Received: 15 December 2021 | Revised: 11 May 2022 | Accepted: 18 May 2022

DOI: 10.1002/hyp.14612

RESEARCH ARTICLE

WILEY

Monitoring phosphorus in the tributaries of a deep lake from the perspective of the receiving water body

Giulia Valerio[1] | **Marco Pilotti**[1] | **Alessandro Scibona**[2] | **Daniele Nizzoli**[2]

[1]DICATAM, Department of Civil, Environmental, Architectural Engineering and Mathematics, Università degli Studi di Brescia, Brescia, Italy

[2]Department of Chemistry, Life Sciences and Environmental Sustainability, Università degli Studi di Parma, Parma, Italy

Correspondence
Giulia Valerio, Department of Civil, Environmental, Architectural Engineering and Mathematics, Università degli Studi di Brescia, Brescia, Italy.
Email: giulia.valerio@unibs.it

Funding information
CARIPLO Foundation, Grant/Award Number: 2015-0241

Abstract

Deep lakes under climate change are experiencing reduced deep mixing, with a consequent increase in nutrients concentration and oxygen deficit in the hypolimnion. From this perspective, greater efforts should be made to reduce the phosphorus load delivered to these lakes by their tributaries. An essential precondition for this action is a reliable estimate of the loads, which is challenging due to the marked temporal variability of the hydrologically-driven diffuse sources. In this article, we used a high-resolution phosphorus time series measured at the mouth of the main tributary of Lake Iseo and a machine learning algorithm to show the dominant role played by the acute, storm-dependent transport. The results emphasize the need to control loads strictly linked to precipitation and runoff in the drained watershed, representing 31% of the observed events but responsible for 64% of the overall load to the lake. We also proved evidence that the current obligations on nutrients monitoring miss the total phosphorus dynamics, leading to a systematic underestimation of the load conveyed by the inflows. Accordingly, we propose a sampling methodology where the timing and the methods for data interpolation are established according to the hydrological conditions in the drained watershed. Accounting for the economical and practical constraints imposed by the monitoring authorities, we propose to integrate, at the mouth of lake tributaries, a monthly manual sampling with an auto-sampler programmed to fill 2 bottles in correspondence of each high flow event. Our simulations showed that in this case the load estimation error was reduced below 1%, implying, on average, 13 field surveys/year and 21 laboratory analyses/year only. It is reasonable to expect comparable performances in hydrologically similar watersheds.

KEYWORDS

deep lake, eutrophication, inflow, load, monitoring, nutrients, phosphorus, sampling

1 | INTRODUCTION

Anthropogenic nutrient loadings to aquatic environments have resulted worldwide in a degradation of surface water quality due to increased biological productivity and decomposition rates (Wurtsbaugh et al., 2019). In lakes, this accelerated eutrophication process has been primarily driven by point and diffuse sources of total phosphorus (TP) and nitrogen (N), often further reinforced by the internal nutrient load released from the sediments (Schindler, 2006) and by the effects of climate driven changes on hydrology, precipitation patterns, water temperature, mixing regime and food webs (e.g., Moss, 2011; Trolle et al., 2011).

Despite noticeable examples of success in North America and Europe in reducing eutrophication, many eutrophication-related problems remain (McCrackin et al., 2017). Reduction in external TP loadings to eutrophic lakes has been considered the best of long-term restoring measures, which in some cases proved effective to slow down internal load too (Schindler, 2006; Schindler et al., 2016) and without which treatments within lakes have minimal lasting results (Lürling et al., 2016; Wetzel, 2001). Starting from the 70s, a rapid decrease of TP external load was obtained by reducing the point sources mainly by improving wastewater treatment and reducing phosphorus in laundry detergents. However, in many cases such effort has not yet led to the desired improvement in water quality, likely because of the presence of unaccounted TP sources such as the accumulation and subsequent release of nutrients in long-term storage pools or the increased contribution of runoff from impervious urban areas and overflowing discharges from combined sewers (Clement, 2001; Goyette et al., 2018; Jarvie et al., 2013; Pilotti et al., 2021). Accordingly, reducing eutrophication in this century will focus on the limitation of the effect of leaching from agricultural environments (Jarvie et al., 2013; Moosmann et al., 2005), of runoff over urban areas and of overflowing discharges by combined sewer weirs (CSWs) present along mixed sewers in urban areas (USEPA, 2009).

This task will be challenging for two main reasons. Firstly, changing the agricultural practices, restoring wetlands, riparian areas and channelized streams is much more difficult than dealing with the point sources (Jarvie et al., 2013; Schindler, 2006). Secondly, diffuse TP loads depend on the interaction between land use and precipitation patterns and show larger temporal variations today than in the past, when wastewater constituted the major input (Moosmann et al., 2005). In contrast to wastewater sources, surface hydrological pathways to the watercourse in agricultural environments, in-stream remobilisation of the sediments, runoff along impervious surfaces associated with urban areas and overflows from combined sewers are hydrologically-driven and discontinuous sources of TP (Withers & Jarvie, 2008). Accordingly, when these diffuse sources are involved, the TP input budget may be dominated by relatively few days of high discharge, especially in its particulate form (PP) (Correll et al., 1999). Although stream-borne PP may not be immediately bioavailable in the receiving lake environment, it feeds lake sediments, and could become a large, long-term nutrient source for aquatic biota (Abell et al., 2013; Carignan & Kalff, 1980).

Experimental studies have provided evidence of a disproportionate increase of river TP loads relative to base load during storm events (e.g., Carpenter et al., 2015, 2018). This evidence questions the suitability of some frequently adopted monitoring practices to estimate the external TP loads from a catchment (Preston et al., 1992). When estimating loads, time series of concentration of specific constituents are combined with time series of flow. Whereas continuous time series of flow are often measured with a frequency that is higher than the natural variability of the signal at most monitoring sites, continuous series of nutrients concentration are rarely available, because conventional laboratory measurements of water chemistry are time consuming and expensive. Moreover, no precise guidance is provided for water quality monitoring programs aimed at computing the nutrient load delivered to a lake (Kirchner et al., 2004). In Europe, for instance, the officially available nutrient data are typically those collected by the environmental authorities to address the ecological status of a river according to the water framework directive (WFD), which recommends a minimum of four measurements per year. These low-frequency monitoring programs could be inadequate for a representative assessment of nutrient load because they likely miss almost all the storm events, thus failing to reveal the very close coupling between the hydrological and chemical dynamics (Kirchner et al., 2004) and to resolve the episodic nature of material fluxes (Abell et al., 2013). Especially in flashy streams, continuous monitoring is essential for accurate estimation of the cumulative flux (Khalil & Ouarda, 2009). In this direction, Kirchner et al. (2004) proposed field-deployable auto-analysers as one of the major technologies for continuous field measurements of individual chemical constituents. However, the overall cost of auto-analysers acquisition and management would make prohibitive to perform continuous measurements at all sampling locations of a regional monitoring network. Accordingly, to extrapolate the high-frequency dataset to longer periods and more stations, some authors used these data to calibrate a physically-based model (Vilmin et al., 2018). If particulate phosphorus is the prevailing TP form, an alternative solution to auto-analysers is represented by surrogate measures such as turbidity, which can be easily measured with high frequency in situ and which may have potential for generating high frequency estimates of TP concentrations after a proper site-specific calibration (Irvine et al., 2019; Robertson et al., 2018). These methodologies were successfully used to measure nutrients at high frequencies in different experimental sites, confirming the relative role of chronic storm-independent and acute storm-dependent contributions on the nutrient load formation (Jordan et al., 2007; Ockenden et al., 2016). They also provided insights on the mobilization and transfer dynamics of nutrients during storm events (Bowes et al., 2015; Jordan et al., 2005; Outram et al., 2014), which are essential to properly designing management strategies in the catchment (Bende-Michl et al., 2013), to determine the optimal timing and frequency of grab samples (Jones et al., 2012; Moatar & Meybeck, 2005; Nava et al., 2019; Vilmin et al., 2018), to quantify nutrient retention in highly flushed reservoirs (Kong et al., 2019) and to calibrate regressive models relating P responses to flow and precipitation (Abell et al., 2013; Rozemeijer et al., 2010). Nevertheless, the unavoidable site-specific character of these studies, which have generated recommendations specific to constituents, watersheds and objectives, makes it difficult to extrapolate to hydrologically different sites an acceptable sampling frequency a priori (Jones et al., 2012).

In this article, we investigate the temporal variability of the TP loading delivered to Lake Iseo (Lombardy, Northern Italy), making use of a high-resolution time series measured by an auto-analyser at the main inflow mouth, and we propose the use of a machine learning algorithm to extrapolate the data over a longer period. Lake Iseo is an emblematic case study of a possible evolution of the other lakes in the subalpine area, where the hypolimnetic waters are suffering a

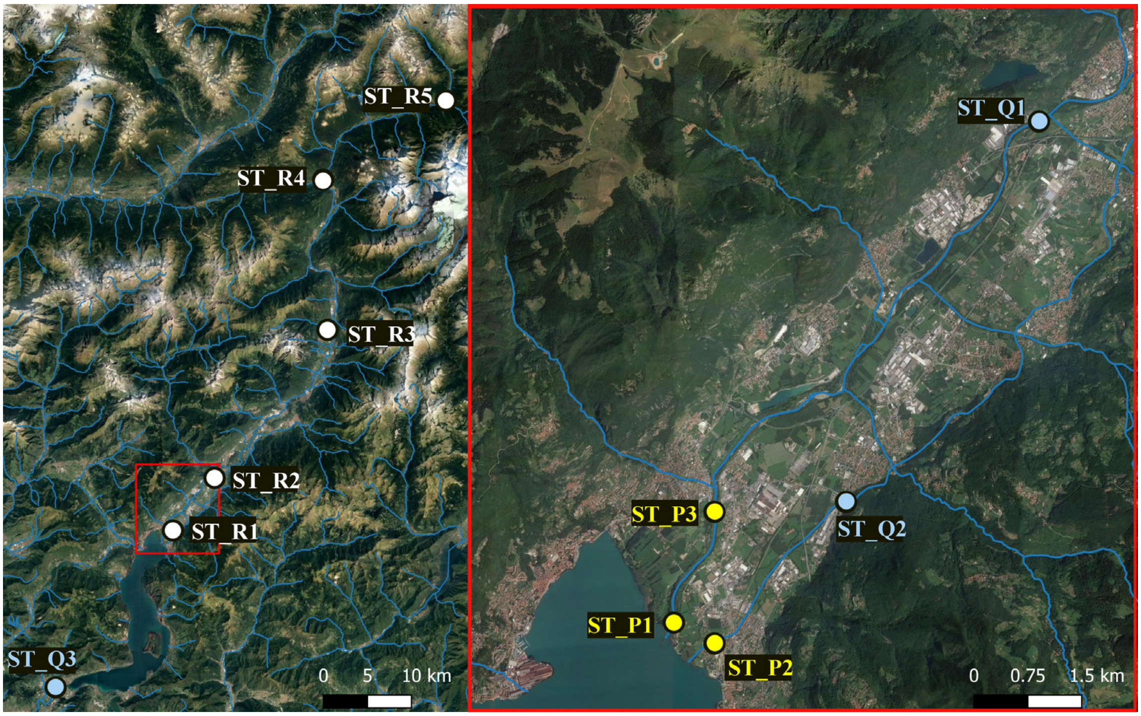


FIGURE 1 Outline of the measurement stations. ST_R refers to rain stations, ST_P to phosphorus stations and ST_Q to discharge stations. The area within the red frame on the left is shown enlarged on the right.

