



ANÁLISE DA TURBIDEZ EM ESTAÇÃO DE TRATAMENTO DE ÁGUA: INFLUÊNCIA DE *OUTLIERS* EM DADOS OPERACIONAIS

Thais Mayara Rodrigues Gomes¹, Elton Santos Franco¹, Rafael Genaro¹, Jairo Lisboa Rodrigues¹, Núbia Aparecida de Aguiar¹

¹Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM), Teófilo Otoni, Brasil (thais.mayara@ufvjm.edu.br)

Resumo: Este trabalho analisa dados de turbidez em estação de tratamento de água, comparando o desempenho do sulfato de alumínio e do policloreto de alumínio. Foi realizada a análise estatística descritiva com e sem a presença de *outliers*. Os resultados indicaram que valores extremos influenciam significativamente a média e a variabilidade, especialmente para o sulfato de alumínio, enquanto o policloreto de alumínio apresentou comportamento mais estável.

Palavras-chave: Policloreto de Alumínio; Sulfato de Alumínio; Turbidez, Estação de Tratamento de Água, Estatística.

INTRODUÇÃO

O tratamento de água para abastecimento público tem como principal finalidade garantir a remoção de impurezas e a adequação da qualidade da água aos padrões de potabilidade estabelecidos pela legislação vigente.

Dentre os parâmetros de controle operacional, a turbidez destaca-se por estar diretamente relacionada à presença de partículas em suspensão, sendo amplamente utilizada como indicador da qualidade da água (Kumar; Balasundaram, 2017).

Além de comprometer o aspecto estético, valores elevados de turbidez podem interferir na eficiência dos processos de tratamento e na qualidade final da água distribuída, uma vez que as partículas suspensas podem conter microrganismos e impurezas que impactam na saúde pública (Libânio, 2016).

A remoção da turbidez está associada, principalmente, às etapas de coagulação, floculação e decantação, sendo a coagulação um processo fundamental para a desestabilização de partículas coloidais e formação de flocos passíveis de remoção (Pandey; Khan, 2022).

Tradicionalmente, o sulfato de alumínio (SA) é amplamente utilizado como coagulante em Estações de Tratamento de Água (ETA). Entretanto, o policloreto de alumínio (PAC) tem sido adotado como alternativa, em função de sua maior eficiência em determinadas condições operacionais, além de favorecer a formação de flocos mais densos e com melhor capacidade de sedimentação (Constantino; Yamamura, 2009).

Estudos indicam que o PAC apresenta melhor desempenho em uma faixa mais ampla de pH, além de

possibilitar menor geração de lodo, contribuindo para maior estabilidade operacional do processo de tratamento (Pavanelli, 2001; Rosalino, 2011).

No contexto das ETA, a análise de dados operacionais de turbidez constitui uma ferramenta importante para avaliação do desempenho dos processos. No entanto, esses dados podem apresentar valores discrepantes (*outliers*), decorrentes de falhas de medição, variações nas condições de operação ou eventos atípicos, os quais podem influenciar, conforme Medri (2011), significativamente os resultados das análises estatísticas.

A presença desses valores pode comprometer a representatividade de medidas como média e desvio padrão, conduzindo a interpretações inadequadas do comportamento do sistema.

Dessa forma, a identificação e o tratamento de *outliers* tornam-se relevantes para garantir maior confiabilidade na análise de dados operacionais, especialmente quando se utilizam métodos estatísticos descritivos.

Nesse contexto, o presente trabalho tem como objetivo analisar o comportamento dos dados de turbidez associados aos coagulantes SA e PAC, com ênfase na avaliação da influência de *outliers* sobre as medidas estatísticas descritivas, por meio da comparação entre conjuntos de dados com e sem a presença desses valores.

MATERIAIS E MÉTODOS

Este estudo caracteriza-se como uma pesquisa aplicada, de abordagem quantitativa, baseada na análise de dados operacionais de turbidez. Os dados foram obtidos junto à ETA Seminário, no município



de Mariana/MG, sendo fornecidos pelo Serviço Autônomo de Água e Esgoto (SAAE).

Os dados abrangem o período de 2018 a 2022 e consistiam em planilhas de controle da estação, nas quais estavam registradas manualmente informações como data, horário, operador, vazão de entrada da água, quantidade de materiais usados no tratamento, turbidez, cor e pH da água bruta e tratada, em intervalos de duas horas ao longo de vinte e quatro horas.

As amostras analisadas correspondem à turbidez da água após o tratamento com os coagulantes SA e PAC. Os dados correspondentes ao tratamento com SA abrangem o período de 1 janeiro de 2018 a 21 de julho de 2020, enquanto os dados referentes ao tratamento com PAC correspondem ao período de 22 de julho de 2020 a dezembro de 2022.

Uma vez que os dados foram registrados manualmente, foi necessário digitalizá-los. Esse processo foi realizado utilizando o *website nanonets.com*, que permitiu a conversão dos dados escritos em tabelas para o formato *Excel*.

