

874 – TRATAMENTO E AVALIAÇÃO ESTATÍSTICA DE VARIÁVEIS DE CONTROLE DO PROCESSO DE LODOS ATIVADOS: ESTUDO DE CASO COM DADOS DE MONITORAMENTO PADRÃO DE ESTAÇÃO DE TRATAMENTO DE EFLUENTES DE GRANDE PORTE

Antônio Carlos de Oliveira Martins Júnior⁽¹⁾

Engenheiro Civil e Ambiental pela Universidade Vale do Rio Doce. Mestrando em Recursos Hídricos e Saneamento Ambiental pelo Instituto de Pesquisas Hidráulicas (IPH) da Universidade Federal do Rio Grande do Sul (UFRGS).

Maria Cristina de Almeida Silva⁽²⁾

Engenheira de Bioprocessos e Biotecnologia pela Universidade Estadual do Rio Grande do Sul. Mestre e Doutora em Recursos Hídricos e Saneamento Ambiental pelo IPH/UFRGS. Professora no IPH/UFRGS.

Antônio Domingues Benetti⁽³⁾

Engenheiro Civil pela UFRGS. Mestre em Engenharia de Recursos Hídricos e Saneamento Ambiental pelo IPH/UFRGS. Doutorado em Civil and Environmental Engineering pela Cornell University. Professor no IPH/UFRGS.

Endereço: Avenida Bento Gonçalves, 9500 – Prédio 44302, Agronomia, Porto Alegre, RS. CEP: 91501-970. e-mail: antoniocomjr@gmail.com⁽¹⁾

RESUMO

Este trabalho avaliou a distribuição e a tendência central de variáveis de controle do processo de lodos ativados utilizando dados de monitoramento padrão de Estação de Tratamento de Esgoto (ETE) de grande porte. Foi analisado um banco com dados de monitoramento de 2018 e 2019 de uma ETE em Porto Alegre/RS. Foram avaliadas as variáveis DBO, DQO, N-NH₃, SST, temperatura, pH e vazão, considerando registros ausentes, censurados e *outliers*. A exclusão dos *outliers* foi condicionada a uma avaliação conjunta das variáveis. Foram calculadas medidas descritivas dos registros e a simetria das amostras de DBO, DQO, N-NH₃ e SST foi indicada pelo coeficiente de distorção (CD). Amostras que se destacaram em relação à simetria foram representadas pela distribuição de frequência de registros. As amostras de registros dos períodos de monitoramento foram comparadas com teste de hipóteses não paramétrico. SST foi a variável que mais apresentou registros ausentes. O mesmo parâmetro e a DBO foram as únicas variáveis que apresentaram registros censurados. *Outliers* de DBO, DQO, N-NH₃ e SST foram mantidos/excluídos considerando o comportamento conjunto das mesmas variáveis. Todos os *outliers* de vazão foram descartados por representarem uma pequena parcela das amostras e os *outliers* de pH foram descartados em registros com variação significativa da variável. Amostras de DBO, DQO, N-NH₃ e SST apresentaram altos CDs indicando assimetria, com exceção de N-NH₃ afluente em 2019. Por outro lado, baixos CDs foram encontrados para amostras de temperatura, pH e vazão em ambos os anos. Não foi observada diferença entre os registros dos dois anos avaliados. A avaliação conjunta dos *outliers* auxiliou na exclusão/permanência dos mesmos. Os resultados reforçaram que dados do monitoramento do tratamento de esgoto (DBO, DQO, N-NH₃ e SST) geralmente se distribuem de forma não simétrica e não se adequam à distribuição normal. Portanto, a mediana das concentrações pode ser a medida de tendência central mais adequada para o banco de dados investigado. Quantidades maiores e mais homogêneas de registros podem contribuir para a realização de uma análise mais representativa das variáveis. A análise estatística deste estudo tem aplicabilidade em demais ETEs para subsidiar uma ampla avaliação do desempenho do processo de tratamento. Análises envolvendo outras medidas de distribuição log-normal também podem auxiliar na interpretação do comportamento central de variáveis de controle de processos de tratamento de esgoto.

