Revista Brasileira de Recursos Hídricos
Brazilian Journal of Water Resources
Versão On-line ISSN 2318-0331
RBRH, Porto Alegre, v. 28, e36, 2023
Scientific/Technical Article

https://doi.org/10.1590/2318-0331.282320230035

Representative water quality parameters in a critical water basin: elements for planning and management

Parâmetros de qualidade da água representativos de uma bacia hidrográfica crítica: elementos para planejamento e gestão

Camila de Carvalho Almeida[1], Luís Otávio Miranda Peixoto[1], Marianne Schaefer França Sieciechowicz[2], Cristovão Vicente Scapulatempo Fernandes[1], Julio César Rodrigues De Azevedo[3] & Jorge Antonio Silva Centeno[1]

[1]Universidade Federal do Paraná, Curitiba, PR, Brasil
[2]Instituto de Tecnologia para o Desenvolvimento, Curitiba, PR, Brasil
[3]Universidade Tecnológica Federal do Paraná, Curitiba, PR, Brasil

E-mails: cami.almeidac@gmail.com (CCA), luisotaviopeixoto@gmail.com (LOMP), marianne.schaefer@gmail.com (MSFS)
cris.dhs@ufpr.br (CVSF), jcrazevedo@hotmail.com (JCRA), centeno@ufpr.br (JASC)

Received: April 09, 2023 - Revised: September 21, 2023 - Accepted: September 22, 2023

ABSTRACT

Changes in water quality are a consequence of water and land use dynamics. Measuring these relationships is challenging, especially when looking at distinct variables such as traditional organic matter constituents and emerging compounds. Although monitoring is necessary for proper water resource planning and management strategies, a comprehensive methodological approach is in general time-consuming and costly, which discourages decision-makers. The goal of the study was to establish consistent water quality elements based on a statistical analysis to identify the most representative parameters of a critical urban basin. The results highlighted BOD, nitrogen, and phosphorus series as chemical representative parameters; and conductivity, turbidity, and total dissolved solids as the physical parameters. In this context, these parameters represent the most significant uses of the studied area and define a more effective monitoring system that will subsidize decision-making and water resources planning and management.

Keywords: Monitoring; Principal component analysis; Urban basins; Water resources management.

RESUMO

As alterações da qualidade da água são reflexo das dinâmicas de usos da água e do solo de uma bacia e é um desafio medi-las apropriadamente, especialmente quando se olha para variáveis bastante distintas, como os constituintes tradicionais da matéria orgânica e os compostos emergentes. Apesar de o monitoramento ser essencial para um planejamento adequado dos recursos hídricos e estratégias de gestão, uma abordagem metodológica abrangente é, em geral, demorada e cara, o que desencoraja o tomador de decisão a fazê-lo. O objetivo deste estudo foi identificar os parâmetros de qualidade da água mais representativos de uma bacia urbana crítica, incluindo concentrações de compostos emergentes. Os resultados destacaram a DBO e as séries de nitrogênio e fósforo como parâmetros químicos; e, condutividade, turbidez e sólidos dissolvidos totais como físicos. Concluiu-se que esses parâmetros representam os usos do solo e da água da bacia e podem contribuir para a definição de um monitoramento mais efetivo, subsidiando a tomada de decisão, o planejamento e a gestão.

Palavras-chave: Monitoramento; Análise de componentes principais; Bacias urbanas; Gestão de recursos hídricos.

INTRODUCTION

Monitoring water quality is essential for the management of hydrological resources because it allows for a proper understanding of what happens within the basin and helps identify the most critical hydrological areas and the parameters that deviate from the optimum distribution. Therefore, it is possible to set actions and goals aiming at improving the water quality of the basin through time and space (Tripathi & Singal, 2019; Alves et al., 2018).

The provision of water with adequate quantity and quality is fundamental for multiple uses, such as energy production, and the development of agricultural or industrial activities (Ma et al., 2020). Clearly, this is challenging and could potentially enhance conflicts under extreme conditions. Thus, the continuous monitoring of the qualitative characteristics of the water should be based on specific water quality parameters that are representative of these activities.

However, there are many variables and feedback between humans and water. Understanding, selecting, and implementing actions from a more consistent and representative approach, as discussed in socio-hydrology studies (Sivapalan & Blöschl, 2015; Ceola et al., 2016; Vollmer et al., 2018; Fischer et al., 2021) is demanding and pressing.

This leads to a complexity in modeling water resources, which is worse in urban areas where land cover and land use are more dynamic than those present in rural scenarios. In this context, it is possible to highlight the discharge of contaminants of emerging concern, such as caffeine and female sex hormones (Peixoto et al., 2022; Mizukawa et al., 2019; Mizukawa, 2016; Kramer et al., 2015). These compounds can be associated directly with human activity and although they are present in aquatic environments in very low concentrations (in the order of nano and micrograms per liter), chronic exposure to these compounds is frequently associated with endocrine disruption, carcinogenicity, and overall environmental imbalance.

The presence of female sex hormones in natural waters has been associated with endocrine disruption and alterations in the metabolic and reproductive behaviors of fish. However, chronic exposure to caffeine does not present a toxicity or endocrine risk for the water biota, due to its environmental stability, low volatility, and photodegradability, caffeine has seen ample use as a chemical tracer for anthropic influence over water resources systems (Reichert et al. 2020).

It is important to highlight, that the Brazilian National Water Resources law (PNRH, officialized by the Law n° 9.433/97, Brasil, 1997) instated six management instruments and involved planning and diagnostics tools that allow for databased monitoring to serve as decision-making devices. Therefore, accurate monitoring is necessary for the formulation of management that is more realistic and indicative of the current issue, taking into account the better water quality measures that indicate human activity in the basin. A more in-depth analysis of distinct water quality parameters is still required, especially those with more organic pollution matrix (Knapik, 2009, Bitencourt, 2018).

In this context, the Brazilian National Water Agency (ANA) developed the National Water Quality Network (RNQA), as a measure to improve upon the National Program for Water Quality Assessment (PNQA). This project established the trimestral monitoring of 22 water quality parameters (physical, chemical, microbiological, and nutrients) for sparsely distributed monitoring stations in each region of the country (Agência Nacional de Águas e Saneamento Básico, 2012). Though annual investments have been carried out, this project was not completely implemented, as there have been several omissions in the water quality monitoring datasets, and some States are yet to start the monitoring efforts. In 2020, due to the Covid-19 pandemic, several stations suffered from interruptions in monitoring (Agência Nacional de Águas e Saneamento Básico, 2021).

Despite the PNQA predicting the monitoring of 22 different parameters, most of the Brazilian basins are only monitored using the Biochemical Oxygen Demand (BOD) parameter. Consequently, the expected classifications of waterbodies were mainly described by BOD (Bitencourt et al., 2019). Furthermore, when a longer monitoring effort is done effectively, the large and complex dataset produced hinders the analysis, as the variables analyzed are diverse and have codependency with each other (Chowdhury & Husain, 2020; Ma et al., 2020; Tripathi & Singal, 2019; Bitencourt, 2018; França, 2009).

One way to assess datasets is through statistical methods, and the application of multivariate statistical techniques would be the most indicated. It is possible to simplify the variability structure of the data by multivariate analysis techniques facilitating their interpretation (Mingoti, 2005; Hair Junior et al., 2009; Chowdhury & Husain; 2020) and it makes this process less subjective (Tripathi & Singal, 2019).

Among these techniques, the Principal Components Analysis (PCA) is one of the standouts. Its application allows for dimensional reduction in datasets while keeping its main characteristics. This method was implemented in several studies relating to water quality monitoring, such as Chowdhury & Husain (2020) in Canada, Tripathi & Singal (2019) and Singh et al. (2004) in India, Bu et al. (2010) in China, and Bitencourt (2018), Alves et al. (2018) and França (2009) in Brazil.

This paper aims to demonstrate that the use of PCA techniques can support the identification of key monitoring parameters in a critical basin. Allowing for better resource allocation in monitoring and contributing to the management of water resources including distinct water quality parameters representative of the human activities developed in the basin.

