

APLICAÇÃO DE MÉTODOS ESTATÍSTICOS ROBUSTOS NA AVALIAÇÃO DA QUALIDADE DA ÁGUA DE VERTEDORES FLUVIOMÉTRICOS

Pedro Fialho Cordeiro^{1*}
Samuel Rodrigues Castro²
Feliciano José Baeta Costa¹
Marcus Manoel Fernandes¹
Marina Andrada Maria¹

1 Centro de Inovação e Tecnologia SENAI FIEMG

2 Universidade Federal de Juiz de Fora - UFJF

* pcordeiro@fiemg.com.br

RESUMO

Os dados provenientes de resultados de amostras ambientais podem apresentar imprecisões relacionadas à amostragem, preparo e análise, bem como ser censurados pelo limite de detecção do método analítico. Devido a estas características, o emprego de métodos estatísticos que não dependem da distribuição dos dados, como os métodos robustos, devem ser utilizados em avaliações ambientais. O presente trabalho apresenta exemplos de aplicação de estatísticas multivariadas robustas, com fins de evidenciar a aplicabilidade e fomentar a difusão da técnica, para o tratamento de dados de qualidade e quantidade de água de 3 vertedores fluviométricos, construídos em área de mineração, no estado de Minas Gerais. O método de tratamento de dados empregado nesse estudo apresentou-se eficiente ao identificar fortes correlações entre variáveis físico-químicas, como ferro dissolvido, manganês total, sólidos dissolvidos totais e condutividade elétrica. Apesar disso, não foi possível verificar a influência de diferentes níveis de descarga líquida no comportamento da qualidade da água dos vertedores fluviométricos, sendo recomendável realizar novas campanhas de monitoramento para tornar o banco de dados mais consistente e mais sensível à influência de diferentes níveis de descarga líquida.

INTRODUÇÃO

Os vertedores fluviométricos são instrumentos hidráulicos utilizados para medir a vazão em cursos de água naturais e em canais. Estes equipamentos são construídos, principalmente, para avaliar o comportamento hidrológico de cursos de água de pequena dimensão (INIGUEZ-COVARRUBIAS et al., 2015). Nos empreendimentos minerários em que a água é um insumo fundamental para atividades como o processo de beneficiamento, operações de lavra, consumo humano e aspersão de vias, o uso de vertedores é imprescindível para monitorar quantitativamente as águas superficiais.

Por meio da aplicação de equações específicas, é possível converter as leituras das cotas dos vertedores em vazões escoadas. Dessa forma, os dados de vazão podem ser obtidos para cada cota registrada (CHOW, 1959). Postos fluviométricos que têm o monitoramento da qualidade da água concomitante à leitura dos seus níveis, permitem verificar a influência de diferentes níveis de vazão no comportamento de variáveis de qualidade de naturezas físico-químicas e biológicas.

O monitoramento ambiental, realizado por amostragens sistemáticas, procura caracterizar aspectos relevantes que permitam diagnosticar mudanças, como no uso e ocupação do solo, possibilitando avaliar o efeito das atividades humanas exercidas sobre os recursos hídricos. Os dados provenientes da análise de amostras ambientais apresentam variáveis espaciais e temporais como, por exemplo, a localização e a data de amostragem. Assim, pode-se afirmar que ocorre dependência entre as amostras, pois os pontos próximos têm mais chance de apresentarem resultados semelhantes que pontos geograficamente distantes (KUPPUSAMY e GIRIDHAR, 2006). Além disso, esses dados podem apresentar imprecisões relacionadas à amostragem, preparo e análise, bem como ser censurados pelo limite de detecção do método analítico.

Estas condições podem impactar a interpretação de análises estatísticas convencionais, ocasionando distorções e avaliações inconsistentes. Reimann et al. (2008) sugerem o emprego de métodos estatísticos que não dependem da distribuição dos dados, como os métodos robustos, os quais devem ser utilizados em avaliações ambientais. O uso de técnicas robustas é uma metodologia consolidada e difundida internacionalmente; porém, no Brasil, prevalece o uso de técnicas convencionais (SABINO et al., 2014). Dessa forma, o presente estudo tem como objetivo aplicar estatísticas multivariadas robustas, com fins de evidenciar a aplicabilidade e fomentar a difusão da técnica, para o tratamento de dados de qualidade e quantidade de água de 3 vertedores fluviométricos construídos em área de mineração, em Minas Gerais, entre janeiro de 2015 e dezembro de 2017.