PALAVRAS-CHAVE: Concentração afluente e efluente, Distribuição de dados ambientais, Eficiência de remoção, Tratamento de esgoto sanitário.

INTRODUÇÃO

O processo de lodos ativados é utilizado para o tratamento de esgoto sanitário de aproximadamente 16,5 milhões de habitantes no Brasil (ATLAS ESGOTOS, 2017). Esse processo requer controle específico, visando

garantir o seu melhor desempenho e a alta qualidade do esgoto tratado, de modo a atender aos padrões de lançamento de efluentes dispostos na legislação ambiental (BRASIL, 2011).

O efetivo controle do processo de lodos ativados em grandes Estações de Tratamento de Esgoto Sanitário (ETE) depende do monitoramento de parâmetros como a Demanda Bioquímica de Oxigênio (DBO), Demanda Química de Oxigênio (DQO), nitrogênio amoniacal (N-NH₃) e Sólidos Suspensos Totais (SST) (TCHOBANOGLIOUS ET AL., 2016). O monitoramento padrão de uma ETE, por sua vez, fornece os denominados dados secundários do processo de tratamento, os quais consistem em registros das concentrações dos parâmetros mencionados, que são geralmente realizados em diferentes frequências de amostragem: diária, semanal, quinzenal ou mensal (VON SPERLING, 2007).

Os registros originados do monitoramento de ETEs necessitam de análises específicas, uma vez que a maior parte dos bancos de dados possuem as seguintes características: falta de dados, dados censurados, existência de *outliers*, grande número de dados com informação limitada, ausência de medições de certas variáveis importantes e muitos erros de amostragem e análise laboratorial (VON SPERLING; VERBYLA; OLIVEIRA, 2020).

Para contornar tais circunstâncias, ferramentas de análise estatística podem ser aplicadas para o tratamento e avaliação desses registros. Assim, podem ser realizadas inferências do provável comportamento das concentrações das variáveis de controle do processo de tratamento ao longo de determinado período de monitoramento. Medidas estatísticas podem ser calculadas para caracterizar diferentes aspectos de uma amostra de dados: tamanho amostral (número de observações), tendência central (média, mediana) e variação (desvio padrão, coeficiente de variação) (EPA, 2006). O conhecimento da forma de distribuição dos registros também é fundamental. Dados de monitoramento do tratamento de esgoto geralmente se distribuem de maneira não simétrica, indicando que a mediana pode ser mais adequada que a média aritmética para representar a tendência central e que hipóteses devem ser investigadas com testes não paramétricos (VON SPERLING; VERBYLA; OLIVEIRA, 2020).

Nesse sentido, o presente estudo teve o intuito de avaliar a distribuição e a tendência central de variáveis de controle do processo de tratamento de esgoto doméstico com lodos ativados, utilizando análise gráfica e estatística de dados provenientes de monitoramento padrão de ETE de grande porte.

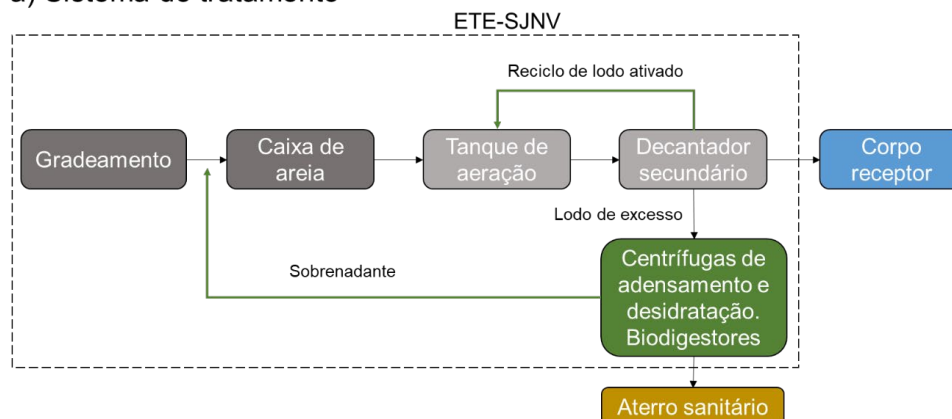
MATERIAIS E MÉTODOS

O estudo foi realizado com dados de monitoramento da ETE São João Navegantes (ETE-SJNV), situada em Porto Alegre/RS (29°59'29"S 51°11'43,5"O). A estação é considerada de grande porte, porque atende cerca de 150 mil habitantes e possui vazão nominal de 444 l/s (VON SPERLING, 2016). O processo de tratamento empregado é lodos ativados com aeração prolongada (Figura 1).

