

## Analysis of a probabilistic approach for modelling and assessment of the water quality of rivers

Marianne Brum <sup>a,\*</sup>, Fernando M. Fan<sup>b</sup>, Marcio R. Salla<sup>c</sup> and Marcos von Sperling<sup>d</sup>

<sup>a</sup> KISTERS AG, BU Water, Aachen, Germany

<sup>b</sup> Instituto de Pesquisas Hidráulicas, Universidade Federal do Rio Grande do Sul (IPH-UFRGS), Porto Alegre, Brazil

<sup>c</sup> Universidade Federal de Uberlândia (UFU), Uberlândia, Brazil

<sup>d</sup> Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil

\*Corresponding author. E-mail: marianne.brum@kisters.de

 MB, 0000-0002-1087-3564

### ABSTRACT

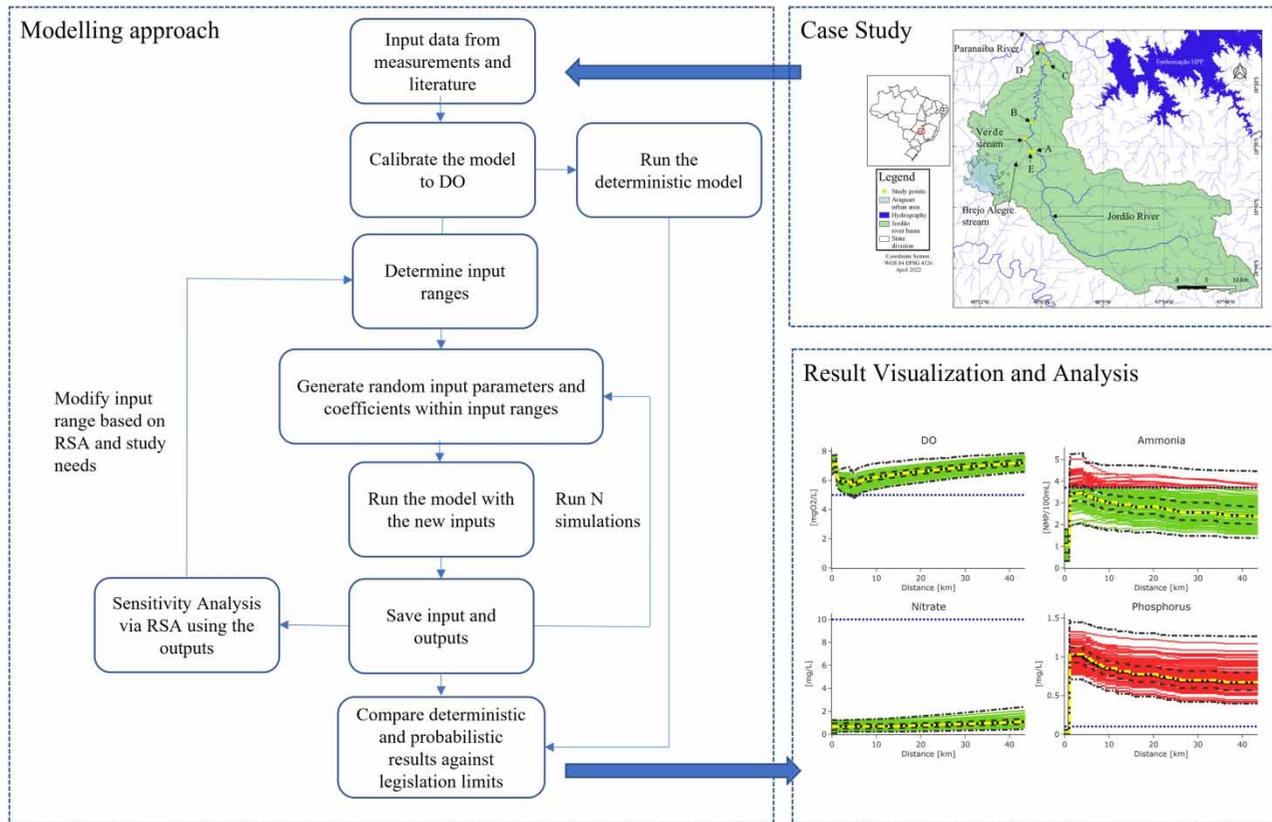
Monitoring the ecological status of water bodies is crucial to guarantee human health and economic development. However, monitoring is often deficient in developing regions due to high installation and maintenance costs, thus it is frequently supported by water quality models, whose results are themselves affected by the lack of detailed input data. A possible solution is to use probabilistic models that consider the inherent uncertainty of the different inputs. In this research, we extended a simple water quality model (QUAL-UFMG, based on Qual2E) through Monte Carlo simulations to generate probabilistic results and applied it to a representative case study in Brazil. Results showed that, depending on the distribution of probabilities and variability of parameters adopted, the outcome of a non-deterministic modelling approach may differ significantly from a deterministic one regarding compliance with water quality standards. Moreover, the probabilistic strategy is more scientifically transparent and robust, as it explicitly communicates the uncertainty in both the measured data and modelling results. We conclude that a probabilistic approach is particularly useful in regions with a low data availability such as developing countries, as uncertainties are high due to insufficient monitoring, and the risk to human health is elevated due to a low prevalence of sanitation.

**Key words:** modelling uncertainty, Monte Carlo simulations, probabilistic analysis, river modelling, water quality

### HIGHLIGHTS

- A deterministic water quality model is expanded through Monte Carlo simulations.
- Outcomes of deterministic and probabilistic approaches to water quality modelling are compared for a case study in Brazil.
- There may be significant differences in modelling results regarding compliance to water quality standards depending on the approach taken; this is key in data scarce regions with deficient monitoring networks.

## GRAPHICAL ABSTRACT



## INTRODUCTION

Access to good-quality water is crucial to guarantee human well-being and economic development. In developing regions, it is a key challenge in the struggle for poverty eradication and social equality, as it is an obstacle to good health, food security, and agricultural and industrial production. Indeed, there is a direct connection between investments in sanitation and hospital admissions due to prevalent waterborne diseases in poorer communities (World Health Organization 2016; Instituto Trata Brasil 2018). According to the latest executive study on sanitation released by the think-tank Instituto Trata Brasil (2018), in 2013, 14.982 million Brazilians had to leave work due to diarrhoea and vomiting. Meanwhile, in 2016, 20% of the population of Brazil still did not have access to potable water, and about 50% were not connected to the sewage collection network (Instituto Trata Brasil 2018). In that context, population growth in major cities of the developing world is expected to stress these issues in the medium- to long-term.

The maintenance and monitoring of water quality are keys to ensure disease prevention, locate pollution sources, and guarantee the good ecological status of water bodies. However, water monitoring networks are often deficient in developing regions due to their high installation and maintenance costs from infrastructure to personnel (Larentis 2004). In that context, water quality mathematical models may be employed as support tools. Modelling is also useful for better management of the available resources through, for example, environmental licensing support, future scenario predictions (particularly climate change), and infrastructure planning.

Though there are many highly complex water quality models available, these are not necessarily best suited for the context of data-scarce environments, such as developing countries. First, emergent regions are often still solving organic pollution issues, which are sufficiently represented by simpler models (von Sperling 2014). These are indeed more pressing problems, as they relate to public health (e.g., waterborne diseases lead to increased hospitalization costs and consequent economic loss) and thus are a serious impediment to social development and poverty reduction. Moreover, due to a lack of monitoring capacity, the required parameters for more comprehensive simulations are unavailable and must be estimated or averaged,

further increasing the uncertainty related to the model. These water quality models rely on good knowledge of water quality parameters, which show significant spread and are difficult to obtain in developing countries due to a lack of reliable and continuous observation time series (e.g., [Ward et al. 2019](#) and [Ngene et al. 2021](#) discuss the difficulties of developing countries to gather and store hydrometric data; [von Sperling 2014](#) thoroughly examines water quality modelling in the Brazilian context). Moreover, other sources of uncertainty inherent to modelling, such as values of reaction rate coefficients, numerical approximation errors, and instabilities must be considered. Thus, the increase in the complexity may not lead to a significant improvement in reliability or accuracy in the results, while increased costs may make a project or evaluation prohibitive.

A possible solution is to use probabilistic modelling for the water quality and explicitly report the associated uncertainty. Non-deterministic (probabilistic or stochastic) simulations rely on the fact that there is inherent variability in the modelling parameters, results, and observation data. They are already used in many water resource applications, notably in meteorological predictions (e.g., [Richardson 2000](#); [Zhu et al. 2002](#); [Brown et al. 2014](#)), hydrological, hydro-meteorological, and water quality forecasting for operational uses such as flood prevention and reservoir management (e.g., [Fan et al. 2015](#); [Anderson et al. 2019](#); [Shamshirband et al. 2019](#); [Asadollah et al. 2021](#)). The literature on ensemble modelling generally concludes that non-deterministic forecast approaches result in more skilful predictions (e.g., [Richardson 2000](#); [Zhu et al. 2002](#); [Fan et al. 2015](#)). Indeed, the US Environmental Protection Agency (EPA) has recommended the usage of non-deterministic modelling when quantifying the uncertainty in deterministic simulations, when using conservative parameters is unnecessary or too expensive, and when the risk of inaccurate simulations due to averaging of parameters is too high ([EPA 1997](#)). However, there are few studies performed in a developing context, where they could potentially be extremely useful.

