, x (m) the longitudinal coordinate, v (m s^{-1}) is the average longitudinal fluid velocity, D_L (m^2 s^{-1}) the longitudinal dispersion coefficient, and R (g m^{-3} s^{-1}) the source/sink term that includes both external sources and chemical and biological reactions.

The solution of this equation can be computationally expensive when applied to a large number of river reaches and multiple species, because the equation must be discretized along each river reach. In long-term simulations (spanning several years), the high computational time and complexity may limit the application of history matching and uncertainty analysis techniques. Therefore, we used a simplification where each reach is discretized into completely stirred reactors (CSTR) connected at network junctions. This tanks-in-series approach simplifies the problem by eliminating the need to compute spatial variations in concentration along individual river reaches. This approach is also employed by widely used models such as US EPA SWMM (James and Rossman, 2010) and US EPA WASP model (Wool et al., 2020). Even with this approach, longer reaches must be discretized into two or more subsections, depending on the heterogeneity of conditions within the reach (e.g., flow, point-source inputs, morphological changes). Consequently, computational time still depends on the chosen discretization. The time-dependent transport equation becomes:

$$\frac{d(Vc)}{dt} = C_{\text{in}}Q_{\text{in}} - cQ_{\text{out}} + R \tag{2}$$

where V (m^3) is the water volume within the reach at time t, C_{in} (g m^{-3}) is the concentration of any inflow to the reach, Q_{in} (m^3 s^{-1}) is the flow rate of the inflow, and Q_{out} (m^3 s^{-1}) is the flow rate leaving the reach. The equation in this form can be solved analytically (James and Rossman, 2010). The transport equation formulated as Eq. (2) readily enables the simulation of mixing processes in lakes and reservoirs by setting the volume V equal to the their mixing volume.

Both Eqs. (1) and (2) have been implemented in MOBIDIC–ADR. The full equation formulation requires dispersion coefficients as input parameters. In contrast, the CSTR approach does not explicitly account for dispersion. Although dispersion is not directly incorporated, an "apparent" dispersion emerges from numerical effects, resulting in an equivalent dispersion coefficient. This coefficient depends solely on the model's discretization and can be quantified as (Socolofsky and Jirka, 2005):

$$D_a = \frac{v\Delta x}{2} \tag{3}$$

where D_a (m^2 s^{-1}) is the "apparent" dispersion coefficient of tanks-in-series model, v (m s^{-1}) is velocity, and Δx (m) is the length of the reach.

The transport model was coupled with the following mass balance equation for constituent concentrations, which is imposed at each junction of the network:

$$c_j = \frac{\sum_i Q_{i,\text{in}}c_i + F_{\text{in}}}{\sum_k Q_{k,\text{out}}} \tag{4}$$

where c_j (g m^{-3}) is the instantaneous concentration of one species in the j-th junction, $Q_{i,\text{in}}$ (m^3 s^{-1}) is the flow rate into the j-th junction from the i-th reach, c_i (g m^{-3}) the concentration at the end on the i-th inflow reach, F_{in} (g s^{-1}) is the mass flow rate of the constituent into the junction (i.e., a point external source), and $Q_{k,\text{out}}$ (m^3 s^{-1}) is the flow rate from the j-th junction to the k-th outlet reach. The concentration c_j was taken as a Dirichlet boundary condition for the integration of the transport equation at each time step.

2.3. Nutrient simulation module

The newly developed BIO–ALGAE module, which simulates the dynamics of water quality constituents, was derived and adapted from the nutrient simulation modules (NSMs) developed by the Engineer Research and Development Center of the US Army Corps of Engineers

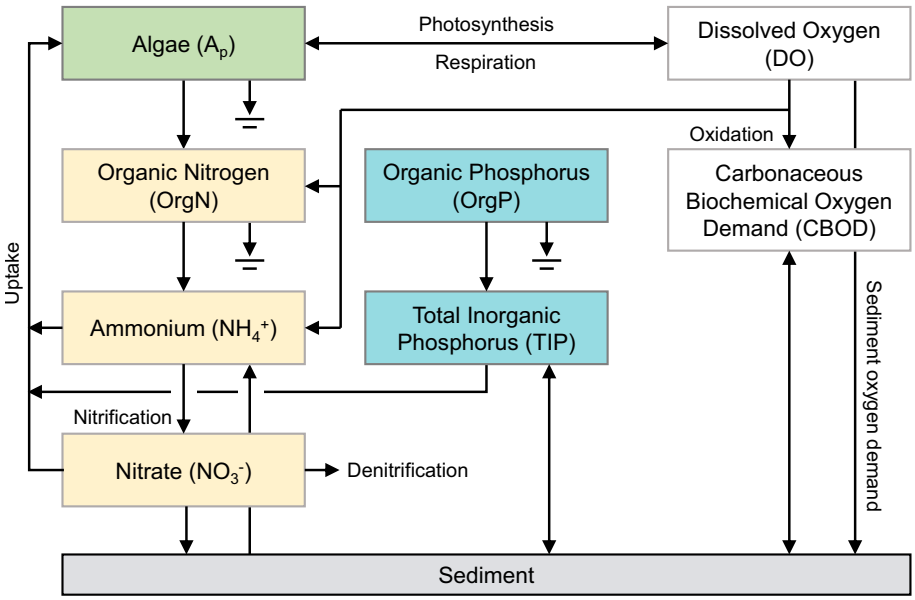


Fig. 2. State variables and transformations included in the BIO–ALGAE water quality module.

(Johnson and Zhang, 2016). The model, which simulates the biochemical reactions and transformations of the constituents is schematically represented in Fig. 2. The model includes the following 8 water quality state variables:

- Dissolved oxygen (DO)
- Carbonaceous biological oxygen demand (CBOD)
- Organic nitrogen (OrgN)
- Ammonium (NH_4^+)
- Nitrate (NO_3^-)
- Organic phosphorus (OrgP)
- Total inorganic phosphorus (TIP)
- Floating algae (A_p)

The equations describing the transformations among the components can be found in Johnson and Zhang (2016). Due to its capability to simulate both biochemical reactions and algae dynamics, the module was designated as "BIO–ALGAE". To simplify the model formulation, all algae species were aggregated into a single constituent (A_p) which is able to move freely with water.

Variables related to phosphorus species were also derived with the following relations:

$$\mathrm{DIP} = f_{\mathrm{dp}}\ \mathrm{TIP} \tag{5}$$

$$\mathrm{TOP} = \mathrm{OrgP} + r_{\mathrm{pa}}\ A_p \tag{6}$$

$$\mathrm{P_{TOT}} = \mathrm{TIP} + \mathrm{TOP} \tag{7}$$

where DIP (g m^{-3}) is the dissolved inorganic phosphorous, f_{dp} (−) is the dissolved fraction of inorganic phosphorous (0–1.0), r_{pa} is the algal phosphorous to Chlorophyll-a ratio (mg-P μg-$Chla^{-1}$), TOP (g m^{-3}) is the total organic phosphorus, and P_{TOT} (g m^{-3}) is the total phosphorous.

We calculated algae limiting factors for light, nitrogen, and phosphorous using a multiplicative formulation, where individual limiting factors (ranging from 0 to 1) are multiplied together to determine overall growth limitation. Light limitation was computed using a half-saturation function (Baly, 1935), while light attenuation followed the Beer-Lambert law (Swinehart, 1962), simulating an exponential decrease in light intensity with increasing depth.

The concentration of dissolved oxygen at saturation as a function of the temperature was calculated with the following equation (Elmore and Hayes, 1960):

$$\mathrm{DO_{sat}} = 14.65 - 0.41T + 0.008T^2 - 0.000078T^3 \tag{8}$$

where T (°C) is the water temperature.

The reaeration coefficient was estimated with the following generalized empirical formula, as a function of water velocity, turbulence, and water head (Chapra, 2008):

$$K_r = \alpha v^{\beta} H^{-\gamma} \tag{9}$$

where K_r (s^{-1}) is the reaeration coefficient, v (m s^{-1}) is the water velocity, H (m) is the water depth, and α, β, γ are three empirical coefficients.

The temperature dependence of the kinetic constants was calculated with the Arrhenius equation (Chapra, 2008):

$$K(T) = K(T_0)\theta^{(T-T_0)} \tag{10}$$

where T_0 is the reference temperature (20 °C), and θ (−) a coefficient that is specific for each parameter to be transformed.

Finally, to include the effect of dissolved oxygen concentration on degradation, nitrification and denitrification kinetics, an inhibition factor was introduced for each of the three processes, using the exponential formulation described in Chapra and Pelletier (2003).

The mathematical description of the processes shown in Fig. 2 resulted in a system of ordinary differential equations (ODEs). The numerical coupling between the transport and reaction modules was implemented using a sequential non-iterative algorithm (SNIA). This operator-splitting procedure separates the transport equations from the water quality module equations (Carrayrou et al., 2004), and solves them independently and sequentially.

2.4. Input loads parametrization

Two types of loads were modelled: the point sources, (e.g., wastewater treatment plants) and the diffuse load sources. For the point sources, the loads are calculated as mass fluxes entering into the network at the junctions (F_{in} of Eq. (4)) and calculated for each species as its concentration times the flowrate of the source flow rate.

The quantification of non-point source loads presents significantly greater challenges due to their diffuse nature and variability in space and time. One of the most commonly used parameterizations for char-

acterizing pollutant concentrations during storm events is the Event Mean Concentration (EMC) method. This approach represents the concentration of a given water quality constituent as constant over the duration of a monitored rainfall-runoff event. It is most frequently applied in urban environments (Behrouz et al., 2024). However, its use has also been extended to larger, catchment-scale studies by associating distinct concentration values with different land uses, and by distinguishing between EMC values for storm events and Dry Weather Concentration (DWC) for baseflow, acknowledging that concentrations can vary substantially between hydrological states (Bartley et al., 2012). Despite its widespread application, the EMC approach is inherently limited, as it assumes that constituent concentrations are independent of flow rate. In contrast, empirical studies have frequently observed non-linear concentration-discharge (C–Q) relationships, including hysteresis and flushing effects (Stewart et al., 2022). To address these limitations, alternative formulations have been developed to better capture solute export dynamics under both runoff and baseflow conditions (Westfall et al., 2025). Among these, the power-law model ($C = aQ^b$, where a and b are empirical constants) is widely used to characterize C–Q relationships and to infer underlying processes related to solute sources, transformation, and transport within the catchment (Speir et al., 2024). In this study we used a modified two-flow component model that separately accounts for solute contributions from quick flow and slow flow pathways, formulated as follows:

$$C_s = a_s q_s^{b_s} \quad , \quad C_q = a_q q_q^{b_q} \tag{11}$$

where C (g m^{-3}) is the solute concentration, a_s and b_s the empirical constants for the slow flow component, a_q and b_q the empirical constants for the quick flow component, q_s (m^3 s^{-1} m^{-2}) and q_q (m^3 s^{-1} m^{-2}) the specific slow and quick flows, respectively, i.e., the slow and quick flows normalized by the sub-catchment area:

$$q_s = \kappa \frac{Q_s}{A_c} \quad , \quad q_q = \kappa \frac{Q_q}{A_c} \tag{12}$$

where Q_s (m^3 s^{-1}) and Q_q (m^3 s^{-1}) are the slow (baseflow) and quick (runoff + ipodermic) flows, A_c (m^2) the contributing sub-catchment area. A dimensionless scaling factor κ (−) is introduced to adjust the magnitude of q, ensuring that the exponent b values in Eq. (11) remain within a numerically stable and interpretable range. Because Eq. (11) may yield unrealistically high concentrations at very low flow rates, we applied a capping mechanism to limit the maximum values of C_s and C_q. Specifically, the maximum allowable concentrations for both slow and quick flow components were set to ten times the maximum observed concentration of each constituent across all monitoring stations in the dataset. This factor was introduced as an empirical safeguard to prevent numerical instability and unrealistic accumulation under low-flow conditions. It represents a conservative upper bound that ensures physical plausibility without constraining model dynamics.

For unmeasured constituents, such as organic phosphorus and organic nitrogen, the reference concentration was derived from the corresponding aggregate species (i.e., total phosphorus and total nitrogen, respectively). These thresholds were applied uniformly to all sub-catchments. The total solute flux entering a river reach from diffuse sources in the associated sub-catchment is thus computed as:

$$F = C_s Q_s + C_q Q_q \tag{13}$$

where F (g s^{-1}) is the lateral mass flux into the river reach from its sub-catchment.

This approach differs slightly from previous formulations (Gunnerson, 1967; Westfall et al., 2025), which typically compute a single in-stream concentration by aggregating the two flow components. Instead, the current formulation treats the fluxes from slow and quick flow pathways separately. As a result, the application of Eq. (13) to each river reach results in a detailed and flexible model structure, although with a high degree of parameterization, as it requires reach-specific coefficients for each segment of the river network.