A partir de registros de monitoramento padrão, foram selecionados dados das seguintes variáveis de controle do processo, para os anos de 2018 e 2019: DBO₅, DQO, N-NH₃ (amônia), SST, pH, temperatura e vazão de esgoto tratado. Essas variáveis também constituem parâmetros estabelecidos para controle de emissão de efluentes líquidos em águas superficiais (RIO GRANDE DO SUL, 2017).

No tratamento dos dados, foram analisadas as ausências de registros e a existência de dados censurados (valores não detectados pelo método laboratorial). Estes foram substituídos pela fração de ½ do Limite de Detecção do Método analítico (VON SPERLING; VERBYLA; OLIVEIRA, 2020). Depois, para cada amostra de afluente e efluente das concentrações das variáveis avaliadas, foram identificados registros que gerassem valores negativos de eficiência de remoção (efluente > afluente). Eficiências negativas podem representar os mesmos erros que levam aos *outliers* ou ainda problemas durante o processo de tratamento que não são comuns, portanto, dispensáveis. No presente estudo, os registros relacionados às eficiências negativas foram excluídos. Igualmente, para registros de afluente e de efluente, valores de DBO superiores a DQO registrados na mesma data foram descartados.

a) Sistema de tratamento



b) Vista aérea



Figura 1 – (a) Esquema e (b) vista aérea da ETE-SJNV.

Em seguida, os *outliers* das amostras de afluente e efluente para cada variável, incluindo vazão de esgoto tratado, e de cada ano, foram identificados utilizando análise gráfica dos dados (gráfico de caixa). Depois, os *outliers* de DBO, DQO, NH₃ e SST foram avaliados em conjunto, mas separadamente para afluente e efluente em cada ano. Os *outliers* do pH e temperatura foram avaliados considerando a variação desse parâmetro, dada pela diferença entre o registro efluente e afluente em um mesmo dia. *Outliers* ligados a altas variações no pH e temperatura foram descartados. A vazão e seus *outliers* também foram avaliados graficamente ao longo do período de monitoramento, em que foram determinados os limites inferior e superior das amostras para identificação de *outliers*.

Na avaliação dos *outliers* foram considerados somente os registros remanescentes das análises anteriores, ou seja, sem registros ausentes, censurados, de eficiência de remoção negativa ou de DBO maior que DQO. As avaliações considerando todas as variáveis e o período de monitoramento visaram auxiliar à detecção de tendências nos comportamentos das variáveis e à busca por justificativas para exclusão ou permanência de determinado *outlier*. O descarte dos mesmos foi baseado em uma avaliação envolvendo todas as variáveis, em que os registros de pH e vazão auxiliaram na decisão de exclusão.

Posteriormente, foi realizada análise estatística dos dados de controle do processo. Calculou-se as medidas descritivas das variáveis para cada ano (n , mínimo, média, mediana, máximo, desvio padrão e coeficiente de variação-CV). Para cada amostra foi assumida a distribuição log-normal dos registros e foi calculado o coeficiente de distorção (CD) ($CD=3*CV+CV^3$). Amostras que se destacaram em relação à simetria/assimetria foram ilustradas através da distribuição de frequência dos registros. A forma de distribuição dos dados foi verificada para inferir qual a melhor medida de tendência central dos registros: média aritmética ou mediana.