Therefore, this paper aims to contribute to the existing practices and literature on the topic by arguing that a probabilistic approach to water quality is advantageous and simple to implement, especially within the context of developing regions that traditionally have more data scarcity. To do so, we upgraded a simple water quality model by extending it through Monte Carlo (MC) simulations to generate probabilistic results and applied to a representative case study in South America. The resulting methodology explicitly evaluates uncertainty with the goal of better communicating risk to non-expert decision-makers.

## MATERIALS AND METHODS

### QUAL-UFMG model

The QUAL-UFMG ([von Sperling 2014](#)) model was chosen as the base model for this study. It is a simple spreadsheet implementation of the globally used Qual2E (Enhanced Stream Water Quality Model) developed by the US EPA ([EPA 1987](#)). Qual-UFMG is computationally inexpensive, open-access, easy to share and set up, and modifiable through Visual Basic for Applications (VBA) code. QUAL-UFMG simulates biochemical oxygen demand (BOD), dissolved oxygen (DO), total nitrogen and its fractions (organic nitrogen, ammonia, nitrite, and nitrate), total phosphorus ( $P_{TOT}$ ) and its fractions (organic and inorganic), and thermotolerant coliforms (*Escherichia coli*). For the sake of simplicity, it does not include the influence of longitudinal dispersion and the effects of algae and its interrelationships. Moreover, the transport equation system is solved via the Euler method, thus small integration steps must be taken. These are acceptable limitations for most rivers, like the one studied here. The model assumes steady-state conditions, which are suitable for models aiming at supporting planning studies ([von Sperling 2014](#)). Basic model formulae are summarized in Supplementary Table S7, and the full model description is found in [von Sperling \(2014\)](#).

### QUAL-PROB probabilistic extension

QUAL-UFMG (River Water Quality Model) was extended using MC simulations to randomly generate multiple possible results within a defined range of probable input parameter intervals. For simplicity, we follow a uniform distribution for the input data. This method was chosen because it is simple to implement and to understand, though it presents limitations, e.g., no real stochasticity is inserted; outcome probabilities may be unstable if not enough simulations are run; the distribution tails are very dependent on the input distribution ([EPA 1997](#)). Moreover, we note that this methodology describes, in fact, a hybrid modelling approach, since no stochastic processes are inserted into the existing deterministic model, but we use the term 'probabilistic' for simplicity. These are considered acceptable within the scope of this study, i.e., to provide an initial uncertainty estimation in data-scarce regions to support decision-makers (for contrast, see e.g., [Sharafati & Pezeshki \(2020\)](#) for uncertainty analysis with true stochasticity). 1,000 simulations are generated for each model run, which was

enough to achieve numerical stability since the configuration employed in this study is sufficiently simple. The collection of simulations is treated as an ensemble, where each simulation is an ensemble member.

We integrated the extension (hereon named QUAL-PROB) through a modification of the base model so that the base is fully responsible for the water quality analysis and its structure is not altered in any way. VBA was chosen to implement the MC so that it can be used directly as an extension to the existing spreadsheet, while Python, as a general-purpose programming language, is applied to enhance the graphical output of the model due to Excel's inherent limitations (e.g., a maximum of 250 traces in a plot). The extension handles the statistical post-processing, data visualization, and output results, as well as the random generation of input parameters. It computes the median, average, minimum, and maximum for the output results, and the frequency of conformity to legislation is calculated according to the Brazilian legal standards; see [von Sperling \(2014\)](#) for a description of the Brazilian water quality legal framework. Additionally, a sensitivity analysis was implemented by the Regional Sensitivity Analysis (RSA, [Spear & Hornberger 1980](#); see also [Spear et al. 2020](#), for an implementation example). It is automatically performed by the extension at each model run and output results are presented in a separate spreadsheet.

### Case study

A representative case study in the Brazilian context was performed regarding the self-purification capacity of the Jordão river to the sewage discharges from the city of Araguari. Araguari is located in the center-west Brazilian state of Minas Gerais and has a population of around 110,000 inhabitants. The climate is tropical with a rainy and a dry season when periods with almost no rainfall are not uncommon. Land usage is mostly urban and water demand is for human consumption. The geology of the region is mostly composed of river sediments over the sandstone of the Marília formation ([CPRM 2003](#)). Despite the city's size, the sewage is discharged directly into the Brejo Alegre creek, which then flows to the River Jordão, a tributary of Rio Paranaíba within the Dourados river basin ([Figure 1](#)). The average sediment discharge of the rivers of the basin is around 5–8 tonnes/year ([Fagundes et al. 2021](#)).

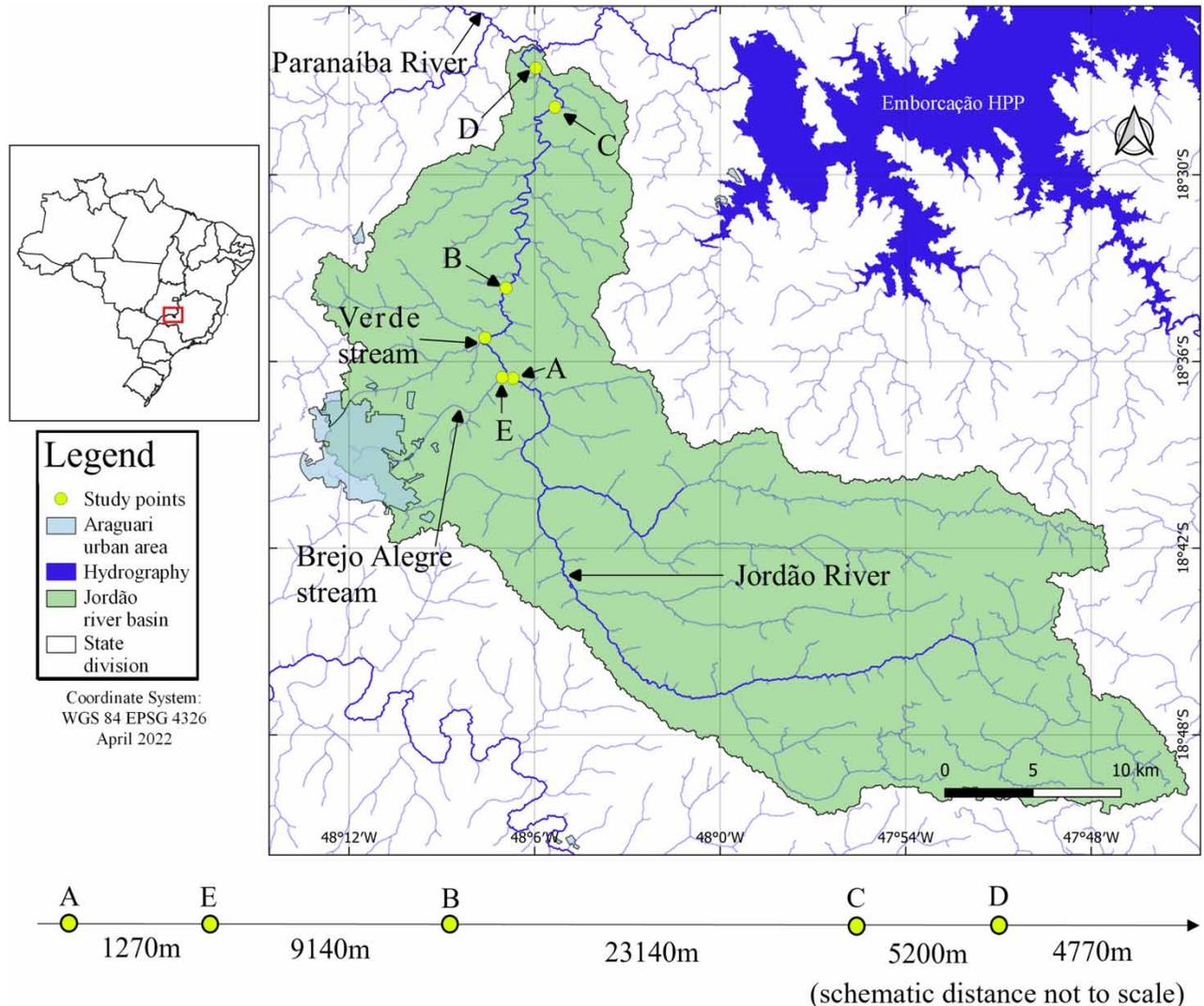
### Data and modelling with QUAL-UFMG

The QUAL-UFMG model was applied to simulate water quality in a 43.5 km reach of the river (from the confluence with the Brejo Alegre to the outflow at the Paranaíba river) under dry weather conditions and calibrated with on-site measured data. From Araguari to the outflow, the Jordão river is joined by 18 other smaller tributaries, modelled as point sources. The modelled parameters were compared against Brazilian water quality standards for Class II fresh water, which may be used for drinking after appropriate treatment ([National Environment Counsel CONAMA 2005](#)). Sequentially, proposed improvements to the sanitation infrastructure of the city were evaluated from two future scenarios of wastewater treatment solutions commonly implemented in Brazilian infrastructure projects: secondary treatment from a combination of an anaerobic lagoon and a facultative; and activated sludge, where the discharge from the treatment plants would flow into the Jordão river. The scenarios were assessed over drought conditions (using the minimum discharge over 7 days with a 10-year return period,  $Q_{7,10}$ , a commonly used reference discharge for dry conditions in the Brazilian legislation) to evaluate the worst-case possibility and verify whether the measures proposed would be enough to improve the overall water quality of the river.