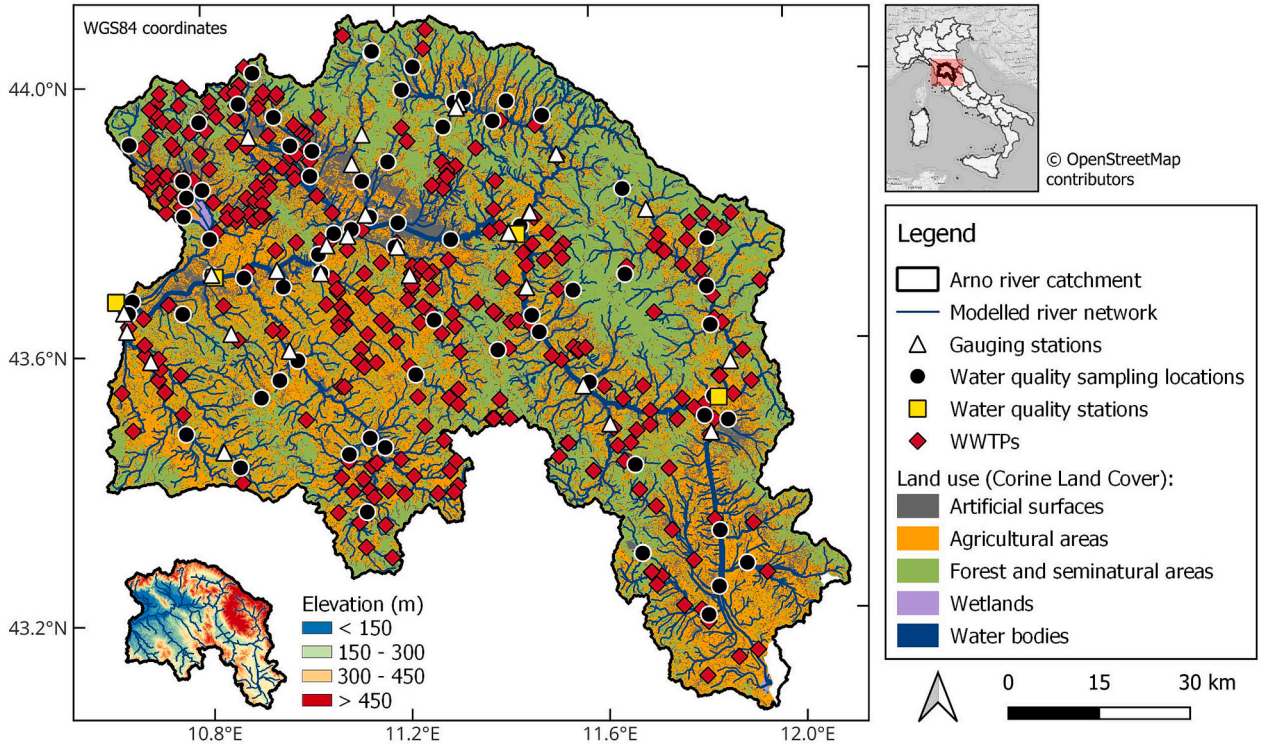


Fig. 3. Overview of the Arno River catchment: modelled river network, hydrometric stations, water quality sampling locations and permanent stations (summer period), locations of the wastewater treatment plants (WWTPs), land use, and terrain elevations.

2.5. Interface to PEST++

Model calibration and uncertainty analysis have been significantly advanced by the development of model-independent inverse modelling tools such as PEST (Doherty, 2015) and PEST++ (White et al., 2020). These tools operate non-intrusively, interfacing with a model via its existing input and output files through a file-based system using template (parameter), instruction (observation), and control files. This architecture facilitated the integration of MOBIDIC-ADR with PEST/PEST++.

In this study, the PESTPP-IES module was employed for history matching and uncertainty quantification. PESTPP-IES is an open-source, scalable, and non-intrusive iterative ensemble smoother designed for high-dimensional environmental models (White, 2018). It combines the Gauss-Levenberg-Marquardt (GLM) algorithm with a Monte Carlo approach to estimate the Jacobian matrix empirically. This empirical Jacobian significantly reduces computational demands, avoiding the need to run the model once per parameter, as in the classical GLM method. Parameter updates are performed using an ensemble of realizations, with the number of model runs is proportional to the ensemble size rather than the number of parameters.

For high-dimensional parametrization, using small ensemble sizes can lead to spurious correlations and "ensemble collapse", where the posterior parameter variance is underestimated. PESTPP-IES addresses this through several regularization techniques. First, Tikhonov regularization introduces preferred parameter values, smoothness, or spatial relationships to stabilize the solution to ill-posed inverse problems. Second, a localization matrix allows expert knowledge to be incorporated to constrain parameter-to-observation relationships (e.g., ensuring downstream processes do not influence upstream water quality). Both techniques are employed in this study and their application to the case study is discussed in later sections. Additionally, Singular Value Decomposition (SVD) is used to reduce the problem's dimensionality, focusing parameter updates on combinations that are well informed by the data.

When Tikhonov regularization is applied, the objective function minimized during calibration is given by Doherty (2015):

$$\Phi = \sum_i (w_i r_i)^2 + \mu^2 \Phi_r \tag{14}$$

where w_i is the weight of the i-th observation, r_i is the residual (difference between a measurement and a model output), Φ_r is the regularization objective function, and μ^2 is a weighting factor that controls the strength of the regularization.

2.6. Case study description

The model framework was applied to the Arno river catchment (Fig. 3), located in central Italy. The Arno River originates from Mount Falterona at 1358 m elevation and flows 241 km to the Ligurian Sea, draining a basin of about 8230 km^2. The river's slopes vary from 20–14% in the Apennine section to 0.3% near the mouth. The basin consists predominantly of impermeable formations (95% of the surface) including clays, marls, argillaceous schists, and compact sandstones.

Precipitation shows a bimodal distribution with abundant rainfall from January–May and October–December, separated by a dry summer period. Annual precipitation varies significantly with topography, ranging from 700–900 mm in lower areas to over 2000 mm on Apennines. Due to the predominantly impermeable terrain, streamflow closely follows precipitation patterns with irregular distribution. The basin produces approximately 3 billion cubic meters of annual runoff with an average discharge of 100 $m^3\ s^{-1}$ at the most downstream gauge (San Giovanni alla Vena), characterized by winter-spring maxima and summer minima.

The study area considered in this work is a subset of the larger catchment, defined by the outlet section located at Pontedera. The catchment area analysed covers 7990 km^2. Based on the 2019 Corine Land Cover data, land use within the study area consists of 10.7% artificial surfaces, 39.3% agricultural land, 49.3% forest and semi-natural areas, 0.2% wetlands, and 0.6% water bodies.

The hydrometric network within the study area includes 29 water level gauging stations, where flow rates are derived from site-specific stage–discharge relationships. In addition, four water quality monitoring stations are installed at fixed locations and operate seasonally from June to October, when low-flow conditions make water quality monitoring most critical. During this period, the stations record continuous measurements at hourly intervals for the following parameters: pH, redox potential, electrical conductivity, water temperature, dissolved oxygen, and turbidity. Data from these stations are available from 2007 to present.

Infrequent water quality sampling has been conducted by ARPAT – Agenzia Regionale per la Protezione Ambientale della Toscana[1] since 1997 and intensified following the implementation of the EU Water Framework Directive (2000/60/EC). For this study, we used data from the period 2011–2020. Approximately 130 sampling locations are distributed across the catchment. From these, 70 stations were selected based on data availability (at least five samples across the full ten-year period 2011–2020) and data quality (excluding values marked as "less than" or "greater than"). The water quality parameters used in this modelling study include biochemical oxygen demand (BOD), ammonium (NH_4^+), nitrates (NO_3^-), and total phosphorous (P_{TOT}). To relate the measured parameters to BIO–ALGAE state variables, the BOD was converted to CBOD using the ratio CBOD/BOD = 0.9 (USEPA, 1995) and Eq. (7) was used to compute the total phosphorous concentrations.

Starting from a total of 12,650 measurements across BOD, DO, NO_3^-, and NH_4^+, and P_{TOT}, the dataset was reduced to 8,151 valid measurements applying the selection criteria mentioned above. The final dataset contained an average of approximately 860 data points per year across all stations, with the exception of 2020, which had 410 data points. Samplings were relatively uniform from February to November, with monthly totals (summed over period 2011–2020) ranging between 550 and 950 data points. January and December were less represented, with 221 and 334 total data points, respectively.

2.7. Case study: model setup

The simulation period spanned approximately ten years, from 2011 to 2020. The modelled river network comprised a total of 3,621 reaches. Each river was subdivided in multiple reaches based on network connectivity, spatial variability in morphology, flow conditions. The discretization of the hydrological model was chosen to balance computational efficiency with an adequate representation of the network and to achieve reliable model performance on simulated discharges. This discretization proved equally suitable for the water quality model, where the primary constraint of having adequate representation of point-source inputs was satisfied. Discharge data for the entire network during the simulation period were obtained from hydrological simulations using the MOBIDIC model. These simulations, conducted as part of a separate study, were officially adopted by the Northern Apennine District Basin Authority (Autorità di Bacino Distrettuale dell'Appennino Settentrionale) to determine water balances across their jurisdiction. Relevant data are publicly available via the Authority's web portal.[2]

[1] Regional Environmental Protection Agency (Agenzia regionale per la protezione ambientale della Toscana, ARPAT) web portal (https://www.arpat.toscana.it/).

[2] Autorità di bacino distrettuale dell'Appennino Settentrionale data platform (https://www.appenninosettentrionale.it).

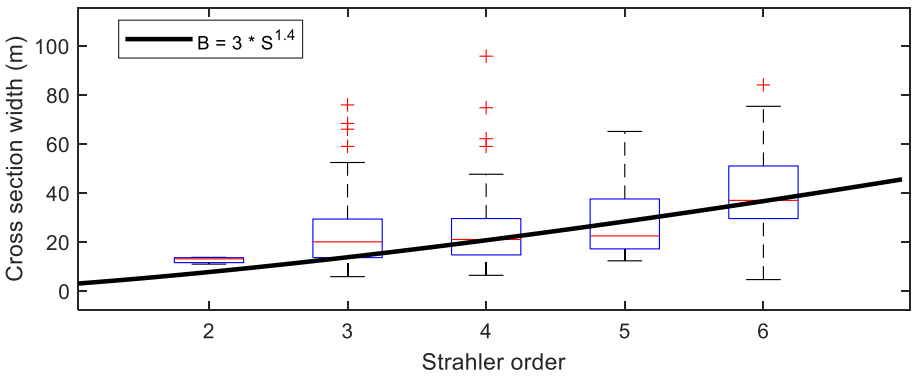


Fig. 4. Box plot of Strahler order vs cross-section widths of the river reaches and fitted power-law relation.

2.7.1. Hydrological data input

The output time series from MOBIDIC, used as input for the present study, included the following variables at a daily resolution for each reach: surface discharge, lateral inflow from slow flow (baseflow), lateral inflow from quick flow. These time series were linearly interpolated to match the simulation time step of 2 h. The transport model and the BIO–ALGAE module were coupled at the same 2-h interval. Within each coupling interval, the BIO–ALGAE module employed adaptive substepping to ensure numerical convergence of the biogeochemical reactions. The MOBIDIC simulations accounted for all known water abstractions and releases, including treated effluent volumes from wastewater treatment plants. Model performance, evaluated against observed discharge data at available gauging stations, showed a mean modified Kling-Gupta Efficiency (KGE) (Kling et al., 2012) of 0.74. The 25th and 75th percentiles were 0.71 and 0.82, respectively. The lowest three performing stations had KGEs of 0.45, 0.63, and 0.64, while the highest performing stations reached values of 0.82, 0.83, and 0.87. Nash-Sutcliffe efficiency (NSE) was also calculated: the 25th and 75th percentiles were 0.63 and 0.76. The lowest three performing stations had NSEs of −0.27, 0.51, and 0.57, while the highest performing stations reached values of 0.76, 0.77, and 0.78. The average percent bias (PBIAS) between the simulations and the observations for all stations is 11.9% (average of absolute values). The lowest three performing stations had PBIAS of 24.7%, 21.7%, and 19.7%, while the highest performing stations had −0.44%, 1.11%, and −1.86%, respectively. The total number of observations was 71,543, among 14 gauging stations.

For MOBIDIC–ADR implementation, water velocities (daily averages) in each reach were estimated using Manning's equation for open channel flow, applied to the discharge time series from MOBIDIC. A constant Manning's coefficient of 3.0×10^{-2} s $m^{-1/3}$ was used, and a rectangular cross-sectional geometry was assumed. Channel widths were derived either from river cross section data provided by the Northern Apennine District Basin Authority (Autorità di Bacino Distrettuale dell'Appennino Settentrionale) or, where unavailable, estimated using an empirical power-law relationship based on the Strahler order of each reach (Fig. 4). A total of 3,555 river cross sections were available, covering the majority of reaches in the lower catchment. The equivalent rectangular widths were automatically calculated using the algorithm developed by Masi et al. (2025).

2.7.2. Meteorological data input

For the temperature datasets, only air temperature time series were fully available. These were obtained by spatially interpolating temperature data from the Tuscany Regional Functional Centre ("Centro Funzionale Regionale", CFR[3]) using Thiessen polygons across the study area. This approach allowed assigning a continuous air temperature time series to each reach of the river network. In contrast, water temperature data were limited to spot measurements at the sampling stations and continuous records from the four permanent monitoring stations. To derive continuous water temperature time series for all river reaches, a linear regression model was developed to relate water temperature to air temperature. The relationship was fitted using a dataset of 5,120 paired air–water temperature observations, resulting in the equation $T_w = 0.71\ T_a + 1.94$, where T_w (°C) is the water temperature and T_a (°C) is the air temperature, with $R^2 = 0.792$, and RMSE = 2.6 °C, as shown in Fig. 5.

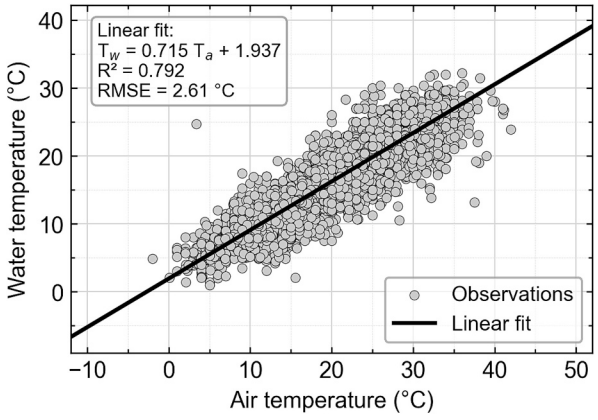


Fig. 5. Linear regression model relating water temperature to air temperature.