Por fim, foi avaliado, separadamente para afluente e efluente, se existe diferença entre os registros de 2018 e 2019 para as seguintes variáveis: DBO, DQO, N-NH₃ e SST. Também foi investigado se as eficiências de remoção das mesmas variáveis diferem entre os anos. Considerando a distribuição log-normal de amostras independentes, foi aplicado o teste não paramétrico teste-*U Mann-Whitney* (VON SPERLING; VERBYLA; OLIVEIRA, 2020). O tratamento dos dados e as análises estatísticas foram feitos no programa *Microsoft Excel*.

RESULTADOS E DISCUSSÃO

Coleta e tratamento de dados de monitoramento

Os dados de monitoramento do processo de lodos ativados foram coletados junto à gerência da ETE-SJNV. O banco de dados da estação contém diversos parâmetros físicos e químicos monitorados para diversos intervalos de tempo (quinzenal, semanal, mensal). Os dados extraídos para o presente estudo abrangem o período de 2018-2019 e consistem em valores de afluente e efluente dos seguintes parâmetros: DBO, DQO, N-NH₃, pH, SST e temperatura do esgoto. A vazão também foi obtida do banco de dados, em que existem registros de produção diária de esgoto tratado. Medidas estatísticas precisas para os dados podem diferir de acordo com o período considerado (um ou dois anos). Sendo assim, as análises de tratamento dos dados foram feitas de forma separada para cada ano: 2018 e 2019.

A maior parte das amostras dos parâmetros mencionados possui registro de afluente e efluente para um mesmo dia, sendo registradas no período da manhã (apenas uma a tarde) e com intervalo de menos de uma hora entre o registro do valor afluente e efluente. Para os mesmos parâmetros, a quantidade de amostras diferiu ao longo do tempo. Em 2018 a distribuição da quantidade de amostras foi mais homogênea, sendo duas amostras por mês, totalizando 24 amostras para o ano. Em 2019, os meses de novembro e dezembro apresentaram apenas uma amostra mensal, enquanto os demais meses do ano variaram; foram obtidos registros de duas (janeiro, fevereiro, março, maio e julho), três (junho, agosto setembro e outubro) e oito (abril) amostras mensais, somando 32 amostras para o ano de 2019.

Tratando-se da vazão, existem registros diários de produção de esgoto tratado, os quais foram usados para o cálculo das estatísticas descritivas de 2018 e 2019. Para essa variável não existe ausência de registros e dados censurados.

Registros ausentes

É comum que existam valores em branco para certas variáveis quando se trabalha com dados secundários, que é o caso de registros de monitoramento padrão. Por isso, após a estruturação dos dados para os anos de 2018 e 2019, foi realizada uma avaliação do impacto da ausência de alguns dados (VON SPERLING; VERBYLA; OLIVEIRA, 2020, p. 117).

O parâmetro SST apresentou falta de registros de sua concentração afluente para uma de cada duas amostras mensais registradas em 2018, obtendo maior número de amostras de concentração efluente. Evitando influenciar uma distorção nas medidas estatísticas descritivas, para o SST foram adotados os registros de pares completos, ou seja, amostras mensais em que existiam ambos registros, afluente e efluente. Foi possível obter uma amostra desse tipo para cada mês, totalizando 12 amostras de SST (afluente e efluente) em 2018. Para o ano de 2019, foi feito o mesmo tratamento, contudo, apenas cinco das 32 amostras foram desconsideradas, restando 27 amostras para SST em 2019.

O mesmo raciocínio adotado para SST também foi empregado para outros parâmetros. Em 2018, registros de afluente de DQO foram desconsiderados diante da ausência de registro do efluente para o mesmo dia. Da mesma forma, em 2019, foi preciso desconsiderar apenas um registro de afluente de DBO. Igualmente, a

exclusão desses valores não impactou de forma significativa a população amostral, uma vez que ainda restaram pelo menos um par de registros completo (afluente e efluente) para cada mês de cada ano.

O pH apresentou seis amostras (afluente e efluente) com falhas, todas em abril de 2019. Como este mês obteve a maior quantidade de amostras (8) do referido ano (32), ainda restavam dois registros desse parâmetro para o mesmo mês. Portanto, foi considerado que tal ausência de dados também não impactaria a população amostral dos registros de pH como um todo, dado que o restante dos meses de 2019 também apresentaram pelo menos duas amostras por mês.