The deterministic study found that in the current scenario for the Jordão river, most water quality parameters except BOD were within the legal limits, despite high pollution levels in the Brejo Alegre tributary. However, the study concluded that the river was incapable of naturally assimilating the current pollutant load generated by the city of Araguari, regardless of the proposed treatment solutions. In both future scenarios, the model results showed that DO, BOD, *E. coli*, and  $P_{TOT}$  levels were outside the legal limits, and the water quality condition in the Jordão river deteriorates. As a result, the analysis, as it was, could be considered unfeasible. The results of this study were originally published by [Salla et al. \(2013\)](#).

### Data and modelling with QUAL-PROB

To compare the deterministic and probabilistic approaches, the same scenarios were modelled with QUAL-PROB. The mean and interval of variation for each parameter were obtained directly from the measured data when available. For all others, it was estimated from the Brazilian literature, resulting in 10–100% of variation range ([Larentis 2004](#); [Costa & Teixeira 2010](#); [Tonon 2014](#)). Furthermore, when the interval derived from measured data was too small (less than 10%), a minimum range of 10% was adopted to guarantee a broader distribution of input values according to the Brazilian literature review. For the alternative scenarios, the minimum pollutant concentration variation was set to 20%, as the uncertainty regarding future conditions is higher. All the input parameters and coefficients are summarized in the Supplementary Material (Tables S1–S6).



**Figure 1** | Situation and schematic diagram of the case study site.

Note that not all QUAL-UFMG parameters were directly available from the study. That was considered acceptable, as the need for indirect (and thus often incomplete) data may also occur in most professional settings. Therefore, we estimated unavailable data and re-calibrated the model. Particularly, the reaeration parameter ( $K_2$ ) was approximated by Owens *et al.*'s relation (Chapra 1997), which is valid for water depths between 0.1 and 0.6 m and velocities of 0.05 and 1.5 m/s, thus applicable to the river Jordão. This was done to facilitate the calibration of the river velocity, depth, and width by letting them be a function of  $K_2$ , which was directly calibrated. Velocity itself is modelled by the QUAL-UFMG model according to the following relation (von Sperling 2014):

$$v = a * Q^b$$

where  $a$  and  $b$  are coefficients and  $Q$  is discharge. In QUAL-PROB, the  $a$  and  $b$  parameters are varied, and not velocity parameter itself. The distributed flow was assumed to be included in the variability of the discharge of tributaries as a simplification because diffusive discharge and load data were also unavailable. For an easier comparison of deterministic and the probabilistic simulations, we also performed a 'best guess' deterministic run (hereafter named DetBG for distinction from the original deterministic simulation) using the averaged input parameters.

The calibration of the missing coefficients was carried out by minimizing the error of the simulated DO parameter throughout the river, as measured by the Nash–Sutcliffe Efficiency (NSE) coefficient (Nash & Sutcliffe 1970). The Nash–Sutcliffe

coefficient is used as a single evaluation metric for simplicity during the calibration phase. The minimization of the objective function was performed independently from the MC approach via a Generalized Reduced Gradient solver, which was already available in Microsoft Excel (based on [Lasdon et al. 1978](#)). Though it does not guarantee an optimal solution, the calibration was found to be sufficiently accurate for the comparative purpose of this study, i.e., evaluating the difference between a probabilistic approach and a deterministic one using the same input data. Other metrics that more heavily penalize bigger errors (e.g., Mean Squared Error) or that evaluate the model accuracy at capturing threshold-based events (e.g., Brier Score) were not applied here but would certainly be relevant in further studies that quantify the difference in the quality of the probabilistic simulation in comparison with the deterministic one.

The sensitivity analysis was evaluated against the DO parameter using a 95% confidence interval. After all simulations are complete, the input coefficients are separated into two groups (those that generated the 50% bigger and 50% smaller DO values, thus balanced by definition). The input samples are compared via *t*-test to verify whether they are statistically different (*p*-value of 0.05); if they are, we consider that they significantly affect the DO concentration in the river. For brevity, the analysis was only performed for DO, which is a key parameter to water quality and affects other model coefficients. Adopting another parameter or a collection of other parameters would certainly affect the output of the sensitivity analysis. Most importantly, as with the original RSA implementation of [Spear & Hornberger \(1980\)](#), here only the univariate marginals are considered, though interactions between the parameters exist. Therefore, an evaluation of parameter correlation would be needed for a complete sensitivity analysis (see, e.g., [Spear et al. 2020](#)). This is beyond the scope of this study; thus, we assume that this simplified analysis provides sufficient estimates of the importance of different parameters and coefficients.

## RESULTS AND ANALYSIS

The semi-automatic calibration resulted in adequate values of the NSE coefficient ([Nash & Sutcliffe 1970](#)) for all parameters evaluated, except BOD, where a steeper degradation rate is evidenced. As a direct consequence, the DetBG and the median of the model runs resembled the original simulation for all parameters, except BOD. The poor performance of the model for BOD was likely due to an incorrect estimation of the organic matter load throughout the river length because the coefficients for BOD modelling are within the range of the original study, but the slope of the BOD decomposition is much steeper in this evaluation. This may be due to the assumption that distributed flow can be modelled through a higher variability interval in the point sources along the river. Still, we decided to maintain this modelling strategy because the NSE score of several runs was positive, there is significant uncertainty in the observed data (up to 30% variability among sampled values), the actual distributed discharge in the region was unknown, and the original implementation also did not conform well to the monitored values.

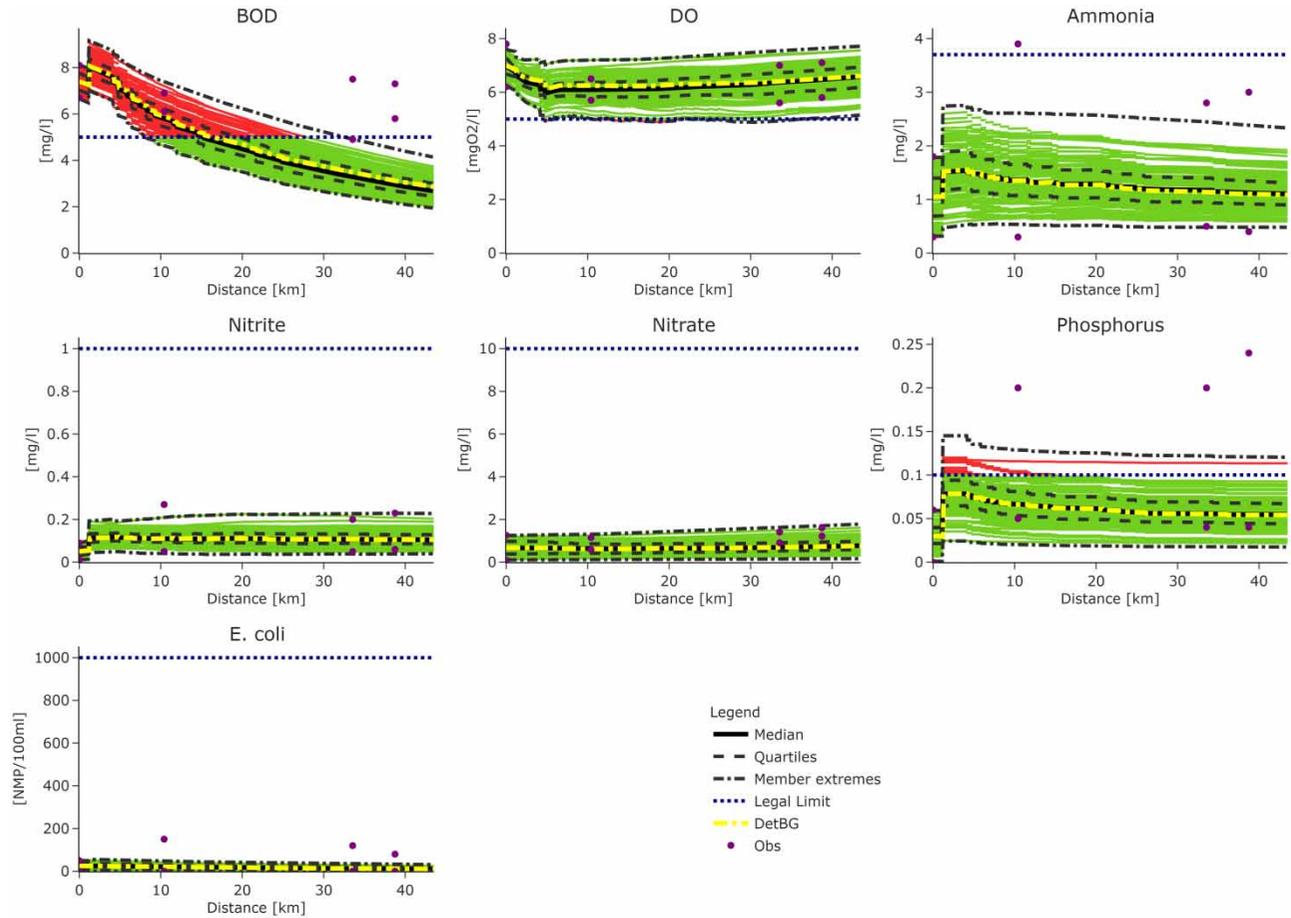
Therefore, hereon the DetBG and the observations will be used as references for easier comparison. The difference between the deterministic and probabilistic results is evaluated by comparing the model outcomes to the observed results and the legislative limits. When evaluating the probabilistic model, both the median and specific members (particularly the extremes and upper and lower quartiles) of the ensemble are considered.

