

XV SIMPÓSIO DE RECURSOS HÍDRICOS DO NORDESTE

AVALIAÇÃO DA QUALIDADE DA ÁGUA DO EIXO LESTE DO PROJETO DE INTEGRAÇÃO DO RIO SÃO FRANCISCO USANDO ANÁLISE ESTATÍSTICA MULTIVARIADA

Hidaiane Fayga Matias Caldas¹; Douglas Nascimento Brandão²; Ariane Silva Cardoso³; Maria do Carmo Sobral⁴ & Davi Tadeu Borges Marwell⁵

RESUMO – Para investigar a qualidade da água ao longo do Eixo Leste do Projeto de Integração do rio São Francisco, foram aplicadas a Correlação de Pearson e a Análise de Componentes Principais - PCA aos dados de 11 variáveis ao longo de 19 pontos de amostragem referentes às campanhas 22 e 23 do Programa de Monitoramento de Qualidade da Água e Limnologia do empreendimento. A correlação de *Pearson* possibilitou a investigação do grau de correspondência entre as variáveis, enquanto o uso da PCA promoveu a redução do número de variáveis significativas e permitiu a identificação de cinco componentes principais que explicam 76,54% da variabilidade dos dados. A maioria das variáveis classificadas como representativas são relacionadas aos efeitos de poluição da água, como DBO, fósforo total, nitrogênio total e clorofila-a. Diante disto, foi possível uma avaliação mais específica dos problemas que afetam as bacias do projeto em relação à qualidade das águas, destacando a importância de medidas de preservação e saneamento básico na região.

ABSTRACT– In order to investigate the water quality along the East Axis of the São Francisco River Integration Project, Pearson's Correlation and Principal Component Analysis - PCA were applied to the data of 11 variables over 19 sampling points for the campaigns 22 and 23 of the project's Water Quality and Limnology Monitoring Program. Pearson's correlation made it possible to investigate the degree of correspondence between variables, while the use of PCA reduced the number of significant variables and allowed the identification of five main components that explain 76,54% of the data variability. Most of the variables classified as representative are related to the effects of water pollution, such as BOD, total phosphorus, total nitrogen and chlorophyll-a. Thus, a more specific assessment of the problems affecting the project's basins in relation to water quality was possible, highlighting the importance of preservation and basic sanitation measures in the region.

Palavras-Chave – PCA; Bacia do Rio São Francisco; Saneamento.

1) Mestranda do Programa de Pós-Graduação em Engenharia Civil/ Universidade Federal de Pernambuco (UFPE). Rua Acadêmico Hélio Ramos, S/N Cidade Universitária CEP 50740-530 - Recife - PE - Brasil. Fone: 81 2126-8744. E-mail: hidaianecaldas@gmail.com

2) Graduando em Engenharia Civil/ Universidade Federal de Pernambuco (UFPE). Rua Acadêmico Hélio Ramos, S/N Cidade Universitária CEP 50740-530 - Recife - PE - Brasil. Fone: 81 2126-8744. E-mail: douglasnascimentobrandao@gmail.com

3) Doutoranda do Programa de Pós-Graduação em Engenharia Civil/ Universidade Federal de Pernambuco (UFPE). Rua Acadêmico Hélio Ramos, S/N Cidade Universitária CEP 50740-530 - Recife - PE - Brasil. Fone: 81 2126-8744. E-mail: arianecardoso8@gmail.com

4) Professora do Programa de Pós-Graduação em Engenharia Civil/ Universidade Federal de Pernambuco (UFPE). Rua Acadêmico Hélio Ramos, S/N Cidade Universitária CEP 50740-530 - Recife - PE - Brasil. Fone: 81 2126-8744. E-mail: mariadocarmo.sobral@gmail.com

5) Coordenador de Gestão Ambiental do Ministério de Desenvolvimento Regional. Esplanada dos Ministérios, Bloco E, Brasília - DF. Fone: 61 2034-4238. E-mail: davi.marwell@integracao.gov.br

1 – INTRODUÇÃO

A água é um bem fundamental para a vida desde as atividades do cotidiano até o ponto de influenciar no desenvolvimento de uma nação, como na sua contribuição para geração de energia e produção de alimentos. O desequilíbrio entre as demandas e as disponibilidades hídricas regionais constitui um dos principais problemas de alocação e distribuição de água que os sistemas de gerenciamento de recursos hídricos necessitam resolver. Para tanto, pode-se tomar como exemplo a região Nordeste do Brasil, que concentra 28% da população do país e, em contrapartida, apenas 3% dos recursos hídricos brasileiros (AZEVEDO *et al.*, 2005).