MATERIAL AND METHODS

Case study

The case study was developed in part of the Iguassu River basin (Figure 1), in a drainage area of about 24,500 km^2 including 68 municipalities (Instituto Brasileiro de Geografia e Estatística, 2010) with an estimated population of just over four million inhabitants (Instituto Brasileiro de Geografia e Estatística, 2021).

According to Bitencourt (2018), though this basin has a significant portion of its area dedicated to vegetation and agricultural production, its urban area is significant, especially in its Southwest region. In this basin, it is located in the Metropolitan Region of Curitiba, the capital of Paraná State in Brazil. The urbanized area has about 914 km^2 (Projeto MapBiomas, 2021), around 28% of Parana's state population and 36% of it is urban population

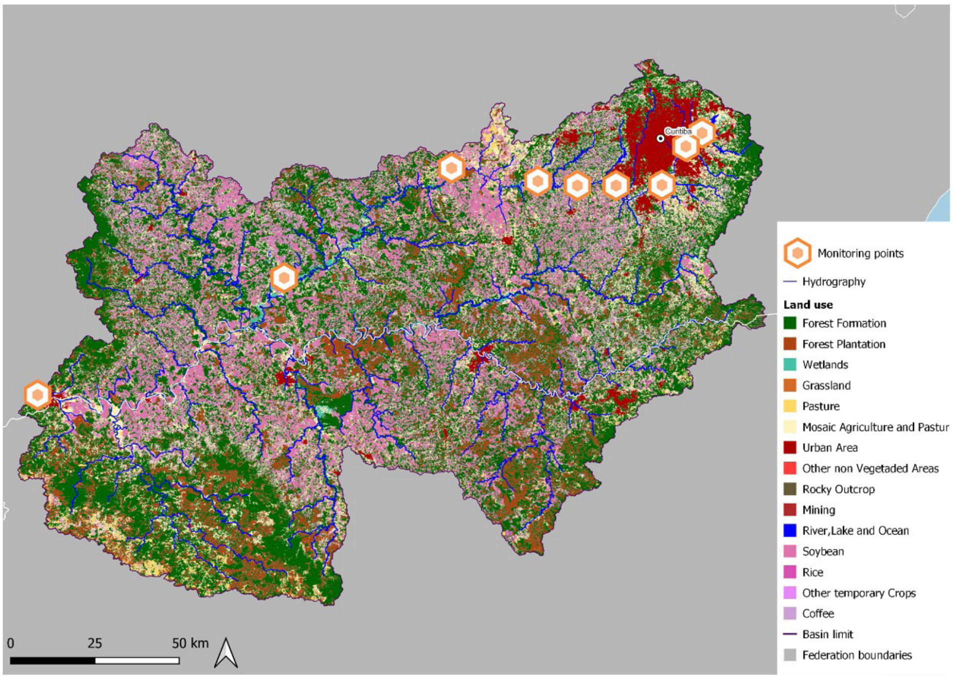

Figure 1. Upper Iguassu Basin and monitoring sites.

(Instituto Brasileiro de Geografia e Estatística, 2021) with inadequate rates of service and treatment of wastewater. Only 40 out of the 68 municipalities encompassed in the study area have information relating to domestic wastewater for the year 2020 (Brasil, 2021), and the data points to only 55% of the population having access to its treatment.

In recent decades, the study area underwent an intense process of industrialization (Stefan & Fernandes, 2020). According to the registry for major water users in Brazil, as 15% of the allowed use of surface water is intended for industrial use, while 71% is dedicated to public water supply (Agência Nacional de Águas e Saneamento Básico, 2022). The majority of the major users are concentrated in its most densely urbanized region, known as the Upper Iguassu Basin, where, historically, high water supply demands for various water uses have been hampered by the deterioration of the water quality (Stefan & Fernandes, 2020; Mizukawa et al., 2019; Bitencourt, 2018; Knapik, 2014; Instituto das Águas do Paraná, 2013; França, 2009; Porto et al., 2007).

According to Bitencourt (2018), this critical water quality scenario is reflected in the Water Body Classification instrument approved for the Upper Iguassu Basin, where the majority of the rivers are classified in the lowest quality. This classification was proposed based on the analysis and modeling results of the BOD, which was defined as the most relevant parameter to identify the two main sources of pressure onto the water resources: industrial and domestic wastewater influx.

The Upper Iguassu River Basin has been the object of study for three interdisciplinary cooperative monitoring projects which accompanied the river's water quality from 2005 until 2019, these projects were the Critical Watershed (Porto et al., 2007), Integra I (Paiva, 2012) and Integra II (which are both still ongoing). These projects monitored nine sampling sites along the Upper Iguassu River, which are presented on Figure 1.

Water quality data

The water quality dataset for each of the nine studied sampling sites is composed of parameters that had their analytical methods previously described in several studies which are presented on Table 1.

Data analysis

In this study, three Water Quality Data Analyses (WQA) were carried out using the same theoretical framework but considering different temporal monitoring conditions. The main goal is to observe the differences introduced by a more robust historical series, considering the monitoring period between each analysis and the natural improvement of technical procedures for statistical analyses, as presented on Figure 2.

The multivariate statistical of Principal Components Analysis (PCA) was the theoretical basis for all analyses. This method is used to investigate the existing relations in a set of, usually correlated, "p" variables, turning them into a new set of non-correlated variables known as the Principal Components (PCs), which are the linear combinations of the "p" original correlated variables X_1, X_2, ..., X_p which have special properties in terms of variance (Johnson & Wichern, 1998; Mingoti, 2005; Ferreira, 2008).

First, what is obtained from the transformation of the initial "p" correlated variables are "p" main compounds. However, even if "p" components are necessary to reproduce a complete variability of the system, the biggest part of this variability may be explained by a number (k) smaller than the number of main components (k < p). Thus, "k" main components explain, practically, the same quantity of information as the original "p" variables, these "p" variables may be replaced by the "k" main components, reducing the number of

Table 1. Past studies with the water quality parameters definition methods.

Project	Time Period	Sources
Critical Watersheds (Bacias Críticas)	2005-2008	Fernandes et al. (2010), França (2009), Knapik (2009)
Integra I	2008-2012	Fernandes et al. (2010), Coelho (2013), Knapik (2014), Kramer et al. (2015)
Integra II	2012-2017	Osawa et al. (2015), Coelho et al. (2017), Ide et al., (2017), Mizukawa et al. (2018), Mizukawa et al. (2019), Coelho (2019)

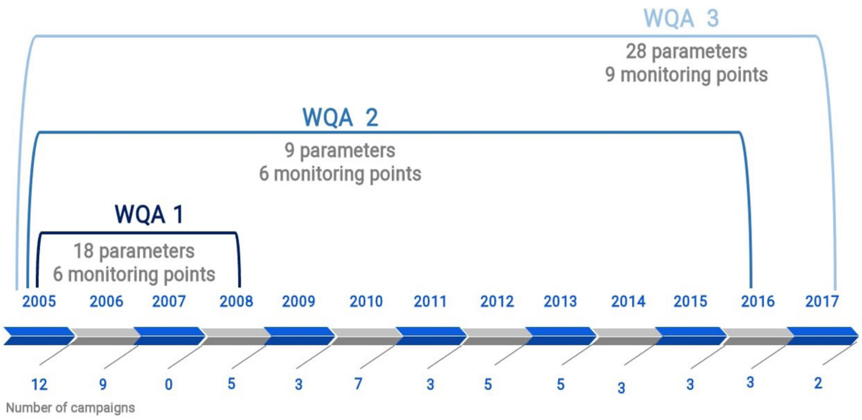

Figure 2. Method Summary.

variables of the problem and losing the lowest amount of information possible. This process is called dimensionality reduction (Johnson & Wichern, 1998; Mingoti, 2005; Ferreira, 2008).

In a dataset, when a variable presents a higher variance than the others do or the units and scales between variables are different, the PCA is started from the correlation matrix. Another similar way of solving this problem is, for example, first normalizing the data (average=0, variance=1) and then starting the analysis from the covariance matrix (Johnson & Wichern, 1998; Mingoti, 2005; Ferreira, 2008).