Os registros de temperatura não apresentaram ausência de valores. Foram considerados os registros de temperatura existentes para as mesmas datas nas quais também existiam registros para as demais variáveis.

Para a vazão de esgoto tratado também não existiam registros ausentes, embora em alguns dias do mesmo mês os registros se igualaram. Isso aconteceu em 2018 nos meses de janeiro (registros do dia 9 até 30 do mês iguais) e fevereiro (todos os dias com registros iguais). Nestes casos, como valores repetidos para tantos dias seguidos não é comum no restante dos dados, para o cálculo de *outliers* e das estatísticas descritivas da vazão de 2018, foi considerado apenas um registro de todos os repetidos para o mesmo mês, evitando fortes distorções nas medidas de tendência central.

Dados censurados

Diferentemente da ausência de dados, também existem os dados censurados, que ocorrem quando não são obtidos valores quantificáveis nas análises laboratoriais de um parâmetro. As concentrações de determinada variável do tratamento de esgoto podem ser muito baixas e, portanto, inferiores ao limite de detecção dos métodos analíticos (LDM). No banco de dados de monitoramento, o valor da concentração para esses casos é informado como sendo menor que o LDM (VON SPERLING; VERBYLA; OLIVEIRA, 2020, p. 117).

No presente estudo, apenas dois parâmetros apresentaram dados censurados. O SST obteve 17% de suas amostras de efluente com valores censurados (indicado como < 10 mg/L) em 2018. Em 2019, 15% das amostras de efluente foram censuradas. Como estes dados censurados representam menos de 20% da população amostral de registros do efluente de SST em cada ano, foi adotado um método de substituição do valor censurado pela fração de ½ do LDM: 50% do intervalo entre zero e o LDM (VON SPERLING; VERBYLA; OLIVEIRA, 2020, p. 119). A partir deste método os valores censurados de SST foram considerados como 5 mg/L. O mesmo foi feito para duas amostras de efluente da DBO em 2019 (6,5% da população amostral), em que o valor substituído consistiu em 4 mg/L, ou seja, metade do LDM igual a 8 mg/L.

Avaliação de outliers

A quantidade de registros que apresentaram eficiência de remoção negativa (efluente>afluente), bem como o número de *outliers* identificados e descartados estão apresentados na Tabela 1. Os *outliers* identificados para cada variável em 2018 e 2019 estão apresentados na Figura 2.

Tabela 1 – Número de registros de concentrações com valor de efluente maior que afluente (eficiência de remoção negativa) e outliers identificados e descartados para cada variável em 2018 e 2019.

Variável	Efluente > Afluente		Outliers identificados				Outliers descartados			
	2018	2019	Afluente		Efluente		Afluente		Efluente	
			2018	2019	2018	2019	2018	2019	2018	2019
DBO	1	0	1	1	1	4	0	1	0	4
DQO	1	2	2	2	1	4	0	2	1	4
Amônia	0	3	6	4	1	4	0	4	1	4
SST	1	6	0	3	0	3	0	2	0	3
Temperatura	-	-	0	0	0	0	0	0	0	0
pH	-	-	0	3	3	2	0	2	1	0
Vazão	-	-	-	-	8	6	-	-	8	6