Nesse contexto, dentre as diversas propostas que foram estudadas para aumentar a disponibilidade hídrica na região, o governo brasileiro considerou como melhor opção levar parte do volume das águas do rio São Francisco para as região mais afetadas pela estiagem de água no Nordeste, dando origem, assim, ao Projeto de Integração do rio São Francisco com as Bacias Hidrográficas do Nordeste Setentrional (PISF), realizado pelo antigo Ministério da Integração Nacional e atual Ministério do Desenvolvimento Regional.

Esse projeto é um empreendimento de infraestrutura hídrica para transferência de água entre bacias hidrográficas, onde dois sistemas independentes, denominados Eixo Norte e Eixo Leste, compostos de canais, estações de bombeamento de água, pequenos reservatórios e usinas hidrelétricas para autossuprimento estão em vias de conclusão de obras. Trata-se de ação com a finalidade de atender o fornecimento de água para vários usos como abastecimento humano, irrigação, dessedentação de animais, criação de peixes e de camarão em municípios do Semiárido, do Agreste Pernambucano e da Região Metropolitana de Fortaleza (BRASIL, 2004).

Sabendo-se que a qualidade de água tem grande impacto na saúde pública e, de modo mais abrangente, na qualidade de vida da população, a garantia de que o recurso chegue com qualidade adequada após a combinação de diferentes bacias é essencial para o equilíbrio e o funcionamento dos ecossistemas. Estudos mostram que as águas bombeadas influirão positivamente na qualidade das águas dos rios e açudes receptores (BRASIL, 2004), mas o estado de conservação dos mananciais e dos ecossistemas em seu entorno são determinantes que isso ocorra realmente.

Diante do exposto, este trabalho tem por objetivo evidenciar possíveis correlações entre amostras e variáveis, além de analisar e identificar os principais parâmetros de qualidade da água e as possíveis fontes poluidoras relacionadas à variação espacial ao longo do Eixo Leste do PISF, através da aplicação das técnicas estatísticas de correlação de *Pearson* e Análise de Componentes Principais.

2 – METODOLOGIA

2.1 – Área de estudo

Segundo o Ministério da Integração Nacional (BRASIL, 2004), o Eixo Leste do PISF foi projetado com o objetivo de levar água para cerca de 4,5 milhões de pessoas em 168 municípios afetados pela seca nos estados de Pernambuco e da Paraíba. O trecho possui 217 km de extensão, e tem seu desenvolvimento a partir do canal de captação no reservatório Itaparica, localizado em Floresta (PE). Ao longo de seu curso, a água passa por 12 reservatórios projetados de pequeno porte, e serão construídas ramificações que devem conduzir para os açudes Barra do Juá e Poço da Cruz, além do Ramal do Agreste, no estado de Pernambuco. Ao ocorrer efetivamente o evento de mistura de bacias, o volume de água é conduzido de forma a incrementar a vazão do rio Paraíba, e distribuído nos açudes Poções, Camalaú, Epitácio Pessoa e Acauã.

Os dados físico-químicos e biológicos para avaliação da qualidade da água que foram utilizados nesta pesquisa são provenientes do Programa de Monitoramento da Qualidade da Água e Limnologia – PBA 22 do PISF, que foi previsto como uma condicionante do IBAMA para o processo de licenciamento ambiental do empreendimento, e é de responsabilidade do Ministério do Desenvolvimento Regional em parceria com a Universidade Federal de Pernambuco – UFPE. Para esta pesquisa, foram selecionadas 19 estações amostrais (Tabela 1), que fornecem dados do trecho entre os reservatórios projetados até o açude Epitácio Pessoa, na bacia do rio Paraíba. A localização dos pontos que foram selecionados neste trabalho são mostrados na Figura 1.

2.2 – Amostragem e análises físico-químicas

Foram analisados ao todo 11 variáveis de qualidade da água: temperatura (Temp), oxigênio dissolvido (OD), pH, turbidez (Turb), clorofila-a (Cl-a), coliformes termotolerantes (Coli), demanda bioquímica de oxigênio (DBO), fósforo total (PT), nitrogênio total (NT), sólidos dissolvidos totais (SDT) e sólidos suspensos totais (SST). Os resultados selecionados são relativos às campanhas 22 e 23 do PBA 22 do PISF, realizadas respectivamente em períodos chuvoso e seco do ano de 2017. As metodologias que fornecem os resultados das variáveis aferidas são descritas na Tabela 1.