Water Quality Analysis 1 (WQA1)

This approach was developed, in detail, by França (2009), with the monitoring data collected from 2005 to 2009, for the sampling sites IG1 to IG6. The analyzed parameters were: dissolved oxygen (DO); turbidity (TURB); conductivity (COND); water temperature (TEMP); pH; Secchi depth; water level; biochemical oxygen demand (BOD); chemical oxygen demand (COD); total organic carbon (TOC); Flow (Q), the nitrogen series; total phosphorus (TP); and, the solid series.

The analysis was performed only on complete samples (no missing data within the example), therefore the final analyzed matrix consisted of 34 samples with 18 different parameters each, without spatial discrimination, therefore all sampling sites were analyzed collectively.

This approach was carried out on the software MATLAB 5.3, using the correlation matrix of the complete examples, after defining an exploratory data analysis of the central tendency parameters. The number of principal components was determined using the Kaiser criterion.

Water Quality Analysis 2 (WQA2)

This approach was developed as part of the study developed by Bitencourt (2018), with the monitoring data collected from 2005 to 2016, for the sampling sites IG1 to IG6. The analyzed parameters were: total dissolved solids (TDS); BOD; TP; nitrate (NO_3); nitrite (NO_2); ammoniacal nitrogen (NH_4); DO; pH; TURB. Amongst the 20 available parameters, these were chosen due to them having limits set for the waterbody classification on the Brazilian CONAMA Resolution n° 357/05 (Brasil, 2005).

Just as presented on WQA1, a matrix was composed of all normalized complete examples, resulting in a grid of 523 samples, covering 9 different parameters. There was also no spatial discrimination for the sampling sites. For the PCA implementation, the software R Core Team (2017) was used (princomp function), based on the data correlation matrix. The Kaiser criterion was also applied to determine the number of principal components.

Water Quality Analysis 3 (WQA 3)

In this approach, two scenarios were analyzed: (i) Composed of 10 traditional chemical water quality parameters: BOD, dissolved organic carbon (DOC), NH_4; NO_2; NO_3; organic nitrogen (N-Org), total nitrogen (TN), TP, orthophosphate (PO_4) and DO; (ii) Composed of 13, mostly physical, traditional water quality parameters: chemical oxygen demand (COD), volatile dissolved solids (VDS), fixed dissolved solids (FDS), volatile suspended solids (VSS), fixed suspended solids (FSS), total volatile solids (TVS), total fixed solids (TFS), Sedimentable solids (SS), total coliforms (TC), fecal coliforms (FC), conductivity, temperature, turbidity and pH.

Both datasets had their starting points on a set of 416 sample examples for all ten sampling sites, yet, not uniformly spread, temporally and spatially, among the sites and the 59 sampling campaigns. As vectors that had any missing data point were dropped from the final analyses, the resulting matrices were composed of 28 and 59 normalized complete examples for Scenarios 1 and 2, respectively.

To implement the PCA, a Python 3.7 script was created, based on the decomposition PCA function from the Scikit-learn open-source library. The Kaiser criterion was also used to estimate the number of principal components.

In this approach, a complimentary analysis was performed as to include the evaluation of caffeine, estradiol and ethinylestradiol. For such, in each scenario, the distribution of the principal components was separated into three binary groups (low and high concentrations). Thus, it would be possible to represent graphically the behavior of distribution of each contaminant about the parameters that best separate the conditions for low or high concentrations of these contaminants (Mizukawa et al., 2019; Ide et al., 2017).

The concentrations that defined the thresholds for low-high concentrations were determined by the median of the concentrations of each of the contaminants of emerging concern. In Scenario 1 these concentrations were: 2.76 ug L^{-1} for caffeine, 0.0 ug L^{-1} for both estradiol and etinylestradiol. In Scenario 2, these concentrations were: 0.83 ug L^{-1} for caffeine and 0.0 ug L^{-1} for both estradiol and etinylestradiol. Though the separation of the concentrations was done through the median, as the dataset for estradiol and ethinylestradiol had many true zeros, the separation of groups was not uniform. For estradiol, the separation for Scenarios 1 and 2, respectively, was 22:6 and 48:11 (low: high), and for ethinylestradiol, the separation was of 25:3 and 48:11.

RESULTS

Water Quality Analysis 1 (WQA 1)

The Table 2 presents the basic exploratory statistics for the data used for WQA1. This data points out that the majority of the water quality parameters presented a significant dispersion, which could be explained by the intrinsic natural variability associated to data that is influenced by temporal and spatial variability. The high dispersion presented by NO_2 could be a consequence of this parameter not being stable under environmental pressures, while for the river's flow, this high variability might have been a result of the distribution of the sampling campaigns that spread over times of low and high precipitation, and in sites that present different characteristics on the river bed.

For the distribution of BOD, the variability might be explained by the spatial distribution of population, urban gatherings, and, industrial and agricultural activity along the river. The correlation matrix for this approach is presented on the Supplementary Material. The variables of SS, NO_3 and TEMP did not present any significant correlation ($> |0.5|$) to any other variable. The resulting eigenvalues of the correlation matrix, taking into account the Kaiser criterion and the accumulated explained variance, presented on Table 3, point out to five principal components which explain 78.27% of the variance in the dataset. The variables that stood out the most for the five principal components, with a correlation higher than 0.7, are also presented on Table 3.

The first principal component expresses the aspects of organic matter degradation (TOC, NH_4), which is a result of wastewater discharge and its dynamic interaction with the transport of solids on water resources. The second principal component underlines the importance of nitrogenated compounds on organic pollution. As the flow does not present an opposite trend to that

Table 2. Basic exploratory data analysis - WQA 1.

Water quality parameter	Mean	Standard deviation	Variance	Coefficient of variation
NO_2 (mg/L)	0.14	0.19	0.04	1.36
Q (m^3/s)	21.9	24.23	587.29	1.11
BOD (mg/L)	15.17	16.55	273.84	1.09
TP (mg/L)	0.58	0.58	0.34	0.99
N-Org (mg/L)	1.26	1.24	1.54	0.98
DO (mg/L)	2.86	2.27	5.16	0.79
NO_3 (mg/L)	0.38	0.3	0.09	0.78
SS (mL/L)	0.16	0.12	0.01	0.75
NH_4 (mg/L)	4.91	3.55	12.58	0.72
TURB (NTU)	15.89	11.5	132.35	0.72
TOC (mg/L)	12.81	8.2	67.22	0.64
TTS (mg/L)	27.41	16.29	265.34	0.59
COD (mg/L)	26.34	14.93	222.78	0.57
TDS (mg/L)	127.25	72.86	5308.69	0.57
COND (µS/cm)	105.08	53.1	2819.3	0.51
Secchi (cm)	43.97	18.7	349.67	0.43
TEMP (ºC)	18.08	2.87	8.25	0.16
pH	6.98	0.32	0.11	0.05

Source: França (2009).

Table 3. Eigenvalues and variance – WQA 1.