### Current scenario

[Figure 2](#) and [Table 1](#) present the results for the current scenario and for all parameters evaluated. [Figure 3](#) reports the probability of non-accordance (PNA) according to Brazilian legal limits for all parameters and scenarios. The DetBG simulation and the median of the ensembles were almost equal for all parameters evaluated. This is expected, as the uniform probability distribution is characterized by the median being the same as the average for sufficiently large samples.

As previously analyzed, the BOD results do not fit well to the monitored (observed) values. Regardless, the DetBG output and QUAL-PROB median behaved differently for the parameter, with DetBG approximating the upper quartile of probabilistic simulations throughout the modelled stretch. Both the deterministic and the median curves dropped below the legal limit of 5 mgO<sub>2</sub>/l at 15 and 20 km, respectively. However, 22.8% (of QUAL-PROB simulations) were above 5 mgO<sub>2</sub>/l at 20 km from the start of the modelled stretch, indicating a significant probability of non-accordance in relation with the deterministic model. At 33 km, all model members are within the Class II definitions. The inter-quartile range is very narrow, which may be due to low input variability in the decomposition coefficient (10%) and river initial concentration (also 10%).

For the DO parameter, the median of the ensembles and DetBG show somewhat different behaviour, which increases with the length of modelled river. At 45 km, the DetBG curve approaches the upper quartile of the probabilistic model. Both are above the legal limit of 5 mg/l and within the measured values, though the QUAL-PROB median is closer to the average of

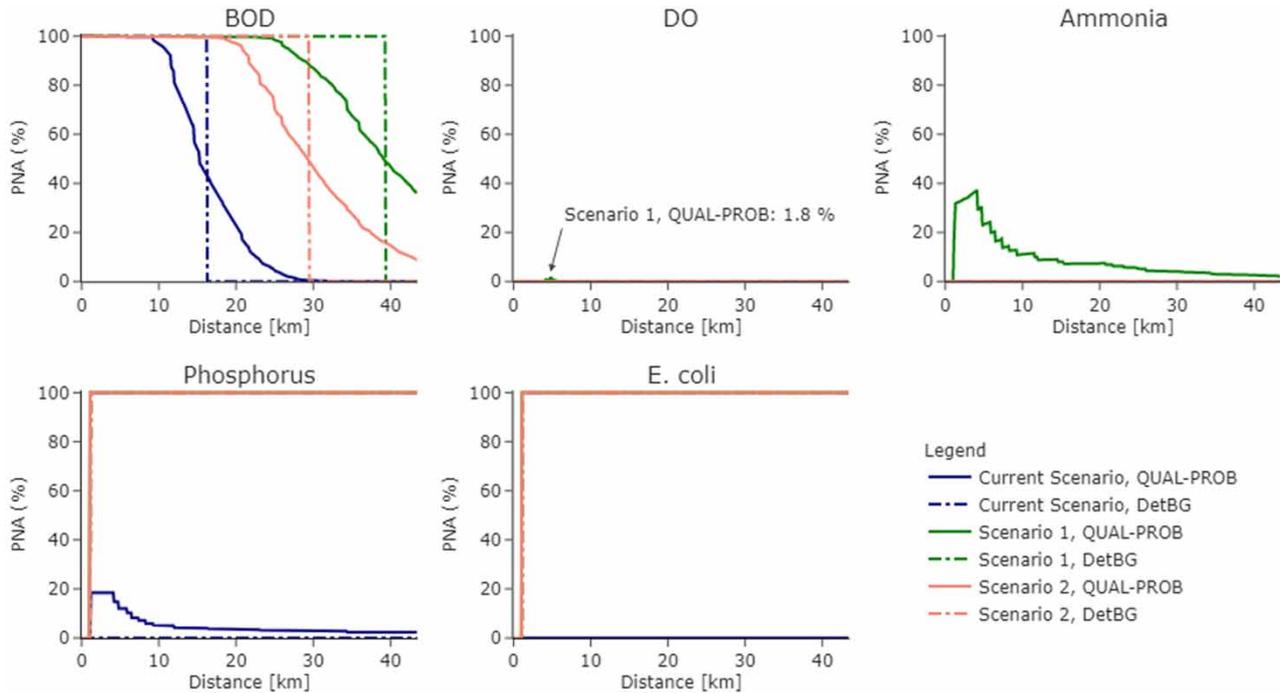


**Figure 2** | Concentration of evaluated parameters throughout the modelled river stretch for the current scenarios. Of 1,000 simulations, 100 are plotted in green (within Class II legal limits) and red (not conforming to Class II legal limits). Median is the median value for all simulations; quartiles represent the 25 and 75% values of the simulations; extremes are the maximum and minimum values of the simulations; DetBG refers to the deterministic run; and Obs are the observed measurements. Please refer to the online version of this paper to see this figure in colour: <http://dx.doi.org/10.2166/hydro.2022.157>.

**Table 1** | Summary results for the current scenario

Parameter	Legal limit (Class II)	Maximum simulated (QUAL-PROB)	Maximum simulated (DetBG)	Maximum observed	Minimum simulated (QUAL-PROB)	Minimum simulated (DetBG)	Minimum observed
BOD (mg/l)	5.00	9.16	8.05	8.10	1.94	2.86	4.90
DO (mg/l)	5.00	7.72	7.00	7.80	4.88	6.12	5.60
Ammonia (mg/l)	3.70	2.75	1.55	3.90	0.32	1.05	0.30
Nitrite (mg/l)	1.00	0.23	0.77	1.61	0.10	0.62	0.09
Nitrate (mg/l)	10.00	1.78	0.12	0.27	0.01	0.05	0.01
Phosphorus (mg/l)	0.10	0.15	0.08	0.24	0.00	0.03	0.00
<i>E. coli</i> (MPN/100 ml)	1.00E+03	5.55E+01	2.81E+01	1.50E+02	4.11E-02	1.11E+01	0.00E+00

the observations. Approximately 2% of the ensemble members showed DO levels below 5 mg/l close to 20 and 30 km. Moreover, the observations are captured within the extremes range, evidencing that the distribution spread is enough to represent the natural uncertainty (Jaun & Ahrens 2009). That is expected, as the model was calibrated based on the DO parameter.



**Figure 3** | Probability of non-accordance (PNA) comparison for QUAL-PROB and DetBG. Results are shown for each scenario and each parameter where PNA > 0 in at least one scenario. The DetBG and QUAL-PROB curves overlap when the PNA is the same for both models in the same scenario.