reduced deep mixing due to climate change, with nutrient accumulation and oxygen consumption (Rogora et al., 2018; Viaroli et al., 2018). This tendency, which is present even in the lakes which best recovered from eutrophication in the past century (Salmaso et al., 2018), requires an urgent control of the external load, that must start from a proper TP monitoring at the mouth of their tributaries. Despite the huge resources invested in monitoring the quality conditions in surface water bodies according to the WDF in Lombardy (367 monitoring points in the rivers and 25 in lakes), the TP load delivered to the subalpine lakes is computed on the basis of monthly grab TP data and, so far, no physically-based study has been provided to verify the suitability of this activity. Accordingly, the aims of this work were: (1) to quantify TP load to Lake Iseo and study its relation with hydrological conditions; (2) to shed light on the relative role of the different modes of transport of phosphorus and the consequent implications in terms of effectiveness of restoration measures in the drained catchment; (3) to compare different monitoring strategies for TP load estimation and the associated bias; and (4) to propose a monitoring strategy that provides reliable load estimates with a sustainable sampling frequency, which could also be replicated in similar watersheds.

2 | MATERIALS AND METHODS

2.1 | Field site

River Oglio and the artificial Canale Industriale are the main tributaries of Lake Iseo, a 256 m deep and 61 km^2 large lake located in the subalpine area of Italy (Figure 1). River Oglio drains Valle Camonica, a large mountain watershed (1439 km^2, constituting 83% of the whole area drained by the lake). Total population amounts to 111 000 inhabitants (77 inh./km^2) with an overall potential TP load of 76 tons/year. Urban wastewater treatment is still incomplete and treated by 61 plants (for a total treatment capacity of 193000 equivalent inhabitants), of which only four serve more than 10000 equivalent inhabitants. The TP load exported by wastewater plants amounts to approx. 17 tons TP/year according to the data provided by the Regional Agency for Environmental Protection (ARPA) of Lombardy Region (ARPA-Sireacque 2014). Agricultural use accounts for 9% of the watershed surface, mainly for pastures or rough grazing (21.4%). 80% of the urbanized area and 70% of the agricultural surface are located in a thin floodplain area within 0.5 km from the course of the main river and lateral streams. Garibaldi et al. (1998) provided a first picture of the external loads delivered to the lake until the mid 90s, characterized by a large overall value (91 tons TP/year) and a dominating role played by the diffuse sources (55%). The authors estimated that the interventions on the sewage connections and on the wastewater treatment plants would have only reduced the load by 7%.

The effects of such a strong external pressure on a relatively small lake have been amplified by the reduced vertical mixing and has determined a progressive worsening of the water quality. Lake Iseo was described as oligotrophic in the first limnological study (Bonomi and Gerletti, 1967), showing an oxygen content of deep water between 7.8 and 8.2 mg/L and an average value of reactive phosphorus at about 6 mg/m^3. Since the 80s, a shift to meromixis has occurred in Lake Iseo, determined by the accumulation of solutes from biomass decomposition, in combination with climatic factors. The deeper water quality has thus deteriorated to the point that Lake Iseo now shows the most worrying trophic condition of all large deep lakes in Northern Italy (Rogora et al., 2018), with an average total phosphorus concentration of 70 mg/m^3, ranging from 10 to 150 mg/m^3 along the water column, and a permanent anoxia below 100 m (Lau et al., 2020). A recent field analysis by Lau et al. (2020) showed that the release of phosphorus by the sediments is fed by sedimentation from the upper layers of the lake. The mobile phosphorous pool in the upper sediments is relatively small and can sustain the actual release for only ~1.6 years without constant renewal. These results overturn the widespread idea that the reduction of the external load is useless in meromictic lakes and indicates the reduction in external loadings as a fundamental long-term corrective measure.

2.2 | Hydrological data

Hourly and daily averaged discharge are available at ST_Q1 and ST_Q2, respectively (see Figure 1). In ST_Q1 the Oglio river discharge (Q1) is obtained through a stage-discharge relationship from an in-stream level logger managed by ARPA. In ST_Q2, along Canale Industriale, the discharge (Q2) is computed from the production measured at an hydropower station. Additionally, the daily average outflowing discharge is measured at ST_Q3 in correspondence with the lake outlet. Together with the level of the lake, the outflowing discharge is used by Consorzio dell'Oglio to provide the whole inflow to the lake (Qin) on the basis of a mass balance. Hourly rainfall data are made available by ARPA at five permanent stations in the watershed (ST_R1-R5 in Figure 1).

For the purpose of connection with typical TP concentration, hydrological conditions at ST_Q1 were classified according to the approach by Haygarth et al. (2004) in different event types, associated to a different nutrient response at ST_P1. A first type of event (E1) is characterized by high discharges in the river, independent from the rainfall. In Lake Iseo these events usually occur once per year, when the water diverted to the Canale Industriale is delivered for about 2 weeks to the Oglio river due to the interruption of the hydropower production for maintenance works. A second event type (E2) is a natural flood event in the Oglio river, when the recorded peak discharge was larger than the 95th percentile of the whole time series ($Q_{95} = 40$ m^3/s); we identified the beginning of the rising limb of the hydrograph when the recorded average rate of change of the discharge exceeded 2 m^3/s/h in the following 4 h. The end point of the event was instead fixed at a time that was followed by at least 6 h of discharge lower than 18 m^3/s and with a recorded rate of change lower than 4 m^3/s/h. In the instance of multi-peaks cases, the corresponding hydrographs were grouped into a single event if the peaks were less than 12 h apart, also considering that the concentration time of ValleCamonica is approximately 10 h. Finally, another event type (E3) is related to limited rainfalls occurring in the watershed, which do not induce a detectable water rise in the river but could locally trigger CSW overflows in the urban areas and induce leaching from agricultural fields and farms. In this case, we refer to the average rain intensities measured at the 5 ST_R1-5 stations, neglecting events with cumulative rain height lower than 2 mm and assuming that the event ends 6 h after the end of the average rainfall. The remaining periods (E4) are dry ones with constant base flow.

2.3 | Phosphorus concentration in river water

The official analysis of TP concentration in the Oglio river is accomplished by ARPA at station ST_P3 (see Figure 1) on a monthly basis, according to the sampling protocol imposed by the Lombardy region at the mouth of lakes tributaries. In all the other upstream river stations, the surveillance monitoring is accomplished once every 3 months. On 2017 we installed a high-resolution auto-analyser (3S-Colorimeter by 3S ANALYZERS s.r.l.) at station ST_P1 (see Figure 1), upstream of the Oglio mouth. The choice of the ST_P1 location was based on the consideration that a CSW and the outlet of an important wastewater treatment plant are present downstream of the ST_P3 location. This instrument extracts stream water samples on an hourly basis and determines their TP content by colorimetry after the conversion of the organic fraction to orthophosphate with a nominal accuracy of ±5 ppb. Organic phosphorus is converted to orthophosphate through digestion with sulphuric acid followed by photochemical oxidation to orthophosphate with UV irradiation and sodium persulfate. Orthophosphate is determined following the bluindophenol method. The monitoring station consists of a sampler, a homogenization, a digestion and a colorimetry unit. The measured data are displayed on a screen, saved in a datalogger, and sent in real-time to a server. The reagents are locally kept in a refrigerator. Overall the auto-analyser measured 6496 data, resulting in 271 days of measurement. At the same time and location, an automatic probe (model CS547A Campbell Scientific) monitored the water temperature and conductivity with an accuracy of 0.1°C and ± 5%, respectively.

Even though the reliability of online chemical analysis is continuously improving, currently it can hardly match the dependability of laboratory analyses on grab samples. For this reason, there is a general agreement in the scientific community on the need for grab samples collection on a periodic basis, to provide benchmarks against instrument drift, cross-checks for unreliable readings detection, and backup measurements during online system failures (Kirchner et al., 2004). For this purpose, three replicated water samples were collected with a Ruttner like bottle at ST_P1 and at the Canale Industriale mouth ST_P2 (see Figure 1) in the period between May 2016 and October 2018. Samplings were conducted monthly during base flow and their frequency was increased during high flow events to take into account the impact of flow variability on water chemistry. Supplementary event-based samples were collected in the Oglio river in correspondence of 3 rain events through an automatic sampler ISCO 3700 equipped with 24 ½-L bottles at a frequency of 1/h. Overall, 112 samples were collected and analysed over the 3-year sampling activity.

In order to assure a high-quality standard and the reproducibility of TP concentrations, the sampling protocol, sample handling and

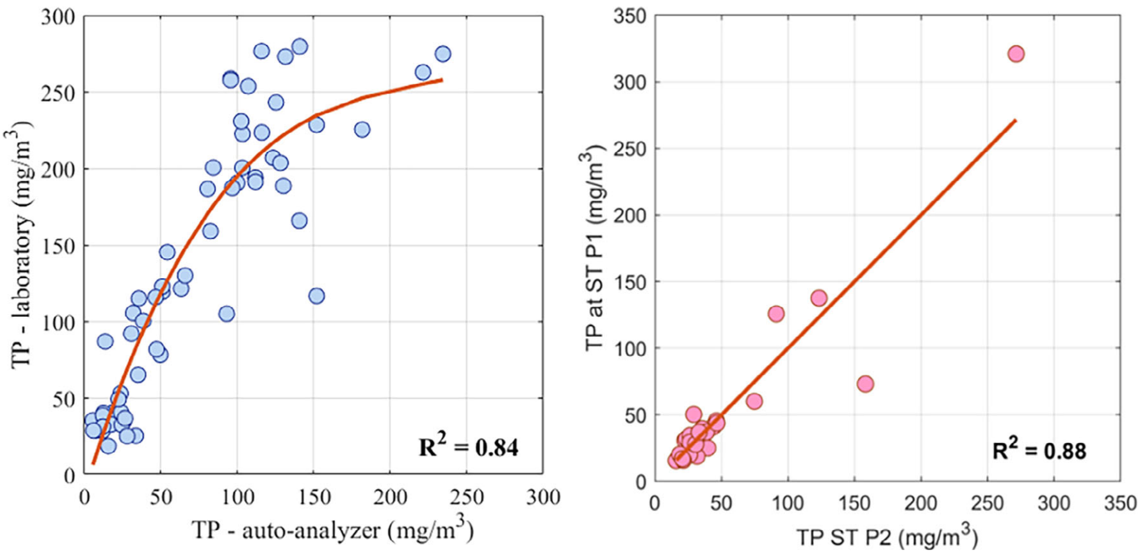


FIGURE 2 Left: Calibration of a regression curve to correct the field data. Right: Correlation between the laboratory measurements at Oglio (ST_P1) and Canale Industriale (ST_P2).

analysis were performed according to standard methods (APHA, 2012; APAT CNR IRSA 2003). After sampling, an aliquot for each replicate was immediately filtered (Whatman GF/F) and stored in glass vials for the determination of total dissolved phosphorus (TDP). A second aliquot for each replicate was filtered on pre-weighed filters (Whatman GF/F), with the material collected on filters being subsequently analysed for the measurement of total PP. TP was estimated as TDP + PP. TDP and PP were analysed within 6 h after collection with standard spectrophotometric (Perkin Elmer, Lambda 35) methods as soluble reactive phosphorus after alkaline peroxydisulfate oxidation and autoclave heating at 120°C for 1 h (Valderrama, 1981). In all samples, TP concentrations were significantly greater than the respective assay detection limits (20 μg/L). Precision (5%) at the relevant concentration ranges was estimated as the relative standard deviation of three replicate analyses of the site water.