MATERIAL E MÉTODOS

Para a avaliação dos resultados, foram considerados os limites e as condições estabelecidas para águas classe 2, de acordo com a Deliberação Normativa Conjunta COPAM/CERH-MG nº 01, de 05 de maio de 2008.

Os dados de qualidade e quantidade da água foram obtidos a partir de amostragens realizadas em 3 vertedores instalados, em área de mineração, no estado de Minas Gerais. Os instrumentos foram construídos em cursos de água de pequenas dimensões, com o objetivo de monitorar as descargas líquidas, com frequência diária, e a qualidade da água, com frequência trimestral, no período de janeiro de 2015 a dezembro de 2017. Ao todo, tem-se 39 coletas realizadas, as quais contemplaram a análise de 23 parâmetros físico-químicos e biológicos, a saber: alumínio total, cádmio total, chumbo total, cloreto, cobre dissolvido, coliformes termotolerantes, condutividade elétrica, cor verdadeira, DBO₅, ferro dissolvido, ferro total, fosfato, manganês total, mercúrio total, nitrato (N), nitrogênio amoniacal, óleos e graxas, ortofosfato, oxigênio dissolvido, pH in situ, sólidos dissolvidos totais, sólidos suspensos totais e turbidez.

A seção transversal dos vertedores supracitados é do tipo trapezoidal, com dimensões da base do trapézio de 40 cm, para os vertedores 1 e 2, e 50 cm, para o vertedor 3. Estes equipamentos têm bacias de contribuição com áreas que variam de 0,8 a 3,5 km². Por meio da aplicação de equações específicas, foram estimadas as suas vazões diárias.

Uma análise de consistência foi empregada sobre as variáveis analisadas, com o objetivo de remover do conjunto de dados *outliers*, parâmetros com mais de 50% de valores não detectados ou abaixo do limite de detecção do método analítico, bem como falhas na série histórica. A análise de consistência foi realizada no *software* ProUCL 5.0.0 (**Tabela 01**).

Para verificar a existência de diferenças significativas nos resultados obtidos das análises dos 3 vertedores, para todos os parâmetros selecionados pela análise de consistência, foi realizado o teste não paramétrico de comparações múltiplas Kruskal-Wallis, no intuito de evidenciar diferenças significativas entre os resultados das análises para os diferentes vertedouros, a 95% de confiança. Para os casos com evidência de diferenças significativas, foram gerados gráficos do tipo box-plot.

Em seguida, procedeu-se com a geração de uma matriz de correlação de Spearman, para verificar a autocorrelação entre todos os parâmetros, contemplando os vertedores individualmente e em conjunto. O objetivo dessa análise foi remover pares de variáveis altamente correlacionadas. Dormann et al. (2012) sugerem o limiar de corte de $r > \pm 0,70$. Nesse sentido, realizou-se o expurgo de um dos parâmetros com evidências significativas de autocorrelação.

Um conjunto reduzido de variáveis não correlacionadas entre si foi selecionado para a realização de análise de componentes principais (PCA), visando aumentar o número de graus de liberdade. Os valores numéricos das componentes principais, denominados *scores*, foram calculados para cada elemento amostral e a sua distribuição espacial possibilitou identificar diferenças no comportamento das variáveis selecionadas nesse estudo.

RESULTADOS E DISCUSSÃO

Mais da metade dos parâmetros selecionados apresentaram elevado percentual de valores não detectados (NDs), bem como falhas em suas séries históricas. A análise estatística descritiva do conjunto de dados dos três vertedores é apresentada na **Tabela 01**.

Tabela 01 – Estatística descritiva do conjunto completo de dados dos 3 vertedores