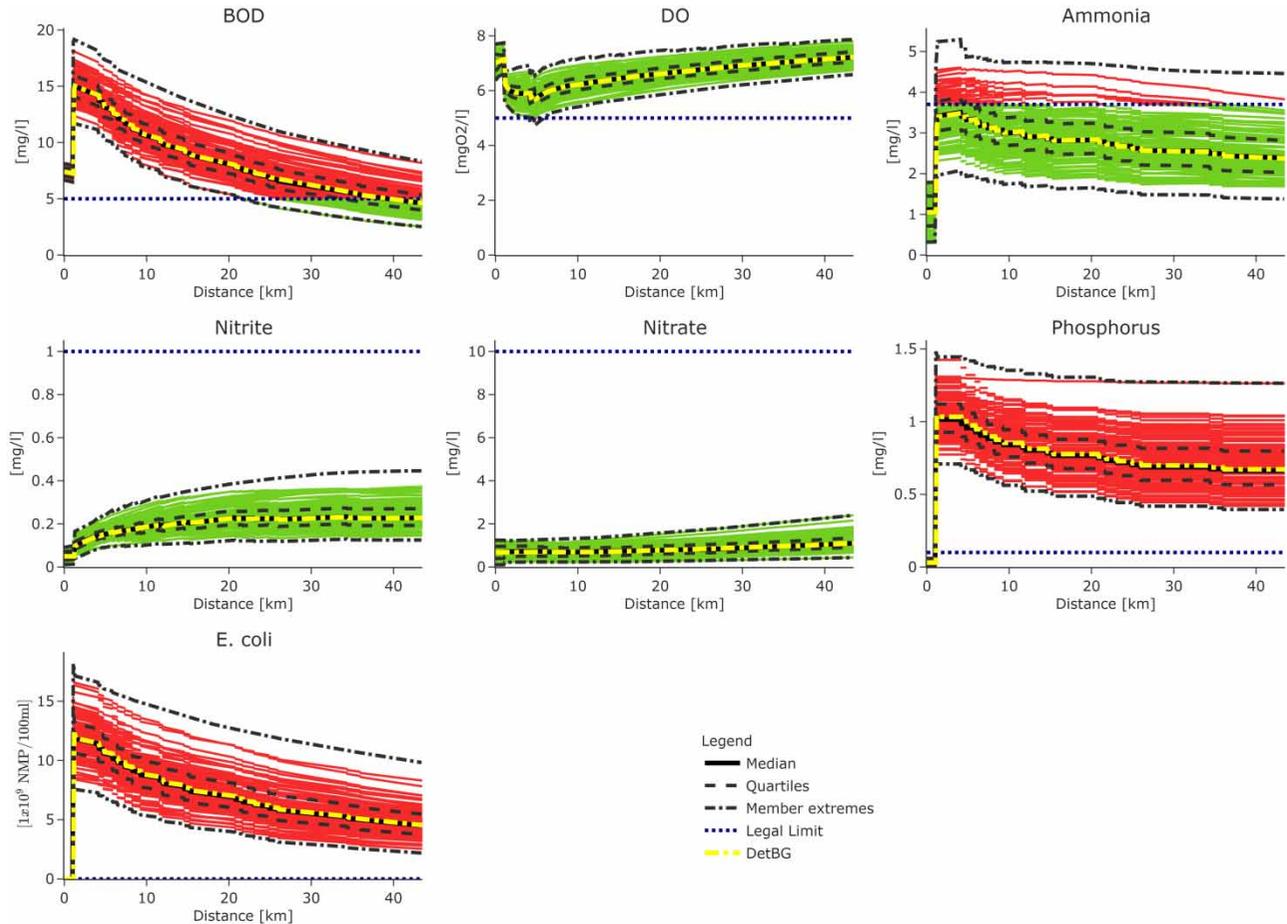
For  $P_{TOT}$ , both model results are in accordance with the Class II limit of 0.1 mg/l and within the maximum and minimum measured values for the parameter. However, a few (18.5%) QUAL-PROB model members reach values above 0.1 mg/l after the confluence with the Brejo Alegre stream, while 2.3% remain above the threshold up to the end of the modelled stretch. Therefore, there is a likelihood of non-accordance to legislation, which is higher just after the inflow from the Brejo Alegre. This is indeed present in the observed data, which shows measured  $P_{TOT}$  concentrations up to 0.14 mg/l at the tributary confluence. The probability of non-compliance is also 18.5% in the whole region, which means that the maximum non-conformity probability happens just after the confluence with the Brejo Alegre tributary. Additionally, the  $P_{TOT}$  plot shows that the model runs have variable paths along the length of the river, representing the significantly different results that different input parameters generate. This is particularly evident in the lower and upper quartiles and in the extremes.

The QUAL-PROB median model results for nitrite, nitrate, ammonia, and *E. coli* were all in accordance with Class II definitions. Likewise, all ensemble members were well below the legal limits for these parameters: at the confluence with Brejo Alegre, the maximum ammonia-N concentration for the upper extreme ensemble was 2.74 mg/l against the limit of 3.7 mg/l; for *E. coli*, it was 55 MPN/100 ml against 1,000 MPN/100 ml (Most Probable Number, method to estimate concentrations); for nitrate and nitrite, the upper extreme at the end of the modelled reach was 1.78 mg/l against 10 mg/l and 0.23 mg/l against 1 mg/l, respectively. Moreover, the ensemble spread was mostly sufficient to capture all the observation variability in nitrite and nitrate, though there is a clear systematic bias between simulated and observed values for nitrate.

Some observations lie outside the extremes range, which means the spread of the simulated runs was not enough to capture all observation uncertainty for *E. coli* and ammonia, which is also evidenced in the BOD and  $P_{TOT}$  parameters. This may simply be due to a too narrow input interval for the model coefficients (as pointed out for BOD) and initial concentrations, evidenced in the larger difference between maximum and minimum ranges of measured values. Regarding variability, the upper quartile members notably exhibit more heterogeneous behaviour than the lower quartile. Nonetheless, these simulations demonstrate that even a simple model may behave very differently depending on the input parameters.

### Future scenarios

In the first alternative scenario (anaerobic lagoon followed by facultative lagoon), there is an overall deterioration of the water quality in the main river (Figure 4; Table 2). The DetBG and probabilistic models do not diverge regarding accordance to



**Figure 4** | Concentration of evaluated parameters throughout the modelled river stretch for future scenario 1. Of 1,000 simulations, 100 are plotted in green (within Class II legal limits) and red (not conforming to Class II legal limits). Median is the median value for all simulations; quartiles represent the 25 and 75% values of the simulations; extremes are the maximum and minimum values of the simulations; DetBG refers to the deterministic run. Please refer to the online version of this paper to see this figure in colour: <http://dx.doi.org/10.2166/hydro.2022.157>.

legislation for DO, nitrite, nitrate, *E. coli*, and  $P_{TOT}$ . The maximum *E. coli* concentration, an indicator of urban sewage, increases to  $18.10^9$  MPN/100 ml which is well above the defined legal limits.  $P_{TOT}$  too surpasses the allowed limit just after the confluence with the sewage system, up to 0.4–1.26 mg/l. There is also an increase in the concentration of nitrite and nitrate, but these parameters remain within legal standards for Class II throughout the Jordão river.

For BOD, both deterministic and probabilistic curves show a significant increase in the concentration of organic matter. However, DetBG indicates that BOD drops below the legal limit at 38 km, while QUAL-PROB indicates that there is only a 45% probability of the parameter being within Class II limits at that location, and 64% for the last modelled point in the river (Figure 3). On the other hand, there is already a non-negligible (10.3%) chance of BOD being in accordance with legislation at 29 km. An even bigger distinction is observed for ammonia, where DetBG indicates that the parameter is compliant throughout the river extension, while the probabilistic model concludes that there is a 37.5% likelihood of that being false. The probability of compliance for ammonia and BOD increases with distance from the confluence as self-purification occurs more rapidly than the inflow of pollutants from the smaller tributaries.

In scenario 2 (Figure 5; Table 2), there is a bigger deterioration in water quality compared to scenario 1 for  $P_{TOT}$ , *E. coli*, and nitrate, while there is an improvement for the other parameters. Nonetheless, nitrate and DO results demonstrate that concentration levels remain within the desired limits throughout the river. Likewise, ammonia concentrations also remain mostly within legislation definitions, with a negligible (0.01%) probability of non-accordance (Figure 3). On the other hand, *E. coli* concentrations are extremely high throughout the river (the maximum is over  $56 \times 10^9$  MPN/100 ml after the confluence with the Brejo Alegre tributary) and so are  $P_{TOT}$  levels (maximum of 1.7 mg/l), though both are within limits

**Table 2** | Summary results for both future scenarios

Parameter	Legal limit (Class II)	Removal efficiency (%)	Maximum simulated (QUAL-PROB)	Maximum simulated (DetBG)	Minimum simulated (QUAL-PROB)	Minimum simulated (DetBG)
<b>Scenario 1</b>						
BOD (mg/l)	5.0	75.0	19.15	15.02	2.53	4.61
DO (mg/l)	5.0	–	7.86	7.22	4.74	5.70
Ammonia (mg/l)	3.7	50.0	5.29	3.49	0.32	1.05
Nitrite (mg/l)	1.0	0.0	0.45	1.11	0.01	0.68
Nitrate (mg/l)	10.0	60.0	2.39	0.23	0.10	0.05
Phosphorus (mg/l)	0.1	35.0	1.48	1.05	0.00	0.03
<i>E. coli</i> (MPN/100 ml)	1,000.0	90.0	1.82E+10	1.25E+10	4.91E–02	2.47E+01
<b>Scenario 2</b>						
BOD (mg/l)	5.0	85.0	14.19	11.82	2.22	3.71
DO (mg/l)	5.0	–	8.05	7.42	5.36	6.18
Ammonia (mg/l)	3.7	85.0	3.76	2.25	0.32	1.05
Nitrite (mg/l)	1.0	0.0	0.38	1.04	0.01	0.68
Nitrate (mg/l)	10.0	25.0	2.45	0.19	0.10	0.05
Phosphorus (mg/l)	0.1	25.0	1.70	1.20	0.00	0.03
<i>E. coli</i> (MPN/100 ml)	1,000.0	60.0	5.66E+10	3.76E+10	7.91E–02	2.47E+01

up to the sewage system inflow. These are like the deterministic output regarding environmental legislation, though the absolute parameter concentrations differ significantly.

