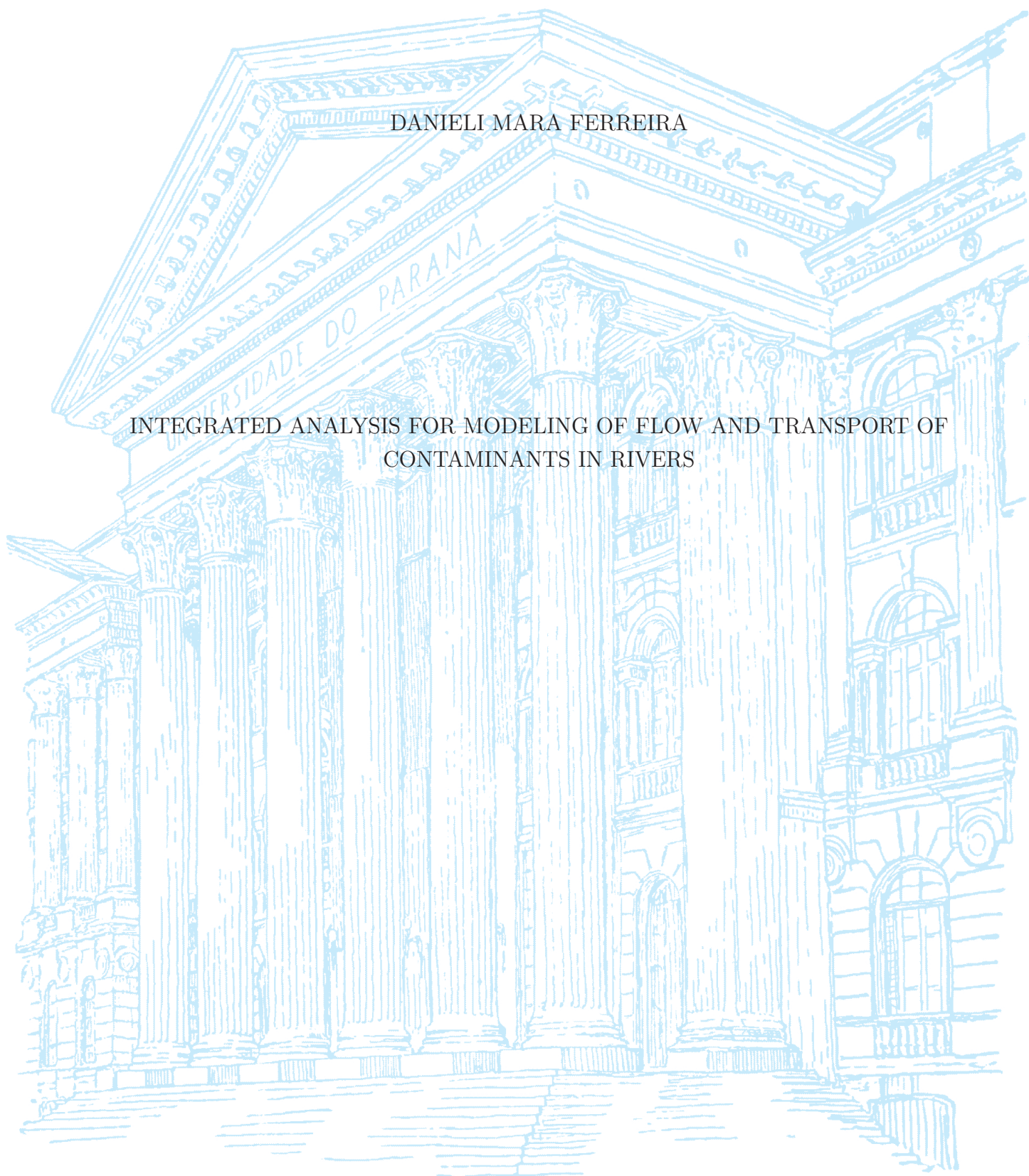UNIVERSIDADE FEDERAL DO PARANÁ

DANIELI MARA FERREIRA

INTEGRATED ANALYSIS FOR MODELING OF FLOW AND TRANSPORT OF
CONTAMINANTS IN RIVERS

CURITIBA

2019

DANIELI MARA FERREIRA

INTEGRATED ANALYSIS FOR MODELING OF FLOW AND TRANSPORT OF
CONTAMINANTS IN RIVERS

Tese submetida como requisito para obtenção do título de Doutora, no Programa de Pós-Graduação em Engenharia de Recursos Hídricos e Ambiental, Setor de Tecnologia, Universidade Federal do Paraná.

Orientador: Prof. Dr. Cristovão Vicente Scapulatempo Fernandes
Coorientador: Prof. Dr. Eloy Kaviski

CURITIBA
2019

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ENGENHARIA DE RECURSOS HÍDRICOS E AMBIENTAL da Universidade Federal do Paraná foram convocados para realizar a arguição da Tese de Doutorado de **DANIELI MARA FERREIRA**, intitulada: **INTEGRATED ANALYSIS FOR MODELING OF FLOW AND TRANSPORT OF CONTAMINANTS IN RIVERS**, após terem inquirido a aluna e realizado a avaliação do trabalho, são de parecer pela sua _____ amavação _____ no rito de defesa.
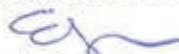
A outorga do título de Doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 24 de Abril de 2019.

CRISTOVÃO VICENTE SCAPULATEMPO FERNANDES
Presidente da Banca Examinadora

ELOY KAVISKI
Coorientador - Avaliador Interno (UFPR)

JULIO CESAR RODRIGUES DE AZEVEDO
Avaliador Interno ()

MONICA FERREIRA DO AMARAL PORTO
Avaliador Externo (UFPR)

DARREL FONTANE
Avaliador Externo (CSU)

DANIEL HENRIQUE MARCO DETZEL
Avaliador Interno (UFPR)

AGRADECIMENTOS

RESUMO

Estratégias recentes para planejamento e gestão de recursos hídricos estabelecem que os conceitos de risco e entradas variáveis devem ser avaliados, a fim de atender a múlti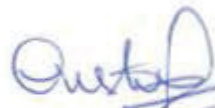plas condições. Isso fica evidente principalmente em um ambiente com mudanças climáticas e usos diversos da água e uso do solo. Nesse contexto, modelagem de vazões e concentrações em rios são estratégias válidas para prever diferentes cenários. Esta pesquisa propõe uma análise integrada para modelagem do fluxo e transporte de contaminantes em rios, baseada em simulações hidrodinâmica, de séries temporais e de qualidade da água. O primeiro módulo estima o volume e a velocidade da água, que têm impacto direto no transporte de poluentes. Séries temporais de concentrações são geradas como polutogramas sintéticos, usando técnicas baseadas em condições de fluxo, tempo e fatores estatísticos de um conjunto histórico de dados de monitoramento; o objetivo é combinar as escalas temporais das condições de contorno, uma vez que dados de qualidade da água geralmente estão disponíveis como amostras irregulares; o terceiro módulo resolve a equação de advecção-dispersão-reação, explorando as diferentes séries sintéticas como condição de contorno de montante, com um processo de calibração baseado na variação temporal dos coeficientes de transformação. Devido ao componente estocástico, um conjunto de mil concentrações diárias são geradas como séries sintéticas, e diferentes critérios são aplicados para selecionar cenários próximos aos dados observados; experimentos para converter conjuntos de dados de monitoramento de amostras discretas em séries contínuas mostram que a preservação de métricas estatísticas de dados históricos, vinculadas a análises de múltiplos cenários e representação de persistência natural, são critérios razoáveis para estimar séries temporais contínuas. Os resultados obtidos ainda evidenciam que o polutograma de entrada, usualmente não explorado em estudos semelhantes, pode ter um papel significativo em modelos de transporte de substâncias em rios em estado transiente, especialmente para predição de métricas de interesse em gestão de recursos hídricos – quartis e concentrações de 10 e 90% de ocorrência. Ao mesmo tempo, as análises conduzidas exploram um procedimento de calibração considerando o aspecto temporal; resultados indicam que essa nova perspectiva melhora simulações para transporte de diferentes parâmetros indicativos de qualidade de água. As contribuições fornecem base para posterior avaliação de sistemas fluviais ligados à dinâmica das bacias hidrográficas, com múltiplos cenários de disponibilidade de dados e condições de entrada.

Palavras-chave: Modelagem hidrodinâmica e de qualidade de água. Condições de contorno. Gerenciamento de recursos hídricos.

ABSTRACT

Recent water resources planning and management strategies state that the concepts of risk and variable inputs should be appraised in order to comply with multiple conditions. This becomes evident especially in an environment with climate change and diverse uses of water and land use. In such a context, modeling of discharges and concentrations in rivers are valuable strategies to predict different scenarios. This research proposes an integrated analysis for modeling of flow and contaminant transport in rivers, based on hydrodynamics, time series, and water quality simulations. The first module estimates water volume and velocity, that have direct impact in pollutants transport. Time series of concentrations are generated as synthetic pollutographs, using techniques based on flow conditions, time and statistical factors of a historical monitoring dataset; the objective is to match temporal scales of boundary conditions, since water quality data is usually available as irregular samples; the third module solves the advection-dispersion-reaction equation, exploring the different synthetic series as upstream boundary condition, and a calibration process based on temporal variation of transformation coefficients. Because of a stochastic component, thousand sets of daily concentrations are generated as synthetic series, and different criterion are applied to select scenarios close to observed data; experiments to convert monitoring datasets from discrete samples into continuous series show that preservation of statistical metrics from historical data, linked to multiple scenarios analysis and representation of natural persistence, are reasonable criterion to estimate continuous time series. The results obtained also evidence that the input pollutograph, usually not explored in similar studies, may have a significant role in models for transport of substance in rivers under unsteady state, specially for prediction of measures of interest in water resources management – quartiles and concentrations of 10 and 90% occurrence. At the same time, the conducted analysis introduces a calibration procedure considering a temporal aspect; results indicate that this new perspective may improve simulations for transport of different parameters indicative of water quality. The contributions lay basis for further assessment of riverine systems linked to watershed dynamics, with multiple scenarios of data availability and input conditions.

Key-words: Hydrodynamic and water quality modeling. Boundary conditions. Water resources management.

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
ADR	Advection-Dispersion-Reaction equation
AR(1)	First-order autoregressive model
BOD	Biochemical Oxygen Demand
BC	Boundary condition
CONAMA	Conselho Nacional do Meio Ambiente[1]
DO	Dissolved Oxygen
DOC	Dissolved Organic Carbon
FTCS	Foward-Time-Centered-Space
HEC-RAS	Hydrologic Engineering Center's River Analysis System
MAPE	Mean Absolute Percentage Error
N-org	Organic Nitrogen
DO	Dissolved Oxygen
PBIAS	Percent Bias
PCHIP	Piecewise Cubic Hermite Interpolating Polynomial
POC	Particulate Organic Carbon
LPOC	Labile Particulate Organic Carbon
LDOC	Labile Dissolved Organic Carbon
RPOC	Refractory Particulate Organic Carbon
RDOC	Refractory Dissolved Organic Carbon
TOC	Total Organic Carbon
R	Pearson correlation coefficient
RMSE	Root Mean Square
SIHQUAL	Simulação Hidrodinâmica e de Qualidade de Água[2]
T1 to T12, I1 to I3	Tests for synthetic series
USACE	United States Army Corps of Engineering
WQ	Water Quality
WRM	Water Resources Management

---

[1]National Environment Council
[2]Hydrodynamic and Water Quality Simulation

LIST OF SYMBOLS

Symbol	Unit	Meaning
α	–	Constant for the Lax diffusive scheme
α_5	$mg\text{-}O_2/mg\text{-}N_2$	Oxygen rate consumed by each unit of oxidized ammonia
α_6	$mg\text{-}O_2/mg\text{-}N_2$	Rate of oxygen consumed by each unit of oxidized nitrite
β_1	d^{-1}	Ammonia oxidation rate
β_2	d^{-1}	Nitrite oxidation rate
γ	–	Courant number
Δx	m	Grid interval in space
Δt	s	Grid interval in time
λ	–	Diffusion number
ρ	–	Correlation coefficient
A	m^2	Cross section area
B	m	Cross section top width
c	m/s	Celerity
C	mg/L	Cross section mean concentration
C_{10}	mg/L	Concentration of 10% of occurrence
C_{90}	mg/L	Concentration of 90% of occurrence
Cr	–	Courant coefficient
CV	%	Coefficient of variation
D	m^2/s	Dispersion coefficient
DI_i	–	Normalized distance
Di	km	Distance between the cell and the river network
E_{ns}	–	Nash-Sutcliffe coefficient
F		Term for mass transformations and external loads
$f(x)$, a_0, N, a_i, b_i, x, w	–	Parameters of Fourier series
g	9.81 m/s^2	Acceleration due to gravity
H	m	Water mean depth in the channel
I	–	Coefficient of sensitivity
$IG2$ to $IG6$	–	Monitoring sections
K_d	d^1	Deoxygenation rate
K_s	d^1	Coefficient for BOD removal by sedimentation
K_{so}	d^1	Sedimentation coefficient of organic nitrogen

K_{oa}	d^1	Coefficient for conversion of organic nitrogen to ammonia
K_a	d^1	Reaeration coefficient
K_4	$gO_2/m^2 d$	Rate for oxygen demand by the sediment
N_a	$mg\text{-}N_2/L$	Ammonia concentration
N_b	$mg\text{-}N_2/L$	Concentration of nitrite
Max	mg/L	Maximum concentration
Min	mg/L	Minimum concentration
K_p	d^{-1}	Particulate organic carbon dissolution rate
K_h	d^{-1}	Dissolved organic carbon hydrolysis
K_{L1}	d^{-1}	Settling of LPOC
K_{L2}	$gm^{-2}d^{-1}$	Resuspension of LPOC
K_{L3}	d^{-1}	Mineralization of LPOC
K_{L4}	d^{-1}	Decay of LPOC to RPOC
K_{L5}	d^{-1}	Decay of LPOC to LDOC
K_{L6}	d^{-1}	Decay of LPOC to RDOC
K_{L7}	d^{-1}	Decay of LDOC to RDOC
K_{L8}	d^{-1}	Mineralization of LDOC
K_{R1}	d^{-1}	Settling of RPOC
K_{R2}	$gm^{-2}d^{-1}$	Resuspension of RPOC
K_{R3}	d^{-1}	Mineralization of RPOC
K_{R4}	d^{-1}	Decay of RPOC to RDOC
K_{R5}	d^{-1}	Mineralization for RDOC
k	–	Constant to estimate the attenuation of diffuse sources
$P10\ to\ P90$	–	Percentiles of series
n	–	Roughness coefficient
T	–	Total number of data
pop_{reach}	$inhab$	Population from each reach
O_s	$mg\text{-}O_2/L$	Dissolved oxygen saturation concentration
Q	m^3/s	Discharge
Q_{obs}	m^3/s	Observed discharge
Q_{sim}	m^3/s	Simulated discharge
Q_{med}	m^3/s	Average of observed discharge
Q_D	L/s	Domestic effluent flow
QPC	$L/inhab.d$	Per capita flow
q	$m^3/s.m$	Lateral contribution
R	–	Pearson correlation coefficient

R_h	m	Hydraulic radius
Rs	–	Coefficient of sewage return
S_0	m/m	Bed slope
S_f	m/m	Friction slope
t	s	Time
U	m/s	Longitudinal velocity
v_L	m/s	Input velocity of lateral contribution
V	m^3	Volume
W	kg/d	Load entering the system
$p,\ w_i,\ y_i,\ s,\ x_i$	–	Smooting spline parameters
x	m	Longitudinal distance
y	m	Water depth
$y_0,\ a,\ r$	–	Parameters of the logarithmic method
z_j	–	Parameter of AR(1) model
μ	–	Mean
σ	–	Standard deviation
S	mg/L	Concentration simulated with unsteady state model
M	mg/L	Observed concentration
M'	t/d	Mean loads
IC	–	Confidence interval
$Q1$	–	First quartile
$Q2$	–	Second quartile
$Q3$	–	Third quartile
S'	mg/L	Synthetic concentration
ST	mg/L	Concentration simulated with steady state model
$P1\ to\ P3,\ r1\ to\ r3$	–	Parameters of the calibration strategy
$v1\ to\ v9$	–	Calibration experiments

CONTENTS

1 Introduction

Research is formalized curiosity. It is poking and prying with a purpose.
— **Zora Neale Hurston**

Efforts to understand and predict water quality dynamics have increased over the years. Events such as floods and droughts, contamination of drinking water and stricter laws have intensified public awareness for proper management of this resource (Chigor et al., 2012; Ascott et al., 2016).

Rivers are the main destination of wastewater from different activities, and integrate important processes that affect environmental degradation in a watershed: atmospheric transport and deposition, surface flow, loads generations and movement, and subsurface flow in groundwater zones (Novotny, 2002). Therefore, these systems reflect the overall conditions of its surrounding environment.

Hydrodynamic simulations linked to water quality give a complete description of transport dynamics in channels, and have been recognized as an important tool in water resources management (Melching et al., 2003; Wang et al. 2013). Despite the fact that numerous models are available today, with reliable numerical solutions, several important questions remain challenging, such as sampling frequency requirements (which affects the quality of modeling results), flow representation (crude estimations might lead to misleading interpretations), and calibration strategies (responsible for uncertainties in results).

Modeling based on mechanistic and deterministic principles (which refers to simulation based on conservation laws) under unsteady state provides a means to connect a wide range of quantitative (flow) and qualitative (concentrations) water aspects. Even though sophisticated techniques are available for time series forecasting, such as artificial intelligence, estimations based on conservation laws offer robustness and easiness to implement different system conditions (Campozano et al., 2014).

The governing expressions (mass and momentum conservation) are solved using numerical approximations, which usually require small time steps, in the order of seconds or minutes. However, contrary to hydrological data, that are often available in the form of continuous daily records, water quality data usually are limited or scarce (Fonseca et al., 2014; Creaco *et al.* 2016).

In this context, synthetic data become a recurrent strategy in modeling studies. Traditionally applied in hydrological sciences, such as streamflow and rainfall forecasts (Efstratiadis et al., 2014; Pereira and Souza, 2014; Lamontagne et al., 2018), this artifice provides controlled experiments that elucidate the significance of possible scenarios.

In water quality modeling, synthetic data was used by McIntyre (2004), who explored this strategy to perform a sensitivity analysis of parameters in the Streeter-Phelps

model, while Jia et al. (2018) applied high frequency synthetic data to study the sensibility of decay rates. Both studies generated synthetic data through the models, applying controlled parameters.

This thesis proposes to use this concept to define boundary conditions for unsteady deterministic simulations. Boundary conditions are key factors in modeling studies, since they can be one of the main sources of errors and uncertainty (Rode and Wriedt, 2006; Ji 2008; Wang et al., 2013).

The major contribution is to provide a comprehensive set of evidence to encourage the integration between deterministic and statistical/empirical approaches for time series prediction, not explored in similar studies of contaminant transport in rivers. This approach becomes an auxiliary strategy in modeling environmental time series in high temporal resolution. At the same time, this research provides new insights about calibration and the significance of boundary conditions in simulations of water quality under unsteady state in urban rivers.

Usual calibration parameters are kinetic coefficients (that represent physical-chemical-biological transformations that substances go through in contact with water). Often unknown and with an important role in mass balances calculation, these processes representation may be responsible for errors or mismatch of reality. This research presents methods to take into account temporal variations of these processes, providing a parsimonious and fast calibration, as an alternative to the often time consuming and uncertain trial and error approach.

The experimental plan consists in evaluating different approaches to define a synthetic pollutograph for the upstream section using a historical dataset based on quarterly monitoring samples in the Iguaçu river, Paraná state. The tests are based on interpolations, Fourier series and first order autoregressive models. Following this step, the series is propagated using the SIHQUAL model (*Simulação Hidrodinâmica e de Qualidade de Água*[3]), that solves Saint-Venant and advection-dispersion-reaction equations. To calibrate the simulations, a set of kinetic rates is defined accordantly to the variation in the boundary condition.

1.1 Thesis Motivation

Sampling frequency and model requirements

Predictions of water quality over time are often performed with monthly data (e. g. Fonseca et al.; 2014; Kanda et al., 2015) and quarterly campaigns (e. g. Larentis et al.; Chang et al., 2015). Data sampled at high frequency, on the other hand, is usually available for brief periods, which may limit the simulations also to short time prediction

---

[3]Hydrodynamic and Water Quality Simulation

(e.g. Mannina and Viviani et al., 2010; Zuo et al., 2015).

Traditional monitoring plans are based on discrete samples, analyzed periodically at particular locations; this information represents snapshots at variable instants, describing specific water quality conditions. If these snapshots are not available sequentially over a longer period of time, the data may fail to capture effects of processes and parameters controlling release, fate, and transport of pollutants that are only identifiable at larger temporal scales, such as seasonal and hydrological variability.

Additionally, monitoring programs have limitations (uncertainties related to sample collection, integrity and laboratory analysis, besides high costs associated with equipment and qualified people), which highlight the need for techniques to interpolate and extrapolate the data necessary to match model requirements. In transient modeling, time series are essential as inputs and for model calibration.

Calibration

To Benedini and Tsakiris (2013), a recurrent difficulty in modeling studies lies in the identification of appropriate reaction coefficients, specific for each pollutant and able to interpret its behavior. A given set of reaction coefficients may overlap the actual effects of transport on the distribution of pollutants in the channel. Additionally, different criteria may produce significantly distinct results, which interfere in the decision making process.

The most common method used for calibration of flow and water quality models is the trial and error iterative process (e.g. Fonseca et al., 2014, Salla et al., 2014, Noh et al., 2015, Kanda et al., 2015, Adams et al., 2016). However, this technique is often time consuming and subjective. When a model has a high number of parameters to be calibrated, it becomes difficult to alter them and still maintain control of the model response (Chau, 2006).

In this case, computational optimization routines are suggested, such as genetic algorithm (Ng and Perera, 2003; Kondageski and Fernandes, 2009) and the Particle Swarm Optimization tool (Wang et al., 2008; Knapik et al., 2016). Nonetheless, these studies consider only a steady analysis of kinetic processes.

The calibration proposed in this research takes into account the temporal variation of kinetic processes using a simplified approach, based on attributes of the case study, characteristic values of the literature, and random fluctuation.

Boundary Conditions

In the problem of pollutant routing, a limited length of interest in the river is selected as control volume; therefore, boundary conditions are the driving forces that cause flow and water quality changes within the domain.

Theoretically, boundary conditions should be placed where its effects are not significant. However, identify such region is not trivial. In addition, it could extend the control volume needed for simulation, requiring further data and computational efforts.

Linden et al. (2015) state that, although sensitivity analysis of parameters in model outputs has been described in the literature, less often the influence of boundary conditions or input data is evaluated. These aspects are usually explored in systems affected by coastal waters (e.g. Fan et al., 2012, Alarcon, 2014). The sensitivity of input condition usually evaluated are those related to meteorological conditions and inflow and outflow volumes into reservoir (Linden et al., 2015), compartment exchanges within the basin (such as export coefficients), loading inputs, and transformation rates (e. g. Zhang et al., 2015; Hankin et al., 2016; Reder et al., 2017; Jia et al. 2018).

Water resources management

To Loucks and Beek (2017), urban and agricultural development, deforestation, climatic variability, and modifications in regional management can alter the distribution of rainfall, stream flow and pollution level over time; combined with challenges faced by mostly developing countries (such as irregular occupation of floodplains and water supply areas, insufficient collection and treatment of wastes etc), these questions usually require an unsteady assessment.

This thesis provides background to evaluate the risk of disagreement with quality standards based on duration curves, complementing the information granted by monitoring programs and steady assessment. The proposed analysis allows to guarantee flexibility for framework classifications, verify the frequency of transgressions, or to estimate critical periods and locations, defining adequate planning for water withdrawals, effluent releases and water use charges throughout the year.

1.2 Hypothesis and objectives

The hypothesis tested in this thesis is that the integration between synthetic pollutographs generation and traditional deterministic modeling under unsteady state is reliable to be used in water resources management, offering superior representativeness than traditional analysis, such as discrete monitoring and evaluations based on steady conditions.

The thesis main goal is to provide a integrated tool for water quality assessment over time and space in rivers, compatible with multiple scenarios of flow, input conditions and data availability. Specific objectives are:

- Generate synthetic pollutographs to convert a historical dataset (monitoring as snapshots during twelve years) into continuous information for a specific period;

- Evaluate if the concept of synthetic pollutographs is suitable as boundary conditions for deterministic modeling;
- Establish simulations under unsteady state for water quality indicators of organic matter and nutrient pollution;
- Assess the behavior of a time series generated through statistical/empirical approaches when propagated in time and space with a model based on conservation laws;
- Assess if upstream boundary condition have a significant role in contaminant transport simulation under unsteady state in rivers;
- Establish calibration based on unsteady behavior.

2   Conceptual aspects of river water modeling

Modelling is a process or procedure intended to focus and force clearer thinking and to promote more informed decision-making. The approach involves problem recognition, system definition and bounding, identification of various goals or objectives, identification and evaluation of various alternatives, and very importantly, effective communication of this information to those who need to know.

– Loucks and Beek

This section presents the main characteristics of water quantity coupled with quality modeling in streams. Water quantity is linked to discharges and the way water flow varies temporal and spatially, while water quality incorporates the mass distribution (concentration) of different constituents in aquatic systems.

Water quality mathematical models have been used since 1925, when Streeter and Phelps first developed a mass balance expression to reproduce the assessment of dissolved oxygen in the Ohio River (Streeter and Phelps, 1925). This model incorporates two primary mechanisms governing the fate of oxygen in rivers that receive sewage: decomposition of organic matter and atmospheric aeration.

Since this pioneering study, several environmental agencies and educational institutions have developed modeling strategies to represent water quantity and quality in natural systems. The differences are mainly due to flow representation and complexity level of interaction between physical, chemical and biological processes affecting the distribution of pollutants.

Even though every modeling evaluation has its own characteristics, they all are configured by three main modules, as presented in figure 1: (*I*) Water quantity, (*II*) Water quality and (*III*) Water resources management.

The first module provides the velocity field and hydraulic features of the channel, which are used to solve the water quality module, along with external loadings, dispersion and kinetic coefficients. The results are concentrations of a substance over space and time, used for water resources planning or management purposes, in the third module. Monitoring plans have an important role in this context, in order to define parameters, boundary and initial conditions, and posterior calibration and verification.
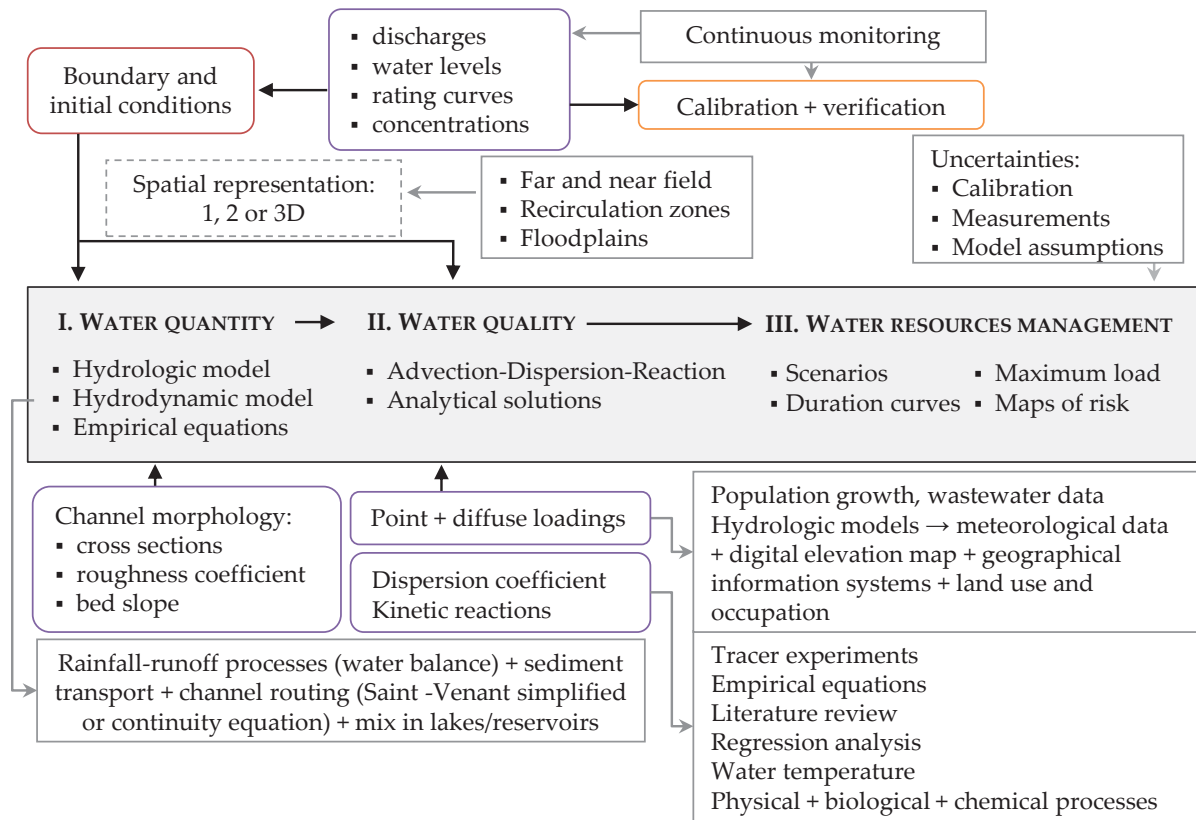
Figure 1: Main components involved in flow and water quality modeling

2.1 Transport of substances

Once a pollutant reach a watercourse, its transport along space and time is result of several processes, as represented in figure 2. Mass transport is a result of advection, diffusion and dispersion processes. The first phenomenon is responsible for the transport of particulate and dissolved material through the fluid movement itself.

Diffusion is the process in which the mass is propagated due to the particles random movement, and may classified as molecular or turbulent. According to Nogueira (1991), molecular diffusion is a fluid property, and occurs due to the errant molecular movement. Turbulent diffusion, on the other hand, is controlled by flow conditions; it is result of random movement of small eddies, and is significantly more relevant than molecular diffusion.

Dispersion represents a process in which the pollutant is mixed in water by the interaction between turbulent diffusion and velocity gradients. In open channels, where there is turbulence due to velocity variability, cross sections geometry and roughness changes, dispersion prevails over diffusion (Chapra, 1997).
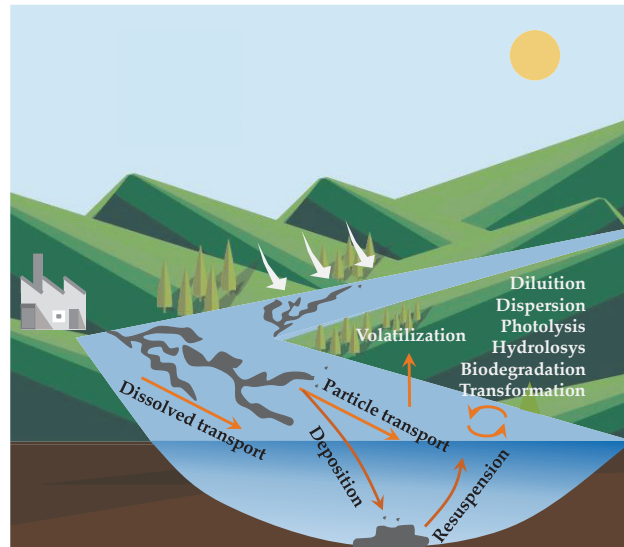
Figure 2: Fate of contaminants in rivers. Source: adapted from Barber *et al.* (1995)

In water quality studies, neglecting diffusion or dispersion can lead to inaccurate results. The authors Martin and McCutcheon (1999) state that disregarding these mixing effects causes overestimation of peak concentrations, and underestimation of arrival time and period of release effects.

A thorough review of the impact of dispersion coefficients can be found in Abderrezzak et al. (2015), who investigate the extent to which existing empirical formulations of longitudinal dispersion coefficient can be used in one-dimensional (1D) numerical modeling of solute transport. The authors affirmed that a proper longitudinal dispersion formula is able to deal with non-uniformities of channel geometry and highly variable flow and solute discharges, that may violate locally the main assumptions of the 1D shallow water equations (e.g. presence of secondary currents, turbulence) and the advection–dispersion equation (e.g. transverse dispersion not negligible, well mixing of the solute over the cross-section not attained).

Besides advection, diffusion and dispersion, other physical (particulate settling, for example), biological (such as algae growth and death) and chemical (chemical reactions in general) activities result in transformations, and affect the transfer of matter through contours of the control volume.

The distribution of a species in the water body is also affected by point and non-point contributions, such as domestic and industrial wastes, tributaries and runoff in the watershed. Point sources are originated from a single waste load location, and nonpoint sources include diffuse pollution loads. While the first can be easily identified, nonpoint loadings are often difficult to attribute to a particular location, and have been recognized as a major threat to water resources throughout the world (Lee et al., 2010; Wang et al., 2012).

2.2 Boundary conditions

Modeling based on deterministic principles (which refers to simulation based on conservation laws) leads to solving differential equations, that mathematically describe natural systems. In order to simulate the mass balance of a control volume, the contours (representing the influence of outside of the domain) must be specified. Figure 3 shows a scheme in which the domain of interest is represented in time and space.
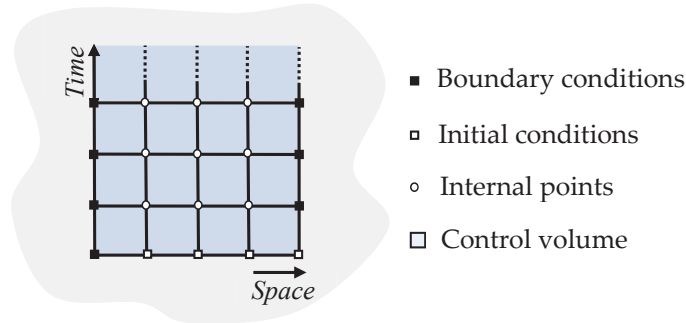
Figure 3: Domain of interest in the solution of partial differential equations

Initial conditions correspond to values at the simulation beginning $(t = 0)$, while boundary conditions correspond to values of the unknowns variables in the contours of spatial domain. A few examples illustrate how critical are the establishment of reliable boundary conditions in modeling contaminant's transport in different systems: In a estuary, Alarcon (2014) showed that nitrate is the constituent most affected by boundary conditions among the parameters studied. Fan et al. (2012) reported a sensitivity analysis regarding boundary conditions for biochemical oxygen demand and dissolved oxygen, although under steady state. They applied a combination between HEC-RAS and a modified Streeter–Phelps model to simulate water quality of a tidal river, showing that the organic matter is significantly affected by pushed-back sea water.

In rivers, Sincock *et al.*, (2003) have shown that boundary conditions can also affect model calibration. The authors identified incorrect rate coefficients when calibrating a water quality model, proposing that the reason behind it is result of very low ammonium concentrations at the model boundaries, reaffirming the role of this aspect in the simulation.

In the same context, Quiel et al. (2011) linked the importance of upstream boundary condition to assessment of a planning scenario in an european river. They showed that even a drastic reduction of phosphorus inputs from anthropogenic sources would not be significant to influence algal biomass, due to the ability of algal cells to store phosphorus. On the other hand, a reduction of phosphorus inputs in the headwaters could compensate possible effects on algal biomass due to climate changes.

2.3   Spatial variation

Flow in natural systems have three-dimensional (3D) characteristics, as consequence of cross sections variations and irregularities, presence of curves, meanders and several others barriers. The expression for mass conservation combined with the Navier-Stokes equations provide a complete description of the velocity field in these cases. This approach is the closest representation to reality, but it requires several data, usually not available. Three-dimensional models have been commonly applied to branched network of rivers or small stretches (e. g. Sinha et al., 2013; Sokolova et al., 2015).

In streams, the longitudinal scale is significantly larger than vertical and transversal scales. Therefore, the flow can be modeled through one-dimensional equations. This approach assumes that longitudinal concentration gradients prevail and there is instant mixture in vertical and transverse directions.

Benedini and Tsakiris (2013) state that the assumptions made for 1D analysis are not valid particularly in large rivers, in which an injected pollutant requires some time before reaching an acceptable uniform distribution in the cross section. Moreover, large rivers may have zones of slow stream, stagnant water or floodplains, in which the substance transport occurs in a different way from that in the main stream. Key components of the flow field, such as flow separations and recirculation zones, should also be reproduced by two-dimensional (2D) models (USACE, 1993; Olsen, 2007).

Nevertheless, Benjankar et al. (2015) call attention to the fact that despite their inability to resolve flow details, 1D models are very useful as they are computationally efficient and allow simulations over much larger stream domains and longer periods than 2D models.

The 1D approach is fundamental when studying analytical solutions, which are considered the most reliable way for solving the fundamental differential equation in simple geometrical configurations, such as channels with uniform flow or prismatic cross sections. Indeed, a common practice is to use this analysis to verify numerical solutions (e. g. Juxiang, et al., 2011; Estabragh et al., 2012), although it can be applied to predict the impact of contaminants in water bodies (e. g. Fan et al., 2013). In spite of the simplifications involved, analytical solutions yield straight and fast results, which is important in cases of accidental spills, for example.

2.4   Flow representation

Flow conditions have a fundamental role in water quality assessment. Water pathway, volume and velocity directly controls the transport of dissolved and particulate substances in water bodies. In addition, advection and morphological characteristics can affect kinetic processes, such as reaeration, volatilization, and photolysis (Ambrose et al., 1988). Flow conditions also may cause ressuspention of sedimented material, that might

promote the release of contaminants (Zuo et al., 2015).

In the transient approach, estimations of advection and cross-sectional areas are closer to reality, since it considers the principles for mass and momentum conservation in space and time. Crude estimations for hydraulic transport, such as using discharge coefficients or the Manning equation (commonly applied in simulation under steady state), may allow kinetics processes to outweigh the effects of advection (during the calibration phase), which can introduce errors in the analysis.

The important link between these aspects is illustrated with the discussion presented by Chapra (1997). Figure 4 shows a sinusoidal flow and the concentration of a pollutant in a channel. Although the curves have a similar shape, the concentration wave moves at about 60% of the flow wave, due to differences in the celerity of each one. Decisions in water resources management might be affect because of the lags; a water user might begin withdraws after the passage of the water wave crest, and consequently collect the critical pollution level, occurring some time latter.
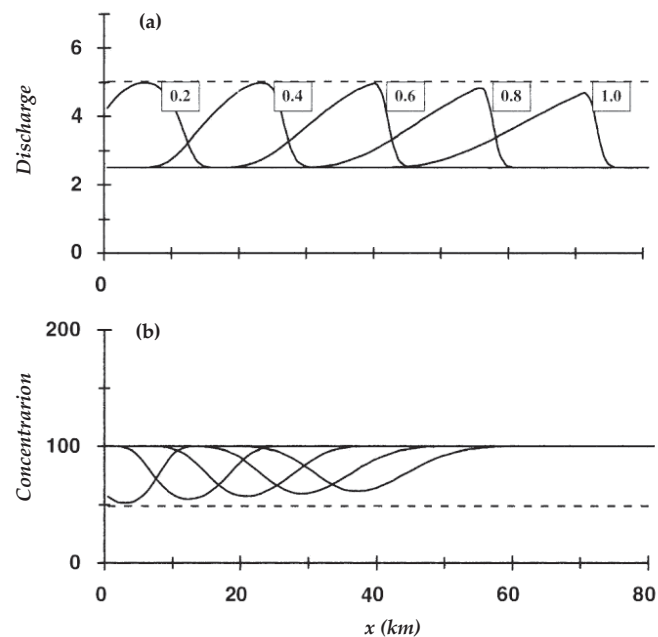
Figure 4: Simulation of (a) a hydrograph and (b) a dilution wave moving through a channel. Source: Chapra (1997)

In steady approaches the flow is estimated using discharge coefficients or the Manning equation (e. g. Parmar and Keshari, 2012; Babbar, 2014). The disadvantage is fewer scenarios possibilities, although it requires less data and is easier for computational implementation. Considering low flow occurrences, this representation have been considered adequate when the cause of degradation is a steady point source (Novotny, 2002), or if the studied reach receive less pollution loads (Ferreira et al., 2016). An alternative strategy is to consider unsteady analysis through a sequence of steady-state simulations,

though this method may lead to a challenging calibration, in order to keep calculations closer to reality (e. g. Noh et al., 2015; Salla et al., 2014).

Considering problems for which the water conditions change rapidly (during a storm event, for example), a transient analysis is required. To represent in time and space a wave characteristic, usually hydrological or hydrodynamical models are applied. The first simulate precipitation-runoff processes, and usually are associated with flow routing methods, that includes the unsteady continuity equation and part of the momentum equation. According to Saleh et al. (2013), this approach is preferable for flow routing at regional scale. These models use simplified techniques, generally based on linear/non-linear reservoirs, that include the Muskingum routing method, the cascade of linear reservoirs, and the Muskingum-Cunge approach, among others.

However, Rode et al. (2010) claim that critical problems arise when modeling catchment-scale water quality. Frequently the basin has a substantial spatial variability (often difficult to represent); in addition, descriptions of mass fluxes and transformation processes in different compartments are required (soil, surface water transportation, stream-aquifer interactions etc).

On the other hand, hydrodynamic models, that include the full one-dimensional unsteady continuity and momentum equations, have the capability to accurately simulate a wide spectrum of waterway characteristics. The continuity equation describes the balance between input, storage and output in a section of river, and the momentum equation relates the change in momentum to the applied forces (Liggett, 1975).

These equations are fit to simulate the downstream propagation of kinematic and diffusive waves, and remain valid when downstream backwater effects or significant tributary inflows are present, or when upstream propagation of a wave can occur, such as from large tides and storm surges (Saleh et al., 2013). In order to gain satisfactory results in hydrodynamic modeling, it is critical a correct description of bed channel (roughness and slope), and reliable rating curves (Saleh et al., 2013; Benjankar et al., 2015). The uncertainties in these later can induce unrealistic estimates for the roughness coefficient (Domeneghetti et al., 2012).

2.5   Calibration

Parameter's values are often unknown, because of spatial and temporal variability, measurement challenges, simplification in model descriptions, or commonly lack of data. Therefore, model calibration is required. Usual parameters calibrated in water quality models are kinetic rates, that describe conversion processes and are often unknown (e. g. Mannina and Viviani, 2010; Tang et al., 2016).

These rates can also be obtained through analytical tests, statistical, conceptual and empirical analysis; common equations relate them to channel hydraulics, such as velocity, water level, slope or discharges (Brown and Barnwell, 1987). These latter differ mainly by

the range of suggested applications, and calculations can be significantly divergent with each one (Ávila, 2014). Fluctuation of coefficients intervals reported in the literature can also be wide, and these are usually estimated through analytical tests for specific areas; a common example is the reaeration rate, that varies from 0 to 100 d^{-1}, according to Brown and Barnwell (1987).

The main strategies for model calibration are trial and error and automatic techniques. The first one is usually subjective and time consuming, while the second often assume criterion of optimization that solves the mathematical issue, but that can generate values without physical meaning (Kondagesk and Fernandes, 2009); furthermore, often a technique identifies multiple datasets that satisfy the optimization problem. Depending on a given combination, these parameters may superimpose transport effects, and therefore generate wrong interpretations (Dortch and Johnson, 1992). Computation time also is increased, due to optimization operations (Razavi et al., 2010).

Automatic routines are commonly applied in hydrological studies, in which many parameters related to watershed scale are involved (e. g. Rode et al., 2007; Preis and Ostfeld 2008). Calibration of river flow modeling under unsteady state also has been broadly investigated through optimization techniques (e. g. Siqueira et al., 2016; Lin 2017); for water quality, model calibration is commonly conducted for steady state analysis (e. g. Ng and Perera, 2003; Wang 2008; Kondagesk and Fernandes, 2009; Knapik et al. 2016).

For transport of pollutants in rivers through a transient analysis, Sincock et al. (2003) and Mannina and Viviani (2010) have applied a Monte Carlo procedure to generate multiple sets of decay rates; the strategy is based on investigation of various runs of the model with different randomly chosen parameter values. In this procedure, usually an objective function is used to discard unrealistic values.

2.6 Sampling frequency

Although sampling frequency is a common concern in researches regarding water quality modeling (e. g. Zhou et al., 2011; Sorribas et al., 2012; Langeveld et al., 2013), there is a lack of attention to a careful assessment of modeling requirements prior to field studies. According to Martin and McCutcheon (1999), the absence of information can result in sampling efforts with missing critical data, loss of important gradients, or failure to close flow and mass balances.

As stated by Meals et al. (2013), 80 to 90% of annual load may be delivered in 10 to 20% of time, indicating that choosing when to sample can be as important as how often to sample.

In this context, Richards (1998), for example, showed in his study that monthly series gave only a very crude representation of the daily load flux, but it was better than expected just because it included the peaks of two of the four major storms of the year;

a monthly series based on dates about 10 days later than these would have included practically no storm observations, and would have underestimated the suspended solids load; quarterly samples resulted in a poor fit on the actual daily flux pattern in this study – however, it was considered only one year of samples, as verified in figure 5.

Richards (1998) also illustrates that different scenarios of suspended solids load emerge from different sampling frequencies: decreasing time resolution tend to miss short-term (but important) events with high flow or high concentrations (figure 5).
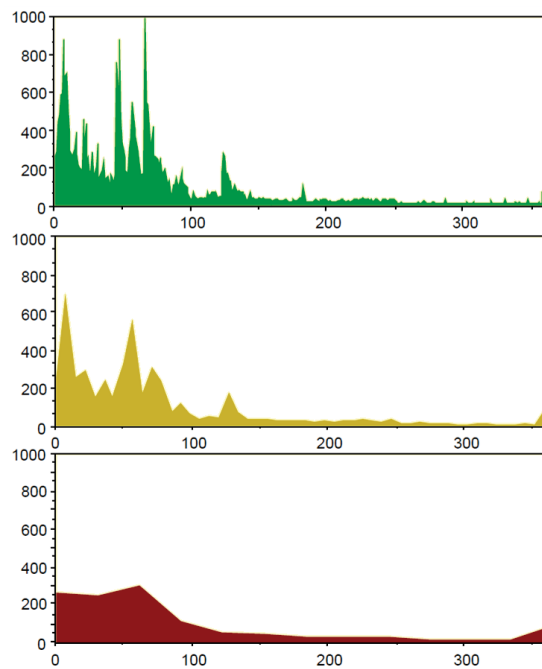
Figure 5: Time series of suspended solids loads for the Grand River from March 1, 1976 to February 28, 1977. Top: daily samples; middle: weekly samples; bottom: monthly samples; weekly and monthly sample values were drawn from actual daily sample data series. Source: Richards (1998)

Table 1 compares several studies of water quality modeling, showing the parameters more often simulated, and data available. This summary shows the potential usability of modeling based on traditional deterministic principles, since they represent watersheds of different characteristics, with multiple input conditions.

Most studies based on high frequency data are also limited to shorter periods of analysis (e. g. Sincock et al., 2003; Langeveld et al., 2013), usually due to financial and logistics challenges. Data as irregular samples, on the other hand, are commonly available for longer periods, which may be an advantage when studying a system's historical behavior (e. g. Larentis et al., 2008; Inthasaro and Wu, 2012).

Table 1: Common parameters simulated in rivers and monitoring data used in water quality modeling under unsteady conditions

Reference	Basin	River length (km)[1]	Data availability	Simulation period
Adams et al. (2016)	125 km^2: 90% grassland	[2]	Phosphorus at each 30 min and suspended solids at 15 min	1 year (hourly time step)
Fonseca *et al.* (2014)	176 km^2: livestock, piggeries and domestic wastewater	[2]	3 water quality stations; monthly data for temperature, fecal coliforms, DO, BOD, total suspended solids, nitrates, orthophosphates and pH	1 year for calibration and another for validation (daily time step)
Hwang et al. (2014)	[3]	18.5	Two days at intervals of 30 min: nitrogen and phosphorous compounds, suspended solids, chemical oxygen demand, BOD and chlorophyll-a	1 day calibration and 1 day validation
Inthasaro and Wu (2012)	[3]	21	$\approx$8 to 80 measured values: ammonia, nitrate, polychlorinated biphenyls, phosphate and biomass	10 years
Kanda *et al.* (2015)	[3]	55	5 sampling points; monthly water quality data: BOD and DO	1 year for calibration and 1 month for validation
Langeveld *et al.* (2013)	[3]	146	1-2 min ammonium and DO	1 month
Larentis *et al.* (2008)	26000 km^2: low population density, livestock, agriculture and industrial activities	550	6 monitoring points; quarterly data for total phosphorus, dissolved oxygen and fecal coliforms	9 years
Launay et al. (2019)	95590 km^2	810	Suspended particulate matter[4]	17 days (flood event)
Mannina and Viviani (2010)	160 km^2: rural ephemeral river	6	BOD and DO at approximately each 10 min	$\approx$8 hours
Nguyen et al. (2018)	25 km^2	11.5	Monthly BOD, DO, temperature and nitrogen compounds	1 year
Quiel *et al.*, (2011)	148268 km^2: 61% of land use is agricultural	700	Biweekly to monthly intervals: phytoplankton and nutrients	8 years
Sincock *et al.* (2003)	2464 km^2	30	Daily data for BOD, nitrate, DO and temperature	4 months
Zhou *et al.* (2011)	9750 km^2: high population and economic density	1600	4 values for chemical oxygen demand	30 h
Zuo *et al.* (2015)	28150 km^2: high population density	$\approx$2.5	Experimental values: permanganate index and ammonia	3 days

[1] Not necessarily the simulated length, since most studies are not specific [2] Model for watershed scale [3] Not available [4] Data frequency is not specified, although authors claim use of high temporal resolution measurements

2.7 Water quality time series

Water quality time series have inherent characteristics, such as nonlinear trends, missing data, outliers, irregular measurement patterns, seasonal behavior, serial correlation, annual and diurnal cycles. In addition, these series are affected by intermittent events such as rain storms (Abaurrea et al., 2011; Milne et al., 2009).

Time series from natural systems usually exhibit autocorrelation, a property described by specific functions. In water quality, autocorrelation often is modeled with classical techniques such as autoregressive, autoregressive moving average or autoregressive integrated moving average models (e. g. Abaurrea et al., 2011; Arya and Zhang, 2015; Chen and Boccelli, 2018).

In the process of infilling missing data, these issues should be addressed. A common strategy is based on linear interpolation using the observed data (e. g. Adams et al., 2016), although this method may not reflect nonlinear trends. In general, trend is modeled using a deterministic function, usually polynomial, power or logarithmic (e. g. Costa and Monteiro, 2015). Another frequent approach to deal with missing data is to use data from near stations (e.g. Quiel et al 2011; Launay *et al.*, 2019).

Hirsh et al. (2010) and Chanat et al. (2016) applied a model in which concentration is a log-linear function of discharge, season, and time. To the authors, this procedure allows flexibility in representations of the long-term trend, seasonal aspects, and discharge-related components.

Additional examples of techniques for water quality time series modeling are: linear regression (Adams et al., 2016), smoothing spline interpolations (Feng et al., 2015), Fourier series (Rodríguez et al., 2013), wavelet analysis (Dokmen and Aslan, 2013; Barzegar et al., 2015), among several others.

A few methods and applications for modeling of concentrations time series are presented in table 2. This review focus on empirical/statistical techniques, that depend mostly on observed inputs and outputs to estimate parameters; these approaches also lack an explicit description of cause and effect relationships, with a minimum of understanding of how the system works.

The review on table 2 highlights that most studies focus on reproduce data with frequencies at least monthly. One of the challenges in this research lies in represent time series at higher frequency than the actual available data, i. e., combine processes of different temporal scales.

Table 2: Comparison of studies for water quality time series modeling through statistical/empirical approaches

Reference	Methods	Parameters and recorded data
Abaurrea *et al.* (2011)	Regression model with a Gaussian autoregressive moving average error, using concentration, temperature and discharge	Conductivity: four monthly series
Adams *et al.* (2016)	Linear regression between phosphorus and turbidity to specify suspended sediment	Turbidity: 15-min intervals; phosphorus: 30-min
Hirsch (2010)	Weighted regressions of concentrations on time, discharge, and season	Total phosphorus: data from 1978 to early 2009 (773 measurements, variable sampling frequency)
Feng et al. (2015)	Smooting (cubic B-splines as basis functions) and interpolation (natural spline interpolation)	Dissolved oxygen, permanganate index and ammonia nitrogen: monthly data
Dökmen and Aslan (2013)	Wavelet	Chlorine, nitrate and pH: monthly data between December 1999 and 2000
Amornsamankul et al. (2012)	Factor analysis, correlation analysis and Fourier series	Dissolved oxygen, biochemical oxygen demand, total phosphorus and suspended solids: monthly data 2002-2007
Barzegar et al. (2017)	Extreme learning machine and wavelet-extreme learning machine hybrid models	Electrical conductivity - monthly data (total of 315 monthly datasets (1984–2011)
Arya and Zhang (2015)	Auto Regressive Integrated Moving Average	Dissolved oxygen and temperature: monthly data 1995-2012
Parmar and Bhardwaj (2014)	Auto Regressive Integrated Moving Average	Monthly average of 10 years: pH, chemical oxygen demand, free ammonia, total Kjeldahl nitrogen, water temperature, BOD and DO
Shao et al. (2018)	Back propagation neural network, optimized by the Cuckoo Search algorithm	188 data for each parameter in 94 days: conductivity, chlorophyll content, dissolved oxygen, dissolved organic matter, pH, permanganate index, turbidity, total nitrogen

2.8 Water resources planning and management

Identification and representation of physical, chemical and biological dynamics gives the decision makers and water managers an opportunity to design and to manage appropriate water policies for different uses. Modeling is an important tool in this context: besides providing a full description of quality conditions, mathematical simulations give an efficient way to predict the response of a water body to interventions and management actions before its implementation. Therefore, it is possible to establish requirements to

attain standards and to identify knowledge gaps, allowing background to conduct future monitoring and research efforts.

For the definition of framework, that is the basis for other water resources management instruments (defined by Law 9433/97 - Brasil, 1997), classification is performed according to concentrations of parameters associated with reference discharges, as defined in CONAMA Resolution n° 357, dated March 17, 2005 (CONAMA, 2005). In this case, mostly the effects of point sources are represented, in which the highest concentration occurs during periods of drought. Critical conditions generated by diffuse sources, however, occur with the beginning of rain, when the surface flow promotes the contribution of a large load of pollutants. Therefore, a more recent approach emphasizes the need to consider hydrological regimes that meet human and ecosystem needs, rather than setting constant reference discharges (Souza et al., 2008).

A wide number of substances are expected to be found in superficial waters, originating from urban and industrial wastewater or natural contact of water with soil and sediment: calcium, magnesium, nitrates, chlorides, carbonates, sulphur, chromium, arsenic, cadmium, nickel and organic compounds. These latter come from the decomposition of organic matter of plants and animals, and include residues from agricultural areas and wastes of domestic and industrial sources (Porto et al, 1991).

Recent studies have recommended the identification of emerging substances – hormones, pharmaceutical, caffeine etc – as indicatives of anthropogenic pollution, as a complement to organic matter (Ide et al., 2013; Santos et al., 2016).

Many of these substances are nonconservative and undergo changes in contact with water. The most frequent pollutants simulated are compounds of oxygen, nitrogen and phosphorus, whose transformations can be described by means of specific biochemical processes (Sincock et al., 2003; Mannina and Viviani, 2010; Salla et al., 2014).

In the aquatic environment, the organic matter plays an important role in production and consumption cycles. This material is composed by protein compounds, carbohydrates, fats, oils, urea, surfactants, phenols, pesticides and other components in lesser amounts. However, in studies of water quality usually only carbonaceous matter is analyzed, since this is the portion that consumes oxygen. Organic matter is traditionally represented by biochemical and chemical demand of oxygen in modeling studies (Zhou et al., 2011; Salla et al., 2014).

Although these parameters are recommended to evaluate pollution by the regulation CONAMA n° 357 (CONAMA, 2005), some analytical limitations may compromise the interpretation of results. The biochemical demand of oxygen test identify only the biodegradable fraction of organic compounds, while the chemical demand of oxygen do not allow to differentiate the sample portion that may be oxidized biologically. In addition, BOD quantifies only the fraction of oxygen consumed by microorganisms adapted; consequently, the presence of heavy metals or other toxic substance may restrict their

action. To overcome such limitations, the organic carbon has been proposed as parameter for the overall determination of organic pollution in water and wastewater (Thomas and Theraulaz, 2007).

In the future it is expected to aggravate the problem of detecting and controlling new pollutants in superficial waters, stressing the role of mathematical models for water quality assessment.

2.9 Overview of previous studies

This section presents an overview of a few researches related to the study of water flow and quality simulations in the last years, as presented in table 3. This compilation summarizes several topics covered in this research, and highlights current gaps in previous studies.

The *Quantity module* column determines the method applied to obtain discharges, whereas *Water quality module* discuss how the concentrations were calculated and if the study presents boundary conditions assessment; *Numerical scheme* indicates if the representations are implicit or explicit; *Calibration* shows the procedures applied to match modeling results and parameters; and *Water Resources Management* considers if the study presented supports water resources management.

Hydrological simulation is a common strategy to obtain flow information and to represent diffuse loads in the watershed scale; results are hydrographs and pollutographs in a specific location of the basin. However, this approach involves a high number of parameters, usually not available, that may compromise model calibration. Moreover, the results often are applied only to a basin portion, because the flow routing phase is neglected or very simplified. This approach has been used to develop water quality indexes (Torres-Bejarano et al., 2011), to analyze the impacts of discharged sewage on downstream water quality (Sokolova et al., 2015), to study scenarios (population growth and industrial sector, installations of hydro-electric power stations) (Larentis et al., 2008), and to develop risk maps (Zhou et al., 2011).

Hydrodynamic modeling, on the other hand, provides flow representation closer to reality, and it is appropriate to simulate changes, such as alterations in water quality through hydraulic modifications in the channel: gate operations (Feng et al., 2012; Hwang et al., 2014), sluice regulation (Zuo et al., 2015) and distinct morphological conditions (Wagenschein and Rode, 2008). This approach, linked to water quality, also have been applied to recommend mechanisms to maintain maximum allowable limit for pollutants, mainly through improvement of treatment efficiencies in wastewater plants and enforcement of effluent standards (Kanda et al., 2015).

Numerical methods are a fundamental aspect of modeling flow and water quality based on conservation laws. Finite difference methods, explicit or implicit, are the most common procedures for one-dimensional problems. These schemes are based on the prin-

ciple of differential equations into algebraic expressions, in which derivatives are converted into finite differences (Chaudhry, 1979). In general, they have first or second order of accuracy, are simple for computational implementation and generate results quickly (Li and Jackson, 2007).

Usual explicit schemes are: FTCS (Foward Time-Centered Space), MacCormack, Quadratic Upstream Interpolation for Convective Kinematics, Leap-Frog and Lax. Similarly, common implicit representations are Backward Time-Centered Space, Crank-Nicolson and Preissman. The explicit method is less complex for implementation, but its solution stability is conditioned to the simulation time interval. To Paiva et al. (2011), hydrodynamic modeling through these schemes have become interesting because of the easiness of programming parallelization. In addition, the algorithm's practicality is an advantage when considering nonlinear problems and simulation of modifications in complex systems, such as urban streams.

In the implicit procedure, the definition of temporal discretization is less restrictive than that of the explicit one, although some studies indicate loss of results quality when increasing the Courant number (e. g. Gajdos and Mandelkern, 1998; Hashemi et al., 2007). In this type of method, the solution involves the resolution of a system of equations, which sometimes causes an increase in the total simulation time due to the size of the matrices (Kalita and Sarma, 2012).

Table 3: Literature overview related to water quality modeling under unsteady conditions

Reference	i	ii	iii	iv	v	vi	vii	viii	ix	x	xi	xii²	xiii	xiv	This thesis
Quantity module															
Hydrologic	✓	✓	✓											6	
Hydrodynamic				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓
Water quality module															
Complete ADR*	✓														
Simplified ADR*					✓	✓	✓	✓	✓	✓	✓	✓	✓		✓
Simple mass balance		✓	✓	✓										✓	
Boundary conditions evaluation		1	1									4			✓
Numerical scheme															
Explicit	✓				✓		1								
Implicit	✓			✓		✓	1	1	✓	✓	✓	3	✓	7	✓
Calibration															
Optimization														✓8	
Manual	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓		
Temporal variation													5		✓
Water resources management															
None	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓			
Scenarios											✓			✓	✓
Duration curves															✓

(*) *Differential Equation of Advection-Dispersion-Reaction*
(1) *Not available*
(2) *Two-dimensional modeling*
(3) *Combine explicit and implicit approaches*
(4) *Input time series were reconstructed with data from near station and power functions relating discharges and concentrations*
(5) *Evaluate the origin of suspended particulate matter and the interaction river-reservoir*
(6) *Routing through a cascade of linear reservoirs – each reservoir represents a river segment*
(7) *Not required*
(8) *Calibration parameters are correction factors, using the least squares method*

(i) Larentis et al. (2008); (ii) Adams (2016); (iii) Fonseca et al. (2014); (iv) Sincock et al. (2003); (v) Mannina and Viviani (2010); (vi) Zhou et al. (2011); (vii) Inthasaro and Wu (2012); (viii) Eisakhani et al. (2012); (ix) Kanda et al. (2015); (x) Hwang (2014); (xi) Langeveld et al. (2013); (xii) Tang et al. (2016); (xiii) Launay et al. (2019); (xiv) Nguyen et al. (2018)

2.10   Summary and thesis contribution

Several questions arise in the representation of flow and contaminant transport in rivers, as summarized in figure 6. The scheme also presents how these aspects are considered in this thesis, and what are the main contributions.

Numerical solutions are associated with numerical diffusion, stability, consistency and convergence limitations. In the hydrodynamic module, channel hydraulic aspects (cross sections, rating curves and spatial representation, for example) and lateral contributions evaluation (tributaries and contributions by hydrological processes) are essential. Water quality, in its turn, depends on external inputs (linked to complex watershed processes and its interaction), advection and dispersion, besides transformation processes (physical, chemical and biological) and other interactions, such as water-sediment and water-atmosphere. In addition, calibration and proper tools for results interpretation are key aspects.

The main contributions of this thesis are related to the subjects: uncertainty due to input data, model requirements versus sampling frequency, and calibration challenges, using an unique integrated analysis for simulations of water quality in rivers. As presented in table 3, although efforts have been made in calibrating models that consider variation over time, there is a lack of reliable and efficient procedures to understand the temporal variation of processes that calibration incorporates. In addition, although boundary conditions has been indicated as an important aspect, its role and forms of description has not been well explored in the proposed context.

This research also draws attention to the need of an analysis regarding risk in water management, using the concept of duration curves. This tool combines the characteristics of a stream throughout the range of variability, without regarding the sequence of occurrence; it may be applied to guarantee flexibility to framework classifications, verify the frequency of transgressions, or to estimate critical periods and locations, defining adequate planning for water withdrawals and effluent releases throughout the year; in the same way, in occasions of high demand, it is possible to regulate fees for water use.

Derived from the hydrological field, this concept has been in general use since 1915 (Searcy, 1959). In water quality, it have generally been performed through simulations under steady state (Brites, 2010; Calmon et al., 2016) and hydrological modeling (Park and Roesner, 2012; Cho and Lee, 2015), monitoring data from rivers (Oliveira et al., 2011; Cunha et al., 2012) and reservoirs (Cunha et al., 2011). Although this product can be generated using these diverse tools, this research provides an unique analysis that complements the information given by duration curves: water quality modeling along time and space in rivers allow to identify and predict, for example, where and when critical events occur.
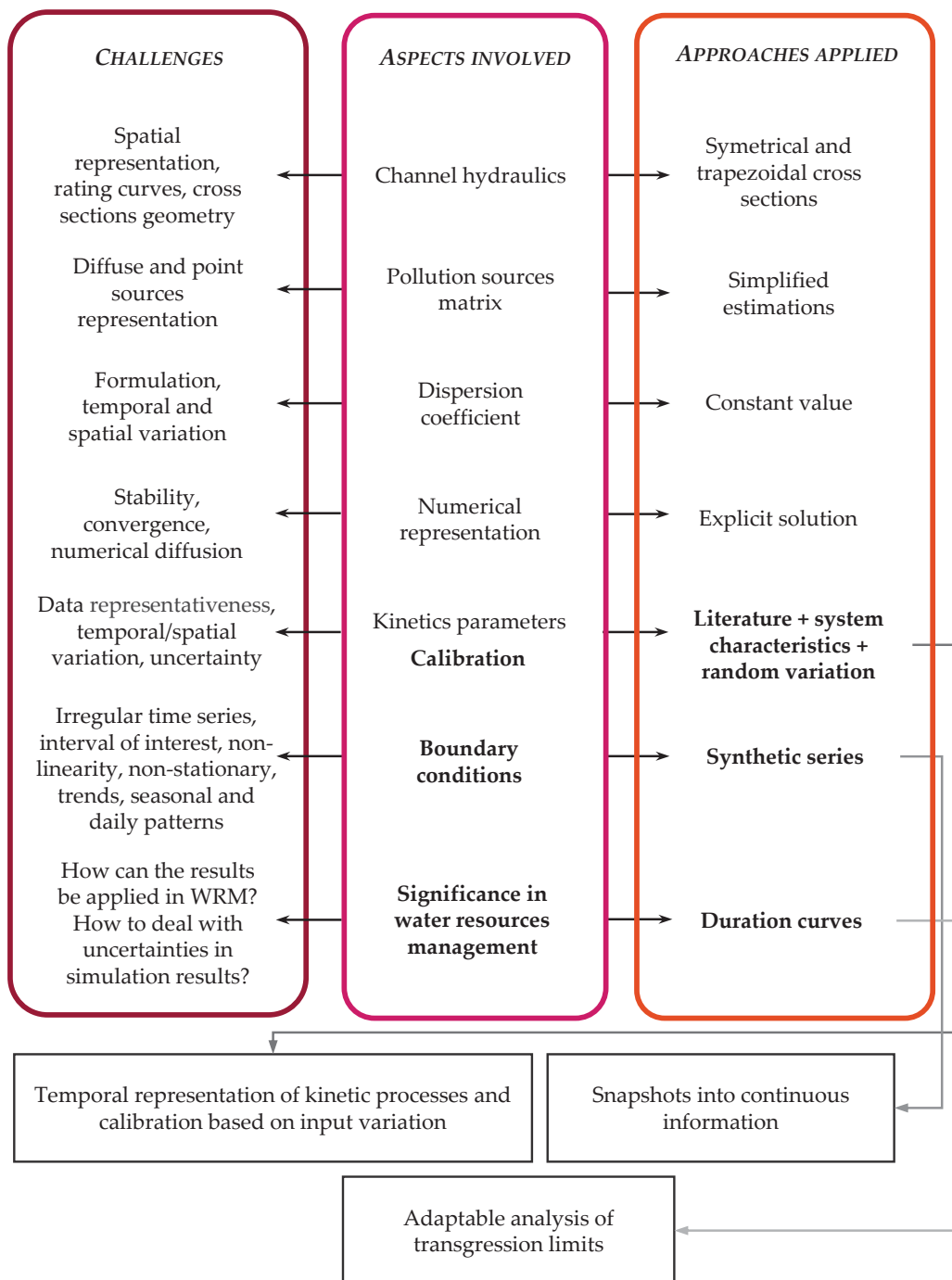
Figure 6: Scheme of development of the thesis and the main aspects involved

3   Integrated modeling

I didn't want to just know names of things. I remember really wanting to know how it all worked.

– Elizabeth Blackburn

The methods applied in this research are based on mathematical and numerical modeling, using the Iguaçu river as case study. The first step is to estimate velocity and cross sections areas. For the water quality modeling, the challenge is to properly represent mass balance for the simulated parameters under unsteady regime, conciliating a calibration procedure and proper representation of boundary conditions.

The main demand in order evaluate this later aspect lies in converting the historical monitoring dataset into the required time series for numerical solution, while accounting for conditions that are continuously changing over time. The problem is schematic in figure 7.
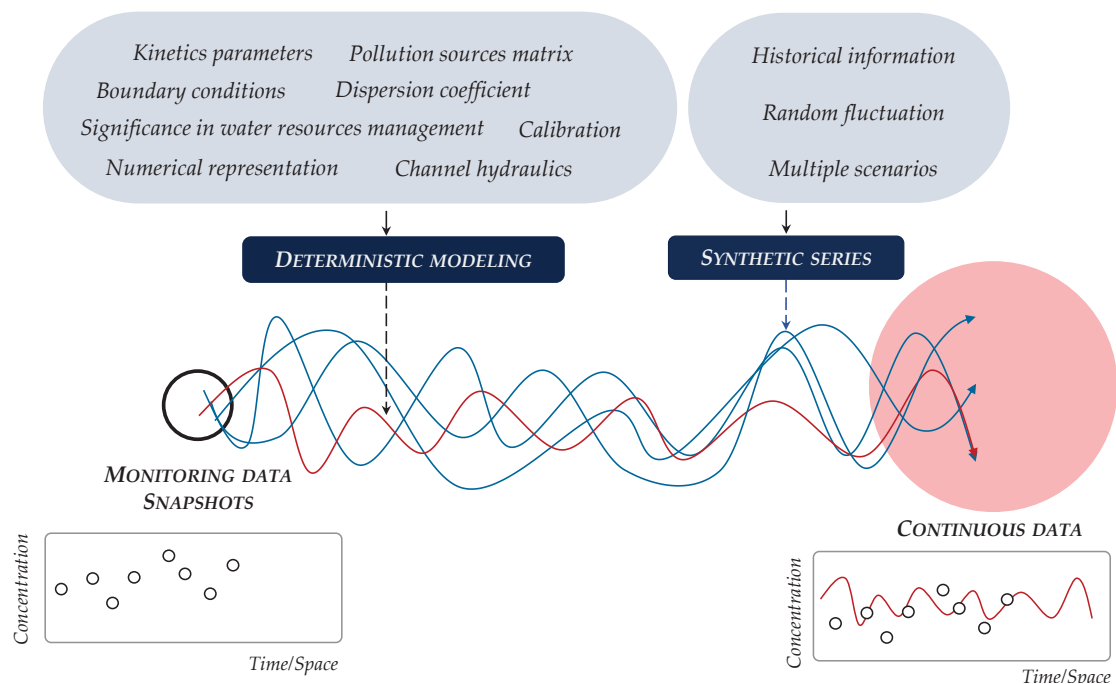
Figure 7: Scheme of the problem in convert discrete sampling data into continuous information; point: observed data (snapshots), continuous line: continuous information

In general, modeling studies aim to interpolate and extrapolate available information, in order to understand or to identify what is not known about a system, and to prepare for future scenarios. In such a context, this thesis proposes a integration between deterministic modeling and synthetic series generation. While the second allows to overcome data limitation, combining uncertainty factors and requiring minimal knowledge about the system (the historical data set provides the necessary information), deterministic modeling accounts for the transfer of matter through contours of a control volume;

several aspects are integrated, such as channel hydraulics, dispersion, transformations processes (due to chemical, biological and physical interactions), point and diffuse pollution dynamics.

Figure 8 represents a scheme of the integrated analysis proposed in this research for the hydrodynamic, time series and water quality modules. The SIHQUAL model is a tool to propagate discharges and concentrations in streams in the longitudinal direction. Developed using Matlab®, the model has a module to solve the Saint-Venant equations, that generates water velocity and cross sections areas to be used in the advection-dispersion-reaction module. The solution in two phases is possible because it is assumed that the studied substance does not affect the river flow or fluid properties.

	Hydrodynamic	Synthetic series	Water quality
Input	Discharges Cross sections Rating curve	Concentration and discharges as snapshots	Diluition Boundary condition Dispersion coefficient Population + Land use
Pre-processing	Rating curve equation Channel geometry	Qqplots Logarithmitic transformation Mean and standard deviation Quantiles	Lateral inputs
Modeling	Saint-Venant equations	Interpolations Fourier AR(1)	Advection + Dispersion + Reaction equation
Calibration	Manning roughness coefficient	Algorithm to select a series	Dispersion coefficient Algorithm to define transformation rates
Output	Advection Cross section areas	Concetration at required numerical time step	Concentration and load over space and time
Post-processing	Discharges	Pollutographs Duration curves Boxplots	Pollutographs Duration curves Boxplots Heatmaps

Figure 8: SIHQUAL modules

The numerical schemes used are finite difference, with explicit representation: Lax diffusive scheme for the hydrodynamic module, and FTCS to solve the water quality module. Appendix A.1 presents the numerical representations and stability criterion.

These numerical techniques have generated stable results, confirming that the association of explicit methods, although not common in similar studies, is feasible and efficient. This solution, despite more sensible to simulation time steps, has the advantage of simple implementation, which may be beneficial when studying complex systems, such as urban watersheds.

3.1 Hydrodynamic module

The Saint-Venant equations are the one-dimensional expressions for mass and momentum conservation. Presented in 1871, by Adhémar Barré de Saint-Venant, the equations are widely applied to represent flow routing (Hwang et al., 2014; Serrano, 2016), flood prediction (Alekseevskii et al., 2014), dam ruptures effects (Peng, 2012), and surface and subsurface runoff (Hughes et al., 2015).

Considering the wide number of variables that characterize flow in rivers, and the complex channel geometry, some assumptions are made in order to apply the conservation principles represented by Saint-Venant (Liggett, 1975):

- uniform velocity in each cross section, varying over the longitudinal direction;
- vertical accelerations not considered; hydrostatic pressure distribution;
- average slope of the channel bottom is sufficiently small to approximate the sine by the tangent of the inclination angle;
- friction losses are not significantly different from those in the steady flow; therefore, the Manning equation or similar can be used;
- invariable bed channel (erosion or sediment depositions neglected);
- longitudinal channel axis represented by a rectilinear reach with low bed slope;
- incompressible flow.