Solar radiation, available at daily interval as raster grids, was computed locally by averaging the values of all raster cells intersecting each reach, and subsequently resampled to the simulation time step using the following equation (Marsili-Libelli, 2016):

$$I(t) = \begin{cases} \frac{J_n}{f}\left[1 + \cos\left((t-0.5)\frac{2\pi}{f}\right)\right] & \text{for } 0.5 - \frac{f}{2} \le t \le 0.5 + \frac{f}{2} \\ I(t) = 0 & \text{otherwise} \end{cases} \quad (15)$$

where I (W m^{-2}) is the instantaneous solar radiation at time t, t is the time of the day normalized between 0 and 1, J_n (W m^{-2}) is the average daily net solar radiation, and f is the photoperiod, defined as:

$$f = \frac{2\arccos(-\tan(\phi)\cdot\tan(\delta))}{2\pi} \quad (16)$$

where ϕ (rad) is the latitude, and δ (rad) is the solar declination angle. Riparian shading was not explicitly represented in the model.

2.7.3. BIO–ALGAE model parameters

The parameters ranges of the BIO–ALGAE model used in generating the prior ensemble (Table 1) were adapted from Johnson and Zhang

[3] Centro Funzionale Regionale (CFR). https://www.cfr.toscana.it/.

Table 1
The parameters ranges of the BIO–ALGAE model for generating the prior ensemble and type of transform used in PESTPP-IES.

Parameter	Description	Lower bound	Upper bound	Transform
K_a	Nitrification rate (day^{-1})	0.1	2	log
K_{ain}	Nitrification inhibition coefficient ($m^3\ g^{-1}$)	0.1	0.95	log
K_d	Denitrification rate (day^{-1})	0.1	0.8	none
K_{din}	Denitrification inhibition coefficient ($m^3\ g^{-1}$)	0.1	0.95	log
K_b	CBOD decay coefficient (day^{-1})	0.02	3.4	log
K_{bin}	CBOD inhibition coefficient ($m^3\ g^{-1}$)	0.1	0.95	log
α	Reaeration coefficient (day^{-1})	3.5	5.5	log
β	Reaeration coefficient (day^{-1})	0.5	1	log
γ	Reaeration coefficient (day^{-1})	1.5	1.85	log
μ_{mxp}	Maximum algal growth rate (day^{-1})	0.8	2	none
K_{rp}	Algal respiration rate (day^{-1})	0.1	0.6	none
K_{dp}	Algal mortality rate (day^{-1})	0.02	0.3	log
AW_n	Algal nitrogen stoichiometry (mg N)	7.2	7.2	fixed
AW_a	Algal Chla stoichiometry (μg Chla)	1000	1000	fixed
AW_c	Algal carbon stoichiometry (mg C)	40	40	fixed
AW_p	Algal phosphorus stoichiometry (mg P)	1	1	fixed
P_N	NH_4^+ preference factor for algal growth	0.5	0.5	fixed
K_{sN}	Half-saturation N limiting constant ($gN\ m^{-3}$)	0.02	4	log
v_{sa}	Algal sedimentation velocity ($m\ day^{-1}$)	0.01	0.4	log
v_{son}	Organic N settling velocity ($m\ day^{-1}$)	0.005	2	log
K_{on}	Decay rate of organic N to NH_4^+ (day^{-1})	0.05	1	log
λ	Light attenuation coefficient (m^{-1})	0.1	5	log
K_L	Light limiting constant for the algal growth	4	40	log
K_{op}	Decay rate of organic P to DIP (day^{-1})	0.01	0.7	log
v_{sop}	Organic P settling velocity ($m\ day^{-1}$)	0.005	0.1	log
v_{sp}	Phosphorous solids settling velocity ($m\ day^{-1}$)	0.01	1	log
f_{pp}	Particulate fractions of inorganic P (0–1)	0.01	0.8	log
K_{sP}	Half-saturation P limiting constant for algal growth (mg P/L)	0.001	0.05	log

(2016).

The parameters of the model related to the sediment oxygen demand, as well as the sedimentation and release rates of CBOD, nitrate (NO_3^-) sedimentation, ammonium (NH_4^+) sedimentation, and inorganic phosphorus exchanges, are not reported in the Table 1, as their corresponding kinetic rates were set to zero. In this configuration, sediment–water interactions were largely neglected, with the exception of sedimentation processes for organic phosphorus (v_{sop}), inorganic phosphorus solids (v_{sp}), algae (v_{sa}), and organic nitrogen (v_{son}). Given this simplified formulation, in which only settling processes were retained, the sediment compartment was excluded as an explicit state variable. This assumption was necessary because sediment-related state variables could not be reliably constrained in the absence of sediment observations. Calibration of the parameters was performed using the iterative ensemble smoother (PESTPP-IES). Depending on their ranges, some parameters were transformed logarithmically, while others were kept in their natural scale. Parameters not adjusted during calibration are indicated as "fixed" in Table 1.

Table 2
Assumed effluent concentrations from wastewater treatment plants (WWTPs) used in the model, based on regulatory limits and typical treatment performance.

Parameter	C (mg/L)	Fraction of regulatory limit
CBOD	13.5	60%
NH_4^+	0.9	60%
NO_3^-	7.2	60%
OrgN	0.9	60%
OrgP	0.42	70%
TIP	0.98	70%
DO	2	–

2.7.4. Point and non-point sources

The point sources considered in the model were effluent discharges from wastewater treatment plants (WWTPs) into the river network. A total of 1,208 WWTPs were included in the model. Average discharge volumes for each WWTP were obtained from the Regional Environmental Protection Agency (ARPAT). At the outlet of the studied part of the catchment, corresponding to the location of Pontedera gauging station (TOS01005181), the total average discharge from WWTPs was estimated at 4.7 $m^3\ s^{-1}$, based on the mean annual treated volumes of all plants within the catchment. Monthly averaged river discharges exhibit a marked seasonal pattern, ranging from high values in winter and early spring (e.g. 161–170 $m^3\ s^{-1}$ in February–March) to pronounced low-flow conditions in summer. The lowest monthly discharges occur in August (10.4 $m^3\ s^{-1}$) and September (12.5 $m^3\ s^{-1}$), during which WWTP effluents constitute a significant share of the total flow at the outlet. This highlights the increased relative importance of point-source inputs under low-flow conditions.

While discharge data were available for all WWTPs, effluent concentration data were available only for a limited number of facilities. Therefore, to define the concentrations of pollutants in the model, reference was made to the emission limits set by the Italian Legislative Decree No. 152 of April 3, 2006. To translate these regulatory limits into model inputs, several assumptions were made. The regulatory emission limits considered were 25 mg/L for biochemical oxygen demand (BOD), 15 mg/L for total nitrogen (TN), and 2 mg/L for total phosphorus P_{TOT}. By analysing the monitoring datasets of WWTP effluent samples collected by the Regional Environmental Protection Agency and made available through their public data portal,[4] we derived the basis for the following assumptions: the data indicate that total nitrogen in WWTP effluents is, on average, partitioned as 80% nitrate (NO_3^-), 10% ammonium (NH_4^+), and 10% organic nitrogen. Likewise, total phosphorus was found to consist of approximately 70% inorganic and 30% organic phosphorus. To reflect the typical operational performance of WWTPs seen in the datasets, effluent concentrations were parameterized as fractions of the regulatory limits: 60% for BOD and nitrogen species, and 70% for phosphorus species. For dissolved oxygen, a concentration of 2 mg/L was assumed, representing typical conditions in effluents from activated sludge processes, which are widely used in the region. The resulting assumed effluent concentrations are summarised in Table 2.

Regarding the distributed input loads, the initial parametrization described in §2.4 would have resulted in a total of 131,388 parameters, calculated as follows: 3,621 sub-catchments × 2 parameters for quick flow × 2 parameters for slow flow × 7 state variables (excluding dissolved oxygen, which is not an input load). This level of detail would

[4] ARPAT (Regional Environmental Protection Agency) data portal: https://sira.arpat.toscana.it.

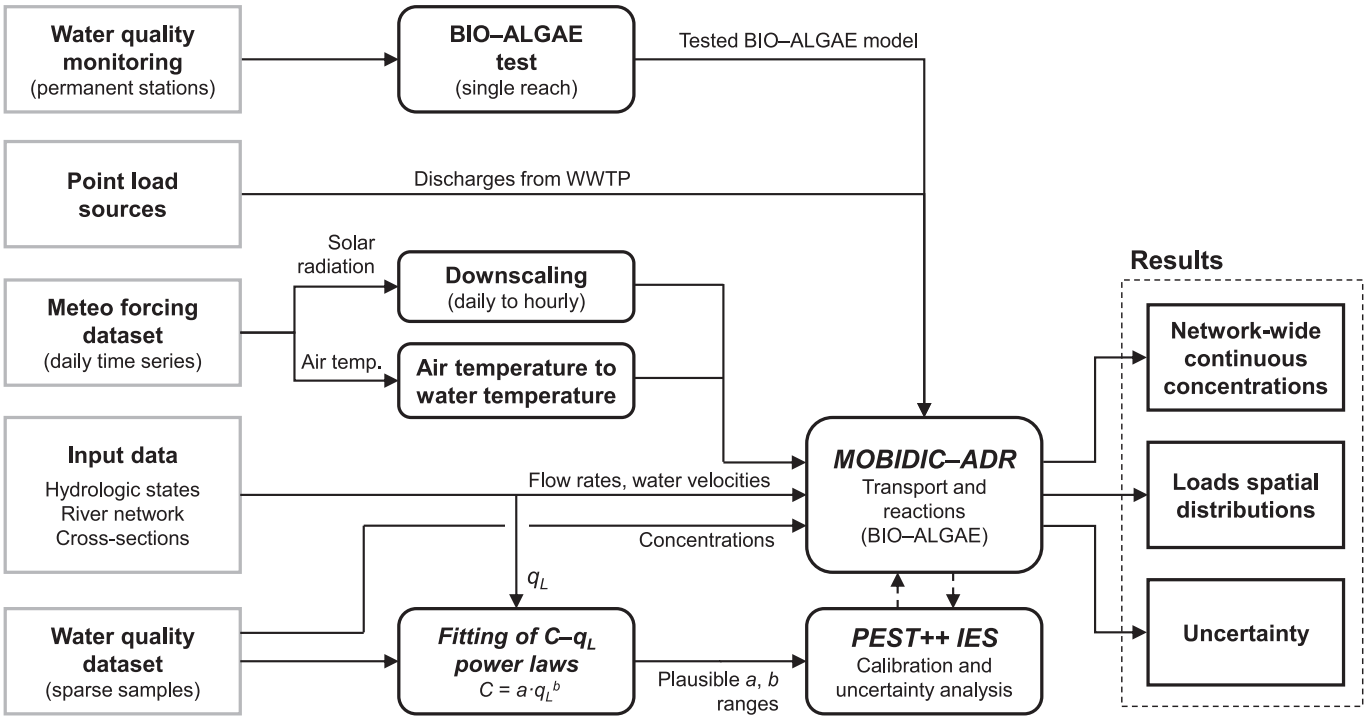


Fig. 6. Workflow of the model application to the case study.

have led to excessive parameterization, particularly in sub-catchments with very small drainage areas. To reduce model complexity and improve parameter identifiability, sub-catchments were aggregated based on their area and hydrologic connectivity, resulting in 198 groups with more uniform spatial distribution. This reduced the number of distributed input load parameters to 5,544 ($198 \times 2 \times 2 \times 7$). The final problem dimensions were as follows: total number of parameters = 5,572, of which 5,567 were adjustable; number of observations = 8,151, with 6,503 used for calibration and 1,648 (approximately 20%) retained for validation.

Three major reservoirs were included in the model with fixed storage volumes: Bilancino (45 Mm^3), La Penna (5 Mm^3), and Levane (2 Mm^3).

For the PEST-IES run, the SVD mode was enabled, with a maximum of five iterations. The ensemble size was set to 300 realizations. A localization matrix was employed to limit parameter-observation updates to hydrologically relevant interactions. This matrix was constructed by reflecting the downstream connectivity of the river network, under the assumption that each parameter only influences downstream processes.

Tikhonov regularization was applied by supplying a prior covariance matrix, which introduces the assumed correlations among parameters into the inversion process. In particular, spatial correlation was imposed on the distributed load parameters based on their geographic proximity. The spatial coordinates were assigned using the centroids of the aggregated sub-catchments associated with each parameter. The prior covariance matrix was derived from an exponential variogram model, assuming a unit sill, isotropic behaviour (i.e. no directional dependence), and a correlation range of 20 km. This variogram defines the expected covariance between any two parameters as a decreasing exponential function of the distance between their centroids. The resulting covariance matrix captures the spatial smoothness of parameter values across the catchment. The regularization weight μ^2 was set to 0.25, balancing the influence of prior information against the observational data in the objective function.

2.8. Workflow

The overall workflow is illustrated in Fig. 6, which summarises the progression from input datasets (left), through the modelling components (centre), to the resulting outputs (right). Although the permanent monitoring stations measured only dissolved oxygen, and only during the summer period, their time series provided a basis for testing the BIO–ALGAE module on a representative single reach. Among the four stations, “Nave di Rosano” was selected because it offered the most complete dataset and was the only one that included solar radiation measurements. This testing step ensures that algal dynamics and oxygen-related processes are reliably represented before extending the model to the full catchment. Meteorological forcing, available as spatially distributed and time-dependent daily air temperature and solar radiation raster grids, is pre-processed by downscaling solar radiation to hourly resolution (§2.7) and converting air temperature to water temperature. Hydrologic states, previously simulated with the MOBIDIC hydrological model, provide flow rates, water velocities, and the partitioning of quick and slow flows required for both transport and load-flux computations. Constituent concentrations from the sparse water quality samples both support calibration through history matching against simulated concentrations, and are used to fit empirical C–q_L power-law relationships, which define plausible a and b ranges to constrain parameter space during calibration. The MOBIDIC–ADR model, which couples hydrodynamic transport with the BIO–ALGAE reaction module, then simulates constituent dynamics across the entire river network. Calibration and uncertainty quantification are performed using PEST++ IES, which adjusts model parameters and evaluates predictive uncertainty. The final outputs include network-wide continuous concentration estimates, spatially resolved load distributions, and uncertainty assessment.