Parâmetro	Obs [1]	Falhas	Num Ds [2]	Num NDs [3]	% NDs [4]	Mínimo	Máximo	Média	Mediana	CV [5]
Alumínio Dissolvido (mg/L)	39	0	6	33	84,62%	0,0495	0,11	0,0699	0,055	0,396
Cádmio Total (mg/L)	39	0	1	38	97,44%	0,002	0,002	0,002	0,002	N/A
Chumbo Total (mg/L)	39	0	0	39	100,00%	N/A	N/A	N/A	N/A	N/A
Cloreto (mg/L)	24	15	6	18	75,00%	1,2	1,86	1,403	1,29	0,175
Cobre Dissolvido (mg/L)	39	0	1	38	97,44%	0,01	0,01	0,01	0,01	N/A
Coliformes Termotolerantes (UFC/100mL)	39	0	39	0	0,00%	3	3300	171,7	55	3,051
Condutividade Elétrica in Situ ($\mu\text{S}/\text{cm}$)	39	0	39	0	0,00%	7,74	110,3	18,41	15,1	0,867
Cor Verdadeira	38	1	19	19	50,00%	6	301	46,96	21	1,468
Demanda Bioquímica de Oxigênio (mg/L)	39	0	1	38	97,44%	3,73	3,73	3,73	3,73	N/A
Ferro Dissolvido (mg/L)	39	0	39	0	0,00%	0,1	1,36	0,403	0,32	0,726
Ferro Total (mg/L)	39	0	39	0	0,00%	0,12	4,89	1,081	0,71	0,9
Fosfato	3	36	0	3	100,00%	N/A	N/A	N/A	N/A	N/A
Manganês Total (mg/L)	39	0	28	11	28,21%	0,0268	0,37	0,0923	0,065	0,86
Mercurio Total (mg/L)	39	0	1	38	97,44%	3,00E-04	3,00E-04	3,00E-04	3,00E-04	N/A
Nitrato (N) (mg/L)	24	15	15	9	37,50%	0,03	0,3	0,105	0,09	0,681
Nitrogênio Amoniacal (mg/L)	39	0	9	30	76,92%	0,07	0,18	0,124	0,13	0,329
Oleos e Graxas Totais (mg/L)	39	0	0	39	100,00%	N/A	N/A	N/A	N/A	N/A
Ortofosfato (mg/L)	33	3	1	32	96,97%	0,06	0,06	0,06	0,06	N/A
Oxigênio Dissolvido (mg/L)	39	0	39	0	0,00%	5,15	8,32	6,991	7,09	0,126
pH in Situ	39	0	39	0	0,00%	5,668	7,89	6,645	6,63	0,0835
Sólidos Dissolvidos Totais (mg/L)	39	0	28	11	28,21%	8	36	15,77	14,5	0,386
Sólidos Suspensos Totais (mg/L)	39	0	7	32	82,05%	2	17	8,071	6	0,679
Turbidez (NTU)	39	0	39	0	0,00%	0,68	49	6,807	3,58	1,414
Vazão (Ls^{-1})	38	1	38	0	0,00%	1,1	217,2	20,34	10,77	1,744

[1] número de observações

[2] número de valores detectados (acima do limite de detecção do método analítico)

[3] número de valores não detectados (abaixo do limite de detecção do método analítico)

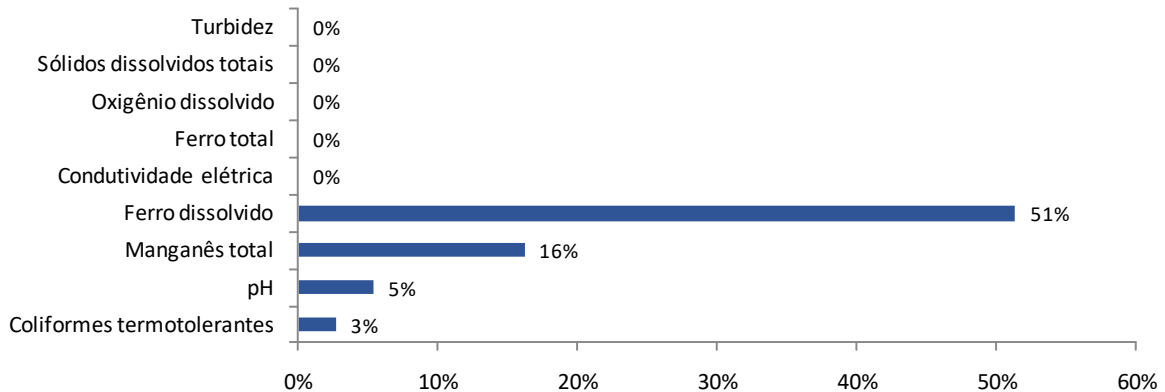
[4] percentual de valores não detectados

[5] coeficiente de variação

Constituídos por 39 observações, os parâmetros alumínio dissolvido, cádmio total, chumbo total, cloreto, cobre dissolvido, cor verdadeira, DBO, fosfato, mercúrio total, nitrato, nitrogênio amoniacal, óleos e graxas, ortofosfato e sólidos suspensos totais apresentaram mais de 50% de NDs e falhas na série histórica de dados, sendo excluídos das análises posteriores para não comprometerem o tratamento estatístico.