Through these simplification, the Saint-Venant equations may be written as (Liggett, 1975):

$$B\frac{\partial y}{\partial t} + UB\frac{\partial y}{\partial x} + A\frac{\partial U}{\partial x} + U\frac{\partial A}{\partial x} = q \tag{3.1}$$

$$\frac{\partial U}{\partial t} + U\frac{\partial U}{\partial x} + g\frac{\partial y}{\partial x} = \frac{q\,(v_L - U)}{A} + g\,(S_0 - S_f) \tag{3.2}$$

where B represents the cross section top width (m) – which varies with flow depth y (m) -, U is the longitudinal velocity of the flow (m/s), A is the cross section area (m^2), q is the lateral contribution per unit of channel length (m^3/sm), g is the acceleration of gravity (m/s^2), v_L is the input velocity of the lateral contribution in the longitudinal direction

(m/s), S_0 is the bottom slope of the channel (m/m) and S_f refers to the friction slope (m/m). The terms of equation (3.2) represent, respectively: (i) local acceleration; (ii) convective acceleration; (iii) pressure force; (iv) momentum flux of the lateral contribution; (v) gravity and friction force.

Equations (3.1) and (3.2) describe the transient flow, gradually varied, in a channel with irregular cross sections and lateral contribution. The mass conservation equation represents the hydrographs damping effects, that occur due to variation of the storage capacity. The expression for momentum conservation, in its turn, considers the balance of forces acting on the mass of water, which includes gravity, friction, pressure and inertia of the flow. In the latter are expressed the translation effects. The use of the hydrodynamic model allows to represent downstream effects, such as backwater and tides.

The lateral contribution q represents the inputs and outputs of flow. Point contributions include releases (domestic and industrial wastewater or tributaries) and withdraws. Examples of diffuse sources are precipitation and evaporation at surface water, infiltration into the soil and runoff.

Although Manning and Chézy's equations were developed for uniform and steady flow, it is accepted that they are well suited for the calculation of resistance in open channels with unsteady regime (Chow, 1959; Liggett, 1975). Therefore, the Manning equation can be used to estimate the friction slope in m/m (Stepien, 1984):

$$S_f = \left(\frac{nU}{R_h{}^{2/3}} \right)^2 \tag{3.3}$$

where n represents the Manning roughness coefficient and R_h the hydraulic radius (m).

According to USACE (1993), other equations can be used to evaluate the term S_f, such as Einstein (1950), Simons and Sentürk (1976), and ASCE (1975). However, they are avoided due to the presence of sediment-related parameters, and the need for iterative solutions.

The Manning roughness coefficient represents the resistance due to friction in the channel, and several factors interfere in its determination, such as: flow event (drought or flood), presence of vegetation and obstructions (bridges and gates, for example), bed material, irregularity of the cross sections, and river alignment (presence of meanders or rectified stretches) (Chow, 1959; Arcement and Schneider, 1984). Given such uncertainties, this coefficient becomes a calibration parameter.

Given this background, in order to solve the hydrodynamic model, the following information is required: geometry of cross sections and distance between them, lateral contributions, Manning roughness coefficient, and initial and boundary conditions. The solution steps are summarized in figure 9.

The cross sections of the monitoring points are represented by a trapezoidal shape,

assuming to be symmetric. Therefore, the information required are bottom width and side slope. It is also considered that these values do not vary over the time interval analyzed.
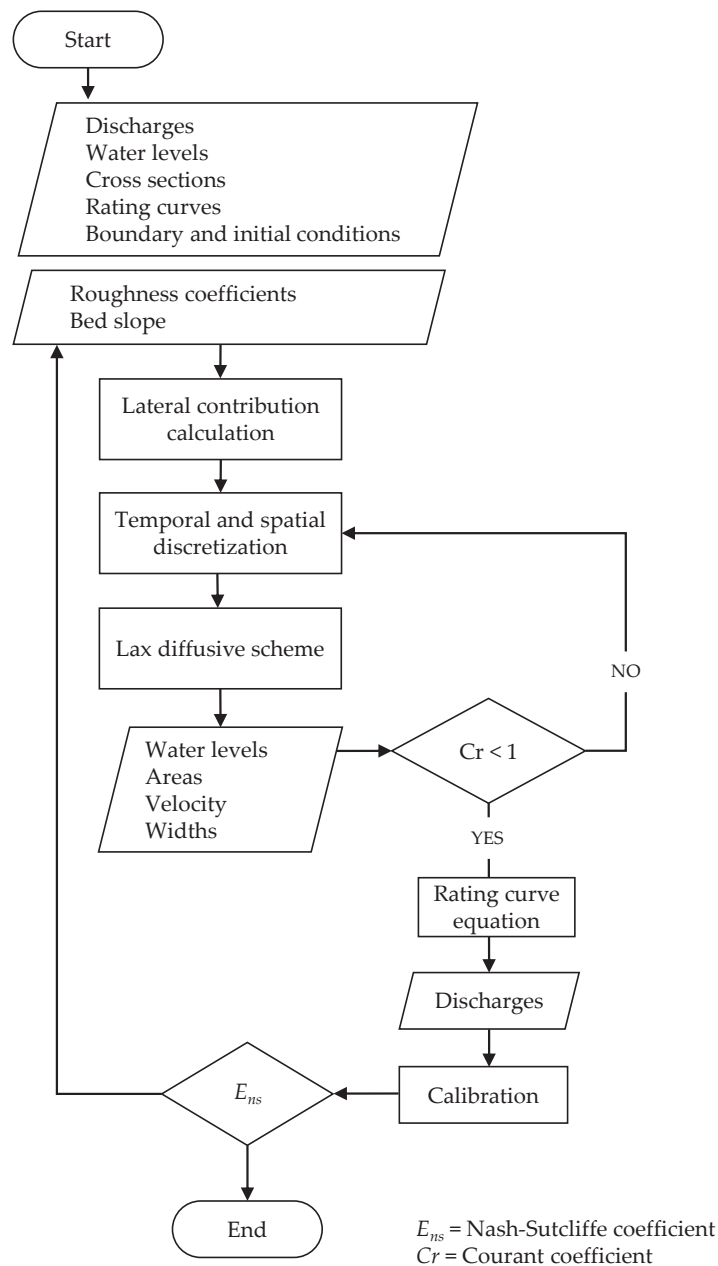
Figure 9: Scheme of the hydrodynamic module

The lateral contribution is estimated by the difference between the hydrographs observed at the known points, divided by the distance between them. In consequence, the total balance between water outlets and inlets in the main channel is evaluated. It is considered that the lateral contribution is evenly distributed along the length of each reach. Because of the lack of information, it is also admitted that the lateral flow

entering or leaving occurs in the same main river velocity, a hypothesis also assumed by Steinstrasser (2005) in a similar study.

Because of numerical stability, the criterion of Courant coefficient (Cr, defined in appendix A.1) defines the required time step. Results are compared as discharges and levels in the monitoring stations, using as measure the Nash-Sutcliffe coefficient (E_{ns}):

$$E_{ns} = \frac{1 - \sum_{t\equiv 1}^{N} [Q_{obs}(t) - Q_{sim}(t)]^2}{\sum_{t=1}^{N} [Q_{obs}(t) - Q_{med}]^2} \tag{3.4}$$

where Q_{obs} is the observed discharge, Q_{sim} represents the discharge simulated with the model and Q_{med} is the average of the flows observed in the period $t = 1,2,...T$ (T is the total number of data); E_{ns} oscillates from $-\infty$ to 1, with an optimum value of 1.

3.2 Synthetic series module

To convert the available data into continuous information, the techniques applied are PCHIP (Piecewise Cubic Hermite Interpolating Polynomial) and smoothing spline interpolation functions, Fourier series, first order autoregressive model and random log-normal series. The first ones explore different techniques to link concentrations to flow conditions and time. The autoregressive model, on the hand, considers statistical indicators of the historical monitoring dataset – mean, standard deviation, quartiles, and concentration with 10 and 90% of occurrence –, dependency on past concentrations, and a random variability (this last part allows to evaluate multiple scenarios). The model of first order is applied because it is a fast and parsimonious method to investigate multiple scenarios; considering the limited dataset available, models of higher orders could lead to more uncertainty, since additional parameters would have to be calibrated.

The Fourier series has the form (MathWorks, 2017a):

$$f(x) = a_0 + \sum_{i=1}^{N} a_1 \cos(ixw) + b_1 \sin(ixw) \tag{3.5}$$

where w is the fundamental frequency of the signal, N is the number of terms (harmonics) in the series.

The smoothing spline s is constructed for the specified smoothing parameter p and the specified weights w_i. This procedure minimizes the function (MathWorks, 2017b):

$$p \sum_{i} w_i(y_i - s(x_i))^2 + (1-p) \int \left(\frac{d^2s}{dx^2}\right)^2 dx \tag{3.6}$$

If the weights are not specified, they are assumed to be 1 for all data points. The parameter p is defined between 0 and 1, automatically set in Matlab.

Figure 10 shows an example of data infilling, using functions for data fitting in Matlab (smoothing spline curve and Fourier series) and the BOD available series.
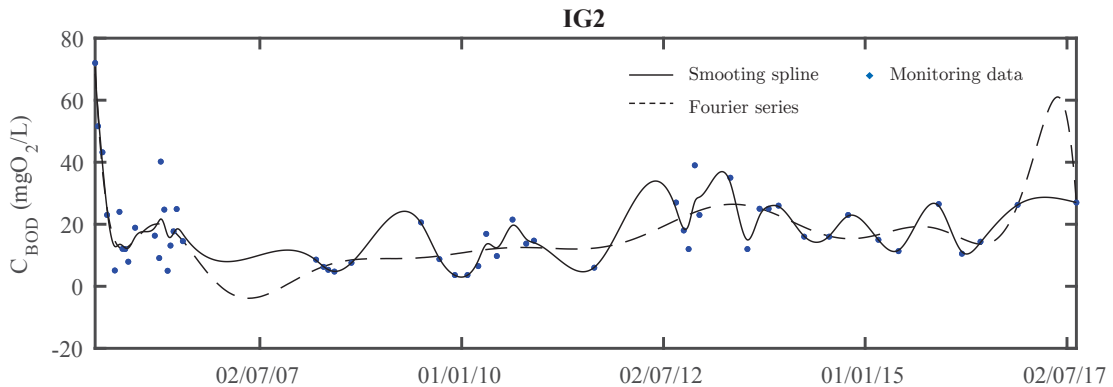
Figure 10: Test for infilling BOD data

The autoregressive process of first order – AR(1) – is represented by a recursive equation determined by random variables, which can be written as (Loucks and Beek, 2017):

$$C_{j+1} = \mu + \rho \left(C_j - \mu \right) + z_j \sigma \sqrt{1 - \rho^2} \qquad (3.7)$$

where C is the concentration at the interval j, ρ is the sample correlation coefficient, which indicates the dependency between time intervals (deterministic component); μ is the mean of C, σ represents the standard deviation of C; z_j is the portion responsible for the random variability in the time series. The distribution of random variables is a key component in this model.

Figure 11 describes the procedure used to estimate the different configurations of AR(1) modeling. This process assumes that the dataset has normal distribution, so, in order to verify this condition, quantile-quantile plots are evaluated (presented in Appendix A.2); they compare quantiles of the sample data with the theoretical values from a normal distribution. The dataset indicate a violation of the normality assumption (qqplots are non-linear); on the other hand, natural logarithms of the dataset are suitable for modeling. Nonetheless, several non-Gaussian first order linear autoregressive models are available in the literature, as reviews by Grunwald et al. (1996). In another step, a value is attributed to the parameter ρ, since the irregular time series does not allow to define sample correlation.

Because AR(1) procedures have a stochastic component (represented by the random distribution with zero mean and unit variance), a thousand different sets of concentrations are generated (except test T10, that generates data at each 50 s; due to processing limitations, only twenty options are evaluated in this test, instead of thousand). In or-

der to eliminate the influence of initial values, the series is generated with extra twenty values, and the first ones are ignored. Because multiple scenarios are generated, different criterion are used to select a series (as presented in figure 11). At last, all time series generated via AR(1) are considered as daily samples.

The fitted models are validated through residual analysis: independence (Portmanteau test), homoscedasticity (Levene test) and normality (Kolmogorov-Smirnov test). The results of each test are presented in appendix A.2.
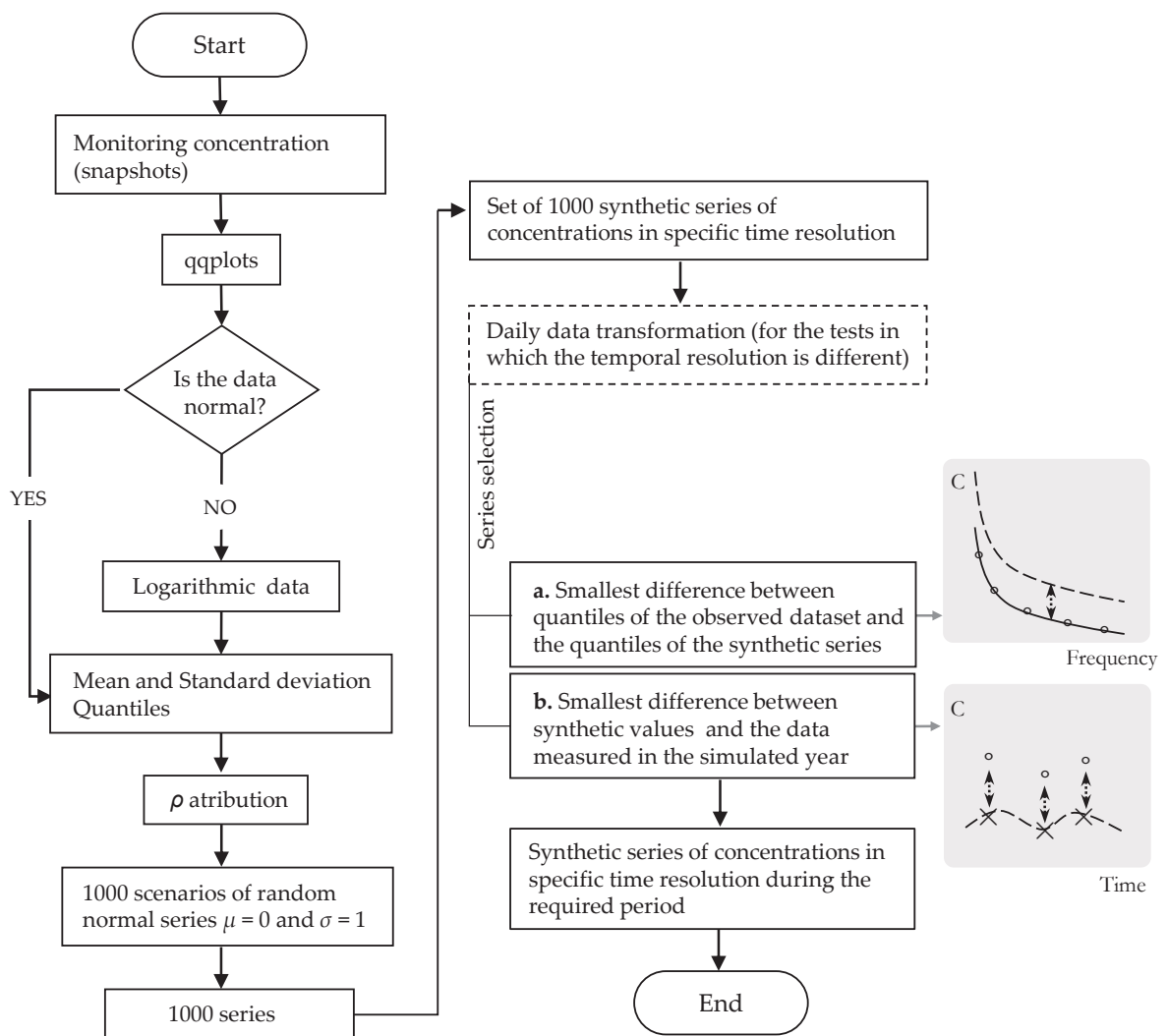
Figure 11: Scheme of time series generation and selection using the AR(1) model

Table 5 summarizes the different configurations for the AR(1) modeling explored in this research, besides tests with random log-normal series and interpolation functions (PCHIP, Fourier series and smoothing spline). The tests presented in this section are applied for BOD in section IG2 (boundary condition).

Test T1 and T2 have the objective to identify the influence of seasonal variation (in south hemisphere) for AR(1) series generation. Tests T3, T8, T9 and T10 explore

effects of temporal interval generation, while T4 explores sampling frequency versus time series prediction with AR(1). T5 shows if in fact the AR(1) is representing the range of observed concentrations. Experiment T6 tests the model ability in simulate the entire period of available data, using the same procedure applied in test T1; besides option *a* and *b* for series selection among the thousand series (as presented in figure 11), this test also compares: *c.* series closer to the data measured in 2013-2015, *d.* euclidean distance between the monitored series and simulated, and *e.* dynamic time warping (an algorithm to compare time series; Mathworks, 2018).

A few indexes are used as reference to compare the fitted series, as presented in table 4; they are calculated considering only the available measured data in the period of simulation and the corresponding simulated data. The exception is test T6: option *c* compares the data measured in the period 03/18/2013 to 03/18/15, *d* and *e* compare all available data.

The coefficient R reflects the linear relationship between the datasets, while the $RMSE$ computes differences between simulated and observed values. The Percent bias ($PBIAS$) demonstrates the average tendency of the simulated data to be larger or smaller than observed values; low-magnitude $PBIAS$ indicate accurate model results, with an optimal value of zero; positive values indicate model underestimation bias, and negative values indicate model overestimation bias (Gupta et al., 1999). $MAPE$ is the average of absolute percentage errors, with optimal value of zero; upper level has no restriction (Parmar and Bhardwaj, 2014).

Table 4: Measures for goodness of fit for the generated time series

Statistic	
Pearson correlation coefficient (R)	$R = 1/(N-1) \sum_{i=1}^{N} \left(\overline{M_i - \mu_M} / \sigma_M \right) \left(S_i - \mu_S / \sigma_S \right)$
Root Mean Square (RMSE)	$RMSE = \sqrt{\sum_{i=1}^{N} (M_i - S_i)^2 \Big/ N}$
Mean Absolute Percentage Error (MAPE)	$MAPE = \sum_{i=1}^{N} \left| \frac{M_i - S_i}{M_i} \right| \Big/ N$
Percent bias (PBIAS)	$PBIAS = \sum_{i=1}^{N} (M_i - S_i) \, 100 \Big/ \sum_{i=1}^{N} M_i$

where N = number of data, S = simulated concentration and M = observed concentration; μ_M and σ_M = are the mean and standard deviation of M, respectively, and μ_S and σ_S = are the mean and standard deviation of S.

Table 5: Summary of the tests performed to estimate synthetic pollutographs

Test	Description	Objective
T1	Seasonal μ and σ, $\rho = 0.5$, generating daily data	Check if seasonal parameters generate a closer series to data than using values of the entire dataset
T2	μ and σ of entire dataset, $\rho = 0.5$, generating daily data	
T3	μ and σ of entire dataset, $\rho = 0.5$, generating weekly data; daily data is generated using linear interpolation	Explore effects of temporal interval generation
T4	μ and σ of dataset for 2010, $\rho = 0.5$, generating daily data	Estimate daily concentration if the monitoring data is available only for the simulated year
T5	Seasonal μ and σ, $\rho = 0.5$, generating daily data (same conditions as T1); the comparisons consider all data as if were measured in 2010	Check if the model fitted in T1 represents the range of observed concentrations
T6	Seasonal μ and σ, $\rho = 0.5$ (same conditions as T1); generating daily data for the entire period	Verify the capacity of the model adjusted in test T1 to represent the entire period of monitoring
T7	μ and σ of entire dataset, $\rho = 0.8$, generating daily data	Explore other persistence scenarios
T8	Seasonal μ and σ, $\rho = 0.5$, generating hourly data (daily concentrations are means)	Explore effects of temporal interval generation
T9	Seasonal μ and σ, $\rho = 0.9$, hourly data; daily data is mean	Explore other persistence of hourly series generation
T10	μ and σ of entire dataset, $\rho = 0.99$, generating data at each 50 s; 20 options	Explore effects of temporal interval generation
T11	Random series as a two - parameters log-normal distribution with same seasonal μ and σ of the observed dataset; generating daily data; no extra data generated	Explore other series distribution
T12	Hydrib AR(1) model with $\mu = 0$, $\sigma = 1$; random component is a two- parameters log-normal distribution with same seasonal μ and σ of the observed dataset; $\rho = 0.5$; no extra data generated	Explore other random component in the AR(1) model
I1	PCHIP (Piecewise Cubic Hermite Interpolating Polynomial); daily data	Explore the relationship discharge and concentration, using PCHIP interpolation
I2	Fitting with a Fourier series truncated in the fifth-term, using as parameters daily discharge and the monitored concentration; daily data	Explore the relationship between discharge and concentration, using a Fourier series
I3	Fitting with a smoothing spline, in which the parameters are date and observed concentrations; daily data	Explore interpolation using dates

3.3   Water quality module

The distribution of a constituent in water is represented by the physical greatness concentration. Considering the processes of advection, dispersion, reaction linked to the principles of mass conservation, concentration of a given substance in a system with predominant flow in the longitudinal direction may be expressed mathematically by:

$$\frac{\partial C}{\partial t} + U\frac{\partial C}{\partial x} - \frac{D}{A}\frac{\partial A}{\partial x}\frac{\partial C}{\partial x} - D\frac{\partial^2 C}{\partial x^2} \pm F = 0 \tag{3.8}$$

where C is the mean cross-sectional concentration of a given constituent (kg/m^3), D represents the longitudinal dispersion coefficient (m^2/s), and F is the term for mass transformations and external loads (kg/m^{3}s).

The term of transformations in equation (3.8) includes variations that occur independently of the transport process. It represents losses or gains of mass due to chemical processes (chemical reactions in general), physical (decantation of particulates, for example) or biological processes (such as growth and death of algae) within the system. In general, such processes are described by first order kinetic reactions and values based on the literature (Chapra, 1997).

Using the explicit method FTCS for numerical solution, the algorithm to solve the water quality module follows the process summarized in figure 12. The numerical solution for the water quality module has been validated comparing a few simplified cases with analytical solutions (Ferreira, 2015).
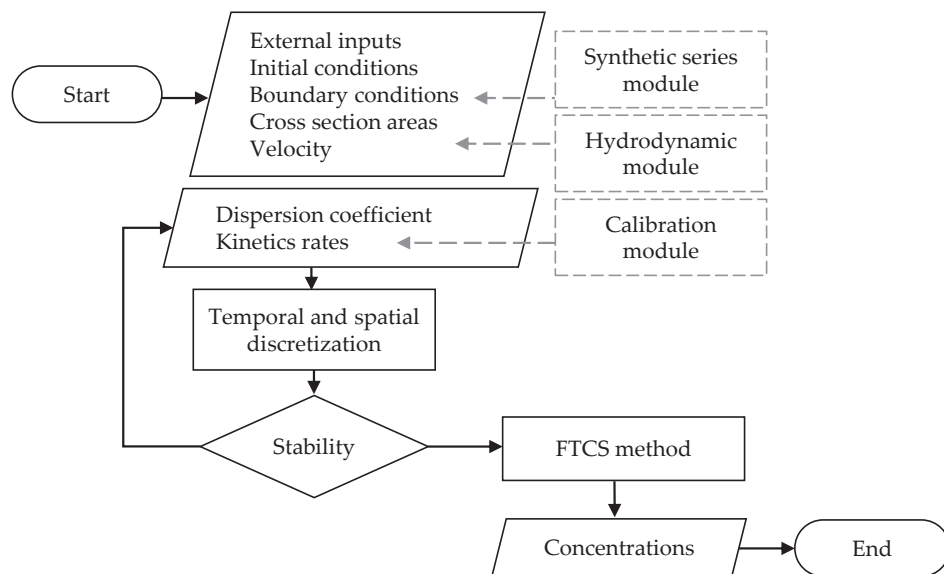
Figure 12: Scheme of the water quality module

The parameters simulated in this research are biochemical oxygen demand (BOD), dissolved oxygen (DO), organic nitrogen (N-org), dissolved organic carbon (DOC), labile and refractory dissolved organic carbon (LDOC and RDOC, respectively), particulate organic carbon (POC), labile and refractory particulate organic carbon (LPOC and RPOC, respectively). The term F for each of these constituents can be expressed by:

(a) BOD (Brown and Barnwell, 1987):

$$F_{BOD} = -(K_d + K_s)C_{BOD} + W_{BOD}/V \tag{3.9}$$

where C_{BOO} is the BOD concentration (mg-O$_2$/L), K_d represents the deoxygenation rate (d^{-1}) and K_s is a coefficient for BOD removal by sedimentation (d^{-1}); W_{BOD} is the BOD load entering the system (kg/d), and V is the volume (m^3).

(b) N-org (Chapra, 1997):

$$F_{N_{org}} = -K_{oa}C_{N_{org}} - K_{so}C_{N_{org}} + W_{N_{org}}/V \tag{3.10}$$

where $C_{N_{org}}$ is the organic nitrogen concentration (mg/L), K_{so} is the sedimentation coefficient of organic nitrogen (d^{-1}), and K_{oa} is a coefficient for conversion of organic nitrogen to ammonia (d^{-1}); $W_{N_{org}}$ is the N-org load entering the system (kg/d).

(c) DO (Brown and Barnwell, 1987):

$$F_{DO} = K_a (O_s - C_{DO}) - K_d C_{BOD} - K_4/H - \alpha_5 \beta_1 N_a - \alpha_6 \beta_2 N_b \tag{3.11}$$

where K_a is the reaeration coefficient (d^{-1}), O_s represents the dissolved oxygen saturation concentration (mg-O$_2$/L), C_{DO} is the DO concentration (mg-O$_2$/L), K_4 is a rate for oxygen demand by the sediment (gO$_2$/m^{2}d), H defines the water mean depth in the channel (m), α_5 represents an oxygen rate consumed by each unit of oxidized ammonia (mg-O$_2$/mg-N$_2$), α_6 is a rate of oxygen consumed by each unit of oxidized nitrite (mg-O$_2$/mg-N), β_1 is ammonia oxidation rate (d^{-1}), β_2 is a nitrite oxidation rate (d^{-1}), N_a is ammonia concentration (mg-N$_2$/L) and N_b defines a concentration of nitrite (mg-N$_2$/L); N_a and N_b are considered as the monitored dataset average in each section.

The equations to represent DOC and POC internal processes are based on an adapted model suggested by Chapra (1997). Originally, these expressions were proposed to represent nutrient/food-chain interactions in a stratified lake. In a second strategy to represent organic carbon in rivers, segmentation of labile and refractory fractions are estimated through expressions proposed by Knapik et al. (2016). This study presented a

model under steady state to represent this four components in the Iguaçu river, including settling and resuspension of particulate fractions, mineralization to inorganic forms (ammonia, phosphate, inorganic carbon), dissolution from particulate do dissolved fractions, and decay between labile and refractory organic carbon (figure 13).
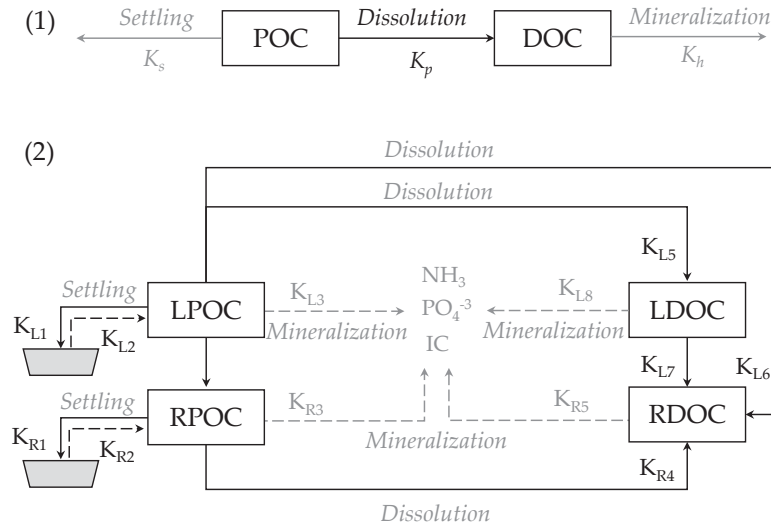
Figure 13: Conceptual segmentation for modeling organic carbon fractions with strategies (1) and (2). Source: Chapra (1997) and Knapik *et al.* (2016)

(d) DOC (Chapra, 1997):

$$F_{DOC} = K_p C_{POC} - K_h C_{DOC} + W_{DOC}/V \tag{3.12}$$

where C_{DOC} is the DOC concentration (mg-C/L), C_{POC} is particulate organic carbon (mg-C/L); K_p is the particulate organic carbon dissolution rate (d^{-1}) and K_h represents the dissolved organic carbon hydrolysis (mineralization) rate (d^{-1}); W_{DOC} is the DOC load entering the system (kg/d); C_{POC} is considered as the monitored dataset average in each section.

(e) POC (Chapra, 1997):

$$F_{POC} = -K_p C_{POC} - K_s C_{POC} + W_{POC}/V \tag{3.13}$$

where K_s is particulate organic carbon sedimentation rate (d^{-1}); W_{POC} represents POC load entering the system (kg/d).

(f) LDOC:

$$F_{LDOC} = K_{L5} C_{LPOC} - K_{L7} C_{LDOC} - K_{L8} C_{LDOC} + \frac{W_{LDOC}}{V} \tag{3.14}$$

(g) RDOC:

$$F_{RDOC} = K_{L6}C_{LPOC} - K_{L7}C_{LDOC} + K_{R4}C_{RPOC} - K_{R5}C_{RDOC} + \frac{W_{RDOC}}{V} \qquad (3.15)$$

(h) LPOC:

$$F_{LDOC} = -K_{L1}C_{LPOC} - K_{L3}C_{LPOC} - K_{L4}C_{LPOC} - K_{L6}C_{LPOC} + \frac{K_{L2}}{H} + \frac{W_{LDOC}}{V} (3.16)$$

(i) RPOC:

$$F_{RPOC} = K_{L4}C_{LPOC} - K_{R1}C_{RPOC} - K_{R3}C_{RPOC} - K_{R4}C_{RPOC} + \frac{K_{R2}}{H} + \frac{W_{RPOC}}{V} \ (3.17)$$

where C_{LDOC}, C_{RDOC}, C_{LPOC}, C_{RPOC} are LDOC, RDOC, LPOC and RPOC concentrations (mg-C/L), respectively; H is depth (m), K_{L1} is settling of LPOC (d^{-1}), K_{L2} represents resuspension of LPOC ($gm^{-2}d^{-1}$), K_{L3} is mineralization of LPOC (d^{-1}), K_{L4} is decay of LPOC to RPOC (d^{-1}), K_{L5} is decay of LPOC to LDOC (d^{-1}), K_{L6} is decay of LPOC to RDOC (d^{-1}), K_{L7} is decay of LDOC to RDOC (d^{-1}), K_{L8} is mineralization of LDOC (d^{-1}); K_{R1} is settling of RPOC (d^{-1}), K_{R3} is mineralization of RPOC (d^{-1}), K_{R4} is decay of RPOC to RDOC (d^{-1}), K_{R2} is resuspension of RPOC ($gm^{-2}d^{-1}$), K_{R5} is mineralization for RDOC (d^{-1}); W_{LDOC}, W_{RDOC}, W_{LPOC}, W_{RPOC} are LDOC, RDOC, LPOC and RPOC loads entering the system (kg/d), respectively.

4 Study case: Iguaçu river

The Upper Iguaçu watershed is located in the metropolitan area of Curitiba, capital of Paraná (Brazil), and has approximately 2 million inhabitants. It covers 3000 km², where 30% of the state population is concentrated. The control volume in this research is 90 km in length, draining regions with different land use: urban occupation, temporary agriculture, meadows and other vegetation categories.

The dataset selected is from 2005 to 2016, with approximately four annual campaigns. The sections of interest (IG2 – *upstream*, IG3, IG4, IG5 and IG6 – *downstream*), shown in figure 14, are located along the main channel, separated by distances between 18 and 24 kilometers. The points IG2, IG3, IG4, IG5 and IG6 are located at the stations: Bridge BR-277 (65009000), Umbarazinho bridge (65017006), Wastewater Treatment Plant Cachoeira (65019980), Guajuvira bridge (65025000) and Balsa Nova (65028000).
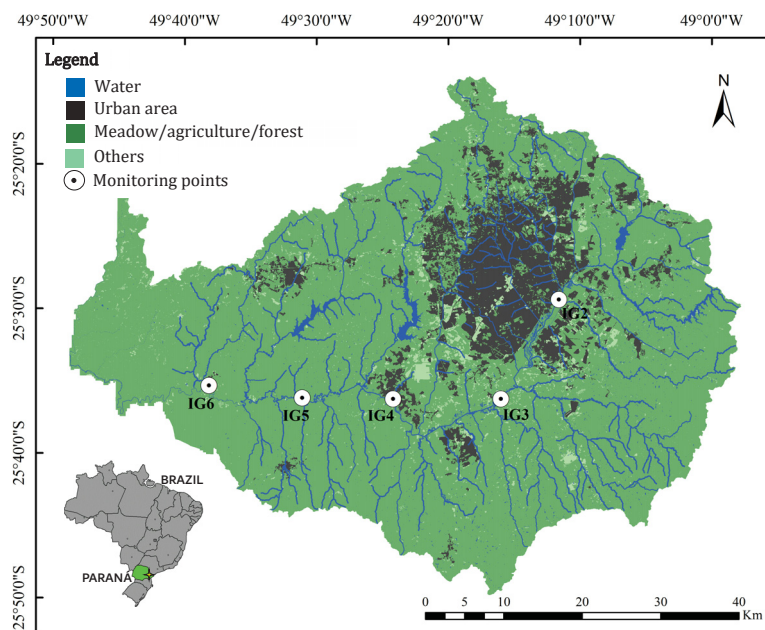
Figure 14: Upper Iguaçu watershed and monitoring points

Sections IG2, IG3 and IG4, inserted in the urban area, receive significant contributions of domestic and industrial wastes, and inputs from surface runoff; points IG5 and IG6, on the other hand, drain an agricultural region, receiving loads from tributaries less impacted than the upstream reach (table 6).

Table 6: Percentage of land occupation by each of the main land classes in the Upper Iguaçu watershed

	Urban	Agricultural	Forest
IG2	57.2	42.8	0
IG3	28.6	63.4	8
IG4	30.7	63.1	6.2
IG5	19.6	68	12.4
IG6	9.7	72.5	17.8

Source: Porto et al. (2007)

4.1 Input data

The dataset for water quality parameters in the period of 2005 to 2017 (concentration and respective discharge), collected at approximately quarterly campaigns, are presented in Figure 15. The number of pairs oscillate between 36 and 55 for each section of interest (details in table 7). Further information about the data (methods used for sample collection, processing, and analysis) are described in Knapik (2014).

Figure 16 presents the BOD dataset as seasonal measures (mean, standard deviation and medians); this parameter is the main study case to investigate synthetic series generation in this research.

The underlying hypothesis assumed for the tests is that the twelve years of monitoring data in the Iguaçu river adequately represent the variability of concentrations. Appendix A.3 shows an analysis exploring this assumption. In a similar analysis, Williams et al. (2014) presumed that a set of five years sufficiently described sample median and log-normal standard deviation of annual concentrations; the number of samples available per year ranged between one and eight in their study.

Flow data is available as daily records, as shown in figure 17 for the periods 01/01/2010 to 12/01/2010, 01/01/2011 to 12/01/2011 and 03/18/2013 to 03/18/2015 (for this latter data for IG3 and IG4 are not available). These periods represent hydrologically contrasting conditions, which is important to test the model ability in reproduce different scenarios.
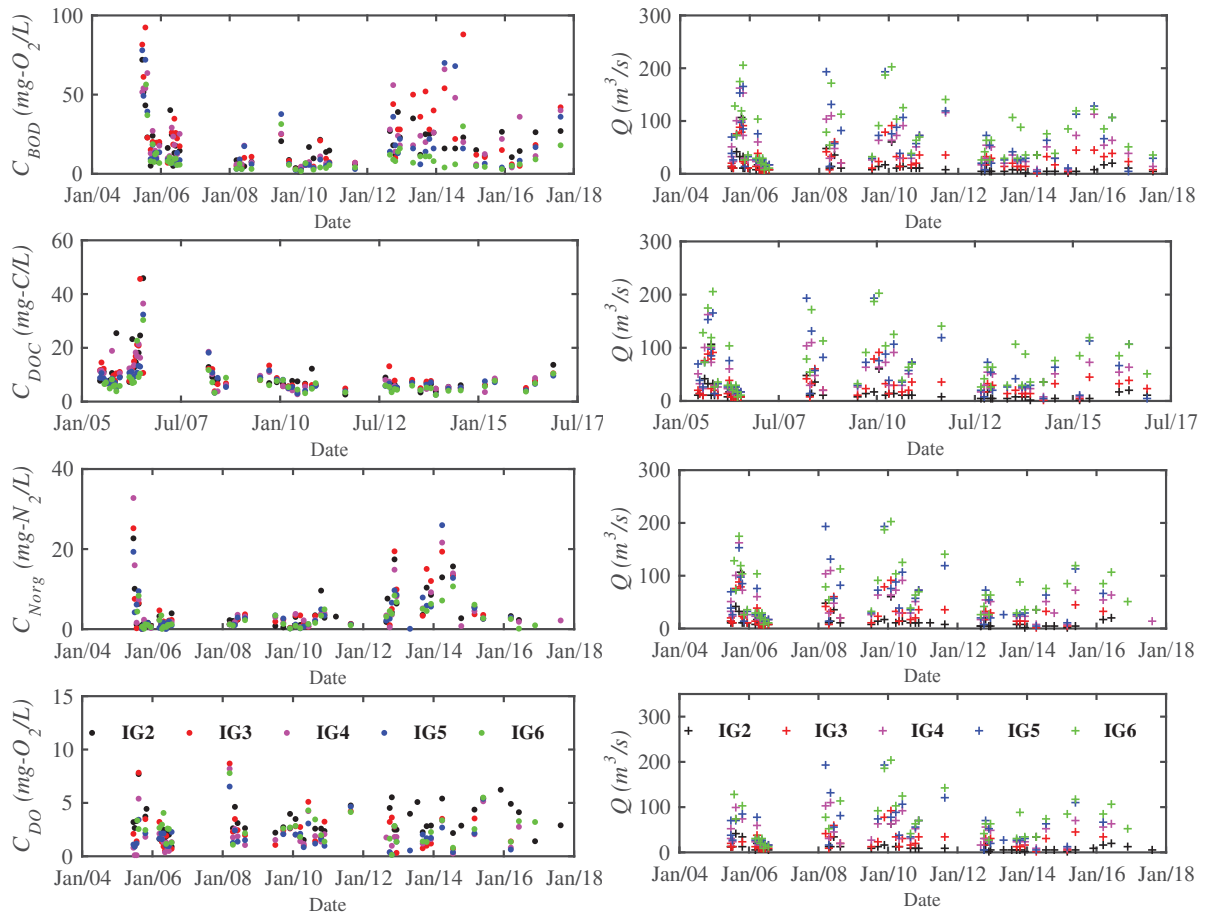
Figure 15: Monitoring data for BOD, N-org, DOC and DO concentrations (circle, first column) and respective estimated discharge (plus sign, second column) in the Iguaçu river

Table 7: Number of samples

	BOD	DOC	N-org	DO
IG2	55	52	50	48
IG3	54	51	46	42
IG4	53	45	47	41
IG5	54	51	47	41
IG6	50	47	43	36
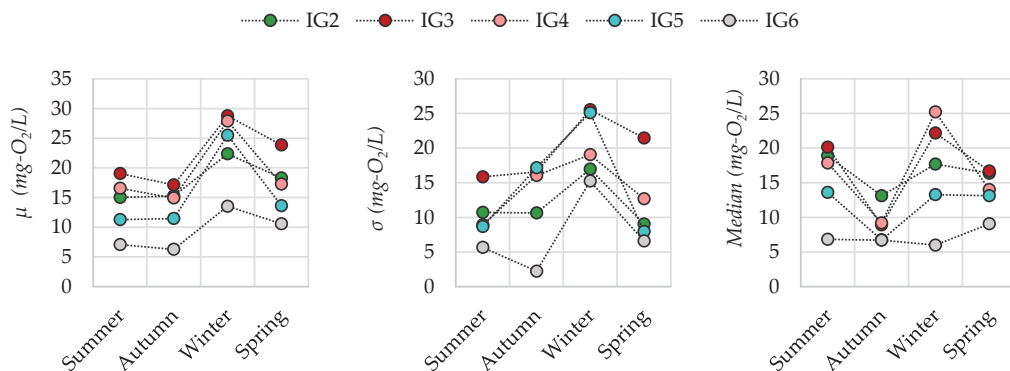
Figure 16: Seasonal monitoring data: means, standard deviation and medians
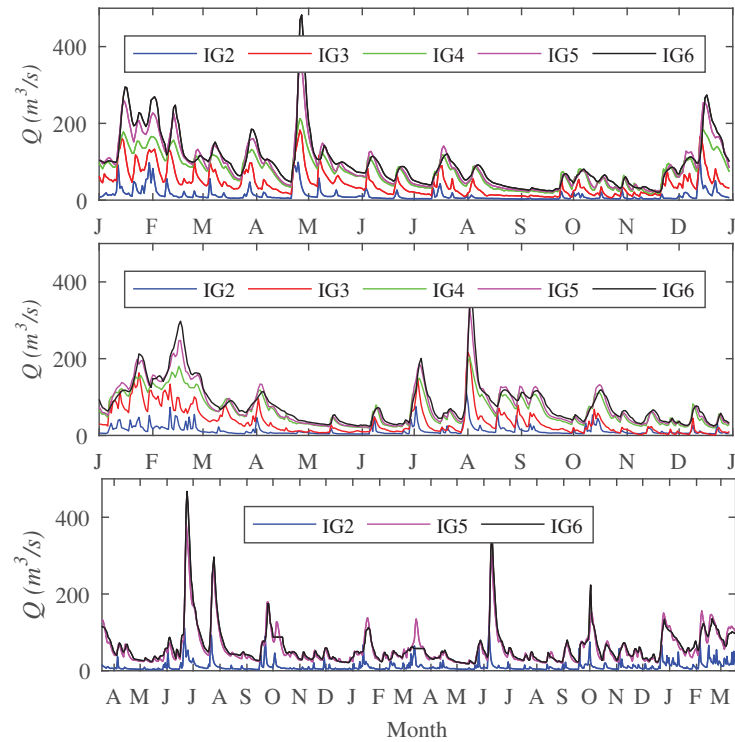
Figure 17: Daily flow data in the Iguaçu river for the years 2010, 2011 and 2013-2015

The observations of level and flow measurements were conducted during the period presented in table 8, although it does cover the amplitude of maximum flows. The rating curve extrapolation is performed through the logarithmic method, which is based on equations of uniform flow, assuming a regular cross section (figure 18).

Table 8: Period of observation level - discharge

Section	Period	Number of observations
IG2	13/12/1973 to 11/12/2013	213
IG3	13/12/1973 to 01/06/2016	134
IG4	16/09/1999 to 12/12/2012	48
IG5	17/08/1973 to 30/03/2016	143
IG6	17/08/1973 to 24/03/2016	120

Source: Adapted from Instituto das Águas do Paraná (2017)
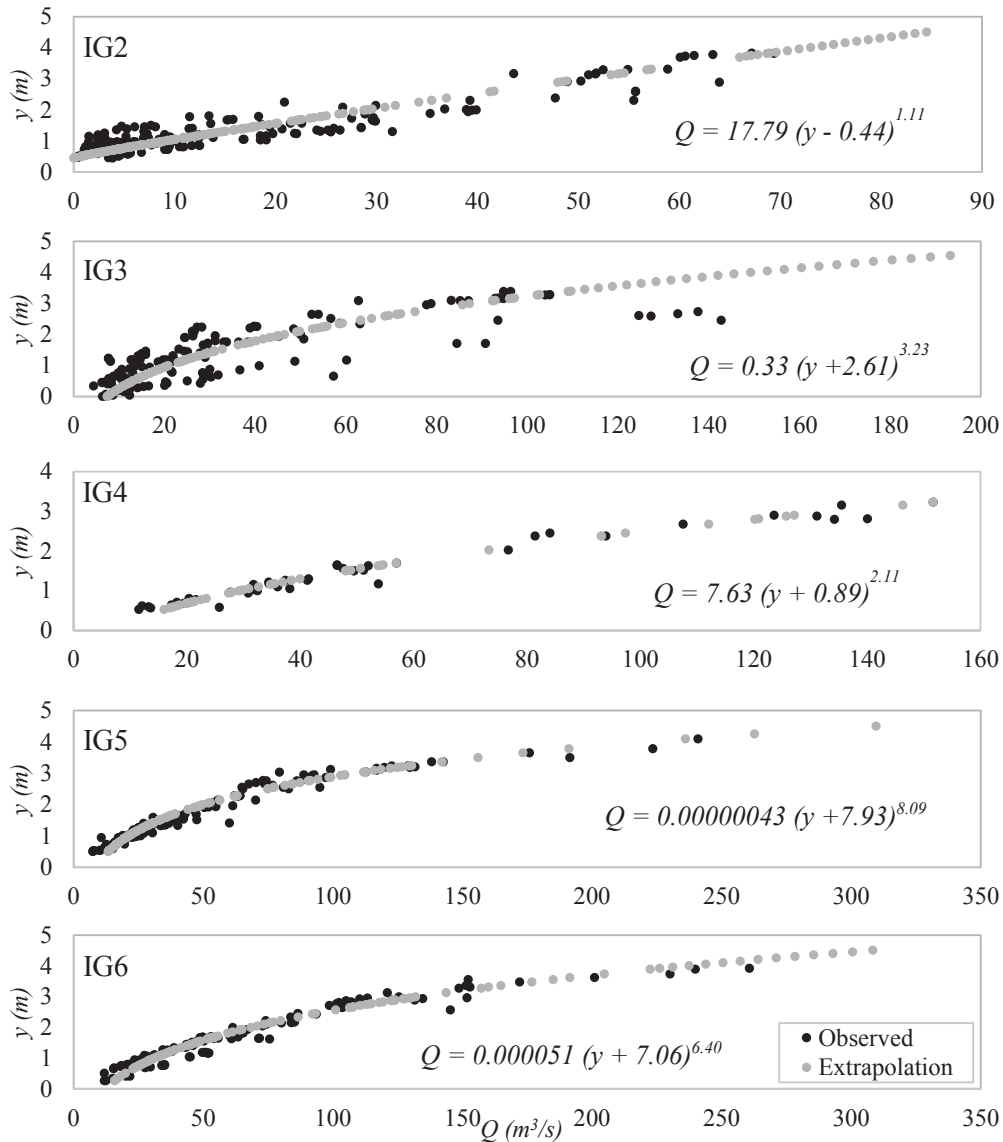
Figure 18: Rating curves for the monitoring sections; fitted equations have the form $Q = a(y - y_0)^r$, where y is elevation corresponding to discharge Q, y_0 represents the elevation for a null flow, a and r are constants determined using the method of least squares.

Pollutant loads released in the main channel are estimated based on population data and land use of each tributary's sub-basin. Population is divided into inhabitants with i. sewage collection and treatment, ii. sewage collection and no treatment, and iii. no sewage collection (data estimated by Porto et al., 2007 in the Upper Iguaçu watershed); the loads not collected are considered as diffuse pollution. Three land use categories are used: urban, agriculture and forest (table 6). Export rates associated with each land use are adopted from literature (table 9). As a way of considering the decay of the pollutants load along the watershed, coefficients of attenuation as a function of distance are applied (Munafo et al., 2005; Cecci et al., 2007) – parameter DI_i in table 10.

A schematic diagram with the main input's location is reproduced in figure 19),

while table 10 presents the equations used to estimate their contributions. The total load in each reach (IG2-IG3, IG3-IG4, IG4-IG5 and IG5-IG6) is considered as constant inputs linearly distributed; its variation over time for unsteady simulations are due to dilution fluctuations, provided by the hydrodynamic module.

Table 9: Export coefficients used in the simulations

Parameter	Assumed values[1]		
	Urban	Agricultural	Forest
DBO	100	100	150
N-org [2]	146	146	110
DOC	500	1000	1000
POC	500	1000	1000

(1) based on literature range suggested by Von sperling (2007), Chapra (1997) and Mattsson *et al.* (2009); (2) export coefficient estimated from total nitrogen values
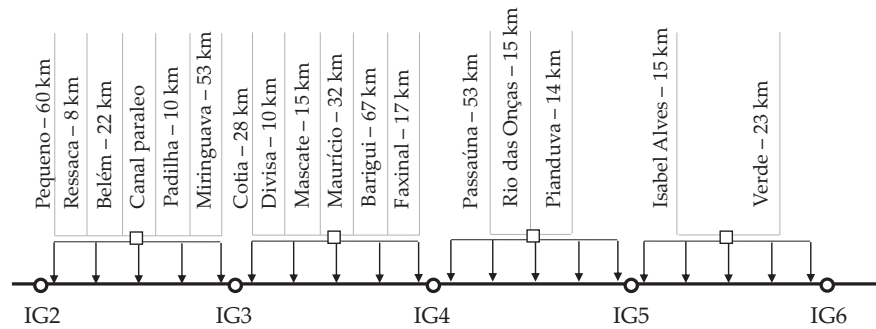
Figure 19: Topological diagram of the Upper Iguaçu Basin, with main tributaries and monitoring sections. Source: adapted from Porto et al. (2007)

Table 10: Equations used to estimate lateral sources

Description	Equation
Domestic wastewater flow (BOD and N-org)	$Q_D = pop_{reach} \times QPC \times R/86400$
Domestic wastewater load (BOD and N-org)	$W_D = Q_D \times C_{capita}$
Domestic wastewater load (DOC)	$W_D = Q_D \times W_{capita}$
Diffuse pollution attenuation	$DI_i = e^{-D_i \times k}$

where Q_D is the domestic effluent flow (L/s); pop_{reach} is the population from each reach (inhab); QPC is per capita flow (L/inhab.d); Rs is the coefficient of sewage return; W_D is the load from non-treated effluent (mg/s); C_{capita} is the per capita concentration (mg/L); W_{capita} is the per capita load (g/inhab.d); DI_i ia a normalized distance between a cell in the watershed and the river network, Di is the distance between the cell and the river network, and k is a constant value set at 0.090533; sources: Cecci et al. (2007) and Von sperling (2007)

The values adopted for each parameters are: $QPC = 150$ L/inhab.d, $Rs = 0.7$, treatment removal efficiency: 80% for BOD and 30% for N-org; C_{capita} for BOD = 280 mg-O$_2$/L and N-org = 15 mg-N/L; for DOC and POC estimation, W_{capita} considered is 3.5 g-C/inhab.d (DOC) and 2 g-C/inhab.d (POC) for treated wastewater and 7 g-C/inhab.d (DOC) and 20 g-C/inhab.d (POC) raw wastewater.

To estimate loads for labile and refractory fraction, it is adopted the same hypothesis by Knapik et al. (2016) – who simulated these parameters under steady state in the Iguaçu river: (i) labile fractions refer to raw effluent and the fraction collected but not treated; (ii) refractory fractions refer to the treated wastewater; (iii) diffuse sources contribute only with refractory fractions into the river. The values adopted for W_{capita} are: LDOC = 5 g-C/inhab.d and LPOC = 4 g-C/inhab.d (raw wastewater), RDOC = 8 g-C/inhab.d and RPOC = 20 g-C/inhab.d (treated wastewater). All values assumed are based on the suggested range by: Von sperling (2007), Chapra (1997) ans Servais *et al.* (1999).

Final results are exhibited in figure 20 for the studied periods, confirming that reaches IG2 to IG4 receive larger amounts of mass flux, where anthropic activity is more intense.
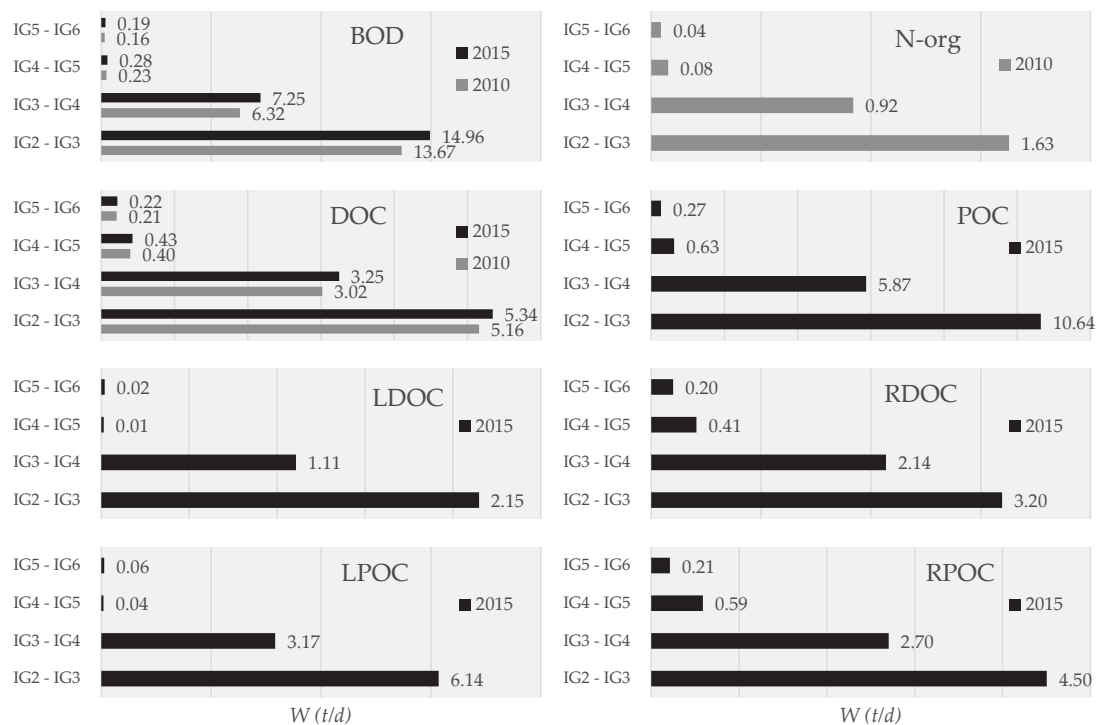
Figure 20: Total lateral load (W) estimated within each reach

4.2 Challenges of modeling the Iguaçu river at Curitiba's metropolitan region

The Upper Iguaçu basin drains an urban area of Curitiba and metropolitan region, daily receiving high polluting loads. This watershed represents a typical case among several cities, where irregular occupation of floodplains and water supply areas, as well as an insufficient collection and treatment of wastes, compromise water quality and generate problems for water supply systems, wastewater treatment and urban drainage (Fernandes, 2013).

This system remains subject of several studies (e. g. Knapik et al., 2016; Coelho et al., 2017; Mizukawa et al., 2017), mostly based upon steady state analysis. Despite the advances in understanding this particular environment, these researches emphasize the need of integrating quali-quantitative aspects in critical basins, in order to generate information to assist models calibration, monitoring strategies and data evaluation for sustainable and adaptable management.

The selected parameters to characterize water quality in this system are traditional indicators of organic matter (BOD), nutrient (N-org) and oxygen distribution (DO). In addition, organic carbon modeling is investigated. Organic carbon is an important aspect in water quality modeling, since its decomposition can significantly affect the oxygen concentration in the environment. Furthermore, because many toxic components associate with organic matter, its dynamics in a system is directly linked to generation, transport and fate of organic carbon (Chapra, 1997).

Although not currently considered in the Brazilian legislation, Knapik et al. (2016) argues that fractions of organic carbon (particulate and dissolved) are less subjective and more reliable than traditional water quality parameters (such as biochemical or chemical oxygen demand). In addition, the study of labile and refractory parts allows to identify areas or periods vulnerable to water quality impairment: labile organic carbon is expected to be found in areas with anthropic interference, while refractory are identified mainly in regions where self-purification is active.

5 Results and discussion

The median isn't the message.

– Stephen Jay Gould

The algorithms are employed using the Matlab software, conducting primary simulations for one year – 2010 (hydrodynamic, synthetic and water quality modules for BOD, DO, N-org and DOC). To validate the definition of boundary condition and the calibration procedure, simulations for a two year period are also presented – 2013 to 2015 (hydrodynamic, synthetic and water quality modules for BOD, DOC, POC, LDOC, RDOC, LPOC and RPOC). These periods are selected due to data availability and to maintain feasible analysis, since larger intervals require extend computational time for processing.

In order to address the objectives and guarantee an answer to the hypothesis, this section is organized in three main modules, as summarized in table 5. Section 5.1 presents hydrodynamic simulations, while item 5.2 compiles results of synthetic series generation. The following item, section 5.3, presents simulated concentrations through deterministic modeling under steady and unsteady state – the latter approach uses a new calibration strategy (discussed in item 5.3.1). At last, sections 5.4 and 5.5 reinforce the role of boundary conditions and calibration challenges in modeling transport of substances in rivers.

Extra tests conducted during the research are presented in the appendix: caffeine modeling, based on steady behavior and first order decay (item A.4); solution of Saint-Venant and advection-dispersion-reaction equations using Fortran language – in order to improve efficiency and allow future extended simulations (item A.5); magnitude of the equations terms (item A.6).

Table 11: Summary of experiments in this thesis

Module	Objective	Thesis item
Hydrodynamic	*Define input data (velocity and cross section areas) for the WQ[1] model*	
Flow routing simulations, comparing water levels and discharges	Explore the effects on water depth and discharge simulations with different boundary conditions	5.1
Sensibility analysis for the Manning roughness coefficient	Study the effects of different calibration sets	A.10
Discharge simulation with the HEC-RAS software	Verify the explicit numerical solution used in SIHQUAL, and validate the channel geometry hypothesis[2]	A.7
Synthetic series	*Define input data (upstream boundary condition) for the WQ model*	
Data fitting through PCHIP, Fourier series and smoothing spline functions	Assess concentration based on flow conditions and time	5.2
Synthetic series generation using multiple configurations based on the first order autoregressive process	Verify if a statistical approach is able to describe concentration time series of river sections	
Calculations of different indexes	Compare the generated time series	
Pre and post synthetic series analysis	Verify the input data normality and check residuals of fitted models	A.2
Application of the AR(1) model in other case studies	Validate the method to estimate synthetic series, using other regions and different data availability	A.8
Monitoring boxplots versus data quantity	Evaluate the input data representativeness assumption	A.3
Water quality module	*Simulate concentrations over time and space*	
Estimation of lateral inputs	Define input data	4.1
Steady state water quality modeling	Comparison with transient results	5.3.2
Development of a procedure to define temporal variation of kinetic rates	Propose a calibration procedure for unsteady modeling	5.3.1
Modeling water quality parameters under unsteady state	Validate the proposed integrated modeling approach, applied to several WQ parameters	5.3.3/5.3.3.6
Comparison of unsteady state water quality modeling using different boundary conditions	Evaluate the effects of deterministic modeling with different synthetic series as input; verify the role of upstream boundary conditions in transport of pollutants	5.3.3/5.4
Unsteady simulation with different scenarios of temporal variability for kinetic processes and lateral loads	Assess time variation of theses aspects and its effect on transport of pollutants	5.5
Sensibility analysis	Quantify the role of kinetic processes, lateral loads, and upstream boundary conditions	A.10

[1] Water Quality [2] HEC-RAS uses an implicit solution – comparison with SIHQUAL shows results very similar, under the same simulation conditions

5.1 Hydrodynamic simulations

Figures 21 to 26 indicate the results for three periods: 01/01/2010 to 12/01/2010, 01/01/2011 to 12/01/2011 and 03/18/2013 to 03/18/2015 (for this latter data for IG3 and IG4 are not available). Through the procedure of trial and error, calibration was performed using the Manning roughness coefficient for the period of 2010 (Ferreira et al. 2016). Results show that the model follow fluctuations over time and space, and that calibration is adequate for other periods (2011 and 2013-2015).

To define upstream boundary condition, two approaches are tested in this research: time series of levels and rating curves. The first procedure is generally better: considering discharge comparisons, for 2010 in IG5, for example, E_{ns} is 0.74 using the first boundary condition, and 0.39 with rating curves (figures 21 (e) and 22 (e), respectively).

Hydraulics modifications occur at upstream from IG5, where the channel have meanders. Because the model propagates information from IG2 to downstream, such modifications may be not properly represented. The bed slope variation or the rating curve, for example, may contribute for the overestimation of some peaks, as observed mainly in section IG5. This is reflected in the comparisons of discharges for IG5, for which was calculated a E_{ns} of 0.39. However, the Nash-Sutcliffe coefficient is sensible to maximum values, that the rating curve seems to not represent adequately. Oliveira et al. (2016) showed that there is a limitation when using the concept of rating curves in transient simulations, finding discharge differences for the same level of 150% in their study.

Additionally, the results for section IG3 show that the level is overestimated by the model, although the discharge simulation is adequate. Despite a test performed with the HEC-RAS model has demonstrated that the hypothesis of symmetric cross sections with trapezoidal shape is coherent (appendix A.7), this result indicates that possibly there is a limitation in the cross section definition (such as bottom width data).
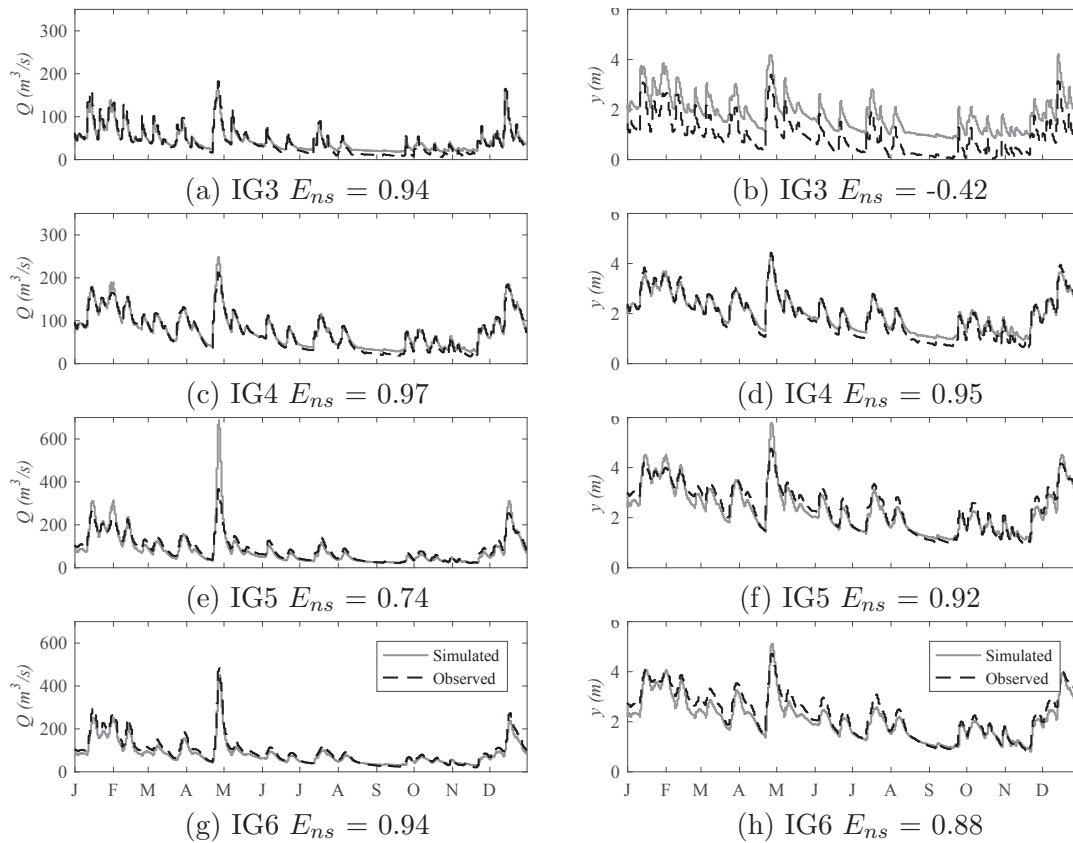
Figure 21: Observed and simulated discharges and levels for 2010; BC: levels time series
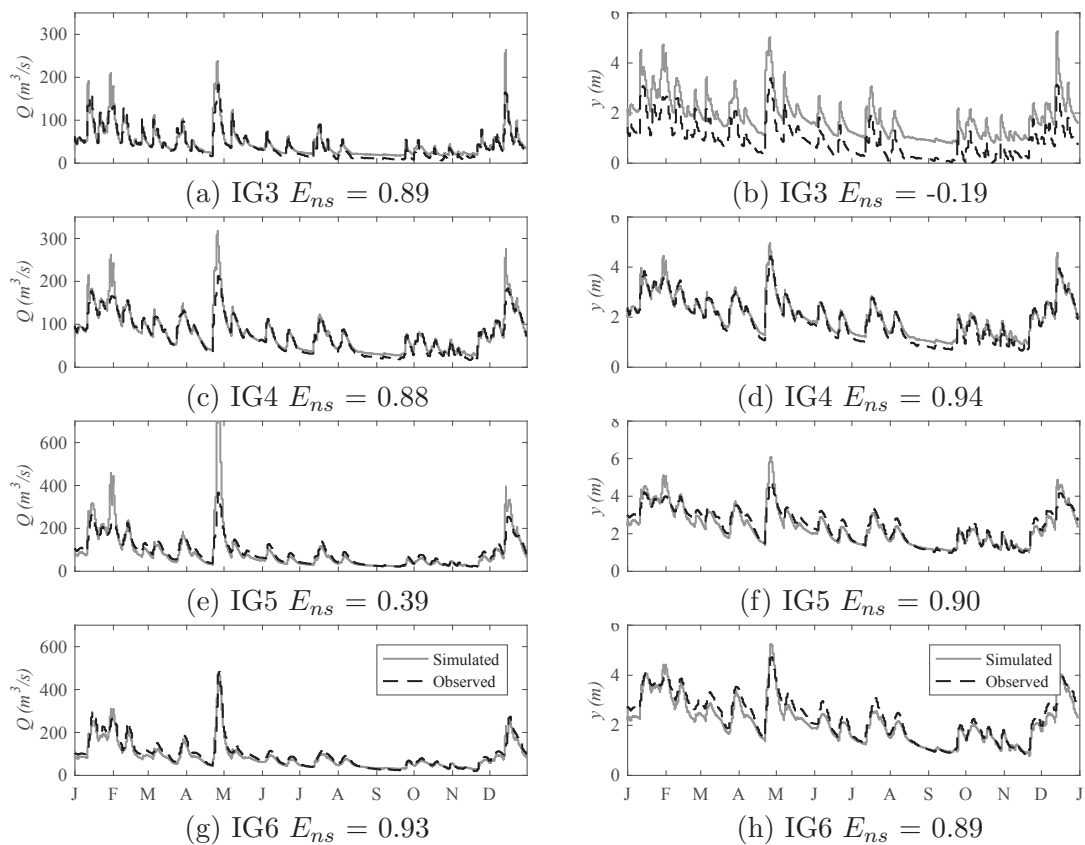
Figure 22: Observed and simulated discharges and levels for 2010; BC: rating curve
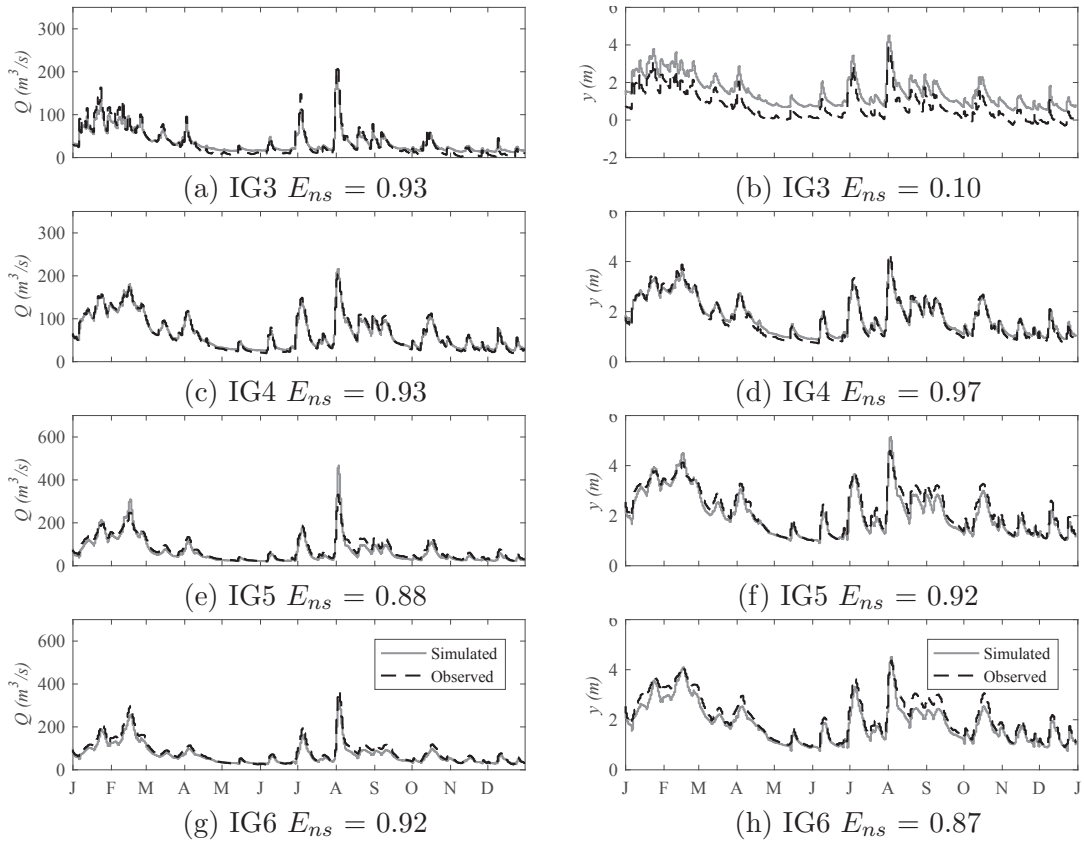
Figure 23: Observed and simulated discharges and levels for 2011; BC: levels time series
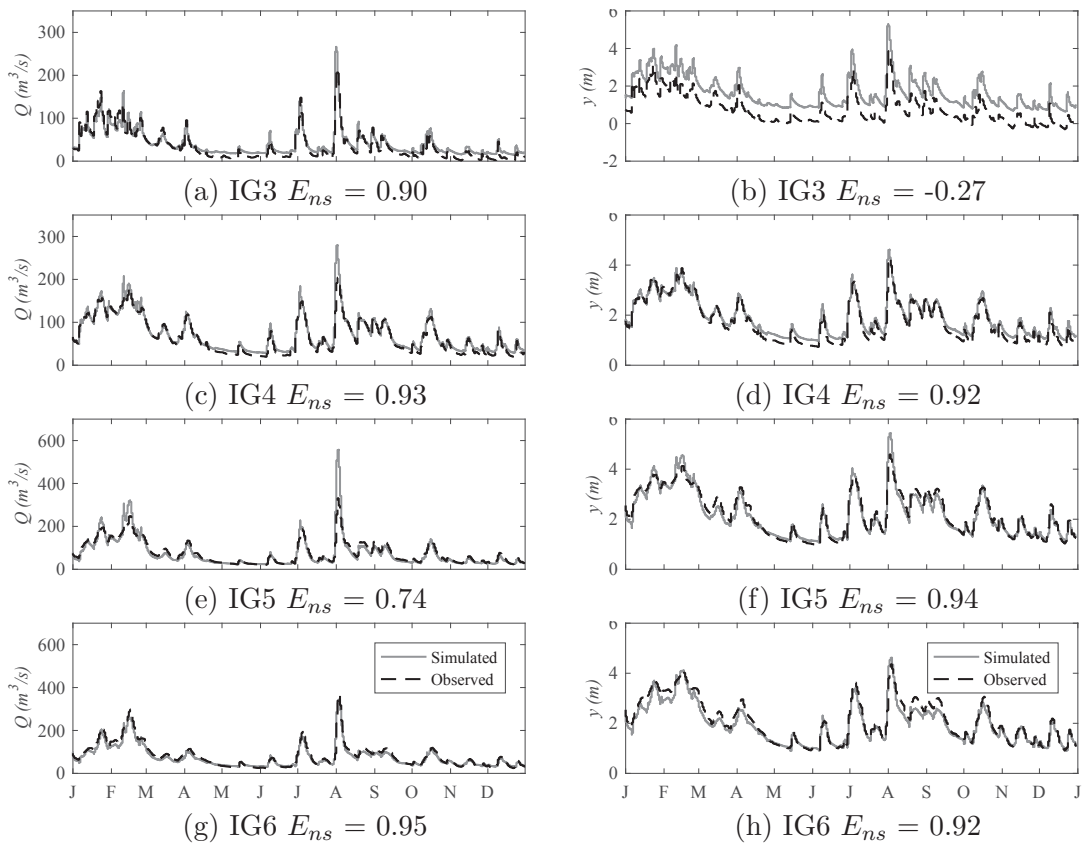
Figure 24: Observed and simulated discharges and levels for 2011; BC: rating curve
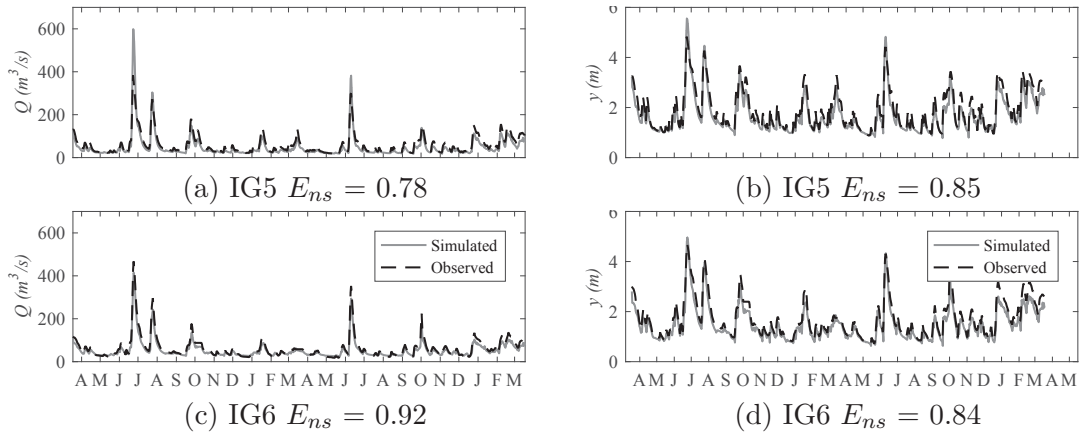
Figure 25: Observed and simulated discharges and levels for 03/18/2013 to 03/18/2015; BC: levels time series
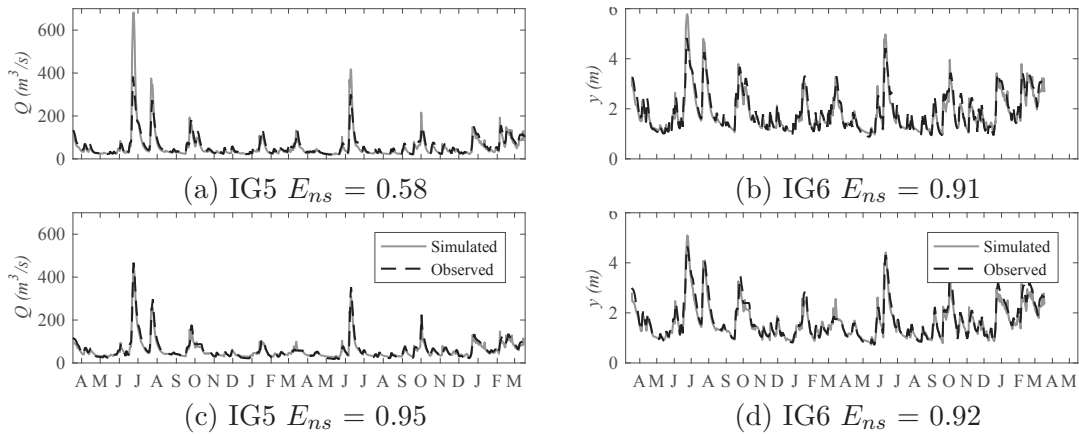
Figure 26: Observed and simulated discharges and levels for 03/18/2013 to 03/18/2015; BC: rating curve

5.2 Synthetic pollutographs

Figure 27 presents the results of daily for BOD concentrations in section IG2 estimated through tests T1 to T12 described in table 5. The procedure is also applied to other study cases in appendix A.8. The thousand series generated are presented as fancharts – a plot of time-varying distribution percentiles (P10, P20, P30, P40, P50, P60, P70, P80, P90) – shown as shaded bands around the median (Deoras, 2016).