PC	Eigenvalue	Explained variance (%)	Accumulated explained variance (%)	Correlation ≥ |0.7|
1	5.40	30.02	30.02	DO (+), TDS (-), NH_4 (-), TP (-), TOC (-), COND (-), pH (-)
2	3.46	19.25	49.27	TSS (+), N-Org (+), TURB (+)
3	2.26	12.53	61.81	TEMP (+)
4	1.70	9.47	71.28	None
5	1.26	7.00	78.27	None

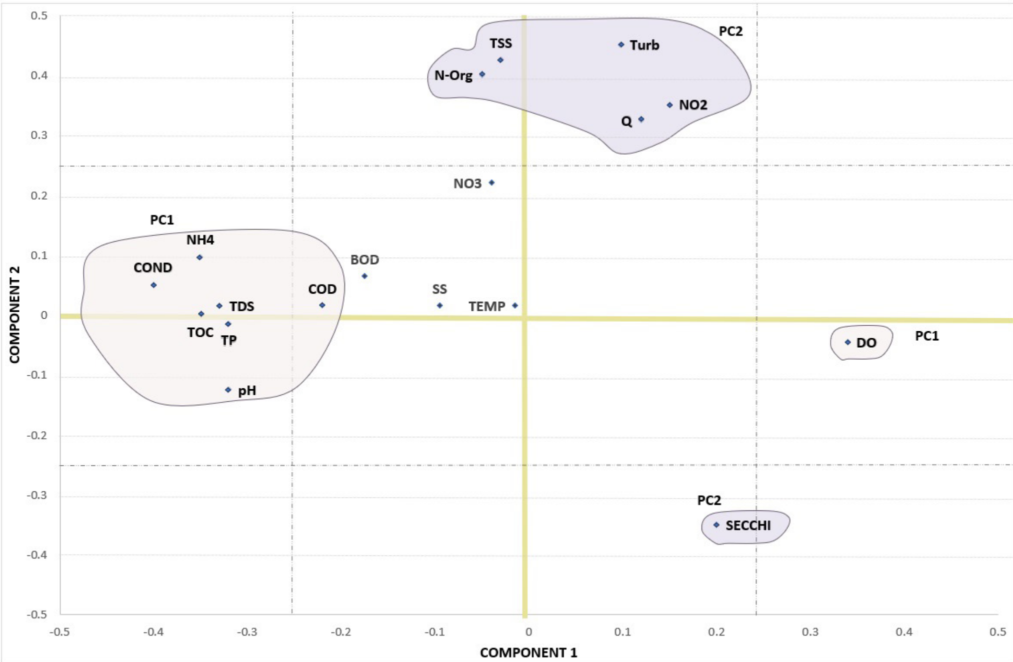

Figure 3. Weights of the variables in the principal components 1 and 2 - WQA 1.
Source: Adapted from França (2009).

of the pollution charge, the dilution of the loads of contaminants are not relevant to their distribution, therefore, this indicates that the source of these pollutants is diffuse.

The reduced nitrogenated compounds are also found within the second principal component, indicating that the influence was recent, which means, pollution sources being located close to the sampling sites. The third principal component only presented a correlation higher than 0.7 for the temperature, while the principal components 4 and 5 did not present any relevant correlation to any other variable.

The observed weights of the principal components 1 and 2 are presented on Figure 3, where groups are easily identifiable. In the group for the first principal component, the variables TDS, DOC, TOC, pH, COND, NH_4 and TP are concentrated on the left-hand side of the graph, and the DO is somewhat isolated on the right-hand side. The variables located close to the origin of the graph are the ones that presented weights that were not significant to the first two principal components. In the group for the second principal component, the variables TSS, N-Org, NO_2, TURB, and flow are group on the top portion of y-axis, while the Secchi depth is isolated on the bottom portion of the graph.

Water Quality Analysis 2 (WQA 2)

The basic exploratory statistics analysis for the data used for WQA 2 are presented on Table 4, where it can be noticed the high coefficients of variation, which was already addressed as being associated to the intrinsic variation of water quality parameters, as well as the uncertainties related to these kinds of data series, as discussed by Coelho (2019).

The correlation matrix for this approach is presented on the Supplementary Material. All resulting correlations are weak, with just the relationship between NH_4 and BOD presenting a linear correlation factor higher than |0.5|. The non-correlation behavior is common when dealing with a variety of environmental parameters, especially when there is no seasonal distinction within the data, as pointed out by Ouyang et al. (2006).

The correlation matrix's eigenvalues resulting by Kaiser criterion application and the total explained variance, are presented on Table 5. The three first principal components are responsible for 74% of the explained variance of the dataset. The variables which had the most importance for each of the five first principal components are also presented on Table 5. The minimum correlation index of just |0.5| was chosen, though defined as weak correlation,

Table 4. Basic exploratory data analysis – WQA 2.

Water quality parameter	Mean	Standard deviation	Variance	Coefficient of variation
BOD (mg/L)	16.22	17.21	296.33	1.06
TDS (mg/L)	143.03	181.82	33057.77	1.27
NH_4 (mg/L)	5.75	8.72	76	1.52
NO_2 (mg/L)	0.28	0.4	0.16	1.45
NO_3 (mg/L)	1.47	2.85	8.15	1.94
TP (mg/L)	1.03	2.12	4.51	2.06
TURB (mg/L)	43.48	53.51	2863.52	1.23
DO (mg/L)	3.76	2.36	5.57	0.63
pH	7.15	0.56	0.32	0.08

Table 5. Eigenvalues and variances – WQA 2.

PC	Eigenvalue	Explained Variance (%)	Accumulated explained variance (%)	Correlations ≥ |0.5|
1	2.21	42.91	42.91	TDS (+), BOD (-), NH_4 (-) and TURB (+)
2	1.49	20.04	62.95	TDS (+) and TURB (+)
3	1.16	10.86	73.81	pH (+)

Source: Adapted from Bitencourt (2018).

as no variable achieved a correlation index of |0.7| or higher, which reflects the high variability of environmental data.

The first principal component contemplated one parameter linked to organic matter and one to nutrient (NH4), which is coherent to the land use distribution within the Upper Iguassu River Basin, as well as the domestic wastewater produced within the urban regions of the basin. Two physical parameters (TDS and TURB) are also part of the first component (which are also presented on the second principal component) which demonstrate the importance of these factors in the variance of the entire dataset. On the third principal component, the correlation to pH once again points out to the nutrients and organic matter transport in natural waters. The observed weights of the principal components 1 and 2 are presented on Figure 4.

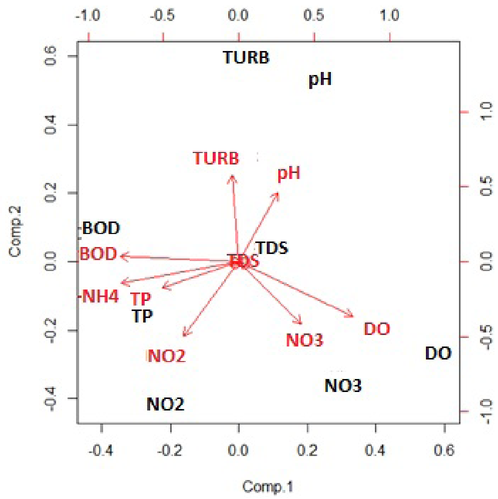

Figure 4. Weights of the variables on PC 1 and PC 2 – WQA 2.
Source: Adapted from Bitencourt (2018).

Water Quality Analysis 3 (WQA 3)

The basic exploratory statistics for the data used for WQA3 are presented on Table 6, which presents a high coefficient of variation, just as indicated on the other approaches. In this case, the largest coefficients are the ones for FC, P04 e TC. This behavior is explained by there not discrimination of sampling sites within the analysis, as these variables presented relatively lower concentrations on the IG1, IG6, IG7, IG8 and IG9 sampling sites, and significantly higher concentrations on the other ones.

The correlation matrices for this approach are presented on the Supplementary Material. In Scenario 1, the matrix only pointed out significant correlations (> |0.7|) in two pairs of variables, TN and NH_4, and, TP and PO_4, which is expected, as both pairs are composed of portions of a total concentration. For the parameters of Scenario 2, similarly, the only significant correlations were found between pairs of portions and its total concentration, these pairs were: FDS and TFS; TVS and VDS; and, TC and FC.

WQA 3 – Scenario 1

The correlation matrix's eigenvalues for Scenario 1, considering the Kaiser criterion, as well as the accumulated explained variance, are presented on Table 7. This analysis originated a total explained variance for the first four principal components of 85.16%. Relevant correlations (>|0.7|) were found between the three first principal components and nutrients, especially with nitrogenated compounds. This result can be associated with the level of urbanization in the Upper Iguassu River Basin, and consequently, with the large demographic density that is one of the main causes of water quality degradation along the course of water. This is also endorsed by the positive correlation to the BOD and the negative correlation to DO.

Table 6. Basic exploratory data analysis - WQA 3.