Following the recommendation suggested by Jordan et al. (2005) in a similar study, we calibrated the measurements provided by the autoanalyser using laboratory data and, consistently with Jordan et al. (2005), we observed an average 45% underestimation of the field data with respect to the laboratory ones likely due to an incomplete oxidation of the particulate component or to the presence of heavy particulates which might be not sampled by the auto-analysers. In order to account for these limitations, we fitted a third degree equation between the two time series ($R^2 = 0.84$). The resulting correlation curve was applied to the continuous time series (TP, mg/m^3), making it consistent with the laboratory results. In the following, we will always make use of the values of TP corrected according to (1).

$$TP^* = 0.000021(TP)^3 - 0.0138(TP)^2 + 3.2366(TP) - 11.9057 \quad (1)$$

Considering that Canale Industriale stems from a diversion of the Oglio river about 20 km upstream of its mouth, we compared the measurements accomplished in the Oglio river and in the Canale Industriale. The two datasets were well correlated ($R^2 = 0.88$, cfr. Figure 2). Therefore, to estimate the phosphorus load to the lake, we assumed that the TP concentrations measured by the auto-analyser at Oglio river were the same in the Canale Industriale.

2.4 | Empirical model

In order to extend the measured TP series over a continuous and longer period and to provide a more reliable picture of the role of the annual and inter-annual variability of the hydrological variables on the TP dynamics, we built a machine learning model that relates the measured TP concentrations at the river's mouth with the E1-E4 hydrological conditions in the watershed.

For this purpose, we followed the approach by Rozemeijer et al. (2010) who identified a set of hydrological events in the measured time series and statistically related the variables characterizing the nutrient-response and the hydrological characteristics, in turn calibrating a model of the NO_3 and TP concentration in wet periods based on the rainfall discharge and groundwater characteristics. In the same line of thought, for each of the E1–E4 event types we firstly identified an analytical relationship of the TP temporal trend, that is, the pollutogram, parametric of several variables ("response variables"), marked in the following equations with the apex "M". In dry periods with low flow (E4), the phosphorus does not show any relevant temporal variation, therefore we assumed it equal to the average measured in the E4 dry periods ($TP_0 = 52.1\,mg/m^3$).

$$TP^{E4} = TP_0 \quad (2a)$$

TABLE 1 Explanatory and response variables included in the model.

Response variables	
TP^M_{max}	Peak concentration of TP (mg/m^3)
t^M_{max}	Time of peak concentration of TP (days)
t^M_{50}	Time when the maximum increase in the TP concentration with respect to the base value decreases below 50% (days)
TP^M_{ave}	Average concentration of TP (mg/m^3)
Explanatory variables	
H_{cum}	Cumulative precipitation height (mm)
I_{max}	Maximum rain intensity (mm/h)
dQ	Maximum increase of the discharge with respect to the initial value (m^3/s)
Q_{ave}	Average discharge (m^3/s)
dist_ev2	Elapsed time from the previous E2 event (days)
dist_ev3	Elapsed time from the previous E3 event (days)
D	Duration of the event (days)
$time\_Q_{max}$	Time of maximum discharge (days)
$time\_I_{max}$	Time of maximum rain intensity (days)
time_Q50	Time when the maximum increase in the discharge with respect to the initial value decreases below 50% (days)
Functional relationships	
$TP^M_{max} = f(H_{cum}, I_{max}, dQ, dist\_ev2, dist\_ev3)$	
$t^M_{max} = f(time\_Q_{max}\text{-}t_0, time\_I_{max}\text{-}t_0)$	
$t^M_{50} = f(D, time\_Q_{max}\text{-}t_0, time\_I_{max}\text{-}t_0, time\_Q50\text{-}t_0)$	
$TP^M_{ave} = f(H_{cum}, I_{max}, Q_{ave}, dist\_ev2, dist\_ev3)$	

We assumed a constant TP response also in dry periods with high flow (E1). Conversely to E4, this value may change from one E1 event to the other, depending on the hydrological conditions that are present in the period when the water is diverted from the Oglio river to the Canale Industriale.

$$TP^{E1} = TP^M_{ave} \tag{2b}$$

Finally, in correspondence of the E2 and E3-type events, we observed a TP response that well fits a linear increase followed by an exponential decay, in agreement with Rozemeijer et al. (2010). We thus modelled the phosphorus dynamics in time t as:

$$TP^{E2,E3} = \begin{cases} TP_0 + \dfrac{TP^M_{max} - TP_0}{t^M_{max} - t_0}(t - t_0) & t_0 < t < t^M_{max} \\ TP_0 + \left(TP^M_{max} - TP_0\right) e^{-\ln 2 \frac{\left(t - t^M_{max}\right)}{t^M_{50} - t^M_{max}}} & t \geq t^M_{max} \end{cases} \tag{2c}$$

where t_0 is the starting time of the event, while TP^M_{ave}, TP^M_{max}, t^M_{max} and t^M_{50} are the response variables of the nutrient concentration detailed in Table 1.

We thus looked for a model to link the response variables with the explanatory ones listed in Table 1, that would characterize the temporal evolution of the hydrological conditions in correspondence with the same event and that were computed from the hydrographs measured at ST_Q1 and from the hyetographs measured at ST_R1-5. There is a wide range of candidate models that can be used for such an analysis: for example, generalized regression models, time series regression models and non-parametric regression models (for a comparison, see Linder et al., 2017). In this study, we successfully employed the Random Forest (RF), a flexible non-linear and non-parametric machine learning algorithm. RF has proven to be among the best performers in many fields of application (Desai & Ouarda, 2021). Each decision tree in the ensemble is trained on a random bootstrap sample of the data (typically two thirds of the observed data) and the results are tested against the remaining observations ("out-of-bag" subset, OOB). This enables RF to estimate the generalization error without using an external test set (Rodriguez-Galiano et al., 2014). The function *RandomForestRegressor* of the Python package Scikit-learn (Pedregosa et al., 2011) was used in this study. The number of variables randomly selected as candidates for each node split was set to the logarithm of the total number of features, while 500 were selected as the number of trees to be fitted.

The RF analysis was set up during the observation period, using the overall 137 events to model the average phosphorus concentration in dry periods and the 50 E2 and E3 events to model the response variables in wet periods. For each response variable, the model was first trained on the entire set of the explanatory variables, in order to identify the ones with the highest explanatory power: the importance of a given variable is determined by applying the feature importance analysis, that computes how much each feature contributes to the decrease of the weighted impurity in each decision tree. Once the most important hydrological variables were selected, the model was then re-built for each response variable and the modelling accuracy was assessed through the coefficient of determination (R^2) computed using the OOB predictions.

The good modelling performance allows to use the resulting model in order to predict the TP time series at ST_P1 in any period of time when the rainfall in the watershed and the discharge at ST_Q1 are available, independently from the functioning of the auto-analyser. In this contribution, we used the model to predict the TP time series from January 2016 to December 2020, providing a more reliable picture of the role of the annual and inter-annual variability of the hydrological variables on the TP dynamics.

2.5 | Computation of the actual TP load to the lake

The hourly ($\Delta t = 1h$) time series of the n measured and modelled TP concentrations (c_j) were used to estimate the actual TP load (L) to the lake delivered by the Oglio river and the Canale Industriale, which provide on average 83% of the overall inflowing waters to the lake:

$$L_{\mathrm{Oglio}} = \sum_{j=1}^{n} Q1_j c_j \Delta t \tag{3a}$$

$$L_{\mathrm{Canale}} = \sum_{j=1}^{n} Q2_j c_j \Delta t \tag{3b}$$

Given the absence of TP measurements in the other minor tributaries, we therein assumed the same TP concentrations of the Oglio river in order to provide an estimation of the overall TP load to the lake:

$$L = \sum_{j=1}^{n} \mathrm{Qin}_j c_j \Delta t \tag{3c}$$

where Q_{in} is the overall discharge delivered to the lake by the tributaries, provided by Consorzio dell'Oglio.

We compared these results with the TP load that can be estimated on the basis of N concentrations C_i of the grab samples, which were manually collected at lower-frequency. Among all the possible methodologies (see for example Quilbé et al., 2006) we computed the TP load associated to each discrete sampling with the following two approaches:

$$\begin{cases} L_i = C_i \left(\sum_{j=u_i}^{f_i} Q_j \Delta t \right) \\ L = \sum_{i=1}^{N} L_i \end{cases} \tag{4a}$$

$$\begin{cases} \overline{C} = \left(\dfrac{\sum_{i=1}^{N} Q_i C_i}{\sum_{i=1}^{N} Q_i} \right) \\ L = \overline{C} \sum_{j=u_1}^{f_N} Q_j \Delta t \end{cases} \tag{4b}$$

where Q_j is the hourly time series of either the Oglio, Canale or the overall discharge to the lake depending on the evaluation under consideration. In Equation (4a), which implements the period-weighted method (PWM in the following), C_i is the concentration measured during the *i-th* sampling. In the same equation (C_i was kept constant for the time interval $\frac{\Delta t_{i-1,i} + \Delta t_{i,i+1}}{2}$, where $\Delta t_{i-1,j}$ and $\Delta t_{i+1,j}$ are the time intervals between the *(i-1)-th*, *i-th* and *(i+1)-th* samping. During the same time interval, the delivered load L_i is obtained using the hourly discharge ranging between Q_{ui} and Q_{fi}, being *ui* and *fi* the end-points of the discharge values corresponding to the same time interval in the measured discharge series. In Equation (4b), which implements the flow-weighted mean concentration method (WAM in the following), Q_i is the discharge flowing at the *i-th* sampling time and $\overline{C}$ is the discharge-weighted mean of the concentration during the whole measurement period.