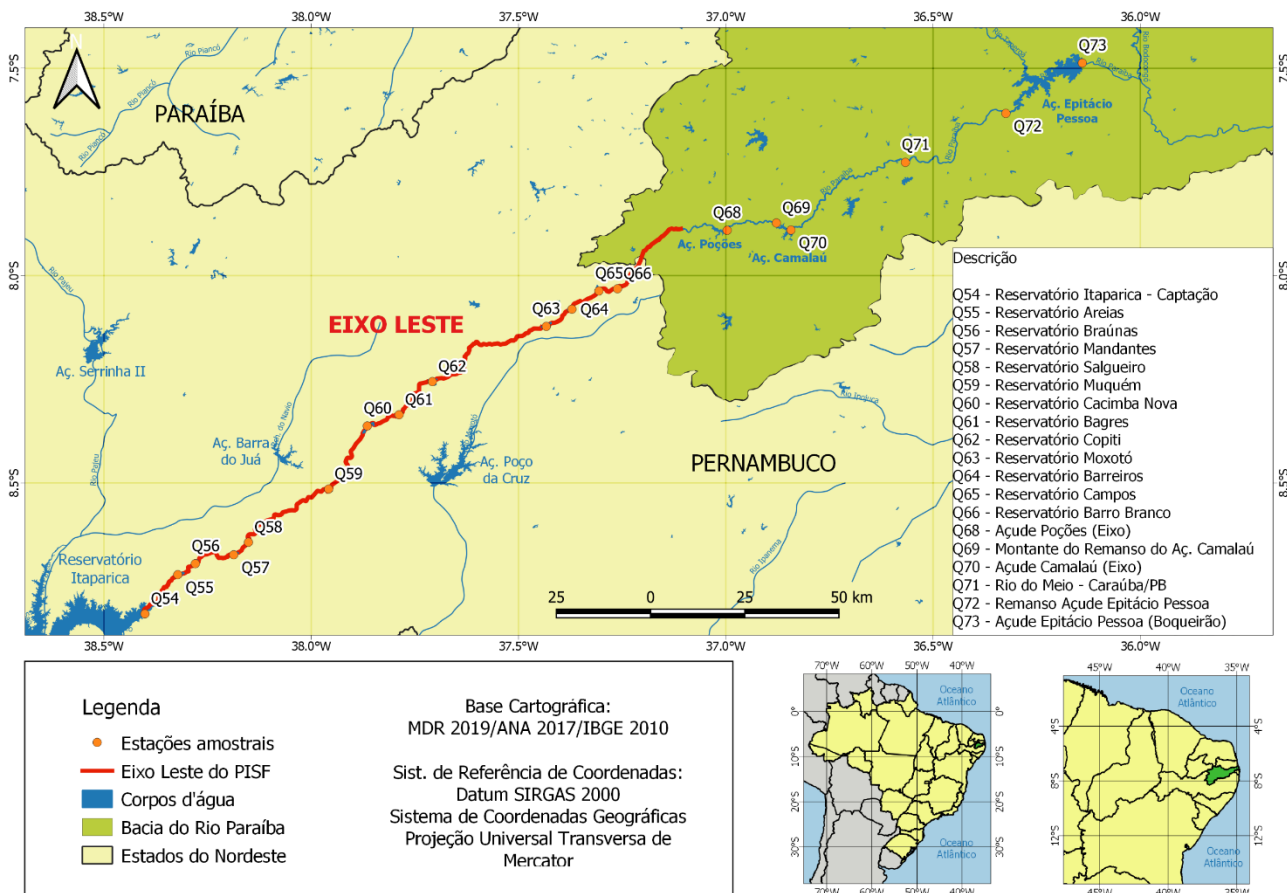


Figura 1 - Localização das estações de monitoramento da qualidade da água avaliadas nesta pesquisa ao longo do Eixo Leste.

Tabela 1 - Variáveis físico-químicas e biológicas com suas respectivas metodologias de análise.