Table 12 shows measures used as reference to compare the tests. Besides indexes introduced in table 4, quartiles (Q1, Q2, and Q3) and concentration of 10 (C_{10}) and 90% (C_{90}) of occurrence are compared; highlighted cells correspond to optimal results.

Overall, it is observed that, for the tests using seasonal parameters (T1, T5, T7, T11 and T12), the set of thousand series have a similar behavior to seasonal mean and standard deviations (exhibited in figure 16) of monitoring data applied for the tests in this study. Therefore, a closer representation to the real time series pattern is expected.

To predict extreme events, T6c shows better performance (difference of observed

and simulated 0.12 mg-O$_2$/L – row 13 in table 12). To predict C$_{90\%}$, most tests generated reasonable results, with the smallest differences between monitored and simulated in test T1b (0.02 mg-O$_2$/L).
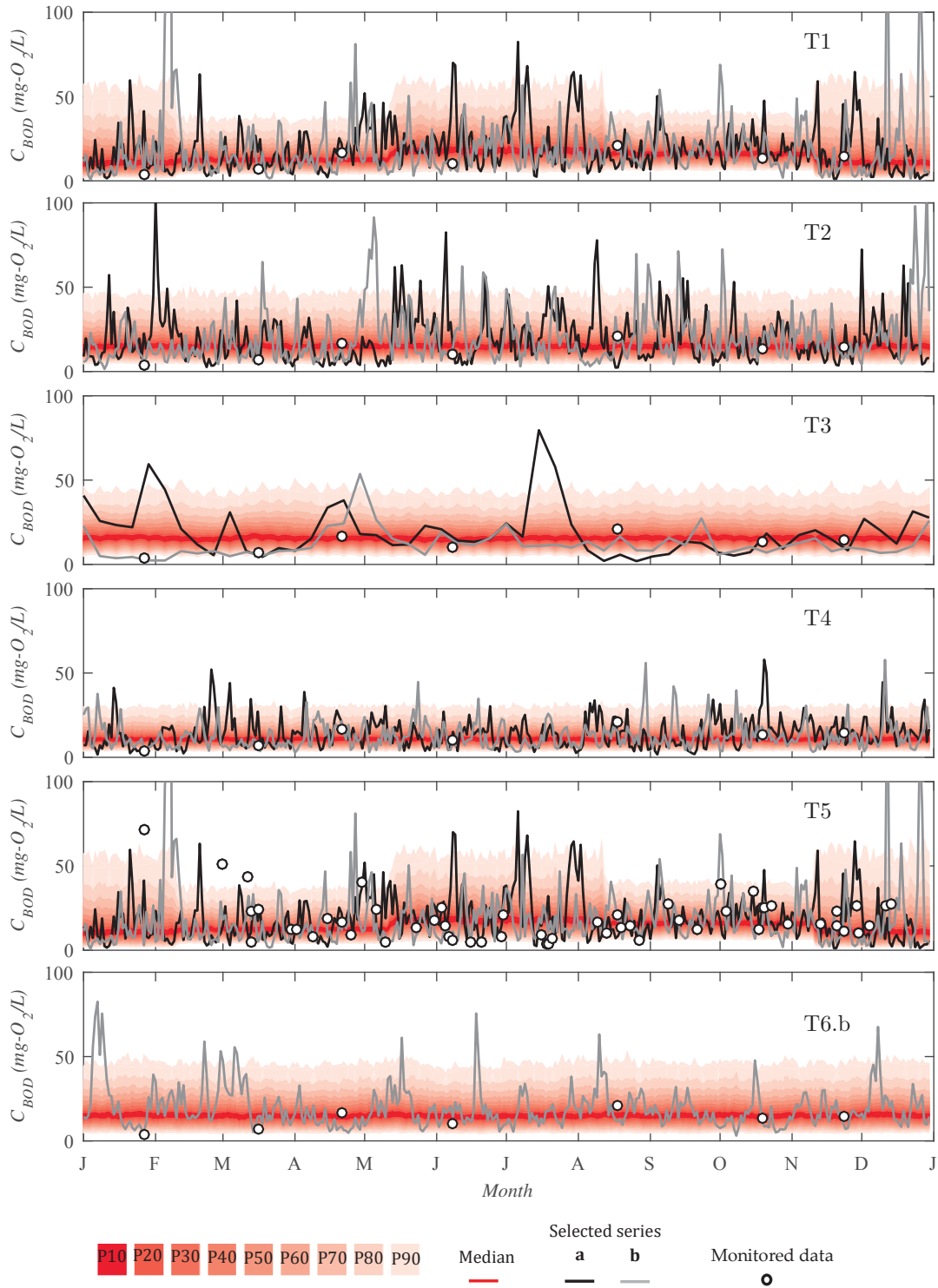
Figure 27: Synthetic series of BOD daily concentrations
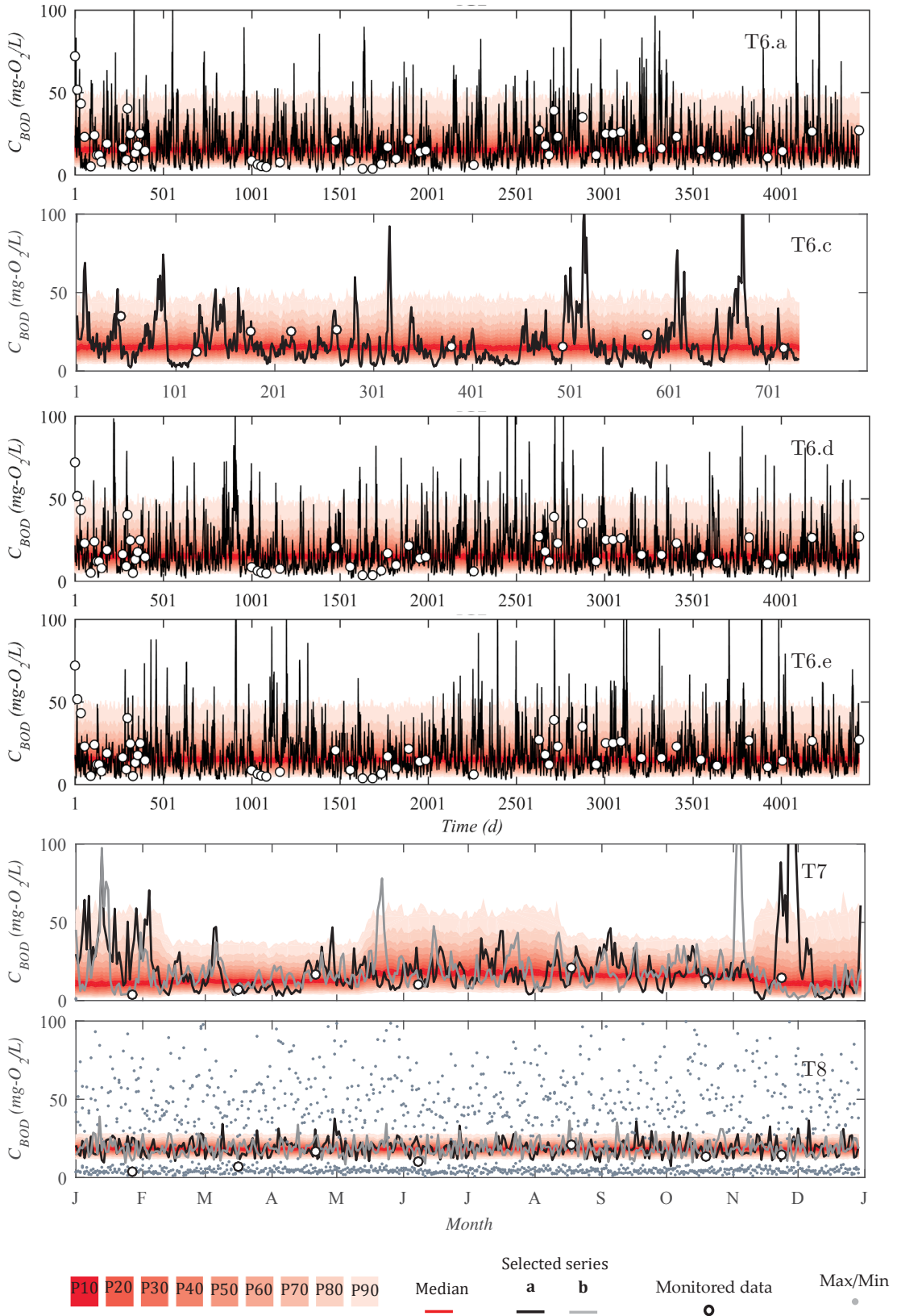
Figure 27: Synthetic series of BOD daily concentrations (continued)
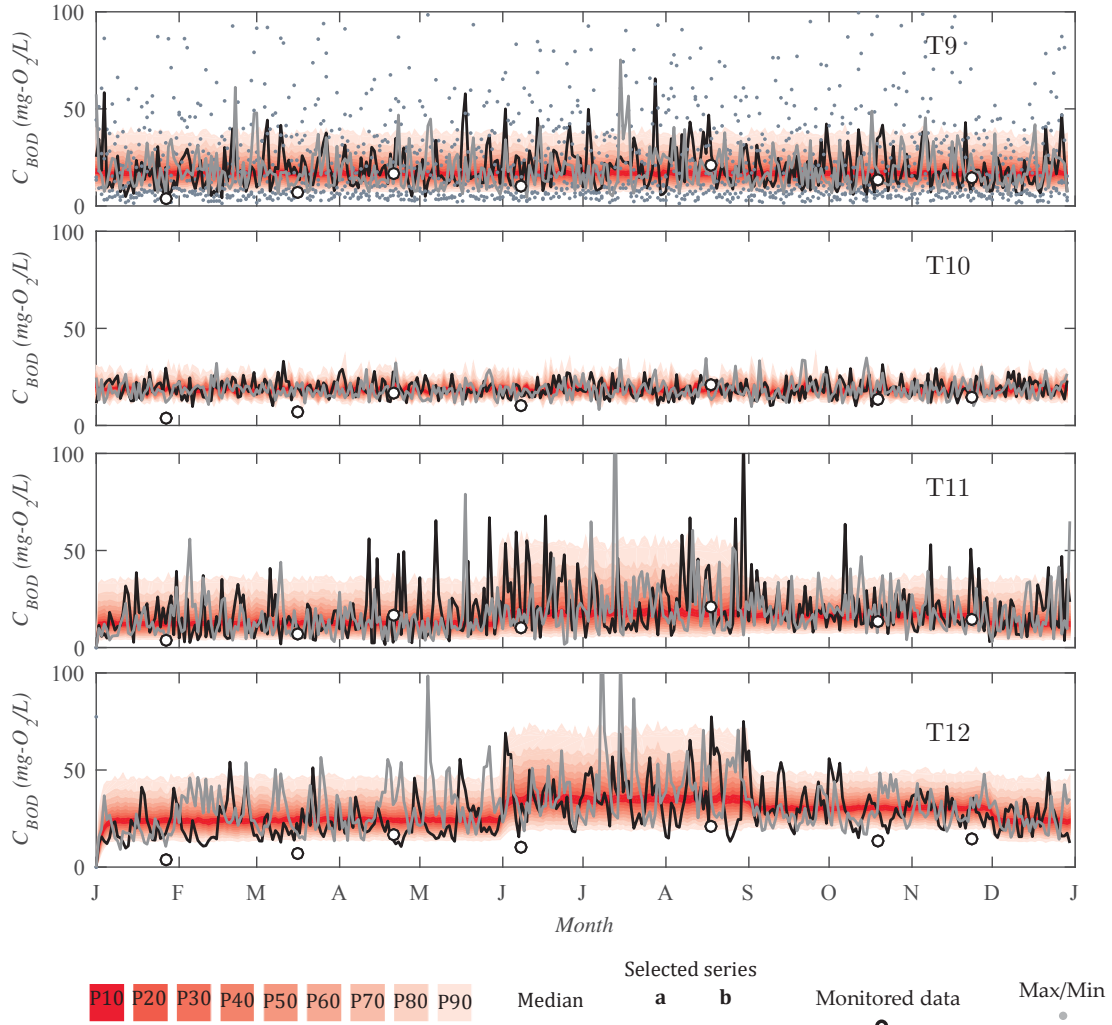
Figure 27: Synthetic series of BOD daily concentrations (continued)

The comparison in figure 27 for T1 and T2 shows that seasonal data generated results closer to the measured data (row 2 of table 12): T1 produced smaller $RMSE$ (0.38 mg/L) and larger R (0.91). Although T1b is closer to the data (smaller error measures), it has higher peaks than T1a, that deviate from the series, as detected in figure 27 (T1) during February and December.

In test T3, after generating weekly data for 2010, daily concentrations are estimated using linear interpolation, which explains the smoother variation over time observed in figure 27 (T3).

Experiment T4 uses only the data available in 2010 – $\mu = 12.39$ mg/L and $\sigma = 6.18$ mg/L (entire dataset has $\mu = 18.57$ mg/L and $\sigma = 12.87$ mg/L). In this test, measured concentrations of 10% and 90% of occurrence are not calculated, since the monitoring data is limited to seven values. Therefore, the comparison required to select a series in the option a is done considering $C_{10\%}$ and $C_{90\%}$ of entire monitored dataset. The series generated in this test do not significantly differ from other experiments, although a smaller variability can be observed in figure 27 (T4).

Experiment T5 shows that the simulated data is inside the range of observed data, while T6 explores the simulation of the entire period of monitoring data – twelve years. Overall, results indicate that quantiles are well estimated.

Test T7 is T1 with higher persistence ($\rho = 0.8$); minor differences are observed in simulated quartiles. Test T8 and T9 explore hourly data generation: series T8 has a smaller persistence ($\rho = 0.5$), so variation from one instant to another is more abrupt than T9 ($\rho = 0.9$). This aspect explains the low fluctuation of daily mean concentrations observed in T8.

Table 12: Goodness of fit between synthetic series and BOD measured concentration in section IG2

*	Test	Series	Measured - Simulated (mg/L)							Error Measures			
			$C_{10\%}$	$C_{90\%}$	Q1	Q2	Q3	σ	μ	RMSE	R	MAPE (%)	PBIAS (%)
1	T1	a	-0.62	-0.05	0.37	0.11	0.28	-0.85	-0.38	4.20	-0.43	19.29	-163.61
2		b	0.75	-0.02	1.19	2.67	1.95	-13.26	-0.86	0.38	0.91	2.68	10.55
3	T2	a	-0.94	-0.12	0.49	0.22	-0.47	-1.57	-0.74	1.49	-0.59	8.02	5.41
4		b	-3.59	-1.26	-0.25	0.76	0.37	-3.16	-1.36	0.56	0.72	3.68	7.18
5	T3	a	0.30	-0.39	-0.84	-0.25	1.25	-0.93	-0.64	2.76	-0.39	11.11	-59.44
6		b	15.04	-0.19	1.76	5.67	10.72	5.23	6.51	0.69	0.69	5.06	7.86
7	T4	a	9.95	0.37	1.84	3.44	7.02	3.93	4.39	1.71	0.05	10.27	-52.64
8		b	14.66	0.08	2.08	5.24	8.48	4.90	5.89	7.19	10.76	16.38	7.97
9	T5	a	-0.62	-0.05	0.37	0.11	0.28	-0.85	-0.38	6.85	0.12	17.50	-5.42
10		b	0.75	-0.02	1.19	2.67	1.95	-13.26	-0.86	6.61	0.07	130.98	15.68
11	T6	a	0.73	-0.69	-0.09	0.85	0.51	-0.84	-0.21	1.78	0.08	0.86	3.33
12		b	0.55	-0.86	-0.22	0.91	1.44	-1.48	-0.32	0.63	0.07	12.80	-65.13
13		c	-0.12	-0.71	-0.05	1.35	1.30	-3.67	-0.61	0.64	-0.05	4.43	19.88
14		d	1.69	-0.88	0.25	1.73	1.91	-1.00	0.28	1.43	0.39	2.45	8.37
15		e	1.93	-0.98	-0.03	1.35	1.70	-0.85	0.22	1.56	0.19	3.84	14.47
16	T7	a	-0.71	-0.35	0.20	-0.32	0.06	-8.99	-2.21	4.12	0.09	15.81	-38.29
17		b	4.55	-1.00	-1.10	0.29	2.56	-1.91	-0.08	0.59	0.76	3.49	-8.73
18	T8	a	10.42	-7.51	-6.19	-2.52	3.35	7.85	-0.36	1.64	-0.17	8.21	-79.47
19		b	10.77	-7.66	-6.06	-2.04	3.26	7.86	-0.17	0.90	0.25	5.12	-22.94
20	T9	a	3.43	-2.68	-2.87	-1.19	0.20	2.78	-0.79	1.58	0.16	7.57	-71.30
21		b	5.55	-3.55	-2.24	-0.07	2.45	2.98	0.05	0.86	0.54	5.42	-6.78
22	T10	a	11.28	23.51	20.76	18.17	14.59	8.45	-0.34	1.68	-0.21	5.73	-67.07
23		b	12.63	23.37	21.29	18.41	15.81	8.38	0.04	1.28	0.13	7.50	-49.69
24	T11	a	-1.09	-1.05	-0.53	0.59	0.87	-1.00	-0.55	2.11	0.27	10.25	-38.02
25		b	3.53	-0.87	-0.29	1.33	2.91	0.10	0.94	0.68	0.60	4.45	13.81
26	T12	a	-8.19	-9.97	-9.74	-9.96	-9.07	0.74	-9.49	3.55	0.65	22.78	-158.36
27		b	-11.67	-12.60	-13.40	-11.96	-12.13	-1.02	-12.82	1.97	0.91	13.79	-110.01
28	I1		1.52	-2.74	-3.58	-2.70	-1.47	-9.22	-4.18	0.05	0.28	13.03	-82.72
29	I2		10.91	-4.68	-8.24	-5.28	-0.11	7.04	-1.29	0.03	0.84	7.63	-72.12
30	I3		17.75	1.16	-1.73	2.43	9.05	8.06	5.84	0.01	0.94	1.81	-0.03

*used for reference in the text

Test T10 is similar to T8, but the data is generated every 50 s – time step for numerical solution in the deterministic model. Because of extended computational time for processing required in this test, only twenty options are generated. Although these tests consider a high persistence ($\rho = 0.99$), as expected in a series with a small time step, some variations are still abrupt from one time step to the next (figure 28). Therefore, the

daily means of series a and b have small variation, detected in figure 27 (T10).
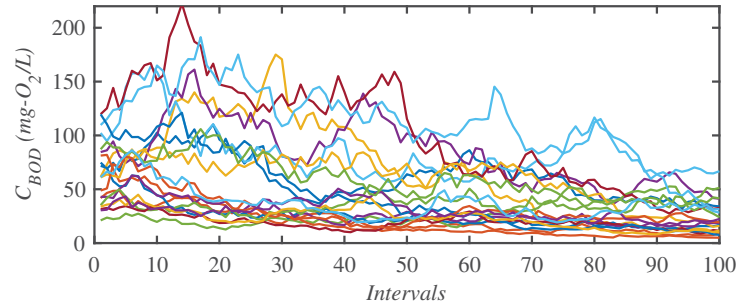
Figure 28: First 100 values of the 20 series options generated in test T10

Tests T11 and T12 differs from the previous tests; the first one is a random and independent series, generated with a two parameters log-normal distribution (with seasonal mean and standard deviation of the observed dataset); it shows higher variability than test T12, a hybrid model with smoother variation over time. Appendix A.9 presents water quality simulations with these tests as boundary condition, showing that simulated concentrations follow the expected behavior.

5.3 Water quality simulations

This section presents the tests in which concentrations of water quality parameters are simulated through the deterministic model, using different series as boundary condition, and a calibration strategy based on the unsteady behavior. Additionally, the methods for synthetic series are applied for the following stations IG3 to IG6, in order to compare the series with those generated by the deterministic approach. Steady state results are compared as duration curves.

Lateral pollution entering the main river is estimated through the procedure described in section 4.1. Tests to verify input load temporal variation and sensibility are discussed in section 5.5 and appendix A.10, respectively.

Another required input, the dispersion coefficient is considered as a constant value during the simulations (20 m^2/s), since it affects numerical solution stability, as verified in Appendix A.1. Advection and cross sections areas are estimated through the hydrodynamic module, assuming levels time series as upstream boundary condition.

5.3.1 Calibration strategy

Table 13 presents the ranges suggested in the literature, and the values defined in this research for all transformation rates involved in BOD, N-org, DOC and POC balances. Variations due to temperature are not considered, as well as the effect of sediment resuspension.

Table 13: List of transformation rates for simulations of BOD, N-org, DOC, POC and DO in SIHQUAL

Parameter	Literature range	Assumed range
K_d (d^{-1})	0.02 – 3.4[1]	0.2 - 1.5
K_s (d^{-1})	0.05 – 0.35[2]	0.2 -0.4
K_{so} (d^{-1})	0.001 – 0.10[2]	0.01 - 0.1
K_{oa} (d^{-1})	0.20 – 0.25[2]	0.18 - 0.25
K_h (d^{-1})	–	0.05 - 0.2[4]
K_p (d^{-1})	–	0.05 - 0.15[4]
$\alpha 5$ (d^{-1})	3.0 – 4.0[1]	3.5
$\alpha 6$ (d^{-1})	1.0 – 1.14[1]	1.14
$\beta 1$ (d^{-1})	0.10 – 1.0[1]	0.12
$\beta 2$ (d^{-1})	0.20 – 2.0[1]	1.0
K_a (d^{-1})	0.00 – 100.0[1]	0.35 - 0.6
K_4 (gO$_2$/m^{2}d)	0.05 – 10.0[3]	1.0 - 1.5

[1]Brown and Barnwell (1987) [2]Von Sperling (2007) [3]Thomann and Mueller (1987) [4] Based on values calibrated by Knapik et al. (2016)

The transformation rates for each water quality constituent are calibration parameters. Attempts to evaluate temporal variation of these kinetic process are conducted using BOD simulations, with series T1a as upstream boundary condition. The basic procedure follows five steps – figure 29: (i) based on traditional values defined by the state-of-art literature, three intervals of variation are defined for each transformation rate; (ii) multiple time series of rates are generated as random series with uniform distribution and variation inside each of the three intervals; (iii) one series in then selected for each control point; (iv) interpolations are applied to define time series rates at the required time step; (v) at last, sets of kinetic rates varying over time are calculated for each interval between the control points, as averages of the selected series.
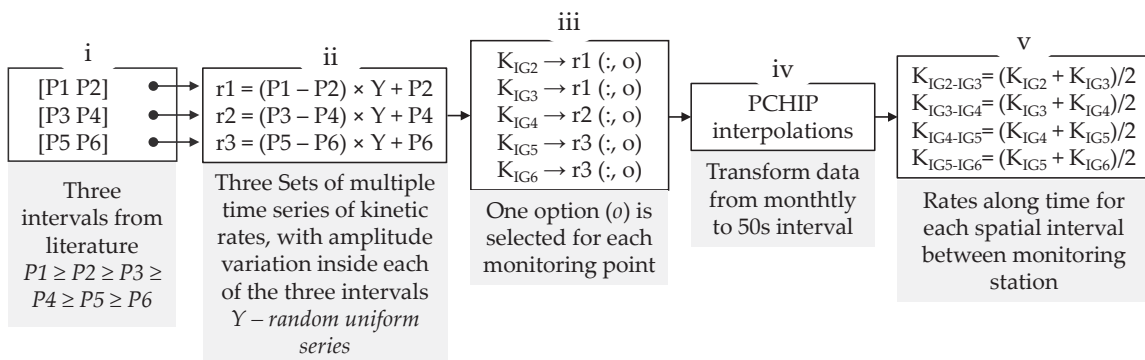
Figure 29: Scheme of the algorithm developed to generate time series of transformation rates

Experiments v1 to v9 considered different criterion mainly on steps (ii) and (iii),

that explore how the random series are generated and how one of them is selected:

- v1: a hundred monthly time series are generated – the option selection among them is arbitrary; transformation rates in IG2 and IG3 are inside of the interval with higher values ($r1$), because this sections receive larger pollution loads; IG4 has values in the second interval ($r2$), while IG5 and IG6 consider the third set (with lower values; $r3$);
- v2: same conditions as v1, but with larger interval ranges;
- v3: same conditions as v1, but, in step (v), the time series are converted into means along time;
- v4: same conditions as v1, but the series are weekly instead of monthly;
- v5: same conditions as v1, but the series are daily instead of monthly;
- v6: same conditions as v1, but different series option (also arbitrary);
- v7: same interval range as v1 in step (i); for step (ii): quartiles (Q1, Q2 and Q3) of the monitored dataset are used to classify the daily pollutograph in IG2 (T1a synthetic series) in four intervals (C_{IG2}[4] $\leqslant$ Q1, Q1 $< C_{IG2} \leqslant$ Q2, Q2 $< C_{IG2} \leqslant$ Q3, $C_{IG2} \geqslant$ Q3); to the days when concentration is higher (Q2 $< C_{IG2} \leqslant$ Q3 and $C_{IG2} \geqslant$ Q3), the time series for kinetic rates is one of the options of $r1$ (arbitrary among thousand datasets); for intermediary concentrations (Q1 $< C_{IG2} \leqslant$ Q2), $r2$; and for smaller concentrations ($C_{IG2} \leqslant$ Q1), $r3$ – with this procedure, also represented in figure 30, a daily time series for rates is generated for section IG2; the rates for the upstream sections are defined as parcels of IG2 – K_{IG3} is 90%, K_{IG4} 60%, K_{IG5} 50% and K_{IG6} 30% of K_{IG2};
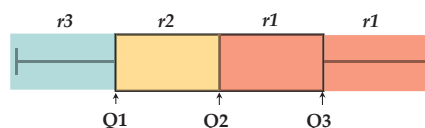
Figure 30: Representation of intervals used in procedure v7; *r1*, *r2* and *r3* are time series of transformation rates attributed to each concentration interval

- v8: same conditions as v7, but the synthetic series is T9a instead of T1a;
- v9: similar conditions to v7, but the time series of IG3 to IG6 are defined similarly to IG2, using the synthetic series instead of being a parcel of IG2.

The simulated BOD series with each calibration strategy are summarized in figures 31 and 32, comparing characteristic measures (C10, C90, Q1, Q2 and Q3 monitored × simulated) and coefficients of variation (CV) – to evaluate the overall variability produced by each calibration scenario. The tests are also supported by the resultant pollutographs, presented in figure A.20.

---

[4]Concentration in IG2
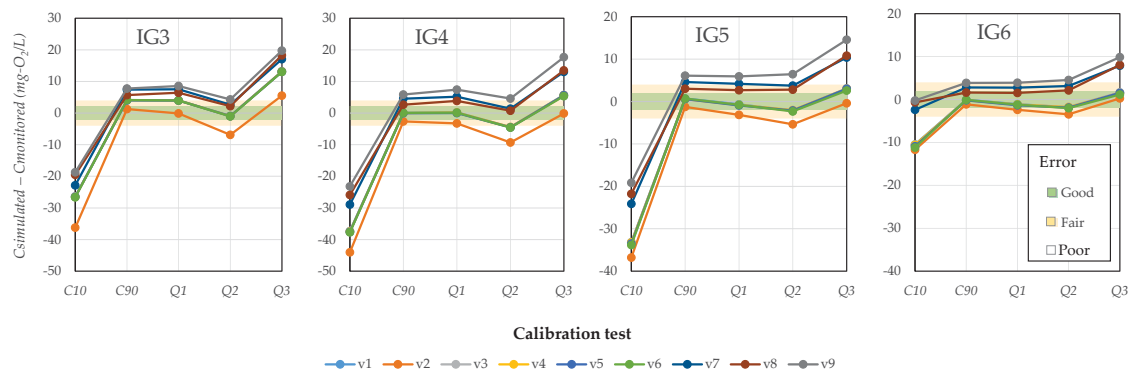
Figure 31: Differences between monitored and simulated measures (C10, C90, Q1, Q2 and Q3); green band indicates difference up to 2 mg-O$_2$/L, while yellow limits 4 mg-O$_2$/L and white space define the tests with larger errors
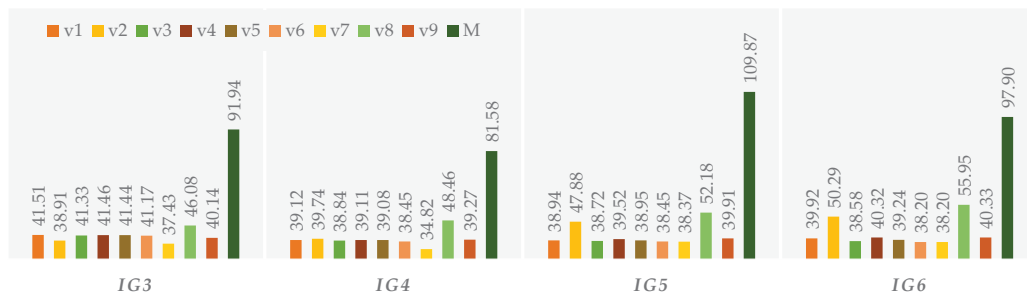
Figure 32: Coefficients of variation (%) for time series simulated with each calibration strategy

Comparison between tests v1 and v2 shows that, although there is an increase in overall temporal variability in IG4 to IG6 (figure 32), the pollutographs suffer more attenuation with v2 (this aspect can be verified in figure A.20 in appendix A.11); in addition, although higher transformation rates resulted in pollutographs more close to the actual data – especially in IG3, the estimation of C10, Q2 and C90 (except for IG3) is more distinct than the previous approach, v1 (verified in figure 31). Therefore, in this research the range of variation for K_d and K_s is based on the values defined in table 13.

Part of the tests do not differ significantly in comparison with v1, such as v3 (constant rates over time – the most common approach in modeling studies), v4 (weekly rates), v5 (daily rates) and v6 (other series from the hundred options).

Procedure v7, on the other hand, generated results with higher temporal variability (figure A.20, v7). Following this test, v8 and v9 investigate the effect of using the synthetic series to calibrate the water quality model. Since T9a is a hourly time series, with higher temporal variability (as previously shown in figure 27 T9), it generates rates with corresponding oscillation over time; consequently, this calibration test produce higher concentration peaks, in comparison with all previous experiments (figure A.20 v8).

Test v9 uses the estimated synthetic series in all sections; results show overestimated BOD distributions, with poor estimates for C_{90} and quantiles; simulated C_{10} , on the

other hand, is closer to C_{10} of the monitored dataset (figure 31). Overall, most strategies overestimated C_{90} and underestimated C_{10}.

Considering all tests, v8 is the selected approach to define calibration for the unsteady water quality model. This test produced reasonable estimates for concentrations quartiles in the studied period (figure 31), and temporal variability produced coefficient of variation closer to the monitored dataset (figure 32). Final results for kinetic rates in this research have a similar shape as K_d, showed as example in figure 33.

It should be stated that the selected approach presumes that less impacted reaches have smaller transformation activities (minor rates to downstream of the Iguaçu river), so concentrations and rate's values are directly proportional.
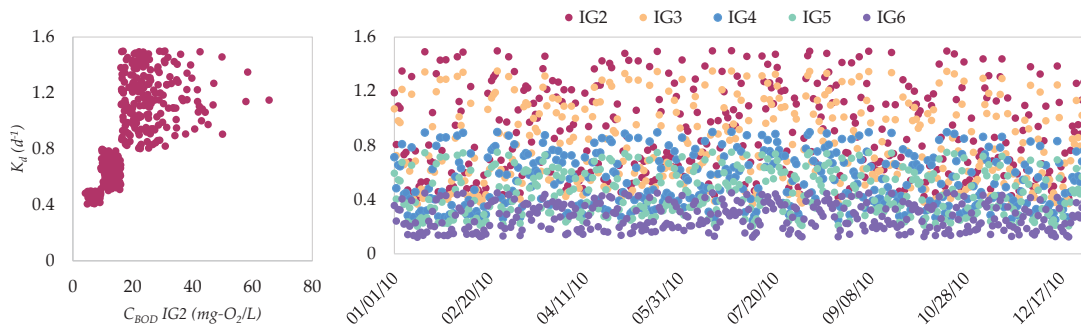
Figure 33: Time series of deoxygenation rates (K_d) in 2010, set as input for BOD simulations under unsteady state

In comparison with test v3, in which the temporal variation is disregard, this methodology adds ten seconds to water quality simulations, although this interval increases with quantity of kinetic rates. Further discussions about temporal variation relevance of transformation rates are presented in section 5.5, while sensibility to results are explored in appendix A.10.

5.3.2 Steady state

The steady water quality model is originated by canceling the term $\partial C/\partial t$ in equation 3.8. The governing equation is solved using centered finite differences, with $\Delta x = 25$ m (Ferreira, 2015).

Averages over time of flow velocity and cross section area are calculated from the hydrodynamic module results, to be used as inputs in the steady module. The hypothesis is that this strategy does not generate results significantly different from usual methods, such flow analysis through Manning's equation. Since these parameters are available every 500 m, linear interpolations are performed to comply with numerical requirements.

Downstream boundary condition is the observed dataset median in the last section of the control volume. As upstream boundary condition, a null concentration flux is adopted

($dC/dx = 0$). The transformation rates used are presented in table 14 – calibrated though trial and error, while other required conditions are the same as those applied for the unsteady model.

Table 14: Kinetic rates calibrated for the steady state module

Rate	IG2	IG3	IG4	IG5	IG6
K_d (d^{-1})	1.30	0.80	0.20	0.20	0.20
K_s (d^{-1})	0.25	0.20	0.20	0.15	0.15
K_h (d^{-1})	2.50	2.00	1.00	0.50	0.10
K_p (d^{-1})	2.50	2.00	0.20	0.10	0.05
K_s (d^{-1})	3.00	2.00	0.20	0.10	0.05
K_a (d^{-1})	4.32	1.66	1.23	0.64	0.84
K_4 (gO$_2$/m^{2}d)	0.80	0.80	1.00	0.70	0.70

other rates constant over space: $\alpha_5 = 3.50$, $\alpha_6 = 1.00$, $\beta_1 = 0.12$ d^{-1}, $\beta_2 = 1.00$ d^{-1}; $K_{so} = 0.10$ (d^{-1}), $K_{oa} = 0.40$ (d^{-1})

Steady state simulations in the Iguaçu river are presented in figure 34, besides boxplots of the data collected between 2005 and 2017. Results indicate that the model follows concentration variation over the main channel, with values close the median of observed data – used as calibration reference. Oscillations along the river, however, are not well represented, since the model considers external inputs as loads diffusely and evenly distributed, and advection effects are invariable in time.

Simulations with data flow from 03/18/2013 to 03/18/2015 are also presented for BOD and DOC (calibration parameters are the same as those calibrated for 2010, and input load are those presented in item 5.3 using data population of 2015). Because POC data is available only from 2012 to 2014, this parameter is simulated only for the period 03/18/2013 to 03/18/2015. Monitored concentration of TOC are also limited to this interval; in figure 34 (h), TOC along space is generated as the sum of POC and DOC simulations (figures 34 (f) and (g), respectively).

The dataset generated for 2013-2015 suggests that locations of point sources need to be better adjusted, or that lateral loads are indeed overestimated for the reach IG2-IG3. Nevertheless, results follow the expected behavior, since concentrations decrease in less impacted regions (BOD, N-org and DOC). For DO, minimum levels occur between IG3 and IG4, due to the presence of organic matter; to downstream of this reach, reaeration increases DO concentrations.

Figure 34 (h) allows to compare the fractions particulate and dissolved in TOC. According to Leenheer and Croué (2003), POC usually is a minor fraction (below 10%) of TOC, proportion that increases with river's size and flow rate; this is verified in the

steady simulation – absolute difference between DOC and POC decreases at points where discharge increases.
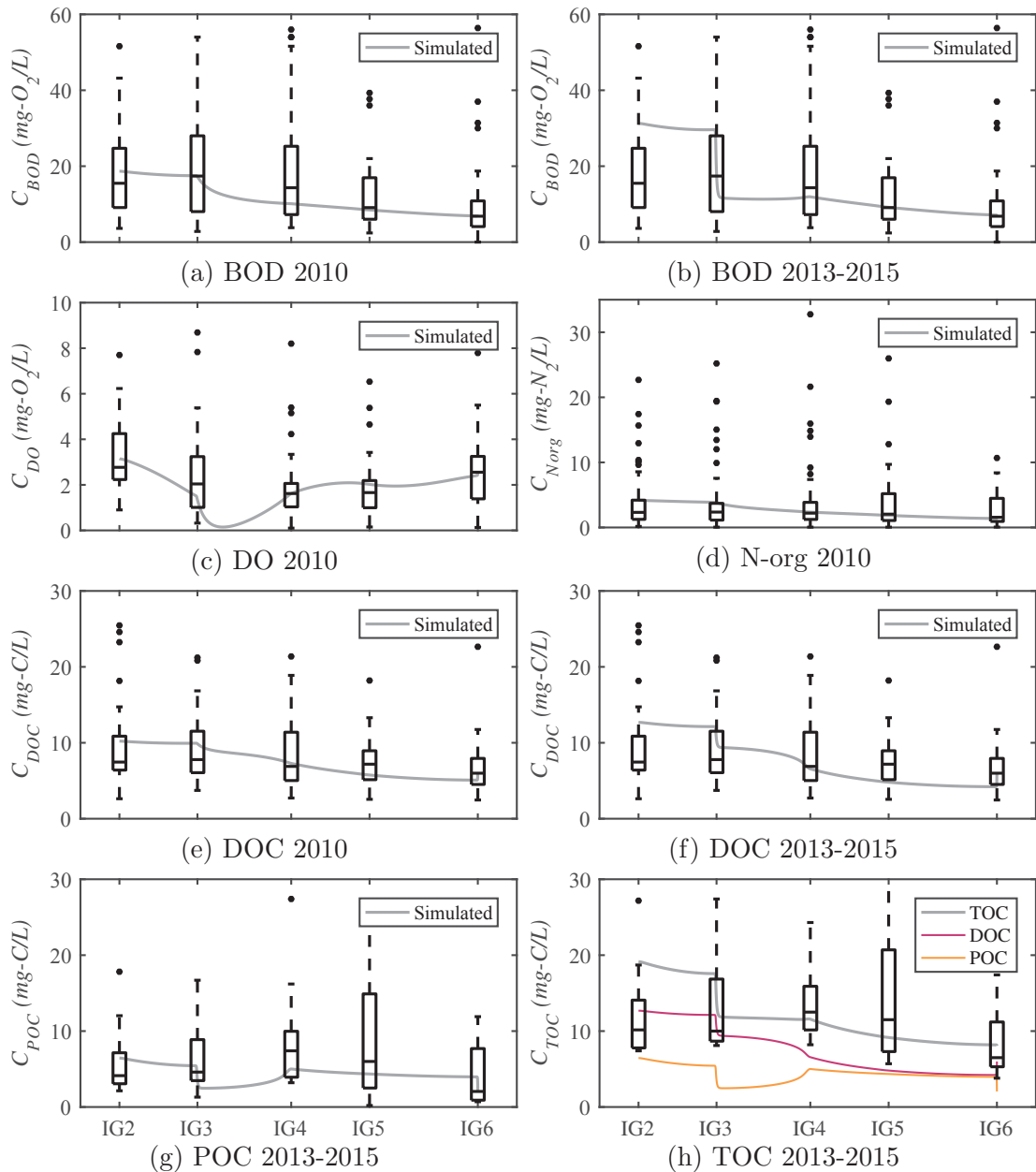
Figure 34: Steady simulations and boxplots of monitoring data

Higher TOC concentrations are reproduced in the areas with larger anthropic influence (IG2 to IG3). The indirectly proportional behavior of DOC/POC in the reach IG3-IG4 exemplifies the importance of knowing the dynamic of kinetic process – since the calibrated K_p in this reach substantially decreases (2.0 to 0.2 d^{-1}), which means that the dissolution process is inhibited, probably POC is accumulating (as dissolution decreases, so it does DOC concentration); at the same time, the sedimentation might be underestimated, considering that in the model it does not change from one section to another (calibrated K_s in IG3 and IG4 = 0.2 d^{-1}).

However, it should be considered that such interpretation is based on a limited dataset, and that the dynamic of total organic carbon and its fractions is not entirely known in riverine environments, so other process might be missing. Literature has shown that the occurrence of particulate and dissolved organic carbon may be significantly variable and depend on system characteristics (e. g. Metcalf and Eddy, 1991; Islam et al., 2019).

5.3.3 Unsteady state

The series selected among the autoregressive tests to be used as upstream boundary condition are T1a, T4b and T9b. The first represents daily concentrations based on seasonal variations; the second is a critical scenario for modeling studies, with a limited set of available data as input; the last series, on the other hand, illustrate a case with high frequency data (hourly time series). In addition, series estimated through PCHIP interpolation, Fourier series and fitting with smoothing spline (tests I1, I2 and I3) are investigated.

The series used as boundary condition are generated as daily samples; therefore, PCHIP interpolations in time are applied, in order to comply with the required time step (50 s). Simulations with direct sub-daily data as input (series T10a and T9b) are also explored (appendix A.12).

Since data is usually sampled at infrequently intervals, and during limited events, it is not expected that models simulate daily averages equal to observed values on particular moments of the day, as mentioned by Kim et al. (2007). Therefore, following a comparison of daily simulated concentrations, boxplots, water quality duration curves and loads estimations assess the predictive performance of the different approaches used in this study. Focus is on the usability of synthetic pollutographs as boundary conditions for the water quality model under unsteady state, and the effects when choosing different approaches for water resources planning and management purposes. The main analysis are based on BOD simulations; other water quality parameters give support to the proposed methodology and reinforce the discussions.

5.3.3.1 Biochemical Oxygen Demand (BOD)

Synthetic series for IG2 originated with I1 and I2, presented in figures 35 and 36, illustrate the definition of concentrations based on discharges. As verified in figure 17, that shows hydrographs in the interval of interest, september and october are periods of low flow; since I1 directly relates flow to concentration, in this period the model predicted high peaks of mass distribution. Test I2, on the other hand, tends to smooth concentration fluctuations in low flow periods. These behaviors are reproduced by the deterministic model in downstream sections.
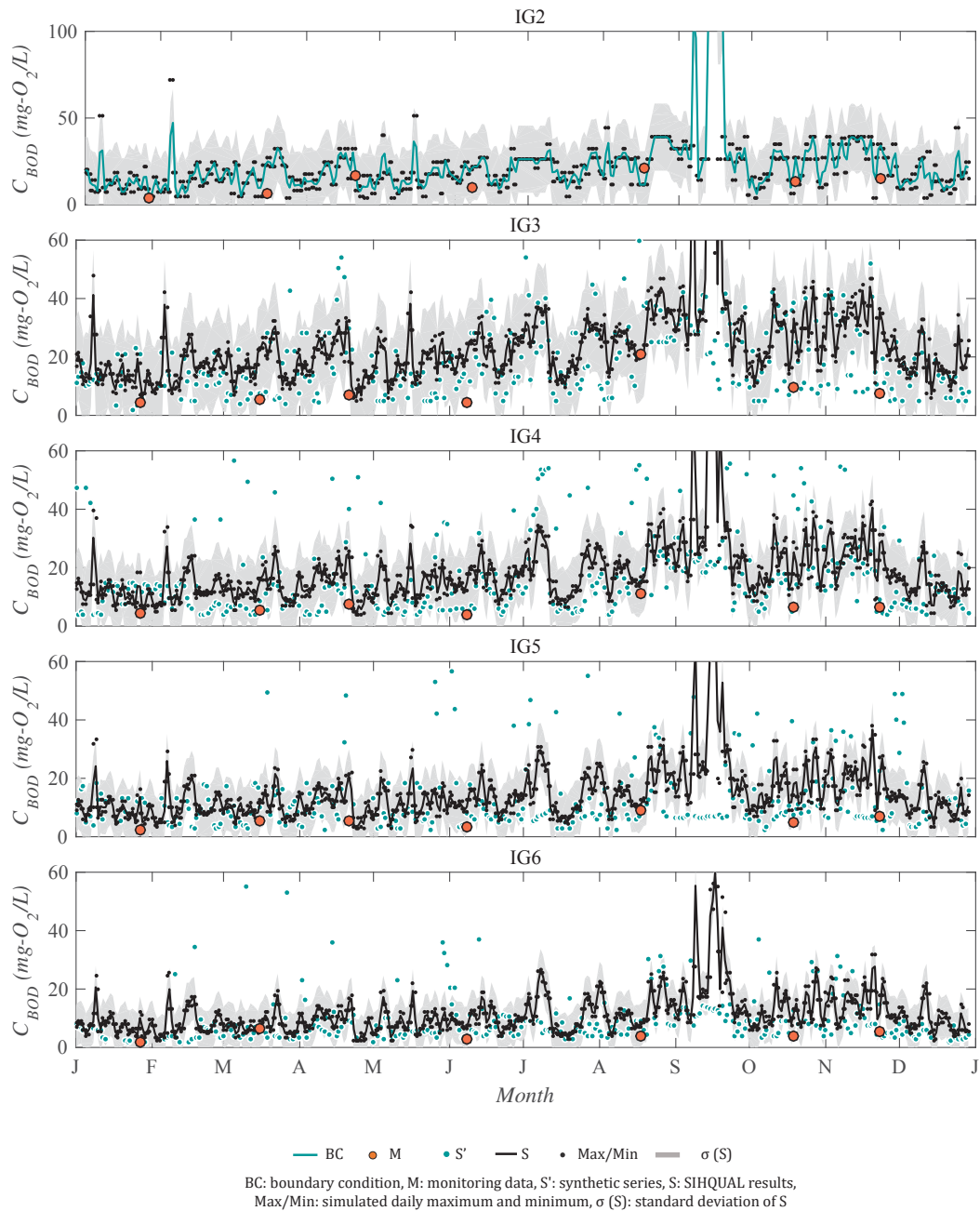
Figure 35: Daily BOD simulated concentrations and monitoring data for the year 2010; BC generated with PCHIP interpolations (test I1)

For IG3 to IG6, synthetic series with I1 and I2 generated unreasonable daily mean concentrations; boxplots, however, illustrated in figure 38, show interquartile values similar to the dataset (outliers are not presented).

This indicates that predictions through PCHIP interpolation and Fourier series are not able to represent extreme values in this case, **and extrapolation is poor**. Such behavior is fair, since it is well established that the dynamic between discharges and concentrations of pollutants in a stream can be highly variable (Zhou et al., 2011; Ramos et al., 2015).

Test I3 is based on spline interpolations using time, showing a smooth and low variation, with similar boxplots of data generated with synthetic series and deterministic modeling (figure 38); the oscillations in daily data observed in figure 37 are due to lateral contributions and calibration parameters, that introduce a certain oscillation along the year (theses effects are further verified in section 5.5). This gradual variation over time is unlikely able to represent the unsteady behavior.
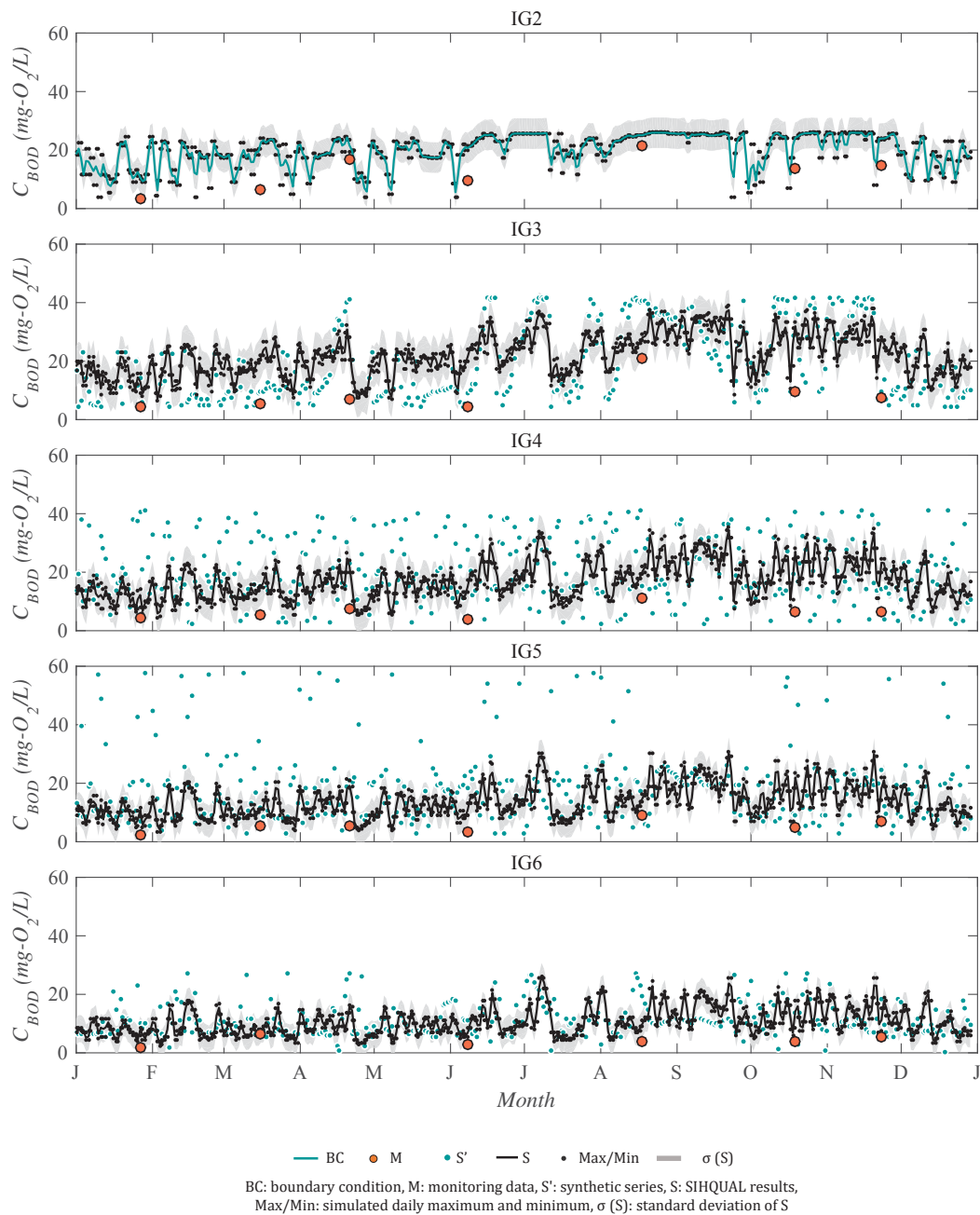
Figure 36: Daily BOD simulated concentrations and monitoring data for the year 2010; BC generated with Fourier series (test I2)
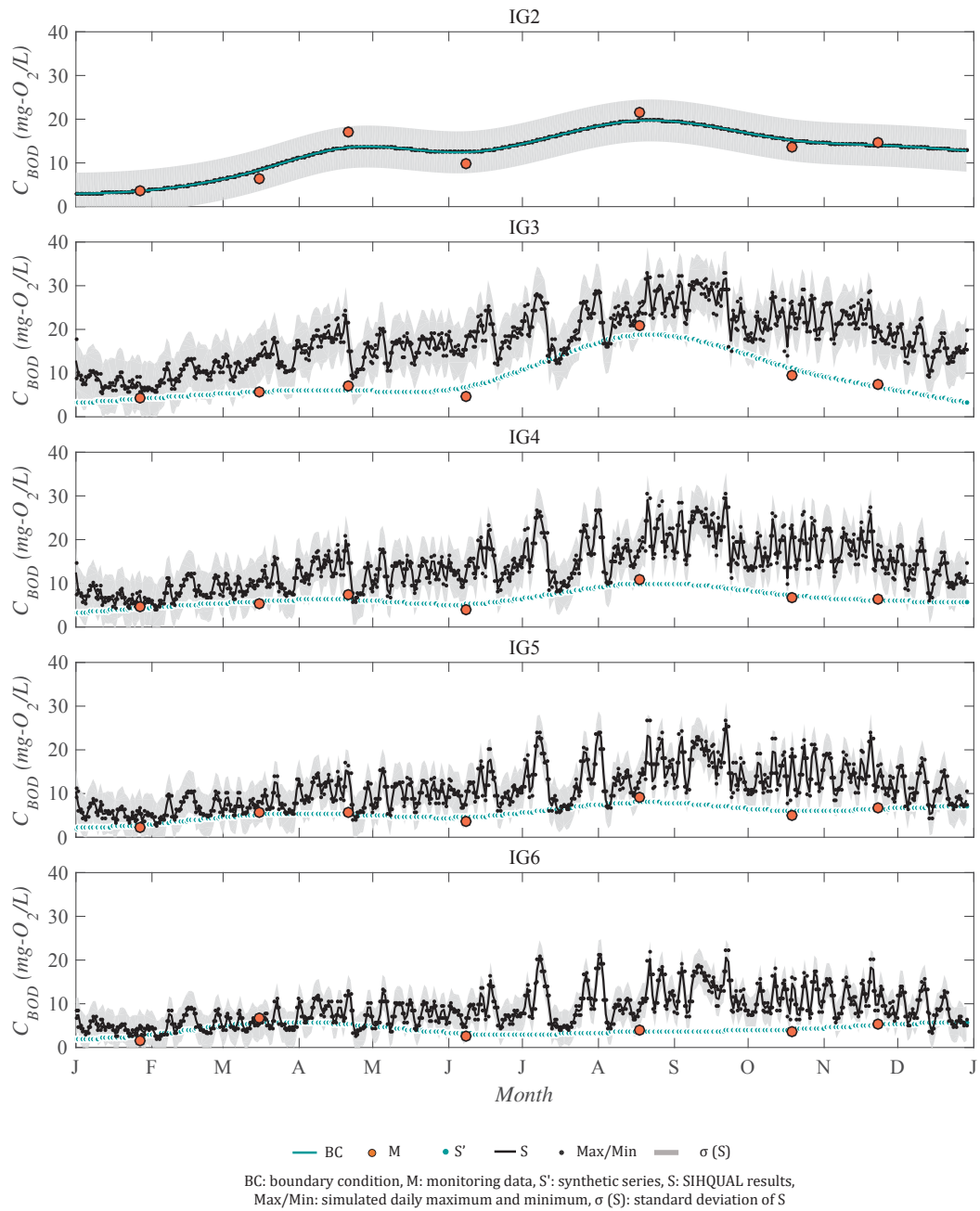
Figure 37: Daily BOD simulated concentrations and monitoring data for the year 2010; BC generated with Spline interpolations (test I3)

Figure 38: Boxplot of measured (M) and simulated BOD concentrations for 2010 with SIHQUAL (S) and synthetic series (S'); I1, I2, I3, respectively in each row

Figures 39 to 42 present the simulations with T1a, T4b and T9b as input condition. Test T1a generated larger concentration peaks during winter, because this period has the highest means and standard deviations of the monitored dataset – supported by the data presented in figure 16; concentrations modeled with SIHQUAL follows this expected seasonal fluctuation (figure 39). Appendix A.13 presents the same simulation conditions also for the period 2013-2015.

Test T4 shows that, even with a small input dataset available, the unsteady model is able to reproduce overall system variability, since boxplots S and M are similar (figure 42 T4b); this is possible due to representation of variability in lateral inputs and kinetic

rates (this statement is also supported by the tests in section 5.5).

Test T9b do not differ significantly from T1a in terms of overall variability, since both time series are generated from seasonal metrics; pollutographs result of T9b show a consistent behavior over time (figure 41), without large peaks as those generated with T1a – this possibly occurs because persistence over time is better represented by hourly data.
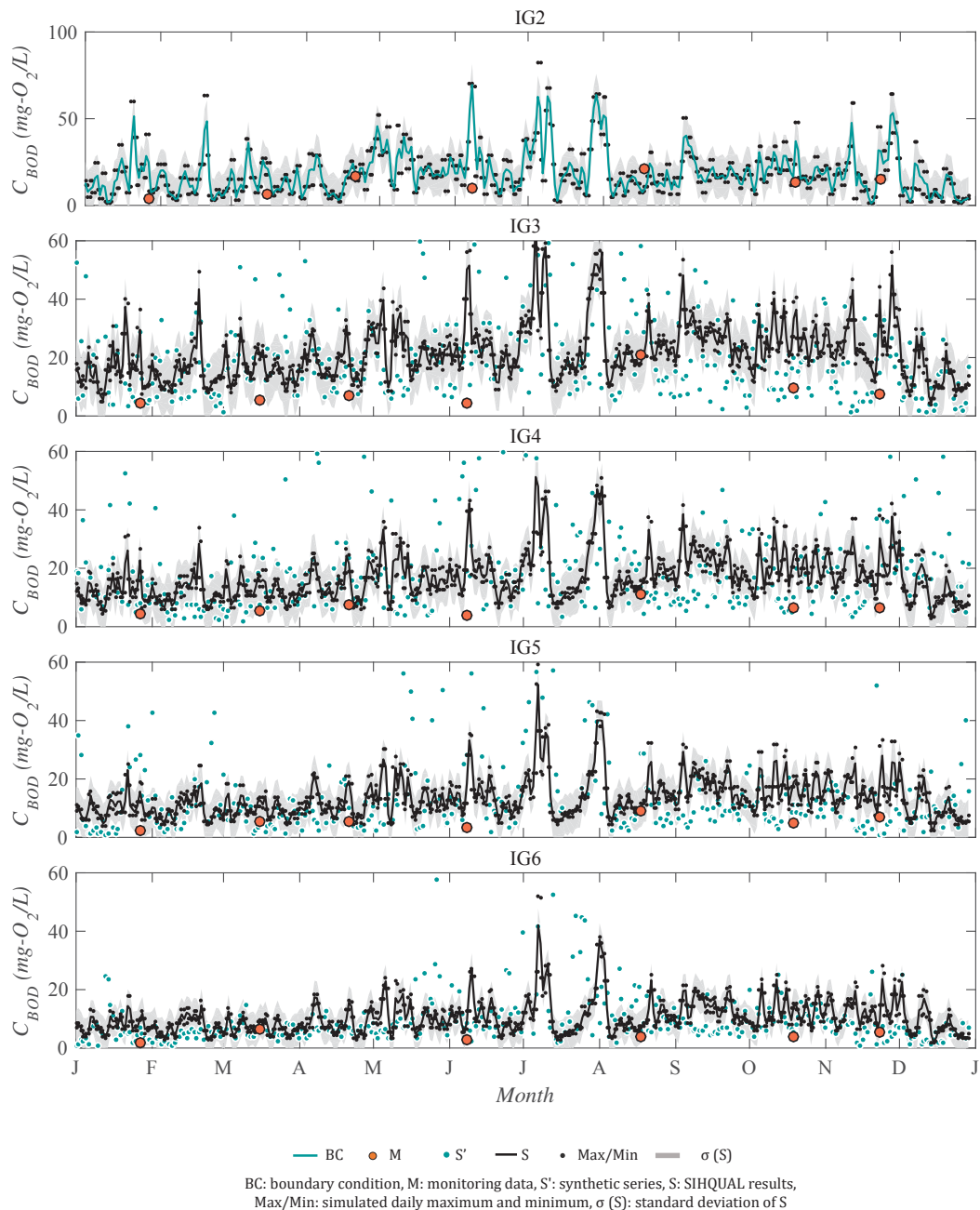
Figure 39: Daily BOD simulated concentrations and monitoring data for the year 2010; BC generated with series T1a
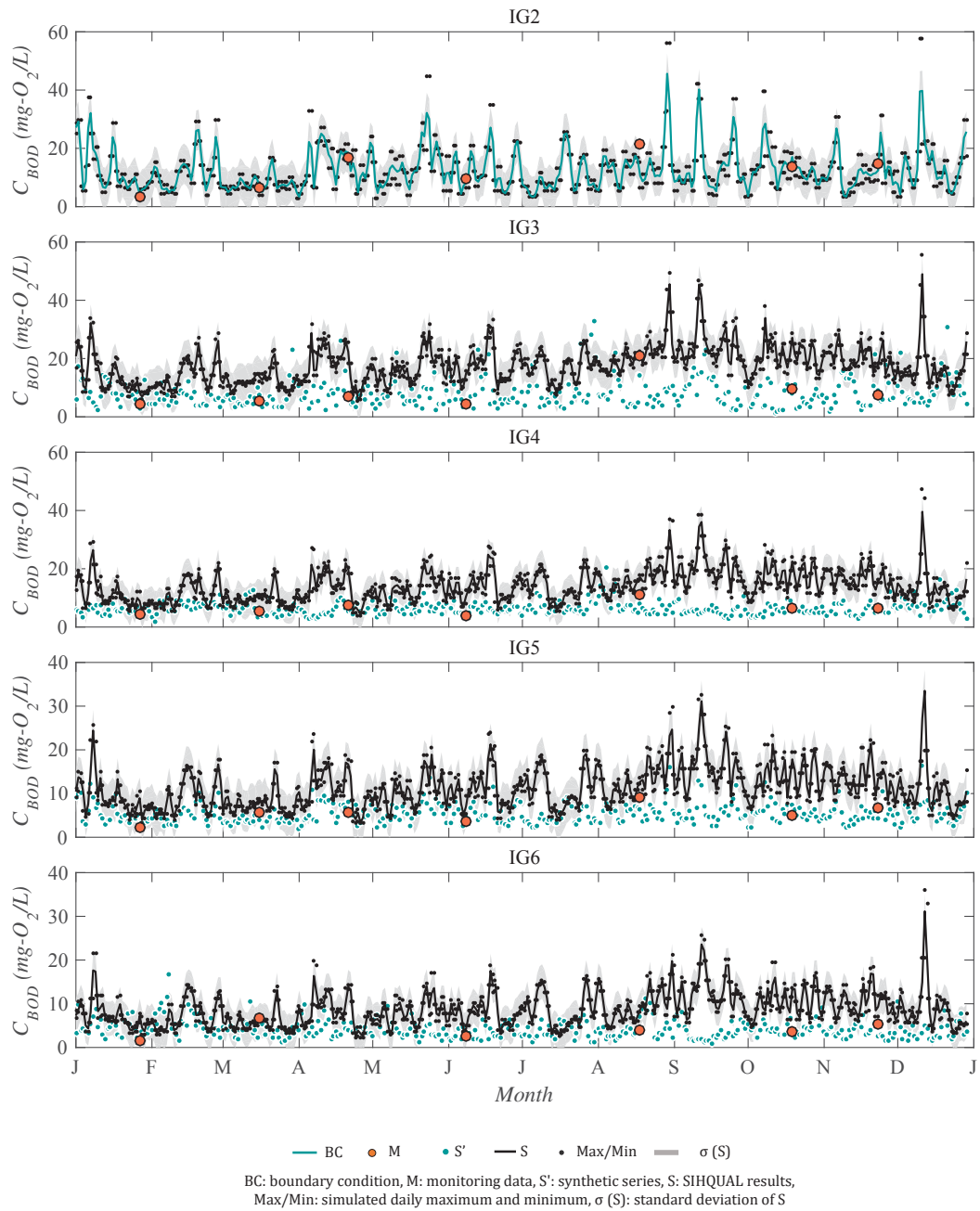
Figure 40: Daily BOD simulated concentrations and monitoring data for the year 2010; BC generated with series T4b
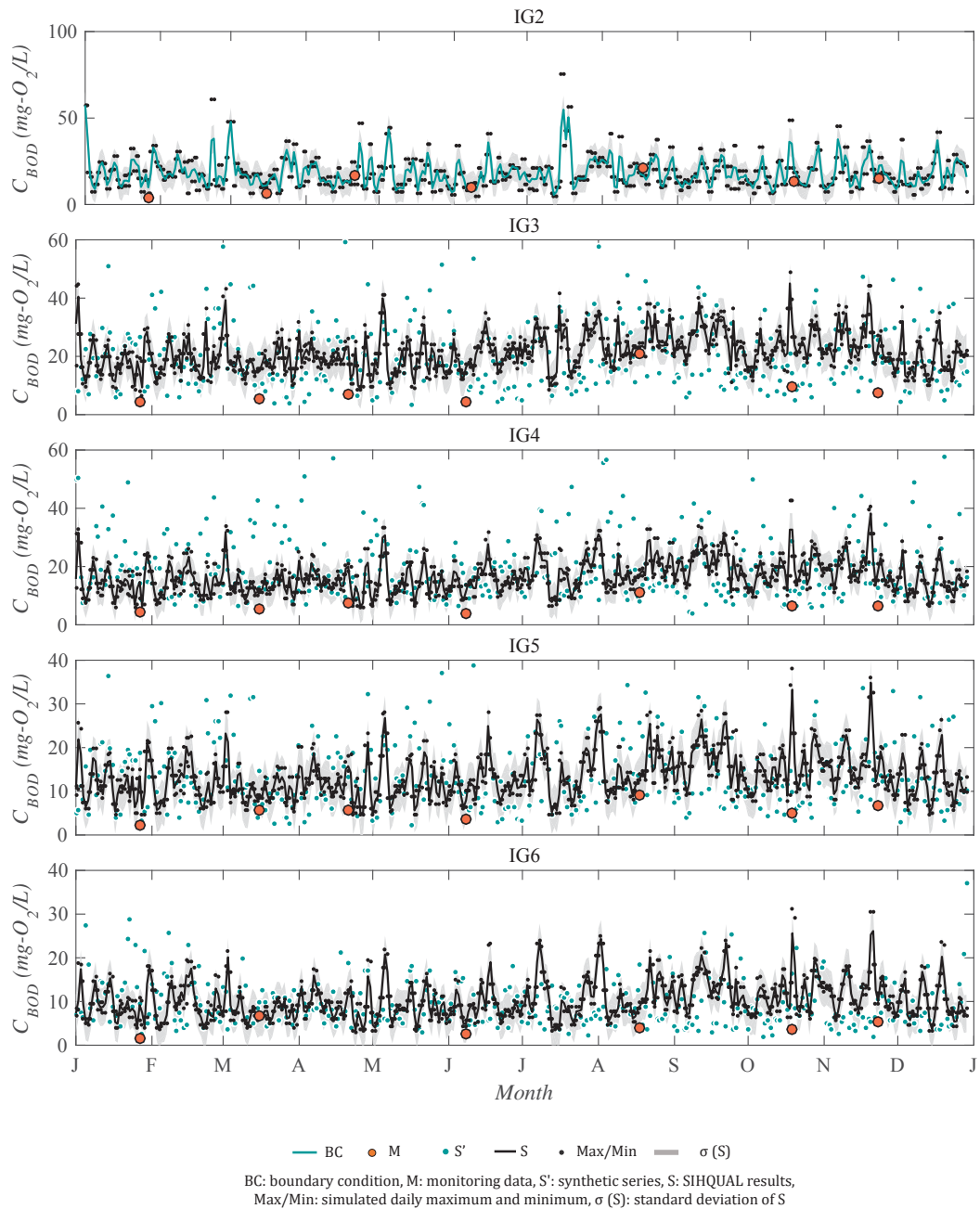
Figure 41: Daily BOD simulated concentrations and monitoring data for the year 2010; BC generated with series T9b

Besides test I1, tests T1b and T9b generated synthetic series with interquartile and median similar to the dataset in IG3 to IG6, as verified in figure 42 (although it should be stated that the series were chosen out of thousand options, and other simulations might suggest different conclusions).
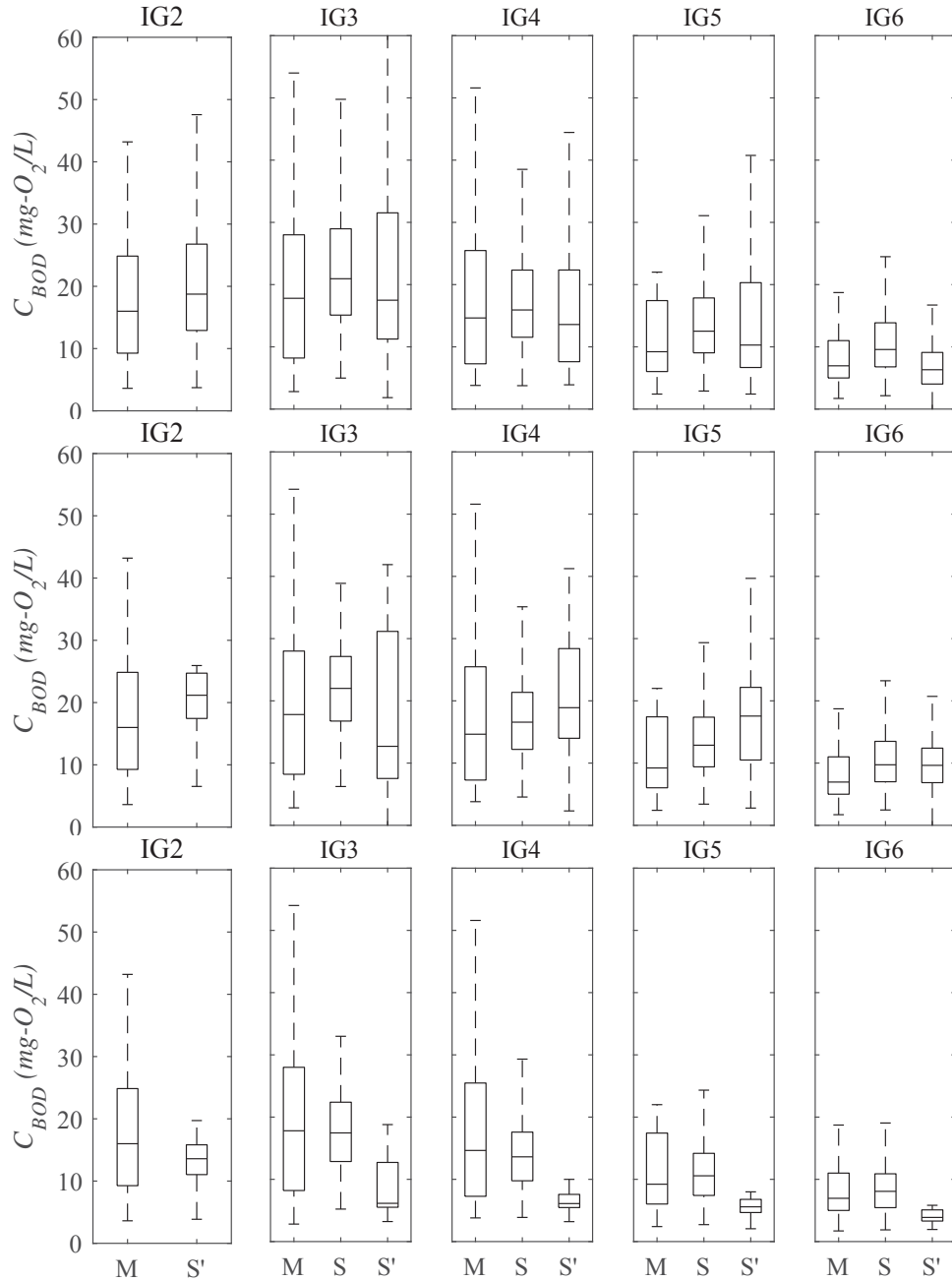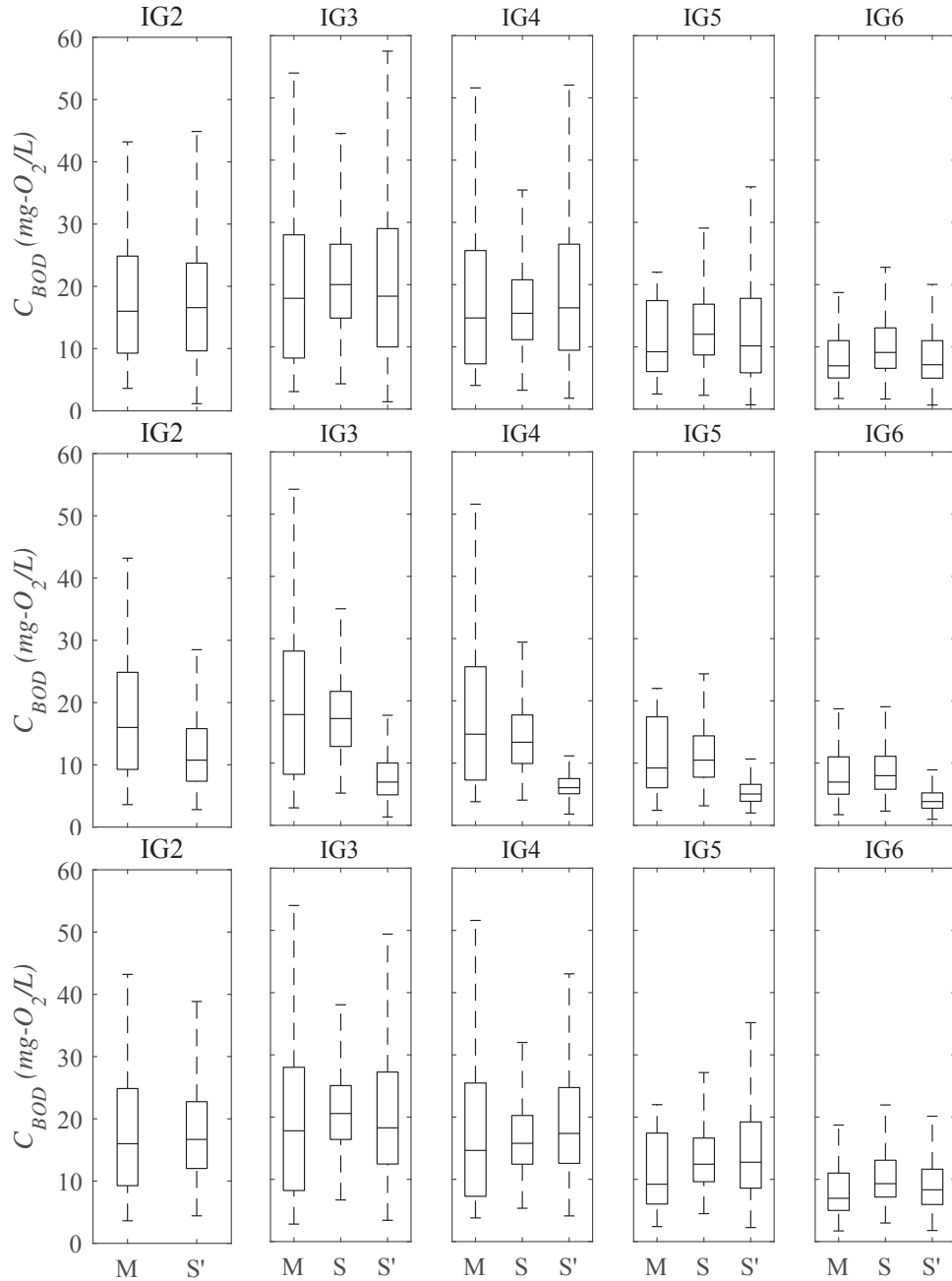
Figure 42: Boxplot of measured (M) and simulated BOD concentrations for 2010 with SIHQUAL (S) and synthetic series (S'); T1a, T4b, T9b, respectively in each row

The synthetic series generated for IG3 to IG6 have higher variability over time than the series generated by the deterministic model, specially for tests I1, I2, T1a and T9b. This is verified in figures 35, 36, 41 and 39. Such patter suggest that the synthetic series might be representing potential conditions not predicted by the approach based on conservation laws; in this case, the conditions are extreme events (high concentrations). In a similar analysis, Siqueira et al. (2016) applied a deterministic hydrological model and an ensemble approach for flood forecasting. The latter technique provides a set of different scenarios for predicted discharges. The authors identified that, in comparison with the first forecast method, the ensemble ones showed higher accuracy and probability

of detection for reference thresholds. Therefore, it is suggested that peaks may be better identified and predicted when different possible solutions are evaluated.