3 | RESULTS

3.1 | Temporal variability of hydrological conditions and TP concentration

Figure 3 provides a picture of the hydrological conditions from February 2017 to September 2018 during the operation of the auto-analyser. Due to operational problems, the auto-analyser provided high-resolution data for only 271 days, with values ranging from 4 to 489 mg/m^3 and with the 10th, 50th and 90th percentile equal to 20.8, 53.4, and 144 mg/m^3, respectively. The operational period of the auto-analyser is shown in Figure 3a–c with black dots superimposed on the continuous lines of discharge, conductivity and rainfall. The TP time series shown in Figure 3d exhibits a strong variability, with peaks occurring frequently and rapidly in response to rainfall events and to large discharges.

The four classes of hydrological events were characterized by a clear difference in TP dynamics. In Figure 4 we show 4 representative types. During E4 events TP concentration remains relatively constant (average value = 52.1 mg/m^3, ranging between 4.0 and 252.9 mg/m^3). During E1 events, characterized by an increased discharge of the Oglio river because of the closure of the Canale Industriale, the larger shear stresses on the riverbed mobilize the sediments, enhancing the suspension of particulate phosphorus, thus keeping the TP in the river higher than 100 mg/m^3 for the whole duration of the event (median value = 218.1 mg/ m^3, ranging between 38.5 and 501.3 mg/m^3). The reduced value of the conductivity is due to the lower salinity of the water ordinarily conveyed by Canale Industriale. During E2 events, typical of river floods, the temporal trend of the TP (median value = 90.7 mg/ m^3, ranging between 6.4 and 311.5 mg/m^3) matches the hydrograph one, having a peak in correspondence of the maximum discharge. It is interesting to observe the opposite decrease of the normalized conductivity, which indicates a dilution of the river flow with less saline rainfall waters. Finally, during the rain events that do not induce any relevant increase of the discharge or variation of the conductivity (E3 events), the trend of the TP (median value = 100.0 mg/ m^3, ranging between 5.2 and 277.0 mg/m^3) appears to be strongly correlated with rainfall, with maximum TP concentrations following the maximum rain intensity.

3.2 | Estimation of the actual TP load to the lake

3.2.1 | Estimation based on the TP data measured by the auto-analyser

The 271-day time series was used to estimate the annual phosphorus load to the lake from all the tributaries, resulting in 129.6 tons TP/year. To quantify the role of the 4 different hydrological periods to the load transfer to the lake, we divided the dataset in 4 separate time series and compared their duration and contribution to the overall TP load. Table 2 highlights the relative duration and load, with the E4 dry periods covering 65% of the whole period and providing only

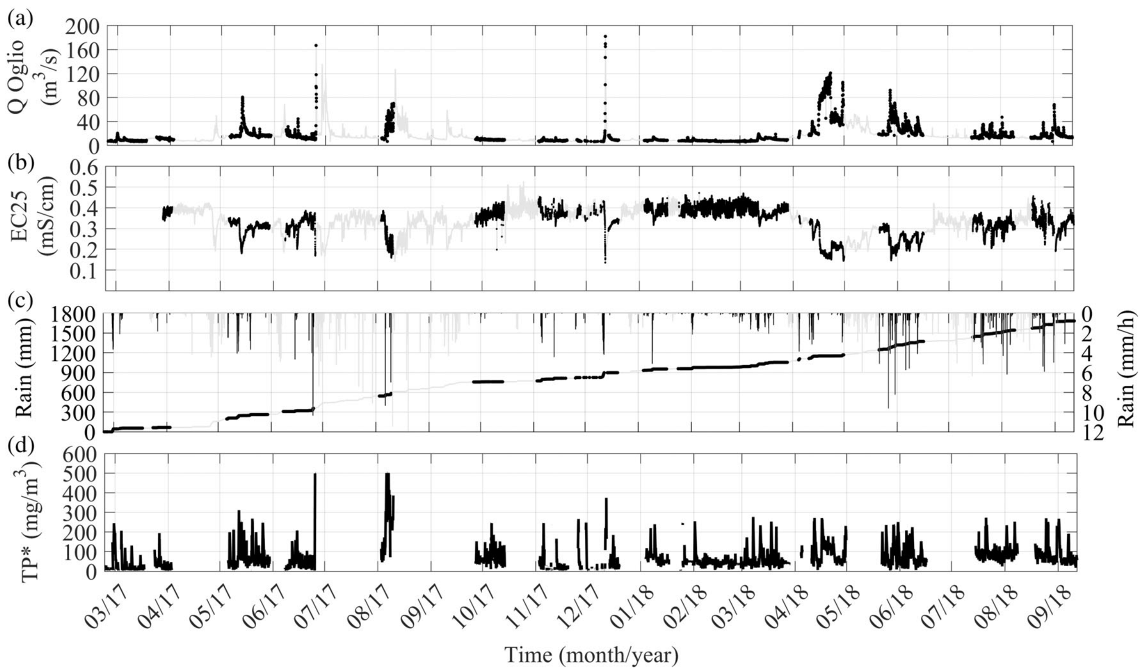


FIGURE 3 Time series of: (a) discharge measured at ST_Q1 in Oglio river, (b) 25°C conductivity at ST_P1, (c) averaged rain intensity (right axis) at the stations ST_R1, ST_R2, ST_R3, ST_R4 and ST_R5 and related cumulative height (left axis), and (d) TP concentration measured at ST_P1. The grey lines indicate the whole time series, while the black dots mark the values that were measured in a period when the auto-analyser was correctly working at ST_P1.

33% of the TP load, while the wet E2 + E3 periods covering 30% of the whole period and providing more than half of the TP load. A large relative contribution is also from E1 periods, which are associated with 15% of the TP load in only 5% of the whole duration.

3.2.2 | Estimation based on the monthly data sampling

The obtained annual phosphorus load was firstly compared with the one obtained from the monthly-based samplings measured by ARPA at ST_P3. In the period February 17 to October 18, ARPA collected 21 samples, at an average distance of 26 days, ranging between 8 and 311 mg/m^3, and having an average value of 44.14 mg/m^3.

PWM and WAM provided an overall TP load of 96.2 and 73.9 tons/year, respectively, being 26% and 43% lower than the one derived from the continuous time series. The computation based on ARPA data strongly underestimated all the contributions from wet events. With reference to 4 representative periods, Figure 4 highlights that the PWM ARPA series completely misses the TP dynamics. In particular, the largest errors occur in correspondence of the high discharge events (E1 and E2), when PWM evaluation misses 30.8 tons TP/year, so almost 24% of the whole load to the lake. Conversely, the estimations during dry periods are in well agreement (45.6 vs. 43.4 tons/year during E4 events by ARPA data and the auto-analyser, respectively).

3.2.3 | Estimation based on the RF model

The RF model was run for each response variable over the measurement period, obtaining a fit between observed and modelled variables shown in Figure 5, where the reported R^2 resulted acceptable according to Moriasi et al. (2007).

As a comparison, the validated RF model was then used to estimate the TP load to Lake Iseo, distinguishing the contributions of the different event types. In the 271 days of operation of the auto-analyser the overall TP resulted 130.7 tons/year and Figure 4, fourth row of each inset, shows the measured time series of TP concentration superimposed to the RF one, with reference to some representative event types. With respect to the evaluation based on the interpolation of the ARPA values, the model succeeded in the reliable estimation of the TP load during wet periods (E2 and E3) when the main features of the TP dynamic are correctly captured (see Figure 4).

Accordingly, the RF model was extended from January 2016 to December 2020, to describe the dynamics of TP in response to a longer times series of hydrological events, extending measurements to ungauged periods. Firstly, we extracted all 4 event types with the same criteria described above. Later, we applied the validated RF model and quantified the different load contributions, as detailed in Table 3. The overall value is 128.5 tons TP/year, only 2% lower than the one obtained during the period of the field campaign. According to this evaluation, the inflows to the lake during wet periods in the watershed (E2 and E3 events), lasting on average 106 days/year,

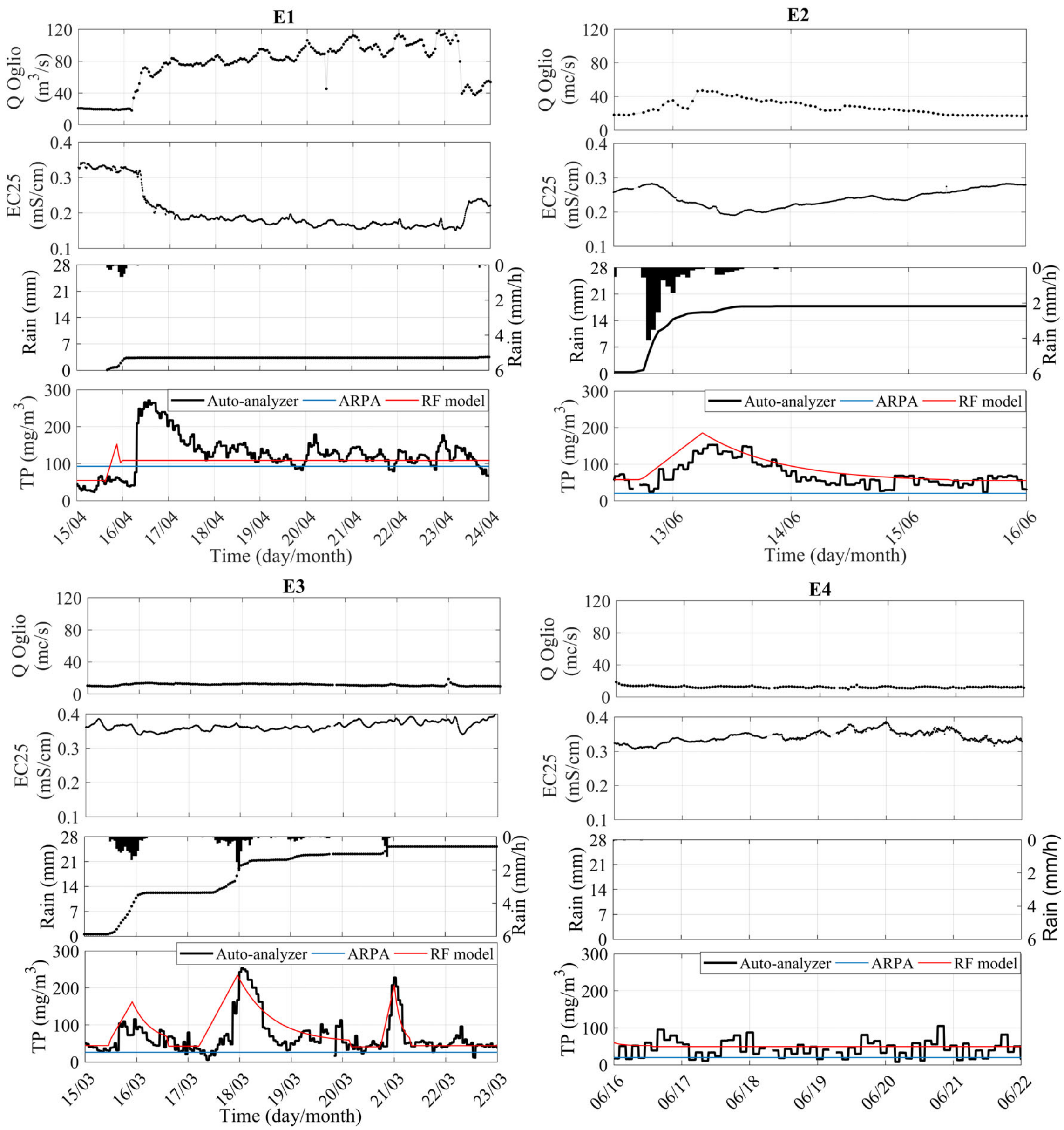


FIGURE 4 Same as Figure 3, but referred to 4 particular time periods which are well representative of event types (E1, high discharges in the river, without rainfall, E2, natural flood event, E3, limited or local rainfalls in the watershed without significative increase of discharge and E4, dry periods with constant base flow). In the last row of each set, the measured TP time series (thick black) is superimposed to that obtained from the interpolation of the sampling data by ARPA (thin blue) and from the RF model (thin red).