Os dados de turbidez tratada correspondentes a cada coagulante foram separados e organizados em planilhas, de acordo com a ordem cronológica de monitoramento, de maneira a facilitar seu uso e consulta. Em seguida, foi realizada uma revisão minuciosa de cada dado para corrigir eventuais erros decorrentes do processo de digitalização dos dados pela inteligência artificial (IA) do *website* utilizado.

A estatística descritiva dos dados foi realizada por meio da ferramenta “Análise de Dados” do *software Excel*, através da qual foram calculados para cada amostra (turbidez tratada com o PAC e turbidez tratada com o SA) o total de dados, a média, a mediana, a moda, o primeiro quartil (Q1), o terceiro quartil (Q3), o intervalo interquartil (IQR), valores mínimo e máximo observados, o percentil 95 (P95), a assimetria, o desvio-padrão e a variância.

Foi produzida uma tabela para apresentação desses valores, de forma a proporcionar uma melhor visualização dos resultados. Nela, constam os resultados obtidos tanto para as amostras completas quanto para as amostras com exclusão dos *outliers*.

Em seguida, procedeu-se à identificação de valores discrepantes (*outliers*), com base em critérios estatísticos, visando avaliar sua influência sobre a distribuição dos dados. Para isso, foram conduzidas análises comparativas considerando dois cenários: o conjunto completo de dados e o conjunto após a remoção dos *outliers*.

Os *outliers* foram detectados por meio do teste de Grubbs e Beck, executado no *Excel*. A robustez das inferências estatísticas realizadas neste estudo é

respaldada pelo tamanho expressivo das amostras, que, mesmo após a exclusão de registros, supera em muito o mínimo necessário para garantir precisão e confiabilidade.

RESULTADOS E DISCUSSÃO

A análise estatística descritiva dos dados de turbidez (Tabela 1) evidenciou diferenças no comportamento das amostras associadas aos coagulantes SA e PAC, bem como a presença de valores extremos que influenciam a distribuição dos dados.

Tabela 1. Resultados da Estatística Descritiva dos Dados de Turbidez Pós-tratamento com o Coagulante SA e com o PAC.

Medidas Descritivas	Resultados			
	Amostra Completa		Amostra Sem Outliers	
	SA	PAC	SA	PAC
Total de Dados	10.903	10.404	10.403	10.403
Média (uT)	0,75	0,29	0,63	0,28
Mediana (uT)	0,55	0,24	0,54	0,24
Moda (uT)	0,48	0,20	0,48	0,20
Q1 (uT)	0,42	0,19	0,42	0,19
Q3 (uT)	0,80	0,32	0,75	0,32
IQR (uT)	0,38	0,13	0,33	0,13
Mínimo (uT)	0,06	0,05	0,06	0,05
Máximo (uT)	73,00	15,10	1,69	13,90
P95 (uT)	1,66	0,49	1,40	0,49
Assimetria	41,54	25,21	1,40	23,09
Desvio-padrão (uT)	1,29	0,33	0,32	0,30
Variância (uT ²)	1,67	0,11	0,10	0,10

Fonte: Autoria própria.

Observa-se que o número total de dados analisados referentes ao parâmetro turbidez tratada foi de 10.903 para o SA e 10.404 para o PAC. Para a amostra de turbidez da água tratada com PAC, foram retirados *outliers* com valores superiores a 13,90 uT, tendo sido excluído apenas 1 dado, totalizando 10.403 dados após a remoção de *outliers*. Já para a amostra tratada com SA, foram excluídos os dados com valores acima de 1,69 uT, o que corresponde à remoção de 500 observações e totaliza 10.403 dados.

Considerando o objetivo deste estudo, a discussão foi direcionada às medidas estatísticas que apresentaram maior sensibilidade à presença de *outliers*, permitindo



avaliar de forma mais consistente seus impactos na interpretação dos dados.

Para a amostra completa, observou-se que o tratamento com SA apresentou média de 0,75 uT, enquanto o PAC apresentou média significativamente inferior, de 0,29 uT. No entanto, a presença de valores máximos elevados, especialmente para o SA (73,00 uT) e para o PAC (15,10 uT), indica a ocorrência de *outliers* que impactam diretamente as medidas de tendência central e dispersão.

Esses valores extremos podem estar associados a eventos operacionais específicos, como variações bruscas na qualidade da água bruta, falhas pontuais no processo de dosagem de coagulantes ou interferências no sistema de medição, evidenciando a natureza dinâmica do processo de tratamento de água.

Após a remoção dos *outliers*, verificou-se redução nas médias para ambos os coagulantes, passando para 0,63 uT no caso do SA e 0,28 uT para o PAC. Essa redução foi mais expressiva para o SA, evidenciando maior influência dos valores extremos nesse conjunto de dados.

Esse comportamento indica que a média, por ser uma medida sensível a valores discrepantes, pode superestimar o nível de turbidez quando *outliers* estão presentes, especialmente em conjuntos de dados com maior variabilidade.