Variável	Unidade	Referência do Método
Temp	°C	Sonda multiparamétrica (Horiba, modelo B-213)
OD	mg/L O ₂	
pH	-	
Turb	UNT	POPMET391, SMWW 22ª Edição 2012 - Método 10200 H
Cl-a	µg/L	
Coli	NMP/100 ml	POPMET397, SMWW 22ª Edição 2012
DBO	mg/L O ₂	POPMET357, SMWW 22ª Edição 2012 - Método 5210 B
PT	mg/L	POPMET375, SMWW 23ª Edição 2017 - Métodos 3030 E e 3120 B
NT	mg/L N	POPMET348
SDT	mg/L	POPMET236, SMWW 22ª Edição 2012 - Método 2540
SST	mg/L	POPMET354, SMWW 22ª Edição 2012 - Método 2540 D

2.3 - Análise estatística dos dados

Para o tratamento estatístico dos dados avaliados neste estudo, foram utilizados recursos de estatística multivariada como correlação de *Pearson* e análise de componentes principais (Principal Component Analysis - PCA), utilizando-se o programa computacional “R” versão 3.6.1 (R DEVELOPMENT CORE TEAM, 2019) para análise dos resultados. Foi utilizado o pacote FactoMineR, que executa automaticamente a padronização dos dados para eliminar o efeito de escala e unidades, através da conversão das variáveis para escores padrões. O nível de significância obtido para a análise estatística foi de $p < 0,05$.

A correlação linear de *Pearson* foi realizada com o objetivo de investigar o grau de dependência ou independência de uma variável em relação à outra. Já com a PCA é possível reduzir a dimensionalidade dos dados através da criação poucas variáveis-chave (compostas pelas variáveis originais) que caracterizam ao máximo possível a variação no conjunto de dados, ou seja, com perda mínima de dados (GOTELLI & ELLISON, 2011).

3 – RESULTADOS E DISCUSSÕES

Na Figura 2 é mostrada a matriz de correlação de *Pearson* entre as variáveis sólidos suspensos totais (SST), oxigênio dissolvido (OD), fósforo total (PT), temperatura (Temp), clorofila-a (Cl-a), turbidez (Turb), nitrogênio total (NT), sólidos dissolvidos totais (SDT), demanda bioquímica de oxigênio (DBO), coliformes termotolerantes (Coli) e pH. Os resultados obtidos indicam que a maioria das correlações obtidas não foram estatisticamente significativas a um nível de significância de 0,05, revelando que há evidências inconclusivas sobre a significância da associação entre as variáveis.

O coeficiente de *Person* é uma medida estatística que tem como finalidade expressar o grau da correlação entre duas variáveis linearmente. Esse coeficiente de correlação pode variar entre -1 e +1. Quanto maior for o valor absoluto do coeficiente, mais forte é a relação entre as variáveis. Nos valores encontrados, se destacam a correlação positiva moderada entre a DBO e os sólidos dissolvidos totais (0,63), e a correlação negativa moderada entre o pH e a clorofila-a (-0,55). A relação negativa entre as variáveis indica que conforme uma aumenta, a outra diminui.

Discorrendo a respeito das relações citadas acima, a exemplo da correlação positiva entre DBO e sólidos dissolvidos totais, temos que a DBO é um parâmetro que determina a quantidade de oxigênio consumida ou necessária na oxidação da matéria orgânica através da decomposição microbiana, enquanto o sólido dissolvido total representa toda a matéria dissolvida, onde entre essas está a matéria orgânica (VON SPERLING, 1996; TUNDISI & MATSUMURA-TUNDISI, 2008).