TABLE 2 Contribution of the identified event-types to the overall TP load during the 271 days of measurements.

	Measurements by the auto-analyser							
			Oglio river		Canale Industriale		Overall inflows	
Event type	Duration (days)	Duration (%)	TP load (tons)	TP load (%)	TP load (tons)	TP load (%)	TP load (tons)	TP load (%)
E1	13.1	5%	13.0	32%	0.0	0%	14.8	15%
E2	35.7	13%	12.4	30%	13.9	33%	27.7	29%
E3	45.6	17%	6.1	15%	9.9	24%	21.4	22%
E4	176.2	65%	9.4	23%	18.0	43%	32.2	33%

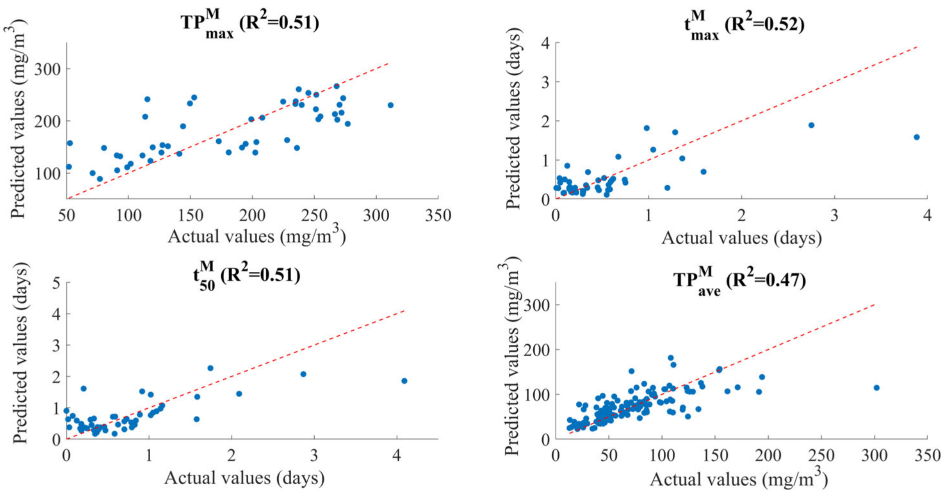


FIGURE 5 Predicted versus actual values of the explanatory variable and resulting R^2 values.

TABLE 3 Contribution of the identified event-types to the overall TP load during the 5-year period.

	Estimation by the RF model 2016–2020							
			Oglio river		Canale Industriale		Overall inflows	
Event type	Duration (days)	Duration (%)	TP load (tons/year)	TP load (%)	TP load (tons/year)	TP load (%)	TP load (tons/year)	TP load (%)
E1	40	2%	5.7	12%	0.0	0%	6.9	5%
E2	255	14%	24.0	49%	20.2	34%	47.6	37%
E3	278	15%	7.1	14%	14.1	24%	28.1	22%
E4	1242	68%	12.4	25%	24.3	41%	46.0	36%

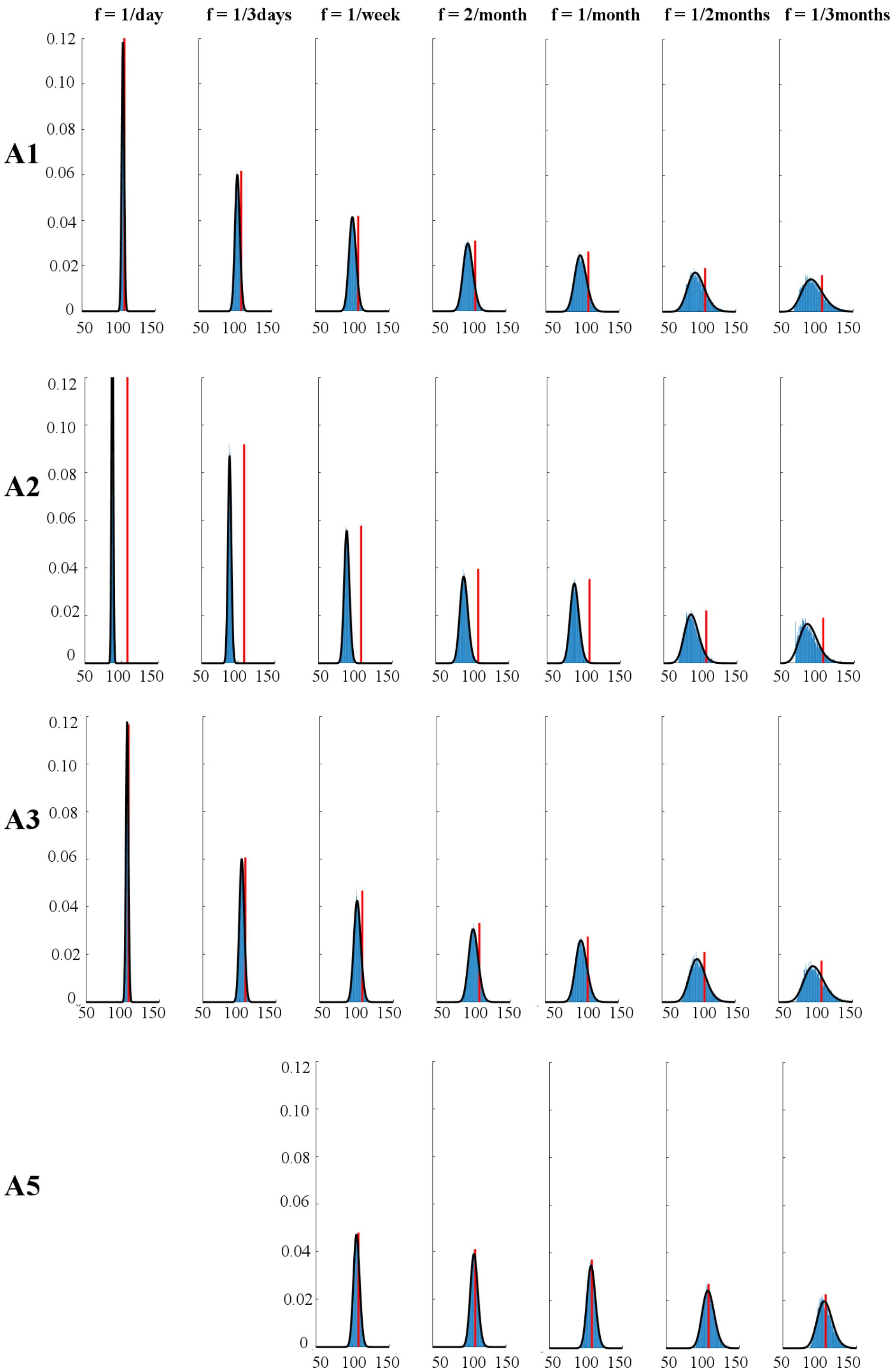
provide 59% of the TP load, with the larger contribution being associated with high-discharge events (E2 events).

3.2.4 | Accuracy of load estimation based on different discrete sampling frequencies

The non-uniform distribution of the phosphorus concentration that characterizes the measured time series highlights the importance of the sampling frequency on the accuracy of the overall load estimation. To better highlight this point, we made use of the modelled time series of TP delivered to Lake Iseo from 2016 to 2020 by the 2 main inflows, along with the measured series of discharge. We then simulated the effect of the discrete grab sampling procedure by generating 10000 different yearly series of discrete TP concentrations, each generated by randomly subsampling the measured series at an assigned frequency, ranging from one grab sample each day to one grab sample every 3 months. To simulate realistic samplings, we imposed a minimum distance of 15 days for the samples collected with frequencies lower than 1/month. Eventually, for each yearly series we computed the corresponding TP load delivered to the lake with the PWM method. In the following, we will refer to this sampling strategy as A1.

The resulting TP load distributions are reported in Figure 6 (first row), showing increasing standard deviation and decreasing mode in correspondence of lower sampling frequencies. As a statistical interpretation of the results, we fitted a log-normal pdf to each distribution (black line in Figure 6), and we used it to quantify the expected error *e* as the normalized difference between the modal value (and accordingly most probable pick) of the computed pdf and the actual load *L*. We also evaluated the probability *p* of estimating the correct load by integrating each pdf distribution in a range ± 10% around *L* (96.9–118.5 tons/year). Both *e* and *p* are reported in Table 4, together with the number of field surveys and laboratory analysis required by each sampling methodology. The errors obtained range from 2% to 14% for a daily and 1/3 months sampling frequency, respectively. The standard monthly sampling, which is applied in Italy by ARPA according to the regulation by Lombardy Region at the mouth of the main lakes' tributaries, underestimates, on average, the TP load by 10% and provides the actual load *L* ± 10% with a 52% of probability only. It is important to stress that if the field surveys are carried out in a "lazy way", limiting sampling to low-flow periods only (A2), the error *e* in the *L* load's estimation systematically increases (e.g., 19% of underestimation in A2 vs. 10% in A1 for a monthly sampling), while the probability *p* to correctly estimate L ± 10% decreases (e.g., from 52% in A1 to 7% in A2 for a monthly sampling), along with the standard deviation, due to the limited temporal variability of the TP concentration in dry periods.

FIGURE 6 Probability distribution of the overall TP load estimation delivered by Oglio and Canale Industriale (tons/year in the x-axis) based on the discrete sampling strategy with an assigned frequency. The first row (A1) makes reference to random surveys with an assigned frequency over the whole observation periods, whereas in the second (A2) the samplings are limited to low flow periods only. The third and the fourth rows show the results of the simulation of the (A3) and (A5) sampling strategies, which include additional data during the E1 and the E2 events. Each column corresponds to the same frequency of the ordinary manual samplings. The red lines mark the value provided by the whole RF time series. The curves refer to the A4 case superimposed to that of A5.