The boxplots for IG2 (figures 38 and 42) show that median concentrations and interquartile interval are more similar to the measured dataset for tests I1, T1b and T9b, which implies that these tests may be reliable as boundary conditions. This is also verified in the duration curves built with data from each test, shown in figure 43.

The duration curves for each test are compared with the monitoring dataset from 2005 to 2017 and with curves of modeling outputs under steady state (figure 43). The latter information is generated using multiple runs of steady solutions, with discharges of different frequencies.
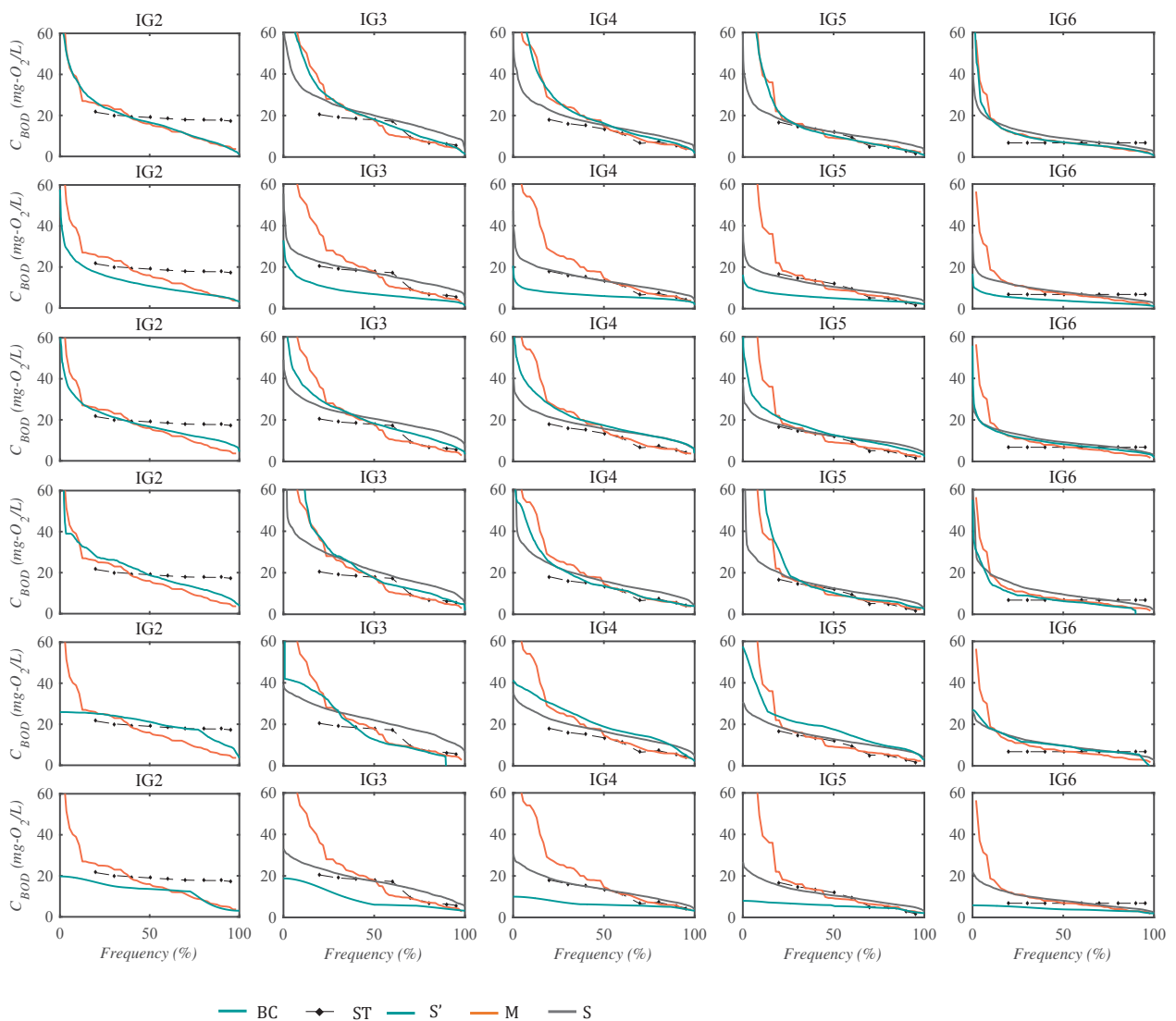
Figure 43: Duration curves of historical monitoring dataset 2005-2017 (M) and simulations for 2010 with synthetic series (S') and SIHQUAL results under unsteady (S) and steady (ST) state: T1a, T4b, T9b, I1, I2, I3 presented respectively in each row

In general, the results show an agreement between methods for concentrations of high frequency occurrence. Representation of critical events, on the other hand, are divergent. The most accurate tests for synthetic pollutographs at IG2 are T1a, T9b and I1, as also suggested in the analysis of boxplots.

Figure 44 compares in detail concentrations with 10% ($C_{10\%}$) and 90% ($C_{90\%}$) of occurrence, besides quartiles of SIHQUAL outputs with each boundary condition (Q1, Q2 and Q3); table 15 complements it, presenting the difference in $mg\text{-}O_2/L$ between concentrations simulated with each test and monitored data in terms of $C_{10\%}$, $C_{90\%}$, Q1, Q2 and Q3. Larger differences are observed for section where temporal variability is higher – mainly IG3 and IG4, that also receive larger amounts of pollution. Concentrations of monitoring data and steady modeling are also presented for reference; steady $C_{10\%}$ concentration is not shown because the numerical solution is unstable with the corresponding discharge of 10% of occurrence.
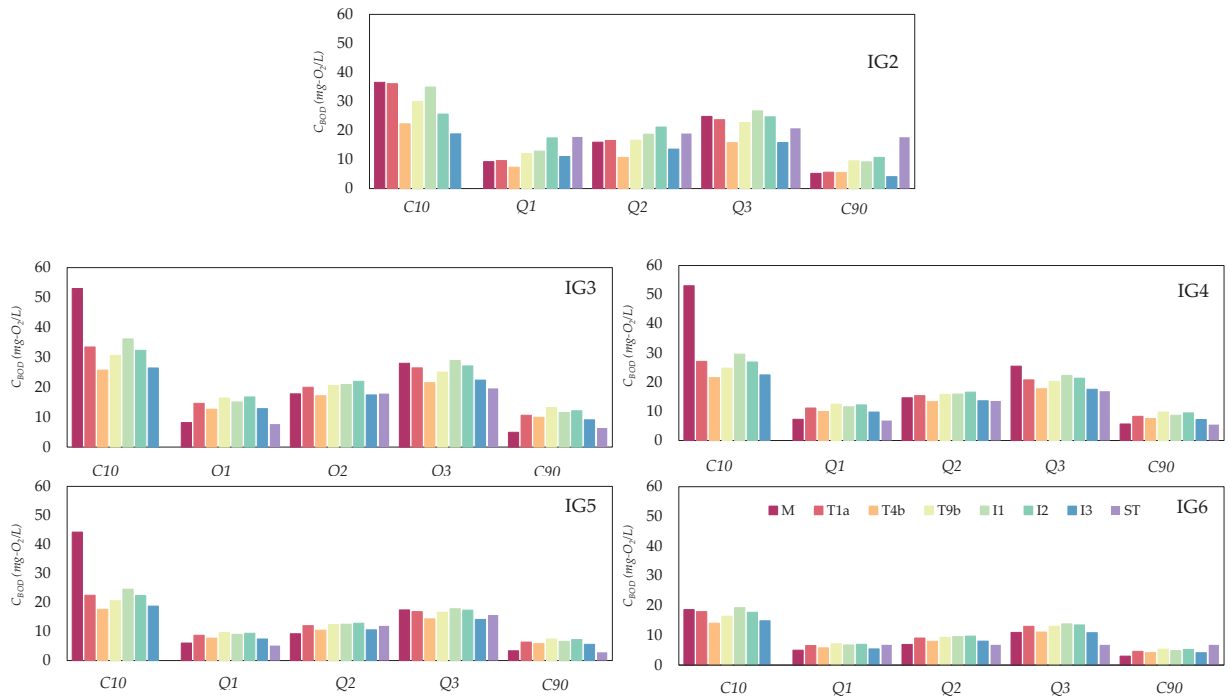
Figure 44: Measures $C_{10\%}$, $C_{90\%}$, Q1, Q2 and Q3 of each simulated and monitored concentrations BOD dataset; M - monitoring dataset; ST - steady simulations

Results reveal the overall effects when choosing different upstream boundary conditions for the unsteady approach, in comparison with the historical dataset; for all sections most tests overestimated $C_{90\%}$ and underestimate $C_{10\%}$. Table 15 shows that the absolute differences for $C_{90\%}$ simulated with SIHQUAL under unsteady state ranges between 1.16 $mg\text{-}O_2/L$ (IG6, I3) and 8.25 $mg\text{-}O_2/L$ (IG3, T9b).

For predicted $C_{10\%}$, absolute differences are between 0.67 (IG6, I1) and 31.49 mg-O_2/L (IG4, T4b). Reproduction of extreme events in this case are more sensible. Differences of simulated medians (Q2), on the other hand, ranged from 0.34 (IG3, I3) to 4.20

mg-O_2/L (IG3, I2). Often used as reference for calibration (e. g. Alaghmand et al., 2011; Wöhling et al., 2013), median outputs may also depend on the boundary conditions.

Table 15 also allows a conceptual comparison between the three AR(1) configurations (different scenarios of sampling frequency), showing that: *i)* the scenario with only quarterly data in the year (T4b) produces fair $C_{90\%}$ estimates in comparison with other tests, but poorest estimation of events with higher concentration ($C_{10\%}$); *ii)* more frequent data (T9b) did not necessarily improved results in terms of C10, Q1,Q2, Q3 or C90 estimates.

Table 15: Difference of $C_{10\%}$, $C_{90\%}$, Q1, Q2 and Q3 between monitored data and simulation results for each test ($C_{simulated}$ – $C_{monitored}$; mg-O_2/L)

Section	Metric	T1a	T4b	T9b	I1	I2	I3	ST
	C_{10}	-0.46	-14.30	-6.58	-1.63	-10.92	-17.76	-
	Q1	0.37	-1.90	2.77	3.65	8.19	1.78	8.53
IG2	Q2	0.58	-5.25	0.69	2.74	5.18	-2.41	2.97
	Q3	-1.12	-9.04	-2.10	1.92	-0.13	-9.03	-4.09
	C_{90}	0.43	0.35	4.34	3.96	5.52	-1.11	12.46
	C_{10}	-19.53	-27.26	-22.34	-16.85	-20.64	-26.55	-
	Q1	6.41	4.47	8.20	6.90	8.55	4.66	-0.50
IG3	Q2	2.20	-0.65	2.77	3.14	4.20	-0.34	0.06
	Q3	-1.52	-6.48	-2.94	0.97	-0.87	-5.61	-8.40
	C_{90}	5.67	5.01	8.25	6.60	7.17	4.16	1.42
	C_{10}	-25.90	-31.49	-28.24	-23.43	-26.14	-30.55	-
	Q1	3.88	2.67	5.16	4.27	4.93	2.49	-0.44
IG4	Q2	0.75	-1.29	1.14	1.28	1.93	-1.03	-1.10
	Q3	-4.70	-7.74	-5.22	-3.14	-4.11	-7.89	-8.55
	C_{90}	2.62	1.90	4.07	2.98	3.78	1.54	-0.24
	C_{10}	-21.76	-26.59	-23.65	-19.67	-21.83	-25.50	-
	Q1	2.67	1.70	3.59	3.01	3.36	1.38	-0.89
IG5	Q2	2.81	1.26	3.21	3.30	3.66	1.35	2.72
	Q3	-0.58	-3.05	-0.80	0.42	-0.05	-3.25	-1.77
	C_{90}	3.00	2.46	4.02	3.29	3.87	2.23	-0.51
	C_{10}	-0.68	-4.57	-2.22	0.67	-0.87	-3.75	-
	Q1	1.55	0.79	2.15	1.78	2.00	0.45	1.82
IG6	Q2	2.15	1.02	2.35	2.62	2.78	1.12	-0.13
	Q3	2.04	0.09	2.07	2.86	2.49	-0.11	-4.18
	C_{90}	1.61	1.18	2.32	1.86	2.27	1.16	3.82

5.3.3.2   Organic nitrogen (N-org)

Simulations for N-org with upstream boundary condition based on interpolation functions – I1, I2 and I3 – have similar behavior to BOD (figures 45 to 47); however, I1 and I2 generated concentrations with larger interquartile interval than expected in IG2 (as verified in boxplot of figure 48). Applied to downstream sections, the interpolation

function I2 also produced unreasonable concentrations (specially in IG4), being more sensible to flow conditions than BOD.

Figure 45: Daily N-org simulated concentrations and monitoring data for the year 2010; BC generated with test I1
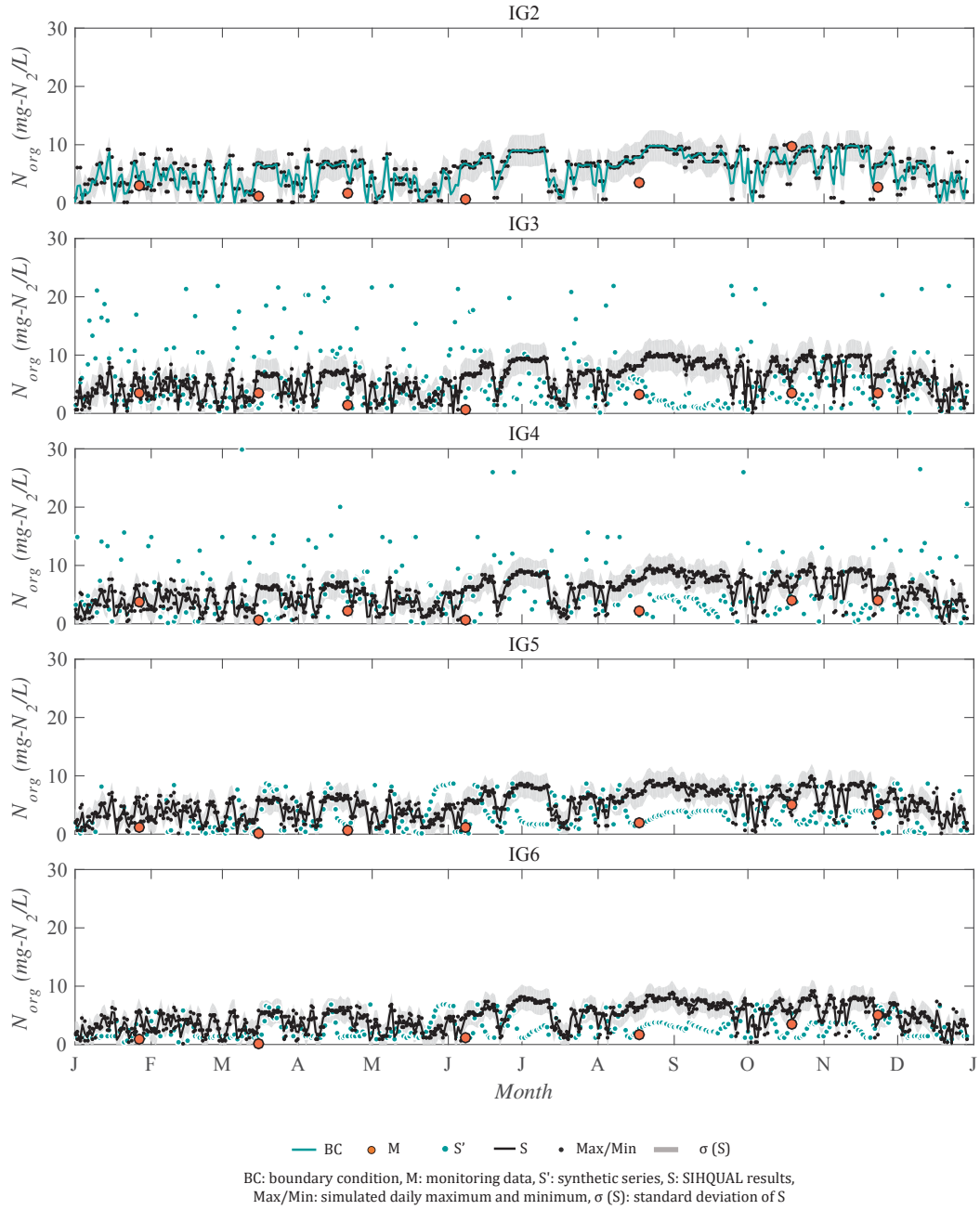
Figure 46: Daily N-org simulated concentrations and monitoring data for the year 2010; BC generated with test I2

Figure 47: Daily N-org simulated concentrations and monitoring data for the year 2010; BC generated with test I3

Figure 48: Boxplot of measured (M) and simulated N-org concentrations for 2010 with SI-HQUAL (S) and synthetic series (S'); I1, I2, I3, respectively in each row

Results obtained with T1a, T4b and T9b as boundary conditions show that simulated concentrations are close to overall variability of measured data (boxplots in figure 52). Large concentration peaks are generated in the daily pollutographs with test T1a and T4b (figures 49 and 50, respectively); series T9b, on the other hand, originate daily N-org concentration with a consistent variation, following the same pattern identified in

BOD modeling. A similar composition of duration curves is also verified (figure 53): as BOD, simulations for N-org differs mainly for estimation of critical events.

Figure 49: Daily N-org simulated concentrations and monitoring data for the year 2010; BC generated with test T1a
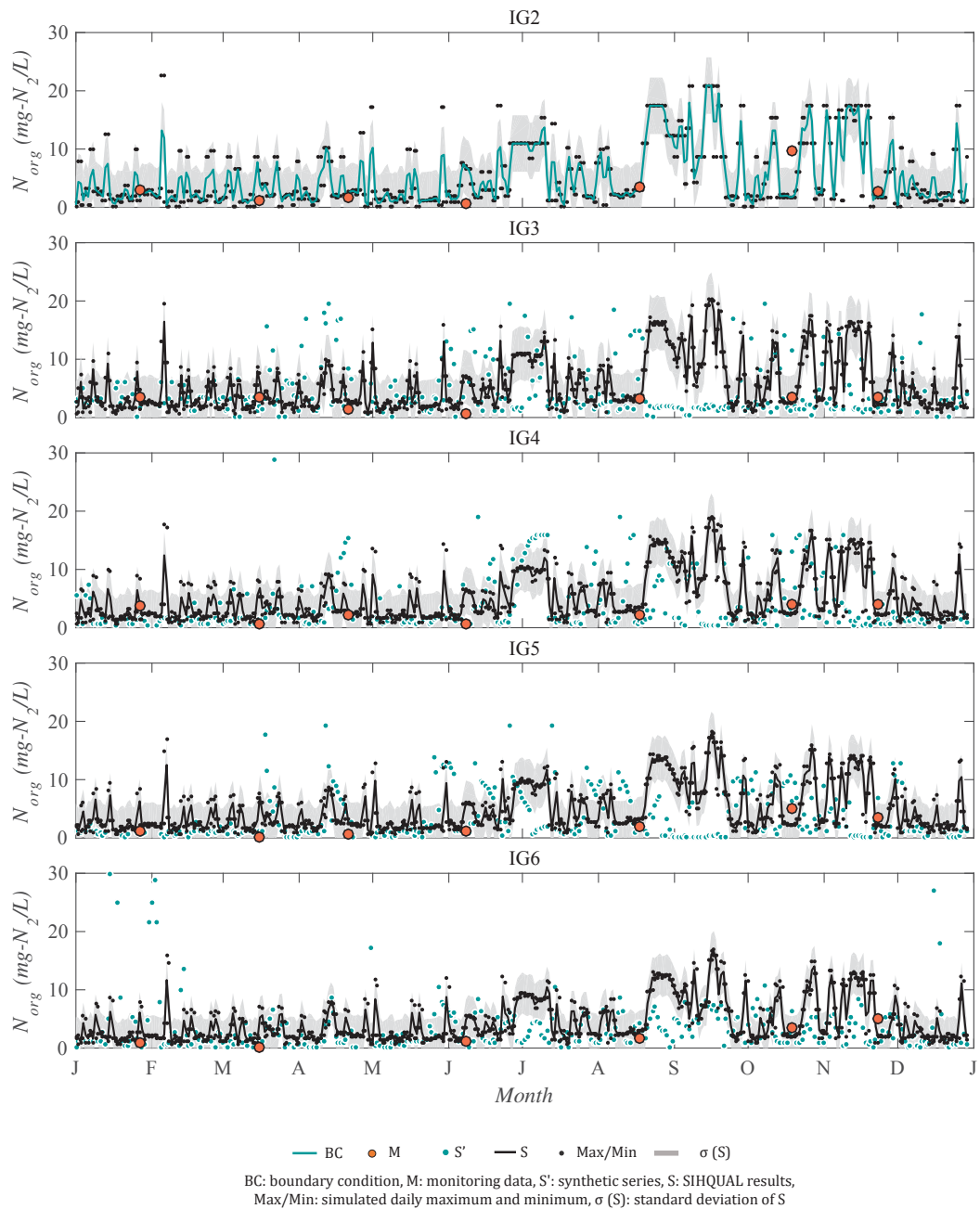
Figure 50: Daily N-org simulated concentrations and monitoring data for the year 2010; BC generated with test T4b

Figure 51: Daily N-org simulated concentrations and monitoring data for the year 2010; BC generated with test T9b
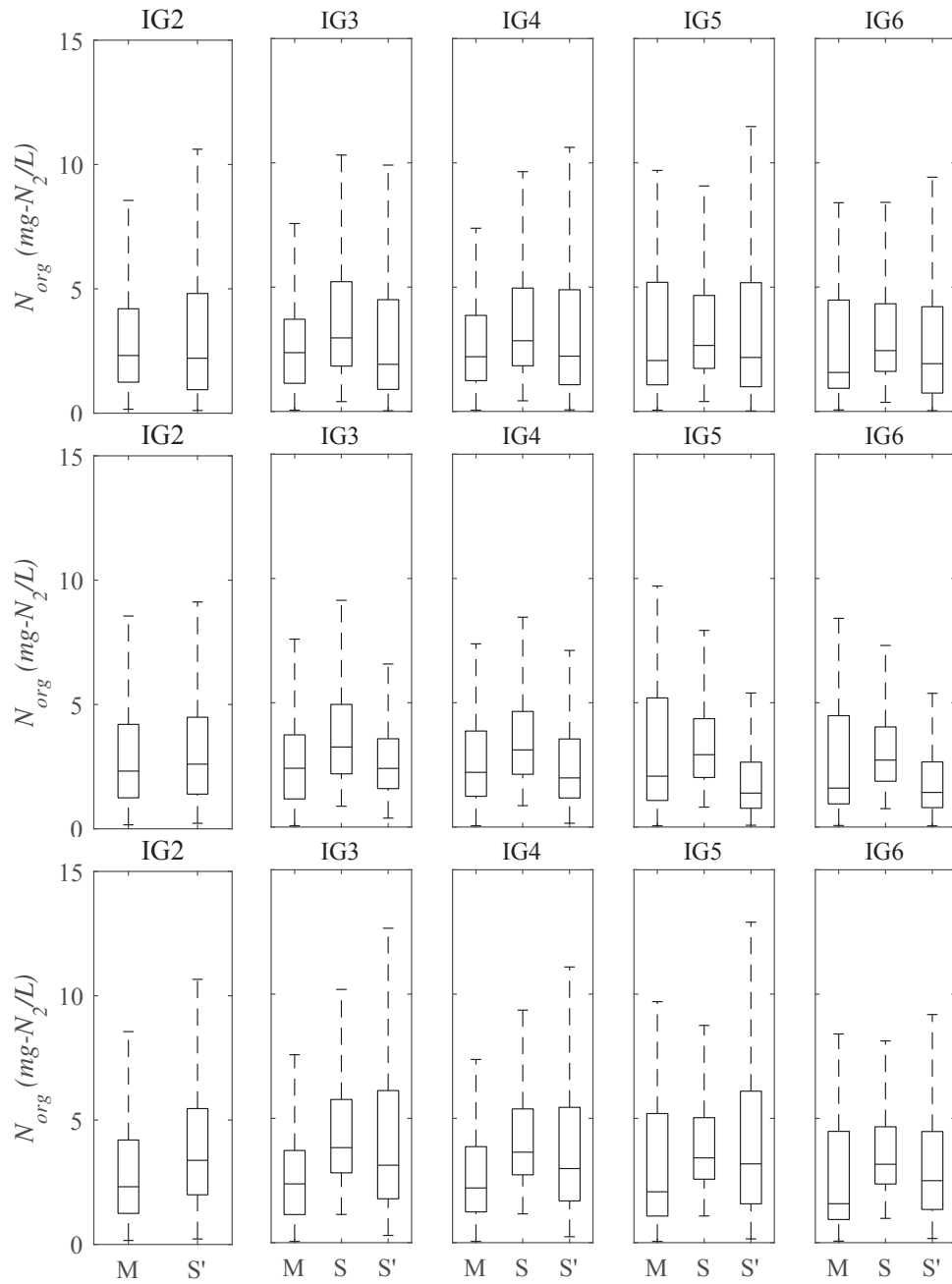
Figure 52: Boxplot of measured (M) and simulated (S) BOD concentrations for 2010; T1a, T4b, T9b, respectively in each row

Figure 53: Duration curves of historical monitoring dataset 2005-2017 (M) and simulations for 2010 with synthetic series (S') and SIHQUAL results under unsteady (S) and steady (ST) state: T1a, T4b, T9b, I1, I2, I3 presented respectively in each row

5.3.3.3 Dissolved Organic Carbon (DOC)

Following the patter of other presented parameters, tests I1 and I2 generated unreasonable daily mean DOC concentrations (figures 54 and 55), although interquartile variation is reasonably similar to data (figure 57).

For BOD and N-org, even though test I3 produced smooth variations over time, a certain oscillation is present in pollutographs, due to lateral contributions and calibration strategy. For DOC, however, this fluctuation is imperceptible (figure 56). The same behavior is observed from results with T4b as boundary condition: the small variability over time in IG2 is reproduced in IG3 to IG6 (figure 59), with boxplots characterized by

smaller interquartile range than expected (figure 61) and duration curves with a shape nearly constant (figure 62). These results suggest that the parameter DOC travels at almost the same speed as water, and lateral inputs and transformation effects cause minimum changes in concentration.

Tests T1a introduced seasonal variation (figure 58), generating duration curves very similar to the dataset (figure 62). Series T9b also produces fair results (figure 60 and 62), although there is some overestimation, verified mainly in the boxplots of IG5 and IG6 (figure 57).

Figure 54: Daily DOC simulated concentrations and monitoring data for the year 2010; BC generated with PCHIP interpolations (test I1)

Figure 55: Daily DOC simulated concentrations and monitoring data for the year 2010; BC generated with Fourier series (test I2)

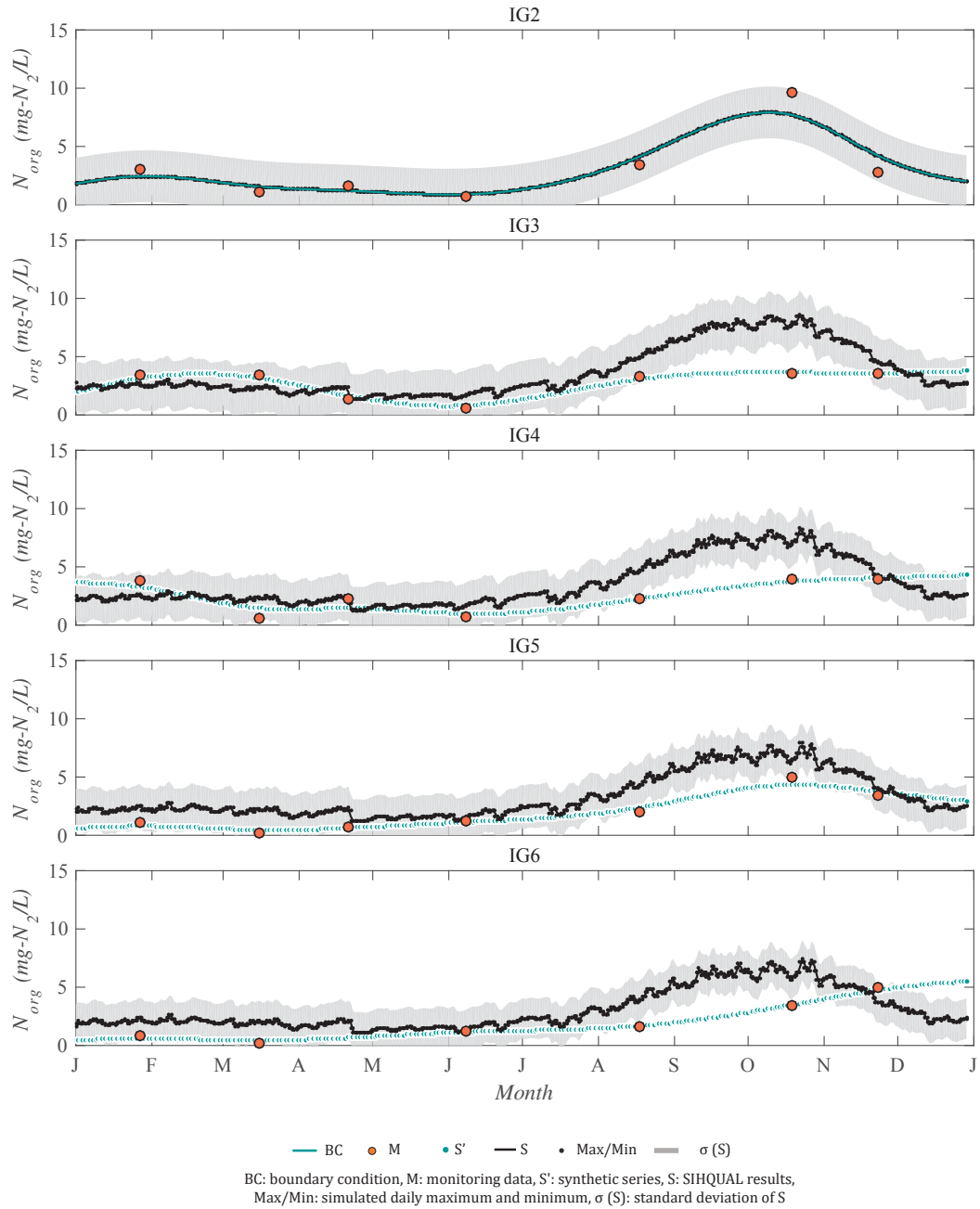Figure 56: Daily DOC simulated concentrations and monitoring data for the year 2010; BC generated with Spline interpolations (test I3)
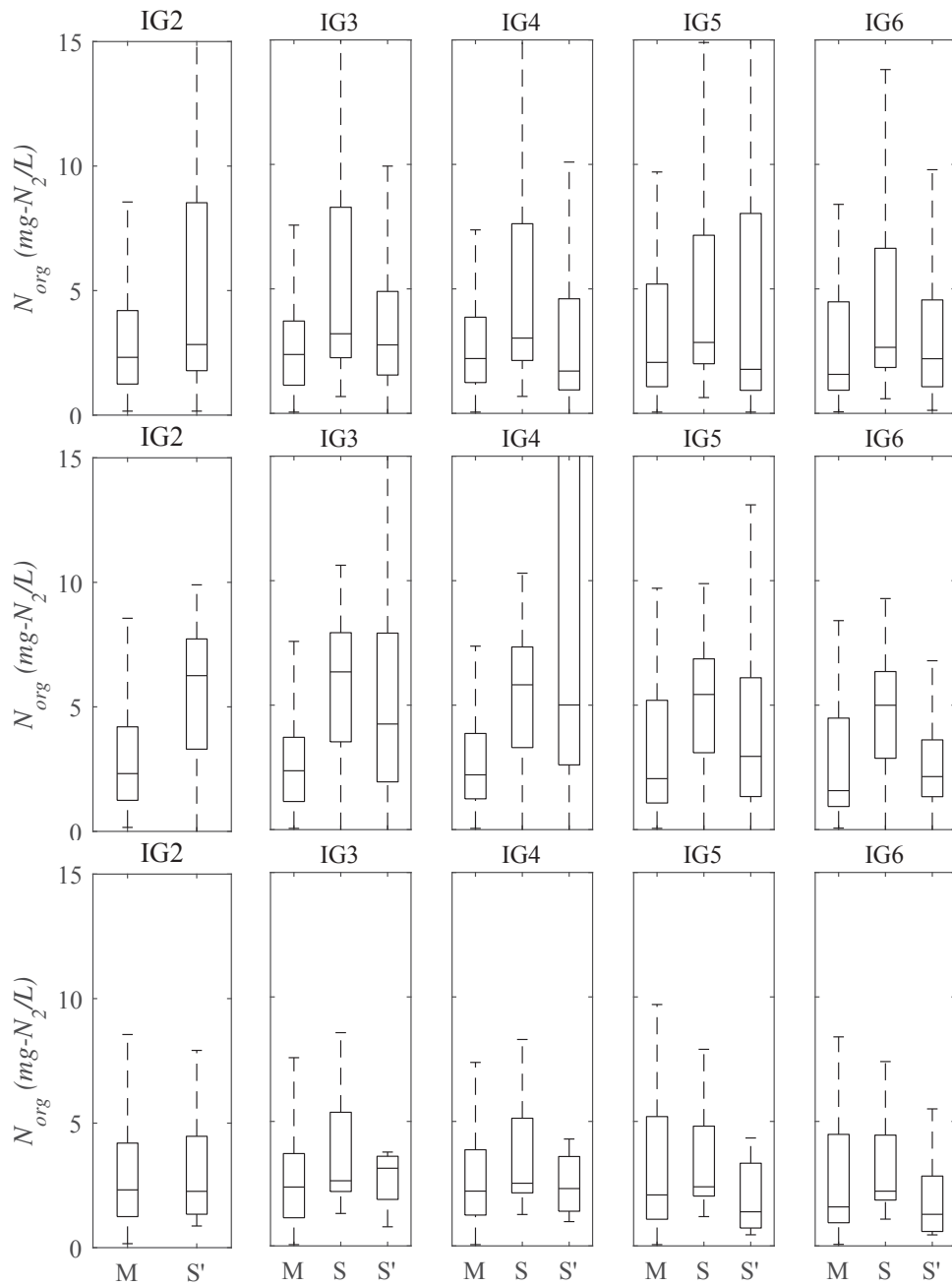
Figure 57: Boxplot of measured (M) and simulated DOC concentrations for 2010 with SIHQUAL (S) and synthetic series (S'); I1, I2, I3, respectively in each row
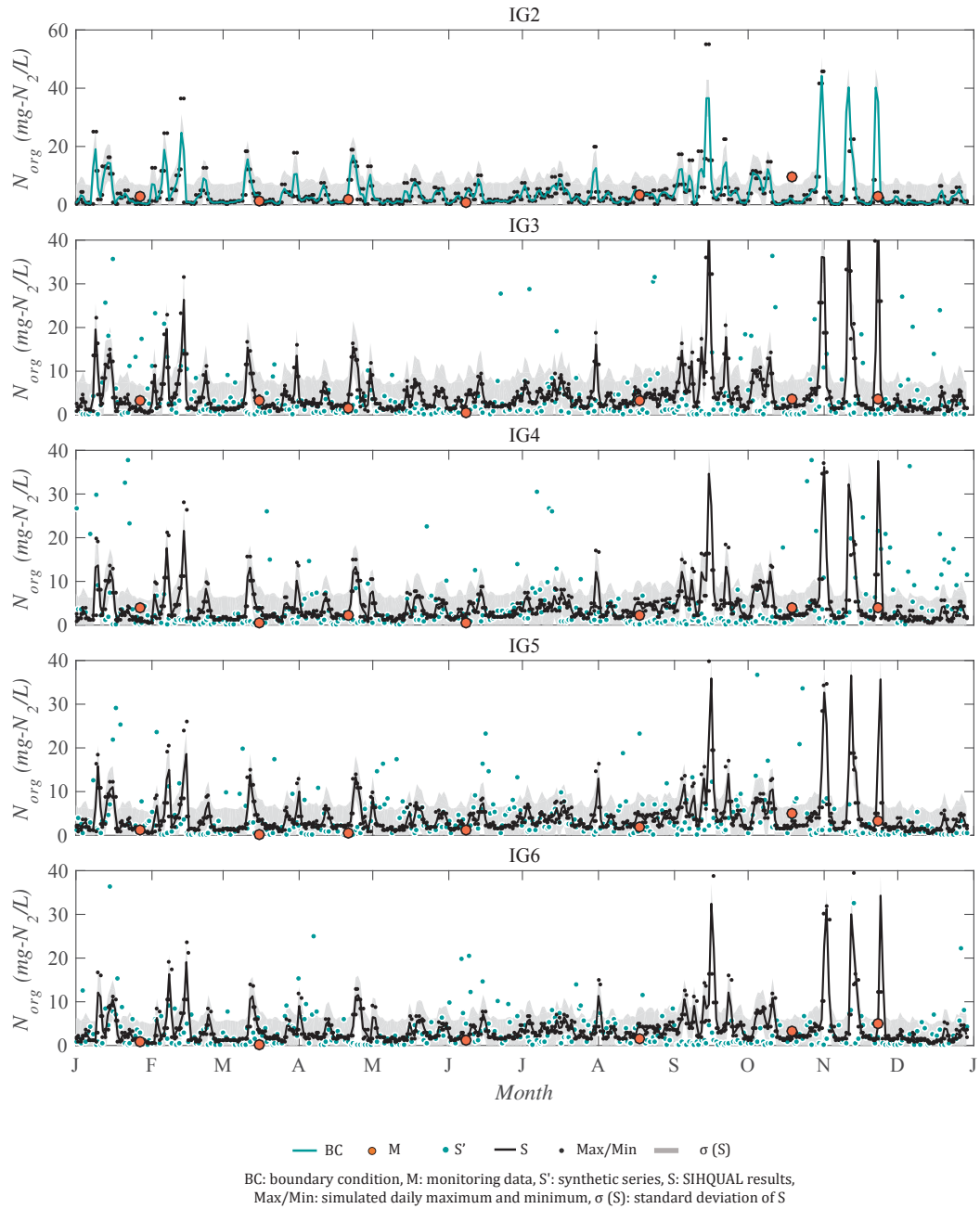
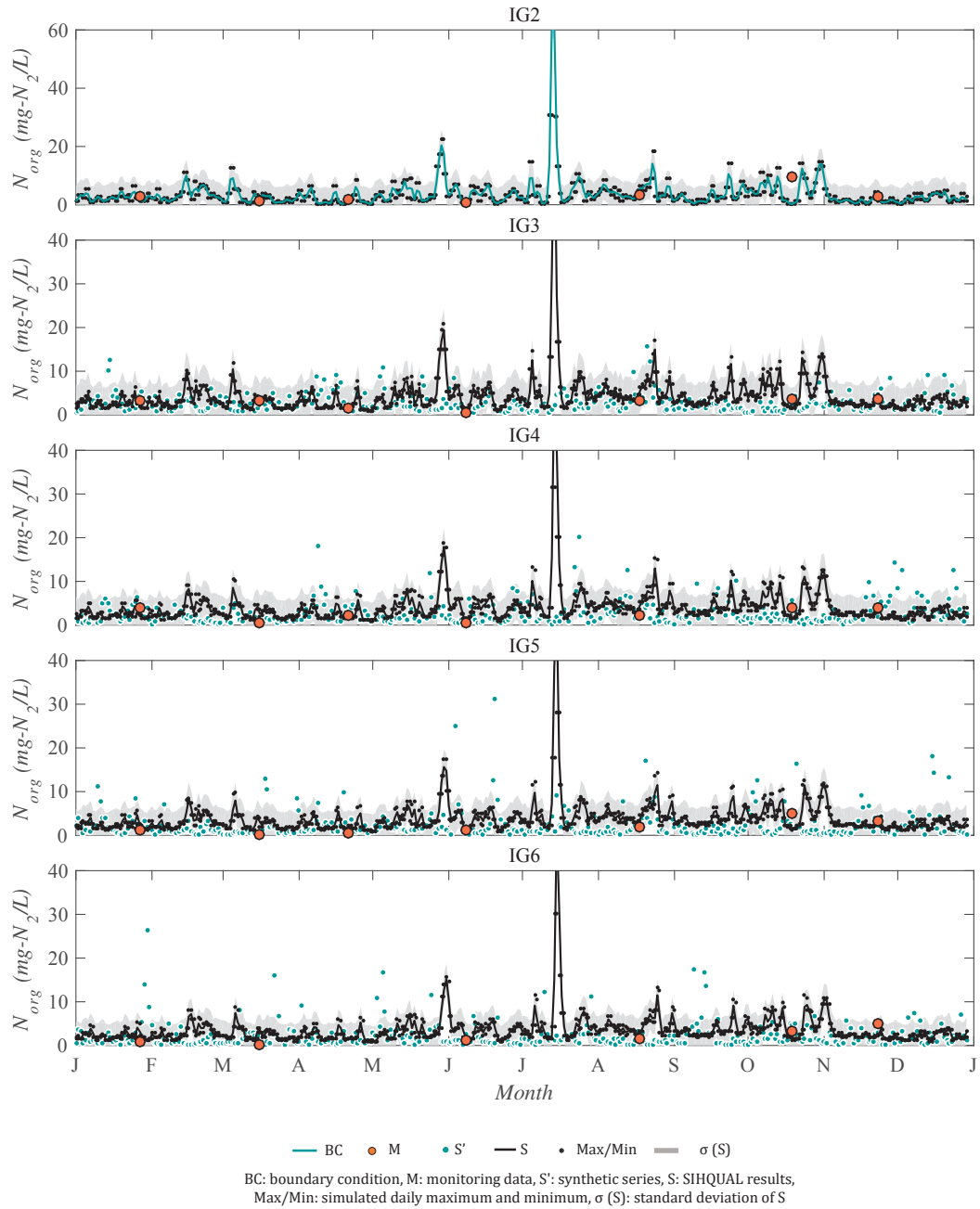Figure 58: Daily DOC simulated concentrations and monitoring data for the year 2010; BC generated with series T1a
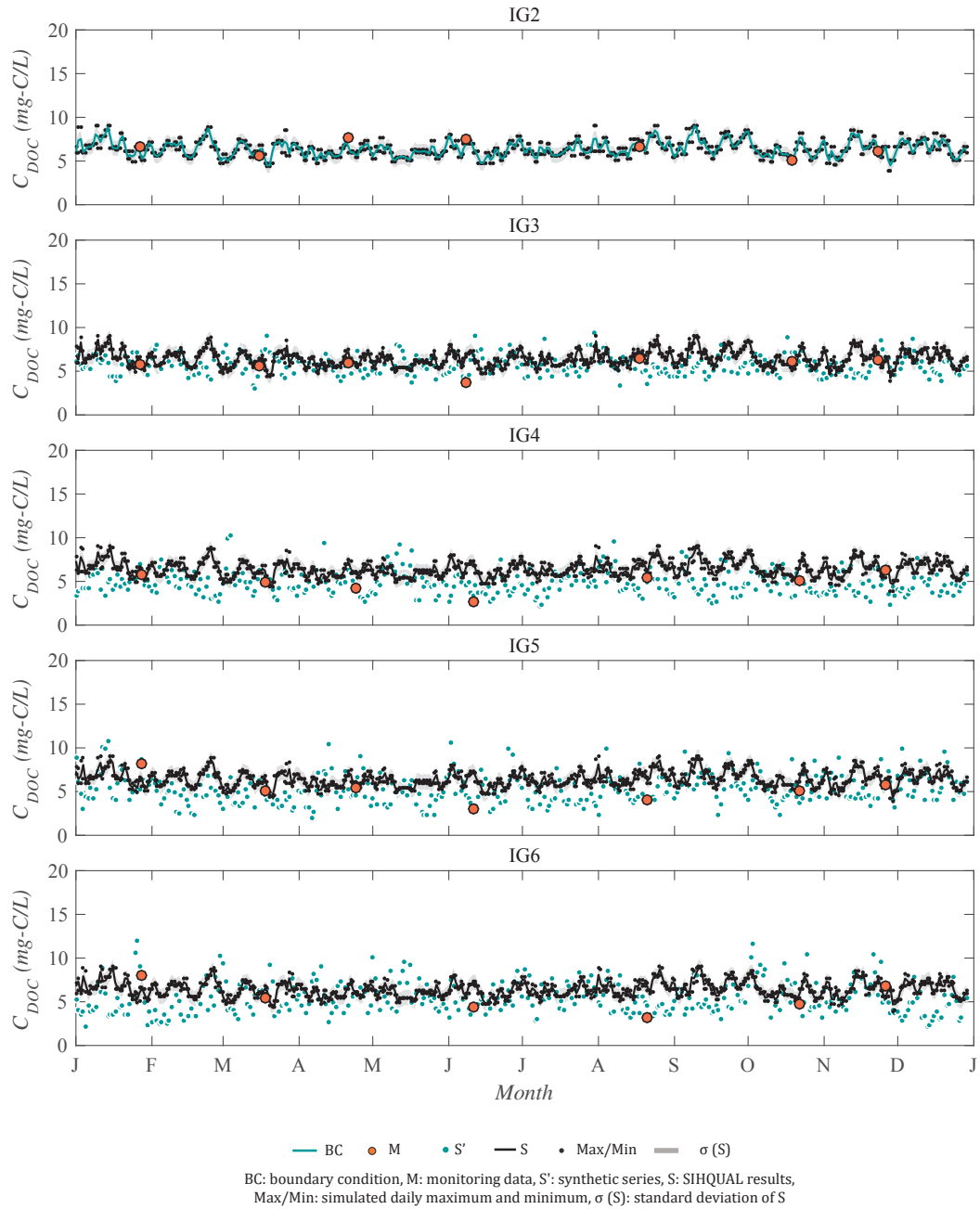
Figure 59: Daily DOC simulated concentrations and monitoring data for the year 2010; BC generated with series T4b
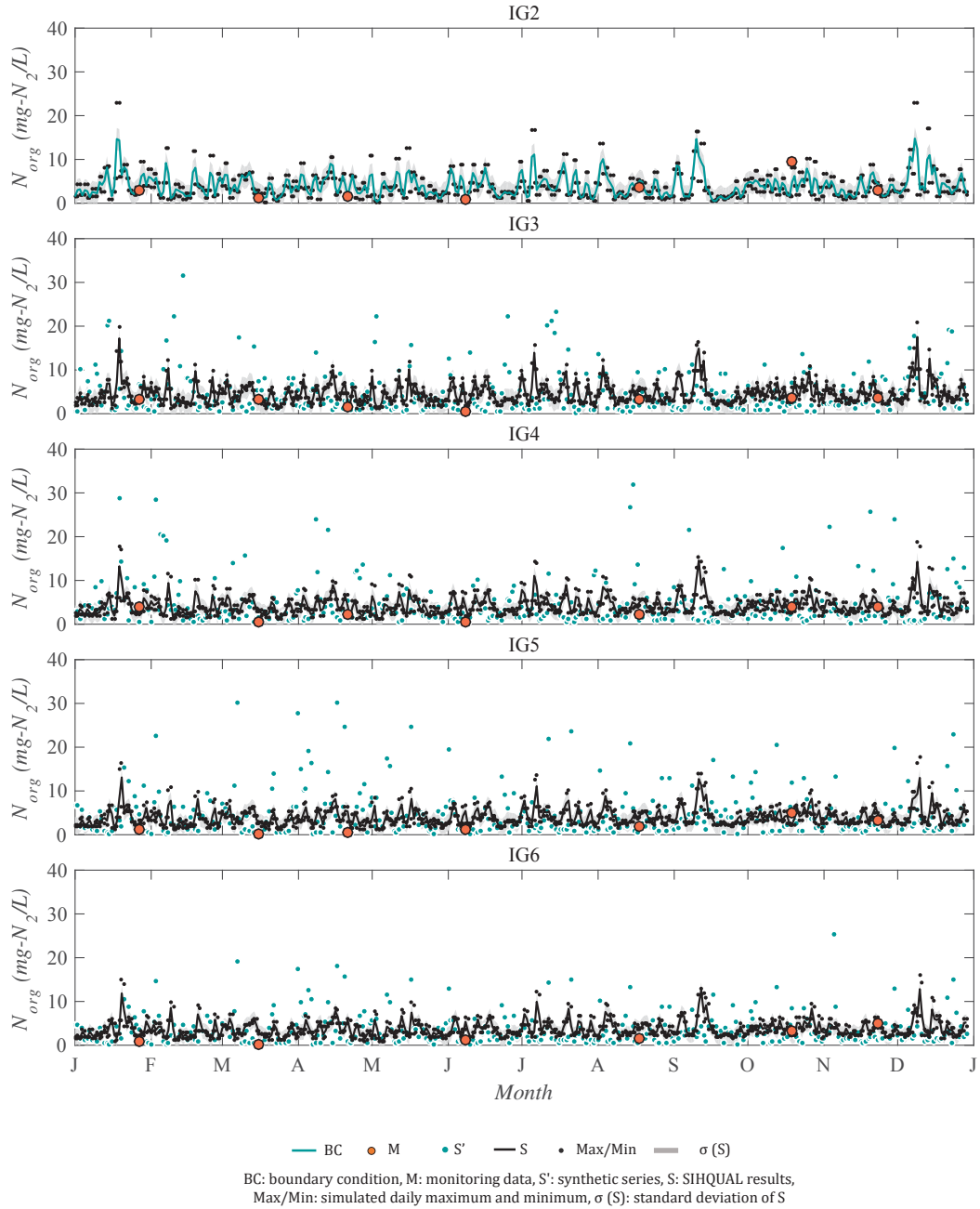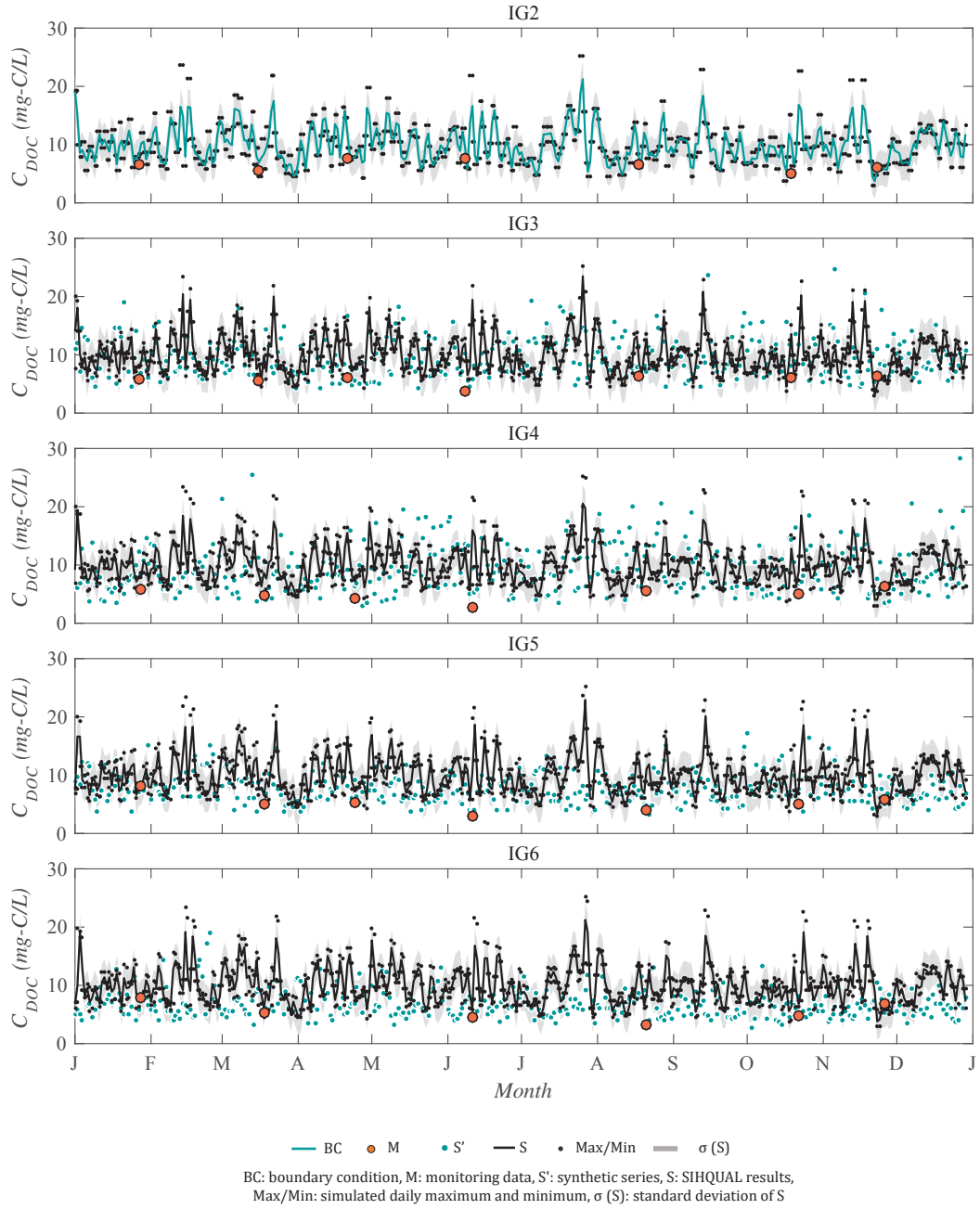
Figure 60: Daily DOC simulated concentrations and monitoring data for the year 2010; BC generated with series T9b
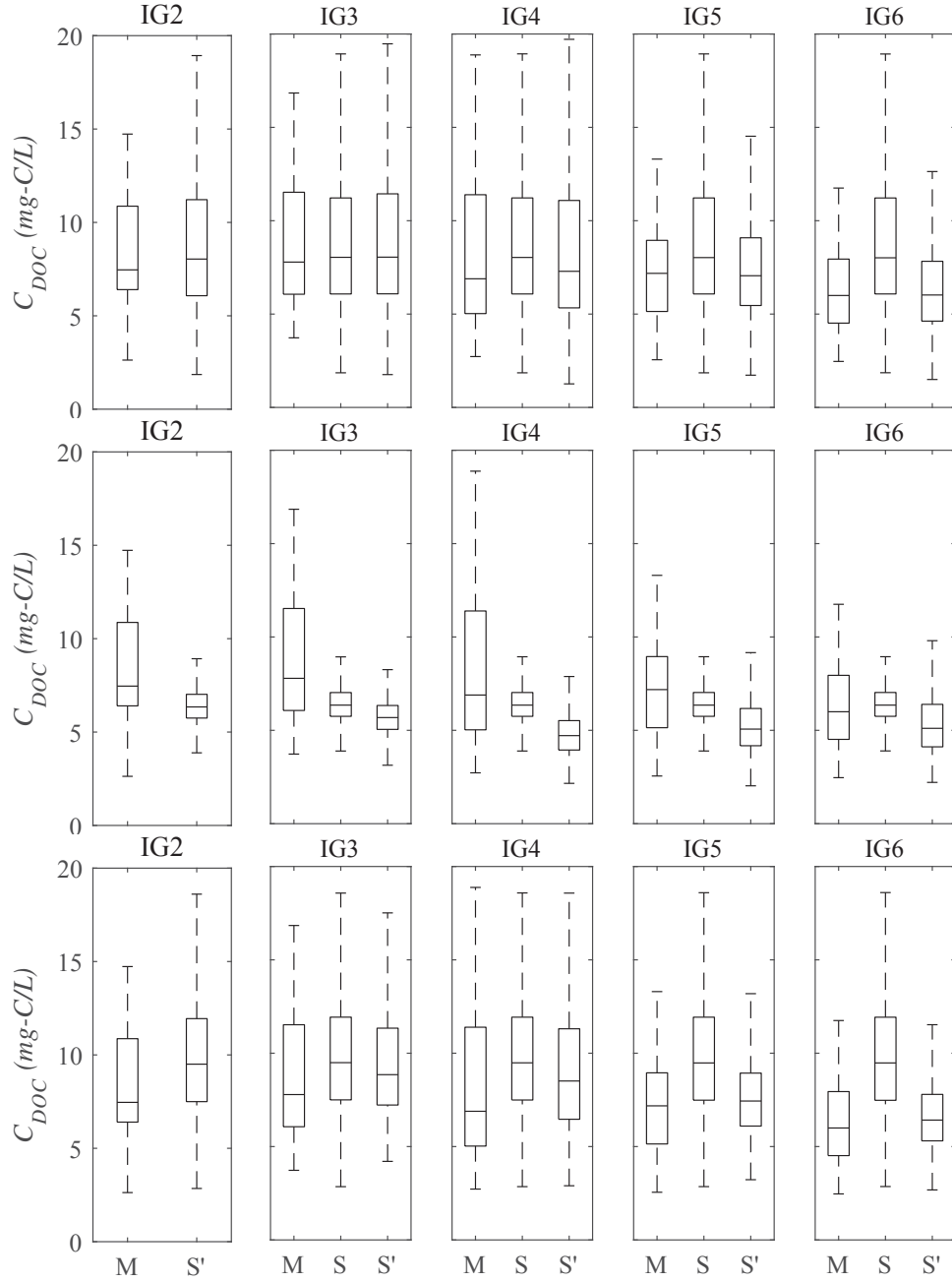
Figure 61: Boxplot of measured (M) and simulated DOC concentrations for 2010 with SIHQUAL (S) and synthetic series (S'); T1a, T4b, T9b, respectively in each row
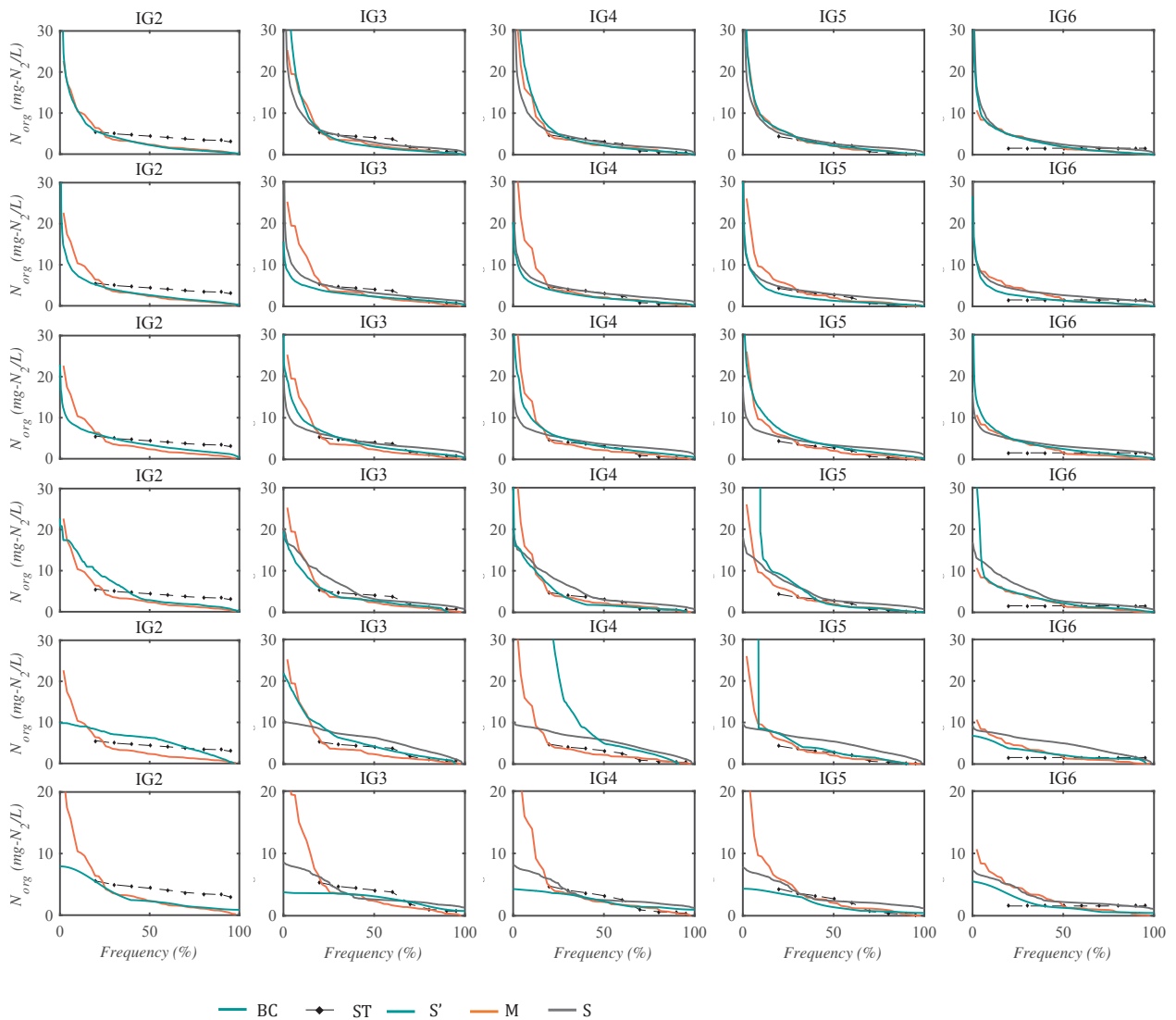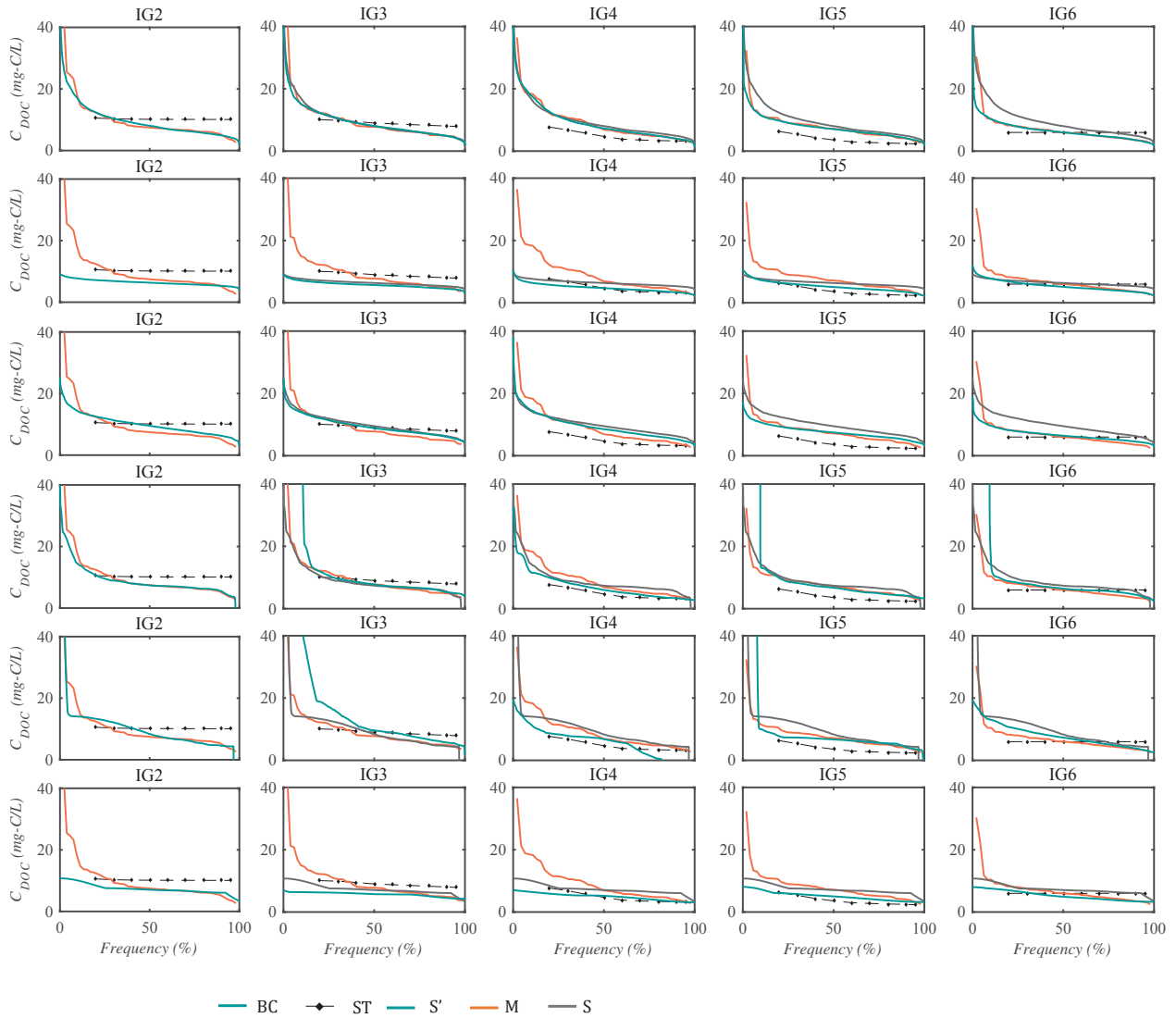
Figure 62: Duration curves of historical monitoring dataset 2005-2017 (M) and simulations for 2010 with synthetic series (S') and SIHQUAL results under unsteady (S) and steady (ST) state: T1a, T4b, T9b, I1, I2, I3 presented respectively in each row

5.3.3.4 Dissolved Oxygen (DO)

Similarly to DOC, synthetic series with I1 and I2 generated unreasonable data for DO concentration in IG2, and are not evaluated as boundary conditions for simulation. Synthetic series in IG2 that generated high concentrations are set to a maximum of 8.5 mg-O_2/L (arbitrated saturation limit), although this value can change due to temperature, salinity and pressure conditions (Chapra, 1997).