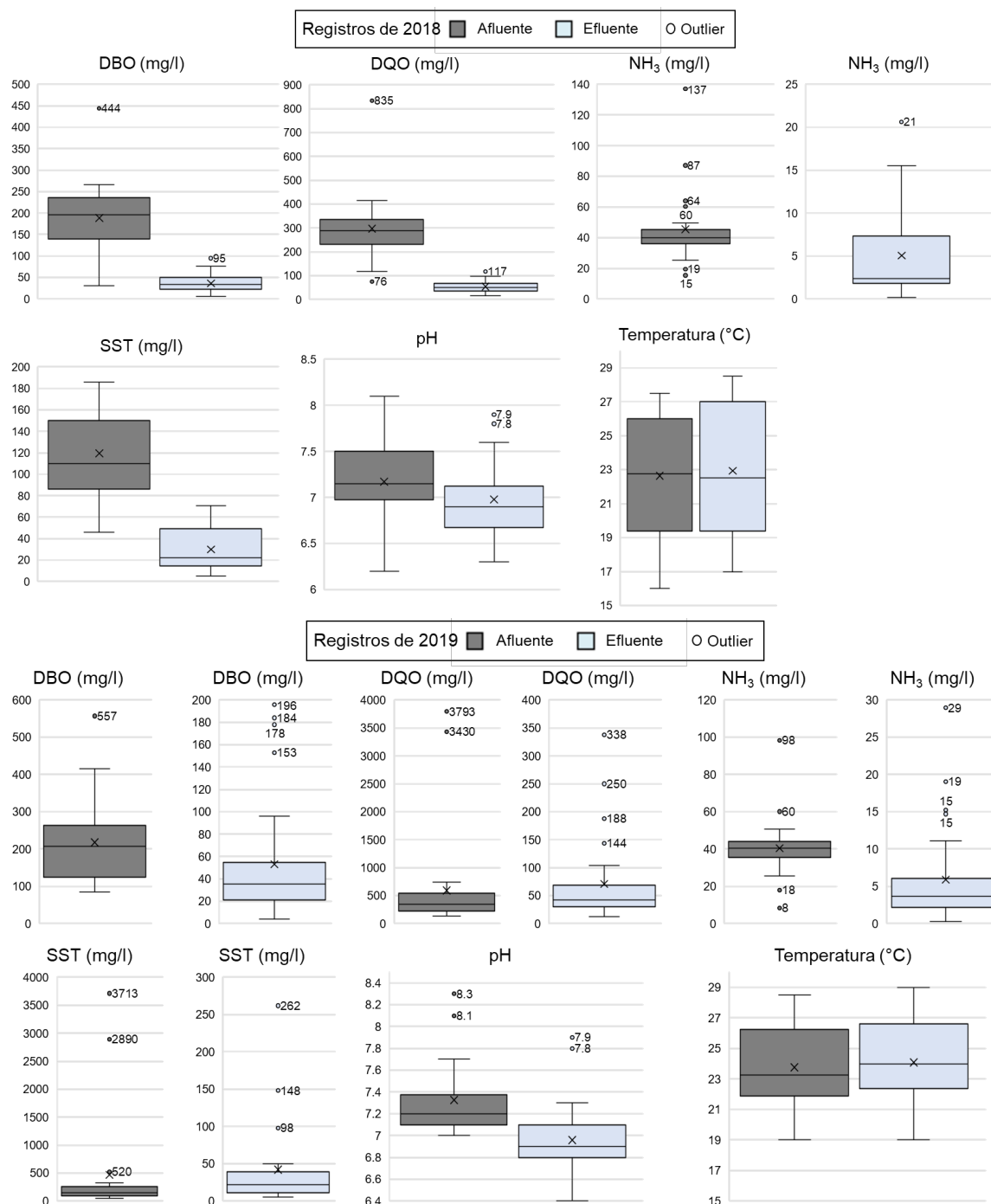


Figura 2 – Outliers identificados nos registros de monitoramento da ETE-SJNV.

Foram encontrados 14 registros em que a eficiência de remoção foi negativa, sendo três em 2018 para DBO, DQO e SST e 11 em 2019 para DQO (2), amônia (3) e SST (6). Enquanto em 2018 não foram observados valores de DBO maiores que DQO, em 2019 foram encontrados dois, em um registro de afluentes em janeiro e de efluente em fevereiro.

Esses resultados podem ser comparados com os achados de outro estudo, que utilizou dados de monitoramento padrão da mesma ETE para as mesmas variáveis, mas para o período de 2006 até 2012 (BAPTISTA, 2020). A autora observou 11 registros com eficiência negativa, envolvendo DBO, DQO e SST, e 30 de DBO superior à DQO. A distinção entre os valores encontrados pode estar associada ao tamanho do período adotado (seis anos de monitoramento). Já a semelhança entre as variáveis com eficiência negativa pode indicar que possíveis

erros de coleta, de registro ou de análises laboratoriais são comuns para as variáveis no monitoramento padrão da ETE-SJNV.