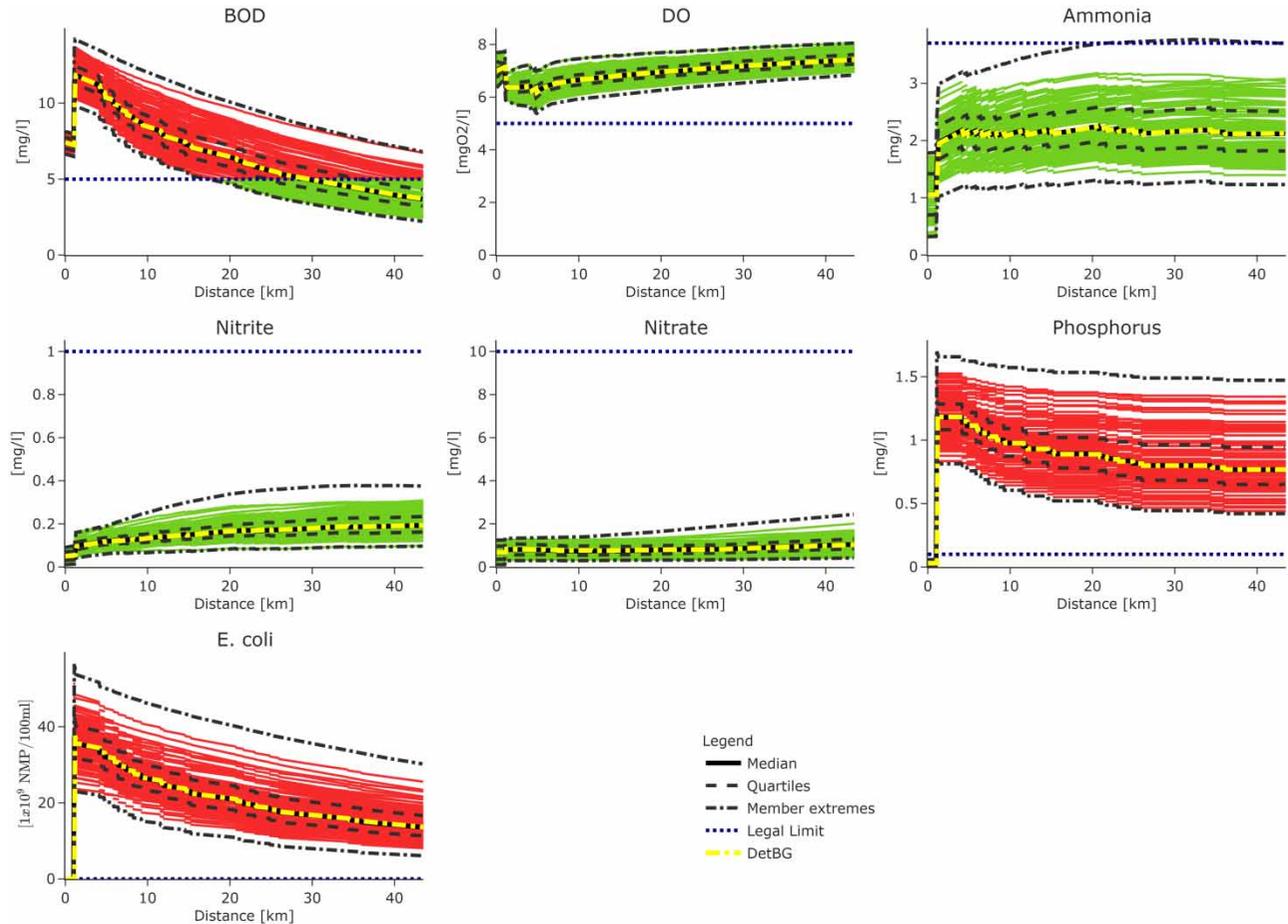
However, probabilistic BOD results show significantly different outcomes: though DetBG predicts that BOD concentrations are below Class II limits (5 mgO<sub>2</sub>/l) after 30 km, QUAL-PROB shows over 46% probability of non-accordance at that point. The probability decreases to 9.1% at the confluence between the Jordão river and Paranaíba river, which is the end of the modelled river stretch. On the other hand, at 20.2 km there is already a 20% chance of accordance according to QUAL-PROB.

In comparison with the current situation, both scenarios show an improvement in DO concentrations due to lower river flows (circa three times less flow than the base scenario), which affect the BOD removal and reoxygenation coefficients and thus artificially increase the self-depuration capacity of the river. Therefore, the distinction in results from scenarios 1 and 2 is due primarily to the variation in efficiency removal of the different wastewater treatment types, but the same cannot be said for the difference in the current and future scenarios. Moreover, sewage becomes proportionally important in comparison with upstream river flow, hence the concurrent increase in the concentration of pollutants, regardless of the treatment method adopted. If average hydro-meteorological conditions were simulated, results would certainly differ.

The variability of parameter concentrations is generally similar in future scenarios, despite the higher interval range of 20% for sewage input. The absolute variation is also higher closer to the outflow into the Paranaíba river than at Araguari. Moreover, P<sub>TOT</sub>, ammonia, and nitrate show most variability at the maximum extremes, with little change at the minimums.

### Sensitivity analysis

Table 3 summarizes the resulting variables deemed important in the sensitivity analysis, where we evaluated the latest point in the modelled river stretch. In all scenarios, temperature, tributary flow ( $Q_t$ ), initial BOD and DO,  $K_1$ , and  $K_D$  were found to be significant. Temperature is a crucial variable for the river self-purification mechanisms, as it impacts the velocity of the chemical and biological processes of nutrient conversion, oxygenation, and deoxygenation and decay of pollutants. Flow in the tributaries is also key to ecological depuration, and in this case study, it represents from 57 to 33% of the total river discharge at the furthestmost point of the current and future scenarios, respectively. Initial BOD and DO show that the initial load of pollutants in the river is significant, even after the confluence with the city's sewage system.



**Figure 5** | Concentration of evaluated parameters throughout the modelled river stretch for future scenario 2. Of 1,000 simulations, 100 are plotted in green (within Class II legal limits) and red (not conforming to Class II legal limits). Median is the median value for all simulations; quartiles represent the 25 and 75% values of the simulations; extremes are the maximum and minimum values of the simulations; DetBG refers to the deterministic run. Please refer to the online version of this paper to see this figure in colour: <http://dx.doi.org/10.2166/hydro.2022.157>.

$K_1$  and  $K_D$  are found to be important, as they are the main variables in the equation of organic matter decay. The  $a$  and  $b$  parameters, which indirectly model the  $K_2$  coefficient, are also singled out in the alternative scenarios, which indicates that the coefficient of reaeration ( $K_2$ ) is significant. This is likely because smaller variations in the discharge of pollutants become more relevant with lower river flows and consequently less dilution capacity. As the  $K_2$  parameter is modelled indirectly, it might be worth increasing the variation interval for the variables that define it.

Notably, it was found that the spread of all parameters except DO is significantly dependent on  $K_D$ , and that the variability is drastically reduced (as much as 80% for BOD) if  $K_D$  is doubled (representing a two-fold decrease in the decay velocity of pollutants). This is likely why the inter-quartile range for BOD was so narrow.