3.2.5 | Accuracy of load estimation based on optimized sampling strategies

A precision below 5% cannot be reached with a manual collection. A possible technological solution is the installation of a refrigerated auto-sampler with a sufficient amount of pre-added H_2SO_4 in storage bottles to bring pH <2 and conserve the samples after collection. We considered the installation of an auto-sampler equipped with 24 bottles, which can be programmed to fill 1 bottle/ day (see A1 results in correspondence of this frequency: $e = 2\%$, $p = 100\%$), to be recovered by the monitoring authorities during the ordinary monthly survey. In terms of feasibility, an auto-sampler is much cheaper and much easier to be installed and managed than an auto-analyser. However, the proposed strategy implies 365 additional analyses of TP in the lab.

To overcome this limitation, we simulated alternative monitoring programs which account for the hydrological conditions occurring in the river. Our hypothesis states that taking into consideration in the sampling strategies the observed hydrologically-driven variability of

TABLE 4 Performance of different sampling strategies in providing the correct TP load delivered by Oglio and Canale Industriale in terms of error *e* and probability *p* (see text for the definition of *e*, *p* and A1-A5).

		Error e (%) – Load							Probability *p* (%) – Load						
	Sampling strategy	1 day	1/3 days	1/1 week	2/ month	1/ month[a]	1/2 month[a]	1/3 month[a]	1 day	1/3 days	1/1 week	2/ month	1/ month[a]	1/2 month[a]	1/3 month[a]
A1	Fixed frequency independent from the hydrology	−2%	−5%	−7%	−9%	−10%	−12%	−14%	100%	97%	75%	58%	52%	43%	38%
A2	Fixed frequency independent from the hydrology; low-flow period only	−18%	−18%	−18%	−18%	−19%	−19%	−20%	0%	0%	1%	7%	7%	20%	23%
A3	Fixed frequency with additional samplings on E.1	−2%	−4%	−7%	−8%	−8%	−9%	−11%	100%	97%	81%	67%	60%	52%	46%
A4	Fixed frequency with additional samplings on E.1 and E.2 (1 sample/event)			−2%	−2%	−1%	−1%	−2%			95%	92%	91%	80%	70%
A5	Fixed frequency with additional samplings on E.1 and E.2 (2 samples/event)			−1%	−1%	−1%	−1%	−2%			98%	97%	95%	82%	73%

		NR lab analysis *n* (%)							NR field survey *n* (%)						
	Sampling strategy	1 day	1/3 days	1/1 week	2/ month	1/ month[a]	1/2 month[a]	1/3 month[a]	1 day	1/3 days	1/1 week	2/ month	1/ month[a]	1/2 month[a]	1/3 month[a]
A1	Fixed frequency independent from the hydrology	365	122	52	24	12	6	4	365	122	52	24	12	6	4
A2	Fixed frequency independent from the hydrology; low-flow period only	318	106	45	21	11	5	4	318	106	45	21	11	5	4
A3	Fixed frequency with additional samplings on E.1	366	123	53	25	13	7	5	366	123	53	25	13	7	5
A4	Fixed frequency with additional			61	33	21	15	12			53	25	13	7	5

TABLE 4 (Continued)

Sampling strategy	NR lab analysis n (%)							NR field survey n (%)						
	1 day	1/3 days	1/1 week	2/month	1/month[a]	1/2 month[a]	1/3 month[a]	1 day	1/3 days	1/1 week	2/month	1/month[a]	1/2 month[a]	1/3 month[a]
samplings on E.1 and E.2 (1 sample/event)														
A5 Fixed frequency with additional samplings on E.1 and E.2 (2 samples/event)			61	33	21	15	12			53	25	13	7	5

Note: In the second box, the strategies are compared in terms of average number of yearly laboratory analysis and field surveys. The yellow and pink boxes highlight the sampling strategy applied in Lombargy by ARPA and the minimum value prescribed by the WFD, respectively. The orange box is the alternative strategy proposed in this article for the Oglio river.

[a]It indicates that a minimum distance of 15 days is imposed between consecutive samplings.

the TP load could improve the actual load estimation, allowing to limit the number of additional samples. A first interesting alternative (A3) to the uniform sampling frequency monitoring program—specific to the test case considered in this article—can be obtained thanks to an additional survey during the shut-down of the Canale Industriale (E1 events). The measured value was then assumed constant for the whole duration of the shut-down and substituted to the interpolated time series for the corresponding time interval. As shown in Figure 6, the estimation of the TP load improves. Despite the continued large standard deviation of the distribution, the underestimation of the TP load associated with the monthly program would decrease from 10% to 8%, while the probability to estimate the correct L with a maximum error of 10% would increase from 52% to 60%.

To further reduce the error on the estimated TP load, we considered the possibility of integrating the samplings A3 with those collected in correspondence of E2 events in the river. This could easily be obtained by means of an auto-sampler, whose activation would be triggered either by water level or turbidity gauge over a certain threshold, or remotely on the basis of the measured discharge at the upstream monitoring stations. We thus introduced other scenarios, where the planned field surveys are completed with the maximum addition of 24 samplings during E2 events. As a first hypothesis, we considered the possibility of one sample/event only, and simulated this scenario by randomly selecting one TP value for each E2 event and assuming it to be constant for the whole duration of the event (A4). In the Oglio case this strategy would imply on average only 8 additional analyses of TP in the laboratory each year and the obtained error in the TP load estimation would be lower than 1% for the monthly frequency. The probability of estimating the correct TP value L with a maximum error of 10% would increase up to 91%.

With specific reference to the case of the Oglio river, a disadvantage of A4 strategy is that, on average, it does not exploit all the available bottles, since the number of E2 events between two regular monthly samplings is always less than 24. We thus evaluated the possibility of taking more samples during each E2 event and analysing their mixture in the laboratory. This strategy provides a more representative value of concentration for a given flood event but could allow the analysis of fewer events due to limited availability of bottles in the auto-samplers. According to the results of several configurations of frequencies and number of samples, the best performance would be obtained by means of a regular monthly manual sampling complemented by an auto-sampler programmed to fill 2 bottles in correspondence of each E2 event (A5). With this strategy and only 8 additional laboratory analyses with respect to A3 strategy, we obtained 1% of average error in the load estimation and 95% of probability to estimate the correct TP value L with a maximum error of 10%.

4 | DISCUSSION

The current experimental study allowed to quantify and update the estimates of TP load delivered to a deep subalpine lake using high temporal resolution measurements. The survey period, while not covering an entire calendar year, included the main hydrological periods such as autumn and summer, allowing for a calibration of a RF model. The modelled concentrations provide a picture of the temporal variability of the water quality over 5 years and of the different loads exported during different hydrological conditions.

In the following, we will discuss the physical reasons and the management implications at a watershed level of temporal variability of TP concentrations which arose in the results. Moreover, we use it as a rational basis to suggest different sampling strategies in order to reduce, in a cost effective way, the bias in the load estimation to a lake.

4.1 | Temporal variability of TP concentrations

The accomplished measurements show a high temporal variability of the TP concentrations and of the corresponding loads. This is not a surprise as previous studies have highlighted that watershed hydrology can have dramatic effects on TP dynamics. With reference to the Mississippi river, Royer et al. (2006) showed that nearly all nutrient export occurred when discharge was larger than the median discharge and that extreme discharges were responsible for more than 86% of the TP export. Carpenter et al. (2015, 2018) showed that most of the TP load to Lake Mendota (~74%) was delivered over a span of 29 days, during heavy rains or snow melt. With reference to an agricultural catchment, Jordan et al. (2007) showed that storm-dependent events transferred more than 90% of the TP load in 39% of the total period. In a similar environment, Ockenden et al. (2016) observed that 80% of the TP and 70% of the soluble reactive phosphorus loads were exported from both catchments during high flows. Studies conducted in the Po river in the period 2003–2007 showed that up to 40% of the total annual load of TP is exported to the Adriatic Sea with short-term flood events and in no more than 40 days (Viaroli et al., 2013). The TP load in the Po river is also subjected to inter-annual variability linked to the hydrological variability with highest peaks (up to 17000 tons) measured in 2000, a year characterized by heavy rainfall and minimum TP loads in dry years (Viaroli et al., 2013).

Our results highlighted that TP concentrations did not follow a seasonal pattern but were subjected to sudden and rapid changes accompanying the increase in flow rate. The analysis of the temporal variability of the TP concentration resulted in 2 opposed modes of transport: an almost steady transport in dry periods (E4), characterized by the lower TP concentrations versus an acute storm-dependent transport in which high TP concentrations occurred over short durations (E1, E2, E3). Over the period January 2016 – December 2020 under analysis, the acute storm-dependent transport, representing 31% of the observed events, was responsible for 64% of the overall TP load.

These trends can be explained in part by the predominantly sedimentary cycle of phosphorus which implies a strong dependence of the TP load on the sediment yield of the river (Viaroli et al., 2013). The high TP load carried during high flow events (E2) can therefore be explained by the average bed shear stress increase, that triggers erosive and resuspension processes and the transport of particulate matter (e.g., Yalin, 1977). However, the strong dependence of TP concentrations on short term rainfall events (E3), that do not significantly increase the water flow, suggests the presence of different P sources. Previous studies have shown that impervious urban surfaces and rainfall intensity are significant predictors of combined sewer overflows and consequently of the water quality of the receiving water body in a subalpine catchment (Salerno et al., 2018) and in lake Iseo sub catchments (Pilotti et al., 2021). If one considers that urbanized surfaces for about 65 km^2, mostly served by mixed sewers, are in the proximity of the Oglio River, one can expect an important contribution from CSWs during local rainfall events. Therefore, a likely explanation of the E3 events is that they heavily depend on runoff from impervious surfaces and overflows from combined sewer systems. The relevance of this contribution also shows that sediment settling in the sewer system could be a systemic problem that currently is not adequately faced. The actual knowledge of the hydraulic behaviour of CSWs along the drainage network is insufficient. An ad-hoc monitoring would be needed to quantify their contribution to the overall nutrient load and to identify the location and types of best management practices to mitigate it.

In the same line of thought, the assessment of the impact of the different types of events on the overall TP load to the lake suggests the need of consideration of the potential effectiveness of possible restoration measures in the drained catchment. Although the improvement of the wastewater treatment plants in Valle Camonica is a mandatory step in the restoration process of Lake Iseo and to reduce the dissolved P fraction that directly feeds its epilimnetic waters, our results suggest that these types of interventions will mostly decrease the concentration of TP during low-flow conditions only. Considering that the E4 events contribute to only 36% of the whole TP (according to the RF model of the 5-year period) and the actual value of 128.5 tons/year, the utopian goal of eliminating the punctual source of TP in dry periods without dealing with the causes of the E3 events will probably be insufficient to lower the load to the 30 tons/year suggested by the model of Vollenweider (1975) for this lake, considering the scenario of restoring the natural conditions described by Bonomi and Gerletti (1967).

Therefore, the action on the wastewater treatment plants is a necessary first step that will not be sufficient to stop eutrophication without an integrated approach in the entire watershed (sewer system and impervious surfaces) to control loads strictly linked to precipitation and runoff. In this perspective, also considering the potential worsening of the meromictic status of Lake Iseo under climate change (Valerio et al., 2015), we believe in the need to implement parallel actions that are aimed at reducing acute storm-dependent transport of TP to Lake Iseo during storms, such as periodic sewer flushing applied to combined sewers in dry periods, first rain tanks or constructed wetlands in correspondence of CSWs and wetlands or riparian corridors in the tributary mouth.