The experiments show that series I3 generated results closer to data than the deterministic approach (figures 63 and 67). Synthetic DO concentrations with T1a and T9b generated higher variability than expected in IG3 to IG6 (boxplot in figure 67); simu-

lation based on these tests as boundary conditions show fair estimation (figures 64 and 66), considering that this parameter is affected by multiple processes in the system, such as nitrification, reaeration and water velocity, for example. Upstream boundary condition defined with T4b also resulted in reasonable concentrations with SIHQUAL (boxplot in figure 67); analysis of boxplots, however, indicates that the AR(1) model with mean and standard deviation of the simulation period predicts overall variability close to the monitored dataset in most sections.
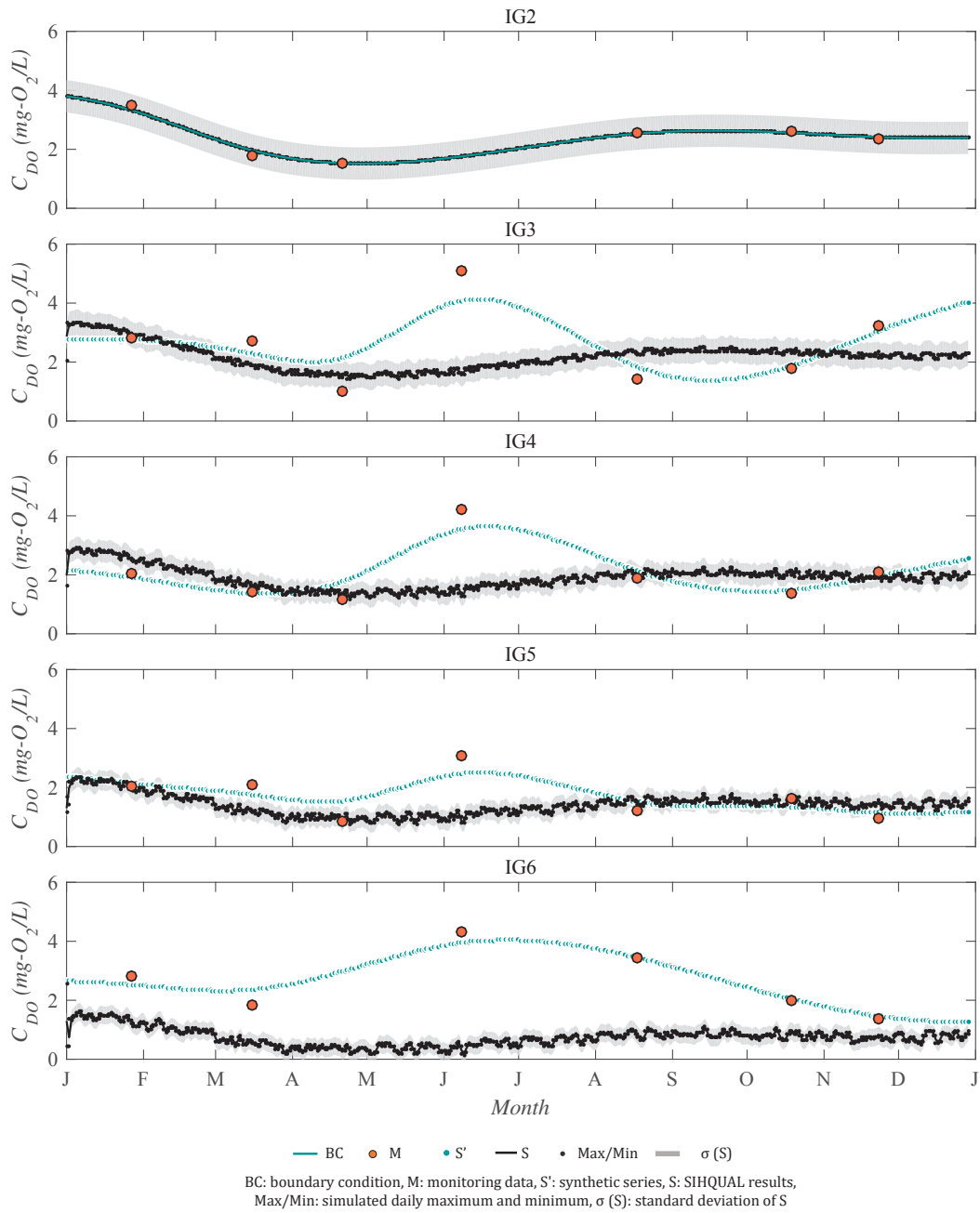
Figure 63: Daily DO simulated concentrations and monitoring data for the year 2010; BC generated with Spline interpolation (test I3)
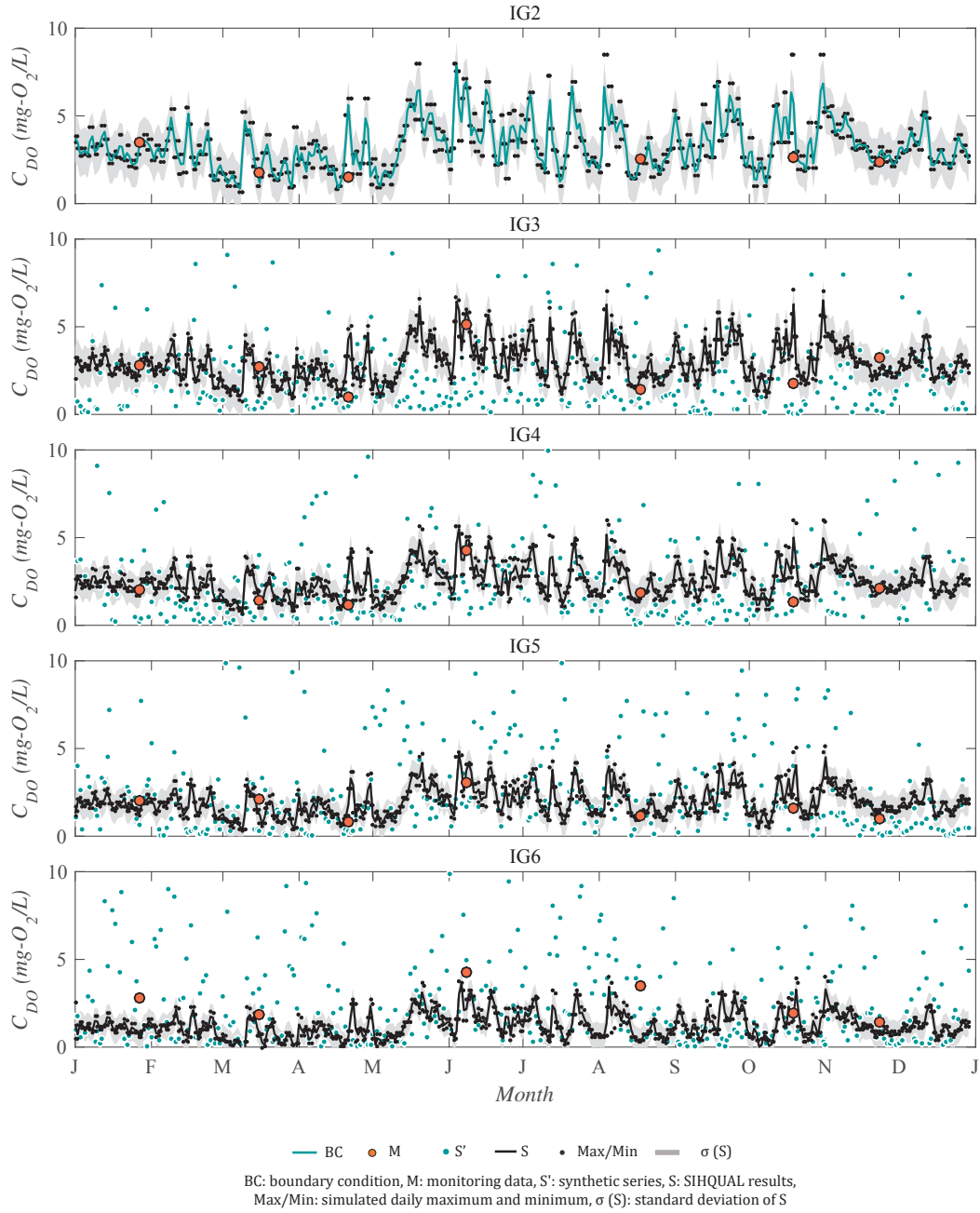
Figure 64: Daily DO simulated concentrations and monitoring data for the year 2010; BC generated with series test T1a
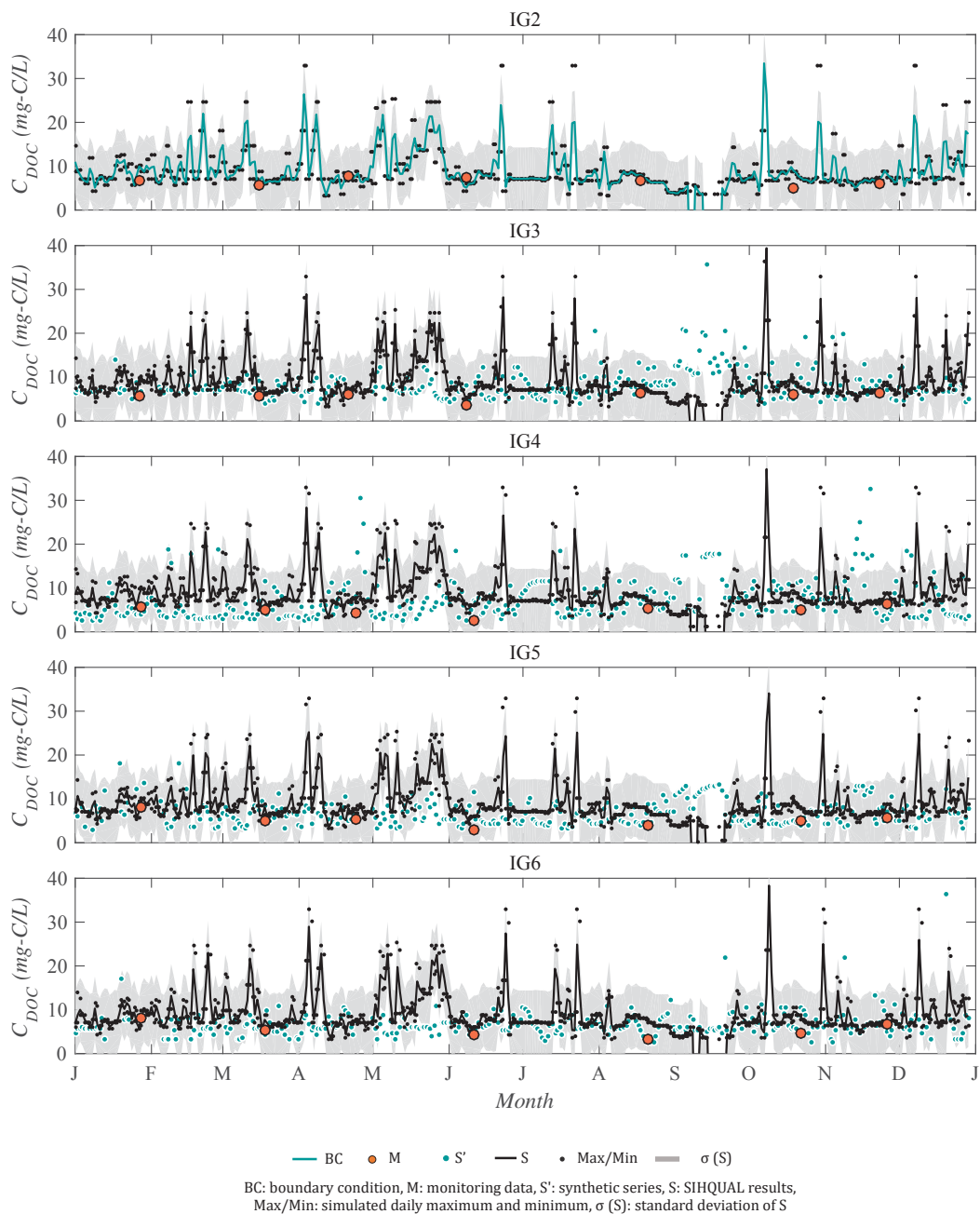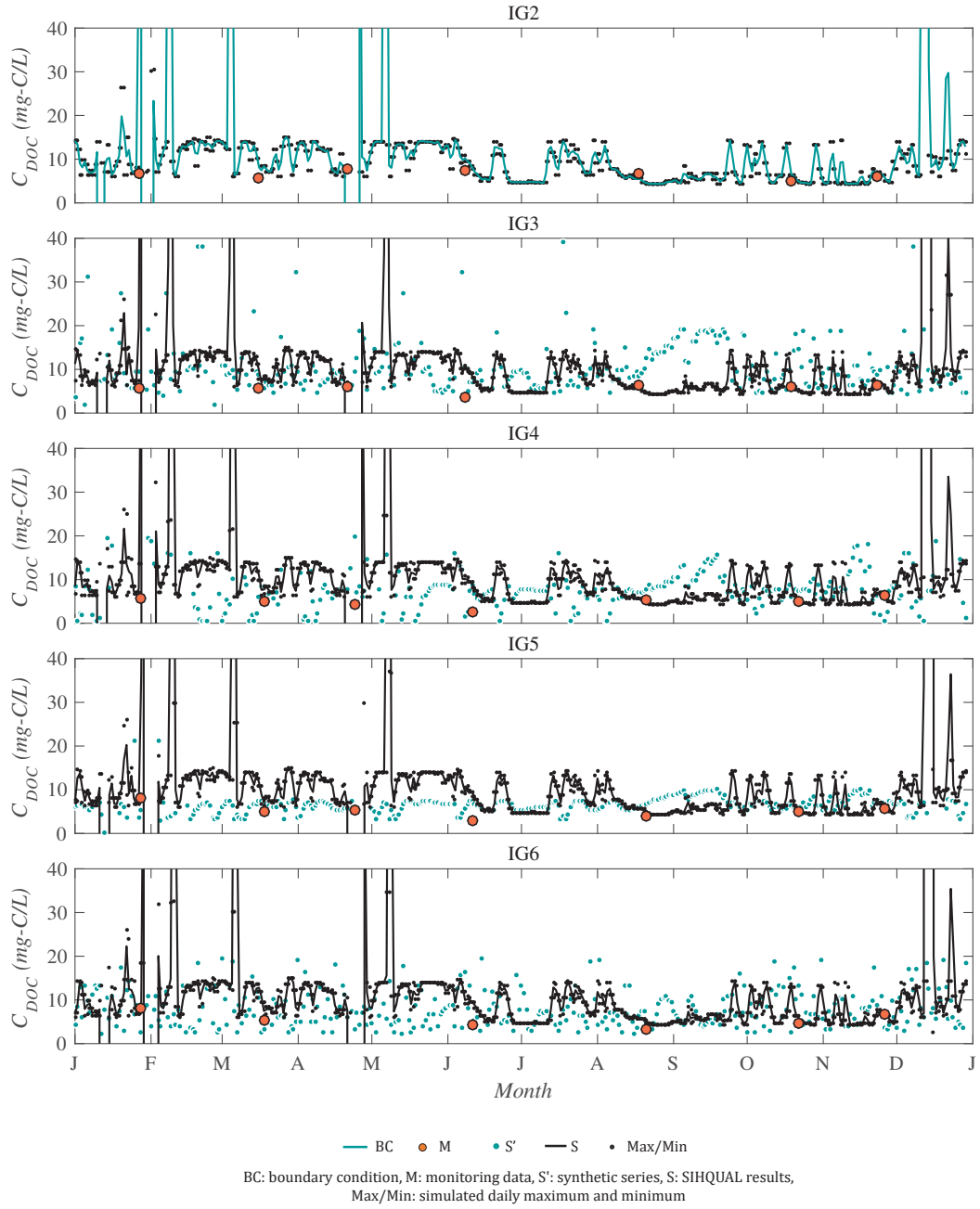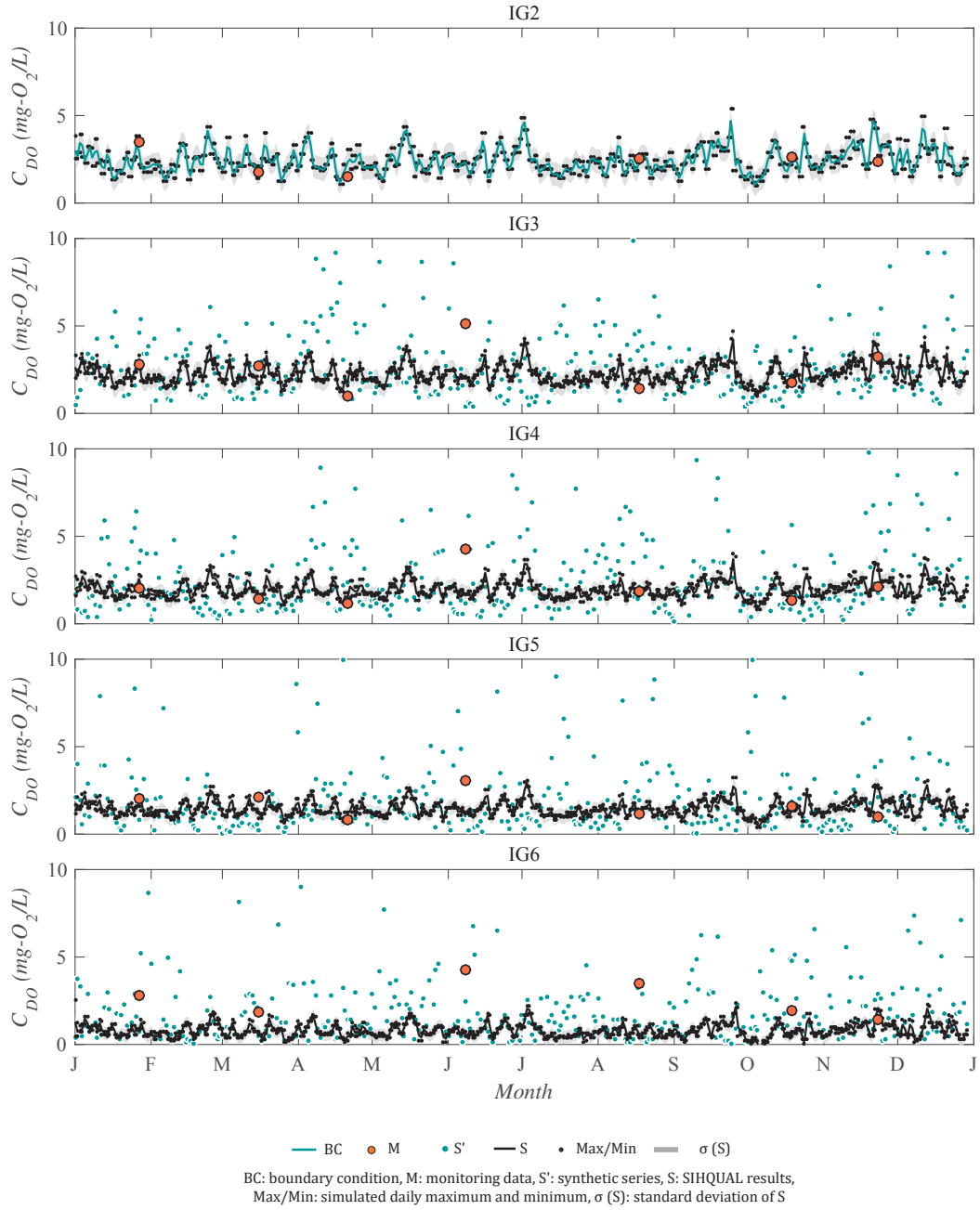
Figure 65: Daily DO simulated concentrations and monitoring data for the year 2010; BC generated with series T4b
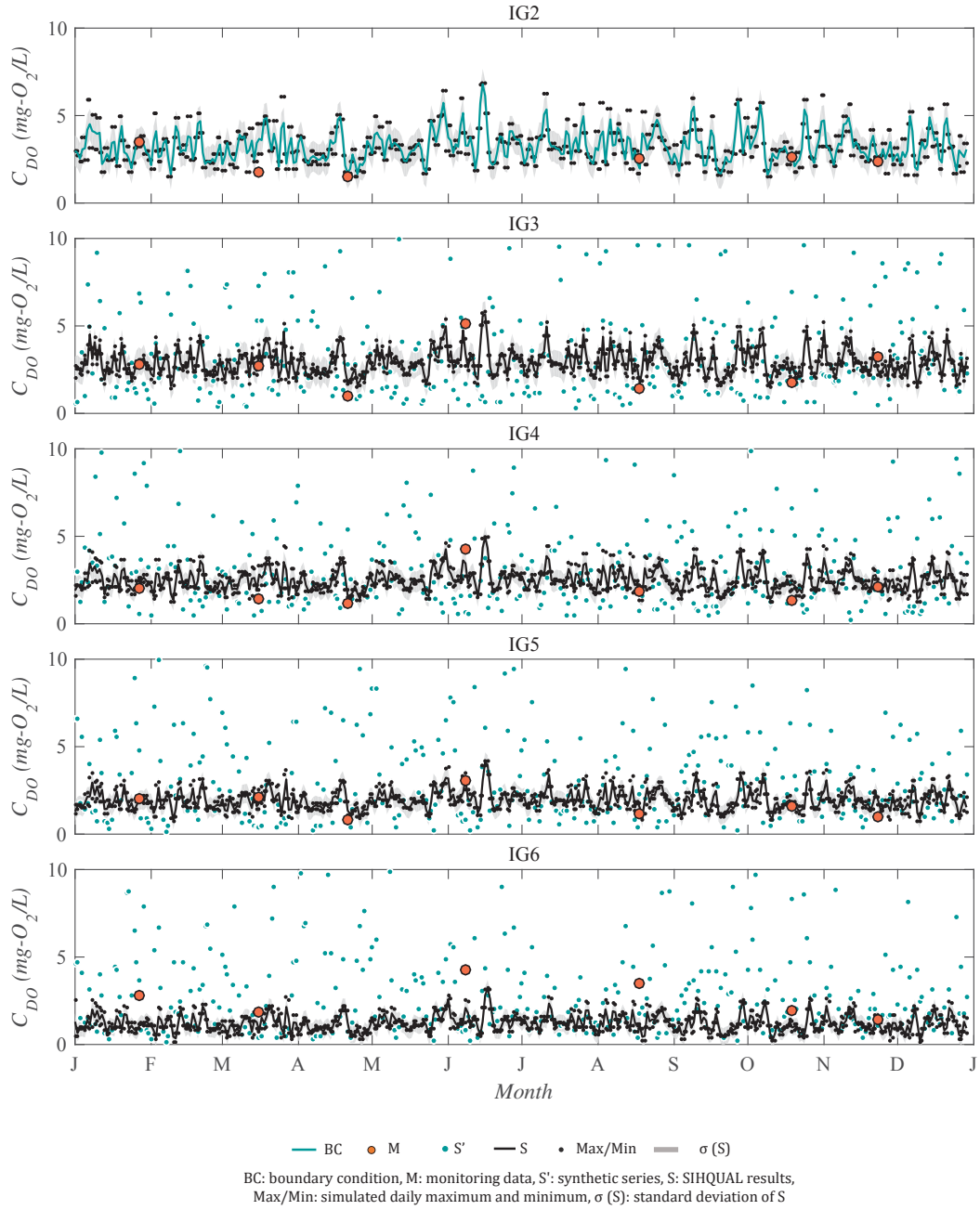
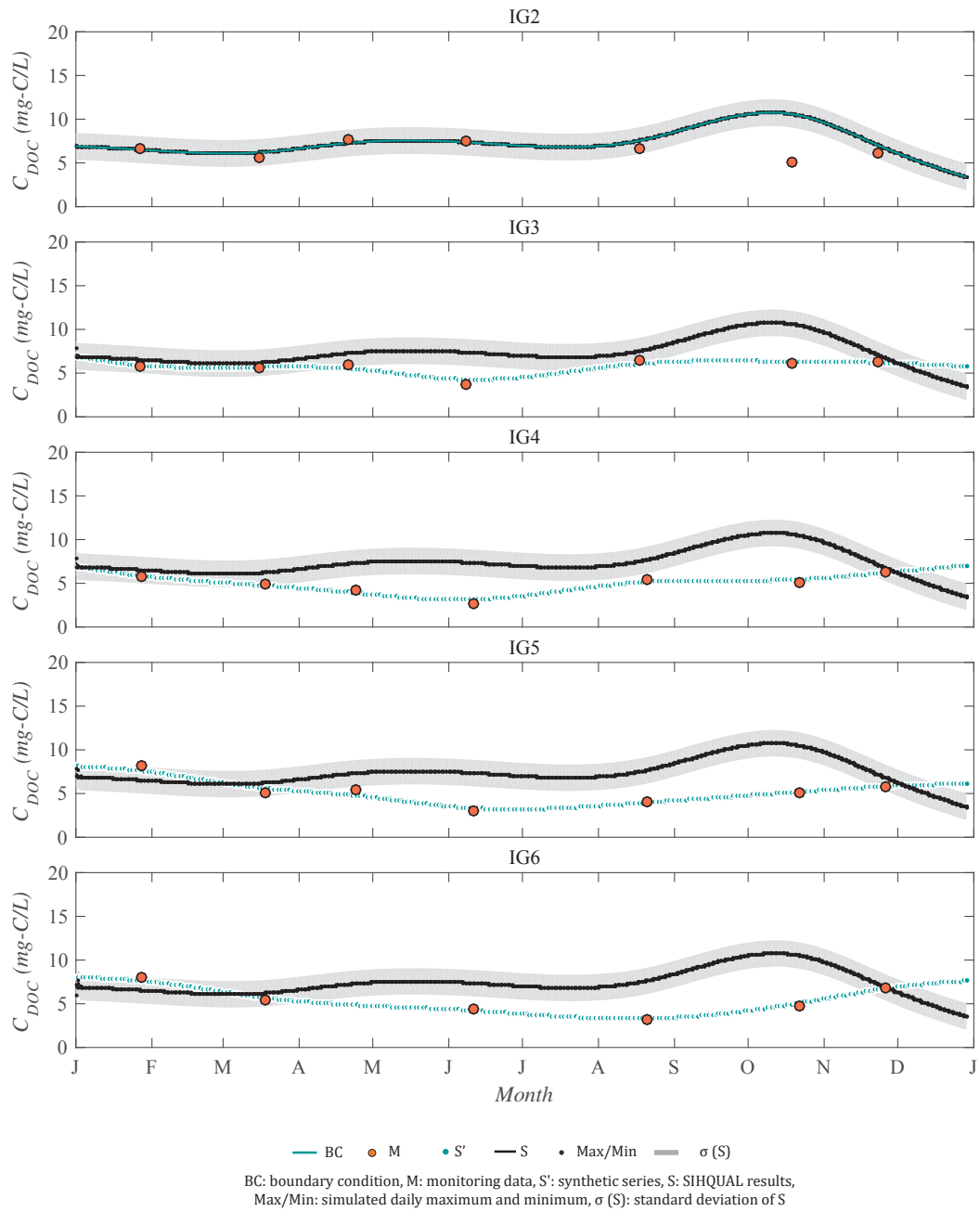Figure 66: Daily DO simulated concentrations and monitoring data for the year 2010; BC generated with series T9b
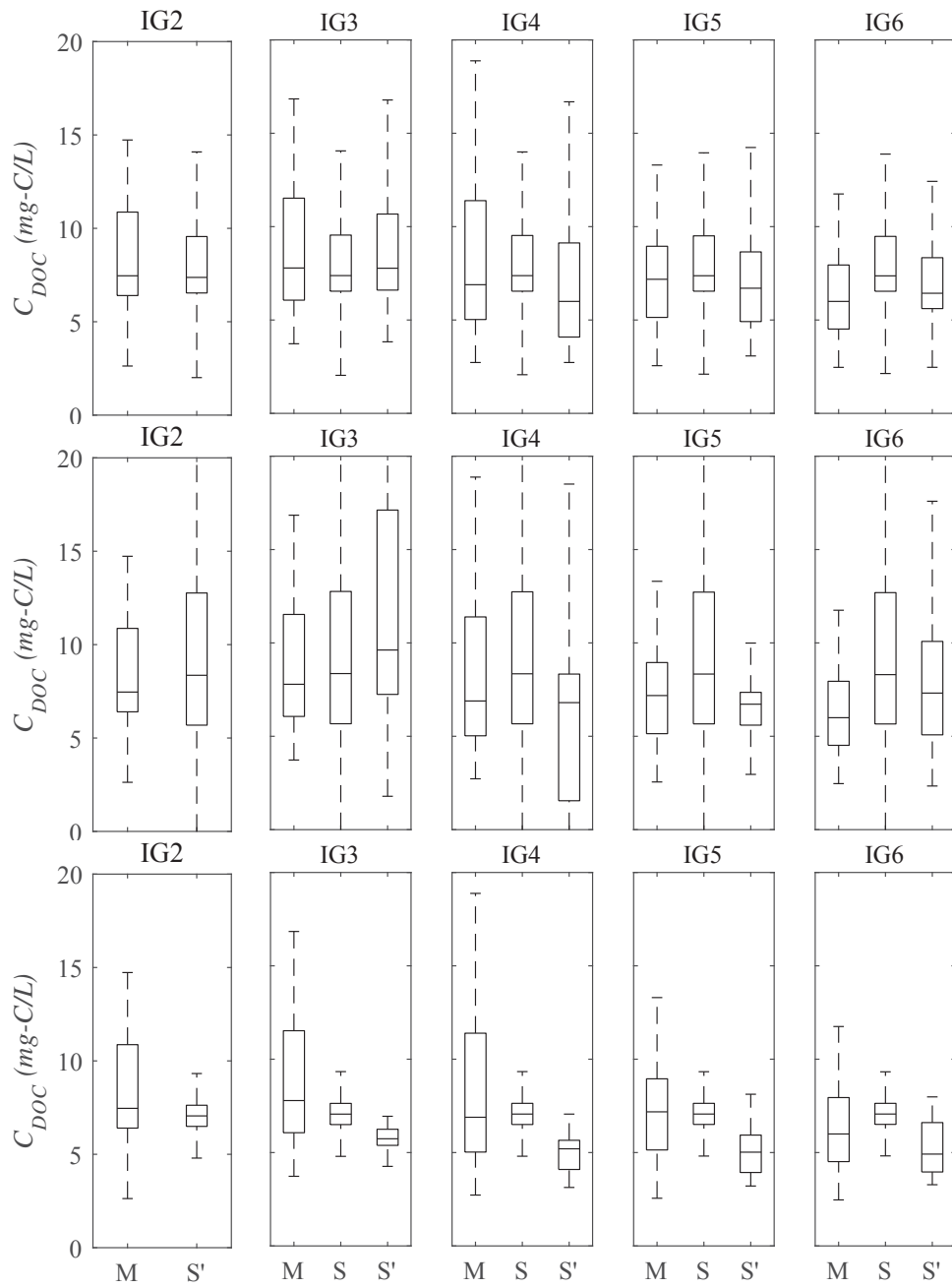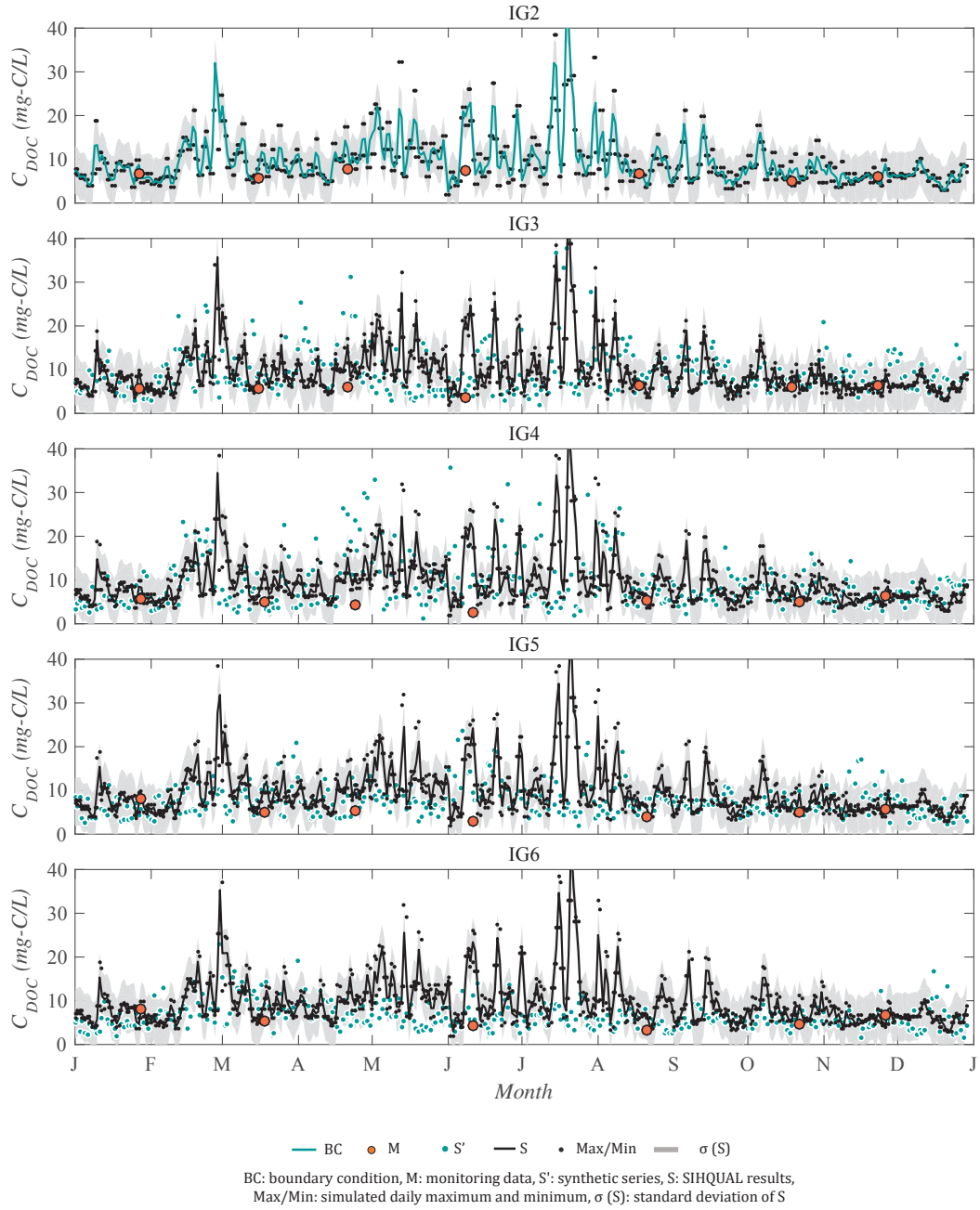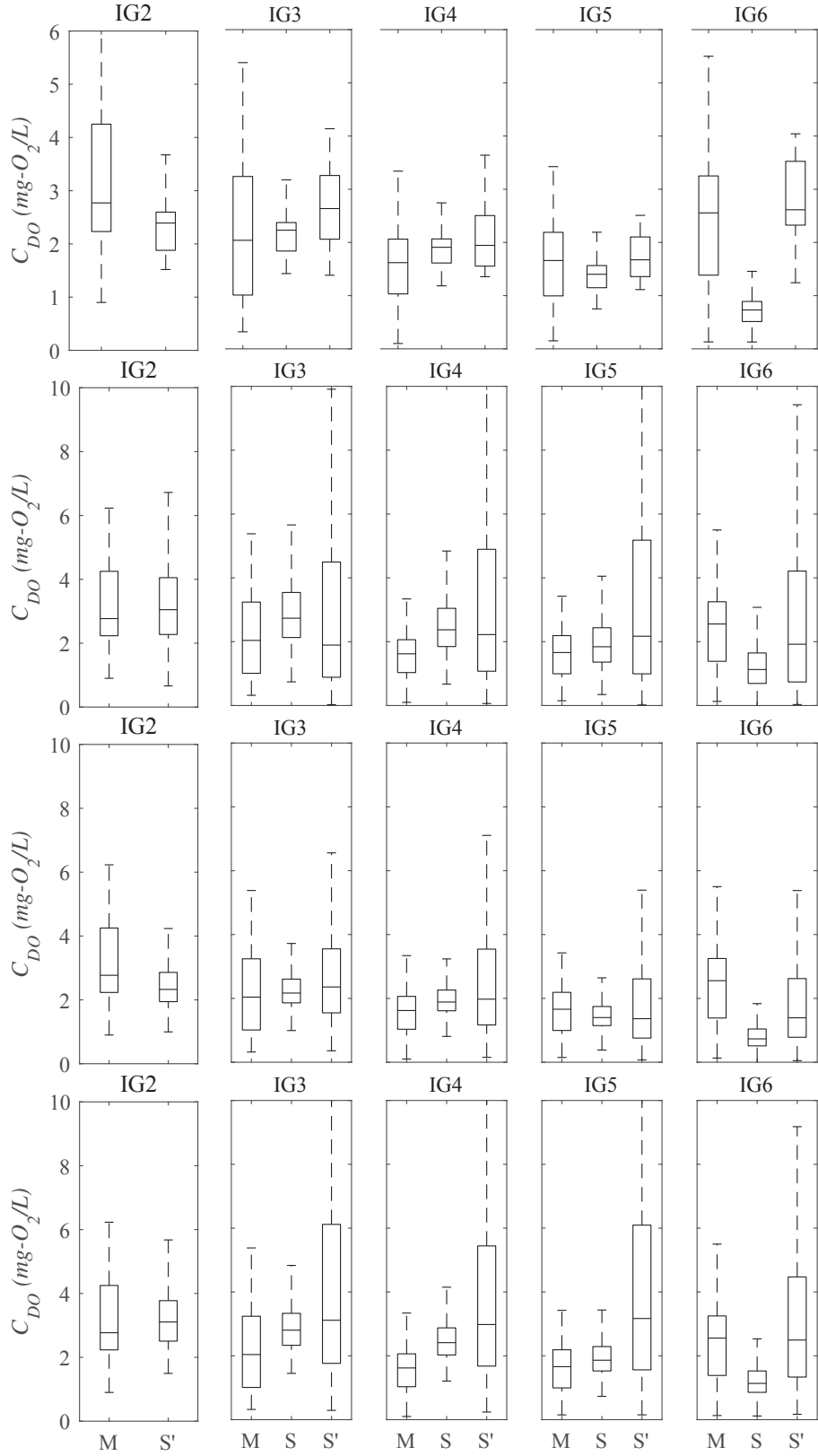
Figure 67: Boxplot of measured (M) and simulated DO concentrations for 2010 with SIHQUAL (S) and synthetic series (S'); I3, T1a, T4b, T9b, respectively in each row
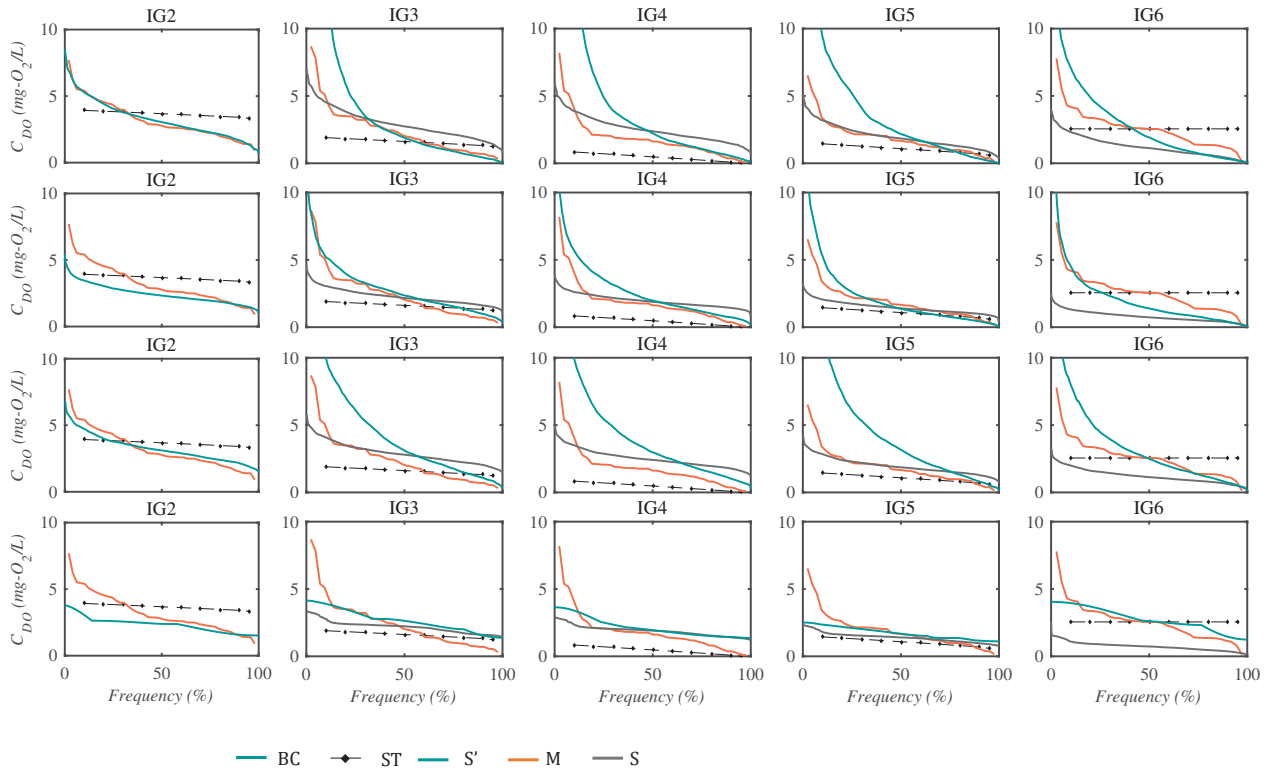
Figure 68: Duration curves of historical monitoring dataset 2005-2017 (M) and simulations for 2010 with synthetic series (S') and SIHQUAL results under unsteady (S) and steady (ST) state: T1a, T4b, T9b, I3 presented respectively in each row

5.3.3.5 Annual loads

Figure 69 presents simulated daily accumulated loads in each section of the Iguaçu river using different boundary conditions, in tons per day, as heatmaps (representation of data as a color-encoded matrix). This analysis shows where, when and for how long critical events occur.

Most of the tests indicate that higher loads pass trough the sections IG4 and IG5, or IG5 and IG6 – maximum values corresponding to the heatmaps are presented in table 16. Although downstream of the Iguaçu river has lower pollutants concentration, larger discharges culminate in higher loads, which explains this behavior.

Estimations with T9b as boundary conditions is the test indicating more events of higher accumulated load (except for N-org, since this test is the one with smaller concentration peaks – figure 51). Series with T4b as input, as expected, underestimate accumulated daily loads.
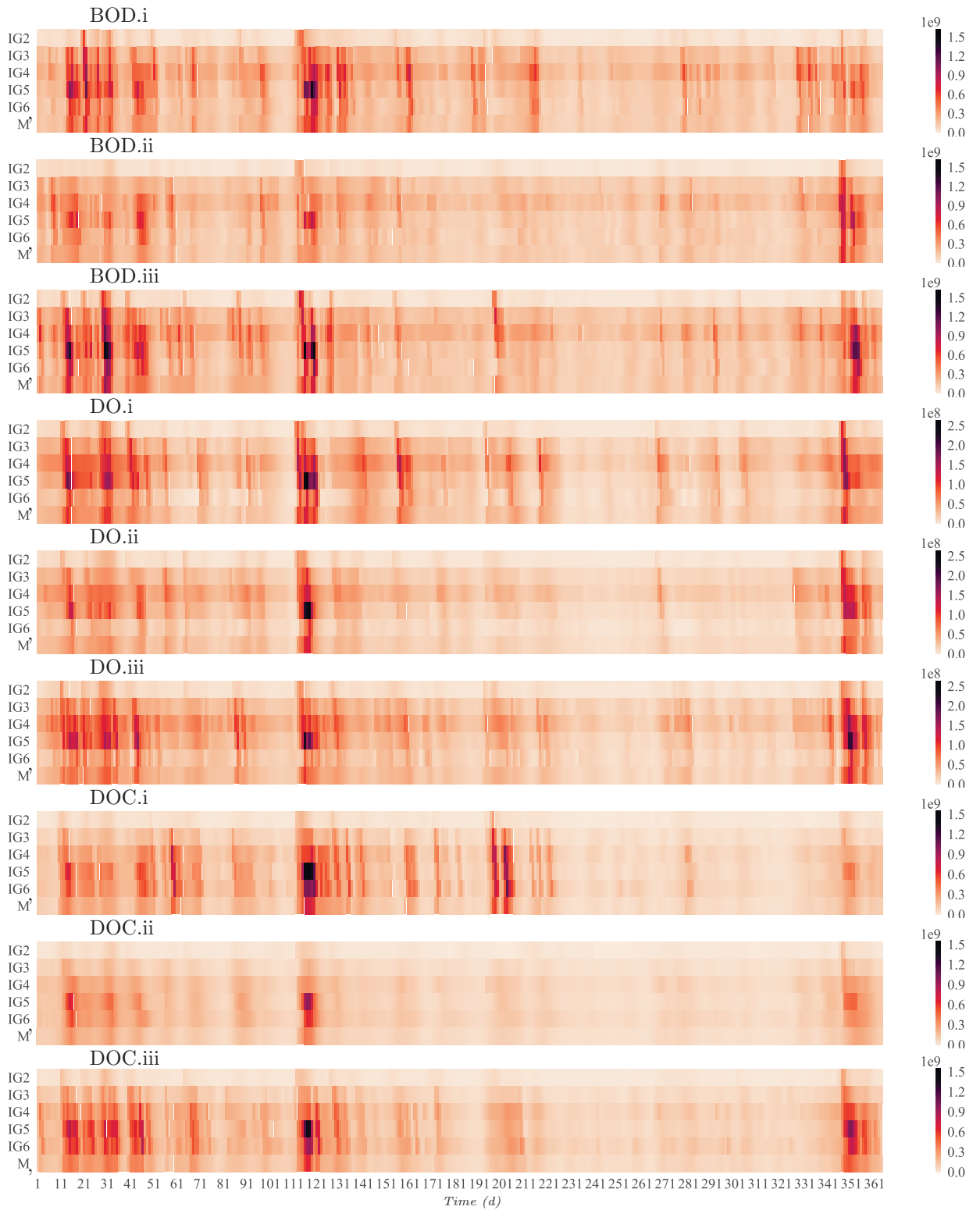
Figure 69: Heatmaps of accumulated daily loads; i: test T1a, ii: test T4b, iii: test T9b; M':
mean of IG2 to IG6
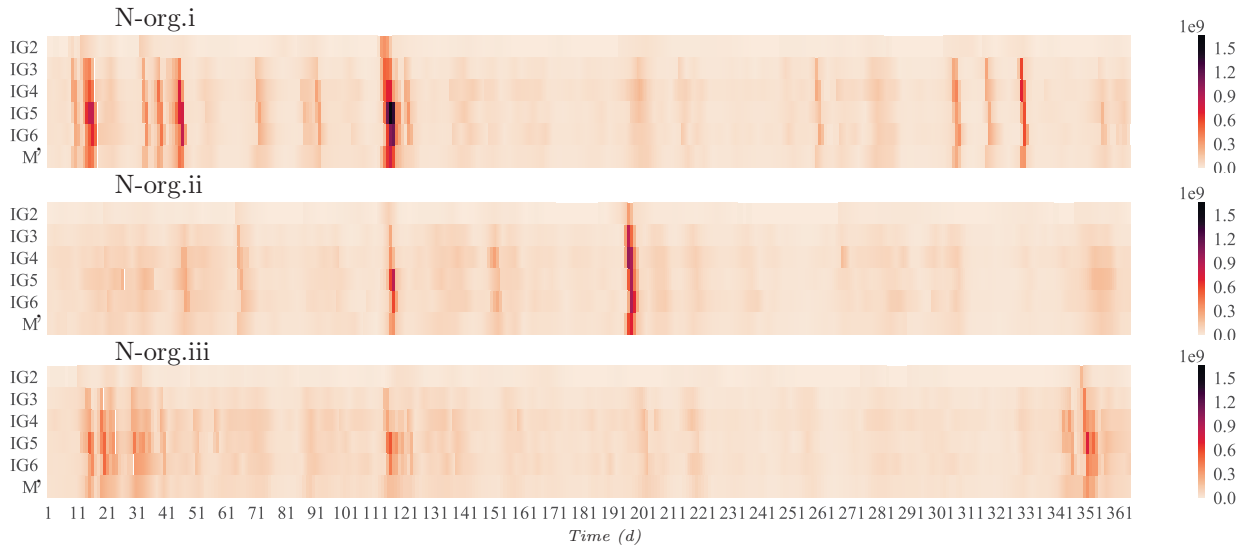
Figure 69: Heatmaps of accumulated daily loads; i: test T1a, ii: test T4b, iii: test T9b; M':
mean of IG2 to IG6 (continued)

Maximum values corresponding to the heatmaps, indicated in table 16, confirm that predictions of magnitude and timing for extreme values varies with the boundary condition, although there is agreement to the location of these scenarios (expect T4b for N-org, that indicates maximum in section IG4 instead of IG5). Even though these extreme events are quickly consumed (they last two or three days for the study case), they might cause critical conditions, such as contamination of drinking water, eutrophication and fauna death.

Table 16: Maximum accumulated daily loads estimated in 2010

Test	T1a	T4b	T9b	T1a	T4b	T9b	T1a	T4b	T9b
Parameter[1]		BOD			DOC			N-org	
Maximum[2]	1482.50	903.37	1614.50	1541.10	972.57	1430.30	1650.60	1007.30	728.60
Day	119	352	116	118	117	117	116	196	350
Section	IG5	IG5	IG5	IG5	IG5	IG5	IG5	IG4	IG5

[1]Extreme conditions for DO correspond to minimum values; because negative concentrations are generated during the deterministic simulations, due to misleading mass balance estimation, this parameter is not considered in this analysis of critical event
[2]in megatons (Mt)

A common measure for comparison in modeling studies is annual load estimation; for tests T1a, T4b and T9b, that can be interpreted as different sampling frequencies, annual loads are compared in figure 70 and table 17 – it is observed agreement between the approaches regarding location of higher/lower loads, as stated with the analysis of heatmaps.

Table 17 explores the percentage difference between annual loads estimated through

each series – T1a and T9b generated annual loads very similar, suggesting that higher sampling frequency (hourly) do not necessarily improves annual load estimation – the same conclusion was verified by Park and Engel (2014) when calculating pollutant load through regression model; the differences are higher regarding the series modeled with T4b as input – lesser data, on the other hand, change annual load estimation, as expected.
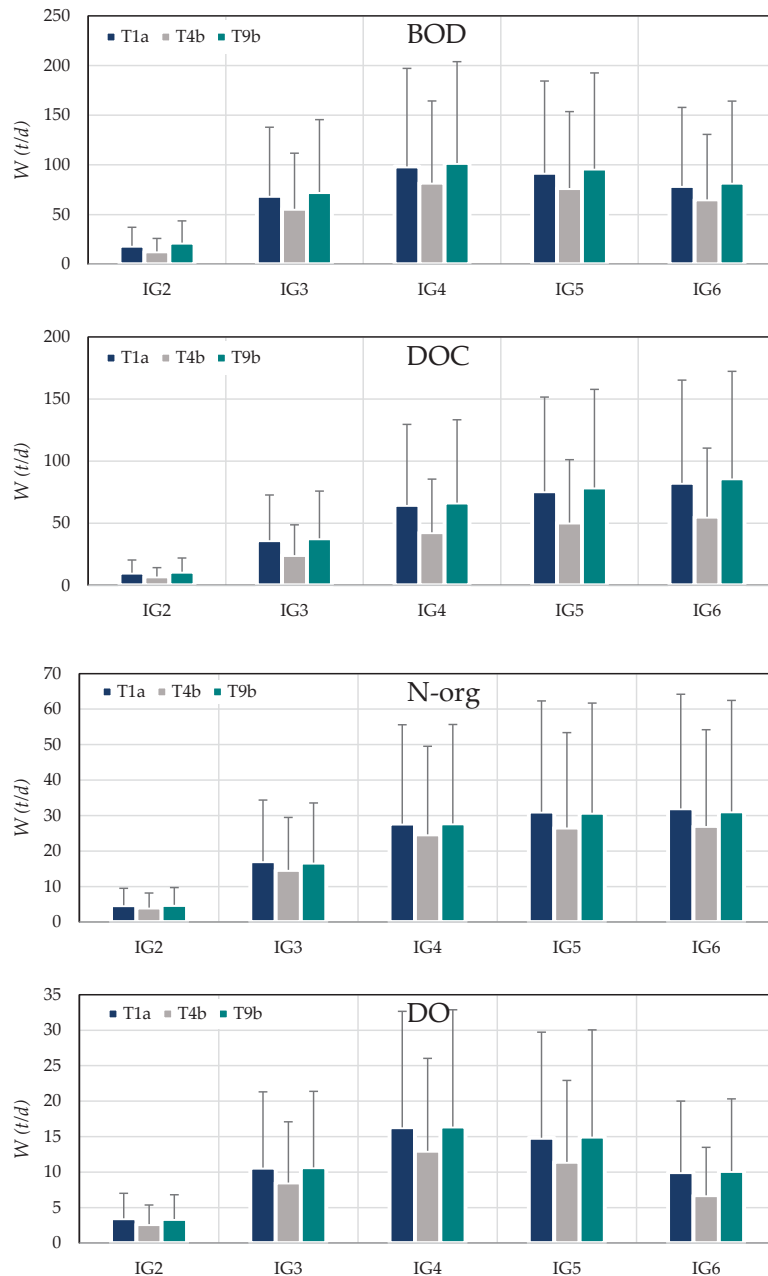
Figure 70: Annual mean loads estimated with simulations using as input series T1a, T4b and T9b; error bars indicate 95% confidence intervals (IC), IC $= \mu_W + 1.96 \times \sigma_W / \sqrt{N}$ – μ_W is mean of daily load, σ_W is standard deviation of daily load, and N is the sample size

Table 17: Percentage difference (%) of annual load estimation with simulations using as input
series T1a, T4b and T9b

	T1a/T9b	T4b/T9b	T1a/T4b	T1a/T9b	T4b/T9b	T1a/T4b
		BOD			DO	
IG2	15	40	30	-3	21	23
IG3	5	23	19	0	20	20
IG4	3	19	17	1	21	20
IG5	4	20	17	1	24	23
IG6	4	20	17	2	34	33
		N-org			DOC	
IG2	2	16	14	7	35	30
IG3	-2	12	14	4	36	33
IG4	0	11	11	3	36	34
IG5	-1	14	14	4	36	33
IG6	-3	13	16	4	36	33

5.3.3.6   Organic Carbon (OC)

Simulations under unsteady state for organic carbon are presented for the period
03/18/2013 to 03/18/2015. Total organic carbon is represented as (i) sum of POC and
DOC (figure 71), and (ii) sum of LDOC, RDOC, LPOC and RPOC (figure 72).

The calibration strategy follows the same procedure applied for other parameters;
however, for labile and refractory fractions, transformation rates variation over time is
neglected, due to processing time limitations – after generating the rates time series,
means over time are used for calculation. Boundary conditions are set as series T2a –
because POC data is available from 2012 to 2014 (ten values total), seasonal metrics are
unreliable.

Results show that, despite the uncertainty in kinetic process dynamics, the model
is able to predict reasonable estimations of organic carbon concentrations distributions
along space and time. The propensity of the ratio POC/DOC in TOC to increase under
higher flow is represented in both strategies.

For the labile and refractory fractions, there is loss of temporal variability mainly
in IG4, IG5 and IG6 (figure 72). While qualitative identification of labile and refractory
organic carbon can be achieved through uv-vis techniques, studies to quantify these frac-
tions are still incipient. Therefore, the attempt to simulate one-dimensional transport
of labile and refractory organic carbon considers arbitrary percentages: labile DOC and
POC part is defined as 70% (IG2 to IG4) and 60% (IG5 and IG6) of total DOC and
POC, while the remaining is set as refractory; although overestimated organic concen-

trations are obtained, the expected behavior from input hypothesis is reproduced by the deterministic model.
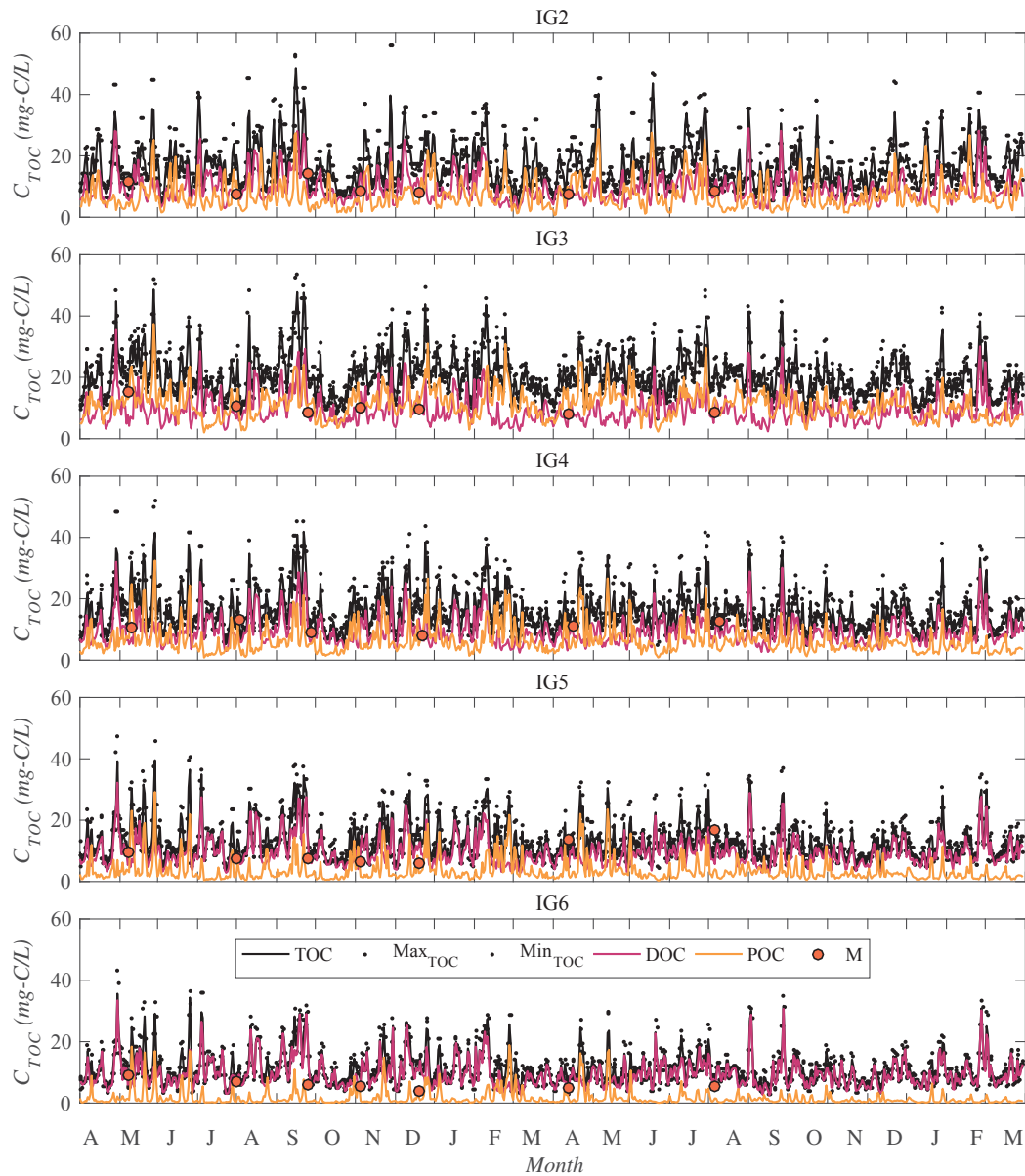
Figure 71: Simulated daily TOC concentrations (POC + DOC) and monitoring data for the years 2013-2015
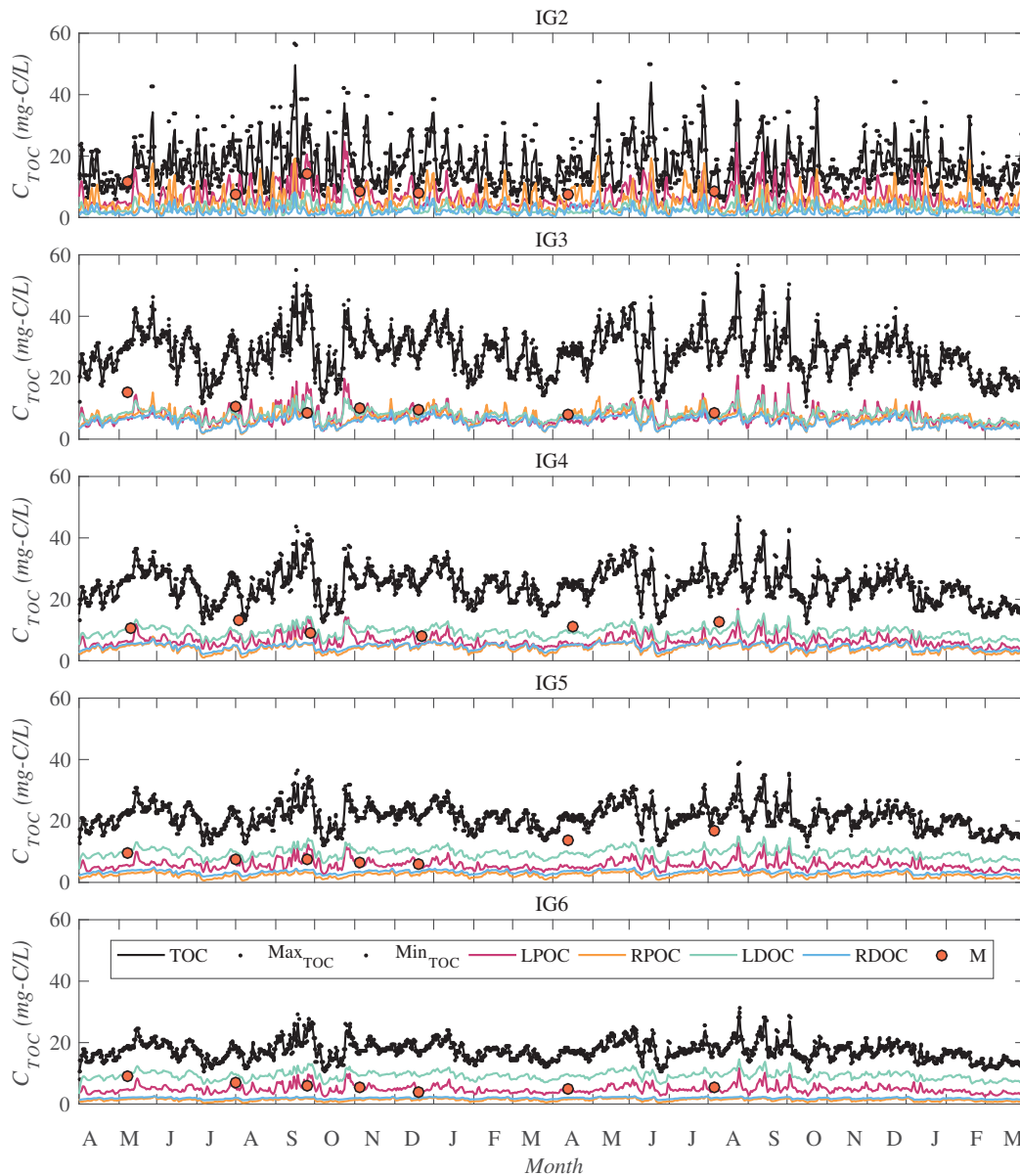
Figure 72: Simulated daily TOC concentrations (LPOC + RPOC + LDOC + RDOC) and monitoring data for the years 2013-2015

Simulation of POC and DOC time series are presented in detail in figures 73 to 76, while appendix A.14 shows LDOC, RDOC, LPOC and RPOC, besides the adjusted corresponding kinetic rates.

Overall, variability of DOC concentrations is well represented, although some overestimation is observed in sections IG5 and IG6; POC simulations, on the other hand, generated overestimated data in IG3, and underestimate downstream concentrations (IG5 and IG6) – as verified in figures 75 and 76.
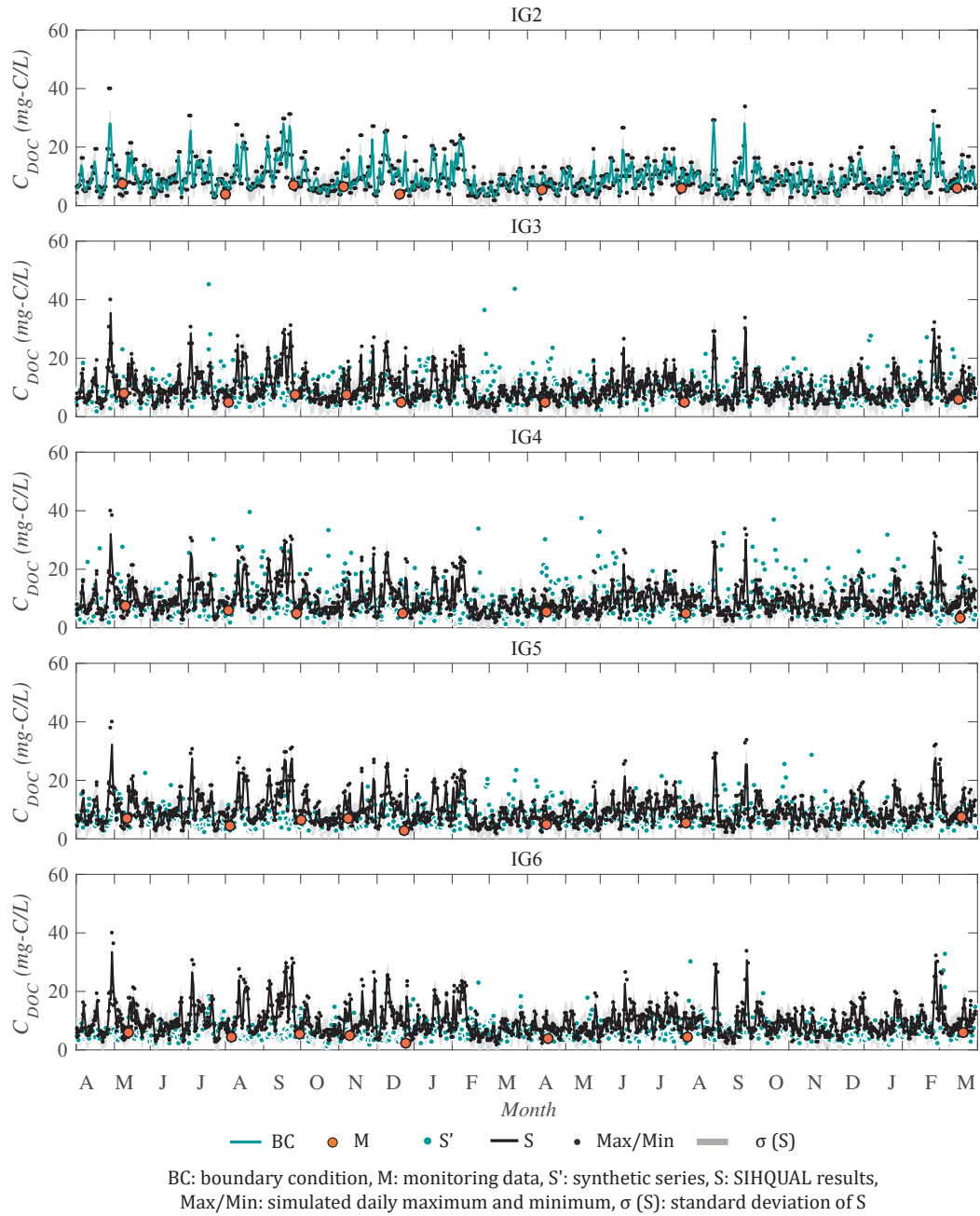
Figure 73: Daily DOC simulated concentrations and monitoring data for the years 2013-2015; BC generated with test T2a
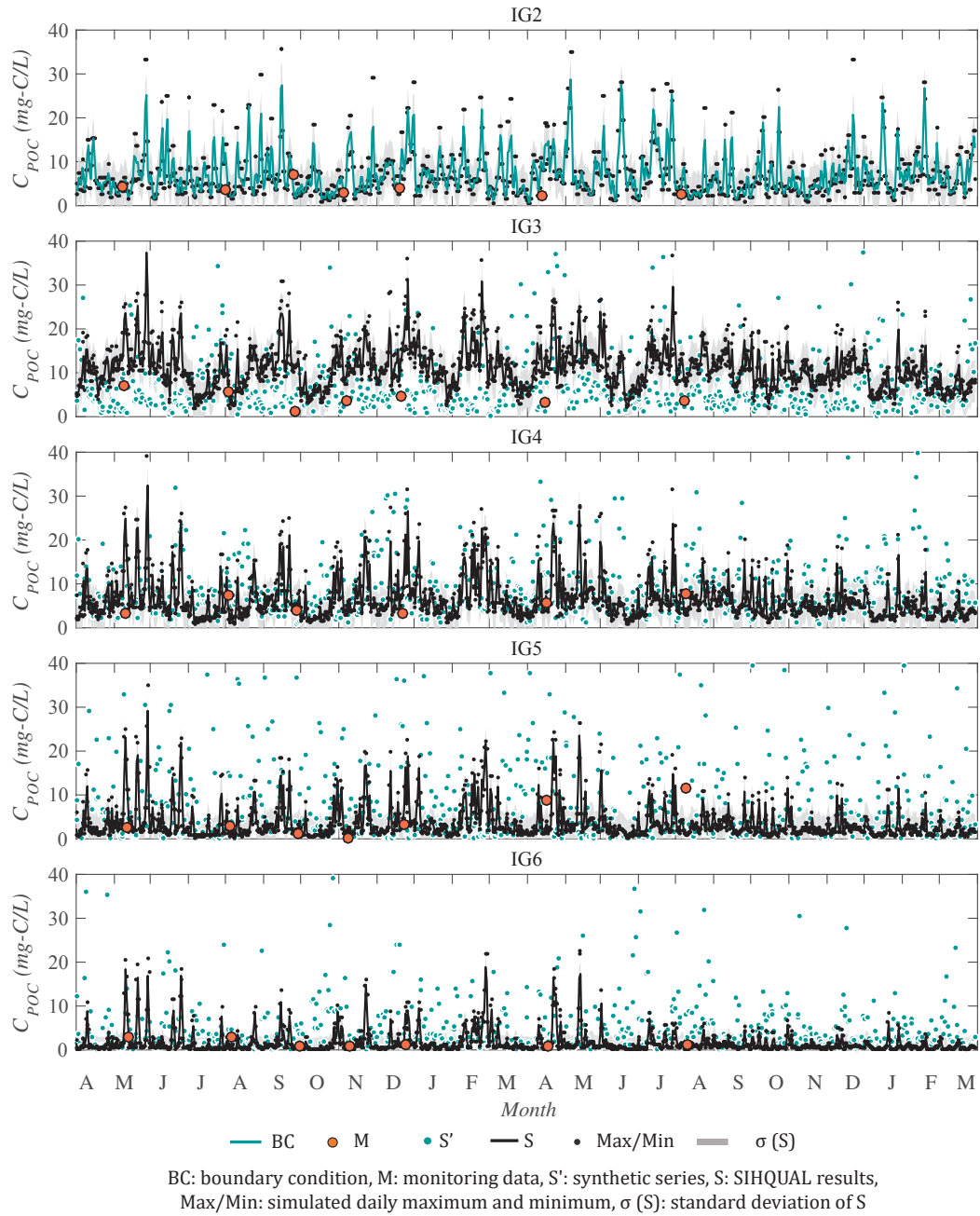
Figure 74: Daily POC simulated concentrations and monitoring data for the years 2013-2015; BC generated with test T2a
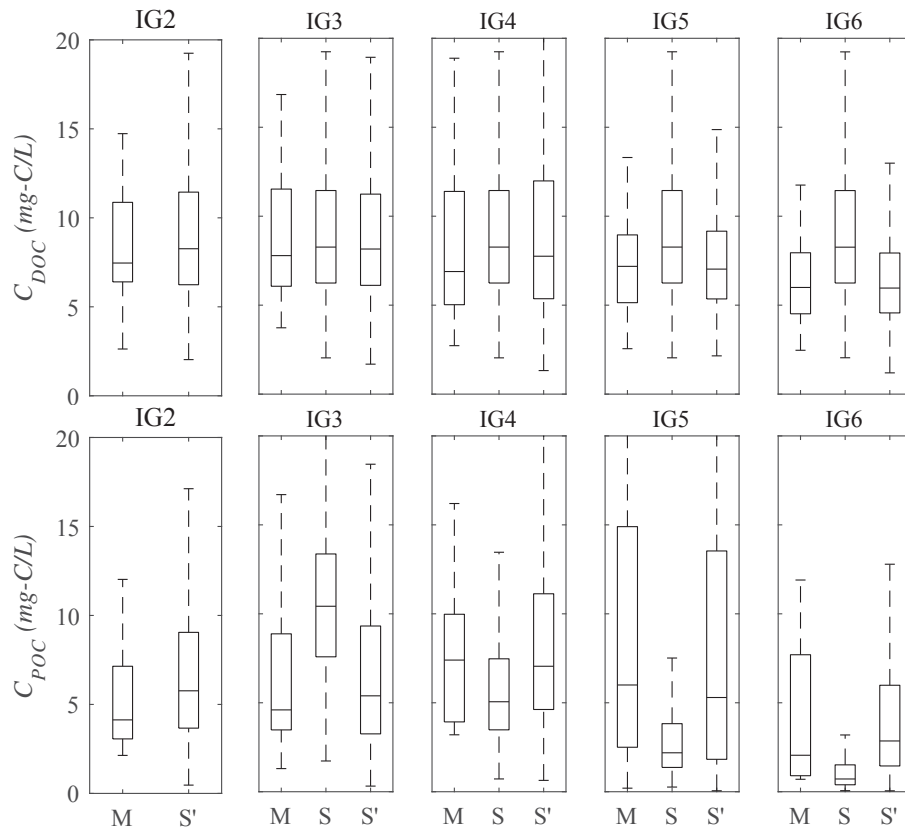
Figure 75: Boxplot of measured (M) and simulated DOC and POC concentrations for years 2013-2015 with SIHQUAL (S) and synthetic series (S'); BC as series T2a
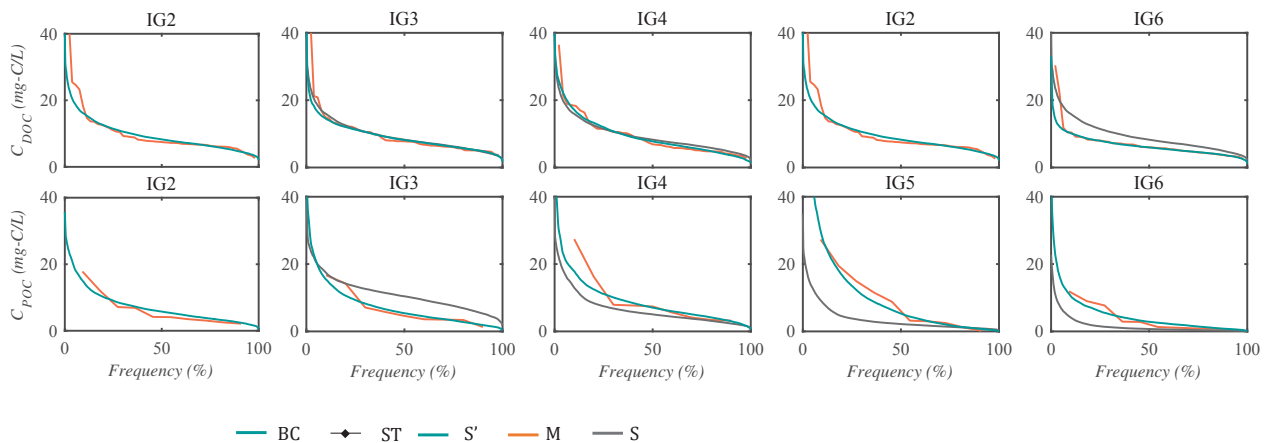
Figure 76: Duration curves of historical monitoring dataset 2005-2017 (M) and simulations for 2013-2015 with synthetic series (S') and SIHQUAL results under unsteady (S) and steady (ST) state; BC as series T2a

5.4   Boundary condition and transport of substances

Without lateral inputs, downstream concentrations are result of pollution transported from upstream through physical processes and kinetic transformations. The following test represents the propagation of a constant concentration, in order to verify how

much of it is in fact transported to downstream (table 18). Tests are conducted for the parameters BOD, N-org and DOC; concentrations in IG2 are 10 mg/L for BOD, 4mg/L N-org and 8 mg/L of DOC.

The organic matter indicated by BOD is consumed almost entirely in the 85 km reach, while organic nitrogen remains for a longer extension. Dissolved organic carbon is expected to be transported almost entirely. Figure 77 shows the spatial distribution of BOD and N-org concentrations with these simulation conditions (the values are averages over time). An additional simplified test of sensibility is presented in section A.10.