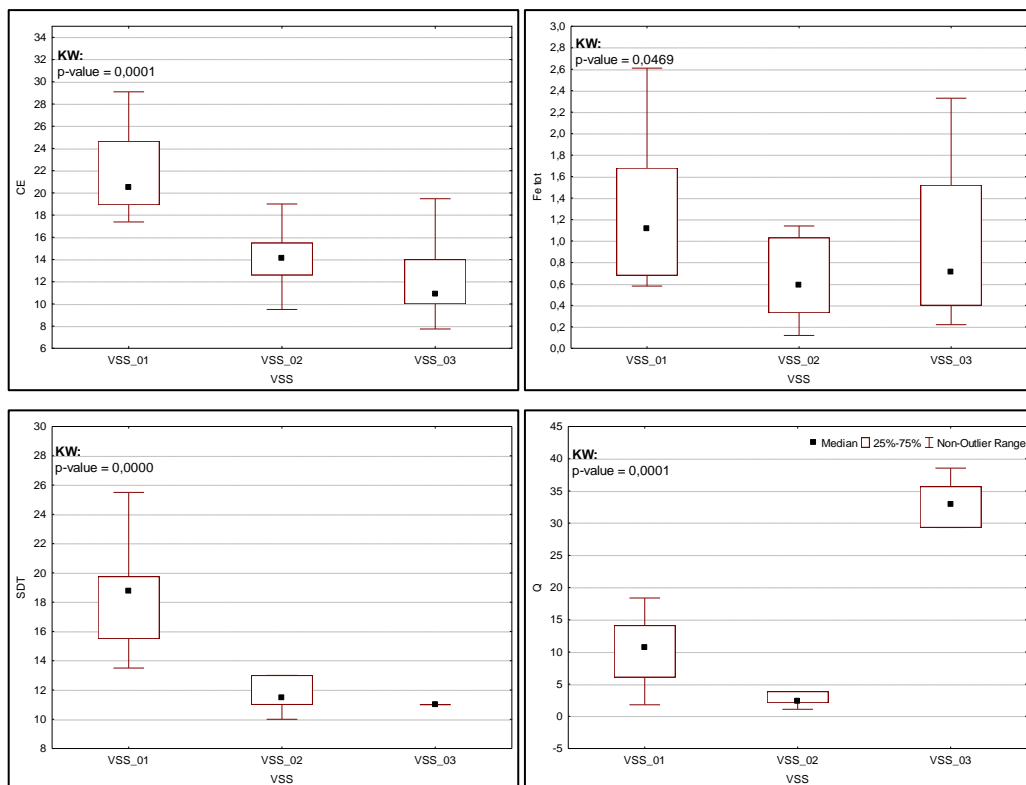
Quatro das 10 variáveis selecionadas apresentaram pelo menos uma não conformidade em relação à DN COPAM/CERH - MG nº 01 / 2008. Os parâmetros ferro dissolvido e manganês total foram os que violaram mais vezes o enquadramento para classe 2, com inconformidades em 51 e 16% dos casos, respectivamente; fato que pode estar relacionado à geoquímica típica da região mineira objeto de estudo (**Figura 01**).

Figura 01 – Variáveis com percentual de não atendimento à legislação DN COPAM / CERH - MG nº 01/ 2008



O teste de comparações múltiplas Kruskal-Wallis indicou haver diferenças significativas entre os resultados das análises para os diferentes vertedores, a 95% de confiança, para os parâmetros condutividade elétrica, ferro total, sólidos dissolvidos totais e vazão (**Figura 02**).

Figura 02 – Teste de comparações múltiplas Kruskal-Wallis



Observa-se que o vertedor 3 possui os menores valores de condutividade elétrica e de concentração de sólidos dissolvidos totais. Por outro lado, esse vertedor apresenta as maiores descargas líquidas entre os 3 instrumentos avaliados, apesar de possuir apenas a segunda maior bacia de contribuição. Esta condição possivelmente está relacionada às contribuições diretas do aquífero de formação ferrífera da área de estudo, que corrobora para maiores vazões específicas, principalmente durante o período de estiagem. Ressalta-se que os cursos de água que drenam sobre formações ferríferas no estado de Minas Gerais são, de modo geral, alimentados por aquíferos em suas vazões escoadas na superfície (GUERRA, 2010).