Water quality parameter	Mean	Standard deviation	Variance	Coefficient of variation
BOD	16.29	16.07	265.6	1.01
DOC	5.37	7.91	28.9	0.68
NH_4	8.89	5.87	79.2	1.52
NO_2	0.33	0.22	0.11	1.51
NO_3	2.10	0.94	0.11	1.51
N-Org	4.97	2.81	24.7	1.77
TN	11.1	9.50	123.4	1.17
TP	1.38	0.79	1.91	1.76
PO_4	1.30	0.46	1.68	2.81
DO	2.27	3.39	5.14	0.67
COD	31.1	21.0	440.7	0.68
TFS	111.8	63.9	4087	0.57
TVS	61.5	37.4	1396	0.61
FDS	77.0	50.8	2575	0.66
VDS	43.3	34.1	1160	0.79
FSS	35.2	41.8	1748	1.19
VSS	18.6	16.9	285.3	0.91
SS	0.32	0.75	0.57	2.38
TC	2556887	5220714	2.10^{14}	2.01
FC	564061	17073333	2.10^{13}	3.03
COND	40.3	38.9	1511	0.96
TURB	182.5	119.8	14339	0.66
TEMP	19.2	3.20	10.2	0.16
pH	7.1	0.54	0.29	0.08
CAF	4.61	2.75	21.3	1.68
E1	0.70	0.65	0.49	1.08
EE1	0.79	0.52	0.63	1.54

Table 7. Eigenvalues and variance – WQA 3 (Scenario 1).

PC	Eigenvalue	Explained Variance (%)	Accumulated explained variance (%)	Correlations ≥ |0.7|
1	1.07	43.64	43.64	BOD (+), N-Org (+), NO2(+), DOC (+), DO (-)
2	0.62	24.26	67.90	NO2(+), NH_4(+), TP (+)
3	0.52	10.21	78.11	NO_3(+)
4	0.33	7.02	85.16	None

For the first three principal components, the variables which presented the highest relevance are the ones related to the discharge of organic matter and its consequent degradation onto the river, through the influx of wastewater, being these: BOD, N-Org, DOC, NH_4 and TP. This is also corroborated to the negative correlation to the DO, as it is consumed during the microbiological degradation of organic matter.

As each datapoint is plotted on a two-dimensional graph (with the x-axis being the first principal component, and, the y-axis, the second) with a color-coded scale which determines if that point is related to an instance of low (blue) or high (red) concentration for each of the contaminants of emerging concern (estradiol, ethinylestradiol and caffeine). These plots are presented on Figure 5.

The Figure 5 allows the graphical analysis concerning the separation of low and high concentrations samples, as well as to understand the most relevant input parameters when taking into account the contaminants of emerging concern. The denoted strong sample separation presented for caffeine concentrations indicates that due to the informational entropy present on the input parameters it could be sufficient to understand the low-high concentration split for this contaminant. Therefore, it could be said that a diminished input data set, with the removal of the least relevant input parameters, would still yield enough information for understanding and interpretation of the caffeine situation within each sample. The redistribution of the hormone concentration, within the same input dataset did not offer such strong separation between divergent results, as there is not a single decision line that could divide the plane into low and high concentrations, which indicates that the relationship between the input dataset and the resulting classification is more complex than a linear behavior. This might be explained by the unbalanced number of examples in each class for both hormones.

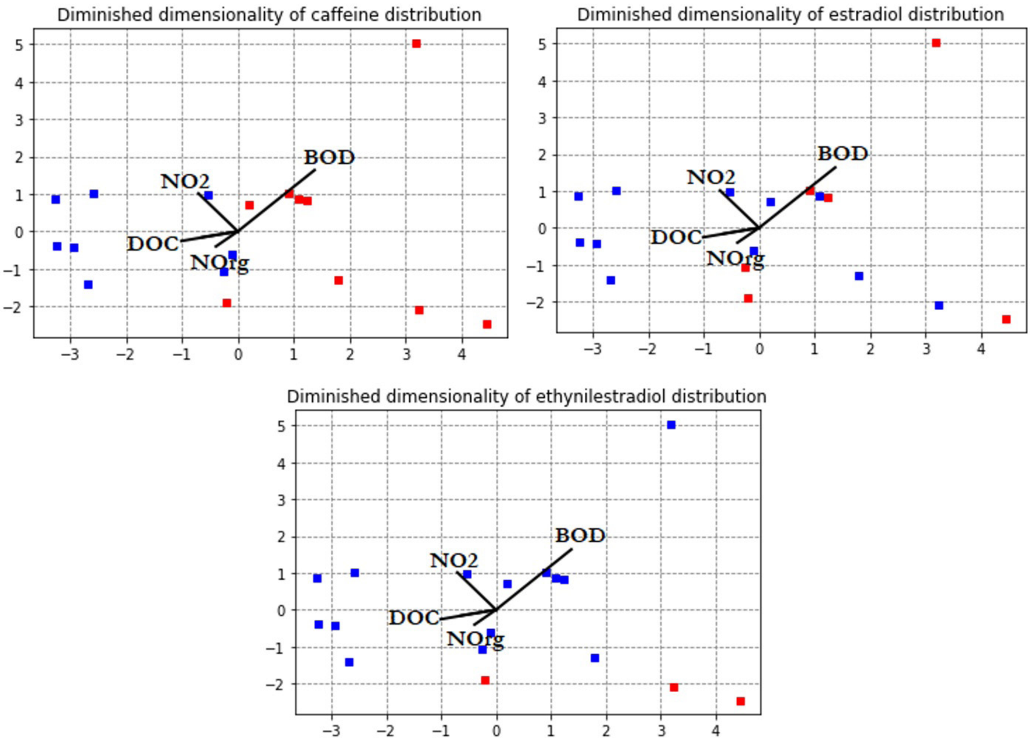

Figure 5. Redistribution of the parameters of Scenario 1, about the concentration of contaminants of emerging concern - WQA 3.

WQA 3 – Scenario 2

The correlation matrix's eigenvalues for Scenario 2, considering the Kaiser criterion, as well as the accumulated explained variance, are presented on Table 8. This analysis originated a total explained variance for the first four principal components of 78.37%. Relevant correlations (>|0.7|) were found between the three first principal components and the different portions of the solids analyzed. This behavior also corroborates the idea that the level of urbanization and the large demographic density are causes for water quality degradation along the basin. This is also endorsed by the positive correlation between the BOD and conductivity.

For the first three principal components, the variables that presented the highest relevance are the ones related to the discharge of organic matter, and its consequent degradation, onto the river, through the influx of wastewater, being these: COD, TVS, FSS and FC.

The same graphical redistribution of the samples was performed on Scenario 1, in which a bi-dimensional graph where its axes are the first and second principal components was drawn and color-coded to the classification of the concentration of each contaminant of emerging concern. These graphs are presented on Figure 6.

Just as presented in Scenario 1, these graphs suggest that only through caffeine it could be possible to create a decision line to separate classes of concentration for the samples. The separations of samples for caffeine strongly indicate that the informational entropy of the input parameters, when transformed into its two first principal components, could be enough for the determination of the state of the sample, relating to caffeine, even when some of the least relevant parameters are removed from the original input dataset. The redistribution of the hormone concentration, within the same input dataset did not offer such strong separation between divergent results, becoming still subjective, as there is not a single decision-line that could divide the plane into low and high concentrations, which indicates that the relationship between the input dataset and the resulting classification is more complex than a linear behavior. This might be explained by the unbalanced number of examples in each class for both hormones.

DISCUSSION

The extended variability of the Iguassu basin water quality datasets is exposed on the three distinct approaches, which indicates the complexity of the water resource system, considering its land use characteristics, represented by the coefficient of variation consistently being larger than 1.

Though the number of principal components varied with each approach, the results of every analysis carried out indicated that the first two principal components were capable of explaining more than 50% of the total data variation. For these principal components, the parameters which presented the highest representativeness were DO, TP, N-Org, COND, for WQA 1 and 3; TDS, NH_4, and TURB for WQA 1 and 2; and, BOD for WQA2 and WQA3.