3. Results and discussion

3.1. Nutrient model test

To validate the developed BIO–ALGAE reactive component of

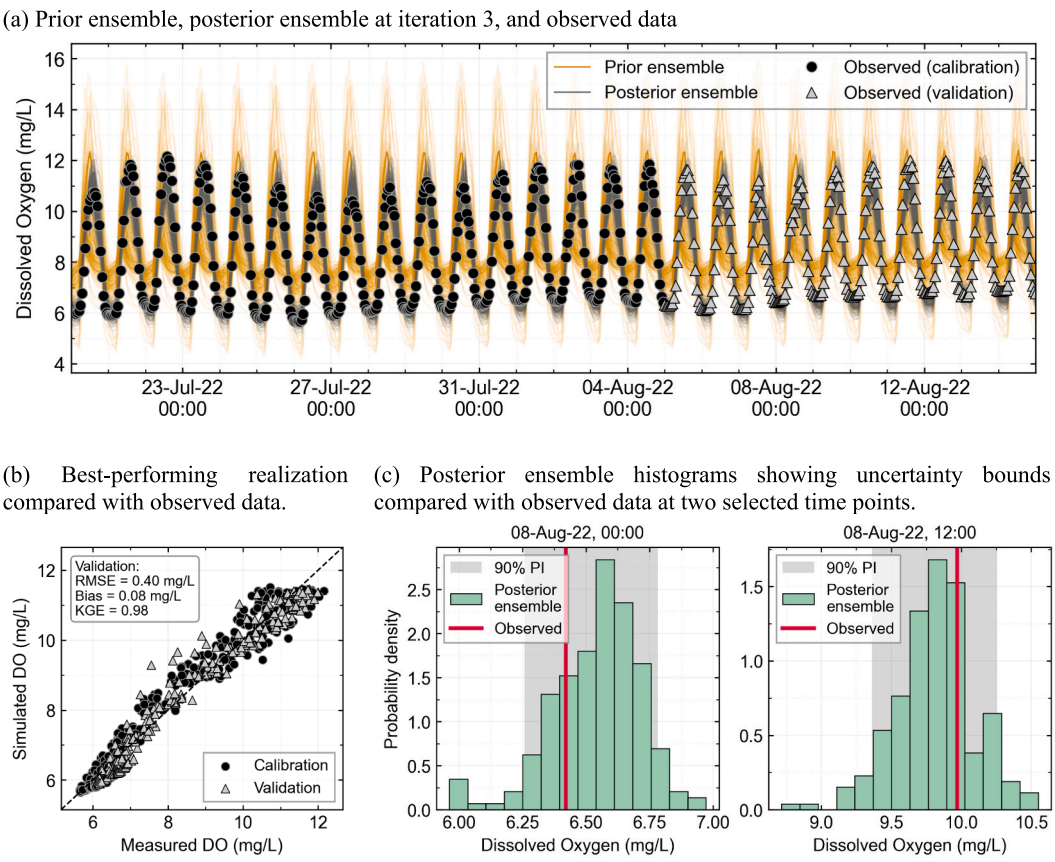


Fig. 7. Results of the BIO–ALGAE model calibration, validation and uncertainty analysis using data measured at "Rosano" station.

MOBIDIC–ADR, the model was applied to simulate a single river reach in the mid part of the Arno River catchment, located at 43.77°N, 11.42°E (WGS84), near the town of Pontassieve. This reach includes one of the four permanent summer monitoring stations (station name: "Rosano"). The objective was to test the reactive module under simplified conditions by assuming steady-state transport. Analysis of monitoring data revealed that from 10 July to 15 August 2022, discharge at the selected reach remained relatively stable i.e., within 4.87 ± 0.18 $m^3\ s^{-1}$, due to the absence of precipitation and constant upstream reservoir releases (Bilancino, Levane, and La Penna). The average water depth at the gauging station was about 0.9 m. The bankfull width of the section is about 70 m, which reduced to about 35 m in the analysed period (dry season). Riparian vegetation along this reach provides limited shading, as the relatively wide channel is largely exposed compared to tree height and density.

History matching of dissolved oxygen (DO) was performed using PESTPP-IES against observations from the monitoring station. During this period, DO concentrations showed strong diurnal variability, with daily peaks typically occurring between 14:00 and 15:00, indicating a significant algal presence. Observed DO values ranged from 5.65 to 12.16 mg/L, while water temperatures varied between 22.4 °C and 30.0 °C.

Hourly time series of water temperature and solar radiation from the monitoring station were used as model inputs, and the simulation time step was set to 1 h. Upstream boundary concentrations were assumed constant and set as follows: $NH_4^+ = 0.4$ mg/L, $NO_3^- = 0.2$ mg/L, OrgN = 0.6 mg/L, OrgP = 0.2 mg/L, TIP = 0.1 mg/L, CBOD = 0.7 mg/L. In the absence of upstream measurements for the selected reach, these values were taken from the nearest sampling point (MAS-503 – "Fiume Arno Anconella"), located approximately 14 km downstream. The upstream boundary condition was defined as a mass flux derived from these concentrations and the reach's flow rate. Since no measurements of algal concentration were available, the initial algal biomass (A_p) was set to 400 μg Chl-a L^{-1}, based on ranges reported by Campolo et al. (2002), while the upstream algal flux was treated as a calibration parameter.

The PESTPP-IES configuration included an ensemble of 300 realizations with SVD mode enabled. The inverse problem involved 33 adjustable parameters and a total of 625 observations, of which 384 (approximately 38%) were reserved for model validation.

The results of the history matching are presented in Fig. 7.

Panel (a) shows the observed and simulated time series. The prior ensemble, generated using parameter distributions based on the bounds listed in Table 1 (derived from literature values), already displayed a reasonable level of agreement with the observed data. However, by iteration 3 of the parameter adjustment process, a substantial reduction in parameter uncertainty was achieved, as reflected in the posterior ensemble. Panel (b) displays a scatter plot of observed versus simulated values for the best-performing realization (i.e., the realization with the lowest objective function value). The model demonstrates strong agreement with observations, achieving a root mean square error (RMSE) of 0.4 mg/L and a Kling–Gupta Efficiency (KGE) of 0.98. Panel (c) illustrates that this uncertainty reduction did not compromise the model's predictive capability. Observed values during the validation period fall within the 90% prediction interval, indicating that the calibrated parameters are both identifiable and robust for forecasting. The successful validation exercise provided confidence in the reactive module's formulation before its application to the entire catchment.

3.2. *Case study: water quality data analysis*

Despite the relatively large number of samples and analytical determinations, the dataset, comprising 8151 measurements, exhibited two main limitations:

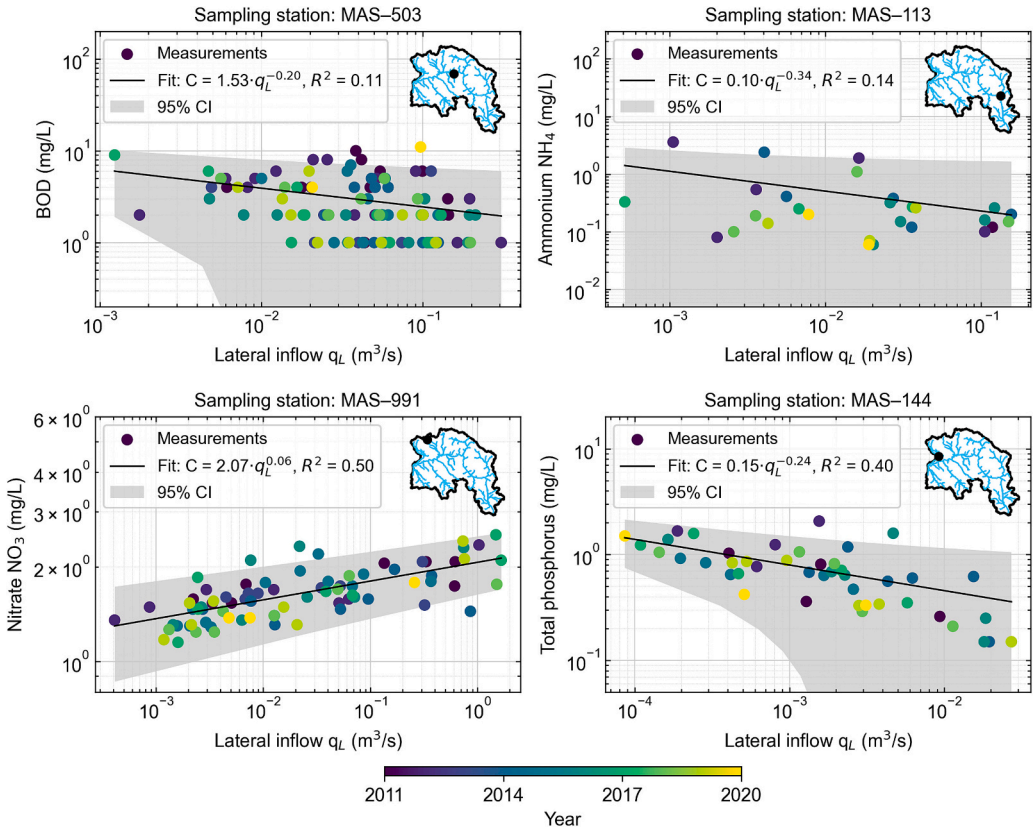


Fig. 8. Results of fitting C–q_L relationships at selected sampling stations.

- Temporal continuity: stations with more complete datasets typically had an average sampling frequency of approximately one sample per month, while others had sparser records.
- Spatial and temporal coverage: data coverage was generally more complete in the lower sections of the catchment.

Estimating pollutant loads at a river section requires not only sufficient concentration data but also accurate discharge measurements. When time series are highly discontinuous and no gauging station is available nearby, the uncertainty of load estimates increases considerably. Under such conditions, directly fitting the parameters of a concentration–discharge (C–Q) relationship becomes particularly challenging. To address this, we attempted to calibrate the load parametrization equations described in §2.4 using simulated flow rates from the MOBIDIC hydrological model for each reach of the catchment. The empirical parameters *a* and *b* were fitted individually for each

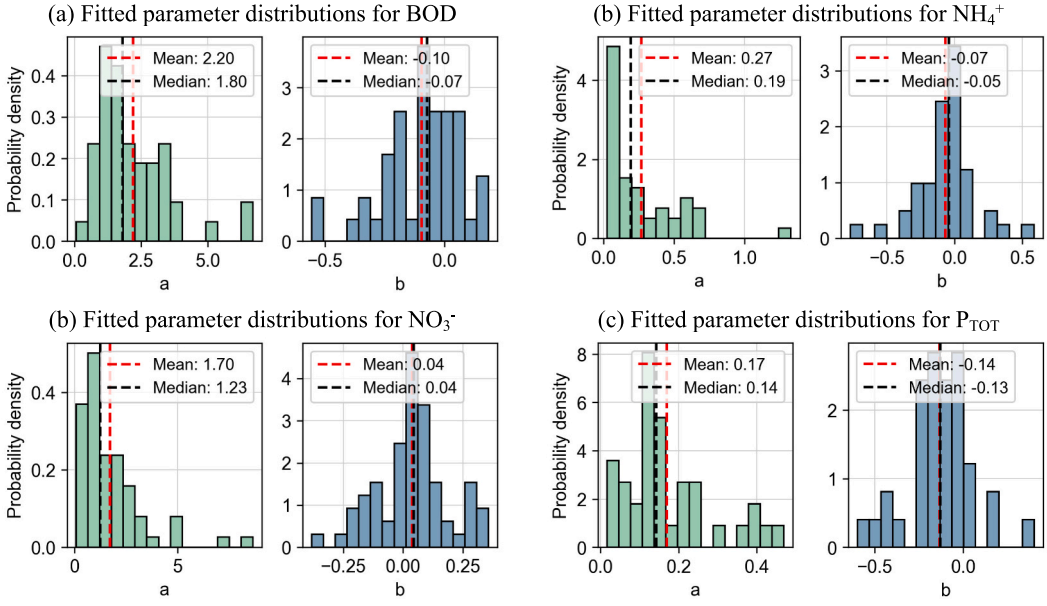


Fig. 9. Distributions of the fitted *a* and *b* parameters of the C–Q relationships.

Table 3
Statistical descriptors of the posterior parameter distributions at the second iteration of PESTPP-IES. Parameters for which only the mean value is reported were held fixed during the inversion process.