Table 18: Percentage of the transported concentration in each section

Sections	%		
	BOD	DOC	N-org
IG2	100	100	100
IG3	61	100	87
IG4	43	100	78
IG5	35	100	73
IG6	27	100	67
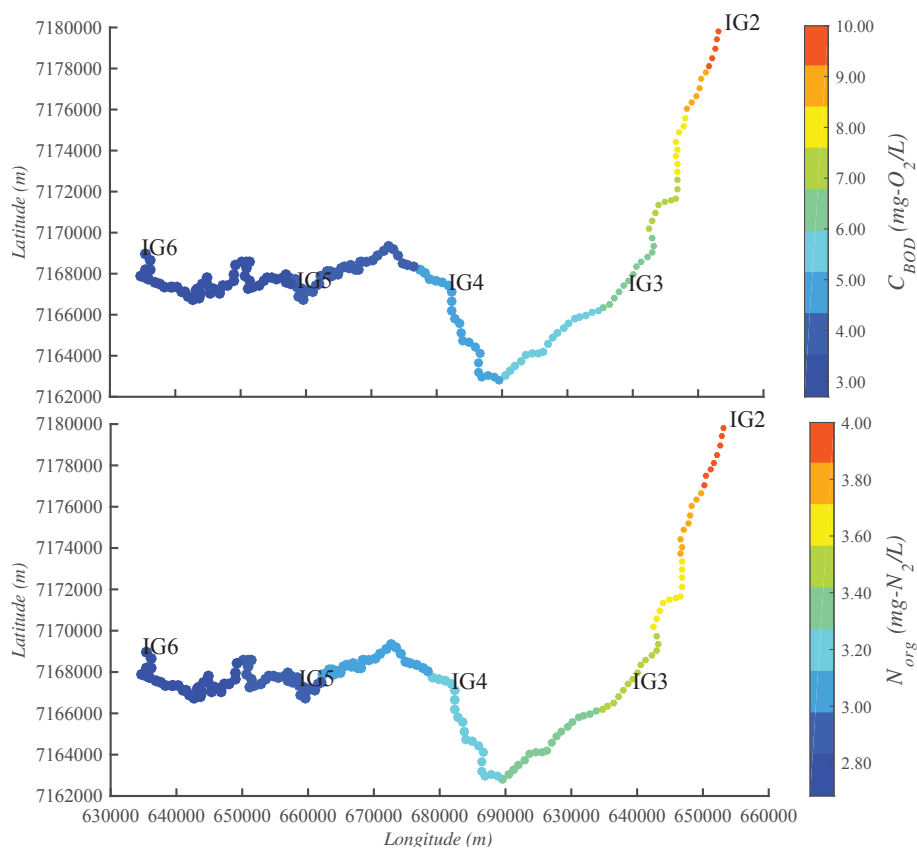
Figure 77: Spatial distribution of BOD and N-org concentrations (average over time) simulated with a constant upstream boundary condition and null lateral input; circle's size is scaled accordingly to simulated cross section top width

Even with an increase of pollution from lateral contribution, the upstream boundary condition still affects results, as verified in figure 78, especially in the identification of critical events along the year. The heatmaps compare simulation with different boundary conditions and the double of the original lateral input estimated along the Iguaçu river.
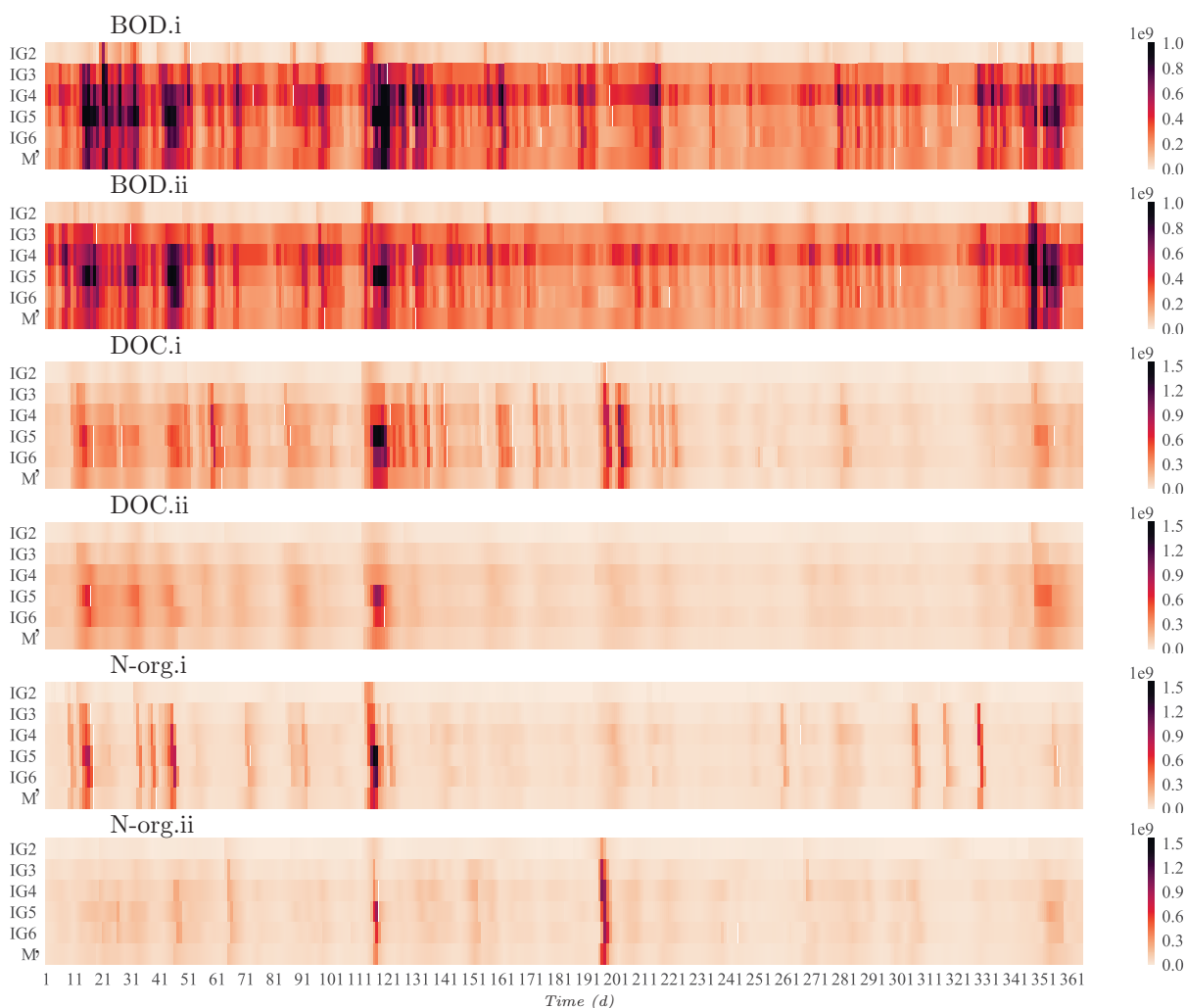
Figure 78: Heatmaps of accumulated daily concentration for BOD, DOC and N-org; boundary condition i: T1a, ii: T4b; M': mean of IG2 to IG6

5.5 Temporal variability of lateral loads and kinetic rates

This section explores the role of temporal variability in lateral inputs and kinetic processes. Experiment A: considering the case with daily concentrations as hourly samples (test T9b) being released at IG2, test (i) compares the original results (presented in section 5.3.3), (ii) represents simulation with kinetic coefficients constant in time, and (iii) is a simulation scenario with both transformation rates and lateral inputs not varying over time. Figures 79, 80, 81 and 82 show the results in terms of boxplots and annual loads estimations (respective pollutographs also support the discussion, and are presented in appendix A.15).
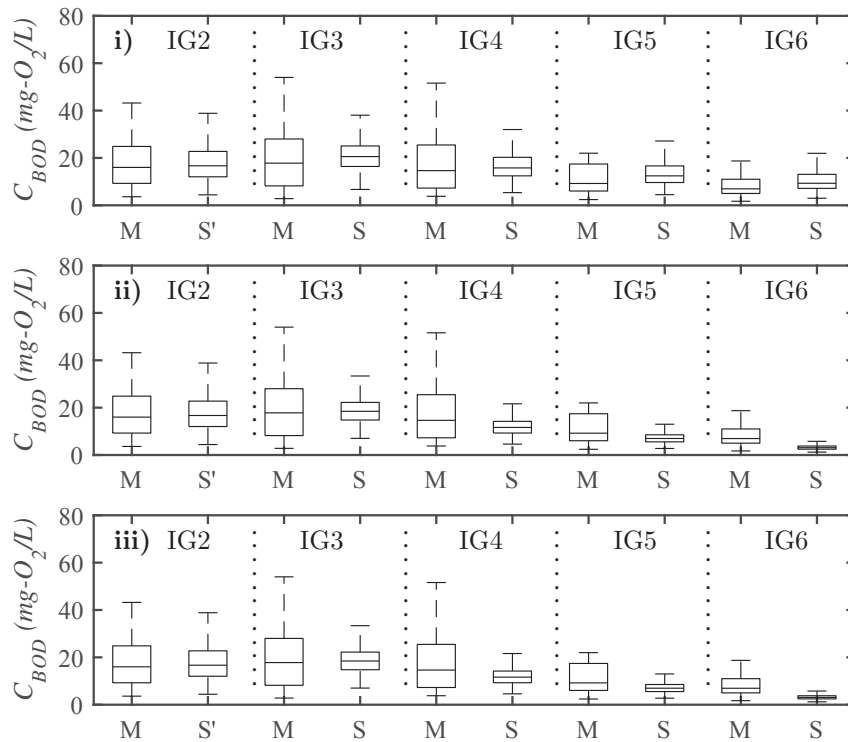
Figure 79: BOD simulations with T9b as boundary condition (S) and monitoring data (M) - boxplots
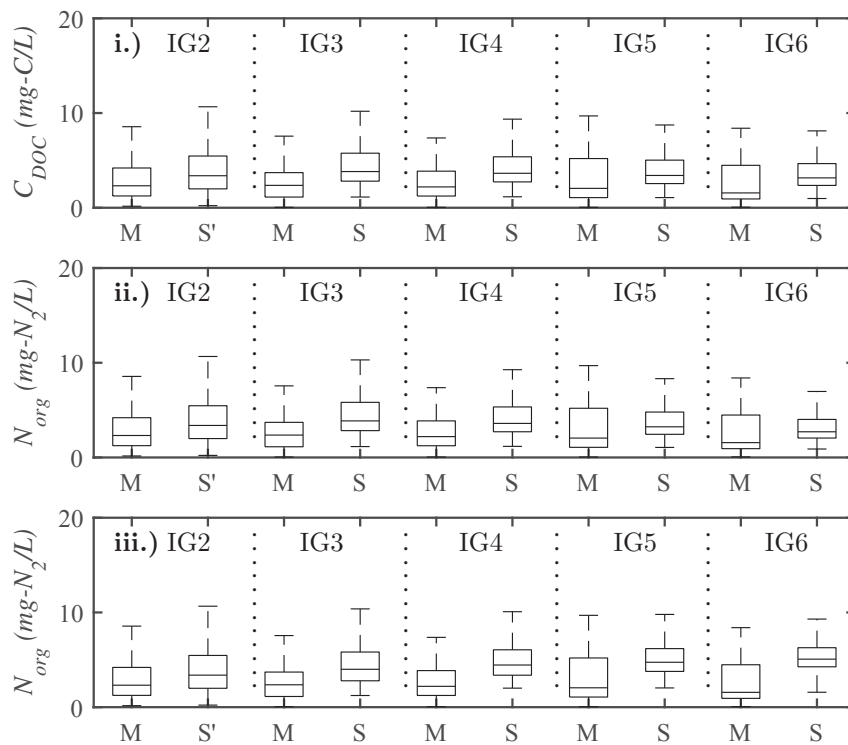
Figure 80: N-org simulations with T9b as boundary condition (S) and monitoring data (M) - boxplots
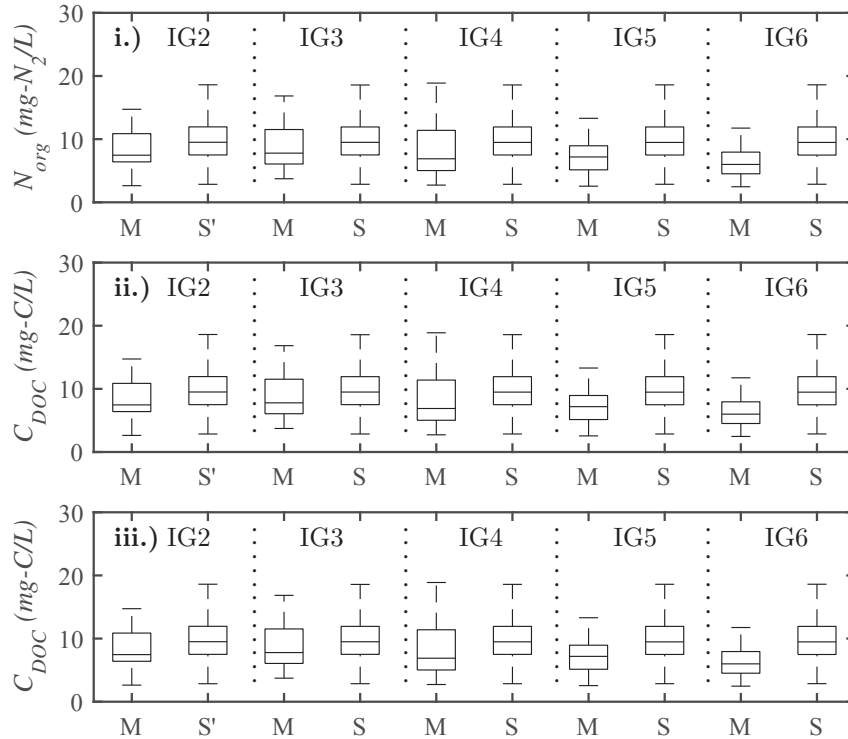
Figure 81: DOC simulations with T9b as boundary condition (S) and monitoring data (M) - boxplots

The comparison of these tests shows that temporal variation of kinetic rates and lateral contribution has an important role in the overall variability of BOD concentrations – N-org also has some variation, but DOC does not show significantly sensitiveness. Disregarding the temporal variability of calibration parameters lead to higher attenuation of pollutant's distribution over time, so boxplots have smaller interquartile intervals (79 ii); test (iii) shows that this behavior is highlighted if dilution effects on lateral inputs are simplified (79 iii), and more likely generate overestimated concentrations.

The same behavior is verified for annual mean loads estimation (figure 82). Results suggest that loss of temporal variability in transformation rates generates smaller annual loads; on the other hand, when lateral input is constant over time, overestimated values are calculated by the model.
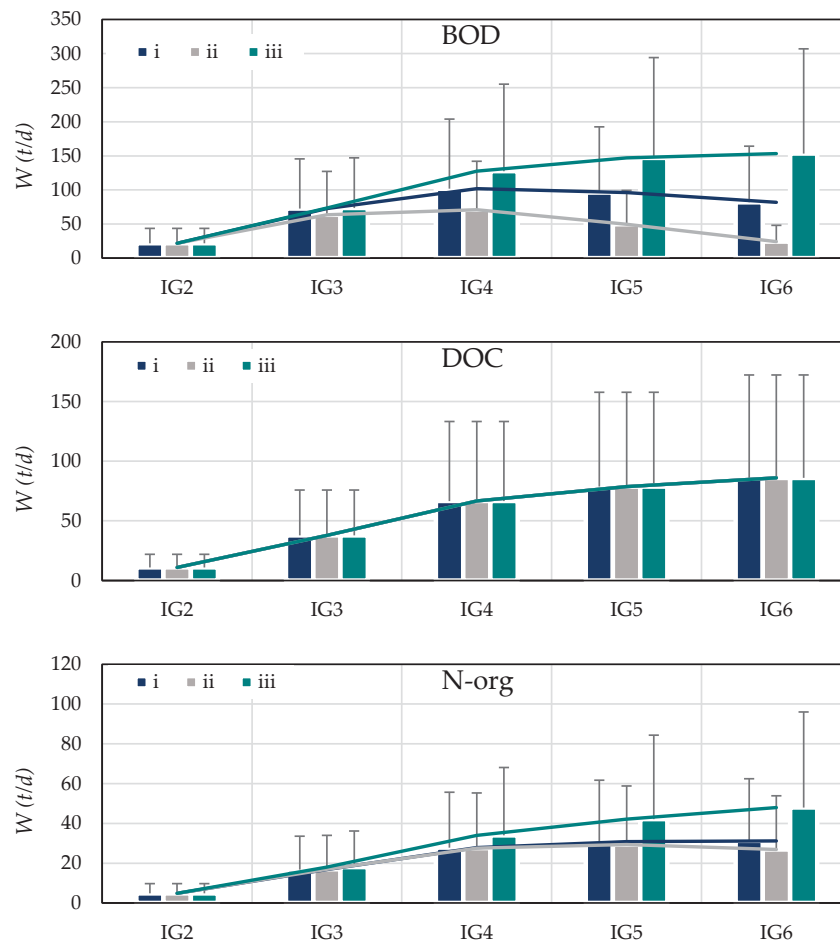
Figure 82: Annual mean loads estimated with simulations exploring temporal variability of lateral loads and kinetic rates

The experiment B is conducted only for BOD – same conditions as the previous ones, but with boundary condition as series T10a; the objective is to show that, even with a misleading upstream boundary condition, results are improved when the unsteady behavior is represented in the calibration phase and in the dilution effects of input lateral loads; when this aspects are neglected, the model is able to reproduce only median values, since natural persistence over time is not being respected.
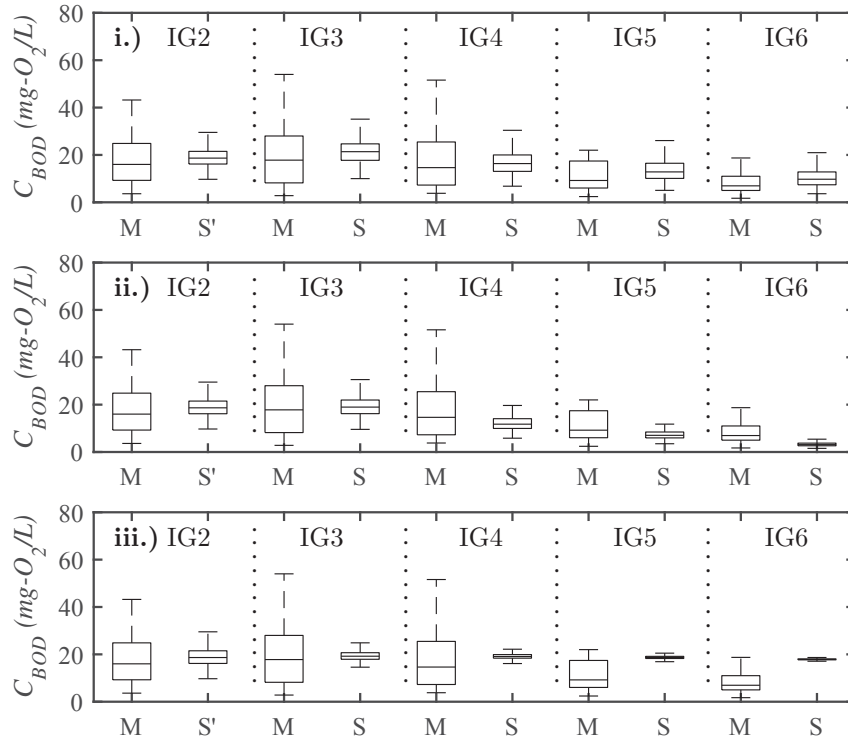
Figure 83: BOD simulations with T10a as boundary condition (S) and monitoring data (M) - boxplots

The importance of temporal variability in kinetic rates and lateral contributions – observed mostly for BOD – can be summarized by the following statements:

1) comparison of tests i, ii and iii: boundary condition with higher daily variability (series T9b with input as daily data from hourly samples); variability of boundary condition is not enough to guarantee variation in downstream sections without variability in kinetic rates and lateral inputs.

2) test iv: boundary condition with small daily variability (series T10a, data every 50s), produces reasonable results due to the developed calibration procedure and due to lateral inputs temporal variation; without these aspects, overall variability is not well represented – this means that, if boundary condition does not respect persistence over time, calibration with temporal variation is even more important.

5.6   Summary

This thesis proposes a complementary tool for river transport analysis, that integrates deterministic and statistical/empirical simulations. The first one refers to the traditional solution of the Saint-Venant equations and advection-dispersion-reaction, while the second approach explores simplified methods of synthetic series generation – this module aims to meet the temporal scales of boundary conditions, since data of water quality are generally not obtained as regular samples. This tool offers steady and unsteady state

simulations for the parameters biochemical demand for oxygen, dissolved oxygen and organic nitrogen, besides conceptual formulations for non-traditional constituents (dissolved, particulate, refractory and labile organic carbon). In addition, the calibration method presents an alternative to the traditional procedures of trial and error and optimization techniques, highlighting the transient approach in the determination of kinetic parameters.

Figure 84 presents a summary of this development: i) hydrodynamic simulations offer flow conditions analysis, while ii) convert discrete ("snapshots") into continuous information as water quality input; the latter is also basis for the iii) calibration procedure, which provides transformation rates for the iv) water quality module. Final results are discharges and concentration over time and space in rivers, granting additional information to those generated by conventional analysis (monitoring and steady state evaluations). In the interest of water resources planning and management, this tool also provides a risk analysis (duration curves), that incorporates system responses to multiple conditions (flow, dynamics of lateral inputs and transformation rates etc).
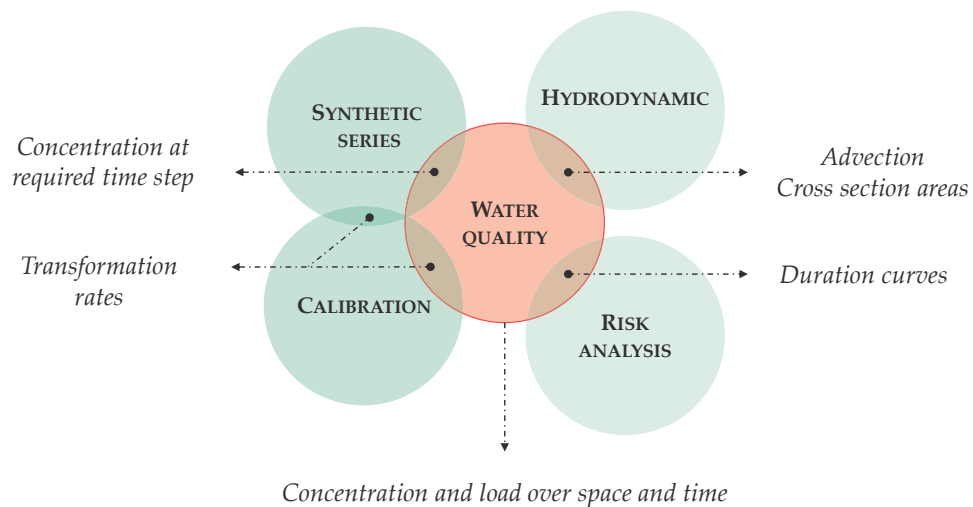
Figure 84: Modules and their link in the SIHQUAL model

6   Conclusions

The goal is to turn data into information, and information into insight.

– Carly Fiorina

For water resources planning and management strategies, the concept of mathematical modeling allows integration of quantity (flow) and quality (concentrations). Input data, however, is a fundamental aspect to interpret results in this type of study. This research provides new contributions aiming to decrease the risk of inadequate decisions in water resources planning and management based on numerical simulations, integrating synthetic pollutographs and deterministic modeling. The proposed methodology aims to support more reliable analysis with available resources, considering that gather appropriate information at minimum cost remains a challenge (Hankin et al., 2016).

Empirical techniques are useful because they combine the many factors responsible for uncertainty in time series (missing data, non-linearity, seasonal and cyclical patterns etc), and demand a minimum description of cause and effect; the historical monitoring dataset provides the required knowledge about the system. The posterior phase, that propagates the information through deterministic equations, incorporates the different processes controlling release, transport and fate of pollutants in rivers. Consequently, strategies for water resources planning and management can be evaluated.

A few techniques to predict time series of concentrations at regular time steps for a continuous period are evaluated, exploring: influence of seasonal variation, effects of temporal interval generation, persistence scenarios, sampling frequency versus time series prediction, association of water quality with flow conditions and time.

Results suggest several important messages. Interpolation with PCHIP function and Fourier series, although accounting for flow conditions, are sensitive to extreme values and did not describe all parameters tested; the function smoothing spline, due to the limited data, generate gradual variation over time, not being able to represent the unsteady behavior. The experiments based on autoregressive principles showed that preservation of statistical metrics from historical information, linked to multiple scenarios analysis and representation of natural persistence, are reasonable criterion to estimate water quality time series; the methodology showed to be effective for all water quality parameters investigated, besides other study cases.

Overall, the experiments with synthetic series demonstrated consistent reproduction of water quality variability or characteristic measures, such as dataset quantiles; when integrated with the deterministic model, as long as the structure of natural persistence over time is consistent, their propagation generate reasonable concentrations over time and space.

The synthetic series also provide plausible different scenarios, allowing a conceptual

analysis of input conditions in deterministic modeling. Although not usually explored in studies of water quality under unsteady state, upstream boundary condition may have an important role in model outputs, and therefore impact decisions in water resources management and planning; this question was addressed by drawing attention to the differences in duration curves and load estimates resulting from multiple input conditions. The analysis suggest that the upstream boundary condition may be responsible for misleading identification mainly of critical events magnitude and estimates of when the same occur. Responses to different inputs were stronger in sections that receive larger amounts of lateral loads, and therefore where water quality have higher temporal variability.

Among the studied water quality parameters, BOD showed to be the constituent most affected by upstream boundary conditions, while organic nitrogen and dissolved oxygen had less impact (since these latter have less overall variability). Dissolved organic carbon also showed to be largely affected by upstream boundary conditions; because of its particular behavior (attenuation over space is very low, and almost all particles from upstream reach downstream), variability in input condition controls nearly all distribution in subsequent sections; this is also supported by the test of section 5.5, since this parameter showed to be almost insensitive to temporal variation of lateral inputs and kinetic rates. However, further analysis are suggested, since the parametrization used in this research was based on a simplified model of kinetic processes.

In this context and to reinforce the model ability in predict different water quality constituents, simulations for non-traditional parameters are investigated (organic carbon fractions dissolved and particulate, labile and refractory). Results highlight that representation of internal processes and interaction between components remain one of the main challenges in transport modeling of non-conservative substances.

In the same context, the proposed method for integrated modeling suggests that temporal variation of kinetic processes may play an import role in transport and fate of pollutants in terms of overall variability and attenuation over time/space, and this aspect can be included in the calibration phase with the proposed method.

The current development of technologies in industry and agriculture, as well as the dynamic in land use and occupation, affect directly the list of pollutants that can reach watercourses. Therefore, the number of parameters indicating water quality is still open, depending on the discovery of new compounds and improvement of analytical methods.

The integrated analysis in SIHQUAL shows a systematic and consistent behavior in comparison with a natural system, being able to integrate and represent multiple scenarios of flow, input conditions and data availability. In such a context, this research provides background to evaluate the risk of disagreement with quality standards, complementing the information granted by monitoring programs and steady assessment.

6.1 Contributions

The contribution of this research is a new approach to evaluate the variation of discharges and concentrations in rivers through one-dimensional modeling. This strategy combine statistical approaches to traditional deterministic analysis, and it can be used as tool for orientation of stakeholders and water resources committees, as complement to traditional mechanisms, such as monitoring and analysis in steady state.

Specific additional original contributions of the research are:

- Fast and parsimonious strategy to generate environmental time series in high temporal resolution: autoregressive model and multiple scenarios analysis/selection allow to: i. overcome lack of data, ii. combine the factors responsible for uncertainty in time series, iii. demand a minimum description of cause and effect – historical monitoring dataset provides the required knowledge about the system, iv. useful to predict results with different input conditions.

- Unique calibration method for modeling water quality under unsteady state, based on system characteristics (that allow to overcome a common issue in calibration – virtual values that solve the mathematical problem, but without physical meaning), random variation (that incorporate the inherent uncertainty), and temporal variation (through a link that relates transformation rates to intervals of concentration); this is different from the traditional association of kinetic rates to empirical equations and other conditions (hydraulic characteristics or composition of wastewater, for example); the procedure also overcome the usual large processing time required in automatic techniques, and it less subjective than traditional trial and error analysis.

- First attempt to simulate non-traditional water quality parameters (organic carbon and its fractions) in rivers under unsteady state.

6.2 Future perspectives

The integrated modeling is robust to predict discharges and overall variability of water quality, which is sufficient for exploratory and planning purposes; however, for regulatory and legal objectives, such as analysis of levels transgression over time, some further improvements might be required.

Better predictions of critical events and actually temporal variation of water quality will be possible with more information, such as lateral input dynamics (schedule of wastewater releases and representation of diffuse pollution linked to hydrological process) and knowledge about kinetic processes (that also depend on water-sediment and water-atmosphere interactions).

Further recommendations for future efforts are discussed in the following items:

i Lateral load inputs: sensibility and temporal variability tests confirmed that lateral contribution is an important aspect in modeling river contaminant transport. Diffuse sources were represented by a simplified procedure, disregarding temporal variability; because this input is directly related to meteorological conditions – that can rapidly change in tropical regions, hydrological models could be associated. For point sources representation, including the location of wastewater treatment plants and tributaries might improve simulated results, especially the estimation of critical events: domestic and industrial waste are usually released during defined periods of the day; therefore, these inputs are characterized by a curve with high peak that decreases until a steady behavior; since the model is not representing this fluctuation, overall results underestimated critical concentrations. Additional characterization of sources can also serve to improve simulations in assisting the identification of proper interactions that should be represented; some industrial waste components, for example, are more susceptible to volatilization or to interact with sediment, which makes critical the representation of such aspects.

ii Calibration: the proposed strategy to estimate temporal variation of kinetic rates presumes two main aspects: 1. daily transformation rates have a uniform distribution; 2. less impacted reaches have smaller activity for sink/sources analysis (minor rates to downstream) – this means that concentrations and rate's values are directly proportional. However, such hypothesis might not be valid for all parameters; simulations of organic carbon, for example – for which transformation processes are not entirely known in riverine systems – might be overestimated because of these assumptions in the calibration phase. In the same context, BOD simulations showed to be more affected by the strategy in defining time series of transformation rates than other water quality parameters; this does not necessarily indicates that temporal variation of kinetic processes are not relevant for the other components, but that further investigations towards other distributions, different processes (such as interaction sediment-water) and analysis to understand temporal scales should be conducted – such sediment resuspension, effects of temperature etc. Additionally, none specific quantitative parameter of calibration fit was assessed (such as difference between measured data and model output), which should be included in future efforts. At last, the calibration procedure also inserts uncertainty in model results, that could be further assessed to increase reliability.

iii Dispersion coefficient: because of numerical solution stability, this parameter was considered as a constant during the simulations; supplementary assessment of this process can generate superior estimations of pollutants transport, especially if contaminant's arrival is the goal.

iv Synthetic series: this is an useful artifice to study multiple possible scenarios, although it might generate values without physical meaning; preservation of statistical moments (mean and standard deviation) might not necessarily be the only criterion to provide reasonable estimates of water quality, especially when time evolution is the main objective – in this case, persistence over time should be represented. To increase the reliability, characteristics of these generated temporal series should be better understand (e. g., autocorrelation, trends, cycles and seasonality), which is possible only with longer term analysis and comparison with data at higher frequency sampling (Coelho et al. 2019). Other methods to generate synthetic data are encouraged; development of techniques for data analysis have been increasing over the years in multiple areas; more sophisticated forecasting methods, such as those based on artificial intelligence, have received more attention in water quality studies (Sengorur et al., 2015; Keshtegar and Heddam, 2017). Furthermore, other traditional approaches, such as regression analysis, remain subject of several studies; these latter have the advantage of empirically relate different conditions, such as land use and loading entering a system, and could be incorporated in the SIHQUAL model.

v Spatial representation: the one-dimensional configuration assumes instant mixing in lateral and vertical dimensions; after mass releases, in reality, uniform distribution in the cross section is achieved some time latter; mainly for compound cross sections, where recirculation or stagnation zones are common, simulation in two or three dimensions might be required.

vi Simulation period/scale: Substances react or interact with other components (with sediment, for example, that can be considered as sink of toxic contaminants); in case of persistent substances, transfer of pollutant to other levels may occur due to bioaccumulation, or they can reach remote regions. In analysis at longer spatial/temporal scales, it is also possible to track patterns in water pollution due to land use modifications, for example. These analysis depend on longer datasets, besides improvement in numerical solution and algorithm aspects, since computational processing time increases significantly for larger scales.

vii Numerical representation: in the same direction, instigations of other numerical strategies are encouraged; advances in computational fluid dynamics have investigated effects on simulations due to numerical diffusion and discretization techniques. that also have impact on model outputs.

viii Sampling frequency and model requirements: this research examined only one sampling strategy that represents common monitoring practice; strategies with different sampling frequencies (e.g., weekly or monthly) and record lengths (e.g., 10 years) should be investigated. Results suggested that more data does not necessarily gen-

erate better estimations of annual loads, which is compatible with other studies. An effective criterion reported to improve simulation results is input data based on flow regime or storm observations (e. g. Park and Engel, 2014; Zhang and Ball, 2017). This aspect also leads to understand the relevance of time intervals required, which is not a trivial aspect and also might lead to numerical instabilities. According to Palmer (2001) and Costa and Monteiro (2015), daily time steps are adequate to represent effects of photosynthesis and respiration, while seasonal predictions may be required for dissolved oxygen or nutrients analysis. To Moeller et al. (1979), comparisons of annual or seasonal dissolved and particulate carbon are indicative of relative amounts of organic loading, and weekly or daily concentrations would be preferable. In the same way, Magness and Raffensperger (2003) state that it is conceptually possible a watershed model to accurately predict annual flow or load, yet incorrectly capture the dynamics of the processes responsible; therefore, the author suggests hourly analysis to predict loading and natural process dynamics. While this research focused on daily averages, Baffaut et al. (2015) argues that some processes happens at a scale of seconds (adsorption/desorption, for example), minutes (such as nitrification and denitrification) and hours (algae growth/eutrophicationfor example). Overall, this definition depends of system conditions, processes and parameters being evaluated, besides model objectives.

ix Monitoring efforts: Experiments with synthetic series show that historical monitoring datasets are undoubtedly valuable, and the relationship between statistical measures should be considered when designing sampling strategies. In fact, expressions to estimate the number of samples required often use as reference medians and standard deviations (e. g. Williams et al., 2014). Technology for data acquirement has been broadly developed, with equipments that collect large amounts of information. Other promising alternative for data acquirement is multi-spectral satellite-based remote sensing; it has been applied to derive information about soil, vegetation, and climate as input data for environmental modeling (e. g. Chang and Imen, 2015; Launay et al. 2019). Despite that, monitoring efforts depend on study objectives. Model calibration could be improved with higher frequency data along the river – specially near the contours of the interval being simulated; to Baffaut et al. (2015), models should be calibrated at the scale at which the results will be analyzed and interpreted.

x Input data uncertainty: In the modeling field, scientists and users usually assume measurement data as the "truth"; however, uncertainty in monitoring data is inherent, since they are susceptible to analytical assumptions and errors due to instruments and operation. In this sense, it is important to take these uncertainties into consideration and communicate them appropriately.

xi <u>River characteristics</u>: system's hydraulic conditions often can be related to changes
in water quality, so investigation of hydraulic structures – such as bridges, con-
tractions, expansions and dams – could provide insides about aspects that alter
water quality conditions; in this sense, presence of dams have been recognized as
traps for some pollutants, such as metals, polychlorinated biphenyls, atrazine and
bisphenol A (Watkins et al 2019; Xu et al 2017), which suggests that the dynamic
river-reservoir is important in mass balances.

xii <u>Hydrodynamic modeling</u>: the simplifications involved to solve the hydrodynamic
module should be further investigated: lateral contribution has been calculated
with the same dataset used in calibration – therefore, results might be biased; for
future prediction scenarios, lateral contribution can be estimated using hydrological
models, for instance (e. g. Paiva, 2009). Results also suggest that uncertainty in
other input data, such as rating curves representation and bathymetry information,
may be responsible for discrepancies in simulations of water level and discharges
– in this sense, further investigations should be towards decreasing uncertainty in
data used as input and calibration, especially if the objective is to predict critical
events (maximum discharges and water level); in the same direction, results to
predict inundations could be improved with 2D or 3D representations, especially in
floodplain regions.

7 Bibliography

ABAURREA, J.; ASÍN, J.; CEBRIÁN, A. C.; GARCÍA-VERA, M. A. Trend analysis of water quality series based on regression models with correlated errors. **Journal of Hydrology**, v. 400, n. 3-4, p. 341–352, 2011.

ABDERREZZAK, K. E. K.; ATA, R.; ZAOUI, F. One-dimensional numerical modeling of solute transport in streams: The role of longitudinal dispersion coefficient. **Journal of Hydrology**, v. 527, p. 978-989, 2015.

ADAMS, R.; QUINN, P. F.; PERKS, M.; et al. Simulating high frequency water quality monitoring data using a catchment runoff attenuation flux tool (CRAFT). **Science of the Total Environment**, p. 1-14, 2016.

ALAGHMAND, S.; ABDULLAH, R. B.; ABUSTAN, I. Selecting the Best Set Value in Calibration Process for Validation of Hydrological Modeling (A Case Study on Kayu Ara River Basin, Malaysia). **Research Journal of Environmental Sciences**, p. 354-365, 2011.

ALARCON, V. J. The role of boundary conditions in water quality modeling. **International Conference on Computational Science and Its Applications**, 2014.

ALEKSEEVSKII, N. I.; KRYLENKO, I. N.; BELIKOV, V. V.; KOCHETKOV, V. V.; NORIN, S. V. Numerial Hydrodynamic Modeling of Inundation in Krymsk on 6 – 7 July 2012. **Power Technology and Engineering**, v. 48, n. 3, p. 179-186, 2014.

AMBROSE, R. B.; WOOL, T. A.; CONNOLLY, J. P.; SCHANZ, R. W. WASP4, **A hydrodynamic and water quality model–Model theory, user's manual and programmer's guide**, 1988.

AMORNSAMANKUL, S.; ROAD, S. A.; ROAD, S. A.; ROAD, S. A. Modified WQI Model using Fourier series and Genetic Algorithm Technique. **Recent Researches in Automatic Control and Electronics Modified**, p. 73-76, 2012.

ARCEMENT, G. J.; SCHNEIDER, V. R. **Guide for selecting Manning's roughness coefficients for natural channels and flood plains**. United States Geological Survey Water-supply, 1984.

ARYA, F. K.; ZHANG, L. Time series analysis of water quality parameters at Stillaguamish River using order series method. **Stochastic Environmental Research and Risk Assessment**, v. 29, n. 1, p. 227-239, 2015.

ÁVILA, L. C. **Modelagem da qualdiade da água utilizando coeficientes bibliográficos e experimentais - aplicação ao rio Vacacaí Mirim**. Dissertação (Mestrado em Engenharia Civil) - Universidade Federal de Santa Maria, 2014.

ASCOTT, M. J.; LAPWORTH, D. J.; GOODDY, D. C.; SAGE, R. C.; KARAPANOS, I. Science of the Total Environment Impacts of extreme fl ooding on riverbank filtration water quality. **Science of the Total Environment**, v. 554-555, p. 89-101, 2016.

BABBAR, R. Pollution risk assessment based on QUAL2E-UNCAS simulations of a tropical river in Northern India. **Environmental Monitoring and Assessment**, v. 186, n. 10, p. 6771-6787, 2014.

BAFFAUT, C. ; DABNEY, S. M.; SMOLEN, M. D.; YOUSSEF, M. A.; BONTA, J. V.; CHU, M. L.; GUZMAN, J. A.; SHEDEKAR, V. S.; JHA, M. K.; ARNOLD, J. G. Hydrologic and water quality modeling: spatial and temporal considerations. **Transactions of the ASABE**, v. 58, p. 1661-1680, 2015.

BARBER, L. B.; LEENHEER, J. A.; PEREIRA, W. E.; NOYES, T. I.; BROWN, G. K.; TABOR, C. F.; WRITER, J. H. **Organic Contamination of the Mississippi River from Municipal and Industrial Wastewater**. United States Geological Survey, circular 1133, 1995.

BARZEGAR, R.; ASGHARI, A.; JAN, M. Multi-step water quality forecasting using a boosting ensemble multi-wavelet extreme learning machine model. **Stochastic Environmental Research and Risk Assessment**, 2017.

BENEDINI, M.; TSAKIRIS, G. **Water Quality modeling for Rivers and Streams**. Texas A & M University, 2013.

BENJANKAR, R.; TONINA, D.; MCKEAN, J. One-dimensional and two-dimensional hydrodynamic modeling derived flow properties: Impacts on aquatic habitat quality predictions. **Earth Surface Processes and Landforms**, v. 40, n. 3, p. 340-356, 2015.

BRASIL. Lei Federal no 9.433, de 8 de janeiro de 1997. **Diário Oficial da República Federativa do Brasil**, Brasília, DF, 8 jan. 1997.

BRAVO, J. M.; COLLISCHOM, W.; TUCCI, C. E. M. Calibração automática de um modelo hidrológico utilizando um algoritmo evolucionário multiobjetivo. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 5, n. 3, p. 1-12, 2001.

BREDA, A.; GONÇALVES, J. E.; SILVEIRA, R. B. Automática Mono-Objetivo na Qualidade e Eficiência do Ajuste de Parâmetros do Modelo Sacramento, **Revista Brasileira**

de Recursos Hídricos, v. 16, n. 2003, p. 89-100, 2011.

BRITES, A. P. Z. **Enquadramento dos corpos de água através de metas progressivas: probabilidade de ocorrência e custos de despoluição hídrica**. Tese (Doutorado em Engenharia) - Universidade de São Paulo, 2010.

BROWN, L, C.; BARNWELL, T. O. **The enhanced stream water quality model QUAL2E and QUAL2E-UNCAS: computer program documentation and user manual**. Athens: United States Environmental Protection Agency, 189 p, 1987.

CALMON, A. P.; SOUZA, J. C.; REIAS, J. A. T.; MENDONÇA, A. S. F. Uso combinado de curvas de permanência de qualidade e modelagem da autodepuração como ferramenta para suporte ao processo de enquadramento de cursos d'água superficiais. **Revista Brasileira de Recursos Hídricos**, v.l, n. 1, p. 118-133, 2016.

CAMPOZANO, L.; SÁNCHEZ, E.; AVILES, A.; SAMANIEGO, E. Evaluation of infilling methods for time series of daily precipitation and temperature: The case of the Ecuadorian Andes. **Maskana**, v. 5, n. 1, p. 99–115, 2014.

CECCHI, G.; MUNAFÒ, M.; BAIOCCO, F.; ANDREANI, P; MANCINI, L. Estimating river pollution from diffuse sources in the Viterbo province using the potential non-point pollution index. **Annali dell'Istituto Superiore di Sanità**, v.43, n.3, 2007.

CHANAT, J. G.; MOYER, D. L.; BLOMQUIST, J. D.; HYER, K. E.; LANGLAND, M. J. **Application of a weighted regression model for reporting nutrient and sediment concentrations, fluxes, and trends in concentration and flux for the Chesapeake Bay Nontidal Water-Quality Monitoring Network, results through water year 2012**. United States Geological Survey, 2016.

CHANG, F.; TSAI, Y.; CHEN, P.; COYNEL, A.; VACHAUD, G. Modeling water quality in an urban river using hydrological factors e Data driven approaches. **Journal of Environmental Management**, v. 151, p. 87–96, 2015.

CHANG, N. B.; IMEN, S. Improving the Control of Water Treatment Plant with Remote Sensing-based Water Quality Forecasting Model. **Proceedings of IEEE 12th International Conference on Networking, Sensing and Control Howard Civil Service International House, Taipei, Taiwan**, 2015.

CHAPRA, S. C. **Surface Water-Quality Modeling**. Boston: Mass: WCB/McGraw-Hill, 1997.

CHAU, K. A review on integration of artificial intelligence into water quality modeling.

Marine Pollution Bulletin, v. 52, p. 726-736, 2006.

CHAUDHRY, M. **Applied Hydraulic Transients**. New York: Van nostrand reinhold company, 1979.

CHEN, J.; BOCCELLI, D. L. Real-time forecasting and visualization toolkit for multi-seasonal time series. **Environmental Modelling & Software**, v. 105, p. 244–256, 2018.

CHIGOR, V. N.; UMOH, V. J. ; OKUOFU, C. A.; AMEH, J. B.; IGBINOSA, E. O.; OKOH, A. I. Water quality assessment: surface water sources used for drinking and irrigation in Zaria. **Nigeria are a public health hazard**, p. 3389-3400, 2012.

CHO, J. H.; LEE, J. H. Watershed model calibration framework developed using an influence coefficient algorithm and a genetic algorithm and analysis of pollutant discharge characteristics and load reduction in a TMDL planning area. **Journal of Environmental Management**, v. 163, p. 2-10, 2015.

CHOW, V. T. **Open-Channel Hydraulics**. McGraw-Hill Book Company, 1959.

CHUCO, T. D. **Dynamic Integrated modeling of Basic Water Water and Fate and Effect of Organic Contaminants in Rivers**. Thesis submitted for the degree of Doctor in Applied Biological Sciences: Environmental Technologfy, Universiteit Gent, 2004.

CREACO, E.; BERARDI, L.; SUN, S.; GIUSTOLISI, O.; SAVIC, D. Selection of relevant input variables in storm water quality modeling by multiobjective evolutionary polynomial regression paradigm. **Water Resources Research**, V. 52, P.2403-2419, 2016.

COELHO, M. FERNANDES, C. V. S.; DETZEL, D. H. M.; Uncertainty analysis in the detection of trends, cycles, and shifts in water resources time series. **Water Resources Management**, 2019.

COELHO, M. FERNANDES, C. V. S.; DETZEL, D. H. M.; MANNICH, M. Statistical validity of water quality time series in urban watersheds. **Brazilian Journal of Water Resources**, v. 22, 2017.

CONAMA (Conselho Nacional do Meio Ambiente). **Resolução n° 357/05**. Estabelece a classificação das águas doces, salobras e salinas do Território Nacional. Brasília, SEMA, 2005.

COSTA, M. A. S.; MONTEIRO, M. S. V. Statistical modeling of Water Quality Time Series – The River Vouga Basin Case Study. Research and Practices in Water Quality,

INTECH, 2015.

CUNHA, D. G. F.; GRULL, D.; DAMATO, M.; BLUM, J. C. R.; LUTTI, J. E. I.; EIGER, S.; MANCUSO, P. C. S. Trophic state evolution in a subtropical reservoir over 34 years in response to different management procedures. **Water Science and Technolog**y, v. 64, p. 2338–2344, 2011.

CUNHA, A. C.; CUNHA, H. F. A.; PINHEIRO, L. A. R. Modelagem e simulaçao do escoamento e dispersão sazonais de agentes passivos no rio Araguari AP: Cenários para o AHE Ferreira Gomes-I-Amapá/Brasil. **Revista Brasileira de Recursos Hídricos**, v. 18, p. 67–85, 2012.

DEORAS, A. Fanchart - Visualize percentiles of time series data <https://www.mathworks.com/matlabcentral/fileexchange/48006-fanchart-visualize-percentiles-of-time-series-data>. **MATLAB Central File Exchange**. Retrieved Nov 06, 2016.

DÖKMEN, F.; ASLAN, Z. Evaluation of the Parameters of Water Quality with Wavelet Techniques. **Water Resources Management**, v. 27, n. 14, p. 4977-4988, 2013.

DOMENEGHETTI, A.; CASTELLARIN, A.; BRATH, A. Assessing rating-curve uncertainty and its effects on hydraulic model calibration. **Hydrology and Earth System Sciences**, p. 1191–1202, 2012.

DORTCH, M.; JOHNSON, B. **Hydrodynamics for Water Quality Models**. Hydraulic Engineering Sessions at Water Forum, Baltimore: American society of civil engineers, 1992.

EFSTRATIADIS, A.; DIALYNAS, Y. D.; KOZANIS, S.; KOUTSOYIANNIS, D. A multivariate stochastic model for the generation of synthetic time series at multiple time scales reproducing long-term persistence. **Environmental Modelling & Software**, p. 139-152, 2014.

ESTABRAGH, A. R.; PERESHKAFTI, M. R. S.; JAVADI, A. A. Comparison Between Analytical and Numerical Methods in Evaluating the Pollution Transport in Porous Media. **Geotechnical and Geological Engineering**, v. 31, n. 1, p. 93–101, 2012.

FAN, C.; WANG, W.-S.; LIU, K. F.-R.; YANG, T.-M. Sensitivity Analysis and Water Quality Modeling of a Tidal River Using a Modified Streeter–Phelps Equation with HEC-RAS-Calculated Hydraulic Characteristics. **Environmental Modeling and Assessment**, v. 17, n. 6, p. 639–651, 2012.

FAN, F. M.; COLLISCHONN, W.; RIGO, D. Modelo analítico de qualidade da água acoplado com Sistema de Informação Geográfica para simulação de lançamentos com duração variada. **Engenharia Sanitária e Ambiental**, v. 18, n. 4, p. 359-370, 2013.

FENG, L.; WANG, D.; CHEN, B. Modeling the effects of gate-controlled on water quality improvement in river network. **Procedia Environmental Sciences**, v. 13, n. 2011, p. 2154-2160, 2012.

FENG, Y.; LING, L.; YANFENG, L.; et al. A dynamic water quality index model based on functional data analysis, **Ecological Indicators**, v. 57, p. 249–258, 2015.

FERNANDES, C. V. S. **INTEGRA 2: Bases Técnicas para a Integração de Instrumentos de Gestão de Recursos Hídricos - Estudo de Caso da Bacia do Alto Iguaçu e Bacia do Alto Tietê**. Universidade Federal do Paraná, Curitiba, 2013.

FERREIRA, D. M. **Simulação hidrodinâmica e de qualidade da água em rios: impacto para os instrumentos de gestão de recursos hídricos**. Dissertação (Mestrado em Engenharia de Recursos Hídricos e Ambiental) - Universidade Federal do Paraná, 2015.

FERREIRA, D. M; FERNANDES, C. V. S.; KAVISKI, E. Curvas de permanência de qualidade da água como subsídio para o enquadramento de corpos d'água a partir de modelagem matemática em regime não permanente. **Revista Brasileira de Recursos Hídricos**, v. 21, n. 3, p. 479-492, 2016.

FONSECA, A.; BOTELHO, C.; BOAVENTURA, R. A R.; VILAR, V. J. P. Integrated hydrological and water quality model for river management: a case study on Lena River.**Science of the Total Environment**, v. 485-486, p. 474-89, 2014.

GAJDOS, A.; MANDELKERN, S. Comparative study of numerical schemes used for one-dimensional transport modeling. **2nd Int. PhD Symposium in Civil Engineering**, p. 1-8, Budapest, 1998.

GRUNWALD, G. K.; HYNDMAN, R. J.; TEDESCO, L. M. **A unified view of linear AR(1) models**. 1996.

GUPTA, H. V.; Sorooshian, S.; YAPO, P. O. Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. **Journal of Hydrologic Engineering**, v. 4, n. 2, p.135-143, 1999.

HANKIN, B.; BIELBY, S.; POPE, L.; DOUGLASS, J. Catchment-scale sensitivity and uncertainty in water quality modelling. **Hydrological Processes**, v. 30, n. 22, p. 4004–4018, 2016.

HASHEMI, M. R.; ABEDINI, M. J.; MALEKZADEH, P. A differential quadrature analysis of unsteady open channel flow. **Applied Mathematical modeling**, v. 31, n. 8, p. 1594–1608, 2007.

HIRSCH, R. M.; MOYER, D. L.; ARCHFIELD, S. A. Weighted Regressions on Time, Discharge, and Season (Wrtds), With an Application To Chesapeake Bay River Inputs 1. **Journal of the American Water Resources Association (JAWRA)**, v. 46, n. 5, p. 857-880, 2010.

HUGHES, J. D.; LANGEVIN, C. D.; WHITE, J. T. MODFLOW-based coupled surface water routing and groundwater-flow simulation. **Groundwater**, v. 53, n 3, p. 452-463, 2015.

HWANG, J. Y.; KIM, Y. DO; KWON, J. H.; et al. Hydrodynamic and water quality modeling for gate operation: A case study for the Seonakdong River basin in Korea. **KSCE Journal of Civil Engineering**, v. 18, n. 1, p. 73-80, 2014.

IDE, A. H.; CARDOSO, F. D.; SANTOS, M. M.; KRAMER, R. D.; MIZUKAWA, A.; AZEVEDO, J. C. R. Cafeína como indicador de contaminação por esgotos domésticos na Bacia do Alto Iguaçu. **Revista Brasileira de Recursos Hídricos**, v. 18, n. 2, p. 201-211, 2013.

ISLAM, M. J.; CHANGWON, J.; JAESUNG, E.; JUNG, S.-M.; SHIN, M.-S.; LEE, Y.; CHOI, Y.; KIM, B. C:N:P stoichiometry of particulate and dissolved organic matter in river waters and changes during decomposition. **Journal of Ecology and Environment**, v. 43., 2019.

INTHASARO, P.; WU, W. A 1-D Aquatic Ecosystem/Ecotoxicology Model in Open Channels. **International Conference on Biomedical Engineering and Biotechnology**, 2012.

INSTITUTO DAS ÁGUAS DO PARANÁ. **Sistema de Informações Hidrológicas**. Available at: <http://www.aguasparana.pr.gov.br/modules/conteudo/conteudo.php ?conteudo=264.> Accessed in june 13th, 2017.

JARUSKOVÁ, D. Some problems with application of change-point detection methods to environmental data. **Environmetrics**, v. 8, p. 469-483. 1997.

JI, Z.G. **Hydrodynamics and water quality: modeling rivers, lakes, and estuaries**. John Wiley & Sons, 2008.

JIA, H.; XU, T.; LIANG, S.; ZHAO, P.; XU, C. Bayesian framework of parameter sen-

sitivity, uncertainty, and identifiability analysis in complex water quality models. **Environmental Modelling & Software**, v. 104, p. 13–26, 2018.

JUXIANG, J.; HAIXIAO, J.; CHANGGEN, L.; JIANHUA, T. Hydrodynamic and Water Quality Models of River Network and its Application in the Beiyun River. **5th International Conference on Bioinformatics and Biomedical Engineering**, 2011.

KALITA, H. M.; SARMA, A. K. Efficiency and performances of finite difference schemes in the solution of Saint Venant's equation. **International Journal of Civil and Structural Engineering**, v. 2, n. 3, p. 950–958, 2012.

KANDA, E. K.; KOSGEI, J. R.; KIPKORIR, E. C. Simulation of organic carbon loading using MIKE 11 model: a case of River Nzoia, Kenya. **Water Practice & Technology**, v. 10, n. 2, p. 298, 2015.

KESHTEGAR, B.; HEDDAM, S. Modeling daily dissolved oxygen concentration using modified response surface method and artificial neural network: a comparative study. **Neural Computing & Applications**, 2017.

KIM,S.; BENHAM, M. B. L.; BRANNAN, K. M.; ZECKOSKI, R. W.; YAGOW, G. R. Water quality calibration criteria for bacteria TMDL development. **Applied Engineering in Agriculture**, v. 23, n. 2, p. 171-176, 2007.

KNAPIK, H. G.; FERNANDES, C. V. S.; AZEVEDO, J. C. R. Modelagem de carbono orgânico em rios urbanos: aplicabilidade para o planejamento e gestão de qualidade de água. **Brazilian Journal of Water Resources**, v.21, n.4, 2016.

KONDAGESKI, J. H.; FERNANDES, C. V. S. Calibração de Modelo Matemático de Qualidade da Água Utilizando Algoritmo Genético: Estudo de Caso do Rio Palmital, PR. **Revista Brasileira de Recursos Hídricos**, v. 14, n.1, p. 63-73, 2009.

LAMONTAGNE, J. R.; STEDINGER, J. R. Generating Synthetic Streamflow Forecasts with Specified Precision. **Journal of Water Resources Planning and Management**, v. 144, n. 4, 2018.

LANGEVELD, J.; NOPENS, I.; SCHILPEROORT, R.; et al. On data requirements for calibration of integrated models for urban water systems. **Water Science and Technology**, v. 68, n. 3, p. 728-736, 2013.

LARENTIS, D. G.; COLLISCHONN, W.; TUCCI, C. E. M. Simulação da Qualidade de Água em Grandes Bacias: Rio Taquari-Antas, **Revista Brasileira de Recursos Hídricos**, v. 13, p. 5-22, 2008.

LAUNAY, M.; DUGUÉ, V.; FAURE, J.-B.; COQUERY, M.; CAMENEN, B; COZ, J. L. Numerical modelling of the suspended particulate matter dynamics in a regulated river network. **Science of the Total Environment**, v. 665, p. 591–605, 2019.

LELIS, T. A.; CALIJURI, M. L.; DA FONSECA SANTIAGO, A.; DE LIMA, D. C.; DE OLIVEIRA ROCHA, E. Análise de Sensibilidade e Calibração do Modelo Swat Aplicado em Bacia Hidrográfica da Região Sudeste do Brasil. **Revista Brasileira de Ciência do Solo**, v. 36, n. 2, p. 623-634, 2012.

LENHART, T.; ECKHARDT, K.; FOHRER, N.; FREDE, H. Comparison of two different approaches of sensitivity analysis. **Physics and Chemistry of the Earth**, v. 27, p. 645-654, 2002.

LEENHEER, J. A.; CROUÉ, J. P. Characterizing dissolved aquatic organic matter. **Environmental Science and Technology**, v. 37, p. 18-26, 2003.

LI, G.; JACKSON, C. R. Simple, accurate, and efficient revisions to MacCormack and Saulyev schemes: high Peclet numbers. **Applied Mathematics and Computation**, v. 186, n. 1, p. 610–622, 2007.

LIN, F.-R.; WU, N.-J.; TU, C.-H.; TSAY, T.-K. Automatic Calibration of an Unsteady River Flow Model by Using Dynamically Dimensioned Search Algorithm. **Mathematical Problems in Engineering**, 2017.

LINDEN, L. V. D. ; DALY, R. I.; BURCH, M. D. Suitability of a coupled hydrodynamic water quality model to predict changes in water quality from altered meteorological boundary conditions. **Water (Switzerland)**, v. 7, n. 1, p. 348-361, 2015.

LINDENSCHMIDT, K. E. **River water quality modeling for river basin and water resources management with a focus on the Saale River, Germany**. Habilitation thesis from the Brandenburgische Technische Universität Cottbus, Germany. 2006.

LIU, S.; BUTLER, D.; BRAZIER, R.; HEATHWAITE, L.; KHU, S.-T. Using genetic algorithms to calibrate a water quality model. **The Science of the total environment**, v. 374, p. 260-272, 2007.

LEE, C.; CHANG, C.; WEN, C.; CHANG, S. Agriculture, Ecosystems and Environment Comprehensive nonpoint source pollution models for a free-range chicken farm in a rural watershed in Taiwan. **Agriculture, Ecosystems and Environment**, v. 139, n. 1-2, p. 23-32, 2010.

LIGGETT, J. A. Basic equations of unsteady flow. In: Mahmood, K.; V. Yevjevich,

Unsteady flow in open channels. Fort collins: water resources publications, 1975.

LINS, R. C.; et al. Integração de um modelo hidrológico a um SIG para avaliação da qualidade da água na bacia do Rio Doce. **Revista Brasileira de Recursos Hídricos**, v. 17, n. 4, p. 171–181, 2012.

LOUCKS, D. P.; BEEK, E. V. **Water Resource Systems Planning and Management: An Introduction to Methods, Models, and Applications**. Springer International Publishing, 2017.

MAGNESS, A. G.; RAFFENSPERGER, J. **Development, calibration, and analysis of a hydrologic and water-quality model of the Delaware Inland Bays Watershed**. Report USGS, 2003.

MANNINA, G.; VIVIANI, G. River water quality assessment: a hydrodynamic water quality model for propagation of pollutants. **Novatech**. 2010.

MARTIN, J. L.; MCCUTCHEON, S. C. **Hydrodynamics and transport for water quality modeling**. Lewis Publishers, 1999.

MATHWORKS. Online MatLab documentation: **Fourier Series**. Available at: <https://www.mathworks.com/help/curvefit/fourier.html>. 2017a.

MATHWORKS. Online MatLab documentation: **Smoothing Splines**. Available at: <https://www.mathworks.com/help/curvefit/smoothing-splines.html>. 2017b.

MATHWORKS. Online MatLab documentation: **Distance between signals using dynamic time warping**. Available at: <https://www.mathworks.com/help/signal/ref/dtw.html>. 2018.

MATTSSON, T.; KORTELAINEN, P.; LAUBEL, A.; EVANS, D.; PUJO-PAY, M.; RÄIKE, A.; CONAN, P. Export of dissolved organic matter in relation to land use along a European climatic gradient. **Science of the Total Environment**, v. 407, p. 1967-1976, 2009.

MEALS, D. W.; RICHARDS, R. P.; TECH, T. Pollutant Load Estimation for Water Quality Monitoring Projects. **National Nonpoint Source monitoring program**, p. 1-21, 2013.

MELCHING, C. S.; ALP, E.; SHRESTHA, R. L.; LANYON, R. Simulation of Water Quality During Unsteady Flow in the Chicago Waterway. **Watershed**, 2003.

METCALF and EDDY, Inc. **Wastewater Engineering: Treatment, Disposal, and**

Reuse. 3rd Edition, McGraw-Hill, 1991.

MCINTYRE, N. **Analysis of uncertainty in river water quality modelling**. PhD diss. London, England: University of London, p. 224, 2004.

MILNE, A. E.; MACLEOD, C. J. A.; HAYGARTH, P. M.; HAWKINS, J. M. B.; LARK, R. M. The wavelet packet transform: A technique for investigating temporal variation of river water solutes. **Journal of Hydrology**, v. 379, n. 1-2, p. 1-19, 2009.

MIZUKAWA, A.; MOLINS-DELGADO, D.; CÉSAR, J.; AZEVEDO, R. Sediments as a sink for UV filters and benzotriazoles: the case study of Upper Iguaçu watershed, Curitiba (Brazil). **Environmental Science and Pollution Research**, p. 18284–18294, 2017.

MOELLER, J. R.; *et al.* Transport of dissolved organic carbon in streams of differing physiographic characteristics. **Organic geochemistry**, v. 1, p. 139-150, 1979.

MUNAFÒ, M.; CECCHI, G.; BAIOCCO, F.; MANCINI, L. River Pollution from non-point source: a new simplified method of assessment. **Journal of Environmental Management**, v. 77, p. 93-98, 2005.

NG, A. W. M.; PERERA, B. J. C. Selection of genetic algorithm operators for river water quality model calibration. **Engineering Applications of Artificial Intelligence**, v. 16, n. 5-6, p. 529-541, 2003.

NGUYEN, T. T.; KEUPERS, I.; WILLEMS, P. Conceptual river water quality model with flexible model structure. **Environmental Modelling & Software**, v. 104, p. 102-117, 2018.

NOGUEIRA, V. P. Q. Qualidade da água em lagos e reservatórios. In: S. B. BRANCO (Ed.); **Hidrologia Ambiental**. São Paulo: Associação Brasileira de Recursos Hídricos, p. 185–210, 1991.

NOH, J.; CHOI, H.; LEE, S. Water quality projection in the Geum River basin in Korea to support integrated basin-wide water resources management. **Environmental Earth Sciences**, v. 73, n. 4, p. 1745-1756, 2015.

NOVOTNY, V.**Water quality: diffuse pollution and watershed management**. Wiley, 2 ed., 2002.

OLIVEIRA, F. A. DE; PEREIRA, T. S. R.; SOARES, A. K.; FORMIGA, K. T. M. Uso de modelo hidrodinâmico para determinação da vazão a partir de medições de nível. **Brazilian Journal of Water Resources**, v. 21, n. 4, p. 707–718, 2016.

OLIVEIRA, P. T. S.; RODRIGUES, D. B. B.; SOBRINHO, T. A.; PANACHUKI, E. Integração de informações quali-quantitativas como ferramenta de gerenciamento de recursos hídricos. **Revista de Estudos Ambientais**, v. 13, n. 1, p. 18-27, 2011.

OLSEN, N.**Numerical modeling and hydraulics**. Department of Hydraulic and Environmental Engineering, Norwegian University of Science and Technology, 2007.

PAIVA, R. C. D. M**odelagem hidrológica e hidrodinâmica de grandes bacias - estudo de caso: bacia do rio Solimões**. Dissertação (Mestrado em Recursos Hídricos e Saneamento Ambiental) - Universidade Federal do Rio Grande Do Sul, 2009.

PAIVA, R.; COLLISCHONN, W.; BRAVO, J. Modelo Hidrodinâmico 1D para Redes de Canais Baseado no Esquema Numérico de MacCormack. **Revista Brasileira de Recursos Hídricos**, v. 16, n. n. 3, p. 151-161, 2011.

PALMER, M. D. **Water Quality Modeling: A Guide to Effective Practice**. 2001.

PARK, D.; ROESNER, L. A. Evaluation of pollutant loads from stormwater BMPs to receiving water using load frequency curves with uncertainty analysis. **Water Research**, v. 46, n. 20, p. 6881–6890, 2012.

PARK, Y. S.; ENGEL, B. A. Use of Pollutant Load Regression Models with Various Sampling Frequencies for Annual Load Estimation. **Water**, v. 6., p. 1685-1697, 2014.

PARMAR, K. S.; BHARDWAJ, R. Water quality management using statistical analysis and time-series prediction model. **Applied Water Science**, p. 425–434, 2014.

PARMAR, D. L.; KESHARI, A. K. Sensitivity analysis of water quality for Delhi stretch of the River Yamuna, India. **Environmental Monitoring and Assessment**, v. 184, n. 3, p. 1487-1508, 2012.

PENG, S. H. 1D and 2D numerical modeling for solving dam-break flow problems using finite volume method. **Journal of Applied Mathematics**, v. 2012, p. 1-14, 2012.

PEREIRA, G. A. A.; SOUZA, R. C. Long Memory Models to Generate Synthetic Hydrological Series. **Mathematical Problems in Engineering**, v. 2014, 2014.

PORTO, R. L.; et al. **Hidrologia ambiental**. São Paulo: ABRH: EDUSP, 1991.

PORTO, M. F. A.; et. al. Relatório parcial nº 12. **Bacias Críticas: Bases técnicas para a definição de metas progressivas para seu enquadramento e a integração com os demais instrumentos de gestão**. FINEP, Curitiba, 2007.

PREIS, A.; OSTFELD, A. A coupled model tree–genetic algorithm scheme for flow and water quality predictions in watersheds. **Journal of Hydrology**, v.349, p. 364-375, 2008.

QUIEL, K.; BECKER, A.; KIRCHESCH, V.; SCHÖL, A.; FISCHER, H. Influence of global change on phytoplankton and nutrient cycling in the Elbe River. **Regional Environmental Change**, v. 11, p. 405-421, 2011.

RAMOS, T. B.; GONÇALVES, M. C.; BRANCO, M. A.; et al. Sediment and nutrient dynamics during storm events in the Enxoé temporary river, southern Portugal. **Catena**, v. 127, p. 177-190, 2015.

RAZAVI, S.; TOLSON, B. A.; MATOTT, L. S.; THOMSON, N. R.; MacLEAN, A.; SEGLENIEKS, F. R. Reducing the computational cost of automatic calibration through model preemption, **Water Resources Research**, v. 46, 2010.

REDER, K.; ALCAMO, J.; FLÖRKE, M. A sensitivity and uncertainty analysis of a continental-scale water quality model of pathogen pollution in African rivers. **Ecological Modelling**, v. 351, p. 129–139, 2017.

RICHARDS, R.P. **Estimation of pollutant loads in rivers and streams: a guidance document for NPS programs**. United States Environmental Protection Agency, 1998.

RODE, M.; ARHONDITSIS, G.; BALIN, D.; et al. New challenges in integrated water quality modeling. **Hydrological Processes**, v. 24, n. 24, p. 3447-3461, 2010.

RODE, M.; WRIEDT, G. Identification of a river water quality model and assessment of data information content. **Proceedings of symposium S7 held during the Seventh IAHS Scientific Assembly at Foz do Iguaçu**, Brazil. p. 108-115, 2006.

RODE, M.; SUHRA, U.; WRIEDTB, G. Multi-objective calibration of a river water quality model-Information content of calibration data. **Ecological Modelling**, p. 129-142, 2007.

RODRÍGUEZ, J. P.; MCINTYRE, N.; DÍAZ-GRANADOS, M.; et al. Generating time-series of dry weather loads to sewers. **Environmental modeling and Software**, v. 43, p. 133–143, 2013.

SALEH, F.; DUCHARNE, A.; FLIPO, N.; OUDIN, L.; LEDOUX, E. Impact of river bed morphology on discharge and water levels simulated by a 1D Saint-Venant hydraulic model at regional scale. **Journal of Hydrology**, v. 476, p. 169-177, 2013.

SALLA, M. R.; ARQUIOLA, L. P.; SOLERA, A.; et al. Integrated modeling of water

quantity and quality in the Araguari River basin, Brazil. **Latin American Journal of Aquatic Research**, v. 42, n. 1, 2014.

SANTOS, M. M., BREHM, F. D. A., FILIPPE, T. C., KNAPIK, H. G., & AZEVEDO, J. C. R. Occurrence and risk assessment of parabens and triclosan in surface waters of southern Brazil: a problem of emerging compounds in an emerging country. **Revista Brasileira de Recursos Hídricos**, v. 21, n. 3, p. 603-617, 2016.

SEARCY, J. K. **Manual of Hydrology: Part 2. Low-Flow Techniques-Flow Duration Curves**. Geological Survey Water-Supply Paper 1542-A, 1959.

SENGORUR, B.; KOKLU, R.; ATES, A. Water Quality Assessment Using Artificial Intelligence Techniques: SOM and ANN—A Case Study of Melen River Turkey. **Journal of Water Quality, Exposure and Health**, 2015.

SERRANO, S. E. Propagation of Nonlinear Flood Waves in Rivers. **Journal of Hydrology Engineering**, v. 21, n. 1, 2016.

SERVAIS, P.; GARNIER, J.; DEMARTEAU, N.; BRION, N.; BILLEN, G. Supply of organic matter and bacteria to aquatic ecosystems through waste water effluents. **Water Research**, v. 33, n. 16, p. 3521-3531, 1999.

SHAO, D. ; NONG, X.; TAN, X.; CHEN, S.; XU, B.; HU, N. Daily Water Quality Forecast of the South-To-North Water Diversion Project of China Based on the Cuckoo Search-Back Propagation Neural Network, **Water**, v. 10, 2018.

SINCOCK, A. M.; WHEATER, H. S.; WHITEHEAD, P. G. Calibration and sensitivity analysis of a river water quality model under unsteady flow conditions. **Journal of Hydrology**, v. 277, n. 3-4, p. 214-229, 2003.

SINHA, S.; LIU, X.; GARCIA, M. H. A Three-Dimensional Water Quality Model of Chicago Area Waterway System (CAWS). **Environmental Modeling and Assessment**, v. 18, n. 5, p. 567-592, 2013.

SIQUEIRA, V. A.; COLLISCHONN, W.; FAN, F. M. CHOU, S. C. Ensemble flood forecasting based on operational forecasts of the regional Eta EPS in the Taquari-Antas basin. **Brazilian Journal of Water Resources**, v. 21, n. 3, p. 587-602, 2017.

SOKOLOVA, E.; PETTERSON, S. R.; DIENUS, O.; et al. Microbial risk assessment of drinking water based on hydrodynamic modeling of pathogen concentrations in source water. **Science of the Total Environment**, v. 526, p. 177-186, 2015.

SORRIBAS, M. V.; COLLISCHONN, W.; MARQUES, M.; et al. Modelagem Distribuída

do Carbono em Bacias Hidrográficas. **Revista Brasileira de Recursos Hídricos**, v. 17, p. 225–240, 2012.

SOUZA, C.; AGRA, S.; TASSI, R.; COLLISCHONN, W. Desafios e oportunidades para a implementação do hidrograma ecológico. **Revista de Gestão de Água da América Latina**, v. 5, n. 1, p. 25-38, 2008.

STEINSTRASSER, C. E. **Método difusivo de Lax aplicado na solução das equações de Saint Venant**. Dissertação (Mestrado em Engenharia de Recursos Hídricos e Ambiental) - Universidade Federal do Paraná, 2005.

STEPIEN, I. On the numerical solution of the Saint-Venant equations. **Journal of Hydrology**, v. 67, p. 1–11, 1984.

STREETER, H. W.; PHELPS, E. B. **A study of the pollution and natural purification of the Ohio river III**. 1925.

SZYMKIEWICZ, R. **Numerical modeling in open channel hydraulics**. Dordrecht: Springer, 2010.

TANG, G.; ZHU, Y.; WU, G.; LI, J.; LI, Z-L.; SUN, J. Modeling and Analysis of Hydrodynamics and Water Quality for Rivers in the Northern Cold Region of China. **International Journal of Environmental Research and Public Health**, v. 13, n. 4, p. 408, 2016.

THOMANN, R. V.; MUELLER, J. A. **Principles of Surface Water Quality Modeling and Control**. New York: Harper Collins Publishers, 1987.

THOMAS, O.; THERAULAZ, F. Agregate Organic Constituints. In. Thomas, O.; Burgess, C. (Editors) **UV-Visible Spectrophotometry of Water and Wastewater**. Techniques and Instrumentation in Analytical Chemistry, v. 27. Elsevier, p. 89 -113. 2007.

TORRES-BEJARANO, F.; DENZER, R.; RAMÍREZ, H.; et al. Development and integration of a numerical water quality model with the geospatial application suite CIDs. **19th International Congress on modeling and Simulation**, Perth, Australia, p.12–16, 2011.

USACE (United States Army Corps of Engineering). **River Hydraulics**. 1993.

VON SPERLING, M. **Estudos e modelagem da qualidade da água de rios**. Belo Horizonte: Editora UMFG, 2007.

WANG, K.; WANG, X.; LV, G.; JIANG, M.; KANG, C.; SHEN, L. Particle Swarm Optimization for Calibrating Stream Water Quality Model. **Second International Symposium on Intelligent Information Technology Application**. p. 682-686, 2008.

WANG, X.; WANG, Q.; WU, C.; et al. A method coupled with remote sensing data to evaluate non-point source pollution in the Xin'anjiang catchment of China. **Science of the Total Environment**, v. 430, p. 132-143, 2012.

WANG, P. F., SUTULA, M., CHADWICK, B., & CHOI, W. Watershed Loading, Hydrodynamic, and Water Quality Modeling in Support of the Loma Alta Slough Bacteria and Nutrient TMDL. **Southern California Coastal Water Research Project**, 2013.

WAGENSCHEIN, D.; RODE, M. modeling the impact of river morphology on nitrogen retention-A case study of the Weisse Elster River (Germany). **Ecological modeling**, v. 211, n. 1–2, p. 224–232, 2008.

WATKINS, L.; MCGRATTAN, S.; SULLIVAN, P. J.; WALTER, M. T. The effect of dams on river transport of microplastic pollution. **Science of The Total Environment**, v. 664, p. 834-840, 2019.

WHITFIELD, P. H.; DOHAN, K. Identification and characterization of water quality transients using wavelet analysis. II. Applications to electronic water quality data. Wavelet Analysis Methodol, **Water Sci Technol**, v. 36, n. 5, p. 337-348, 1997.

WILLIAMS, R. E., ARABI, M., LOFTIS, J., ELMUND, G. K. Monitoring Design for Assessing Compliance with Numeric Nutrient Standards for Rivers and Streams Using Geospatial Variables. **Journal of Environmental Quality**, v. 43, p. 1713–1724, 2014.

WÖHLING, T.; SAMANIEGO, L.; KUMAR, R. Evaluating multiple performance criteria to calibrate the distributed hydrological model of the upper Neckar catchment. **Environmental Earth Sciences**, p. 453-468, 2013.