## DISCUSSION

In this research work, the median of the ensembles and the deterministic outputs are almost identical. The median of the probabilistic model is the outcome of a statistical treatment of the results and thus is distinguished from the deterministic concentrations, which are only reliant on the input data. Thus, the median reflects the probabilistic results but is more easily compared with the deterministic model. In that regard, the close approximation of the QUAL-PROB median and the DetBG results is expected, due to the chosen distribution of probabilities. However, the statistical approach here presented is not limited to uniform distributions, and various Probability Density Functions (PDF) (e.g., normal and log-normal distributions) could have been applied depending on the use-case, thus resulting in different outputs. System analysts and experienced decision-makers are encouraged to explore different statistical distributions that best fit their data. Still, the

**Table 3** | Sensitive parameters for DO for each modelled scenario

Stretch	Parameter	p-value		
		Current scenario	Scenario 1	Scenario 2
Main river	River discharge ( $Q_r$ )	2.73E-39	*	*
	Initial river DO concentration (DOR)	7.27E-04	*	2.42E-02
	Initial river BOD concentration (BODr)	1.54E-17	3.00E-06	1.67E-06
	Temperature ( $T$ )	1.07E-35	2.19E-53	6.08E-145
	Deoxygenation coefficient ( $K_1$ )	4.02E-13	5.39E-27	4.82E-23
	BOD removal coefficient ( $K_D$ )	4.01E-13	2.50E-16	7.90E-27
	Initial river <i>E. coli</i> concentration (Colir)	*	6.59E-03	
	Velocity coefficient ( $a$ ), where $v=a.Q^b$	*	1.56E-03	1.71E-02
	Velocity coefficient ( $b$ ), where $v=a.Q^b$	*	1.85E-02	
	Coefficient of $N_{ORG}$ sedimentation ( $K_{so}$ )	*	4.27E-02	4.92E-02
	Coefficient of nitrification inhibition due to low DO (Knitr)	*	4.66E-03	2.61E-02
	Fraction of free ammonia (fNH <sub>3</sub> )	*	1.43E-02	*
	Coefficient of sedimentation of $P_{ORG}$ ( $K_{spo}$ )	*	2.64E-02	*
	Initial river $N_{ORG}$ concentration (Norgr)	*	*	4.36E-02
	Initial river $P_{ORG}$ concentration (Porgr)	*	*	1.56E-02
	Coefficient of liberation of $P_{INORG}$ from bottom sediments (Spinorg)	*	*	8.82E-03
	Coefficient of coliform decay ( $K_b$ )	*	*	3.42E-02
Small tributaries	Initial tributary discharge ( $Q_t$ )	1.96E-55	3.56E-14	1.43E-33
	Initial tributary BOD concentration (BODt)	3.83E-02	*	*
	Initial tributary $P_{ORG}$ concentration (Porgt)	*	3.70E-02	*
	Initial tributary nitrate concentration (Namont)	*	*	4.37E-03
Brejo Alegre tributary	Initial tributary $N_{ORG}$ concentration (Norgt)	4.49E-02	*	*
	Initial tributary $P_{TOT}$ concentration (Porgt)	3.62E-02	*	*
	Initial tributary BOD concentration (BODt)	3.60E-02	*	6.45E-04
	Initial tributary nitrate concentration (Nnitratot)	*	*	2.20E-02
Sewage inflow	Sewage discharge ( $Q_s$ )	NA	4.00E-25	9.31E-15
	Initial tributary $N_{ORG}$ concentration (Norgt)	NA	7.08E-03	*
	Initial sewage BOD concentration (BODs)	NA	1.09E-21	8.86E-09

Non-sensitive parameters are marked with an asterisk.

NA, not applicable.

median curve may be viewed as a form of consensus among various users since it does not depend on subjective choices regarding risk. Ideally, further research should be carried out to better evaluate other statistical aggregation methods or verification criteria that could act as measures of the overall agreement.

It should be noted that the ‘worst-case’ scenario for discharges still leaves room for multiple worse water quality outcomes than the deterministic run depending on the choice and variability of coefficients and/or the initial concentration of parameters, even when there is an agreement between the ensemble mean and the deterministic run. In fact, the possible heterogeneity between ensemble members (e.g.,  $P_{TOT}$ ) illustrates how a particular set of model parameters can lead to significantly different results than the ensemble mean. One must verify whether the coefficient and parameter intervals are physically meaningful for each study case. Assuming deterministic parameters centred in the ensemble distribution is appropriate here, when the monitoring was carried out under normal river conditions; if sampling was done during a drought period, the deterministic run could be set as a minimum and the input distribution accordingly modified.

In that regard, QUAL-PROB results evidenced that many parameters ( $P_{TOT}$ , DO, and BOD in the current scenario, BOD and ammonia in the first future scenario, and BOD in the second) present a risk of non-compliance which differs from the deterministic output. Therefore, depending on the input parameters, variability, and probability distribution adopted, the probabilistic and deterministic results may differ significantly regarding legislation accordance.

In this case study, the maximum  $P_{TOT}$  concentrations in the current scenario reach 0.14 mg/l, which is almost the limit for Class 3 waters (0.15 mg/l) and would ultimately prohibit the river from being used for fishing or a drinking source for livestock (CONAMA 2005). Meanwhile, ammonia levels have a 37.5% chance of remaining above legal limits throughout the

river stretch. While small variations of pollutant concentration from the limits might not affect the aquatic environment instantaneously, bigger divergences (even if not highly likely) may have drastic consequences. On the other hand, for BOD, QUAL-PROB demonstrates that there is a chance of conformity to the legislation closer to the city than what the deterministic output predicts. This illustrates how the probabilistic approach may also prevent the usage of overly conservative parameters in situations where the risk is negligible and eventually save money by avoiding over dimensioning.

Indeed, even when both methodologies agree, they do not provide the same amount of information. The probabilistic approach can better capture the uncertainty in the observations (e.g., nitrate and nitrite) and is ultimately more scientifically transparent, as it explicitly communicates those results are not absolute.

It has already been shown that there is value in the explicit usage and communication of uncertainty and variability. Probabilistic results allow for a wider range of decision thresholds and relate to how different users have different sensitivities to risk (e.g., *Zhu et al. 2002*; *Roulin 2007*). There is a broader range of information provided by the probabilistic model, which can be interpreted in various ways depending on the risk aversion of the end-user. For example, a municipality that relies heavily on eco-tourism for revenue may be inclined to reject a new factory that, even at low probabilities, might worsen the water quality in the region to the point recreational use is no longer permitted, because the risk of loss of income is too high. On the other hand, a city looking to expand its industrial activities might approve it. Indeed, even for the same situation, different risk profiles may lead to different decisions. For the current scenario of this study, DO has a small (about 2%) probability of dropping below the legal limit of 5 mg/l, which would restrict the permitted water uses in the river. A conservative decision-maker might find this unacceptable, while another administrator might consider the risk sufficiently low. We must stress that no strategy (e.g., risk aversion or tolerance to higher risk) is inherently better or worse than the others. Decision-makers (who are frequently non-experts) will over time develop their subjective perception of risk and of possible losses to each situation depending on the uncertainty in the modelling results. However, risk-based decision-making is only feasible when results produced by a probabilistic model are considered.

Moreover, though legislated limits are not absolute thresholds of ecological behaviour of water environments, they do affect the legal tools that decision-makers have at their disposal to act such as financial sanctions and refusal of permits. Likewise, such financial tools can be purposefully curtailed by agents that on bad faith choose advantageous parameters within the available literature. For example, the variability range for ammonia is 1.94–5.23 mg/l in the first future scenario after the confluence point (*Figure 4*). Therefore, the chosen initial parameter concentration and the values for the different decay coefficients (e.g., fraction of free ammonia, coefficient of nitrification) may shift the reported ammonia concentration at the city water intake to either within or outside the legal limit of 3.7 mg/l if no uncertainty is detailed. That is less likely to happen when values vary within a given range. In that regard, the probabilistic approach is more trustworthy, because apart from allowing decision-makers to make risk-informed decisions, it dilutes the risk of the operator so that the results are less reliant on the parameter choice of each modeller. In that case, perhaps adopting confidence intervals instead of absolute limits would be useful to regulate polluters more easily.

Although some input data were not available from other studies in the area, the probabilistic model was able to generate similar central tendency results for most parameters, which evidences the viability of using literature values as input in cases of sparse data availability if the inherent uncertainty in them is considered. In fact, the bigger the significance of a parameter to the overall model behaviour, the bigger the interval of variability must be, so that the model is able to capture the true river nature.

We believe that further value is generated by a probabilistic approach since it supports the execution of a sensitivity analysis. For experts, it is a key point to understand how the model output is affected by the different input parameters and which ones must be handled with care. Moreover, if the model application is to support environmental studies or scenarios, then knowledge of parameters sensitivity is important when allocating resources to data collection, or for fine-tuning and calibrating the model. For brevity, in this proof-of-concept study, the sensitivity analysis was only carried out for DO (i.e., the variables that cause a significant change in their concentration were found), but it could be extended for other parameters depending on the final use-case, and the result of the analysis might then be different. Indeed, a complete sensitivity analysis should include all parameters that are relevant to the study. For example, an ecologist concerned with algae blooms would focus on determining the relevant variables for nitrate and phosphorus output to optimize their model.

Finally, the adherence of QUAL-PROB results to the observed data could be improved by adopting variable distributed discharge inputs along the river and by increasing the input variability for the initial pollutant concentrations and chemical reaction coefficients, notably those found to be significant. Again, though the scope of this initial study was limited, we

encourage further work and applications to tackle these issues. Regardless, the results of this study demonstrated that a probabilistic approach may be especially useful in the context of developing regions, where uncertainty in the observed data and model parameters is generally higher, while low-water quality potential damage to population health is also elevated. It should be noted that here a simple model was applied, but this methodology could be extended to any complex water quality model, including forecasting applications.