A correlação de Spearman das variáveis selecionadas, com destaque para as correlações significativas ($r > \pm 0,70$; $p < 0,05$), destacadas em vermelho, é apresentada na **Tabela 02**.

Tabela 02 - Matriz de correlação de Spearman

	Colif_term	CE_in_situ	Fe_dissolvido	Fe_total	Mn_total	OD	pH	SDT	Turbidez	Vazão
Colif	--	-0,070596	-0,217401	0,068031	0,088995	0,365676	0,023206	0,080818	0,057085	0,244974
CE	-0,070596	--	0,547735	0,506566	0,511608	-0,204946	-0,073649	0,705771	0,352812	-0,066046
Fe dissol	-0,217401	0,547735	--	0,849521	0,699961	-0,463822	0,101123	0,539036	0,723035	0,142442
Fe tot	0,068031	0,506566	0,849521	--	0,763164	-0,187459	-0,089471	0,551555	0,854712	0,277906
Mn tot	0,088995	0,511608	0,699961	0,763164	--	-0,037099	0,006054	0,515218	0,804488	-0,146651
OD	0,365676	-0,204946	-0,463822	-0,187459	-0,037099	--	-0,131210	-0,101743	-0,134931	-0,162870
pH	0,023206	-0,073649	0,101123	-0,089471	0,006054	-0,131210	--	-0,043783	-0,064244	-0,135995
SDT	0,080818	0,705771	0,539036	0,551555	0,515218	-0,101743	-0,043783	--	0,526116	-0,051088
Turb	0,057085	0,352812	0,723035	0,854712	0,804488	-0,134931	-0,064244	0,526116	--	0,109322

Correlações significativas ($r \geq \pm 0,70$) indicadas em vermelho

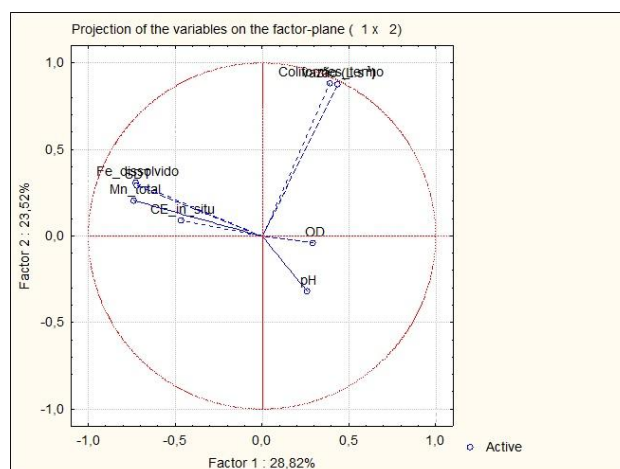
A avaliação comparativa dos parâmetros elencados indicou a presença de variáveis altamente correlacionadas entre si, com destaque para o ferro total e a turbidez, com valor superior de r igual a 0,85. Ambas apresentaram, ainda, outras duas correlações fortes, com $r > 0,70$, e foram excluídas das análises posteriores (DORMANN et al., 2012).

Ao final das exclusões, restaram as variáveis: coliformes termotolerantes, condutividade elétrica, ferro dissolvido, manganês total, oxigênio dissolvido, pH, sólidos dissolvidos totais e vazão, que foram levadas à análise de PCA. Os *scores* foram calculados para cada elemento amostral e a sua distribuição espacial possibilitou identificar diferenças no comportamento das variáveis selecionadas nesse estudo (Tabela 03 e Figura 03).