XU, T.; VALOCCHI, A. J; YE, M.; LIANG, F.; LIN, Y.-F. Bayesian calibration of groundwater models with input data uncertainty. **Water Resources Research**, v. 53, n. 4., 2017.

ZHANG, Q.; HARMAN, C. J.; KIRCHNER, J. W. Evaluation of statistical methods for quantifying fractal scaling in water-quality time series with irregular sampling. **Hydrology and Earth System Sciences**, p. 1175–1192, 2018.

ZHANG, W.; LI, T.; DAI, M. Uncertainty assessment of water quality modeling for a small-scale urban catchment using the GLUE methodology: a case study in Shanghai,

China. **Environmental Science and Pollution Research**, v. 22, n. 12, p. 9241–9249, 2015.

ZHANG; Q.; BALL, W. P. Improving riverine constituent concentration and flux estimation by accounting for antecedent discharge conditions. **Journal of Hydrology**, v. 547, p. 387-402, 2017.

ZHOU, N.; WESTRICH, B.; JIANG, S.; WANG, Y. A coupling simulation based on a hydrodynamics and water quality model of the Pearl River Delta, China. **Journal of Hydrology**, v. 396, n. 3-4, p. 267-276, 2011.

ZETTERQVIST, L. Statistical estimation and interpretation of trends in water quality time series. **Water Resources Research**, v. 27, n. 7, p. 1637-1648, 1991.

ZUO, Q.; CHEN, H.; DOU, M. Experimental analysis of the impact of sluice regulation on water quality in the highly polluted Huai River Basin. **Environmental Monitoring Assessment**, 2015.

A Appendix

A.1 Numerical solutions

Hydrodynamic model

Applying the Lax diffusive method to the expression for continuity, represented by equation (3.1), we have an explicit equation for calculating the channel depth:

$$y_i^{k+1} = \alpha y_i^k + (1 - \alpha)\, y^* - \frac{U^* \Delta t}{2\Delta x} \left(y_{i+1}^k - y_{i-1}^k\right) - \frac{U^* \Delta t}{B^* 2\Delta x} \left(A_{i+1}^k - A_{i-1}^k\right) - $$
$$\frac{A^* \Delta t}{B^* 2\Delta x} \left(U_{i+1}^k - U_{i-1}^k\right) + \frac{q_i^k \Delta t}{B^*} \tag{A.1}$$

where:
$B^* = (B_{i-1}^k + B_{i+1}^k)/2$, $y^* = (y_{i-1}^k + y_{i+1}^k)/2$, $U^* = (U_{i-1}^k + U_{i+1}^k)/2$ and $A^* = (A_{i-1}^k + A_{i+1}^k)/2$.

Applying the same scheme to equation (3.2), for momentum conservation, and re-arranging the terms, an expression is obtained for the calculation of the velocity at the instant t_{k+1} based on values from the time t_k:

$$U_i^{k+1} = \alpha U_i^k + (1 - \alpha)\, U^* - \frac{U^* \Delta t}{2\Delta x} \left(U_{i+1}^k - U_{i-1}^k\right) - \frac{g\Delta t}{2\Delta x} \left(y_{i+1}^k - y_{i-1}^k\right) + $$
$$\frac{q_i^k \left(v_L - U^*\right)}{A^*} \Delta t + g\Delta t \left(S_0 - S_f^*\right) \tag{A.2}$$

where $S_f^* = (S_{fi-1}^k + S_{fi+1}^k)/2$, e S_0 varies along space.

The stability of the method depends on Courant's condition, given by (Liggett and Cunge, 1975):

$$\frac{\Delta t}{\Delta x} \leq \frac{1}{|U + c|} \tag{A.3}$$

The variable c in equation (A.3) represents the celerity (m/s).

Water quality model

The solution of the unsteady water quality model is based on the application of the FTCS method to equation (3.4):

$$C_i^{k+1} = C_i^k - \frac{U_i^k \Delta t}{2\Delta x} \left(C_{i+1}^k - C_{i-1}^k \right) + \frac{D\Delta t}{A_i^k} \frac{A_{i+1}^k - A_{i-1}^k}{2\Delta x} \frac{C_{i+1}^k - C_{i-1}^k}{2\Delta x} + \tag{A.4}$$

$$\frac{D\Delta t}{(\Delta x)^2} \left(C_{i+1}^k - 2C_i^k + C_{i-1}^k \right) \pm F\Delta t \tag{A.5}$$

The requirements for stability of the solution for the advection-dispersion equation are given by (Chapra, 1997):

$$\lambda = \frac{D\Delta t}{(\Delta x)^2} < \frac{1}{2} \tag{A.6}$$

$$\gamma = \frac{U\Delta t}{\Delta x} < 1 \tag{A.7}$$

where λ is the diffusion number e γ is the Courant number.

A.2 Synthetic series: data verification and residual check
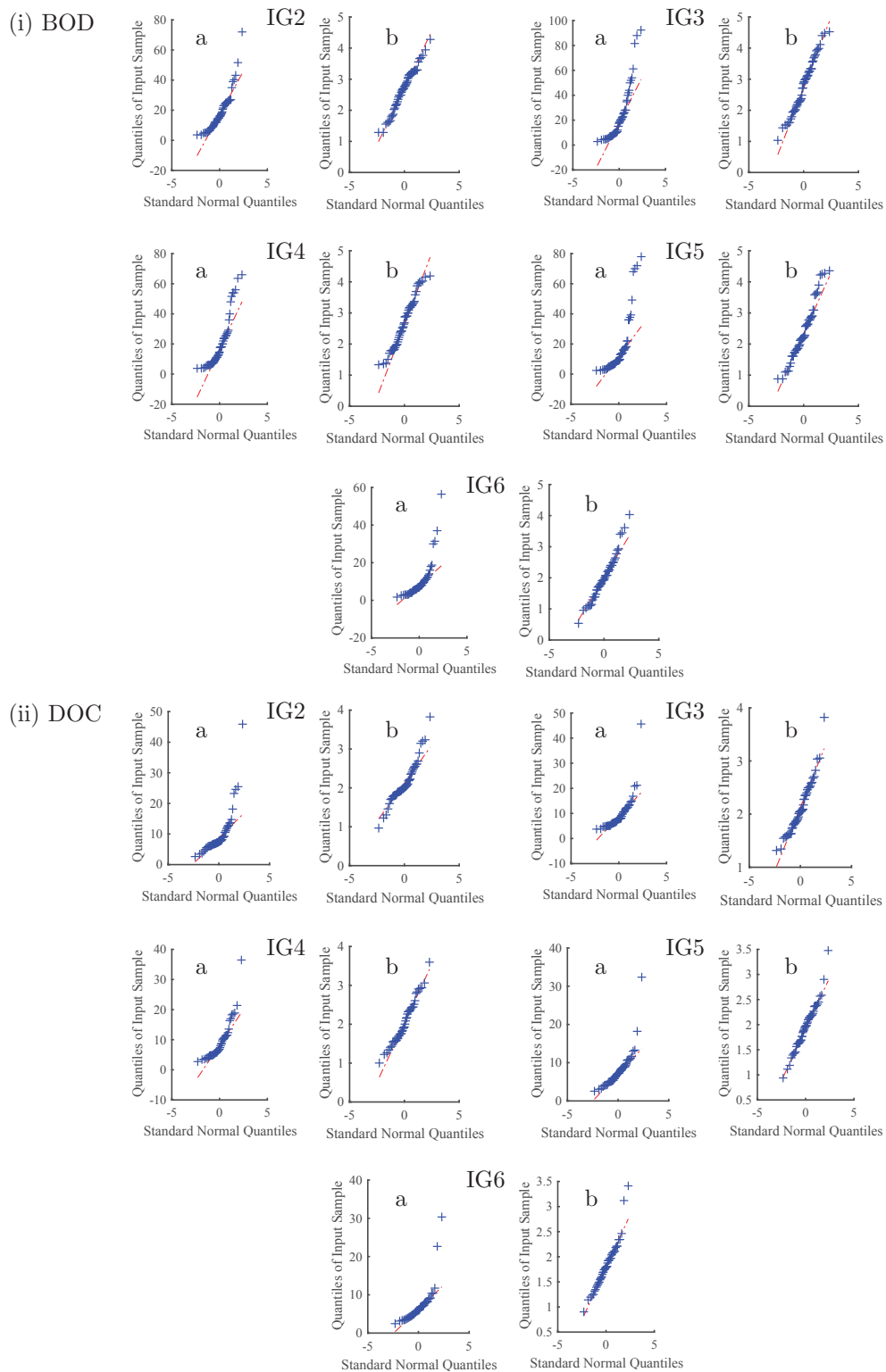
(i) BOD

(ii) DOC

Figure A.1: Quantile-quantile plots for the concentration sample in each section (a), and corresponding natural logarithm (b)

(iii) N-org
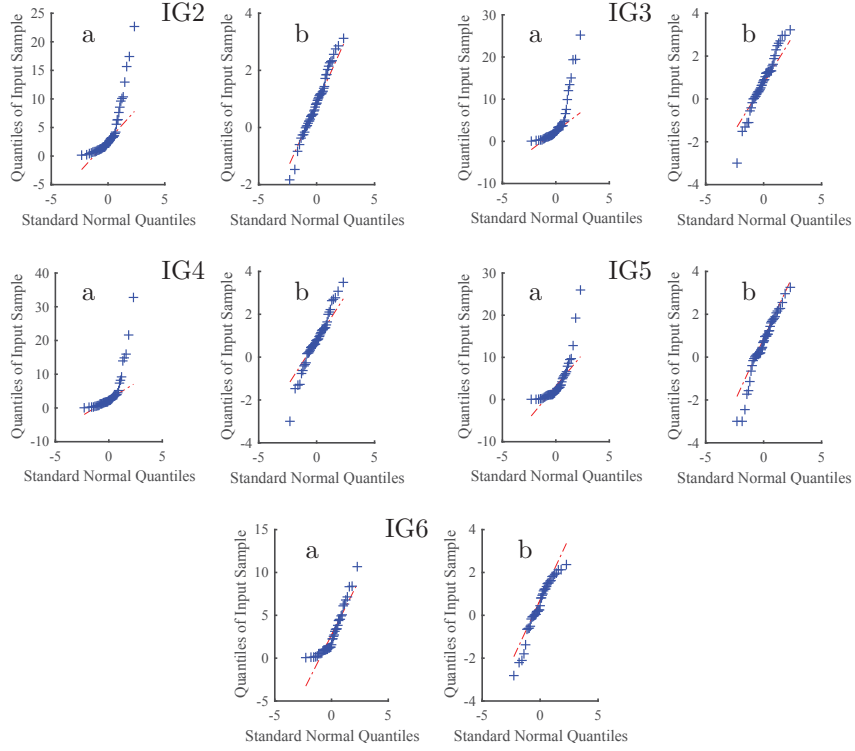
(iv) DO
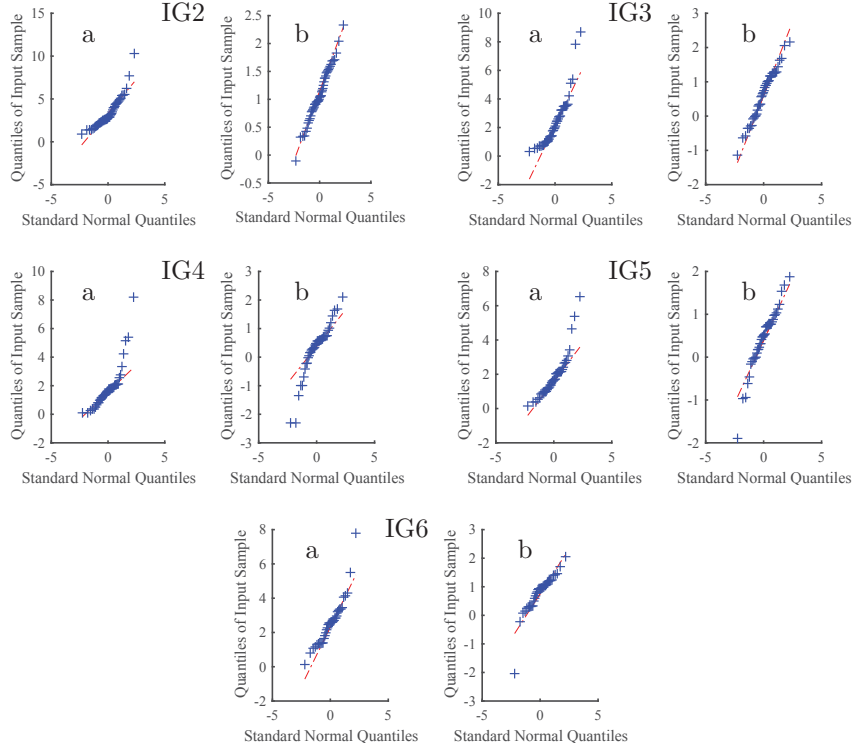
Figure A.1: Quantile-quantile plots for the concentration sample in each section (a), and corresponding natural logarithm (b) (continued)

Table A.1 shows the results for residuals check of the fitted AR(1) models. Further detail about this theory is presented in Box et al. (2008)[5]. Since the p-values are larger than 0.05, the hypothesis of independence, homoscedasticity and normality are accepted.

Table A.1: AR(1) validation tests

Test	Portmanteau	Levene	Kolmogorov-Smirnov
		p-values	
T1	0.254	0.162	0.948
T2	0.254	0.162	0.948
T3	0.254	0.162	0.948
T4*	0.043	–	0.876
T5	0.254	0.162	0.948
T6	0.066	0.06	0.316
T7	0.066	0.06	0.316
T8	0.254	0.162	0.948
T9	0.059	0.08	0.712
T10	0.058	0.115	0.511

*the unsatisfactory results for this test might be caused by lack of data, since only seven values are considered in this computation; further investigations are recommended

[5]BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Time Series Analysis: Forecasting and Control**. Wiley, 4 ed., 2008.

A.3   Representativeness of monitoring data

Figure A.2 shows boxplots formed with different monitoring campaigns added, following the concept presented by Coelho et al. (2019). The first one represents only the first five data available, and the subsequent boxplots are built adding values. Results corroborate the assumption that the historical monitoring dataset is representing the natural range of concentrations in the system, since boxplots, mean and standard deviations have not been significantly altered in the last campaigns.
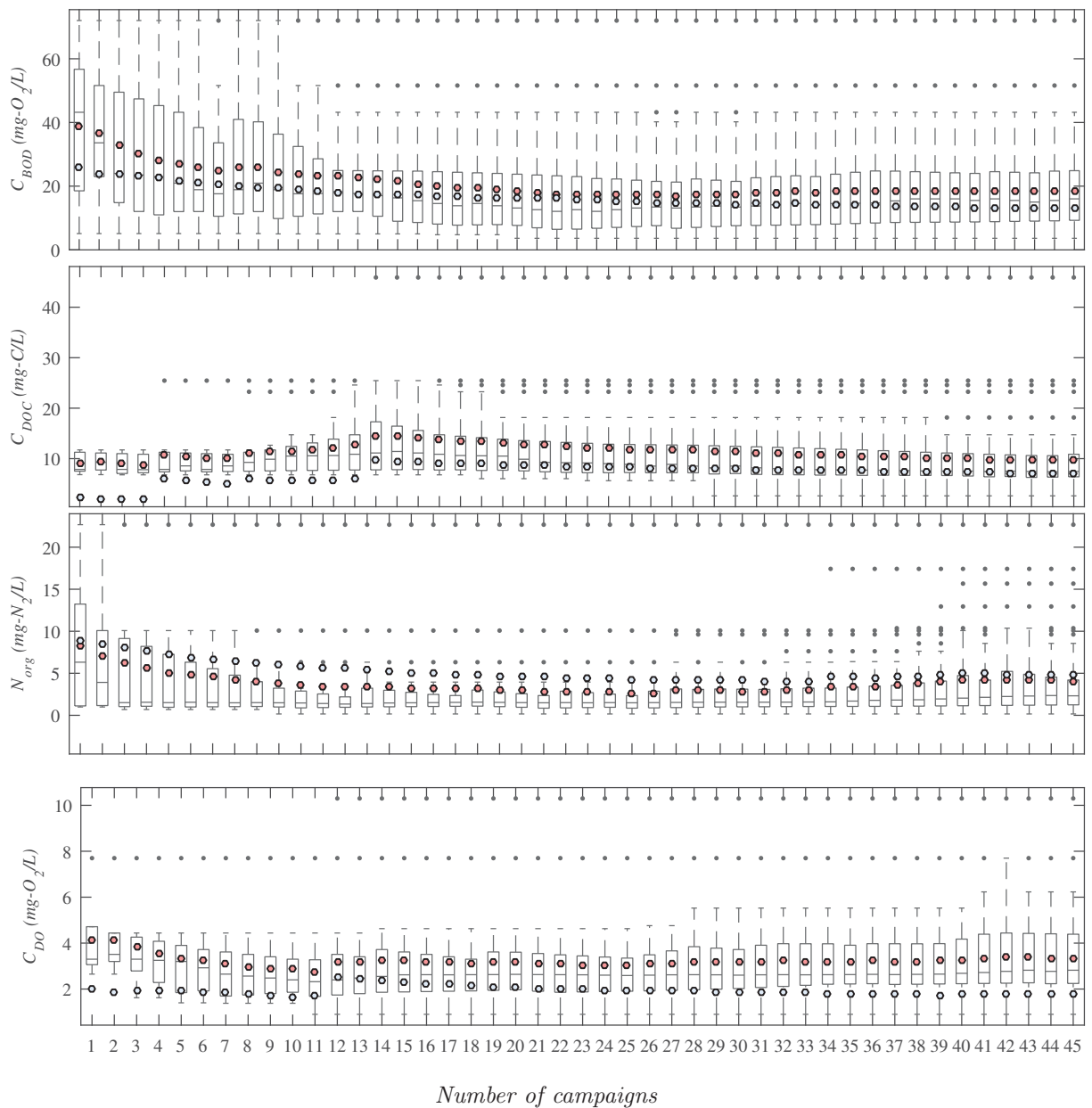
Figure A.2: Monitoring boxplots, means (blue circle) and standard deviation (pink circle) at section IG2

A.4 Caffeine model

Caffeine is recognized as one of the pharmacologically active chemical substances of higher consumption worldwide (100-200 mg/person), and may be found in the composition of food, medicines, flavorings, among others. This substance is indicative of potential pollution of water bodies in areas of intense urbanization, since anthropogenic sources are the main responsible for the presence of this compound in aquatic systems (Seiler et al., 1999).

Dombroski et al. (2013) call attention to the fact that few studies have been dedicated to analyze the impact that the presence of this compound may have on the aquatic ecosystem and in its transport and decomposition processes. Many of the researches involving modeling of this parameter have a strong investigative character of decay processes, adsorption and transformation in the water column of aquatic systems, since the mechanisms are not completely known.

According to Canela et al. (2014), in continental surface waters, caffeine levels follow aspects such as seasonality, proximity of sources, hydrological conditions and pattern of consumption. Therefore, it may be an adequate indicator of the organic pollution coming from domestic wastewater to be used in unsteady simulations. Temporal variations of caffeine concentrations have been identified by Buerge et al., (2006) in Switzerland, while Busse and Nagoda (2015) detected this substance during wet and dry seasons in the San Diego Region. According to these authors, caffeine may undergo sorption, chemical transformations, phototransformations, and biotransformations under aerobic and anaerobic environments.

The half-life of this substance in surface waters has been reported to range from 5.3 to 24 hours (Bradley et al., 2007). However, Thomas and Foster (2005) argued that even a quickly degradable drug can act as a persistent chemical. Moore et al. (2008) state that if caffeine is profusely discharged from anthropogenic sources into an environment, it could constantly replenish levels regardless the amount degraded, creating a dynamic equilibrium.

Assuming that caffeine is consumed according to a first order reaction ($C = C_0 e^{-kx/U}$), figure A.3 presents an estimation of caffeine levels along the Iguaçu river. Discharge at the different reaches is estimated through the Manning equation; Inputs due to domestic waste are calculated using the population data, considering a per capita contribution of 16 mg/person.day, and efficiency removal in the wastewater treatment plant of 90%; three values of half-life (t_{50}) are arbitrated (i) 5h, (ii) 7h and (iii) 12h, and decay rate (k) is estimated through the equation $k = 0.693/t_{50}$ (Chapra, 1997).
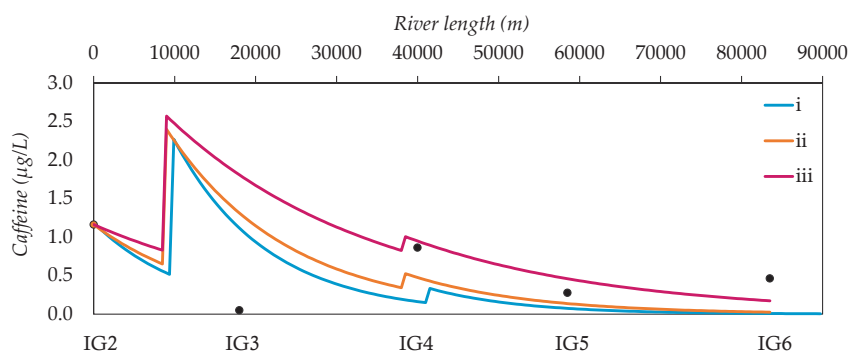
Figure A.3: Caffeine levels along the Iguaçu river (i) $k = 3.33\ d^{-1}$, (ii) $2.38\ d^{-1}$, (iii) $1.40\ d^{-1}$; data is from monitoring campaign in 06/06/2016 (black dots)

References: BRADLEY, P. M.; BARBER, L. B.; KOLPIN, D. W.; MCMAHON, P. B.; CHAPELLE, F. H. Biotransformation of caffeine , cotinine , and nicotine in stream sediments: implications for use as wastewater indicators. **Environmental Toxicology and Chemistry**, v. 26, n. 6, p. 1116–1121, 2007.

BUERGE, I. J.; POIGER, T.; MU, M. D.; BUSER, H. Combined Sewer Overflows to Surface Waters Detected by the Anthropogenic Marker Caffeine. **Environmental Science and Technology**, v. 40, n. 13, p. 4096–4102, 2006.

BUSSE, L.; NAGODA, C. **Detection of Caffeine in the Streams and Rivers within the San Diego Region - Pilot Study**. California Regional Water Quality Control Board San Diego Region. San Diego, 2015.

CANELA, M. C.; JARDIM, W. F.; SODRÉ, F. F.; GRASSI, M. T. **Cafeína em águas de abastecimento público no Brasil**. Instituto Nacional de Ciências e Tecnologias, Editora Cubo, São Carlos, 2014.

DOMBROSKI, L. F.; FERNANDES, C. V. S.; KNAPIK, H. G.; MUHLENHOFF, A. P.; FROEHNER, S. J. A importância da cafeína como parâmetro de qualidade de água para o gerenciamento de recursos hídricos. **Anais do 11º Simpósio de Hidráulica e Recursos Hídricos dos Países de Língua Oficial Portuguesa**, v. 1. p. 1–12, 2013.

MOORE, M. T.; GREENWAY, S. L.; FARRIS, J. L.; GUERRA, B. Assessing Caffeine as an Emerging Environmental Concern Using Conventional Approaches. **Archives of Environmental Contamination and Toxicology**, p. 31–35, 2008.

SEILER, R. L.; ZAUGG, S. D.; THOMAS, J. M.; HOWCROFT, D. L. Caffeine and pharmaceuticals as indicators of wastewater contamination in wells. **Groundwater**, v. 37, n. 3, p. 405-410, 1999.

THOMAS, P. M.; FOSTER, G. D. Tracking acidic pharmaceuticals, caffeine , and triclosan through the wastewater treatment process. **Environmental Toxicology and Chemistry**, v. 24, n. 1, p. 25–30, 2005.

A.5 Fortran code tests

In figures A.4 and A.5 are presented some preliminary results. It can be observed that are a few differences, that may arise from precision aspects. Mean absolute difference for discharges between Fortran and Matlab range from 0.53 m^3/s (IG6) to 2.83 m^3/s (IG5). For BOD concentrations, the mean absolute difference ranges from 0.05 mg/L in IG2 to 0.61 mg/L in IG4.

For the simulation of one year (2010), Matlab solves the equations in approximately 2 min, while Fortran uses 45s, using the same equations construction.
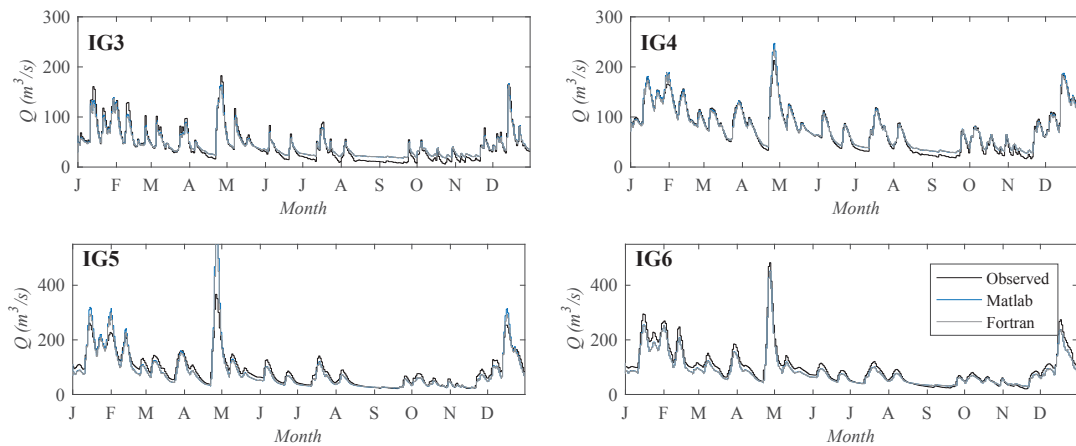
Figure A.4: Comparison of observed discharges and simulated with Fontran and Matlab
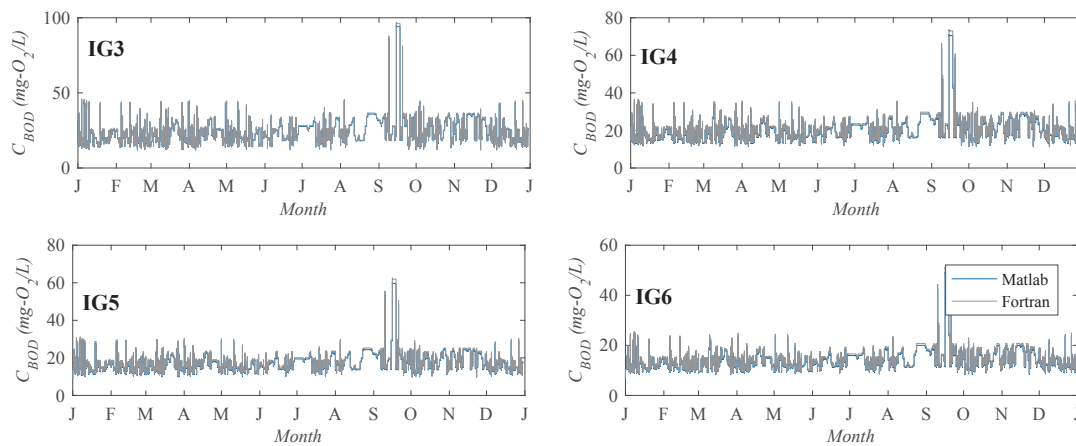
Figure A.5: Comparison of simulated BOD concentrations with Fontran and Matlab

A.6 Terms magnitude of the governing equations

This section presents means values over space of each term in the Momentum Conservation and Advection-Dispersion-Reaction equations, solved for the year 2010, investigating the role of the different components in overall results (figures A.6 and A.7). Simulations consider the BOD parameter, with the proposed calibration strategy (kinetic rate varies with time and space).

Table A.2: Terms of Saint-Venant and Advection-Dispersion-Reaction

	Saint-Venant	**Advection-Dispersion-Reaction**
Term 1	$U_i^{k+1} - \alpha U_i^k - (1-\alpha)U^*$	$C_i^{k+1} - C_i^k$
Term 2	$\dfrac{U^*\Delta t}{2\Delta x}\left(U_{i+1}^k - U_{i-1}^k\right)$	$\dfrac{U_i^k \Delta t}{2\Delta x}\left(C_{i+1}^k - C_{i-1}^k\right)$
Term 3	$\dfrac{g\Delta t}{2\Delta x}\left(y_{i+1}^k - y_{i-1}^k\right)$	$\dfrac{D\Delta t}{A_i^k}\dfrac{A_{i+1}^k - A_{i-1}^k}{2\Delta x}\dfrac{C_{i+1}^k - C_{i-1}^k}{2\Delta x} +$ $\dfrac{D\Delta t}{(\Delta x)^2}\left(C_{i+1}^k - 2C_i^k + C_{i-1}^k\right)$
Term 4	$\dfrac{q_i^k\left(v_L - U^*\right)}{A^*}\Delta t$	$C_i^k(-k)\Delta t$
Term 5	$g\Delta t\left(S_0 - S_f^*\right)$	$\dfrac{W_i}{V}\Delta t$
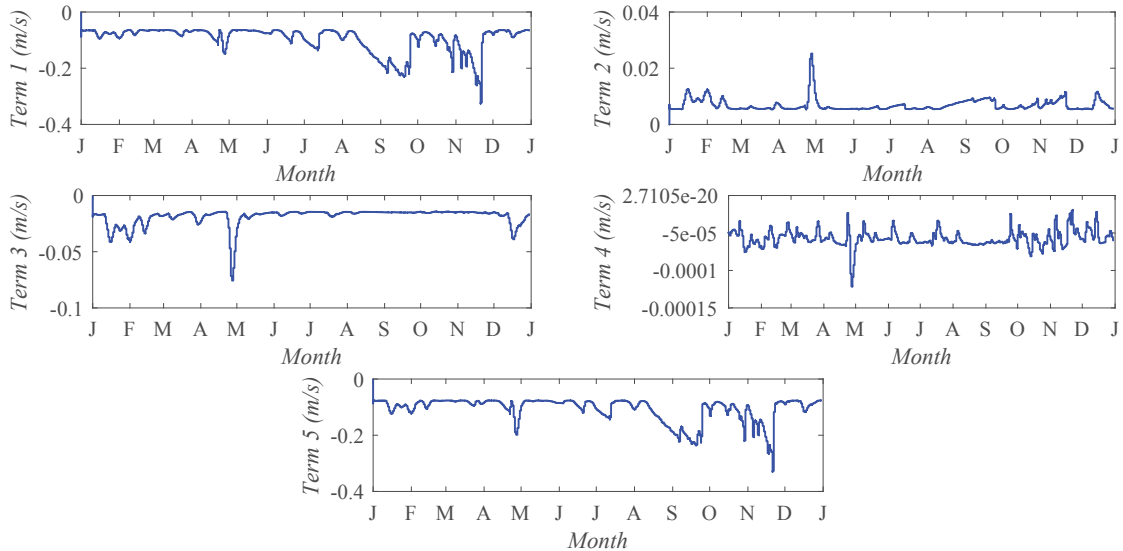
Figure A.6: Magnitude of terms in momentum conservation – simulation 2010
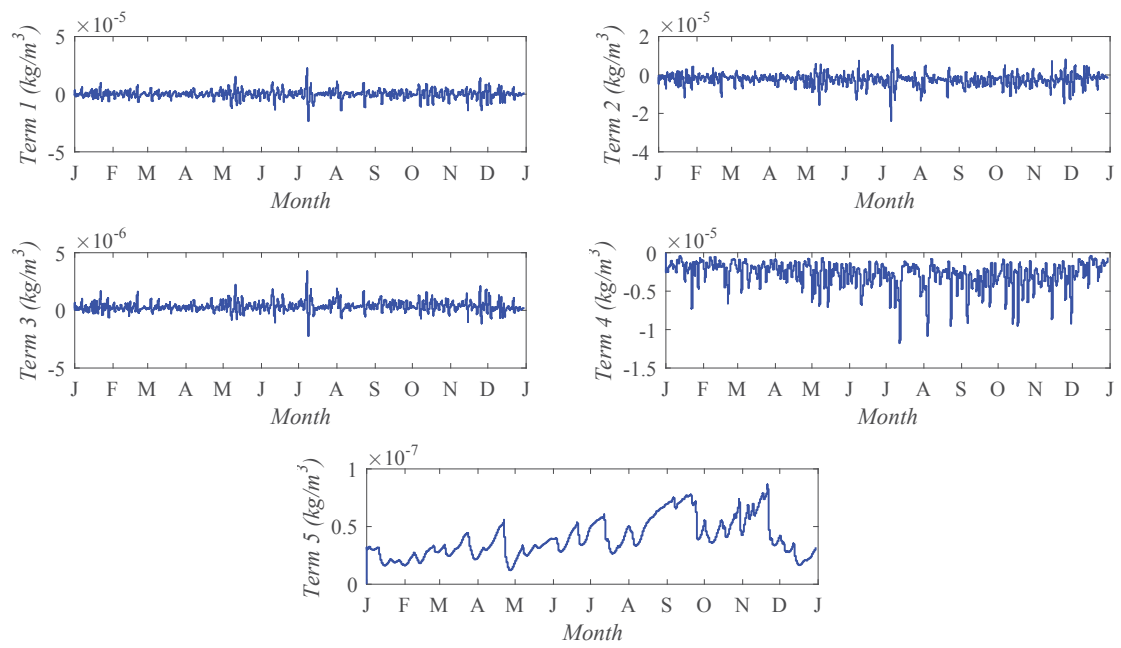
Figure A.7: Magnitude of terms in Advection-Dispersion-Reaction – BOD simulation for 2010

A.7 Verification of Saint-Venant equations solution

Figure A.8 shows modeling results using the HEC-RAS software, using the same configurations that SIHQUAL (as presented in sections 3.1 and 5.1). Table A.3 compares other simulation conditions. More details in Ferreira *et al.* (2017)[6].
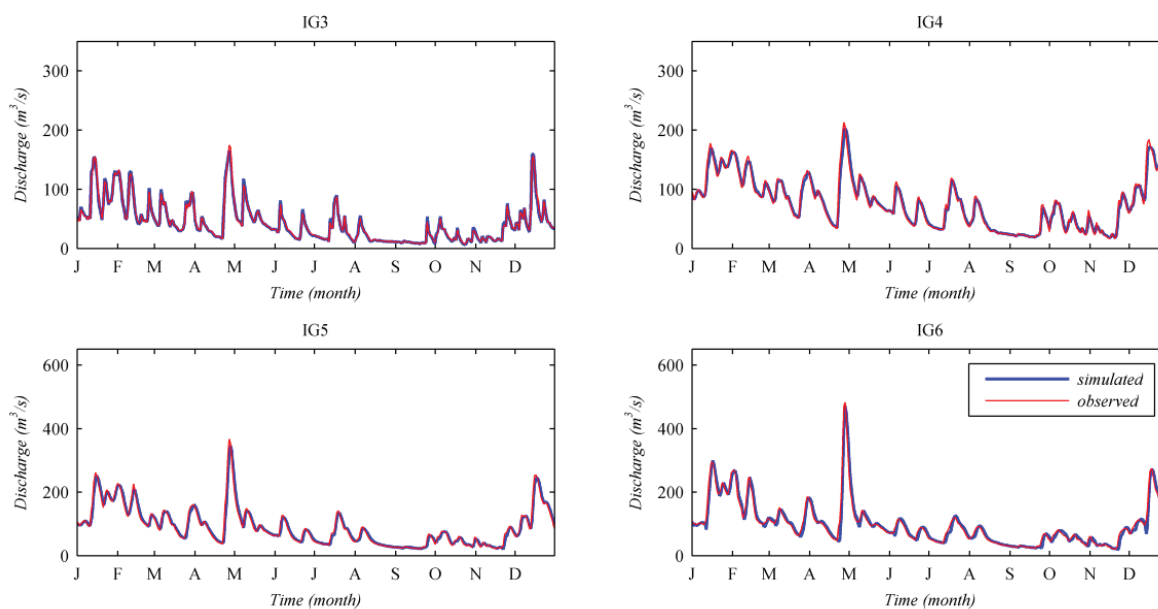
(a)   Measured and simulated discharges for 2010 (calibration) - HEC-RAS solution
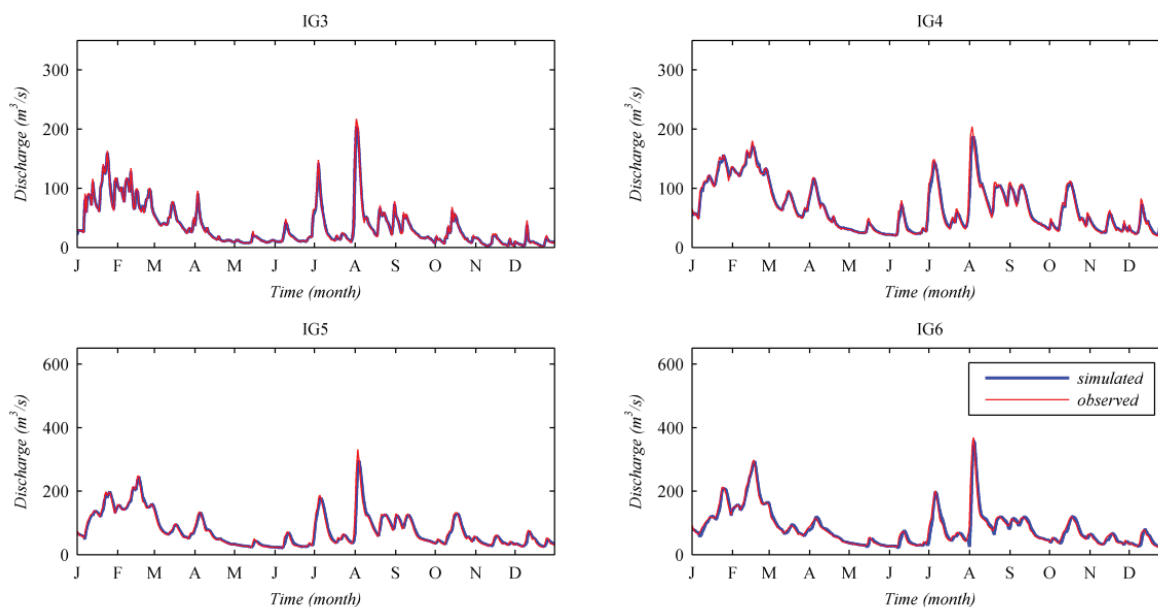
(b)   Measured and simulated discharges for 2011 (verification) - HEC-RAS solution

Figure A.8: Simulated discharges with HEC-RAS

[6]FERREIRA, D. M.; FERNANDES, C. V. S.; GOMES, J. Verification of Saint-Venant equations solution based on the lax diffusive method for flow routing in natural channels. **Brazilian Journal of Water Resources**, v. 22, 2017

Table A.3: Nash-Sutcliffe (E_{ns}) coefficients for different tests

Tests	Lax solution	HEC-RAS solution (E_{ns})			
		IG3	IG4	IG5	IG6
Natural section	–	0.997	0.959	0.957	0.939
$\varDelta t = 30$ min	Unstable	0.999	0.976	0.971	0.955
$\varDelta t = 10$ min	Unstable	0.999	0.976	0.971	0.955
$\varDelta t = 5$ min	Unstable	0.999	0.977	0.971	0.955

A.8   Synthetic series: validation

In order to validate the methodology of AR(1) modeling and the multiple solution analysis, other study cases are investigated (i. and ii.). Results indicate that the strategy is robust and can be applied for different systems, parameters and data availability.

(i.) Series of dissolved oxygen in Wyoming, United States: station USGS 09258980 Muddy Creek. The simulations are performed based on daily monitoring data, from 09/16/2011 to 11/14/2011. Conditions for the test are: mean and standard deviation of entire dataset, $\rho = 0.98$ (correlation of the series). Results are presented in table A.4 and figure A.9, with options b and d generating best fit.

(ii.) Series of total nitrogen (N_{tot}) in Colorado, United States: station USGS 06714000 South Platte river. The simulations are performed based on infrequent monitoring data, from 01/01/2005 to 12/31/2007. Conditions for the test are: mean and standard deviation of entire dataset, $\rho = 0.8$ (arbitrated). Results are presented in table A.4 and figure A.9, with option b as best fit.

Table A.4: Differences between synthetic series and monitoring data in USA

Case study	Series[1]	RMSE (mg-O_2/L)	R	MAPE (%)	PBIAS (%)
	a	0.40	0.90	11.05	14.37
i.	b	0.53	0.95	3.82	-2.35
	d	0.53	0.95	3.82	-2.35
	e	0.56	0.94	3.93	-1.13
	a	0.33	-0.03	22.61	11.41
ii.	b	1.07	0.82	11.03	2.70
	d	1.43	0.68	12.12	1.37
	e	2.68	-0.20	24.19	4.40

[1] criterion a to d to select the series are those presented in section 3.2
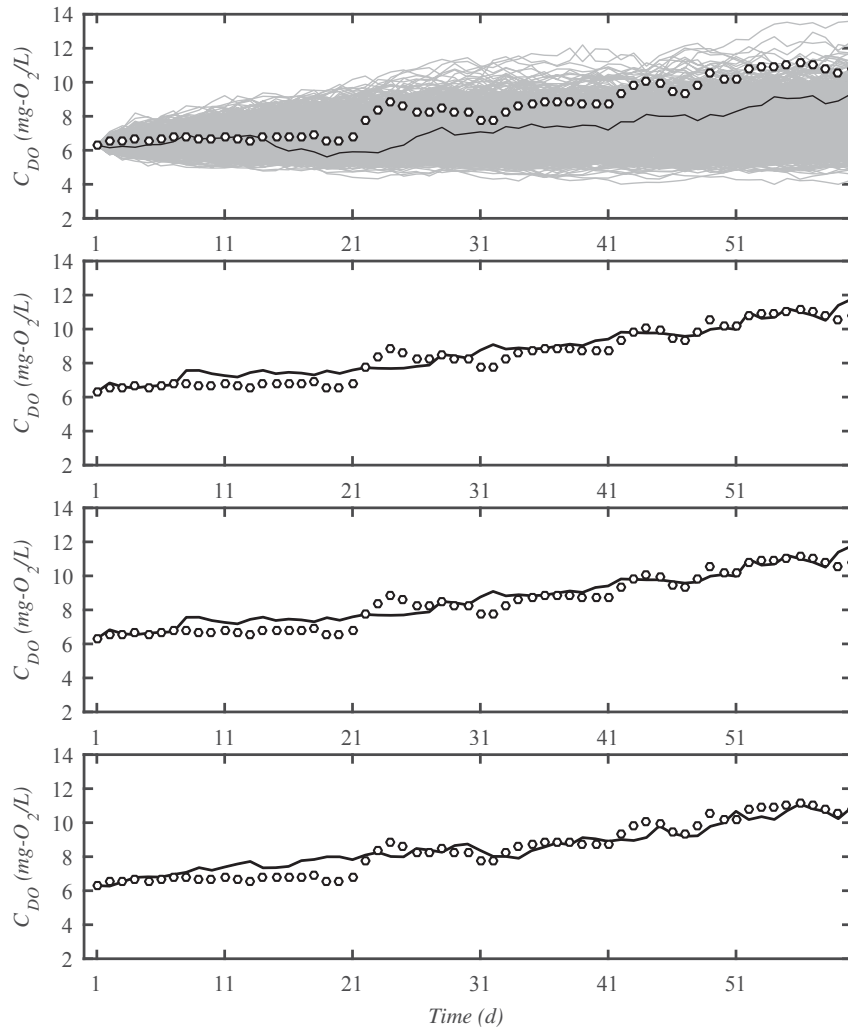
Figure A.9: Daily synthetic series (line) and monitored data (circles) in USA; series *a*, *b*, *c* and *d*, respectively; shaded area in *a* indicates the thousand options
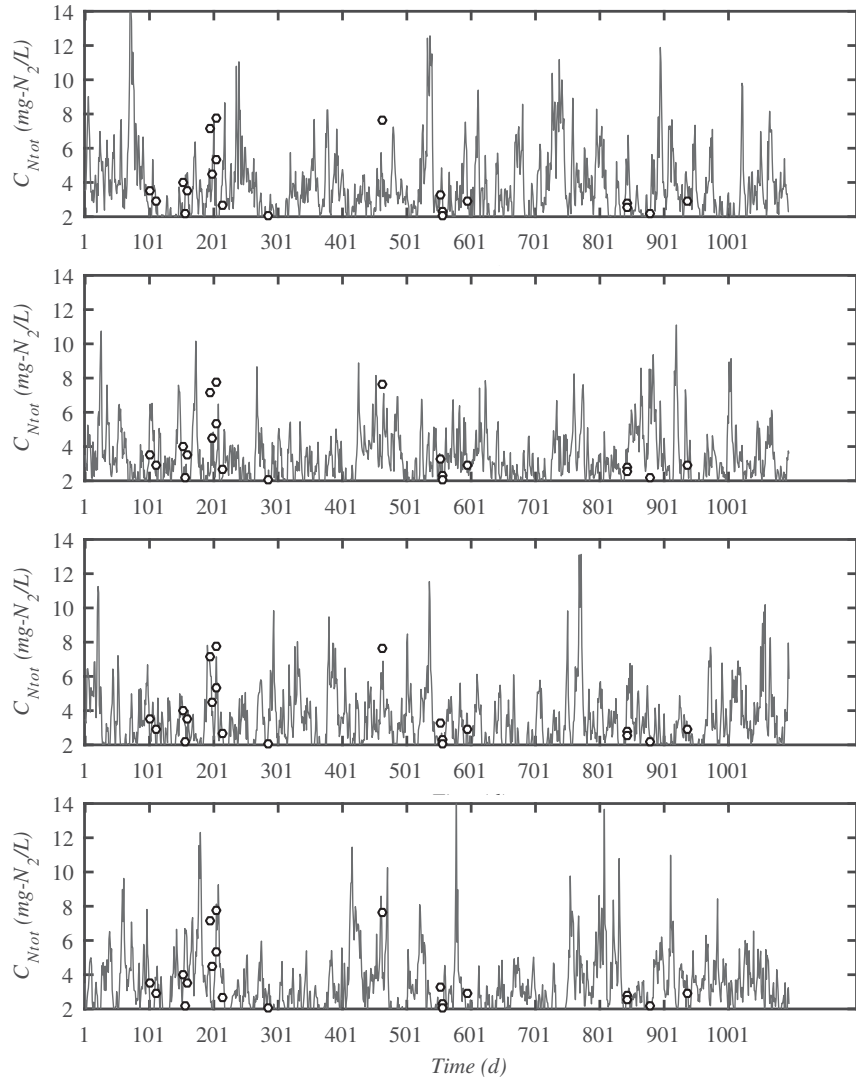
Figure A.9: Daily synthetic series (line) and monitored data (circles) in USA; series *a*, *b*, *c* and *d*, respectively (continued)

A.9 Water quality simulations with alternative boundary condition

Simulations with series T11b and T12b as boundary conditions are presented, since these tests consider other distribution in the autoregressive equation – two parameter log-normal (figure A.10).
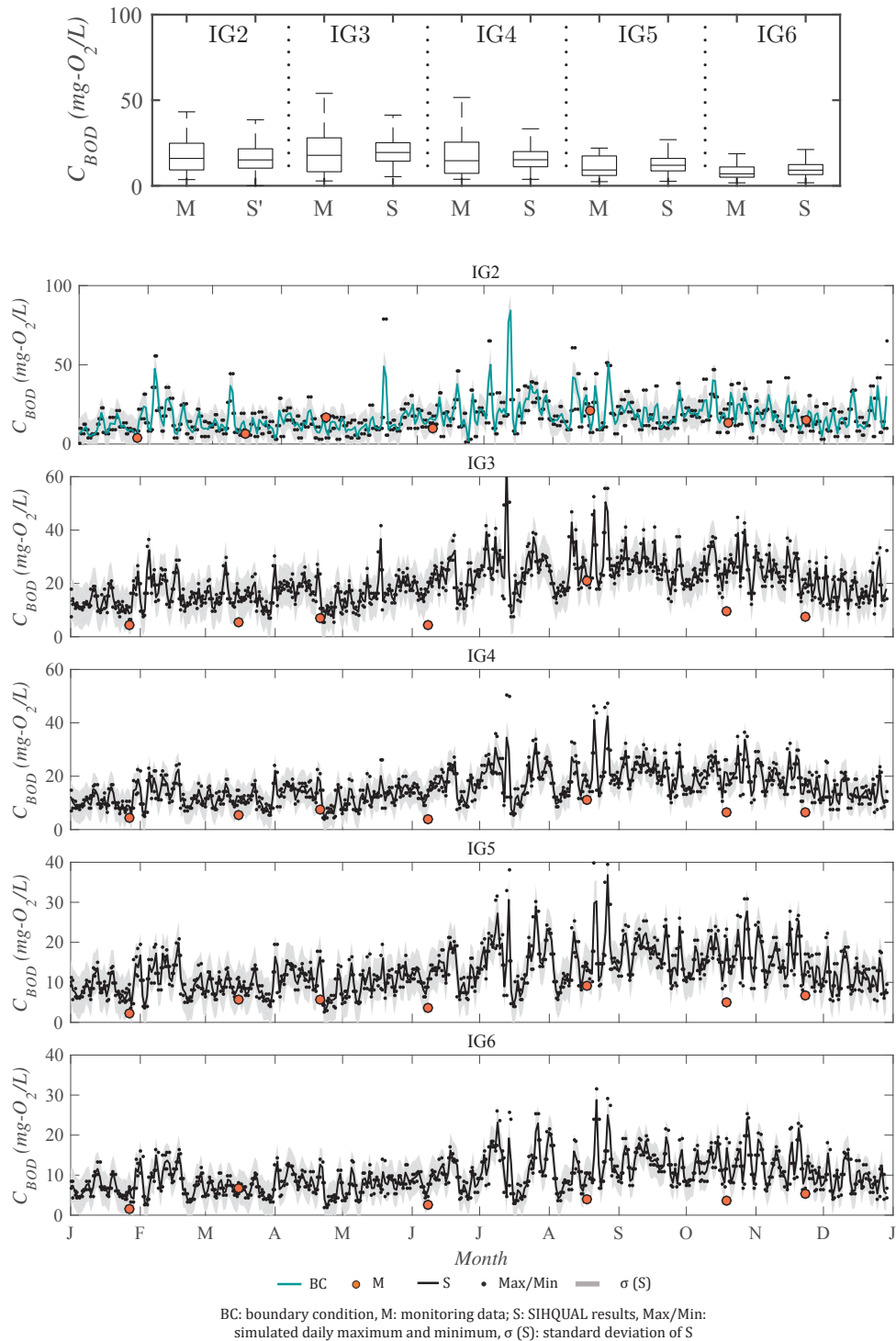
Figure A.10: BOD simulations with hydrid AR(1) as BC – T11b
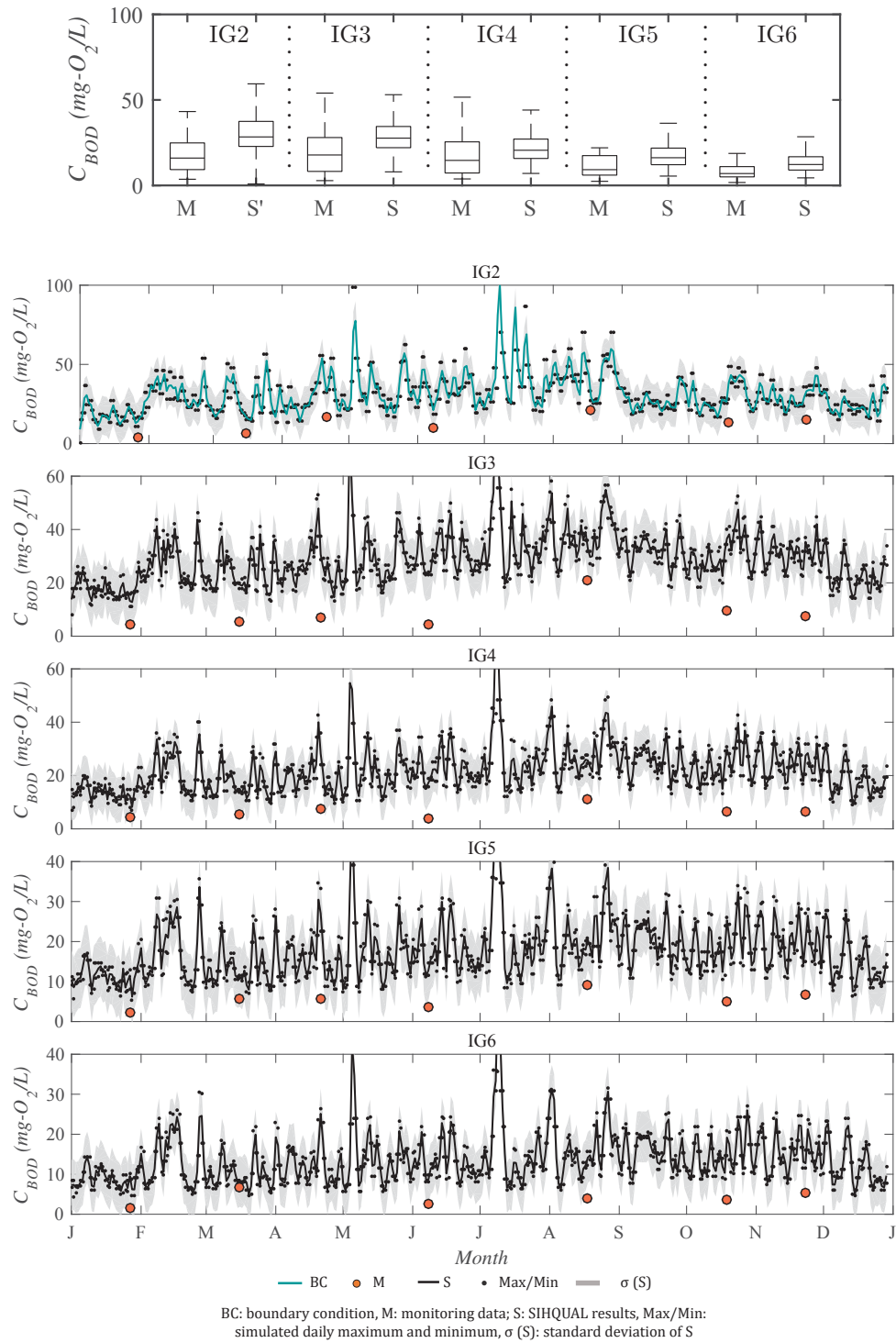
Figure A.10: BOD simulations with hydrid AR(1) as BC – T12b (continued)

A.10 Sensibility analysis

In the hydrodynamic module, the parameter assessed is the Manning coefficient. For the water quality module, simulations with different kinetic parameters and external loads are performed; the results are compared using a sensibility coefficient proposed by Lenhart et al. (2002) and a normalized analysis.

The dependence of a variable y of a parameter x can be expressed mathematically by the partial derivative $\partial y/\partial x$, approximated by a finite difference scheme. The result provides a coefficient of sensitivity I of y to x, which in its normalized form is presented as:

$$I = \frac{(y_2 - y_1)/y_0}{2\Delta x/x_0} \tag{A.8}$$

where y_0 represents the model output (simulated concentrations), result of the initial parameter x_0 (original kinetic rates); the initial parameter is varied by Δx, yielding $x_1 = x_0 - \Delta x$ and $x_2 = x_0 + \Delta x$.

Table A.5 describes the sensibility classes suggested by Lenhart et al. (2002), according to absolute values of I.

Table A.5: Sensibility classes

Class	Index	Sensibility
I	$0.00 \leq \mid I \mid < 0.05$	Small to negligible
II	$0.05 \leq \mid I \mid < 0.20$	Medium
III	$0.20 \leq \mid I \mid < 1.00$	High
IV	$\mid I \mid \geq 1.00$	Very high

Source: Lenhart et al. (2002)

Hydrodynamic module

Table A.6 and figure A.11 presents the values of E_{ns} when varying the Manning coefficient, considering the difference between simulated and observed discharges. The simulations are performed for 2010, with level series as boundary condition (results presented in figure 21). Results show that the discharges in station IG5 is the most sensitive one, probably due to the equation that represents the rating curve, as discussed in section 5.1.

Table A.6: Sensibility tests for hydrodynamic simulations

IG3		IG4		IG5		IG6	
n	E_{ns}	n	E_{ns}	n	E_{ns}	n	E_{ns}
0.040	0.919	0.045	0.786	0.050	-0.084	0.055	0.978
0.030	0.940	0.035	0.974	0.040	0.738	0.045	0.935
0.020	0.790	0.025	0.798	0.030	0.707	0.035	0.837

(*)The second line corresponds to calibrated values
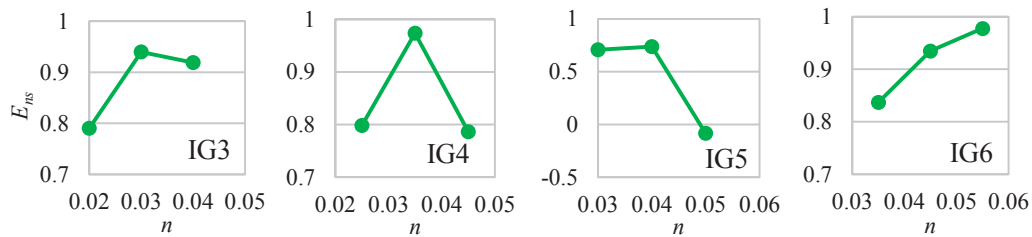
Figure A.11: Sensibility of hydrodynamic simulations to the Manning coefficient

Water quality module

For BOD simulation, different deoxygenation rates are initially tested, since this variable is usually more sensible for calibration (Kondageski and Fernandes, 2009). For N-org, the K_{so} rate is varied, since its range is wide – 0.001 to 0.10 d^{-1}, according to Von Sperling (2007). The analysis are based on simulations of the year 2010.

The results for computational experiments are presented in figure A.12 for BOD and figure A.13 for N-org. These simulation are produced with series T1a as boundary condition; the parameter of interest is the median of the time series produced (presented in section 5.3.1). The same calibration presented in section 5.3.1 are applied, though the kinetic rates are taken as means along the time (since constant values in time are required to calculate the I index). The boundary conditions are not affected by the variation of reaction coefficients, which explains the convergence in IG2 for the transient simulation.

The parameter K_d' is 50% larger and K_d'' is 50% smaller than the original value in each section (K_d). The analysis for N-org has the same configurations.

According to the classification by Lenhart et al. (2002), BOD simulations are highly affected by the K_d coefficient (figure A.12). The simulations of organic nitrogen, on the other hand, show a medium sensibility to K_{so} (figure A.13).
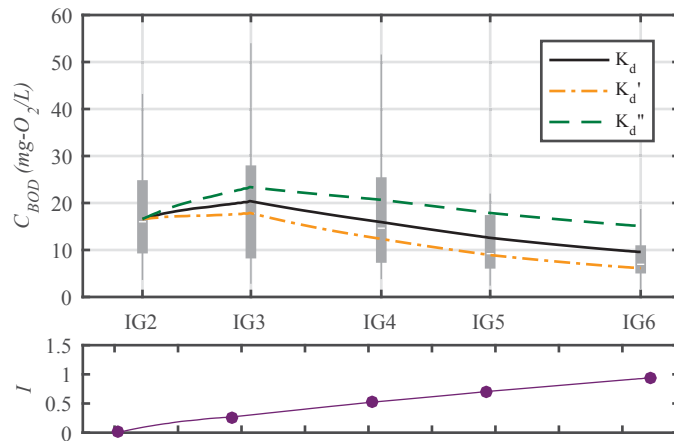
Figure A.12: Sensibility to K_d – BOD simulation 2010. Boxplots: set of monitored data, lines: median of simulated concentrations
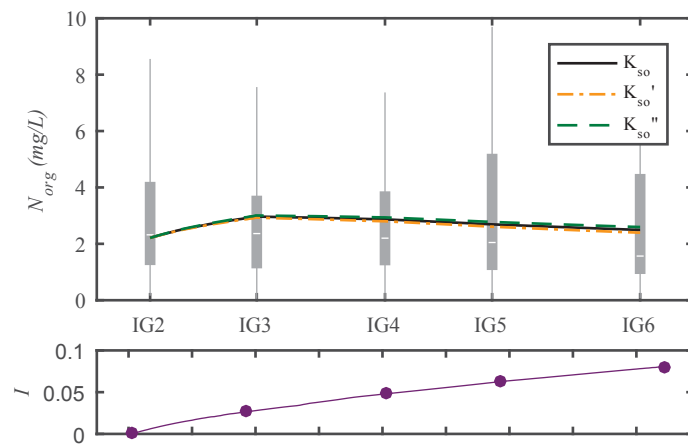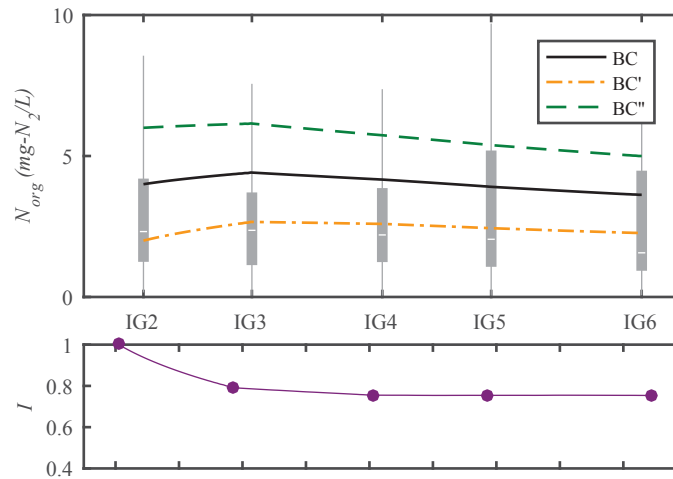

Figure A.13: Sensibility to K_{so} – N-org simulation 2010. Boxplots: set of monitored data, lines: median of simulated concentrations

The loads of each reach are set as: IG2-IG3 = 5000000 kg/d, IG3-IG4 = 250000 kg/d, IG4-IG5 = 85000 kg/d, IG5-IG6 = 60000 kg/d. These values are variated by 50 %, using the simulations in 2010. Figure A.14 shows the medians simulated compared with the boxplots of observed data, and coefficient I calculated. Considering the classification proposed by Lenhart et al. (2002), BOD simulations are highly sensible to external inputs, with a maximum coefficient I of 0.6.
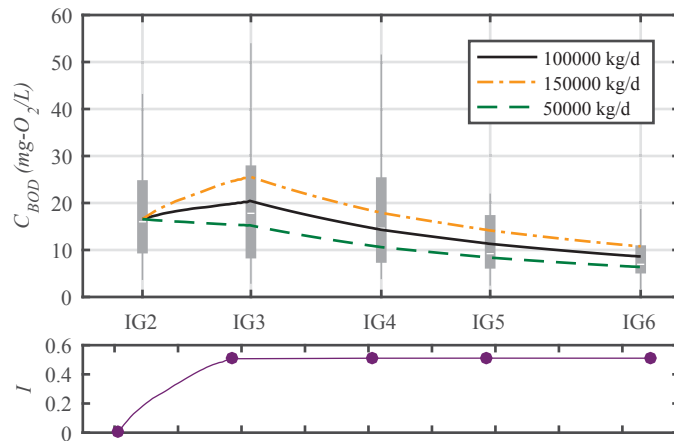
Figure A.14: Sensibility for loadings (BOD simulation – 2010); Boxplots: set of monitored data, lines: median of simulated concentrations

The same simulation conditions of section 5.4 are used to evaluate the sensibility of boundary conditions, comparing median concentrations. Results are presented in appendix A.10, giving support to confirm the role of this aspect in water quality modeling under unsteady state.
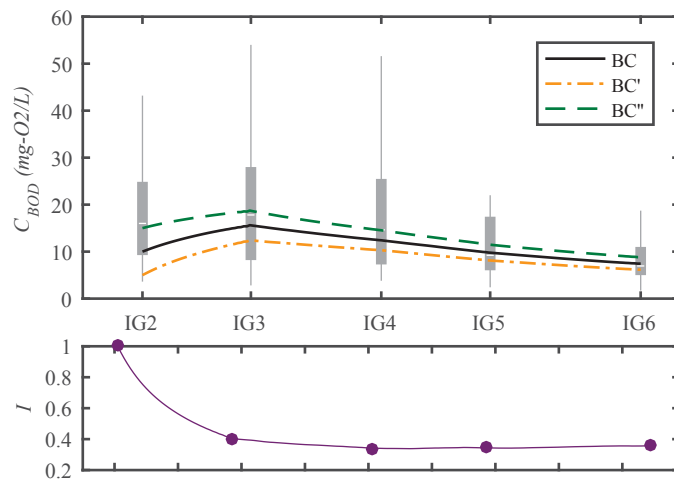
Figure A.15: Sensibility for BC (BOD simulation – 2010); Boxplots: set of monitored data, lines: median of simulated concentrations

Figure A.16: Sensibility for BC (N-org simulation – 2010); Boxplots: set of monitored data, lines: median of simulated concentrations
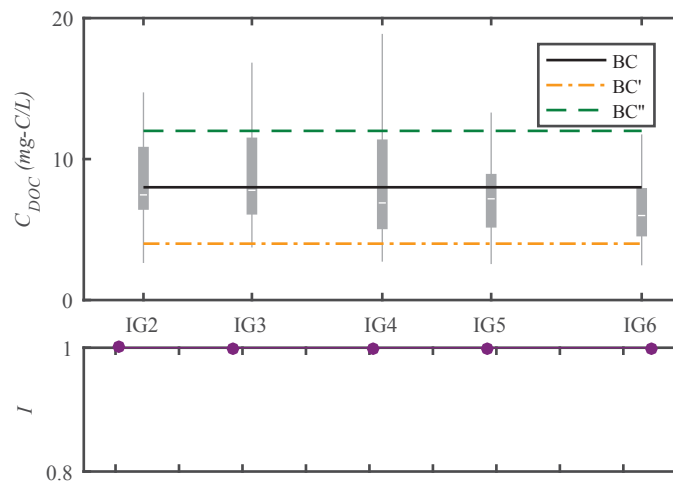
Figure A.17: Sensibility for BC (DOC simulation – 2010); Boxplots: set of monitored data, lines: median of simulated concentrations

Normalized analysis

In the normalized analysis, the increments for K_d and K_{so} values range from -0.2 to +0.6 (figures A.18 and A.19). The coefficient R_N is defined by the ratio between simulated concentration with the varied rate and the value obtained with the original reaction coefficient.
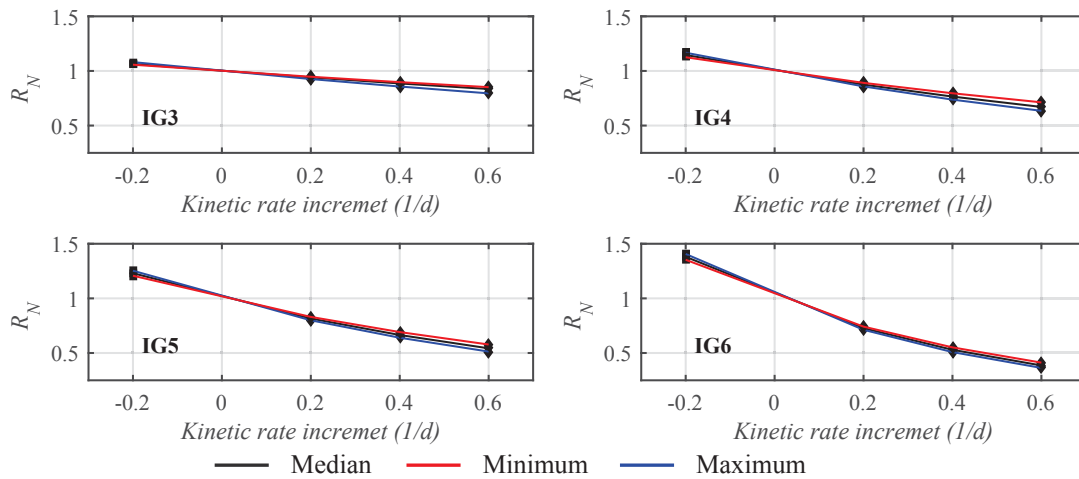
Figure A.18: Sensibility to K_d – BOD simulation 2010
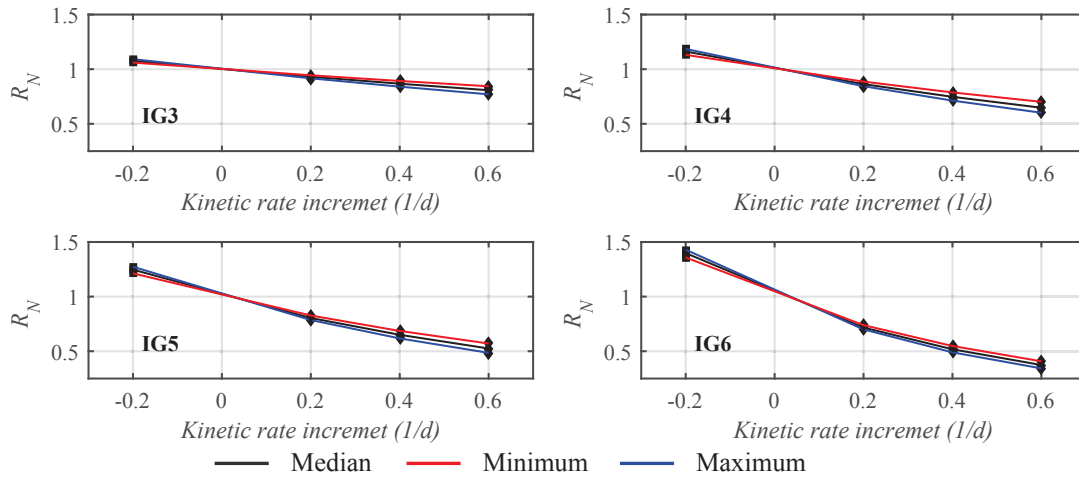
Figure A.19: Sensibility to K_{so} – N-org simulation 2010

Further investigations are suggested, since these analysis are only rough estimations – the tests consider traditional calibration strategy (constant rates), and are evaluated comparing only median simulation results.