Par.	Description	Mean	Median	5th pct.	95th pct.
K_a	Nitrification rate (day^{-1})	0.36	0.37	0.30	0.42
K_{ain}	Nitrification inhibition coefficient ($m^3\ g^{-1}$)	0.54	0.54	0.38	0.72
K_d	Denitrification rate (day^{-1})	0.34	0.34	0.29	0.42
K_{din}	Denitrification inhibition coefficient ($m^3\ g^{-1}$)	0.17	0.17	0.14	0.20
K_b	CBOD decay coefficient (day^{-1})	0.77	0.77	0.71	0.83
K_{bin}	CBOD inhibition coefficient ($m^3\ g^{-1}$)	0.51	0.51	0.38	0.66
α	Reaeration coefficient (day^{-1})	5.02	5.04	4.78	5.21
β	Reaeration coefficient (day^{-1})	0.64	0.64	0.58	0.72
γ	Reaeration coefficient (day^{-1})	1.85	1.85	1.85	1.85
μ_{mxp}	Maximum algal growth rate (day^{-1})	1.01	1.01	0.90	1.15
K_{rp}	Algal respiration rate (day^{-1})	0.23	0.23	0.18	0.29
K_{dp}	Algal mortality rate (day^{-1})	0.15	0.15	0.11	0.21
AW_n	Algal nitrogen stoichiometry (mg N)	7.2	–	–	–
AW_a	Algal Chla stoichiometry (μg Chla)	1000	–	–	–
AW_c	Algal carbon stoichiometry (mg C)	40	–	–	–
AW_p	Algal phosphorus stoichiometry (mg P)	1	–	–	–
P_N	NH_4^+ preference factor for algal growth	0.5	–	–	–
K_{sN}	Half-saturation N limiting constant ($gN\ m^{-3}$)	0.17	0.15	0.09	0.30
v_{sa}	Algal sedimentation velocity ($m\ day^{-1}$)	0.1996	0.2000	0.1987	0.2000
v_{son}	Organic N settling velocity ($m\ day^{-1}$)	0.036	0.032	0.015	0.075
K_{on}	Decay rate of organic N to NH_4^+ (day^{-1})	0.072	0.070	0.055	0.102
λ	Light attenuation coefficient (m^{-1})	3.79	4.00	2.84	4.00
K_L	Light limiting constant for the algal growth	12.4	12.3	9.2	16.1
K_{op}	Decay rate of organic P to DIP (day^{-1})	0.10	0.09	0.05	0.18
v_{sop}	Organic P settling velocity ($m\ day^{-1}$)	0.013	0.013	0.009	0.019
v_{sp}	Phosphorous solids settling velocity ($m\ day^{-1}$)	0.101	0.095	0.060	0.160
f_{pp}	Particulate fractions of inorganic P (0–1)	0.57	0.60	0.45	0.60
K_{sP}	Half-saturation P limiting constant for algal growth (mg P/L)	0.0012	0.0011	0.0010	0.0018

monitoring station based on available data. Lateral inflow was chosen for the fitting, as it generally provided a better predictive performance and yielded more reliable parameter ranges for the subsequent modelling exercise, where baseflow and quick flow were employed (lateral inflow being their sum). Results for selected stations are presented in Fig. 8, focusing on stations MAS-503 (Arno river), MAS-113 (Canale Maestro della Chiana), MAS-991 (Torrente Vincio), and MAS-144 (Canale Usciana), which represent diverse catchment settings in terms of river geometry, discharge magnitude, land use, and topography. The exponent *b* provides insight into the dominant pollutant generation mechanism (Speir et al., 2024): negative values suggest dilution-dominated behaviour, whereas positive values indicate flushing dynamics. For BOD at MAS-503, the fitted relationship revealed a clearly negative *b*, although with high uncertainty and a low coefficient of determination ($R^2 = 0.11$), indicating a likely dilution pattern. BOD measurements at this and most other stations were also strongly quantized, with concentrations often reported in truncated integer values (e.g., 1 mg/L, 2 mg/L), especially at low levels, which added further uncertainty for this constituent.

Finally, examining the temporal distribution of the fitted parameters (indicated by the colour scale from blue, 2011, to yellow, 2020) revealed no systematic shift in the C– q_L relationship over time. This stability is a critical observation, as the modelling framework relies on the assumption of stationary pollutant export: while pollutant export depends on flow magnitude, the underlying export mechanisms are assumed not to vary over time.

The distributions of the fitted *a* and *b* parameters across all available measurement stations are presented in Fig. 9, providing a broader perspective on pollutant export dynamics across the catchment. Distinct patterns emerge for the different constituents. For BOD and NO_3^-, the parameter *a* exhibits wider distributions, indicative of strong spatial heterogeneity in baseline concentrations and site-specific export behaviour. In contrast, NH_4^+ and P_{TOT} show narrower and more compact ranges, indicating more uniform conditions across the catchment. The exponent *b* is approximately centred near zero in all cases, yet with constituent-specific tendencies: BOD and P_{TOT} lean toward negative values, suggesting dominating dilution behaviour, whereas NH_4^+ and NO_3^- spread around neutral values, reflecting weak or mixed flow-concentration responses. Overall, caution should be exercised when interpreting these patterns, as measurement non-uniformity, q_L estimate errors, and measurement noise may partly mask the underlying C– q_L relationships.

These fitted distributions were subsequently used to define physically plausible parameter bounds in the PESTPP-IES calibration setup. Though the results highlight the wide range of uncertainty in these estimates, this exercise provided a basis for deriving plausible distributions of the *a* and *b* parameters, which were then used to define the parameter bounds for the full-catchment application of the MOBIDIC–ADR model.

3.3. Case study: implementation and history matching

The main statistical descriptors of the posterior parameter distributions for the BIO–ALGAE model, obtained from the second iteration of PESTPP-IES, are reported in Table 3. The calibrated values are generally consistent with ranges reported in the literature for similar models (Flynn et al., 2015; Johnson and Zhang, 2016). The inversion process significantly reduced the uncertainty for most parameters, as indicated by the narrow 5th–95th percentile range compared to the prior bounds in Table 1.

The results of the history matching are shown in Fig. 10 for ten stations located across different reaches of the river network, selected for having the most continuous time series. Specifically, the CBOD results are plotted for station MAS-140 (Canale del Capannone) and MAS-129 (Ombrone Pistoiese river upstream section); the ammonium concentrations for MAS-130 (Ombrone Pistoiese river downstream section), and MAS-2011 (Pescia river); the nitrates for MAS-106 (Arno river), and MAS-503 (Arno river); the dissolved oxygen for MAS-141 (Torrente Nievole stream), and MAS-552 (Bisenzio river); the total phosphorous for MAS-510A (Torrente Cessana stream), and MAS-512 (Torrente Brana stream). The station codes (e.g., "MAS-XYZ") reflect the original naming conventions used by ARPAT, the environmental agency responsible for the measurements.

The CBOD and total phosphorus datasets are the most affected by data scarcity, as also reflected by the limited number of observations (orange points) shown in Fig. 10. Across all parameters, no clear long-term trend is evident over the 10-year observation period. For all parameters and monitoring stations reported in Fig. 10, the model successfully reproduces the correct order of magnitude. Temporal

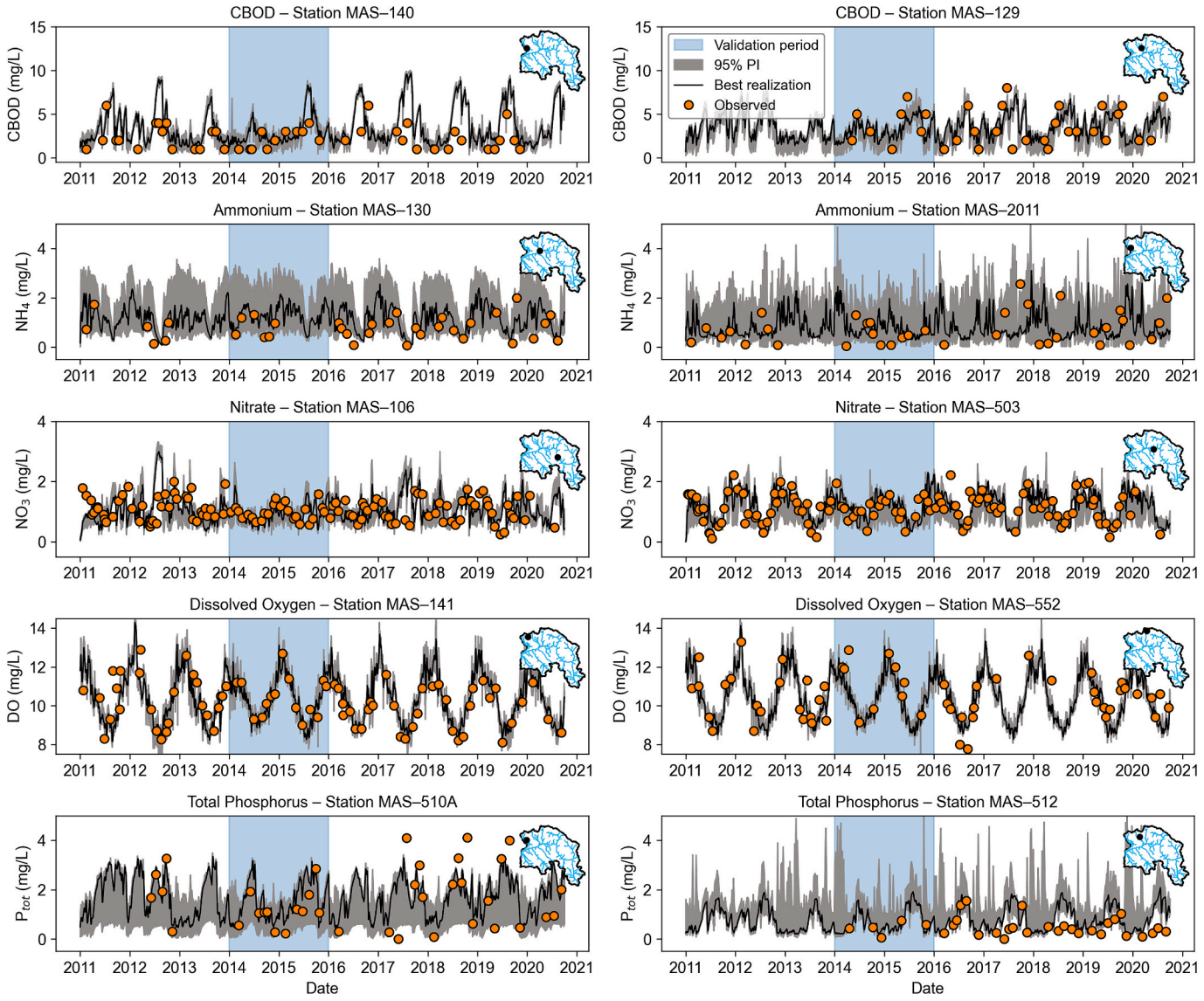


Fig. 10. Results of model calibration, validation and uncertainty analysis for selected stations.

variability is generally well captured for CBOD, nitrate, and dissolved oxygen, whereas agreement is weaker for ammonium and phosphorus. Nevertheless, for all parameters, the majority of observations fall within the model prediction uncertainty bounds.

Fig. 11 presents the overall model performance as scatter plots comparing observed concentrations with those simulated by the best-performing realization. The scatter plots indicate that the validation dataset shows a level of agreement with the model comparable to, and in most cases even better than, that achieved for the calibration dataset. This suggests that the calibration is robust and not overfitted.

Model performance for dissolved oxygen (DO) is generally good at concentrations above 7 mg/L, while performance decreases at lower concentrations, where the model tends to overestimate observed values. For the remaining parameters, the largest prediction errors are associated with reaches of lower Strahler order. CBOD concentrations are typically overestimated in reaches with Strahler orders 2–3, whereas underestimation predominantly occurs in higher-order reaches (>4). For ammonium, a small number of pronounced outliers are observed exclusively in Strahler order 2 reaches. Similarly, the phosphorus scatter plot shows several outliers, all associated with reaches of the same Strahler order (Strahler = 3).

3.4. Analysis of distributed loads

The spatial discretization of non-point sources enabled the calculation of pollutant loads for each sub-catchment associated with individual reaches in the river network. During the PESTPP-IES inversion process, the empirical parameters *a* and *b* (for both quick and slow flow components) were adjusted for each aggregated group of sub-catchments, while accounting for spatial smoothness constraints imposed by regularization.

The load for each constituent was computed by summing the mass influxes from both quick and slow flows into each reach, normalizing by the contributing sub-catchment area, and averaging the results annually. The resulting spatial load maps, presented in Fig. 12, highlight both the presence and magnitude of pollutant sources across the catchment, effectively identifying major source areas.

To quantify uncertainty in the estimated parameters, we defined a simple proxy based on the ratio of the standard deviation of the posterior distribution to that of the prior. A proxy value near 0 indicates strong constraint by the data, with a significantly narrowed posterior distribution and thus low uncertainty. Conversely, a value near 1 implies that the parameter could not be constrained by the available observations, and its uncertainty remains largely unchanged.

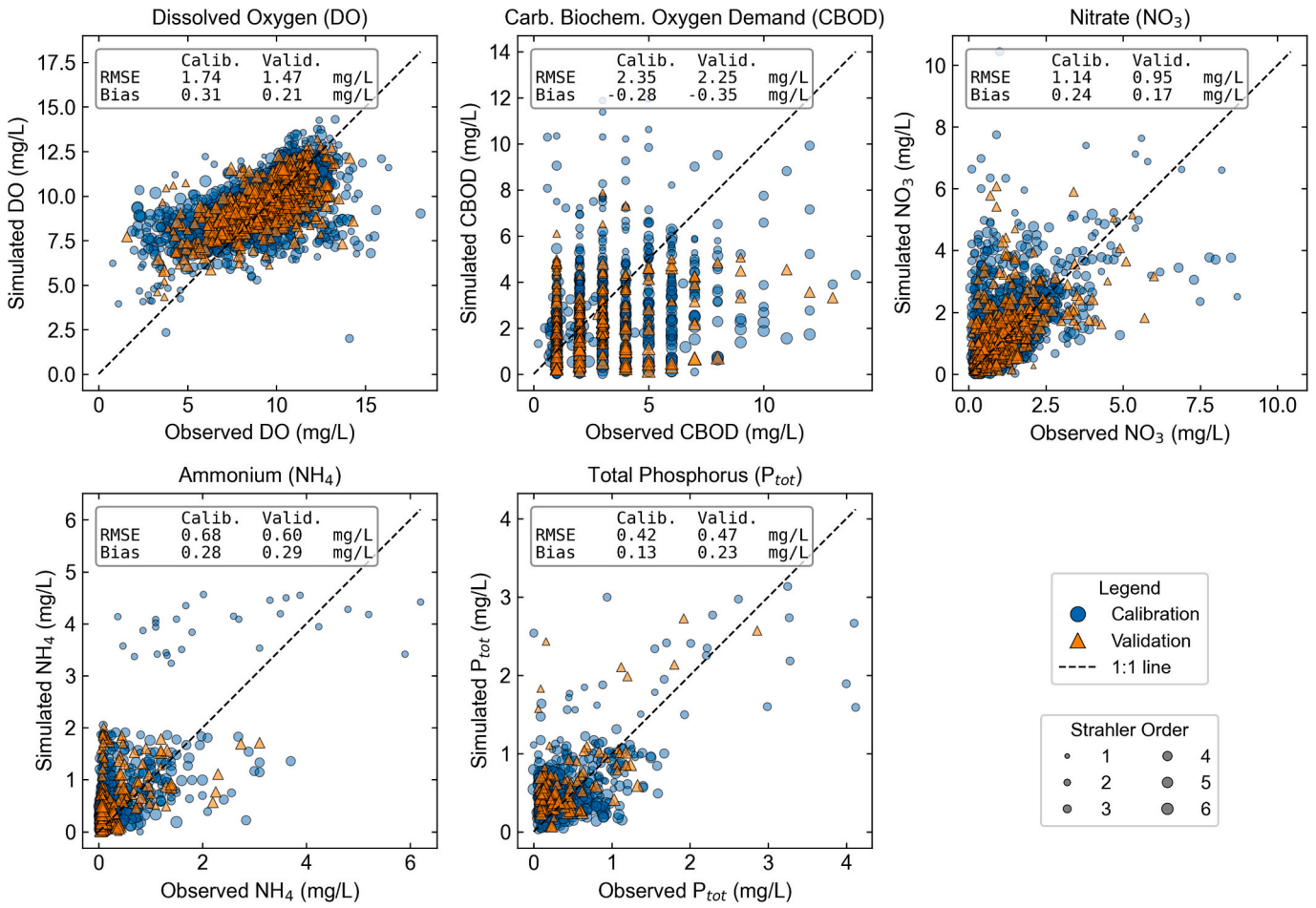


Fig. 11. Scatter plots between the observed data and the best-performing realization.