## CONCLUSION

In this research, we carried out an evaluation of a non-deterministic water quality modelling approach and compared it with a deterministic strategy through an application of a simple probabilistic model to a case study in Brazil. The main conclusions of this study are that a probabilistic approach:

- may lead to different results regarding adherence to legislation and skill at predicting observations compared to a deterministic one depending on the defined probability distribution and input parameter intervals. In that context, the variability of the parameters must be proportional to their uncertainty and significance, and sensitivity analysis must be carried out to define which parameters are sensitive to each use-case;
- seems more scientifically transparent and robust because it is explicit regarding uncertainties in measurements and choice of parameters. Moreover, it allows decision-makers to act depending on their risk aversion towards a certain situation;
- may be more trustworthy in the long-term as it dilutes the risk of the model operator: the model findings are not as reliant on the parameter choice of each modeller. In that regard, instead of legal hard limits, perhaps governmental agencies could define intervals of confidence for water quality parameters;
- is particularly useful when the risk is significantly high, such as when human health is concerned; or very low, to avoid overly conservative parameters that may prohibitively increase the cost of a project.

As this was an initial analysis, this work included several simplifications and limitations, namely the sensitivity analysis and the calibration were limited to DO; there is no real stochasticity included in the simulations; diffusive load was modelled as inflow variability; and the uniform distribution was used as a PDF for the input variability. Further work should address these issues and apply this methodology to different use cases.

Finally, we believe that the probabilistic approach may be particularly useful in developing regions because uncertainties are high due to low-resolution monitoring networks, funds are limited, and low-quality water is a serious risk to human health and ecological status due to a low prevalence of sanitation.

## DATA AVAILABILITY STATEMENT

All relevant data are available from an online repository or repositories. <https://github.com/bpbrum/qual-prob>.

## REFERENCES

- Anderson, S. R., Cole, S. J., Csima, G., Moore, R. J. & Mittermaier, M. 2019 [Towards operational joint river flow and precipitation ensemble verification: considerations and strategies given limited ensemble records](#). *Journal of Hydrology* **577**. doi:10.1016/j.jhydrol.2019.123966.
- Asadollah, S. B. H. S., Sharafati, A., Motta, D. & Yaseen, Z. M. 2021 [River Water Quality Index prediction and uncertainty analysis: a comparative study of machine learning models](#). *Journal of Environmental Chemical Engineering* **9** (1), 104599.
- Brown, J. D., Wu, L., He, M., Regonda, S., Lee, H. & Seo, D. J. 2014 [Verification of temperature, precipitation, and streamflow forecasts from the NOAA/NWS Hydrologic Ensemble Forecast Service \(HEFS\): 1. Experimental design and forcing verification](#). *Journal of Hydrology* **519**, 2869–2889. doi:10.1016/j.jhydrol.2014.05.028.
- Chapra, S. C. 1997 *Surface Water-Quality Modelling*. McGraw-Hill, New York.
- Conselho Nacional do Meio Ambiente (CONAMA) 2005 *Resolução n. 357 (Resolution n. 357)*. Brasília, Brazil.
- Costa, D. J. & Teixeira, D. 2010 [Aplicação de Modelo de Autodepuração para Avaliação da Qualidade da Água do Ribeirão do Ouro, Araraquara-SP \(Application of a self-purification model to evaluate the water quality of Ribeirão do Ouro river in Araraquara-SP\)](#). *Revista Uniara* **13** (1), 49–62.
- Fagundes, H. O., Fan, F. M., Paiva, R. C. D., Siqueira, V. A., Buarque, D. C., Kornowski, L. W., Laipelt, L. & Collischonn, W. 2021 [Sediment flows in South America supported by daily hydrologic-hydrodynamic modeling](#). *Water Resources Research* **57**, e2019WR026379.
- Fan, F. M., Schwanenberg, D., Collischonn, W. & Weerts, A. 2015 [Regional studies verification of inflow into hydropower reservoirs using ensemble forecasts of the TIGGE database for large scale basins in Brazil](#). *Journal of Hydrology: Regional Studies* **4**, 196–227.

- Instituto Trata Brasil 2018 Benefícios Econômicos e Sociais da Expansão do Saneamento Brasileiro (Economic and Social Benefits of Brazilian Sanitation Expansion). Available from: <http://tratabrasil.org.br/estudos/estudos-itb/itb/beneficios-economicos-e-sociais-da-expansao-do-saneamento-brasileiro> (accessed 26 April 2021).
- Jaun, S. & Ahrens, B. 2009 Evaluation of a probabilistic hydrometeorological forecast system. *Hydrology and Earth System Sciences Discussions* **6** (2), 1843–1877. doi:10.5194/hessd-6-1843-2009.
- Larentis, D. G. 2004 *Modelagem Matemática da Qualidade da Água em Grandes Bacias: Sistema Taquari-Antas – RS (Mathematical Modelling of the Water Quality of Large Basins: Taquari-Antas System – RS)*. Master Thesis, Universidade Federal do Rio Grande do Sul.
- Lasdon, L. S., Waren, A. D., Jain, A. & Ratner, M. 1978 Design and testing of a generalized reduced gradient code for nonlinear programming. *ACM Transactions on Mathematical Software* **4** (1), 34–50. doi:10.1145/355769.355773.
- Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models part I – a discussion of principles. *Journal of Hydrology* **10** (3), 282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6).
- Ngene, B. U., Nwafor, C. O., Bamigboye, G. O., Ogiye, A. S., Ogundare, J. O. & Akpan, V. E. 2021 Assessment of water resources development and exploitation in Nigeria: a review of integrated water resources management approach. *Heliyon* **7** (1). doi:10.1016/j.heliyon.2021.e05955.
- Richardson, D. S. 2000 Skill and relative economic value of the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society* **126**, 649–667.
- Roulin, E. 2007 Skill and relative economic value of medium-range hydrological ensemble predictions. *Hydrology and Earth System Sciences* **11** (2), 725–737. doi:10.5194/hess-11-725-2007.
- Salla, M. R., Pereira, C. E., Alamy Filho, J. E., de Paula, L. M. & Pinheiro, A. M. 2013 Estudo da autodepuração do Rio Jordão, localizado na bacia hidrográfica do Rio Dourados (Study of the self-purification of the Jordão river, located in the Dourados river basin). *Engenharia Sanitária e Ambiental* **18** (2), 105–114.
- Serviço Geológico do Brasil (CPRM) 2003 Geologia, tectônica e recursos mine-rais do Brasil: texto, mapas & SIG (Geology, Tectonics and Mineral Resources of Brazil: Text, Maps & GIS). Rio de Janeiro, Brazil.
- Shamshirband, S., Nodoushan, J. E., Adolf, J. E., Manaf, A. A., Mosavi, A. & Chau, K. 2019 Ensemble models with uncertainty analysis for multi-day ahead forecasting of chlorophyll *a* concentration in coastal waters. *Engineering Applications of Computational Fluid Mechanics* **13** (1), 91–101.
- Sharafati, A. & Pezeshki, E. 2020 A strategy to assess the uncertainty of a climate change impact on extreme hydrological events in the semi-arid Dehbar catchment in Iran. *Theoretical and Applied Climatology* **139** (1–2), 389–402.
- Spear, R. C. & Hornberger, G. M. 1980 Eutrophication in peel inlet – II. Identification of critical uncertainties via generalized sensitivity analysis. *Water Research* **14**, 43–49.
- Spear, R. C., Cheng, Q. & Wu, S. L. 2020 An example of augmenting regional sensitivity analysis using machine learning software. *Water Resources Research* **56** (4).
- Tonon, K. 2014 *Modelagem da qualidade da água utilizando os modelos Streeter-Phelps e QUAL-UFMG na bacia do Rio Lambari – Poços de Caldas – MG (Water Quality Modelling Using the Streeter-Phelps and QUAL-UFMG Models in the Lambari River Basin – Poços de Caldas – MG)*. Master Thesis, Universidade Federal de Alfenas.
- U.S. Environmental Protection Agency (EPA) 1987 *The Enhanced Stream Water Quality Models QUAL2E and QUAL2E-UNCAS*. Documentation and user model, EPA/600/3-87/007. Washington, DC, USA.
- U.S. Environmental Protection Agency (EPA) 1997 *Guiding Principles for Monte Carlo Analysis*. EPA/630/R-97/001. Washington, DC, USA.
- von Sperling, M. 2014 Estudos e modelagem da qualidade da água de rios (Study and Modelling of River Water Quality), 2nd edn. Editora UFMG, Belo Horizonte.
- Ward, S., Borden, D. S., Kabo-Bah, A., Fatawu, A. N. & Mwinkom, X. F. 2019 Water resources data, models and decisions: international expert opinion on knowledge management for an uncertain but resilient future. *Journal of Hydroinformatics* **21**, 32–44. doi:10.2166/hydro.2018.104.
- World Health Organization 2016 *Sanitation Fact Sheet, Reviewed November 2016*. Available from: <https://reliefweb.int/report/world/sanitation-fact-sheet-reviewed-november-2016> (accessed 1 February 2022).
- Zhu, Y., Toth, Z., Wobus, R., Richardson, D. & Mylne, K. 2002 The economic value of ensemble-based weather forecasts. *Bulletin of American Meteorological Society* **83**, 73–83.

First received 13 November 2021; accepted in revised form 24 April 2022. Available online 25 May 2022