Tabela 03 – Matriz de cargas das variáveis em cada componente principal

	Fator 1	Fator 2	Fator 3	Fator 4
Coliformes_termo	0,390583	0,883715	0,004996	-0,081302
CE_in_situ	-0,464187	0,090608	-0,461719	-0,660948
Fe_dissolvido	-0,727471	0,308082	0,484956	0,110088
Mn_total	-0,738800	0,206826	0,027300	0,471832
OD	0,294585	-0,038636	-0,770911	0,296026
pH	0,258958	-0,320183	0,486881	-0,454491
SDT	-0,722706	0,290556	-0,292723	-0,288846
Vazão	0,431936	0,875361	0,106843	-0,057034

Figura 03 – Gráfico das componentes principais CP1 e CP2



A componente 1, explicando cerca de 28% da variabilidade dos dados, sugere uma relação entre o ferro dissolvido, o manganês total, sólidos dissolvidos totais e condutividade elétrica. Esta condição é esperada uma vez que a presença de íons na água influencia diretamente os valores de condutividade elétrica e de sólidos dissolvidos totais. Segundo Sabino et al. (2014), resultados para parâmetros como ferro e manganês (cargas negativas) podem estar relacionados com a presença natural desses elementos nos solos da região, podendo ser intensificado por atividades minerárias.

A componente 2, explicando cerca de 23% da variabilidade dos dados, indica relação entre a vazão e os coliformes termotolerantes, sugerindo influência das atividades agrosilvopastoris nos cursos de água, principalmente no período chuvoso, quando ocorrem os maiores deflúvios.

Os parâmetros OD e pH possivelmente não estão relacionados com nenhuma das demais variáveis selecionadas.

Embora seja possível identificar vazões de referência para as estações de seca e chuva, não foi possível verificar a influência de diferentes níveis de descarga líquida no comportamento da qualidade da água nos vertedores fluviométricos.

CONSIDERAÇÕES FINAIS

Os vertedores fluviométricos são equipamentos imprescindíveis para monitorar a qualidade da água e o comportamento hidrológico de cursos de água em áreas de influência de empreendimentos minerários.

Os métodos de tratamento de dados empregados neste estudo, como a estatística descritiva, a avaliação de consistência de série histórica e de valores não detectados, além da verificação de *outliers* e exclusão de variáveis altamente correlacionadas entre si, apresentou-se eficiente para que fossem empregadas as melhores ferramentas estatísticas. Em muitos casos, esse tratamento dos dados é ignorado por pesquisadores em diversos estudos e pode impactar negativamente toda a análise dos resultados.

A inclusão de novas campanhas de monitoramento poderá tornar o banco de dados mais consistente e mais sensível à influência de diferentes níveis de descarga líquida no comportamento de variáveis de qualidade da água.

REFERÊNCIAS BIBLIOGRÁFICAS

- CHOW, V. T. Open Channel Hydraulics. New York: McGraw-Hill. 1959. 680 p.
- DORMANN, C.; F. ELITH, J.; BACHER, S.; BUCHMANN, C.; GUDRUN, C., G.; GARCÍA MARQUÉZ, J., R.; GRUBER, B.; LAFOURCADE, B.; LEITÃO, P., J.; MÜNKEMÜLLER, T.; MCCLEAN, C.; OSBORNE, P., E.; BJÖRN, R.; SCHRÖDER, B.; SKIDMORE, A., K; ZURELL, D.; LAUTENBACH, S. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, v. 35, p. 001–020, 2012.
- GUERRA, J., K. Aplicação de método geofísico em estudo hidrogeológico no município de Nova Lima – MG. Dissertação de mestrado, Universidade Federal de Minas Gerais – UFMG. Belo Horizonte – MG. 2010.
- INIGUEZ-COVARRUBIAS, M.; OJEDA-BUSTAMANTE, W.; DIAZ-DELGADO, C. Mejoras de eficiencia hidráulica en vertedores con canal de descarga libre en presas: propuesta metodológica. *Tecnol. cienc. agua, Jiutepec*, v. 6, n. 1, p. 69-79, 2015.
- KUPPUSAMY, M.R.; GIRIDHAR, V.V. Factor analysis of water quality characteristics including trace metal speciation in the coastal environmental system of Chennai Ennore. *Environmental International*, v.31, n. 2, p. 174-179, 2006.
- REIMANN, C.; FILZMOSER, P.; GARRETT R.; DUTTER R. Statistical data analysis explained. *Applied environmental statistics with R*. 1 ed. Chichester. Ed. John Wiley & Sons, 2008. 362 p.
- SABINO, V., S., C.; LAGE, L., A.; C., B., A, K. Uso de métodos estatísticos robustos na análise ambiental. *Engenharia Sanitária e Ambiental*, v. 19, p. 87-94, 2014.