O total de *outliers* identificados no afluente (22) e no efluente (23) no presente estudo foram semelhantes em relação às variáveis (Tabela 1), mas menores que a quantidade encontrada por Baptista (2020). A autora encontrou mais *outliers* para DBO e SST afluentes e menos para amônia afluente. DQO e pH afluente apresentaram semelhança na quantidade de *outliers*. No efluente, a autora registrou menos *outliers* para a maior parte das variáveis (DQO, amônia e pH), mas 22 *outliers* para DBO efluente, sendo a maior quantidade de *outliers* do estudo. Para a temperatura, a identificação de *outliers* foi semelhante aos resultados da autora, sendo que ela encontrou apenas um *outlier* nos registros de temperatura analisados.

Diferentemente de Baptista (2020), que descartou todos os *outliers* identificados, no presente estudo alguns foram mantidos com base em avaliações envolvendo os *outliers* de todas as variáveis. Os resultados dessas avaliações serão apresentados para 2018 e 2019.

- **Registros de 2018**

Foram identificados nove *outliers* nos registros do afluente em 2018 (Tabela 1). Apesar do SST não apresentar *outlier* no afluente, seus registros variaram ao longo do ano. Em contrapartida, embora o pH também não apresentou *outliers* no afluente de 2018, a variação de seus valores foi menor, sendo próxima a uma unidade, acima ou abaixo, em situações específicas que se relacionam com as demais variáveis e são discutidas posteriormente.

A partir da avaliação conjunta dos *outliers* do afluente em 2018 (Figura 3a), optou-se pelo não descarte dos *outliers* identificados, considerando as seguintes observações. Na segunda metade de março de 2018, junto ao baixo valor de amônia (N-NH₃), também foi observada redução de DBO. O mesmo aconteceu em junho, incluindo também a DQO. Por outro lado, na segunda metade de abril, um grande aumento foi observado nas concentrações afluentes de DBO, DQO e N-NH₃. Em julho, a mesma situação de aumento das três variáveis foi identificada, mesmo que com menor intensidade. Por fim, na segunda metade de setembro, NH₃ aumentou a um nível aproximado ao de julho, acompanhada pelo aumento de DQO.

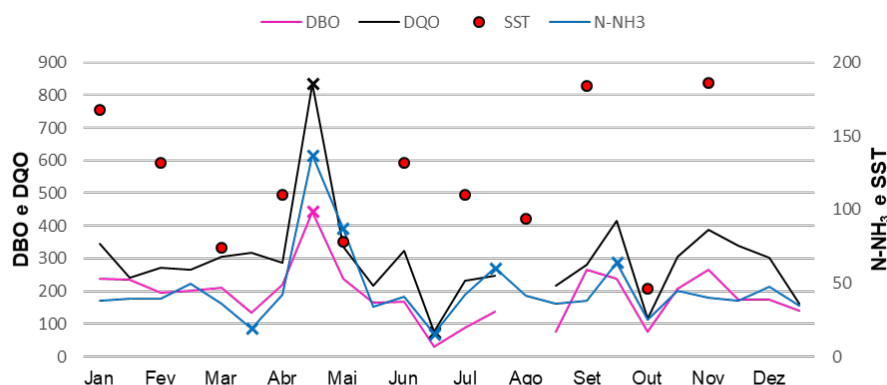
Essa tendência de redução ou aumento simultâneo de DBO, DQO e amônia representa possíveis alterações na composição do esgoto afluente, o que justifica manter os *outliers* na análise. Retirá-los poderia gerar medidas de tendência central que não representam a variação real que aconteceu no ano.

A variação do pH reforça essa tendência, sendo que no afluente o menor valor de pH (6,2) foi obtido em março e os maiores valores (8,1 e 8,0) em abril e setembro. Vale ainda ressaltar que todos os *outliers* identificados nesses dados correspondem a registros feitos sempre na segunda metade do mês, o que poderia estar ligado a algum fator operacional da estação. Contudo, a gerência da ETE-SJNV relatou não haver manutenções ou problemas no processo de tratamento que sejam recorrentes na segunda metade dos meses de 2018.