4.2 | Accuracy in TP load estimation

The advent of the WFD and its requirement to evaluate the ecological status of rivers has greatly increased the amount of available data of P concentration, but in our opinion the monitoring plans are not often tuned to the different hydrological phases of the watercourses in question. Our results highlight that a "hydrologically blind" grab sampling leads to a systematic underestimation of the P loads. This result is in line with the outcomes of several studies in other lakes' tributaries. For instance, a simulation study for some Great Lakes tributaries (Richards & Holloway, 1987) revealed that data from a monthly sampling program gave load estimates which were biased low by 35% or more half of the time. With reference to the quantification of the P load to Lake Champlain, Meals et al. (2013) concluded that quarterly and monthly concentration observations are generally inadequate, and even weekly observations may not be satisfactory if very accurate load estimates are required. Similarly, Cassidy and Jordan (2011), with reference to the inflows to Lough Neagh in Ireland, showed average underestimation of the P load, up to 60%, and high variability for all feasible sampling approaches independently from the grab sampling strategy (either random or storm-event triggered). On the other hand, Rozemeijer et al. (2010) for the Hupsel catchment in The Netherlands showed that an improvement of the load estimates from 63% to 5% bias for P was possible taking into account the sub-daily variability during flood measurements.

Although there is growing evidence that the acquisition frequency should be controlled by the watershed hydrology, it should be equally evident that there cannot be a single optimal acquisition frequency for every watershed. This can be easily shown considering the E2 hydrologic events, when the TP concentration is related to a rainfall event that presumably acts at the synoptic scale. In this case the time scale of the rising limb of the hydrograph (where the bed shear stress is growing in time) is provided by the concentration time T_c of the watershed. In turn, T_c is proportional to the main stream length l and inversely proportional to the average flow velocity or, according to Manning's equation, to the square root of the main stream bed slope, S_b. Accordingly, one can expect that a suitable frequency of grab sampling for an E2 event is a fraction of $\frac{1}{T_c} = K\sqrt{S_b}/l$, where K is approximately a constant. Accordingly, the optimal frequency of grab sampling will be lower for long rivers and/or for small average bed slopes. For the Oglio river $T_c \cong 10$ h and, consistently, the averaged time of maximum TP concentration foreseen by the model ($t^M_{\max}$) is 0.54 days. According to our measured data, the TP concentration along the rising limb is faithfully reproduced with a sampling frequency $f = 0.2$ [1/h]. Therefore, $fT_c \cong 6$ so that, as a first temptative guess for similar basins, one could use $f = \frac{6}{T_c}$. This value could be used as a rule of thumb for programming a continuous auto-analyser.

Finally, our simulated discrete sampling strategies clearly demonstrate that the underestimation associated with the sampling procedure grows when measurements concentrate in low flow periods. In absence of flood events, the variability in the TP concentrations is much lower, so that the resulting load is less affected by the frequency in the manual sampling (see A2 results). Paradoxically, in these periods the reduced standard deviation of the time distribution of the loads would erroneously lead to conclude that the temporal variability of the load is not relevant. Conversely, the estimation is affected by a systematic underestimation, regardless of the sampling frequency. Therefore, when field surveys are constrained to dry days, a monitoring strategy alternative to manual samplings should be seriously considered.

4.3 | Alternative sampling strategies at the inlet of a lake

According to the results shown in previous sections, sampling frequencies in rivers comparable to the minimum value imposed by the WFD to surveillance monitoring (1/3 months) are unsuitable to correctly estimate the TP load. For this reason we propose a sampling methodology specific for the tributaries of lakes, where a special attention must be dedicated to the monitoring of the cumulated nutrient load. These stations must be treated differently with respect to the riverine ones, as the measured values are essential to evaluate loads and the effectiveness of the restoration measures in the drained watershed from the point of view of the receiving water body (Jordan et al., 2005). The installation of auto-analysers in these sites comes up against very high costs both for purchase, management and complex calibration and stability procedures, therefore we do not believe that it could be a generalized choice in a short-term perspective. The statistical analysis of our modelled time series allowed to evaluate the potential benefits and constraints of alternative monitoring strategies, taking into account the economical and practical constraints imposed by real applications.

According to our results, an effective solution would be that of integrating the monthly samplings with those collected in correspondence of high-discharge events in the river by means of an autosampler. We expect that the same sampling strategy will be successful in all tributaries characterized by an acute storm-dependent transport of TP due to leaching and sediment resuspension. In particular, we recommend introducing this practice in the sampling activities at the mouth of all the tributaries of the subalpine lakes in Italy and in other similar situations, wherever a careful monitoring of the incoming nutrient load is imposed by the reduced deep mixing and the consequent accumulation of nutrients in deeper waters (Rogora et al., 2018; Viaroli et al., 2018). The number of bottles/event cannot be generalized to other watersheds, although characterized by a similar hydrological regime, due to the different transport time scales. When a site-specific evaluation based on the frequencies and duration of the hydrographs is not possible, a first attempt to extrapolate the results is made to scale it to the concentration time T_c, [h], so that the number of samples to be programmed for each E2 events could be $\frac{2}{10}T_c$.

5 | CONCLUSIONS

The WFD dictates that a good quality status should be obtained for all inland waters. Deep lakes are fundamental resources in Europe and elsewhere for which the WFD target is particularly difficult to meet. Therefore, we believe that it is fundamental to tailor suitable sampling strategies of the pollutants delivered to these fragile ecosystems, focusing the effort on the river stations at the entrance of a lake.

A field activity conducted at the mouth of the main inflows of deep Lake Iseo sheds light on the temporal variability of the TP load delivered to the lake. Over 5 years, we provided evidence of a disproportionate increase of TP loads relative to base flow during storm events, with wet or high discharge periods covering 31% of the investigated period and contributing to 61% of the load. These discontinuous sources of TP are in-stream remobilisation of the sediments, runoff along impervious surfaces associated with urban areas and overflows from combined sewers. This result emphasized that the restoration of Lake Iseo ecosystem requires an integrated management approach in the entire watershed, which needs to include the control of the diffuse loads linked to precipitation and runoff.

On a statistical basis we showed that this temporal variability leads to an unavoidable underestimation of the TP load when based on manual sampling collected at a fixed monthly (or lower) frequency, with an average error in the range between 10% and 14% and further growing when measurements are concentrated in dry periods. Accordingly, the sampling protocol imposed by Lombardy region at the River Oglio mouth, where the frequency of the surveillance monitoring was tripled (1 sample/1 month vs. 1 sample/3 months), is still unsuitable to properly quantify the TP load to lake Iseo and its evolution in time.

Our results showed that the intensification of random and "hydrologically blind" nutrient samplings could be ineffective to improve the TP load estimation at the entrance of a lake. Conversely, an improvement can be obtained by prescribing an ad-hoc protocol, where both the sampling timing and the methods for data interpolation are defined on the basis of the watershed hydrology. With regards to situations similar to the subalpine lake considered in the study, our results suggest integrating a monthly manual sampling with an auto-sampler programmed to fill 2 bottles in correspondence to each high flow event. In the case of Lake Iseo, this sampling strategy reduces the error in the load estimation to 1%, implying, on average, 13 field surveys/year and only 21 laboratory analyses/year.

FUNDING INFORMATION

This research is part of the Improving the lake Status from Eutrophy to Oligotrophy (ISEO) project and was made possible by a CARIPLO Foundation grant number 2015-0241.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are made available for scientific purposes from the corresponding author upon reasonable request from the date of publication.

ORCID

Giulia Valerio https://orcid.org/0000-0003-2287-4641

REFERENCES

Abell, J. M., Hamilton, D. P., & Rutherford, J. C. (2013). Quantifying temporal and spatial variations in sediment, nitrogen and phosphorus transport in stream inflows to a large eutrophic lake. *Environmental Science. Processes & Impacts*, *15*, 1137–1152.

Bende-Michl, U., Verburg, K., & Cresswell, H. P. (2013). High-frequency nutrient monitoring to infer seasonal patterns in catchment source availability, mobilisation and delivery. *Environmental Monitoring and Assessment*, *185*, 9191–9219.

Bonomi, G., & Gerletti, M. (1967). Il Lago d'Iseo: primo quadro limnologico generale (termica, chimica. plancton, e bentos profondo). *Memorie dell'Istituto Italiano di Idrobiologia*, *22*, 149–175.

Bowes, M. J., Jarvie, H. P., Halliday, S. J., Skeffington, R. A., Wade, A. J., & Loewenthal, M. (2015). Characterising phosphorus and nitrate inputs to a rural river using high-frequency concentration-flow relationships. *Science of the Total Environment*, *511*, 608–620.

Carignan, R., & Kalff, J. (1980). Phosphorus sources for aquatic weeds: Water or sediments?*Science*, *207*(4434), 987–989.

Carpenter, S. R., Booth, E. G., Kucharik, C. J., & Lathrop, R. C. (2015). Extreme daily loads: Role in annual phosphorus input to a north temperate lake. *Aquatic Sciences*, *77*, 71–79.

Carpenter, S. R., Booth, E. G., & Kucharik, C. J. (2018). Extreme precipitation and phosphorus loads from two agricultural watersheds. *Limnology and Oceanography*, *63*, 1221–1233.

Cassidy, R., & Jordan, P. (2011). Limitations of instantaneous water quality sampling in surface-water catchments: Comparison with near-continuous phosphorus time-series data. *Journal of Hydrology*, *405*, 182–193.

Clement, A. (2001). Improving uncertain nutrient load estimates for Lake Balaton. *Water Science and Technology*, *43*, 279–286.

Correll, D. L., Jordan, T. E., & Weller, D. E. (1999). Transport of nitrogen and phosphorus from rhode river watersheds during storm events. *Water Resources Research*, *35*, 2513–2521.

Desai, S., & Ouarda, T. B. M. J. (2021). Regional hydrological frequency analysis at ungauged sites with random forest regression. *Journal of Hydrology*, *594*, 125861.

Garibaldi, L., Mezzanotte, V., Bizzo, M. C., Varallo, A., & Mosello, R. (1998). Apporti di fosforo al sebino. Confronto fra misure sperimentali e teoriche. *Acqua Aria*, *9*, 105–110.

Goyette, J. O., Bennett, E. M., & Maranger, R. (2018). Low buffering capacity and slow recovery of anthropogenic phosphorus pollution in watersheds. *Nature Geoscience*, *11*, 921–926.

Haygarth, P., Turner, B. L., Fraser, A., Jarvis, S., Harrod, T., Nash, D., Halliwell, D., Page, T., & Beven, K. (2004). Temporal variability in phosphorus transfers: Classifying concentration–discharge event dynamics. *Hydrology and Earth System Sciences*, *8*, 88–97.