Thus, the parameters that are directly correlated to urban pollution, especially domestic wastewater discharge, such as phosphorus, nitrogen, and organic matter (including BOD and DO), are relevant. Regarding the BOD and the nitrogen series, considering that they presented a high coefficient of variation, it was expected that they would significantly represent the variance in

Table 8. Eigenvalues and variances – WQA 3 (Scenario 2).

PC	Eigenvalue	Explained Variance (%)	Accumulated explained variance (%)	Correlations ≥ |0.7|
1	1.19	36.63	36.63	COD (+), TVS (+), COND (+), SS (+), VSS (+)
2	0.35	15.24	51.87	VDS (-), TC (+)
3	0.30	10.61	62.48	FDS (+)
4	0.30	9.64	72.12	None
5	0.26	6.25	78.37	None

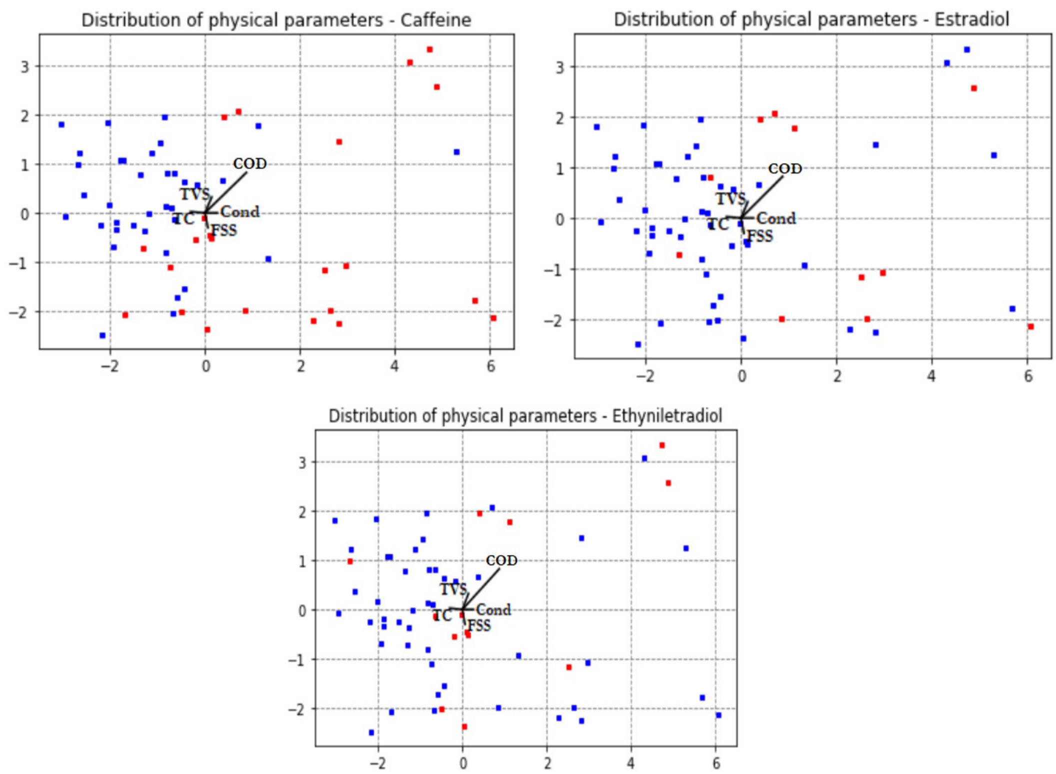

Figure 6. Redistribution of the parameters of Scenario 2, about the concentration of contaminants of emerging concern - WQA 3.

water quality in the basin, because of their unique PCA properties. The importance of nitrogen series and organic matter in the evaluation of water quality in urban areas was also identified by Miguntanna et al. (2010) and Giao et al. (2021).

The chemical parameters that the PCA indicated as the most representative of the water quality variability are among those foreseen for monitoring in the PNQA, so they are included in the Brazilian water resources management plan. Despite that, BOD was the only parameter used in the elaboration of the proposal for water body classification in the Upper Iguassu basin and this is the parameter most widely used for the planning of water resources in Brazil, as discussed by Bitencourt et al. (2019)

From the physical water quality parameters, the most relevant ones were: COND, TURB and TDS, which indicate the presence of diffuse pollution in the basin. Turbidity and TDS were also found as important parameters in PCA application realized by Alves et al. (2018) and in this case was observed that they are important in dry and rainy seasons. It is interesting to highlight that conductivity and turbidity are generally measured by the same equipment and if compared to parameters determined by laboratory analysis, they are simpler and cheaper.

Miguntanna et al. (2010) discussed that TDS and nitrogen have a high correlation and that TDS could be used as an indicator for nitrogen concentrations, both parameters were observed as significant in Iguassu River Basin and this consideration can be explored in future studies (Kozak & Fernandes, 2022; Kozak et al., 2021).

A linear relationship between BOD and the nitrogenated compounds and the DO about the chemical water quality parameters, as well as the distribution of the contaminants of emerging concern, again points to the typical relationship between these water characteristics and the discharge of domestic wastewater onto the river.

While on WQA 1 and 2 the turbidity and pH distribution showed an important relevance to the dataset variance, this was not noticed during WQA 3. This might be due to the intrinsic characteristics of the contaminants of emerging concern. For caffeine, which is used as a chemical marker for anthropic influence over water resources systems (Mizukawa et al., 2019), while pH and turbidity variations might be an indication of wastewater discharge onto water resources, these parameters may also suffer influence over other external activities such as geomorphological and sediment variations of the river (Ha et al., 2020). For the female sex hormones, while it still indicates the influx of domestic wastewater (Ide et al., 2017), the dataset for these parameters presented remarkably low concentrations throughout the study's sampling period, having the majority of the data points being zero, thus the relationship between these contaminants and the traditional parameters probably was corrupted as of the entropy extraction from the PCA analysis.

Therefore, for WQA 3, the parameters presented on Scenario 2 still point out to an internal correlation between the traditional parameters and the contaminants of emerging concern. The parameters were more intrinsically correlated to the caffeine and the hormones, thus the COD, the conductivity and the coliform series may be useful in the modeling of these contaminants.

CONCLUSION

In this study, PCA was carried out using three distinct approaches that depict the development of monitoring of a critical basin while taking into account the complex characteristics of an urban basin to identify the more representative parameters of the basin.

The use of PCA revealed that BOD and the nutrient series (phosphorus and nitrogen) are the most representative chemical parameters of this basin, while conductivity, turbidity, and total dissolved solids account for the majority of the physical dataset variance.

The weights discovered for the principal components reflect the basin's multiple uses, such as the potential for point source pollution and diffuse loads along the river's course given that, despite its significant urban sprawl, the basin's land use and land cover are not organized or uniformly distributed.

Such findings could be used to optimize monitoring for the Iguassu basin by improving resource allocation and cost reduction, in addition to contributing to more effective decisions and law-making for a more precise understanding of the basin situation. The more relevant parameters may then be used as estimators for the remaining ones, which, for this study, were satisfactory when understanding the behavior of the contaminants of emerging concern.

The results obtained demonstrated that the variance in water quality parameters concerns parameters highly associated with the land use in the studied basin and can help to understand the complex comprehension of water-human interactions. Case studies are important in this context because they can contribute to understanding how to extrapolate the variables and the interactions associated with the water-human system and how it behaves in different space and time scales (Sivapalan & Blöschl, 2015; Blair & Buytaert, 2016; Vollmer et al., 2018; Di Baldassarre et al., 2019; Fischer et al., 2021).

Finally, the implementation of water resource management depends on the acquisition of data that helps to classify and characterize the river basin water quality, which, in turn, is the reflection of the way that the population inhabiting the basin utilizes its land and water. Thus, approaches such as the ones presented in this study may optimize the definition of important parameters to be monitored, besides imbuing, in a more direct way, social behavior aspects, by the inclusion of the study of contaminants of emerging concern within the analyzed dataset.

REFERENCES

Agência Nacional de Águas e Saneamento Básico – ANA. (2012). *Panorama da qualidade das águas superficiais do Brasil* (264 p.). Brasília: ANA.