Fig. 12 shows the spatial distribution of the calculated non-point source loads for the analysed water quality constituents, with one panel for each contaminant. Each panel is composed of two subfigures: the left subfigure displays the spatial distribution of the loads together with the locations of the monitoring stations, with areas where the uncertainty proxy exceeds 0.9 highlighted using hatching, while the right subfigure shows the spatial distribution of the uncertainty proxy across the entire study area.

For CBOD, several hotspot areas are identified, with export values reaching more than 6000 kg km^{-2} yr^{-1}, while ammonium exports exceed 2000 kg km^{-2} yr^{-1} in some areas. Export values for the remaining constituents are substantially lower, with less pronounced hotspots, most of which occur in areas characterized by high uncertainty. Load estimates with the lowest uncertainty are obtained for ammonium, nitrate, organic phosphorus, and total inorganic phosphorus (TIP), whereas the highest uncertainties are associated with CBOD and organic nitrogen. A consistently uncertain area is identified across all constituents, corresponding to the westernmost part of the catchment, i.e., the most downstream subcatchments.

3.5. Discussion

The calibration approach used in this study successfully produced posterior parameter distributions that are narrower than the prior ranges. The highly parameterized spatial load parameters were generally well identified, with lower uncertainties than the initial guesses, indicating that the regularized Tikhonov calibration scheme is appropriate for this purpose. In contrast, the BIO–ALGAE parameters were applied parsimoniously, using a single global set across all network reaches. As shown in Table 3, spatially regularized calibration effectively reduced the posterior uncertainty ranges (5th–95th percentiles) relative to the initial bounds, suggesting that the available data provide sufficient information to constrain these parameters even when applied uniformly. An alternative approach could involve Strahler-order-specific parameterization of the water-quality model to capture additional process heterogeneity. However, this would substantially increase the number of calibration parameters and exacerbate parameter identifiability issues, motivating our choice of a common parameter set for the entire network. While the distributed load parameters benefit from the explicit spatial regularization, imposing spatial constraints that enable effective calibration even for highly parameterized variables, no straightforward regularization exists for water-quality parameters grouped by stream order. Overall, our results indicate that this parsimonious approach still effectively reduces the range of posterior parameter distributions (Table 3).

Although we assumed that the model parameters are stationary, neglecting possible temporal trends or shifts that could result from major land use changes or other modifications in pollutant sources, this represents a strong assumption and a notable limitation of the model. Nevertheless, the observed water quality time series over the ten-year period do not exhibit any discernible trends (Fig. 10).

The scatter plot (Fig. 11) highlights two different aspects, related to DO and the reaches with lower stream orders. In the case of dissolved oxygen, model deviations are more pronounced at lower observed DO concentrations. This suggests reduced model accuracy under conditions of strong oxygen depletion, possibly due to the underestimation of oxygen-consuming processes, such as ammonium and CBOD oxidation, or the presence of unmodeled substances that contribute to additional

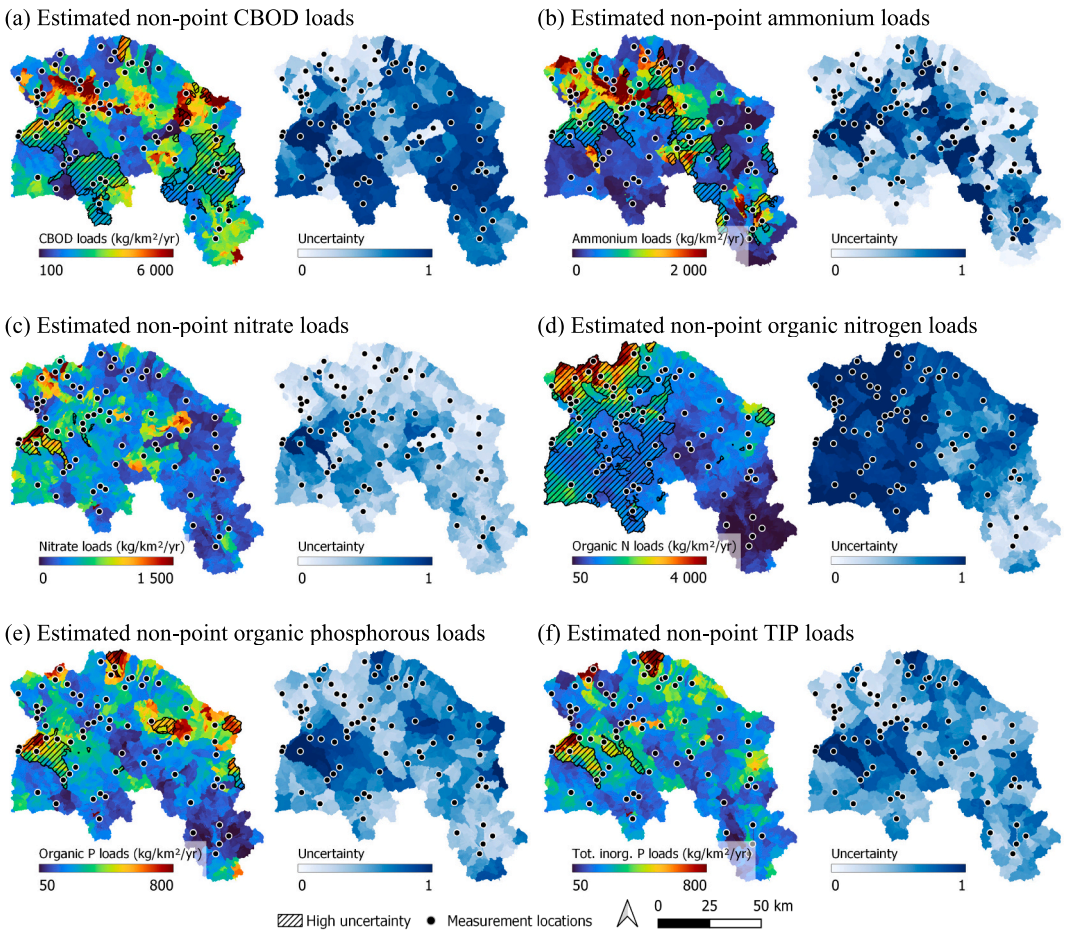


Fig. 12. Spatial distribution of the estimated non-point source loads obtained from the inversion. Within each panel, the left subfigure displays the spatial distribution of the loads together with the locations of the monitoring stations; hatched areas indicate where the uncertainty proxy exceeds 0.9. The right subfigure shows the spatial distribution of the uncertainty proxy.

oxygen demand. This limitation may be addressed by extending the model to include spatially distributed parametrization for the BIO–ALGAE module, not just for the pollutant loads. However, while we recognize that assuming uniform biochemical reaction rates and parameters across all river reaches is a major simplification, implementing distributed kinetic parameters could increase the risk of generating implausible spatial patterns through the inversion process, potentially leading to biased results (Moore and Doherty, 2005). Therefore, the current approach reflects a trade-off between model complexity and data limitations.

The model exhibits its poorest performance in lower-order reaches (e.g., see ammonium and phosphorus in Fig. 11), suggesting that some processes are not fully captured at the catchment scale. Headwater reaches are typically the most biologically active and sensitive to environmental forcings, such as intermittent flow, shading by riparian vegetation, spatial heterogeneity within a reach, and riverbed or morphological features that influence reoxygenation. Fig. 13 presents the spatial overlay of estimated load exports and land use. The highest non-point CBOD loads are primarily concentrated in highly urbanized areas. Panel (a) highlights the densest urban zones of the catchment, including the cities of Florence, Prato, and Pistoia. This spatial correspondence indicates that urban land use, likely through runoff from impervious surfaces, combined sewer overflows (CSOs), and diffuse domestic sources, plays a major role in contributing to catchment-wide pollution.

A similar pattern is observed for ammonium loads (panel b), where areas of high contributions largely coincide with urbanized zones. This reinforces the link between human activities, such as sewage leakage, stormwater runoff, and inadequately treated wastewater, and higher nutrient concentrations in surface waters.

Panel (c) shows the spatial distribution of total inorganic phosphorus (TIP) loads. While higher phosphorus loads are still associated with urban areas, elevated contributions are also evident in predominantly agricultural regions, suggesting that practices such as fertilizer application represent an additional and significant source of non-point phosphorus inputs. Overall, the spatial patterns of phosphorus reflect the combined influence of both urban and agricultural land uses on phosphorous pollution.

These results also highlight distinct export regimes among pollutants. The dilution-dominated behaviour of CBOD, ammonium, and phosphorus indicates the dominance of point sources or rapidly depleted diffuse inputs, particularly in urbanized areas. In contrast, nitrate exhibited flushing behaviour, reflecting mobilization of subsurface stores, a finding consistent with broader evidence from catchment biogeochemistry (Basu et al., 2010; Stewart et al., 2022). Together, these findings provide a quantitative basis for linking land use to pollutant generation and offer critical insights into the spatial targeting of management actions.

A key methodological contribution of this work is the application of a highly-parameterized, spatially-distributed inversion approach. By parameterizing the non-point source loads at the sub-catchment level and using Tikhonov regularization to ensure spatial smoothness, we were able to extract spatial information from sparse data, a common challenge in water quality modelling (Tscheikner-Gratl et al., 2019). The

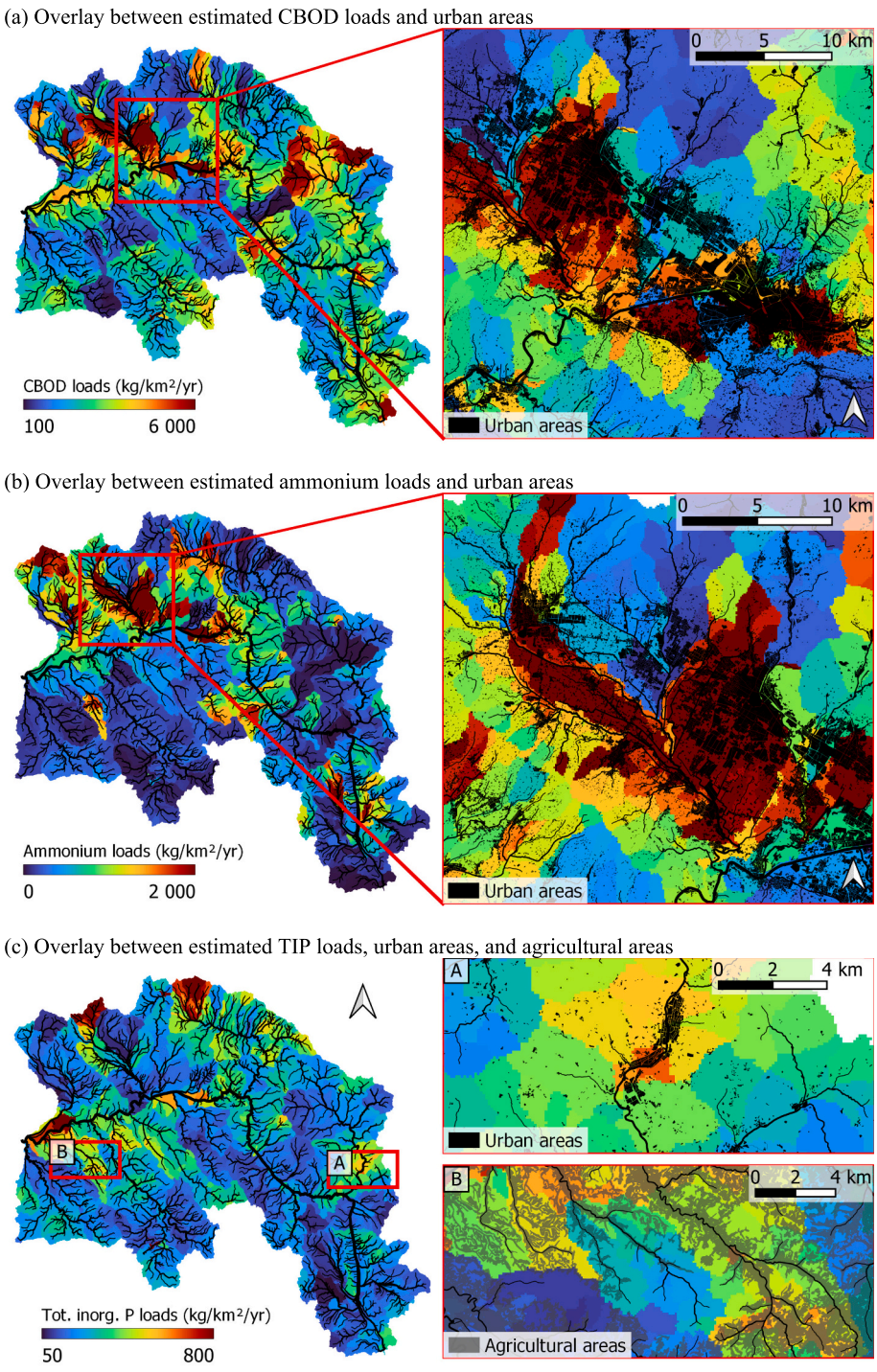


Fig. 13. Overlay between estimated loads spatial distributions and land cover (i.e., urban areas and agricultural surfaces).

reliance on low-frequency "grab sample" data further complicates pollutant load estimation in large-scale systems (Fu et al., 2020), particularly when assessing annual or seasonal cumulates (Moatar and Meybeck, 2005). Compared with traditional lumped calibration, this framework offered a more realistic representation of the heterogeneity of pollutant generation.