Irvine, C. A., Backus, S., Cooke, S., Dove, A., & Gewurtz, S. B. (2019). Application of continuous turbidity sensors to supplement estimates of total phosphorus concentrations in the Grand River, Ontario, Canada. *Journal of Great Lakes Research*, *45*, 840–849.

Jarvie, H. P., Sharpley, A. N., Withers, P. J. A., Scott, J. T., Haggard, B. E., & Neal, C. (2013). Phosphorus mitigation to control river eutrophication: Murky waters, inconvenient truths, and "postnormal" science. *Journal of Environmental Quality*, *42*, 295–304.

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 loads1. *Journal of the American Water Resources Association*, *48*, 1258–1275.

Jordan, P., Arnscheidt, J., McGrogan, H., & McCormick, S. (2005). High-resolution phosphorus transfers at the catchment scale: The hidden importance of non-storm transfers. *Hydrology and Earth System Sciences*, *9*, 685–691.

Jordan, P., Arnscheidt, A., McGrogan, H., & McCormick, S. (2007). Characterising phosphorus transfers in rural catchments using a continuous bank-side analyser. *Hydrology and Earth System Sciences*, *11*, 372–381.

Khalil, B., & Ouarda, T. B. M. J. (2009). Statistical approaches used to assess and redesign surface water-quality-monitoring networks. *Journal of Environmental Monitoring*, *11*, 1915–1929.

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. *Hydrological Processes*, *18*, 1353–1359.

Kong, X., Zhan, Q., Boehrer, B., & Rinke, K. (2019). High frequency data provide new insights into evaluating and modeling nitrogen retention in reservoirs. *Water Research*, *166*, 115017.

Lau, M. P., Valerio, G., Pilotti, M., & Hupfer, M. (2020). Intermittent meromixis controls the trophic state of warming deep lakes. *Scientific Reports*, *10*, 12928.

Linder, H. L., Horne, J. K., & Ward, E. J. (2017). Modeling baseline conditions of ecological indicators: Marine renewable energy environmental monitoring. *Ecological Indicators*, *83*, 178–191.

Lürling, M., Mackay, E., Reitzel, K., & Spears, B. M. (2016). Editorial – A critical perspective on geo-engineering for eutrophication management in lakes. *Water Research*, *97*, 1–10.

McCrackin, M. L., Jones, H. P., Jones, P. C., & Moreno-Mateos, D. (2017). Recovery of lakes and coastal marine ecosystems from eutrophication: A global meta-analysis. *Limnology and Oceanography*, *62*, 507–518.

Meals, D. W., Richards, R. P., & Dressing, S. A., 2013. Pollutant load estimation for water quality monitoring projects. Tech Notes 8, April 2013. Developed for U.S. Environmental Protection Agency by Tetra Tech, Inc., Fairfax, VA, 21 p.

Moatar, F., & Meybeck, M. (2005). Compared performances of different algorithms for estimating annual nutrient loads discharged by the eutrophic river Loire. *Hydrological Processes*, *19*, 429–444.

Moosmann, L., Müller, B., Gächter, R., Wüest, A., Butscher, E., & Herzog, P. (2005). Trend-oriented sampling strategy and estimation of soluble reactive phosphorus loads in streams. *Water Resources Research*, *41*, W01020. https://doi.org/10.1029/2004WR003539

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*, 885–900.

Moss, B. (2011). Allied attack: Climate change and eutrophication. *Inland Waters*, *1*(2), 101–105.

Nava, V., Patelli, M., Rotiroti, M., & Leoni, B. (2019). An R package for estimating river compound load using different methods. *Environmental Modelling & Software*, *117*, 100–108.

Ockenden, M. C., Deasy, C. E., Benskin, C. M. W. H., Beven, K. J., Burke, S., Collins, A. L., Evans, R., Falloon, P. D., Forber, K. J., Hiscock, K. M., Hollaway, M. J., Kahana, R., Macleod, C. J. A., Reaney, S. M., Snell, M. A., Villamizar, M. L., Wearing, C., Withers, P. J. A., Zhou, J. G., & Haygarth, P. M. (2016). Changing climate and nutrient transfers: Evidence from high temporal resolution concentration-flow dynamics in headwater catchments. *Science of the Total Environment*, *548–549*, 325–339.

Outram, F. N., Lloyd, C. E. M., Jonczyk, J., Benskin, C. M. W. H., Grant, F., Perks, M. T., Deasy, C., Burke, S. P., Collins, A. L., Freer, J., Haygarth, P. M., Hiscock, K. M., Johnes, P. J., & Lovett, A. L. (2014). High-frequency monitoring of nitrogen and phosphorus response in three rural catchments to the end of the 2011–2012 drought in England. *Hydrology and Earth System Sciences*, *18*, 3429–3448.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., & Cournapeau, D. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Pilotti, M., Barone, L., Balistrocchi, M., Valerio, G., Milanesi, L., & Nizzoli, D. (2021). Nutrient delivery efficiency of a combined sewer along a lake challenged by incipient eutrophication. *Water Research*, *190*, 116727.

Preston, S. D., Bierman, V. J., & Silliman, S. E. (1992). Impact of flow variability on error in estimation of tributary mass loads. *Journal of Environmental Engineering*, *118*, 402–419.

Quilbé, R., Rousseau, A. N., Duchemin, M., Poulin, A., Gangbazo, G., & Villeneuve, J. P. (2006). Selecting a calculation method to estimate sediment and nutrient loads in streams: Application to the Beaurivage River (Québec, Canada). *Journal of Hydrology*, *326*(1–4), 295–310.

Richards, R. P., & Holloway, J. (1987). Monte Carlo studies of sampling strategies for estimating tributary loads. *Water Resources Research*, *23*, 1939–1948.

Robertson, D. M., Hubbard, L. E., Lorenz, D. L., & Sullivan, D. J. (2018). A surrogate regression approach for computing continuous loads for the tributary nutrient and sediment monitoring program on the Great Lakes. *Journal of Great Lakes Research*, *44*, 26–42.

Rodriguez-Galiano, V., Mendes, M. P., Garcia-Soldado, M. J., Chica-Olmo, M., & Ribeiro, L. (2014). Predictive modeling of groundwater nitrate pollution using random forest and multisource variables related to intrinsic and specific vulnerability: A case study in an agricultural setting (southern Spain). *Science of the Total Environment*, *476*, 189–206.

Rogora, M., Buzzi, F., Dresti, C., Leoni, B., Lepori, F., Mosello, R., Patelli, M., & Salmaso, N. (2018). Climatic effects on vertical mixing and deep-water oxygen content in the subalpine lakes in Italy. *Hydrobiologia*, *824*, 33–50.

Royer, T. V., David, M. B., & Gentry, L. E. (2006). Timing of riverine export of nitrate and phosphorus from agricultural watersheds in Illinois: Implications for reducing nutrient loading to the Mississippi River. *Environmental Science & Technology*, *40*, 4126–4131.

Rozemeijer, J. C., van der Velde, Y., van Geer, F. C., de Rooij, G. H., Torfs, P., & Broers, H. P. (2010). Improving load estimates for NO3 and P in surface waters by characterizing the concentration response to rainfall events. *Environmental Science & Technology*, *44*, 6305–6312.

Salerno, F., Viviano, G., & Tartari, G. (2018). Urbanization and climate change impacts on surface water quality: Enhancing the resilience by reducing impervious surfaces. *Water Research*, *144*, 491–502.

Salmaso, N., Boscaini, A., Capelli, C., & Cerasino, L. (2018). Ongoing ecological shifts in a large lake are driven by climate change and eutrophication: Evidences from a three-decade study in Lake Garda. *Hydrobiologia*, *824*, 177–195.

Schindler, D. W. (2006). Recent advances in the understanding and management of eutrophication. *Limnology and Oceanography*, *51*, 356–363.

Schindler, D. W., Carpenter, S. R., Chapra, S. C., Hecky, R. E., & Orihel, D. M. (2016). Reducing phosphorus to curb lake eutrophication is a success. *Environmental Science and Technology*, *50*(17), 8923–8929.

Trolle, D., Hamilton, D. P., Pilditch, C. A., Duggan, I. C., & Jeppesen, E. (2011). Predicting the effects of climate change on trophic status of three morphologically varying lakes: Implications for lake restoration and management. *Environmental Modelling & Software*, *26*, 354–370.

U.S. Environmental Protection Agency (USEPA), 2009. National Lakes Assessment: A collaborative survey of the Nation's lakes. EPA 841-R-09-001. U.S. Environmental Protection Agency, Office of Water and Office of Research and Development, Washington, D.C.

Valderrama, J. C. (1981). The simultaneous analysis of total nitrogen and total phosphorus in natural waters. *Marine Chemistry*, *10*(2), 109–122.

Valerio, G., Pilotti, M., Barontini, S., & Leoni, B. (2015). Sensitivity of the multiannual thermal dynamics of a deep pre-alpine lake to climatic change. *Hydrological Processes*, *29*(5), 767–779.

Viaroli P, Bartoli M, Castaldelli G, Naldi M, Nizzoli D, & Rossetti G., 2013. Recent evolution and expected changes of nutrient loads in a heavily exploited watershed: the Po River, Italy. Understanding Freshwater Quality Problems in a Changing World Proceedings of H04, IAHS-IAPSO-IASPEI Assembly, Gothenburg, Sweden, July 2013 (IAHS Publ. 361, 2013).

Viaroli, P., Soana, E., Pecora, S., Laini, A., Naldi, M., Fano, E. A., & Nizzoli, D. (2018). Space and time variations of watershed N and P budgets and their relationships with reactive N and P loadings in a heavily impacted river basin (Po river, northern Italy). *Science of the Total Environment*, *639*, 1574–1587.

Vilmin, L., Flipo, N., Escoffier, N., & Groleau, A. (2018). Estimation of the water quality of a large urbanized river as defined by the European WFD: What is the optimal sampling frequency?*Environmental Science and Pollution Research*, *25*, 23485–23501.

Vollenweider, R. A. (1975). Input - output models with special reference to phosphorus loading concept in limnology. *Schweizerische Zeitschrift fur Hydrologie - Swiss Journal of Hydrology*, *37*, 53–84.

Wetzel, R. G. (2001). *Limnology: Lake and river ecosystems* (3rd ed.). Academic.

Withers, P. J. A., & Jarvie, H. P. (2008). Delivery and cycling of phosphorus in rivers: A review. *Science of the Total Environment*, *400*, 379–395.

Wurtsbaugh, W. A., Paerl, H. W., & Dodds, W. K. (2019). Nutrients, eutrophication and harmful algal blooms along the freshwater to marine continuum. *WIREs Water*, *6*, e1373, 1 of 27.

Yalin, M. S. (1977). *Mechanics of sediment transport*. Pergamon Press.

How to cite this article: Valerio, G., Pilotti, M., Scibona, A., & Nizzoli, D. (2022). Monitoring phosphorus in the tributaries of a deep lake from the perspective of the receiving water body. *Hydrological Processes*, *36*(7), e14612. https://doi.org/10.1002/hyp.14612