Agência Nacional de Águas e Saneamento Básico – ANA. (2021). *Conjuntura dos Recursos Hídricos no Brasil 2021*. Retrieved in 2022, July 15, from http://relatorio-conjuntura-ana-2021.webflow.io

Agência Nacional de Águas e Saneamento Básico – ANA. (2022). *Cadastro Nacional de Usuários de Recursos Hídricos - CNARH*. Retrieved in 2022, July 15, from https://www.gov.br/ana/pt-br/assuntos/regulacao-e-fiscalizacao/cadastro-nacional-de-usuarios-recursos-hidricos-cnarh

Alves, J. P. H., Fonseca, L. C., Chielle, R. S. A., & Macedo, L. C. B. (2018). Monitoring water quality of the Sergipe River basin: an evaluation using multivariate data analysis. *Brazilian Journal of Water Resources*, *23*, e27. http://dx.doi.org/10.1590/2318-0331.231820170124.

Bitencourt, C. C. A. (2018). *Integração de análise estatística multivariada e geoprocessamento: aplicação para estudos de enquadramento de corpos d'água* (Dissertação de mestrado). Universidade Federal do Paraná, Curitiba.

Bitencourt, C. C. A., Fernandes, C. V. S., & Gallego, C. E. C. (2019). Panorama do enquadramento no Brasil: uma reflexão crítica. *Revista de Gestão de Água da América Latina*, *16*, e9. http://dx.doi.org/10.21168/rega.v16e9.

Blair, P., & Buytaert, W. (2016). Socio-hydrological modelling: a review asking "why, what and how?". *Hydrology and Earth System Sciences*, *20*(1), 443-478.

Brasil. (1997). Lei Federal 9.433 de 08 de janeiro de 1997. Institui a Política Nacional de Recursos Hídricos, cria o Sistema Nacional de Gerenciamento de Recursos Hídricos, regulamenta o inciso XIX do art. 21 da Constituição Federal, e altera o art. 1º da Lei nº 8.001, de 13 de março de 1990, que modificou a Lei nº 7.990, de 28 de dezembro de 1989. *Diário Oficial [da] República Federativa do Brasil*, Brasília.

Brasil. Conselho Nacional de Meio Ambiente - CONAMA. (2005). Resolução nº 357 de 17 de março de 2005. Dispõe sobre a classificação e diretrizes ambientais para o enquadramento dos corpos de águas superficiais, bem como estabelece as condições e padrões de lançamentos de efluentes. *Diário Oficial [da] República Federativa do Brasil*, Brasília.

Brasil. Ministério do Desenvolvimento Regional - MDR. (2021). *Sistema Nacional de Informações sobre Saneamento. Série Histórica*. Retrieved in 2022, July 15, from http://app4.mdr.gov.br/serieHistorica

Bu, H., Tan, X., & Li, S. (2010). Water quality assessment of the Jinshui River (China) using multivariate statistical techniques. *Environmental Earth Sciences*, *60*, 1631-1639.

Ceola, S., Montanari, A., Krueger, T., Dyer, F., Kreibich, H., Westerberg, I., Carr, G., Cudennec, C., Elshorbagy, A., Savenije, H., Van Der Zaag, P., Rosbjerg, D., Aksoy, H., Viola, F., Petrucci, G., MacLeod, K., Croke, B., Ganora, D., Hermans, L., Polo, M. J., Xu, Z., Borga, M., Helmschrot, J., Toth, E., Ranzi, R., Castellarin, A., Hurford, A., Brilly, M., Viglione, A., Blöschl, G., Sivapalan, M., Domeneghetti, A., Marinelli, A., & Di Baldassarre, G. (2016). Adaptation of water resources systems to changing society and environment: a statement by the International Association of Hydrological Sciences. *Hydrological Sciences Journal*, *61*(16), 2803-2817.

Chowdhury, S., & Husain, T. (2020). Reducing the dimension of water quality parameters in source water: an assessment through multivariate analysis on the data from 441 suplly systems. *Journal of Environmental Management*, *274*, 111202. http://dx.doi.org/10.1016/j.jenvman.2020.111202.

Coelho, M. (2013). *Estratégia de monitoramento da qualidade da água para a gestão de recursos hídricos em bacias urbanas* (Dissertação de mestrado). Universidade Federal do Paraná, Curitiba.

Coelho, M. (2019). *Uncertainty analysis in the statistical and stochastic context of water quality time series.* (Tese de doutorado). Universidade Federal do Paraná, Curitiba.

Coelho, M., Fernandes, C. V. S., Detzel, D. H. M., & Mannich, M. (2017). Statistical validity of water quality time series in urban watersheds. *Brazilian Journal of Water Resources*, *22*, e51. http://dx.doi.org/10.1590/2318-0331.0217160071.

Di Baldassarre, G., Sivapalan, M., Rusca, M., Cudennec, C., Garcia, M., Kreibich, H., Konar, M., Mondino, E., Mård, J., Pande, S., Sanderson, M. R., Tian, F., Viglione, A., Wei, J., Wei, Y., Yu, D. J., Srinivasan, V., & Blöschl, G. (2019). Sociohydrology: scientific challenges in addressing the sustainable development goals. *Water Resources Research*, *55*(8), 6327-6355.

Fernandes, C. V. S., Knapik, H. G., Brites, A. P. Z., Dombroski, L. F., Ribeiro, D. A., & Formigoni, Y. (2010). *Projeto Integra - Disponibilidade Hídrica para Aplicação dos Instrumentos de Gestão de Recursos Hídricos: Quantidade e Qualidade de Água. Relatório de Avaliação* (No. 01). Curitiba.

Ferreira, D. F. (2008). *Estatística multivariada.* Lavras: Universidade Federal de Lavras.

Fischer, A., Miller, J. A., Nottingham, E., Wiederstein, T., Krueger, L. J., Perez-Quesada, G., Hutchinson, S. L., & Sanderson, M. R. (2021). A systematic review of spatial-temporal scale issues in sociohydrology. *Frontiers in Water*, *3*, 730169. http://dx.doi.org/10.3389/frwa.2021.730169.

França, M. S. (2009). *Análise multivariada dos dados de monitoramento de qualidade de água da Bacia do Alto Iguaçu: uma ferramenta para a gestão dos recursos hídricos* (Dissertação de mestrado). Universidade Federal do Paraná, Curitiba.

Giao, N. T., Anh, P. K., & Nhien, T. H. (2021). Spatiotemporal analysis of surface water quality in dong thap province, vietnam using water quality index and statistical approaches. *Water (Basel)*, *13*, 336. http://dx.doi.org/10.3390/w13030336.

Ha, N. T., Nguyen, H. Q., Truong, N. C. Q., Le, T. L., Thai, V. N., & Pham, T. L. (2020). Estimation of nitrogen and phosphorus concentrations from water quality surrogates using machine learning in the Tri An Reservoir, *Vietnam. Environmental Monitoring and Assessment*, *192*: 789.

Hair Junior, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2009). *Análise Multivariada de Dados* (6. ed.). Porto Alegre: Bookman.

Ide, A. H., Osawa, R. A., Marcante, L. O., Da Costa Pereira, J., & De Azevedo, J. C. R. (2017). Occurrence of pharmaceutical products, female sex hormones and caffeine in a subtropical region in Brazil. *Clean (Weinheim)*, *45*, 1600437.

Instituto Brasileiro de Geografia e Estatística – IBGE. (2010). *Censo Populacional 2010*. Retrieved in 2022, July 15, from http://www.ibge.gov.br/censo/

Instituto Brasileiro de Geografia e Estatística – IBGE. (2021). *Estimativa Populacional 2021*. Retrieved in 2022, July 15, from https://www.ibge.gov.br/estatisticas/sociais/populacao/9103-estimativas-de-populacao.html?=&t=downloads

Instituto das Águas do Paraná – AGUASPARANÁ. (2013). *Finalização do Plano das Bacias do Alto Iguaçu e Afluentes do Alto Ribeira*. Retrieved in 2022, Jan 16, from https://www.iat.pr.gov.br/sites/agua-terra/arquivos_restritos/files/documento/2020-05/finalizacao_plano.pdf

Johnson, R. A., & Wichern, D. W. (1998). *Applied multivariate statistical analysis* (4th ed.). New Jersey: Prentice Hall.