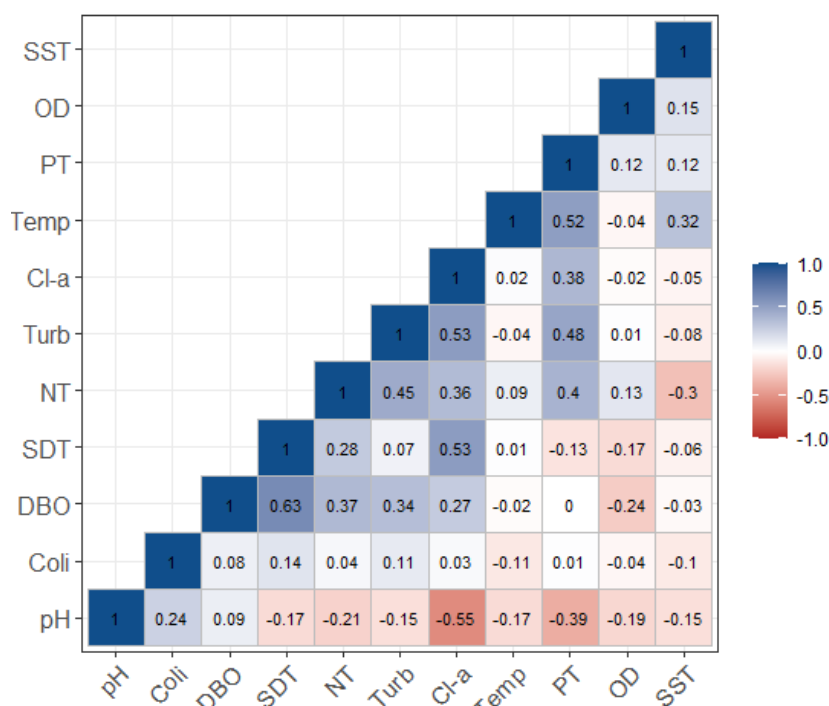


Figura 2 - Matriz de correlação de Pearson das variáveis estudadas.

Com a aplicação da Análise de Componentes Principais à matriz de dados, identificou-se que o melhor comportamento das variáveis relacionadas à qualidade da água foi aquele composto pelas cinco componentes iniciais. A Tabela 2 mostra que esses grupos, ao todo, explicam 76,54% da variância total dos dados, ou seja, as informações concentradas nessas cinco dimensões representam 76,54% das informações antes diluídas em 11 dimensões (número de variáveis analisadas). Em dados ecológicos típicos, a maioria da variação nos dados é capturada pelas primeiras componentes principais, podendo-se descartar as demais que explicam a variação residual (GOTELLI & ELLISON, 2011). Isso implica que a maior parte da variabilidade e das componentes pode ser usada para indicar o processo que afeta a qualidade das águas, sem perda de características significativas.

Para analisar as cargas encontradas, deve ser levado em consideração que os valores com maior proximidade a 1 ou -1 são aqueles que possuem maior influência na variabilidade dos dados, sendo então classificados como os mais representativos. A primeira componente (PC1) explica 26,51% da variabilidade dos dados, sendo representada pelas seguintes variáveis: DBO, pH, turbidez, fósforo total, nitrogênio total, sólidos dissolvidos totais e clorofila-a. A maioria dessas variáveis são relacionadas aos efeitos de poluição da água, como DBO, fósforo e nitrogênio totais e a clorofila-a, que é considerada a principal variável indicadora de estado trófico dos ambientes aquáticos (CETESB, 2019). De acordo com Ferreira (2016), os corpos hídricos da bacia do Paraíba (Pontos Q68, Q69, Q70, Q71, Q72 e Q73) sofreram nos últimos anos com diminuição no nível de água devido à seca na região, além da poluição por esgotos domésticos e agrotóxicos.

Tabela 2 - Cargas das componentes principais, porcentagens de variância explicada e acumulada para as amostras das campanhas 22 e 23 do monitoramento da qualidade da água do PISF.

Variáveis	PC1	PC2	PC3	PC4	PC5
OD	0,00	-0,44	-0,38	-0,34	0,50
DBO	0,53	0,57	0,33	0,16	0,03
Temp	0,22	-0,55	0,44	0,50	-0,15
pH	-0,52	0,48	-0,12	0,52	0,01
Turb	0,70	0,02	-0,37	0,20	0,03
PT	0,61	-0,58	-0,16	0,38	-0,07
NT	0,70	0,09	-0,33	0,07	-0,16
SDT	0,54	0,52	0,48	-0,22	0,17
SST	-0,07	-0,49	0,57	0,10	0,47
Cl-a	0,82	0,02	0,06	-0,29	0,08
Coli	0,06	0,35	-0,24	0,44	0,63
Autovalor	2,92	2,02	1,34	1,18	0,96
% de variância explicada	26,51	18,4	12,19	10,74	8,70
% de variância acumulada	26,51	44,91	57,1	67,84	76,54

A bacia doadora do PISF presumivelmente também tem sua parcela de contribuição para a qualidade da água no Eixo Leste, uma vez que são instaladas pisciculturas que aportam nutrientes à água do Reservatório Itaparica (Ponto Q54), além da existência de perímetros irrigados nas suas proximidades e carência de tratamento de esgotos nas cidades por onde passa o rio São Francisco (GUNKEL *et al*, 2018; SILVA *et al*, 2018; SILVA, 2019). Os reservatórios projetados estão em condições de recém instalação, e ainda demandam estudos para sua avaliação.