The use of the iterative ensemble smoother (PESTPP–IES) was crucial to handle the computational demand of this high-dimensional problem. Unlike the classical Gauss–Levenberg–Marquardt approach, which requires explicit calculation of large sensitivity matrices (Jacobian matrices) and becomes computationally prohibitive with thousands of parameters, PESTPP–IES infers sensitivities from ensembles of model runs, substantially reducing computational costs while also providing posterior uncertainty estimates (White et al., 2021).

Despite the model's strengths, the following limitations must be acknowledged. The sparse temporal resolution and uneven spatial coverage of water quality data limited both load estimation and calibration robustness, leading to elevated uncertainty in poorly monitored areas that can propagate errors downstream. Additionally, the assumption of spatially uniform reaction kinetics across all reaches represents a simplification that may overlook critical local-scale heterogeneities. The model tended to underperform during episodes of low

DO concentrations, potentially indicating an underestimation of oxygen-demanding processes such as ammonium oxidation, CBOD degradation, or other transformation pathways not currently included in the model structure. The assumption of temporally constant load inputs over the ten-year simulation period appeared to be acceptable. No significant temporal trends or abrupt shifts were detected in the observations that would have contradicted this assumption, suggesting that the long-term average behaviour of the system was adequately captured by the model. While the current modelling framework effectively captured first-order reactive processes, it does not yet account for more complex dynamics such as sediment–water interactions, or algal species competition. Addressing these aspects in future work could further improve model accuracy and applicability.

4. Conclusions

This study demonstrated the application of the MOBIDIC–ADR model, enhanced with the newly developed BIO–ALGAE reactive component, to simulate nutrient dynamics across a complex river network of a large catchment. Supported by the PESTPP–IES framework, the approach enabled the derivation of spatially distributed pollutant loads and identification of key source areas, despite limitations in monitoring data.

Key conclusions are:

- The modelling framework successfully reproduced observed patterns of carbonaceous biochemical oxygen demand (CBOD), dissolved oxygen (DO), nitrogen species, and phosphorus compounds, across all the Arno River catchment.
- Distinct pollutant export regimes were identified: CBOD, ammonium, and phosphorus were dilution-dominated and linked mainly to urban areas, while nitrate displayed flushing behaviour indicative of subsurface storage release.
- Spatially explicit load maps highlighted pollution hotspots, revealing strong links between CBOD/ammonium and urban activities, and mixed urban–agricultural contributions to phosphorus.
- The uncertainty metrics, based on the ratio of posterior to prior standard deviation, provided insight into the spatial variability of information content across the catchment. In general, areas with more monitoring coverage showed well-constrained parameters, whereas parameters associated with data-scarce regions remained largely unconstrained.
- The framework proved computationally feasible for highly parameterized problems and provided uncertainty estimates, thereby offering a robust tool for decision support.

Overall, the study illustrates the potential of combining process-based reactive transport modelling with spatially distributed inverse calibration to assess catchment-scale water quality dynamics. It also underscored the critical importance of high-resolution, spatially representative observational datasets which are essential for informing targeted management strategies and guiding investments in monitoring infrastructure, ultimately supporting more effective water quality improvements at the catchment scale. Further research should address the incorporation of spatially-variable reaction kinetics and additional biogeochemical processes, e.g., oxygen-demanding pathways, sediment–water interactions, shading by riparian vegetation, and multiple algal species.

CRediT authorship contribution statement

Matteo Masi: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Data curation, Conceptualization. **Maryam Barati Moghaddam:** Writing – original draft, Software, Data curation. **Fabio Castelli:** Writing – review & editing, Software. **Chiara Arrighi:** Writing – original draft, Supervision, Resources, Project administration, Funding acquisition, Formal analysis, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was carried out within RETURN Extended Partnership and received funding from the European Union Next-GenerationEU (National Recovery and Resilience Plan – NRRP, Mission 4, Component 2, Investment 1.3 – D.D. 1243 2/8/2022, PE0000005). The authors also wish to acknowledge Fondazione Cassa di Risparmio Firenze for co-funding this work within the project ECO-C – "ECOidrologia dei corsi d'acqua urbani nel contesto dei Cambiamenti climatici e socioeconomici".

Data availability

Most datasets analysed in this study are publicly accessible, and the sources are referenced in the text. The remaining data will be made available on request.

References

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

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E., Rasmussen, J., 1986. An introduction to the European Hydrological System — Systeme Hydrologique Europeen, "SHE", 1: history and philosophy of a physically-based, distributed modelling system. J. Hydrol. 87 (1–2), 45–59. https://doi.org/10.1016/0022-1694(86)90114-9.

Arheimer, B., Olsson, J., 2003. Integration and coupling of hydrological models with water quality models. In: Applications in Europe, Report of the Swedish Meteorological and Hydrological Institute, Norrköping Sweden, p. 49.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large area hydrologic modeling and assessment part I: model development. J. Am. Water Resour. Assoc. 34 (1), 73–89. https://doi.org/10.1111/J.1752-1688.1998.TB05961.X.

Ator, S.W., 2019. Spatially referenced models of streamflow and nitrogen, phosphorus, and suspended-sediment loads in streams of the northeastern United States. US Geol. Surv. Sci. Investig. Rep. 2019–5118, 57. https://doi.org/10.3133/SIR20195118.

Bai, J., Zhao, J., Zhang, Z., Tian, Z., 2022. Assessment and a review of research on surface water quality modeling. Ecol. Model. 466, 109888. https://doi.org/10.1016/J.ECOLMODEL.2022.109888.

Baly, E.C.C., 1935. The kinetics of photosynthesis. Proc. R. Soc. Lond. B Biol. Sci. 117 (804), 218–239. https://doi.org/10.1098/RSPB.1935.0026.

Bartley, R., Speirs, W.J., Ellis, T.W., Waters, D.K., 2012. A review of sediment and nutrient concentration data from Australia for use in catchment water quality models. Mar. Pollut. Bull. 65 (4–9), 101–116. https://doi.org/10.1016/J.MARPOLBUL.2011.08.009.

Basu, N.B., Destouni, G., Jawitz, J.W., Thompson, S.E., Loukinova, N.V., Darracq, A., et al., 2010. Nutrient loads exported from managed catchments reveal emergent biogeochemical stationarity. Geophys. Res. Lett. 37 (23). https://doi.org/10.1029/2010GL045168.

Behrouz, M.S., Sample, D.J., Kisila, O.B., Harrison, M., Nayeb Yazdi, M., Garna, R.K., 2024. Parameterization of nutrients and sediment build-up/wash-off processes for simulating stormwater quality from specific land uses. J. Environ. Manag. 358, 120768. https://doi.org/10.1016/J.JENVMAN.2024.120768.

Beven, K., Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. Hydrol. Process. 6 (3), 279–298. https://doi.org/10.1002/HYP.3360060305.

Bin Mamoon, W., Zhang, K., Luhar, M., Parolari, A.J., 2025. Stream nutrient load and concentration estimation from minimal measurements. Geophys. Res. Lett. 52 (8), e2025GL114935. https://doi.org/10.1029/2025GL114935.

Blaen, P.J., Khamis, K., Lloyd, C.E.M., Bradley, C., Hannah, D., Krause, S., 2016. Real-time monitoring of nutrients and dissolved organic matter in rivers: capturing event dynamics, technological opportunities and future directions. Sci. Total Environ. 569–570, 647–660. https://doi.org/10.1016/J.SCITOTENV.2016.06.116.

Campolo, M., Andreussi, P., Soldati, A., 2002. Water quality control in the river Arno. Water Res. 36 (10), 2673–2680. https://doi.org/10.1016/S0043-1354(01)00483-3.

Cao, M., Gao, W., Zhang, Y., Guo, F., 2025. Riverine nitrogen and phosphorus output in an arid watershed: combined effects of anthropogenic nutrient inputs, land use, and