Knapik, H. G. (2009). *Reflexões sobre monitoramento, modelagem e calibração na gestão de recursos hídricos: Estudo de caso da qualidade da água da bacia do Alto Iguaçu* (Dissertação de mestrado). Universidade Federal do Paraná. Curitiba.

Knapik, H. G. (2014). *Organic matter characterization and modeling in polluted rivers for water quality planning and management* (Tese de doutorado). Universidade Federal do Paraná, Curitiba.

Kozak, C., & Fernandes, C. (2022). A influência dos eventos de precipitação e poluição difusa em rios: implicações para os instrumentos de gestão dos recursos hídricos. *REGA - Revista de Gestão de Água da América Latina*, *19*, e23.

Kozak, C., Leithold, J., Do Prado, L., Knapik, H., & De Rodrigues Azevedo, J. C., Braga, S. M., & Fernandes, C. V. S. (2021). Adaptive monitoring approach to assess dissolved organic matter dynamics during rainfall events. *Environmental Monitoring and Assessment*, *193*(7), 423.

Kramer, R. D., Mizukawa, A., Ide, A. H., Marcante, L. J. O., Santos, M. M., & Azevedo, J. C. R. (2015). Determinação de anti-inflamatórios na água e sedimento e suas relações com a qualidade da água na bacia do Alto Iguaçu, Curitiba-PR. *Revista Brasileira de Recursos Hídricos*, *20*(3), 657-667.

Ma, X., Wang, L., Yang, H., Li, N., & Gong, C. (2020). Spatiotemporal analysis of water quality using multivariate statistical techniques and the water quality identification Index for the Qinhuai River Basin, East China. *Water (Basel)*, *12*, 2764. http://dx.doi.org/10.3390/w12102764.

MapBiomas. (2021). *Projeto MapBiomas – Coleção 06 da Série Anual de Mapas de Uso e Cobertura da Terra do Brasil*. Retrieved in 2022, May 15, from http://mapbiomas.org

Miguntanna, N. S., Egodawatta, P., Kokot, S., & Goonetilleke, A. (2010). Determination of a set of surrotage parameters to assess urban stormwater quality. *The Science of the Total Environment*, *408*(40), 6251-6259. http://dx.doi.org/10.1016/j.scitotenv.2010.09.015.

Mingoti, S. A. (2005). *Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada*. Belo Horizonte: Editora UFMG.

Mizukawa, A. (2016). *Avaliação de contaminantes emergentes na água e sedimento na bacia do Alto Iguaçu* (Tese de doutorado). Universidade Federal do Paraná, Curitiba.

Mizukawa, A., Filippe, T. C., Peixoto, L. O. M., Scipioni, B., Leonardi, I .R., & Azevedo, J. C. R. (2019). Caffeine as a chemical tracer for contamination of urban rivers. *Revista Brasileira de Recursos Hídricos*, *24*, e29. http://dx.doi.org/10.1590/2318-0331.241920180184.

Mizukawa, A., Reichert, G., Filippe, T. C., Brehm, F. D. A., & Rodrigues de Azevedo, J. C. (2018). Occurrence and risk assessment of personal care products in subtropical urban rivers. *Environmental Engineering Science*, *35*(11), 1263-1272.

Osawa, R. A., Sampaio, N. M. F. M., & Azevedo, J. C. R. (2015). Determinação de fármacos anti-hipertensivos em águas superficiais na região metropolitana de Curitiba. *Revista Brasileira de Recursos Hídricos*, *20*(4), 1039-1050.

Ouyang, Y., Nkedi-Kizza, P., Wu, Q. T., Shinde, D., & Huang, C. H. (2006). Assessment of seasonal variations in surface water quality. *Water Research*, *40*, 3800-3810. http://dx.doi.org/10.1016/j.watres.2006.08.030.

Paiva, J. B. D., Fernandes, C. V. S., Porto, M. F. A., Knapik, H. G., Ribeiro, D. A., Bassanessi, K., Gonçaves, V. D., Muhlenhoff, A. P., & Almeida, C. C., 2012. *Projeto Integra: disponibilidade hídrica para aplicação dos instrumentos de gestão de recursos hídricos - quantidade e qualidade de água*. Retrieved in 2022, May 15, from http://jararaca.ufsm.br/websites/integra/0b027011f84c872a78b6ebf127a00a09.htm.

Peixoto, L. O. M., Marques, L. M. T., Mizukawa, A., & Azevedo, J. C. R. (2022). Indirect caffeine modeling in urban rivers. *Ambiente & Água - An Interdisciplinary. Journal of Applied Sciences*, *17*(2), e2810.

Porto, M. F. A., Fernandes, C. V. S., Knapik, H. G., França, M. S., Brites, A. P. Z., Marin, M. C. F. C., Machado, F. W., Chella, M. R., Sá, J. F., & Masini, L. (2007). *Bacias Críticas: Bases Técnicas para a definição de Metas Progressivas para seu Enquadramento e a Integração com os demais Instrumentos de Gestão.* Curitiba: UFPR – Departamento de Hidráulica e Saneamento. (FINEP/ CT-HIDRO). Projeto concluído.

R Core Team. (2017). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved in 2022, May 15, from http://www.R-project.org/.

Reichert, G., Mizukawa, A., Antonelli, J., De Almeida Brehm, F., Filippe, T., & Rodrigues De Azevedo, J. (2020). Determination of parabens, triclosan, and lipid regulators in a subtropical urban river: effects of urban occupation. *Water, Air, and Soil Pollution*, *231*(3), 1-11.

Singh, K. P., Malik, A., Mohan, D., & Sinha, S. (2004). Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India) – a case study. *Water Research*, *38*(18), 3980-3992.

Sivapalan, M., & Blöschl, G. (2015). Time scale interactions and the coevolution of humans and water. *Water Resources Research*, *51*(9), 6988-7022. http://dx.doi.org/10.1002/2015WR017896.

Stefan, E., & Fernandes, C. V. S. (2020). Water availability and water reuse: a new approach for water resources management. In: UNESCO. (Org.), *Water reuse within a circular economy context* (Vol. 2, pp. 71-83). Paris: UNESCO.

Tripathi, M., & Singal, S.K. (2019). Use of principal component analysis for parameter selection for development of a novel water quality index: a case study of River Ganga India. *Ecological Indicators*, *96*(Pt 1), 430-436.

Vollmer, D., Shaad, K., Souter, N. J., Farrell, T., Dudgeon, D., Sullivan, C. A., Fauconnier, I., MacDonald, G. M., McCartney, M. P., Power, A. G., McNally, A., Andelman, S. J., Capon, T., Devineni, N., Apirumanekul, C., Ng, C. N., Rebecca Shaw, M., Wang, R. Y., Lai, C., Wang, Z., & Regan, H. M. (2018). Integrating the social, hydrological and ecological dimensions of freshwater health: The Freshwater Health Index (2018). *The Science of the Total Environment*, *627*(18), 304-313. http://dx.doi.org/10.1016/j.scitotenv.2018.01.040.

Authors contributions

Camila de Carvalho Almeida: Conceptualization, organization, and writing. Carrying out the Water Quality Analysis 2.

Luís Otávio Miranda Peixoto: Carrying out the Water Quality Analysis 3, writing and editing.

Marianne Schaefer França Sieciechowicz: Carrying out the Water Quality Analysis 1, review and editing.

Cristovão Vicente Scapulatempo Fernandes: Supervision, writing, and review.

Julio César Rodrigues De Azevedo: Supervision and review.

Jorge Antonio Silva Centeno: Supervision and review.

Editor in-Chief: Adilson Pinheiro

Associated Editor: Stephan Fuchs

SUPPLEMENTARY MATERIAL

Supplementary material accompanies this paper.

Supplementary material

This material is available as part of the online article from https://doi.org/10.1590/2318-0331.282320230035