A Figura 3 representa o *biplot* onde se visualizam simultaneamente as variáveis de qualidade da água e as unidades de estudo sobre o espaço das componentes principais PC1 e PC2. As unidades de estudo se mostram em cinza para melhor visualização, e os números são etiquetas que coincidem com a ordem do dado na matriz do banco de dados, enquanto as flechas correspondem às variáveis (representadas como vetores). A posição de cada unidade de estudo no espaço de componentes principais é dada por suas coordenadas, denominadas *scores*.

Segundo Palacio *et al.* (2020) para interpretação do *biplot*, tem-se as seguintes regras: (1) as unidades de estudo que são próximas no espaço, têm características similares nas suas variáveis; (2) o cosseno do ângulo entre dois vetores corresponde a sua correlação; (3) o cosseno do ângulo entre um vetor e uma componente principal corresponde a sua correlação; (4) uma unidade de estudo na direção e sentido de uma variável e afastada do centro de origem tem um valor alto para essa variável.

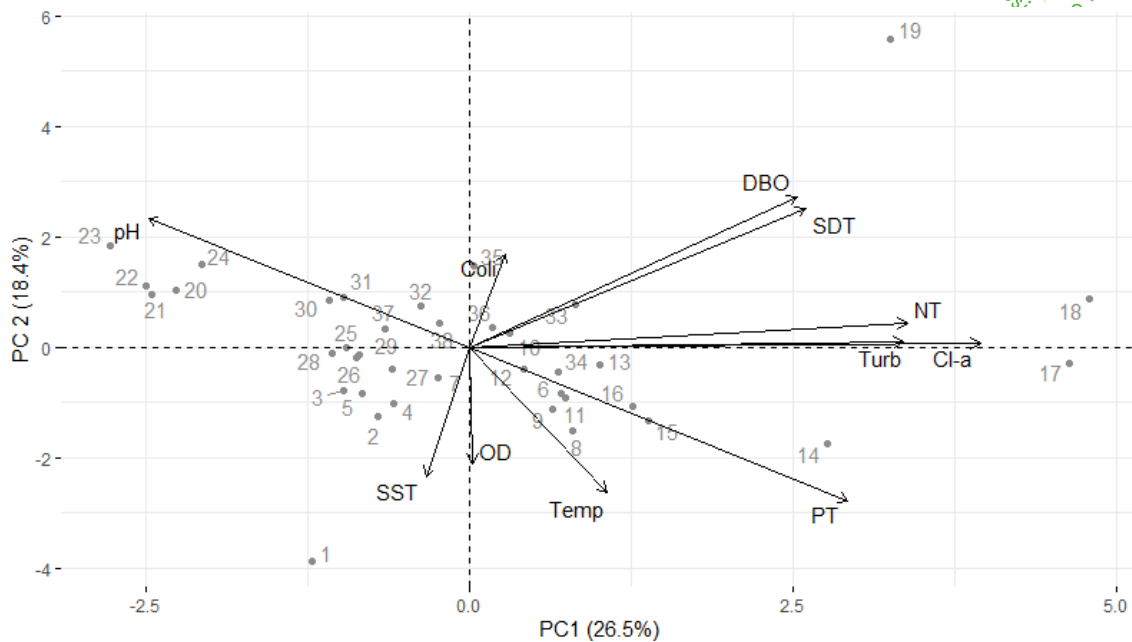


Figura 3 - Biplot da PC1 e PC2 correspondente à matriz de dados das amostras das campanhas 22 e 23 do monitoramento da qualidade da água do PISF.

Assim, por exemplo, as variáveis clorofila-a, turbidez e nitrogênio total se encontram muito correlacionadas entre si e de forma positiva, enquanto ambas se encontram levemente correlacionadas e de forma negativa com o pH. Da mesma forma, clorofila-a, turbidez e nitrogênio total são as variáveis que se encontram mais correlacionadas com a PC1 e também as que mais contribuem com este componente. Interpretando a relação entre as variáveis e as unidades de estudo de forma conjunta, percebe-se que as unidades posicionadas na extremidade esquerda do biplot têm características similares entre si, além de altos valores para a variável pH (etiquetas 20, 21, 22, 23 e 24). O mesmo ocorre com as unidades da extrema direita, posicionadas de forma central na direção dos vetores correspondentes a clorofila-a, turbidez e nitrogênio total (etiquetas 17 e 18). As unidades que se encontram mais isoladas no gráfico apresentam os maiores valores identificados nas duas campanhas para as variáveis sólidos suspensos totais (etiqueta 1), fósforo total (etiqueta 14) e, sólidos dissolvidos totais e demanda bioquímica de oxigênio (etiqueta 19).