A.11 Pollutographs resulting from calibration strategies
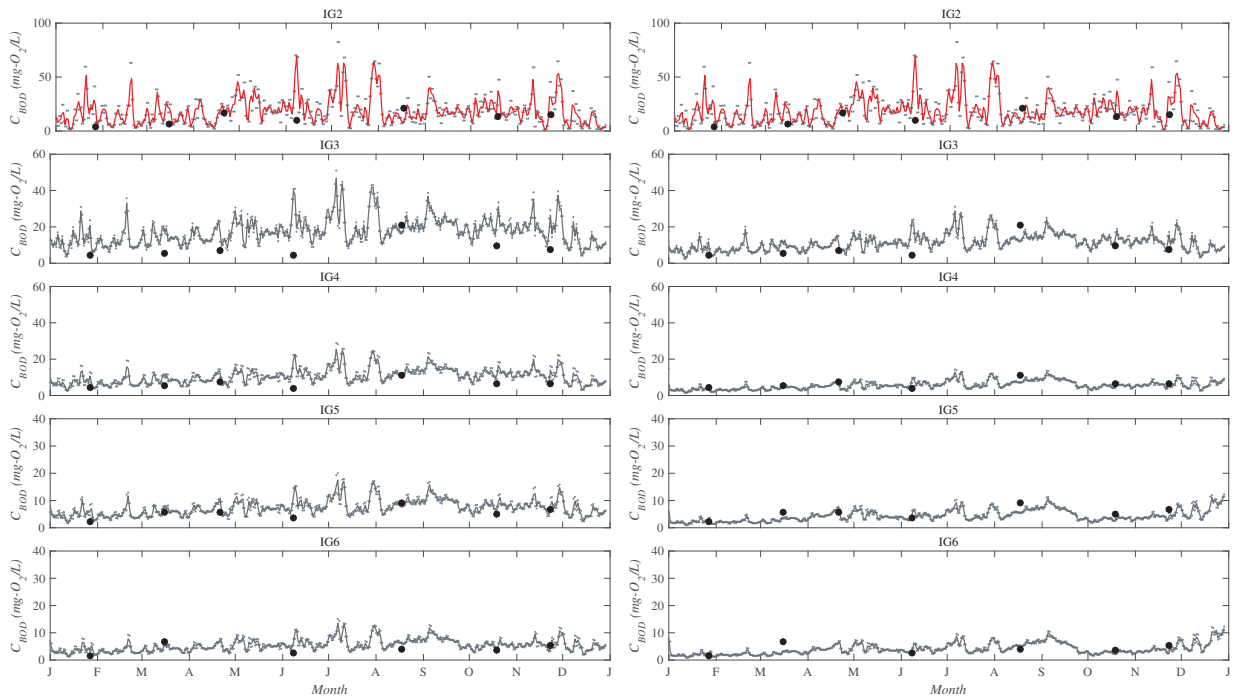
Figure A.20: (a) Daily BOD simulated concentrations and monitoring data for the year 2010 with T1a as BC and different calibration strategies – v1 (left) and v2 (right)
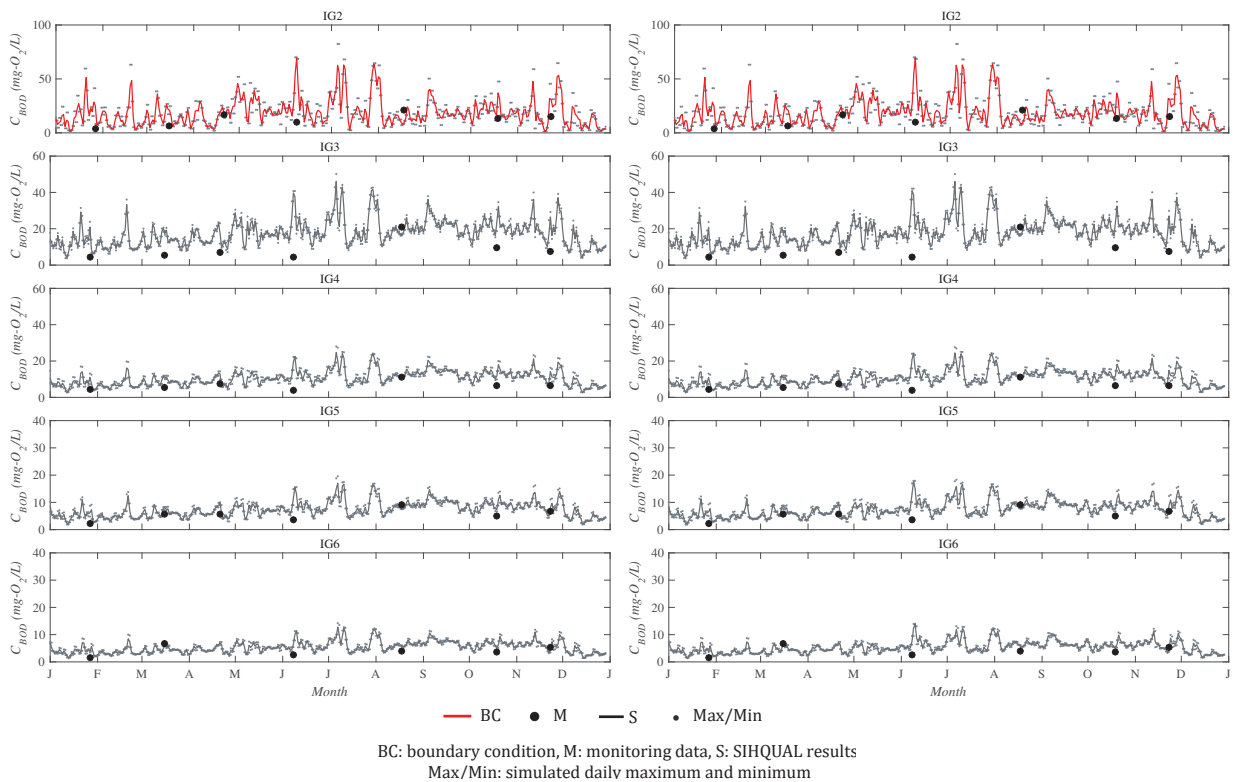
BC: boundary condition, M: monitoring data, S: SIHQUAL results
Max/Min: simulated daily maximum and minimum

Figure A.20: (a) Daily BOD simulated concentrations and monitoring data for the year 2010 with T1a as BC and different calibration strategies – v3 (left) and v4 (right) (continued)
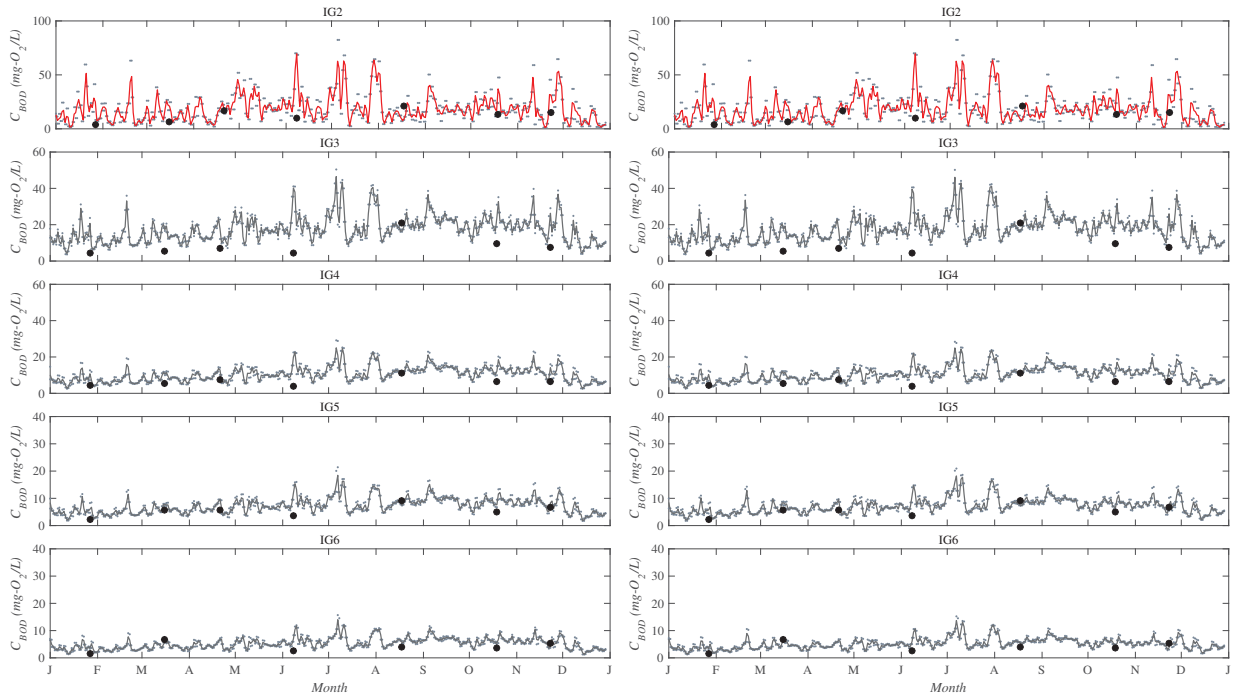
Figure A.20: (a) Daily BOD simulated concentrations and monitoring data for the year 2010 with T1a as BC and different calibration strategies – v5 (left) and v6 (right) (continued)
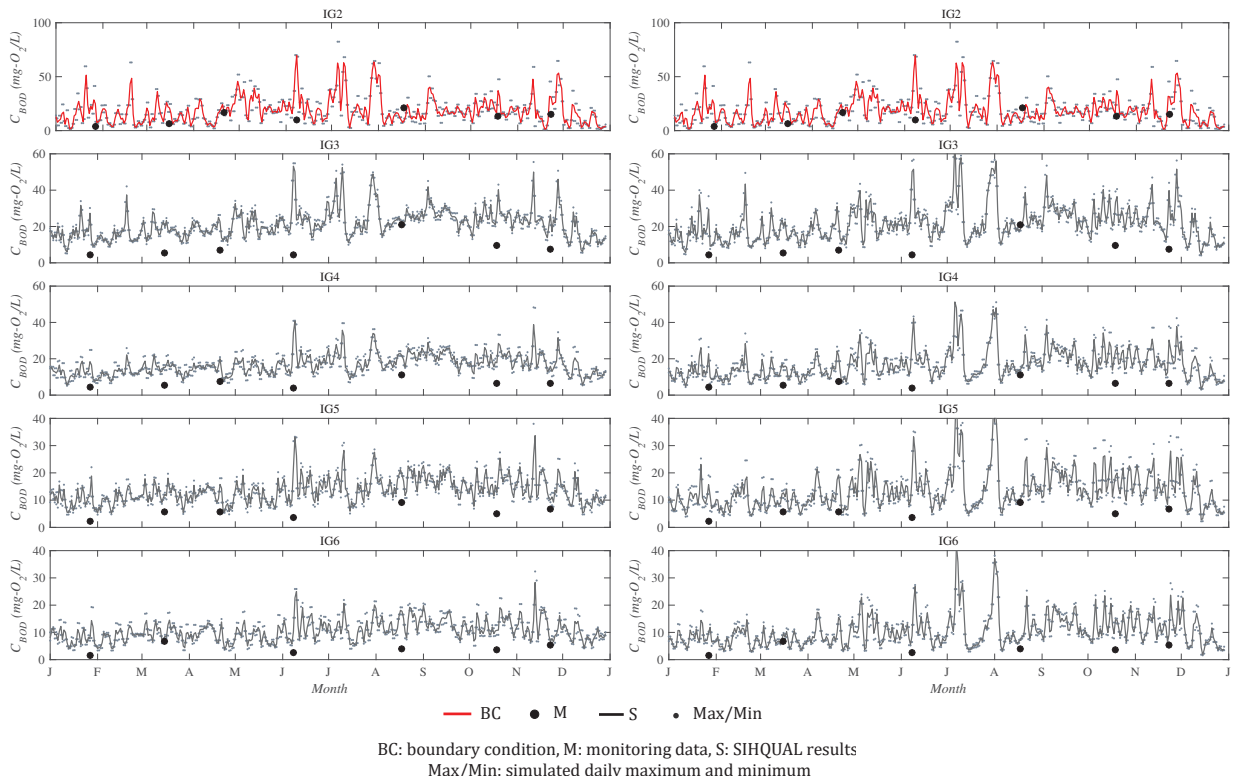
BC: boundary condition, M: monitoring data, S: SIHQUAL results
Max/Min: simulated daily maximum and minimum

Figure A.20: (a) Daily BOD simulated concentrations and monitoring data for the year 2010 with T1a as BC and different calibration strategies – v7 (left) and v8 (right)
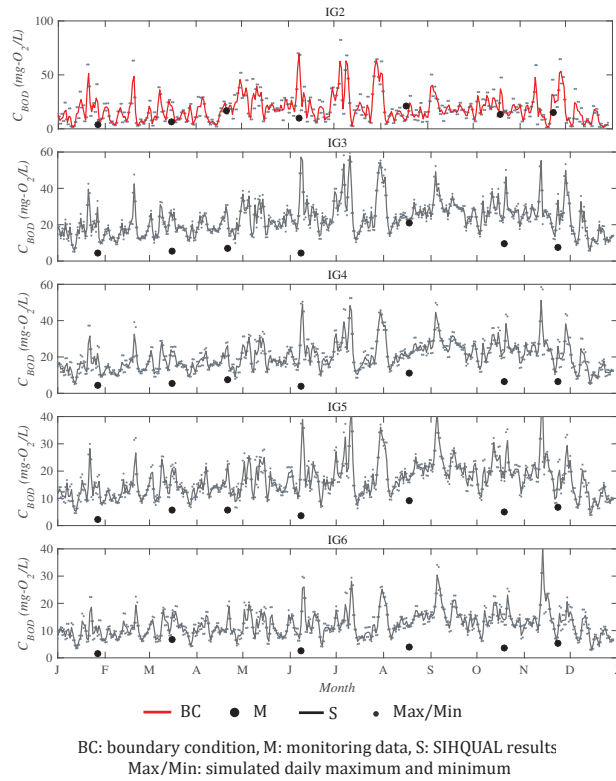
Figure A.20: (a) Daily BOD simulated concentrations and monitoring data for the year 2010 with T1a as BC and different calibration strategies – v9 (continued)

A.12   Sub-daily water quality boundary condition

This item shows simulations with sub-daily boundary conditions: T10a as concentration at each 50 s and series T9b as hourly data (figure A.21).
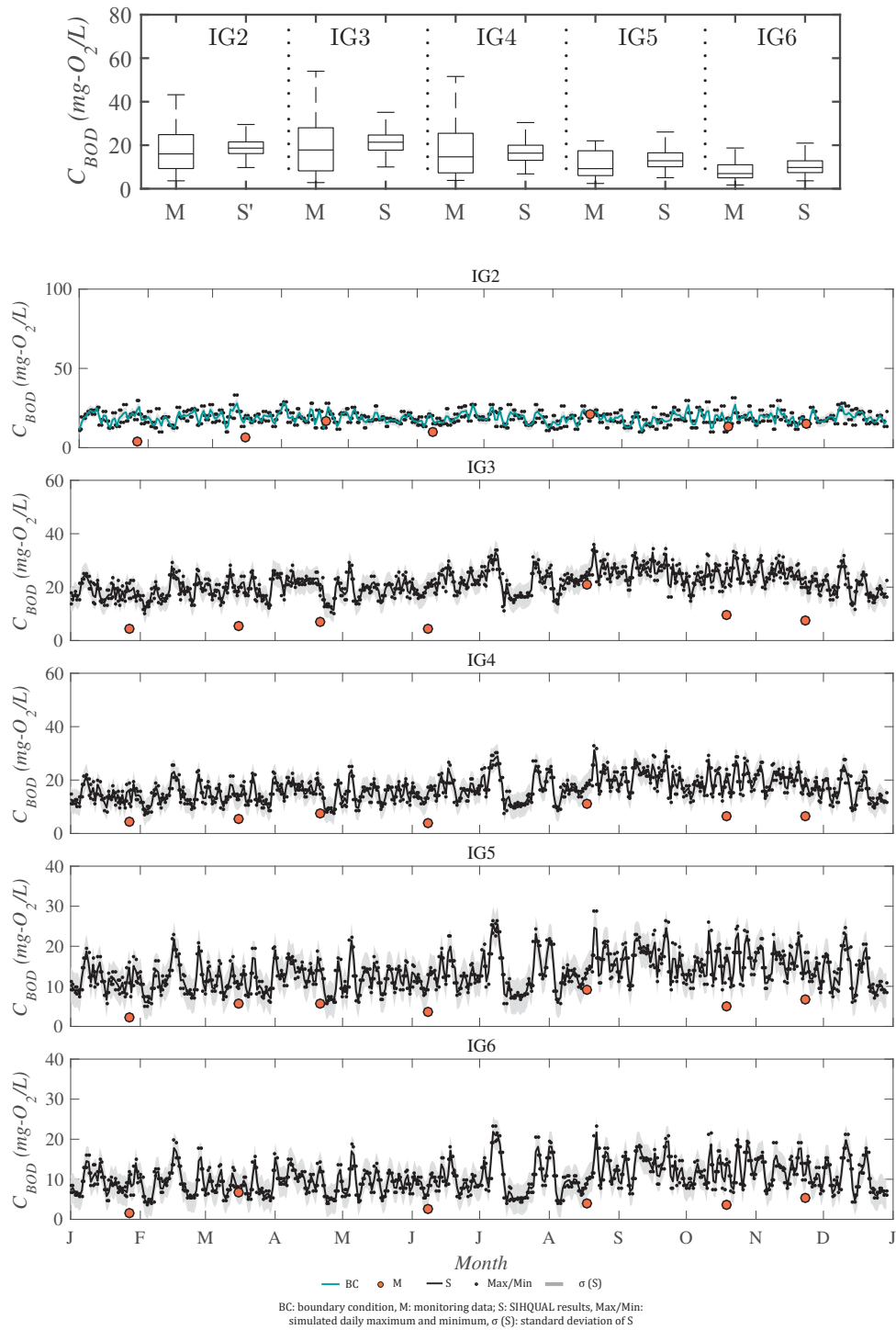
Figure A.21: Daily concentrations generated with different BC as sub-daily data – T10a
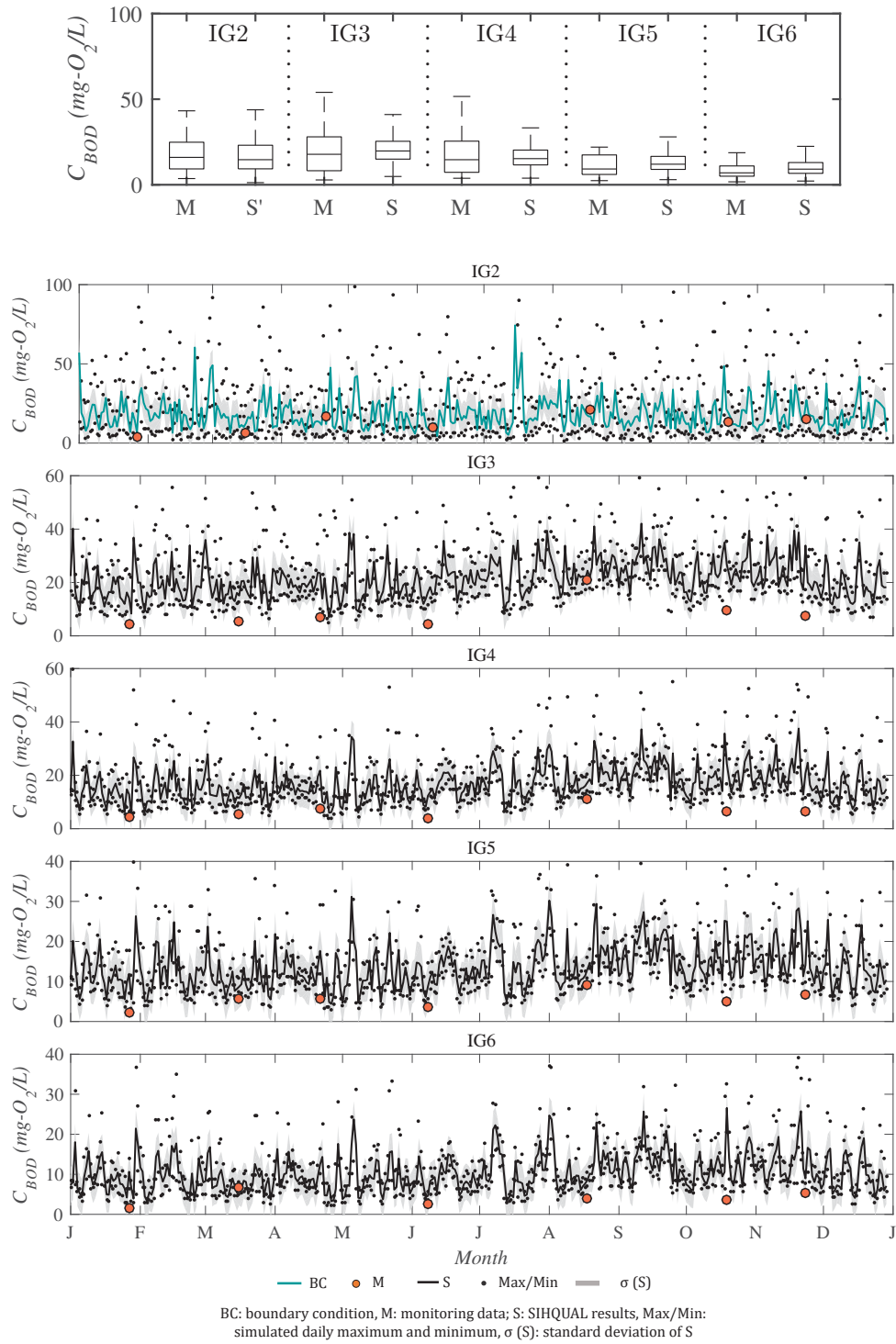
Figure A.21: Daily concentrations generated with different BC as sub-daily data – T9b hourly data (continued)

A.13 Validation of BOD modeling

The dataset for comparison in boxplots (figure A.23) and duration curves (figure A.24) is the same used in the simulations for 2010; pollutographs, however, are compared with monitored data in 2013-2015 (figure A.22). Lateral loads between IG2 and IG3 are overestimated, as identified in the steady simulation; therefore, the duration curve for steady state is built with data from 2010 simulation; flow simulations are those presented in section 5.1 for the period 2013-2015.

BC: boundary condition, M: monitoring data, S': synthetic series, S: SIHQUAL results,
Max/Min: simulated daily maximum and minimum, σ (S): standard deviation of S
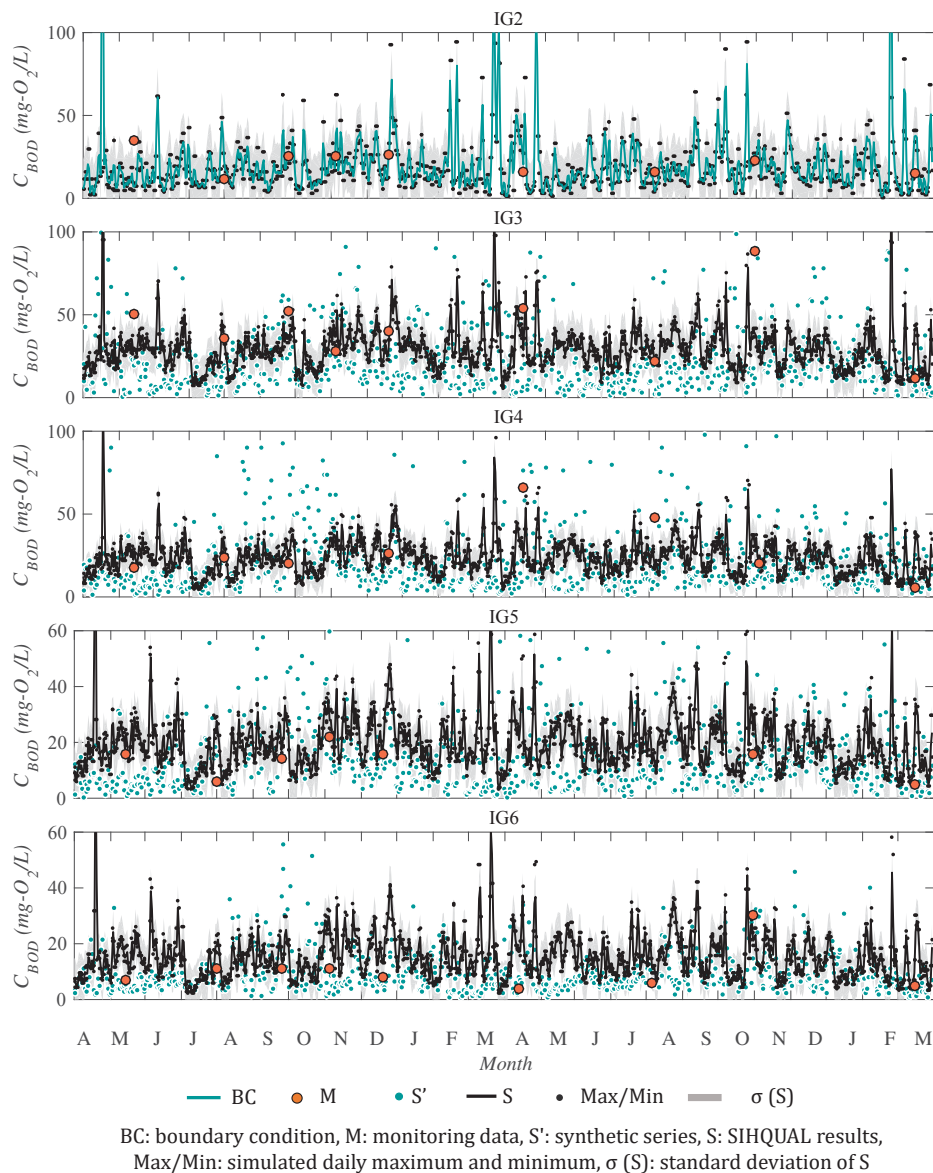
Figure A.22: Daily BOD simulated concentrations and monitoring data for the years 2013-2015; BC generated with test T1a
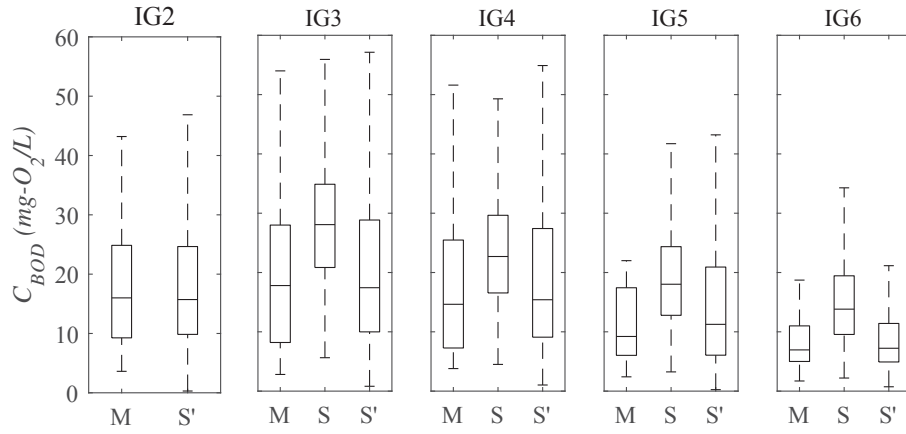
Figure A.23: Boxplot of measured (M) and simulated BOD concentrations for years 2013-2015 with SIHQUAL (S) and synthetic series (S'); T1a as BC
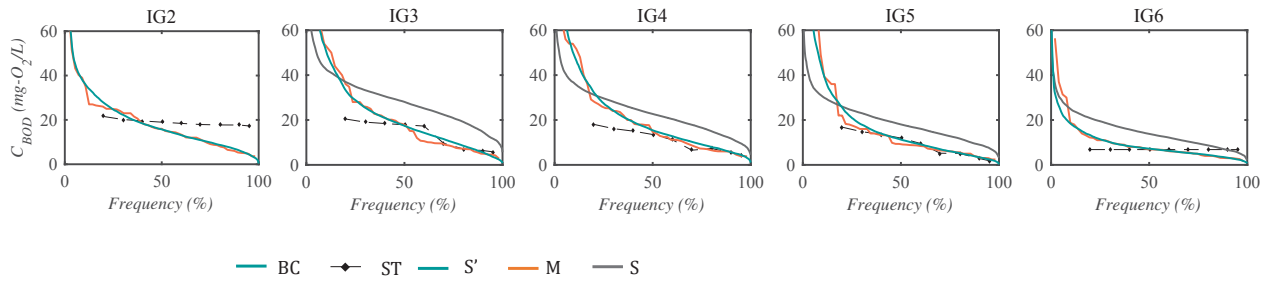
Figure A.24: Duration curves of historical monitoring dataset 2005-2017 (M) and simulations for 2013-2015 with synthetic series (S') and SIHQUAL results under unsteady (S) and steady (ST) state: T1a as BC

A.14 Labile and refractory carbon

Estimations of boundary condition and monitoring data to simulate labile and refractory carbon are the assumed fractions of the information available for POC and DOC. Monitoring data in pollutographs are POC/DOC values, and the error bars represent the attributed fraction of labile/refractory for each constituent; the monitoring data in boxplots and duration curves are only the attributed fraction.

Table A.7: Kinetic rates estimated in simulation for labile and refractory organic carbon

Interval of variarion	RPOC		LPOC		RDOC		LDOC	
	K_{L4} (d^{-1})		K_{L1} (d^{-1})		K_{L6} (d^{-1})		K_{L5} (d^{-1})	
int1	1.00 –	0.50	1.00 –	0.50	0.50 –	0.35	0.50 –	0.35
int2	0.50 –	0.10	0.50 –	0.10	0.35 –	0.10	0.35 –	0.10
int3	0.10 –	0.01	0.10 –	0.01	0.08 –	0.01	0.08 –	0.01
	K_{R1} (d^{-1})		K_{L2} (g/m^2.d)		K_{L7} (d^{-1})		K_{L7} (d^{-1})	
int1	1.00 –	0.50	0.10 –	0.06	0.50 –	0.35	0.50 –	0.35
int2	0.50 –	0.10	0.06 –	0.04	0.35 –	0.10	0.35 –	0.10
int3	0.10 –	0.01	0.04 –	0.01	0.08 –	0.01	0.08 –	0.01
	K_{R2} (g/m^2.d)		K_{L3} (d^{-1})		K_{R4} (d^{-1})		K_{L8} (d^{-1})	
int1	0.10 –	0.06	1.00 –	0.50	1.00 –	0.50	1.00 –	0.50
int2	0.06 –	0.04	0.50 –	0.10	0.50 –	0.10	0.50 –	0.10
int3	0.04 –	0.01	0.10 –	0.01	0.10 –	0.01	0.10 –	0.01
	K_{R3} (d^{-1})		K_{L4} (d^{-1})		K_{R5} (d^{-1})			
int1	1.00 –	0.50	1.00 –	0.50	1.00 –	0.50		
int2	0.50 –	0.10	0.50 –	0.10	0.50 –	0.10		
int3	0.10 –	0.01	0.10 –	0.01	0.10 –	0.01		
	K_{R4} (d^{-1})		K_{L5} (d^{-1})					
int1	1.00 –	0.50	0.50 –	0.35				
int2	0.50 –	0.10	0.35 –	0.10				
int3	0.10 –	0.01	0.08 –	0.01				
			K_{L6} (d^{-1})					
int1			0.50 –	0.35				
int2			0.35 –	0.10				
int3			0.08 –	0.01				
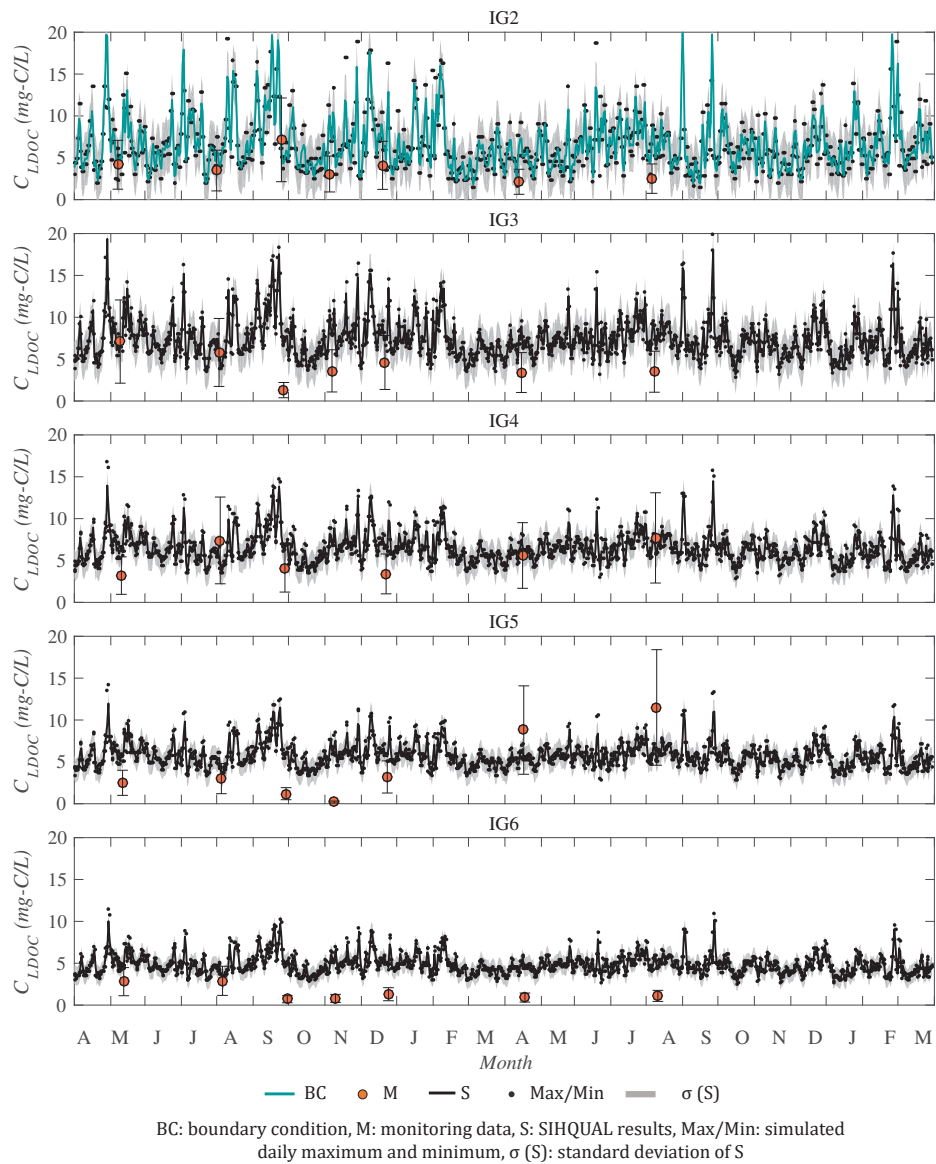
Figure A.25: Daily organic carbon fraction simulated for the years 2013-2015 – BC generated with test T2a: LDOC
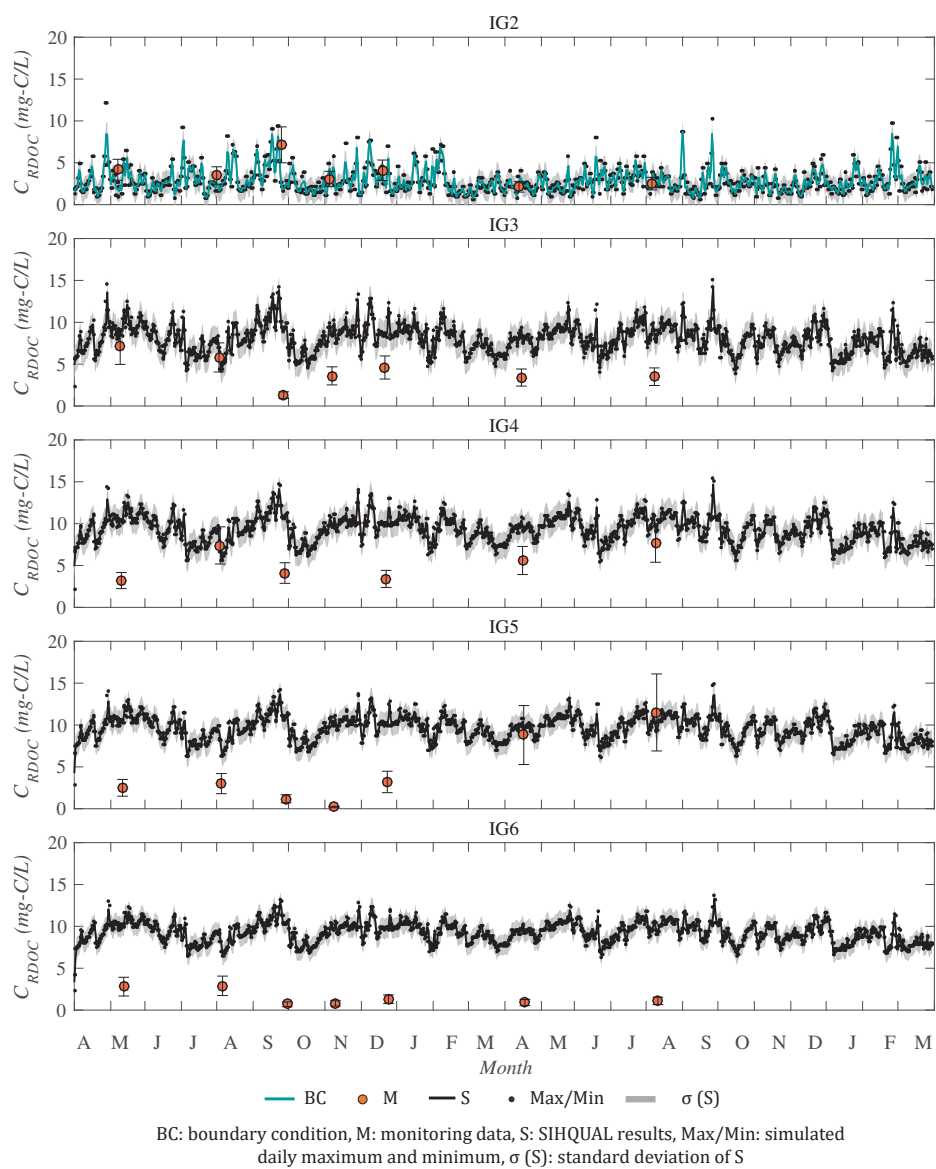
Figure A.25: Daily organic carbon fraction simulated for the years 2013-2015 – BC generated with test T2a: RDOC (continued)
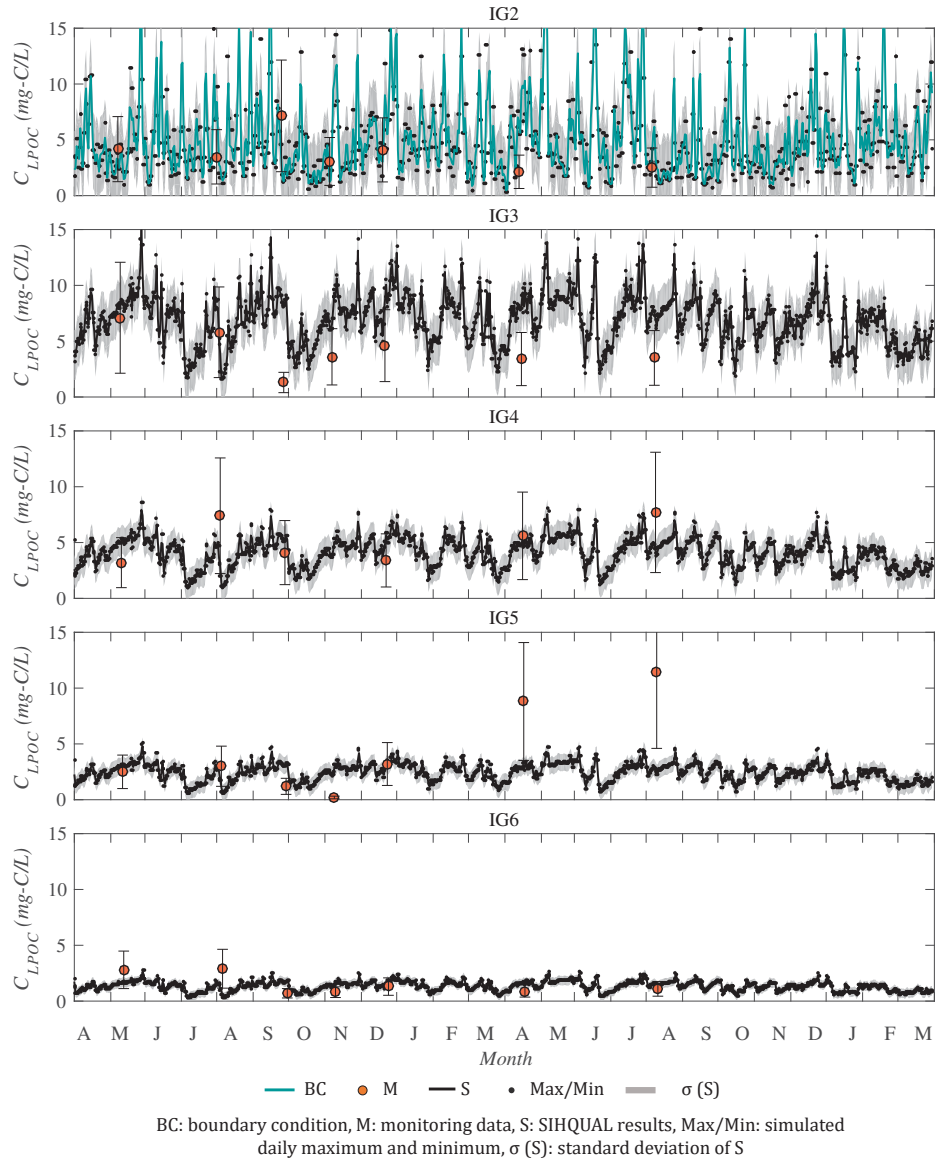
Figure A.25: Daily organic carbon fraction simulated for the years 2013-2015 – BC generated with test T2a: LPOC (continued)
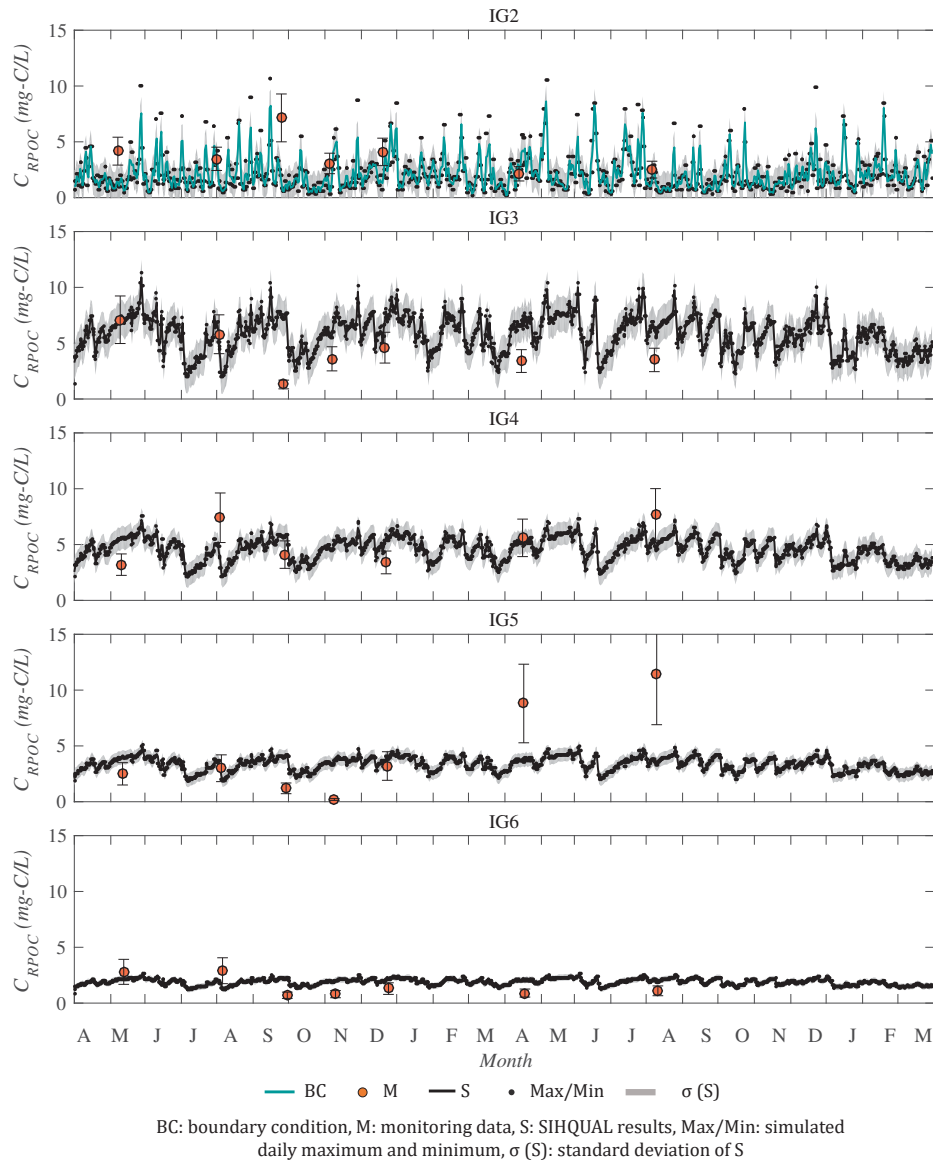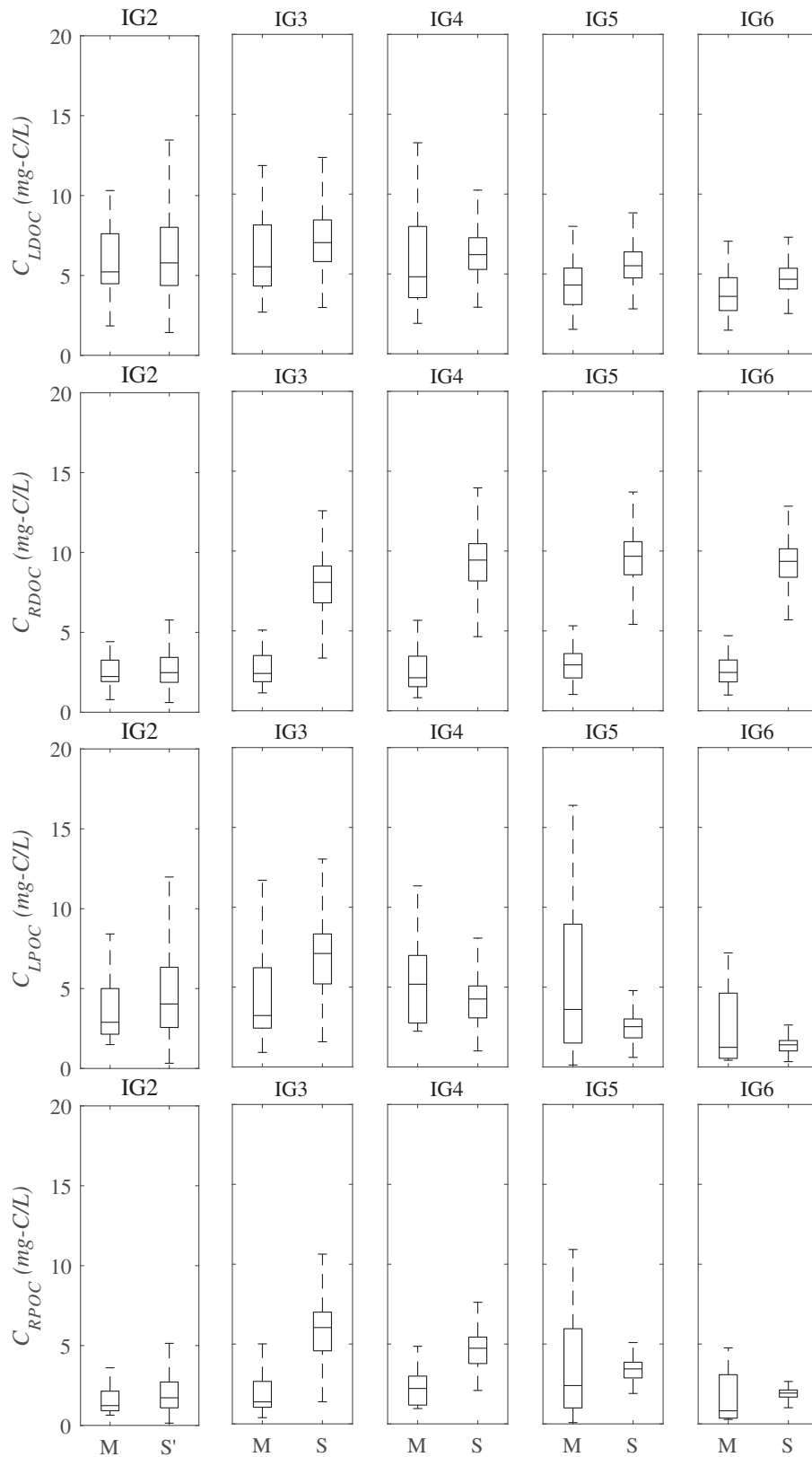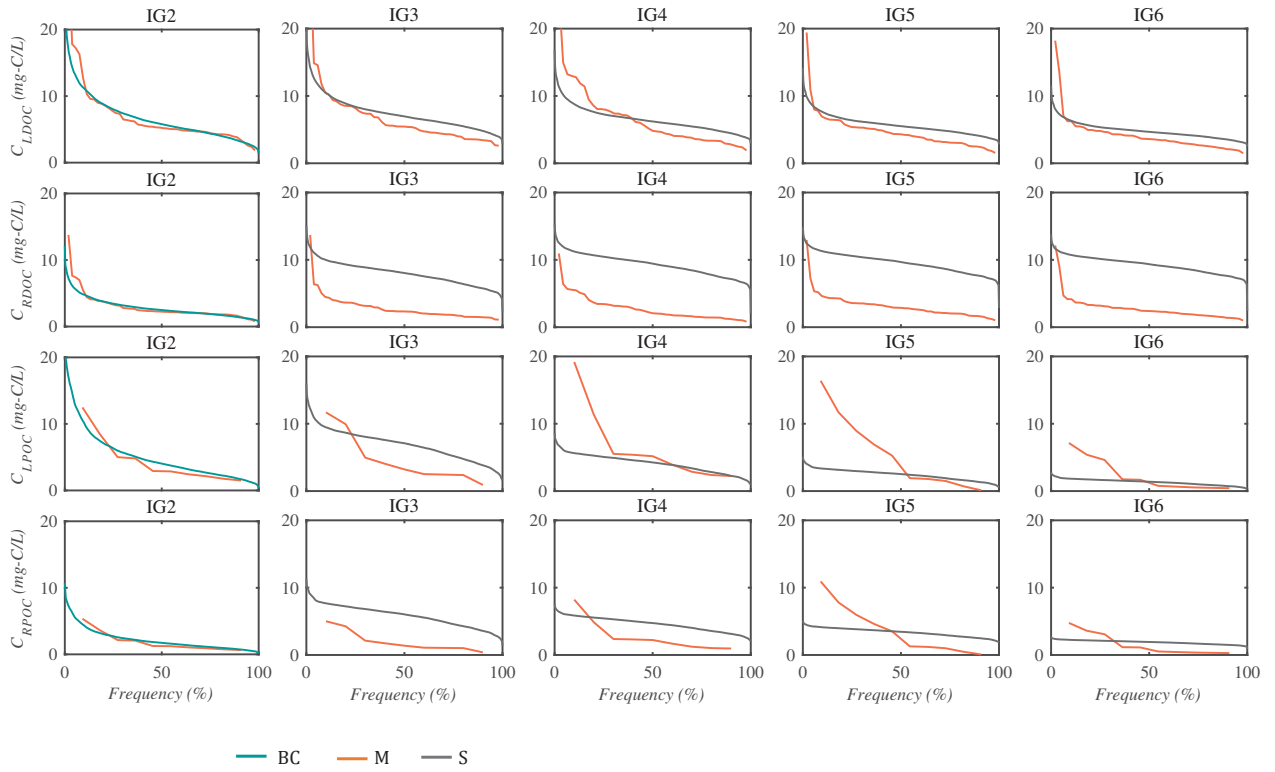
Figure A.25: Daily organic carbon fraction simulated for the years 2013-2015 – BC generated with test T2a: RPOC (continued)

Figure A.26: Boxplot of measured (M) and simulated RPOC concentrations for years 2013-2015 with SIHQUAL (S) and synthetic series (S'); T2a as BC; LDOC, RDOC, LPOC and RPOC, respectively in each line

Figure A.27: Duration curves of historical monitoring dataset 2005-2017 (M) and simulations for 2013-2015 with synthetic series (S') and SIHQUAL results under unsteady (S) and steady (ST) state: BC as T2a; LDOC, RDOC, LPOC and RPOC, respectively in each row

A.15   Pollutographs of temporal variability tests
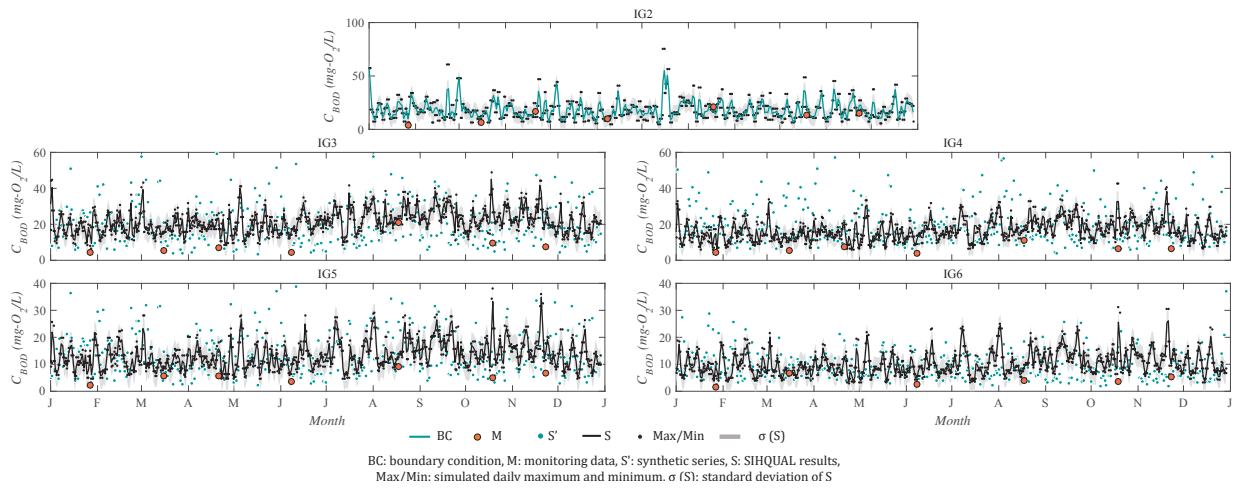
Figure A.28: Daily concentrations generated with different input temporal variability: Experiment *A*: BOD i)
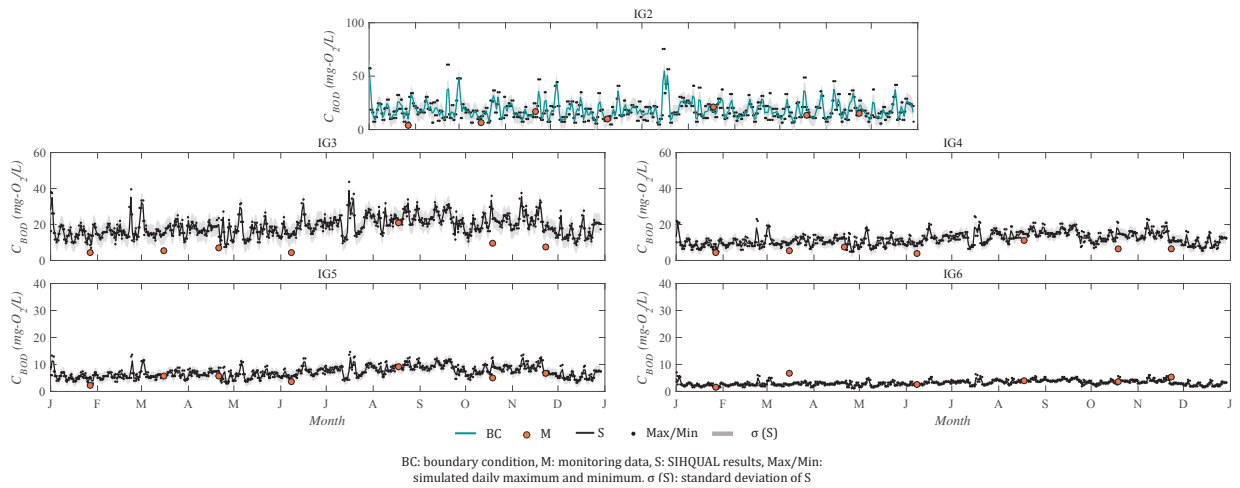
Figure A.28: Daily concentrations generated with different input temporal variability: Experiment *A*: BOD ii) (continued)
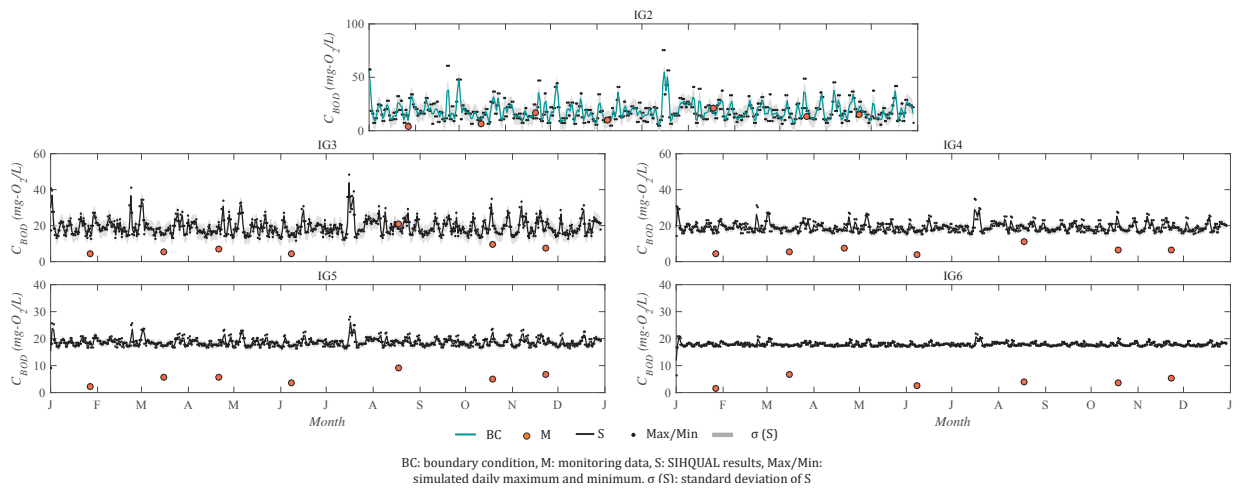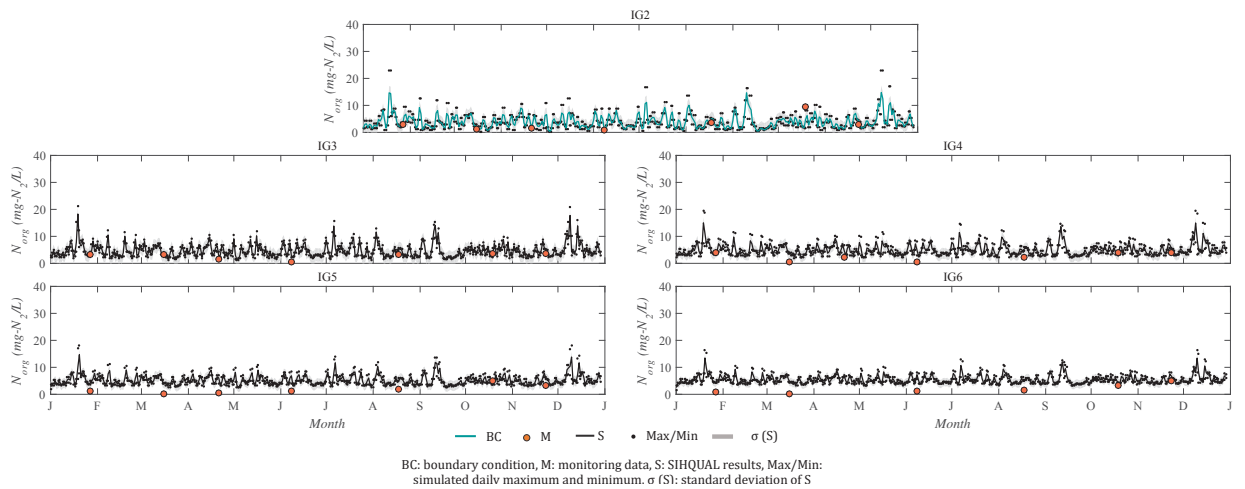

Figure A.28: Daily concentrations generated with different input temporal variability: Experiment *A*: BOD iii) (continued)
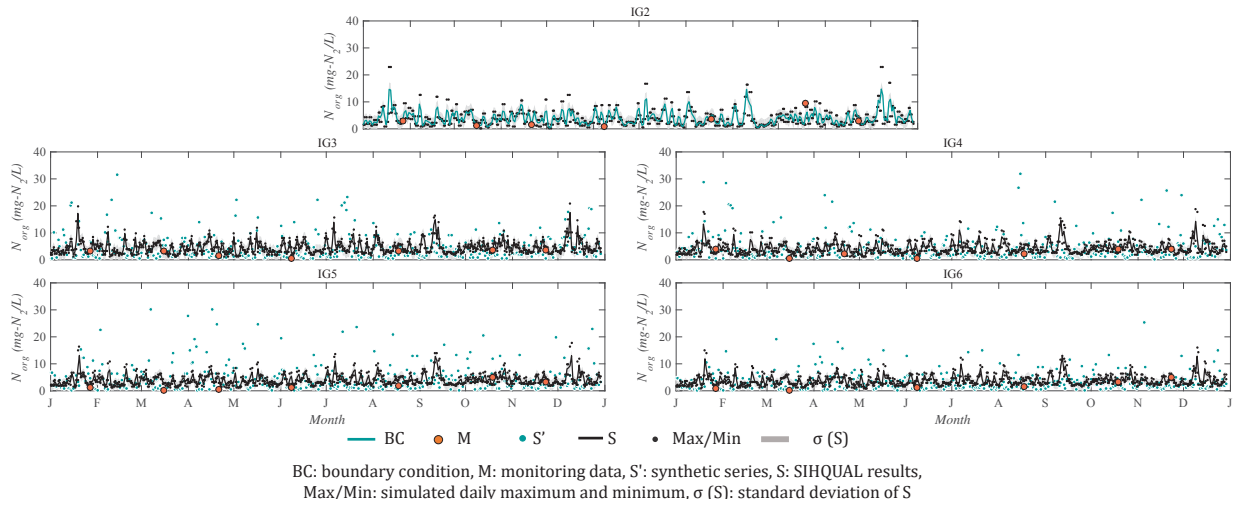
Figure A.28: Daily concentrations generated with different input temporal variability: Experiment *A*: N-org i) (continued)
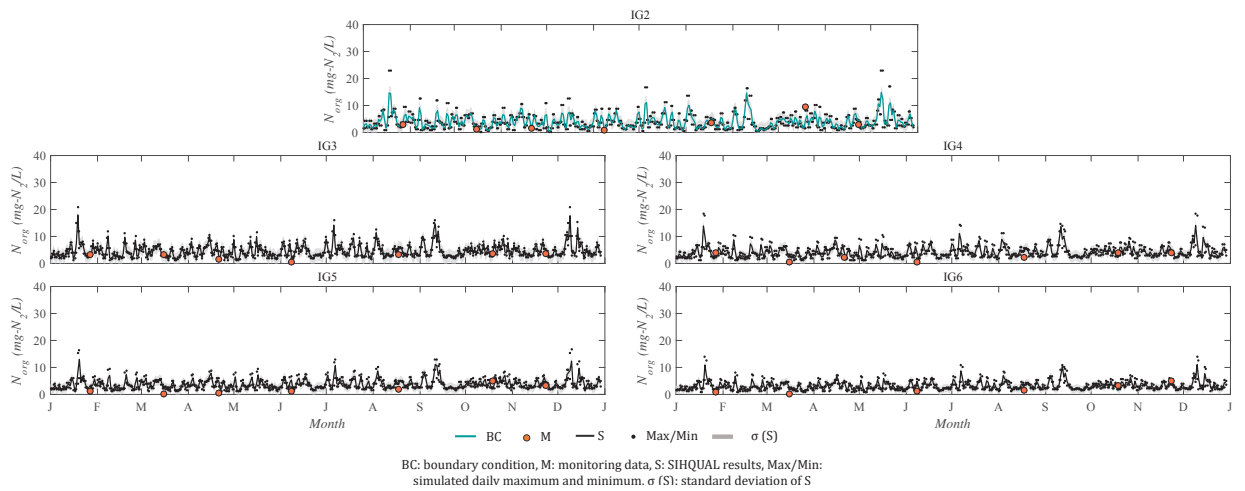
Figure A.28: Daily concentrations generated with different input temporal variability: Experiment *A*: N-org ii) (continued)

Figure A.28: Daily concentrations generated with different input temporal variability: Experiment *A*: N-org iii) (continued)
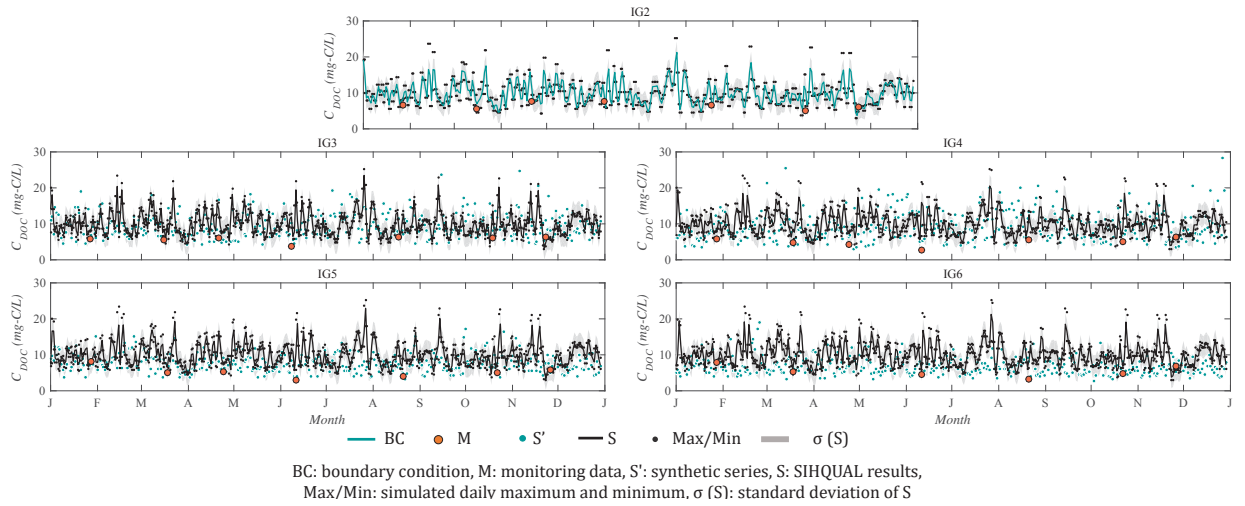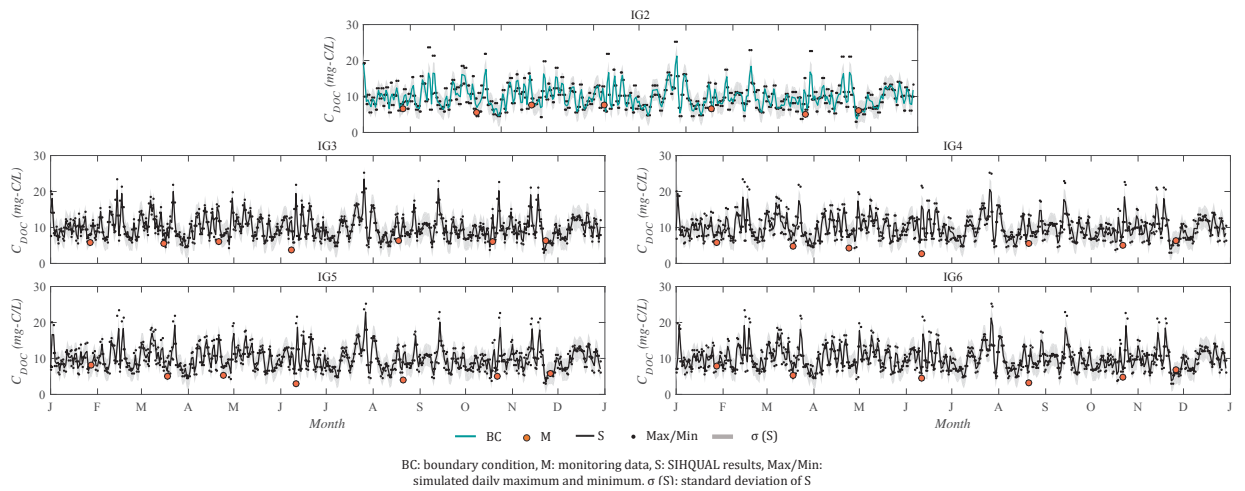
Figure A.28: Daily concentrations generated with different input temporal variability: Experiment *A*: DOC i) (continued)

Figure A.28: Daily concentrations generated with different input temporal variability: Experiment *A*: DOC ii) (continued)
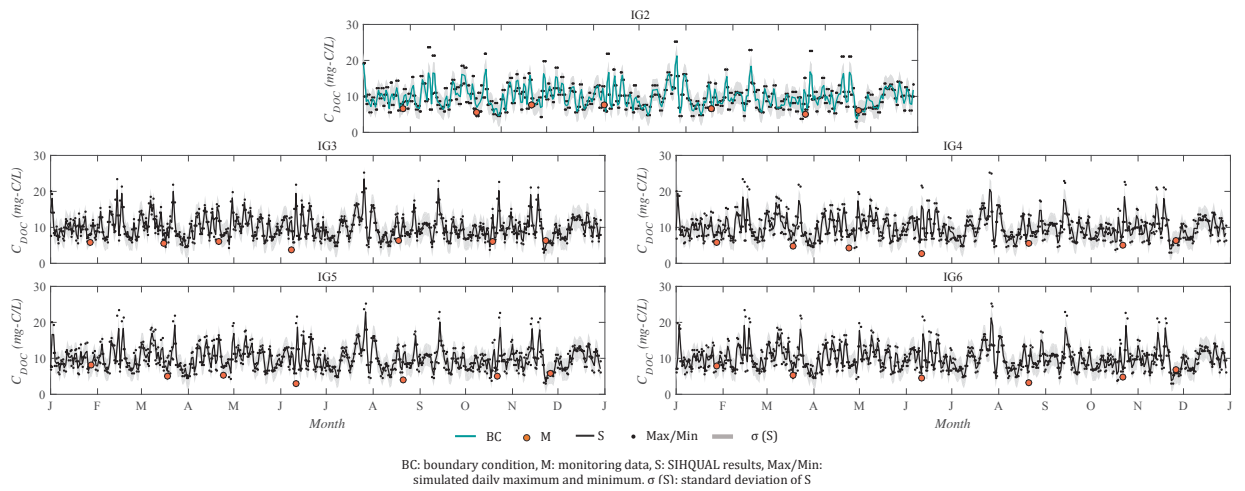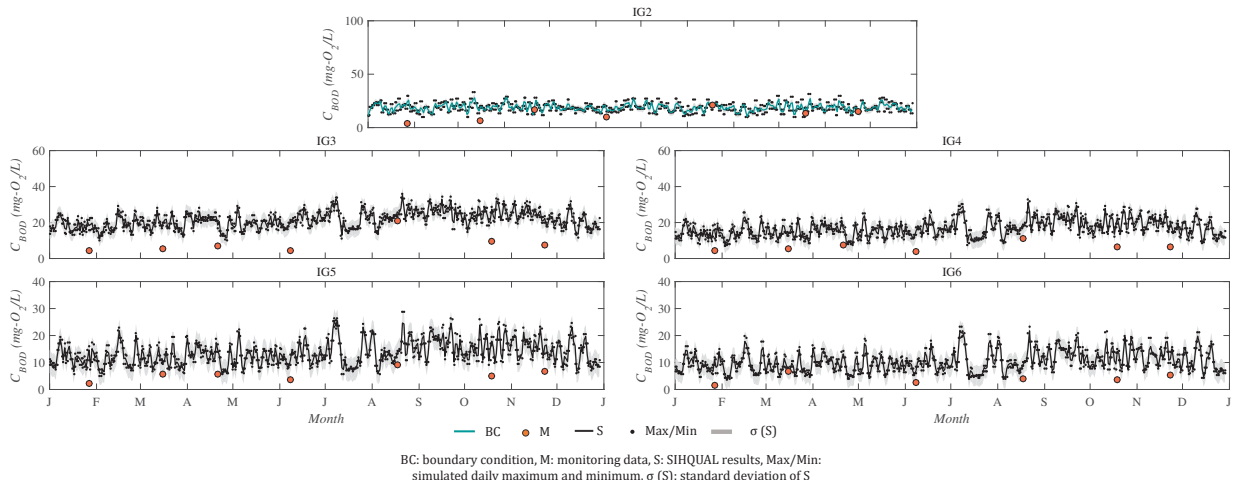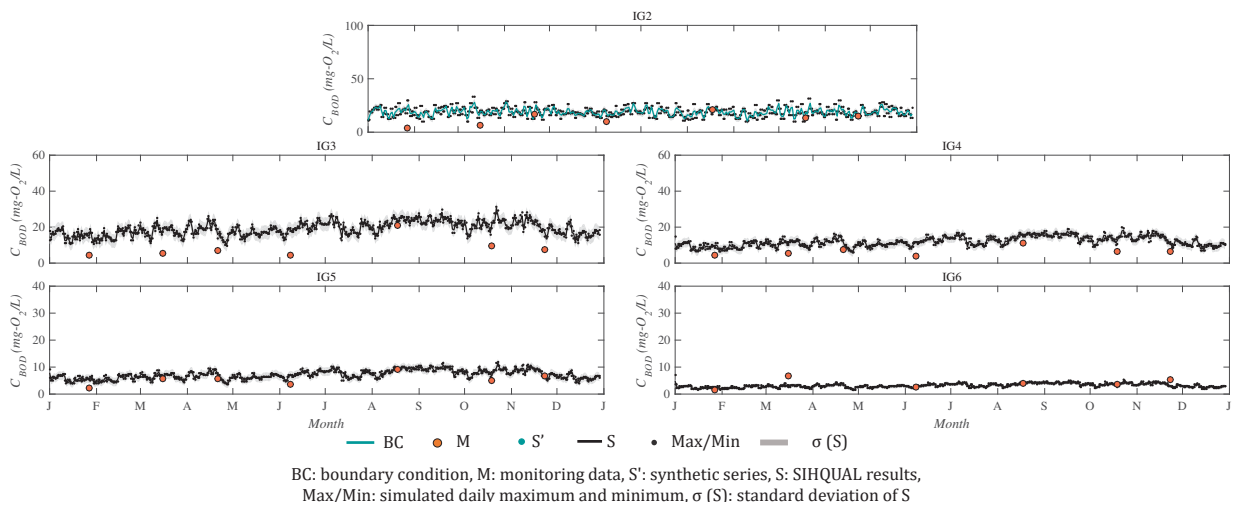

Figure A.28: Daily concentrations generated with different input temporal variability: Experiment *A*: DOC iii) (continued)

Figure A.28: Daily concentrations generated with different input temporal variability: Experiment B: BOD i) (continued)

Figure A.28: Daily concentrations generated with different input temporal variability: Experiment B: BOD ii) (continued)
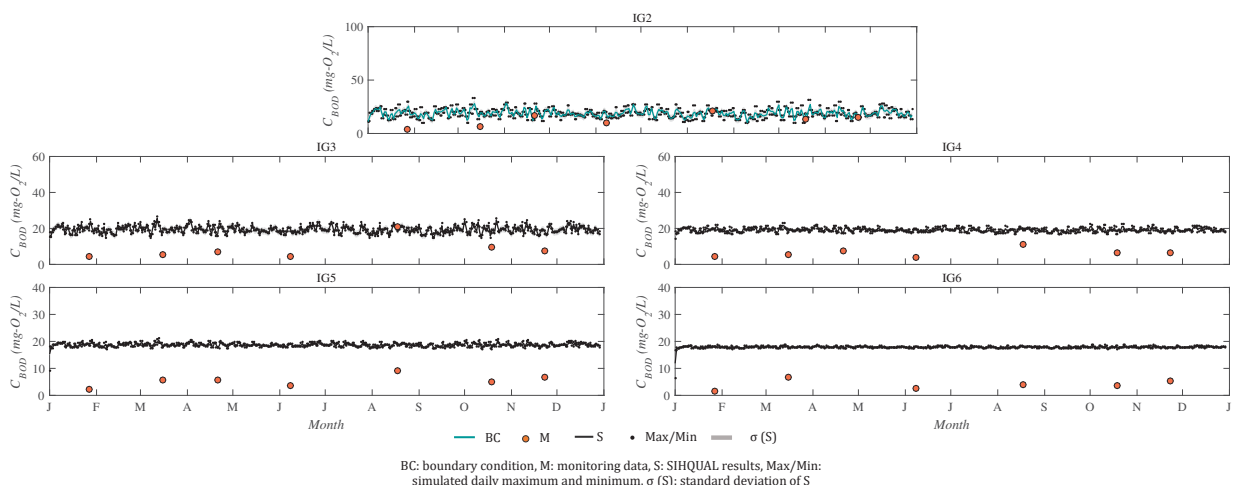

Figure A.28: Daily concentrations generated with different input temporal variability: Experiment B: BOD iii) (continued)