A mediana apresentou menor variação entre os cenários analisados, mantendo-se próxima de 0,55 uT (amostra completa) e 0,54 uT (sem *outliers*) para o SA, e estável em 0,24 uT para o PAC. Esse comportamento confirma a menor sensibilidade dessa medida à presença de *outliers*.

O valor encontrado para o primeiro quartil (Q1) foi de 0,42 uT para a amostra de turbidez após o tratamento com o SA e de 0,19 uT com o PAC, o que indica que 25% dos dados se localizam abaixo desses valores para cada coagulante.

O valor do terceiro quartil (Q3) para o tratamento com SA foi de 0,80 uT para a amostra completa e 0,75 uT para a amostra sem *outliers*, enquanto para o PAC foi de 0,32 uT, estes valores representam o limite abaixo do qual se encontram 75% dos dados para cada tratamento. Esses resultados revelam que a dispersão dos dados é menor para a amostra do tratamento com PAC.

Conforme apresentado na Tabela 1, observou-se redução significativa do desvio padrão após a remoção dos *outliers*, especialmente para o SA, que passou de 1,29 uT para 0,32 uT. Para o PAC, essa redução foi menos acentuada, passando de 0,33 uT para 0,30 uT. Esses resultados indicam que os valores extremos influenciam diretamente a variabilidade dos dados, sobretudo no caso do SA.

Os valores de assimetria também evidenciam a influência dos *outliers*, com destaque para a amostra completa do SA (41,54) e do PAC (25,21), reduzindo-se após sua remoção para 1,40 e 23,09, respectivamente. Esses resultados indicam que a presença de valores discrepantes contribui para distorções significativas na distribuição dos dados, especialmente no caso do SA.

A comparação entre os cenários demonstrou que a consideração dos *outliers* pode levar a interpretações distintas sobre o comportamento da turbidez, sobretudo quando se utilizam medidas sensíveis, como a média e o desvio padrão.

De modo geral, os resultados indicam que o conjunto de dados referente ao SA apresenta maior variabilidade e maior influência de *outliers* quando comparado ao PAC, enquanto o PAC apresenta comportamento mais estável. Nesse contexto, destaca-se a importância da identificação e análise criteriosa desses valores, uma vez que sua presença pode comprometer a interpretação dos dados e a avaliação do desempenho dos processos de tratamento.

CONCLUSÃO

Os resultados obtidos evidenciam que a presença de *outliers* influencia significativamente a análise estatística dos dados de turbidez, especialmente em relação às medidas de média e dispersão. Observou-se que o conjunto de dados associado ao SA apresentou maior sensibilidade aos valores extremos, enquanto o PAC demonstrou comportamento mais estável.

A remoção dos *outliers* contribuiu para uma representação mais consistente dos dados, reduzindo a variabilidade e permitindo melhor interpretação do comportamento da turbidez. Nesse contexto, destaca-se a importância da identificação e avaliação criteriosa desses valores em análises de dados operacionais, de modo a evitar interpretações distorcidas e apoiar decisões mais confiáveis no tratamento de água.

Os resultados também indicam que o tipo de coagulante pode influenciar não apenas a eficiência do tratamento, mas também a estabilidade dos dados operacionais, sendo o PAC associado a menor variabilidade e menor presença de valores discrepantes.

Adicionalmente, reforça-se a relevância da aplicação de ferramentas estatísticas na análise de sistemas reais, contribuindo para o aprimoramento do controle operacional em estações de tratamento de água.

REFERÊNCIAS

CONSTANTINO, A. F.; YAMAMURA, V. D. Redução do gasto operacional em estação de tratamento de água utilizando o PAC. In: SIMPÓSIO DE PÓS-GRADUAÇÃO EM ENGENHARIA



URBANA, 2., 2009, Maringá. **Anais [...]**. Maringá: UEM, 2009.

KUMAR, N.; BALASUNDARAM, N. Efficiency of PAC in water treatment plant & disposal of its sludge. **International Journal of Applied Engineering Research**, v. 12, n. 12, p. 3253-3262, 2017.

LIBÂNIO, M. **Fundamentos de qualidade e tratamento de água**. 4. ed. Campinas: Átomo, 2016.

MEDRI, W. **Análise exploratória de dados**. Londrina: Universidade Estadual de Londrina, 2011. Disponível em: https://www.uel.br/pos/estatisticaquantitativa/textos_didaticos/especializacao_estatistica.pdf. Acesso em: 25 jan. 2025.

PANDEY, P.; KHAN, F. Moringa oleifera Plant as potent alternate to Chemical Coagulant in Water Purification. **Brazilian Journal of Pharmaceutical Sciences**, v. 58, 2022.

PAVANELLI, G. **Eficiência de diferentes tipos de coagulantes na coagulação, floculação e sedimentação de água com cor ou turbidez elevada**. 2001. Dissertação (Mestrado em Hidráulica e Saneamento) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2001.

ROSALINO, M. R. R. **Potenciais Efeitos da Presença de Alumínio na Água de Consumo Humano**. 2011. Dissertação (Mestrado em Engenharia do Ambiente) – Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Lisboa, 2011.