4 – CONCLUSÕES

A aplicação da técnica de estatística multivariada Análise de Componentes Principais aos dados do monitoramento da qualidade da água do PISF permitiu a identificação das variáveis que explicam 76,5% da variância dos dados, distribuídas em cinco componentes principais.

Os resultados demonstraram que a maioria das variáveis representativas para a área estão relacionadas aos efeitos da poluição de origem antropogênica, como demanda bioquímica de oxigênio, fósforo total, clorofila-a e nitrogênio total. Essa resposta confirma a importância da adoção efetiva de medidas em relação ao saneamento básico da região, além do manejo de atividades que aportam nutrientes causadores de eutrofização, como a piscicultura e agricultura.

AGRADECIMENTOS

Os autores agradecem ao Ministério do Desenvolvimento Regional pela concordância no uso dos dados referentes ao PISF, bem como à Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco – FACEPE e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPQ pela concessão de bolsas de estudo aos primeiros autores.

REFERÊNCIAS

AZEVEDO, L.G.T., PORTO, R.L.L., MÉLLO JUNIOR, A.V., PEREIRA, J.G., ARROBAS, D.L.P., NORONHA, L.C., PEREIRA, L.P. (2005). *Série Água Brasil - Transferência de águas entre bacias hidrográficas*. Banco Mundial. Brasília - DF.

COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO (2019). Apêndice E: Significado ambiental e sanitário das variáveis de qualidade das águas e dos sedimentos e metodologias analíticas e de amostragem. In: CETESB. *Qualidade das Águas Interiores no Estado de São Paulo 2018*. São Paulo – SP, pp. 1-58

FERREIRA, A. L. N. (2016). Análise integrada da qualidade de água dos corpos hídricos do projeto de integração do Rio São Francisco no nordeste do Brasil. Tese (Doutorado) – Universidade Federal de Pernambuco. Programa de Pós-Graduação em Engenharia Civil, 2016.

GOTELLI, N.J.; ELLISON, A.M. (2011). *Princípios de estatística em ecologia*. Arned Porto Alegre-RS, 528 p.

GUNKEL, G.; SELGE, F.; KEITEL, J. et al. (2018). “Water management and aquatic ecosystem services of a tropical reservoir (Itaparica, São Francisco, Brazil)”. *Regional Environmental Change* 18, 1913–1925. <https://doi.org/10.1007/s10113-018-1324-8>

MINISTÉRIO DA INTEGRAÇÃO NACIONAL (2004). *Relatório de Impactos Ambientais do Projeto de Integração do São Francisco*. Brasil, 136 p.

PALACIO, F. X.; APODACA, M. J.; CRISCI, J. V. (2020). *Análisis multivariado para datos biológicos: teoría y su aplicación utilizando el lenguaje R*. 1ª ed. Fundación de Historia Natural Félix de Azara, Ciudad Autónoma de Buenos Aires, 268p.

R DEVELOPMENT CORE TEAM. (2010). *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing, Disponível em: <https://www.r-project.org/>. Acesso em: 01 set. 2019.

SILVA, G.M.N; CARVALHO, R.M.C.M.O.; EL-DEIR, A.C.A. *et al.* (2018). “*Artisanal fisheries of the Itaparica reservoir, São Francisco River, Brazil: socioeconomic profile, environmental dynamics, and management recommendations*”. *Regional Environmental Change* 18, 1889–1899. <https://doi.org/10.1007/s10113-018-1293-y>

SILVA, G. M. N. (2019). Avaliação do aporte de nutrientes proveniente da piscicultura na alteração da qualidade da água diante de cenários de redução de vazão em reservatório no semiárido. Tese (Doutorado) - Universidade Federal de Pernambuco. Programa de Pós-Graduação em Engenharia Civil, 2019.

TUNDISI, J. G.; MATSUMURA-TUNDISI, T. (2008). *Limnologia*. Oficina de Textos São Paulo - SP, 631p.

VON SPERLING, M. (1996). *Introdução à qualidade das águas e ao tratamento de esgotos*. DESA-UFMG Belo Horizonte- MG, 243 p.