hydrology. Environ. Syst. Res. 14 (1), 1–14. https://doi.org/10.1186/S40068-025-00399-2/TABLES/1.
Capell, R., Bartosova, A., Tonderski, K., Arheimer, B., Pedersen, S.M., Zilans, A., 2021. From local measures to regional impacts: modelling changes in nutrient loads to the Baltic Sea. J. Hydrol. Reg. Stud. 36, 100867. https://doi.org/10.1016/J.EJRH.2021.100867.
Carrayrou, J., Mosé, R., Behra, P., 2004. Operator-splitting procedures for reactive transport and comparison of mass balance errors. J. Contam. Hydrol. 68 (3–4), 239–268. https://doi.org/10.1016/S0169-7722(03)00141-4.
Castelli, F., Menduni, G., Mazzanti, B., 2009. A distributed package for sustainable water management: a case study in the Arno basin. In: Role of Hydrology in Water Resources Management, vol. 327. IAHS-AISH Publication, pp. 52–61.
Castillo, A., Castelli, F., Entekhabi, D., 2015. Gravitational and capillary soil moisture dynamics for distributed hydrologic models. Hydrol. Earth Syst. Sci. 19 (4), 1857–1869. https://doi.org/10.5194/HESS-19-1857-2015.
Chapra, S.C., 2008. Surface Water-Quality Modeling, p. 844.
Chapra, S., Pelletier, G., 2003. A Modeling Framework for Simulating River and Stream Water Quality Documentation and Users Manual the Mystic River at Medford, MA.
Costa, C.M.D.S.B., Leite, I.R., Almeida, A.K., de Almeida, I.K., 2021. Choosing an appropriate water quality model—a review. Environ. Monit. Assess. 193 (1), 1–15. https://doi.org/10.1007/S10661-020-08786-1/FIGURES/5.
Crossman, J., Bussi, G., Whitehead, P.G., Butterfield, D., Lannergård, E., Futter, M.N., 2021. A new, catchment-scale integrated water quality model of phosphorus, dissolved oxygen, biochemical oxygen demand and phytoplankton: INCA-Phosphorus Ecology (PEco). Water 13 (5), 723. https://doi.org/10.3390/W13050723.
Doherty, J., 2015. Calibration and Uncertainty Analysis for Complex Environmental Models. Watermark Numerical Computing.
Elmore, H.L. t, Hayes, T.W., 1960. Solubility of atmospheric oxygen in water. In: Proc. Am. Sot. Civil Engrs, vol. 86, pp. 41–53.
Ercolani, G., Castelli, F., 2017. Variational assimilation of streamflow data in distributed flood forecasting. Water Resour. Res. 53 (1), 158–183. https://doi.org/10.1002/2016WR019208.
Flipo, N., Even, S., Poulin, M., Théry, S., Ledoux, E., 2007. Modeling nitrate fluxes at the catchment scale using the integrated tool CAWAQS. Sci. Total Environ. 375 (1–3), 69–79. https://doi.org/10.1016/J.SCITOTENV.2006.12.016.
Flynn, K.F., Suplee, M.W., Chapra, S.C., Tao, H., 2015. Model-based nitrogen and phosphorus (nutrient) criteria for large temperate rivers: 1. Model development and application. J. Am. Water Resour. Assoc. 51 (2), 421–446. https://doi.org/10.1111/JAWR.12253.
Fonseca, A., Botelho, C., Boaventura, R.A.R., Vilar, V.J.P., 2024. Evaluating the effects of parameter uncertainty on river water quality predictions. Resources 13 (8), 106. https://doi.org/10.3390/RESOURCES13080106/S1.
Fu, B., Horsburgh, J.S., Jakeman, A.J., Gualtieri, C., Arnold, T., Marshall, L., et al., 2020. Modeling water quality in watersheds: from here to the next generation. Water Resour. Res. 56 (11), e2020WR027721. https://doi.org/10.1029/2020WR027721.
Gassman, P.W., Reyes, M.R., Green, C.H., Arnold, J.G., 2007. The soil and water assessment tool: historical development, applications, and future research directions. Trans. ASABE 50 (4), 1211–1250. https://doi.org/10.13031/2013.23637.
Ghaffar, S., Zhou, X., Jomaa, S., Yang, X., Meon, G., Rode, M., 2024. Toward a data-effective calibration of a fully distributed catchment water quality model. Water Resour. Res. 60 (9), e2023WR036527. https://doi.org/10.1029/2023WR036527.
Grizzetti, B., Bouraoui, F., De Marsily, G., 2008. Assessing nitrogen pressures on European surface water. Glob. Biogeochem. Cycles 22 (4), 4023. https://doi.org/10.1029/2007GB003085.
Gunnerson, C.G., 1967. Streamflow and quality in the Columbia River Basin. J. Sanit. Eng. Div. 93 (6), 1–16. https://doi.org/10.1061/JSEDAI.0000767.
James, W.R.C., Rossman, L.A., 2010. User's Guide to SWMM 5. Computational Hydraulics International.
Johnson, B.E., Zhang, Z., 2016. Aquatic Nutrient Simulation Modules (NSMs) Developed for Hydrologic and Hydraulic Models. Environmental Laboratory (U.S.). Retrieved from. https://hdl.handle.net/11681/10112.
Jones, A.S., Horsburgh, J.S., Mesner, N.O., Ryel, R.J., Stevens, D.K., 2012. Influence of sampling frequency on estimation of annual total phosphorus and total suspended solids loads. J. Am. Water Resour. Assoc. 48 (6), 1258–1275. https://doi.org/10.1111/J.1752-1688.2012.00684.X.
Keller, A.A., Garner, K., Rao, N., Knipping, E., Thomas, J., 2023. Hydrological models for climate-based assessments at the watershed scale: a critical review of existing hydrologic and water quality models. Sci. Total Environ. 867, 161209. https://doi.org/10.1016/J.SCITOTENV.2022.161209.
Khorashadi Zadeh, F., Nossent, J., Woldegiorgis, B.T., Bauwens, W., Van Griensven, A., 2022. A fast and effective parameterization of water quality models. Environ. Model Softw. 149, 105331. https://doi.org/10.1016/J.ENVSOFT.2022.105331.
Kirchner, J.W., Feng, X., Neal, C., Robson, A.J., 2004. The fine structure of water-quality dynamics: the (high-frequency) wave of the future. Hydrol. Process. 18 (7), 1353–1359. https://doi.org/10.1002/HYP.5537.
Kling, H., Fuchs, M., Paulin, M., 2012. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. J. Hydrol. 424–425, 264–277. https://doi.org/10.1016/J.JHYDROL.2012.01.011.
Lannergård, E.E., Sandström, S., Nilsson, E.W., Lewan, E., Futter, M.N., 2024. High expectations and open questions: using high-frequency proxy data to calibrate a catchment-scale phosphorus model. ACS ES&T Water 4 (5), 2135–2143. https://doi.org/10.1021/ACSESTWATER.3C00775/ASSET/IMAGES/LARGE/EW3C00775_0007.JPEG.
Leigh, C., Kandanaarachchi, S., McGree, J.M., Hyndman, R.J., Alsibai, O., Mengersen, K., Peterson, E.E., 2019. Predicting sediment and nutrient concentrations from high-frequency water-quality data. PLoS One 14 (8), e0215503. https://doi.org/10.1371/JOURNAL.PONE.0215503.
Lindström, G., Pers, C., Rosberg, J., Strömqvist, J., Arheimer, B., 2010. Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales. Hydrol. Res. 41 (3–4), 295–319. https://doi.org/10.2166/NH.2010.007.
Lucas, L.V., Brown, C.J., Robertson, D.M., Baker, N.T., Johnson, Z.C., Green, C.T., et al., 2025. Gaps in water quality modeling of hydrologic systems. Water 17 (8), 1200. https://doi.org/10.3390/W17081200.
Marsili-Libelli, S., 2016. Environmental Systems Analysis With MATLAB. CRC Press.
Masi, M., Masseroni, D., Castelli, F., 2025. Coupled hydrologic, hydraulic, and surface water quality models for pollution management in urban–rural areas. J. Hydrol. 657, 133172. https://doi.org/10.1016/J.JHYDROL.2025.133172.
McIntyre, N., Jackson, B., Wade, A.J., Butterfield, D., Wheater, H.S., 2005. Sensitivity analysis of a catchment-scale nitrogen model. J. Hydrol. 315 (1–4), 71–92. https://doi.org/10.1016/J.JHYDROL.2005.04.010.
Minaudo, C., Dupas, R., Gascuel-Odoux, C., Fovet, O., Mellander, P.E., Jordan, P., et al., 2017. Nonlinear empirical modeling to estimate phosphorus exports using continuous records of turbidity and discharge. Water Resour. Res. 53 (9), 7590–7606. https://doi.org/10.1002/2017WR020590.
Moatar, F., Meybeck, M., 2005. Compared performances of different algorithms for estimating annual nutrient loads discharged by the eutrophic River Loire. Hydrol. Process. 19 (2), 429–444. https://doi.org/10.1002/HYP.5541.
Moore, C., Doherty, J., 2005. Role of the calibration process in reducing model predictive error. Water Resour. Res. 41 (5), 1–14. https://doi.org/10.1029/2004WR003501.
Namugize, J.N., Jewitt, G.P.W., Clark, D., Strömqvist, J., 2017. Assessment of the Hype model for simulation of water and nutrients in the upper uMngeni river catchment in South Africa. Hydrol. Earth Syst. Sci. Discuss. 1–37. https://doi.org/10.5194/HESS-2017-365.
Neilson, B.T., Chapra, S.C., 2003. Integration of water quality monitoring and modeling for TMDL development. Water Resour. IMPACT 5 (1), 9–11. Retrieved from. http://www.jstor.org/stable/wateresoimpa.5.1.0009.
Pellerin, B.A., Bergamaschi, B.A., Gilliom, R.J., Crawford, C.G., Saraceno, J., Frederick, C.P., et al., 2014. Mississippi River nitrate loads from high frequency sensor measurements and regression-based load estimation. Environ. Sci. Technol. 48 (21), 12612–12619. https://doi.org/10.1021/ES504029C/SUPPL_FILE/ES504029C_SI_001.PDF.
Pellerin, B.A., Stauffer, B.A., Young, D.A., Sullivan, D.J., Bricker, S.B., Walbridge, M.R., et al., 2016. Emerging tools for continuous nutrient monitoring networks: sensors advancing science and water resources protection. J. Am. Water Resour. Assoc. 52 (4), 993–1008. https://doi.org/10.1111/1752-1688.12386.
Phillips, A.K., Mandal, S., Mohamed, M., Sorichetti, R.J., Ross, C.A., Thomas, J.L., Wellen, C.C., 2024. Is comprehensive event sampling necessary for constraining process models of water quality? A comparison of high and low frequency phosphorus sampling programs for constraining the HYPE water quality model. J. Hydrol. 639, 131502. https://doi.org/10.1016/J.JHYDROL.2024.131502.
Rankinen, K., Turtola, E., Lemola, R., Futter, M., Bernal, J.E.C., 2021. Nutrient load mitigation with wintertime cover as estimated by the INCA model. Water 13 (4), 450. https://doi.org/10.3390/W13040450.
Salim Aoubid, H., Opp, C., 2023. Nitrogen and phosphorus discharge loads assessment using the SWAT model: a case study of the Shatt Al-Arab river basin. Appl. Sci. 13 (14), 8376. https://doi.org/10.3390/APP13148376.
Saravanan, S., Singh, L., Sathiyamurthi, S., Sivakumar, V., Velusamy, S., Shanmugamoorthy, M., 2023. Predicting phosphorus and nitrate loads by using SWAT model in Vamanapuram River Basin, Kerala, India. Environ. Monit. Assess. 195 (1), 1–13. https://doi.org/10.1007/S10661-022-10786-2/FIGURES/7.
Socolofsky, S.A., Jirka, G.H., 2005. CVEN 489-501: Special Topics in Mixing and Transport Processes in the Environment Engineering-Lectures.
Speir, S.L., Rose, L.A., Blaszczak, J.R., Kincaid, D.W., Fazekas, H.M., Webster, A.J., et al., 2024. Catchment concentration–discharge relationships across temporal scales: a review. WIREs Water 11 (2), e1702. https://doi.org/10.1002/WAT2.1702.
Spill, C., Gassmann, M., 2025. High resolution simulation of nitrate and ammonium from point and diffuse sources in a small headwater catchment. Hydrol. Process. 39 (2), e70081. https://doi.org/10.1002/HYP.70081.
Stewart, B., Shanley, J.B., Kirchner, J.W., Norris, D., Adler, T., Bristol, C., et al., 2022. Streams as mirrors: reading subsurface water chemistry from stream chemistry. Water Resour. Res. 58 (1), e2021WR029931. https://doi.org/10.1029/2021WR029931.
Strokal, M., Wang, M., Micella, I., Janssen, A.B.G., 2024. Building trust in large-scale water quality models: 13 alternative strategies beyond validation. Discov. Water 4 (1), 1–11. https://doi.org/10.1007/S43832-024-00149-Y.
Stutter, M., Baggaley, N.J., Davies, J., Gagkas, Z., Janes-Bassett, V., Laudon, H., et al., 2023. The riparian reactive interface: a climate-sensitive gatekeeper of global nutrient cycles. Front. Environ. Sci. 11, 1213175. https://doi.org/10.3389/FENVS.2023.1213175/BIBTEX.
Swinehart, D.F., 1962. The Beer-Lambert law. J. Chem. Educ. 39 (7), 333–335. https://doi.org/10.1021/ED039P333.
Tscheikner-Gratl, F., Bellos, V., Schellart, A., Moreno-Rodenas, A., Muthusamy, M., Langeveld, J., et al., 2019. Recent insights on uncertainties present in integrated catchment water quality modelling. Water Res. 150, 368–379. https://doi.org/10.1016/J.WATRES.2018.11.079.
Udías, A., Efremov, R., Galbiati, L., Cañamón, I., 2014. Simulation and multicriteria optimization modeling approach for regional water restoration management. Ann. Oper. Res. 219 (1), 123–140. https://doi.org/10.1007/S10479-012-1101-X/TABLES/4.

USEPA, 1995. Technical Guidance Manual for Developing Total Maximum Daily Loads, Book II: Streams and Rivers, Part 1: Biochemical Oxygen Demand/Dissolved Oxygen and Nutrients/Eutrophication (Vol. EPA823-B-95–007). US EPA, Washington, DC.

Vrugt, J.A., Ter Braak, C.J.F., Diks, C.G.H., Robinson, B.A., Hyman, J.M., Higdon, D., 2009. Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. Int. J. Nonlinear Sci. Numer. Simul. 10 (3), 273–290. https://doi.org/10.1515/IJNSNS.2009.10.3.273/PDF/FIRSTPAGE.

Vystavna, Y., Paule-Mercado, M.C., Schmidt, S.I., Hejzlar, J., Porcal, P., Matiatos, I., 2023. Nutrient dynamics in temperate European catchments of different land use under changing climate. J. Hydrol. Reg. Stud. 45, 101288. https://doi.org/10.1016/J.EJRH.2022.101288.

Wade, A.J., Durand, P., Beaujouan, V., Wessel, W.W., Raat, K.J., Whitehead, P.G., et al., 2002. A nitrogen model for European catchments: INCA, new model structure and equations. Hydrol. Earth Syst. Sci. 6 (3), 559–582. https://doi.org/10.5194/HESS-6-559-2002.

Westfall, T.G., Peterson, T.J., Lintern, A., Western, A.W., 2025. Slow and quick flow models explain the temporal dynamics of daily salinity in streams. Water Resour. Res. 61 (6), e2024WR039103. https://doi.org/10.1029/2024WR039103.

White, J.T., 2018. A model-independent iterative ensemble smoother for efficient history-matching and uncertainty quantification in very high dimensions. Environ. Model Softw. 109, 191–201. https://doi.org/10.1016/J.ENVSOFT.2018.06.009.

White, J.T., Hunt, R.J., Fienen, M.N., Doherty, J.E., 2020. Approaches to highly parameterized inversion: PEST++ version 5, a software suite for parameter estimation, uncertainty analysis, management optimization and sensitivity analysis. In: Techniques and Methods. https://doi.org/10.3133/TM7C26.

White, J.T., Hemmings, B., Fienen, M.N., Knowling, M.J., 2021. Towards improved environmental modeling outcomes: enabling low-cost access to high-dimensional, geostatistical-based decision-support analyses. Environ. Model Softw. 139, 105022. https://doi.org/10.1016/J.ENVSOFT.2021.105022.

Whitehead, P.G., Wilson, E.J., Butterfield, D., 1998. A semi-distributed integrated nitrogen model for multiple source assessment in catchments (INCA): part I — model structure and process equations. Sci. Total Environ. 210–211, 547–558. https://doi.org/10.1016/S0048-9697(98)00037-0.

Whitehead, P.G., Wilby, R.L., Butterfield, D., Wade, A.J., 2006. Impacts of climate change on in-stream nitrogen in a lowland chalk stream: an appraisal of adaptation strategies. Sci. Total Environ. 365 (1–3), 260–273. https://doi.org/10.1016/J.SCITOTENV.2006.02.040.

Whitehead, P.G., Sarkar, S., Jin, L., Futter, M.N., Caesar, J., Barbour, E., et al., 2015. Dynamic modeling of the Ganga river system: impacts of future climate and socio-economic change on flows and nitrogen fluxes in India and Bangladesh. Environ. Sci. Process Impacts 17 (6), 1082–1097. https://doi.org/10.1039/C4EM00616J.

Wool, T., Ambrose, R.B., Martin, J.L., Comer, A., 2020. WASP 8: the next generation in the 50-year evolution of USEPA's water quality model. Water 12 (5), 1398. https://doi.org/10.3390/W12051398.

Yang, J., Castelli, F., Chen, Y., 2014a. Multiobjective sensitivity analysis and optimization of distributed hydrologic model MOBIDIC. Hydrol. Earth Syst. Sci. 18 (10), 4101–4112. https://doi.org/10.5194/HESS-18-4101-2014.

Yang, J., Entekhabi, D., Castelli, F., Chua, L., 2014b. Hydrologic response of a tropical watershed to urbanization. J. Hydrol. 517, 538–546. https://doi.org/10.1016/J.JHYDROL.2014.05.053.

Zhi, W., Shi, Y., Wen, H., Saberi, L., Ng, G.H.C., Sadayappan, K., et al., 2022. BioRT-Flux-PIHM v1.0: a biogeochemical reactive transport model at the watershed scale. Geosci. Model Dev. 15 (1), 315–333. https://doi.org/10.5194/GMD-15-315-2022.