Para os registros de concentração no efluente de 2018, foram identificados apenas um *outlier* para DBO, DQO e amônia (Tabela 1). O *outlier* da DBO em janeiro está associado a um aumento na DQO e a uma aproximação entre as duas concentrações, embora o valor de DQO ainda tenha sido maior (Figura 3b). Portanto, optou-se por não excluir o *outlier* da DBO. Os demais *outliers* de amônia e DQO, em agosto e dezembro, respectivamente, foram descartados por não haver alguma relação com as demais variáveis nos mesmos registros.

Para os registros de vazão em 2018, foram observados picos anormais. Um deles ocorreu no final do mês de setembro, equivalente a 2.571,64 l/s, valor superior a cinco vezes a capacidade da ETE estudada (444 l/s). Um pico maior ainda aconteceu no final de novembro. Estes dois picos consistem nos maiores *outliers* para vazão em 2018.

a) afluente



b) efluente

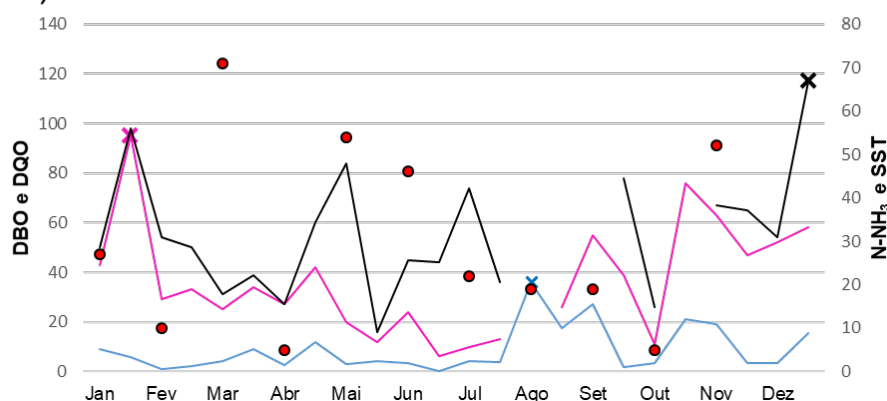


Figura 3 – Registros (mg/l) de (a) afluente e (b) efluente de parâmetros monitorados pela ETE-SJNV em 2018, referentes ao tratamento de esgoto sanitário com lodos ativados. *Outliers* identificados com o símbolo “X”. A descontinuação da linha representa ausência de registros da variável em determinada data.

Uma justificativa para picos de vazão pode estar relacionada aos registros de altos valores de SST no afluente nos mesmos meses (Figura 3a), porém esses valores não constituem *outliers* do SST, ou seja, não representam anormalidade na concentração do parâmetro em 2018. Outra explicação para os altos picos de vazão pode ser a ocorrência de chuvas que foram registradas pela ETE na segunda metade de setembro. Contudo, em novembro não teve registro de chuva nas coletas realizadas. Inobstante, é bem provável que o efeito de chuvas na região não seja tão forte sobre o esgoto afluente à ETE, conforme relatado pela equipe de operação da estação. Isso se deve ao fato da estação possuir um grande poço de acumulação no início do sistema, fazendo com que esse tipo de efeito seja amortecido antes da medição de vazão e do tratamento biológico. Com base nessas evidências, optou-se pela exclusão desses picos de vazão.

Em seguida, foram determinados os *outliers* de vazão para o restante dos dados de 2018, sendo obtido oito *outliers* (Tabela 1). Os registros de vazão de esgoto tratado em 2018, sem os dois altos picos de setembro e novembro, estão apresentados na Figura 4. Os *outliers* desses registros consistem nos valores abaixo do limite inferior (13.893 m³/d) e acima do limite superior (38.861 m³/d), os quais representaram cerca de 0,6% e 2,5% do total de registros, respectivamente. Dada as baixas proporções, optou-se também por desconsiderá-los na análise estatística, de modo a evitar maiores distorções nas medidas de descrição da vazão em 2018.

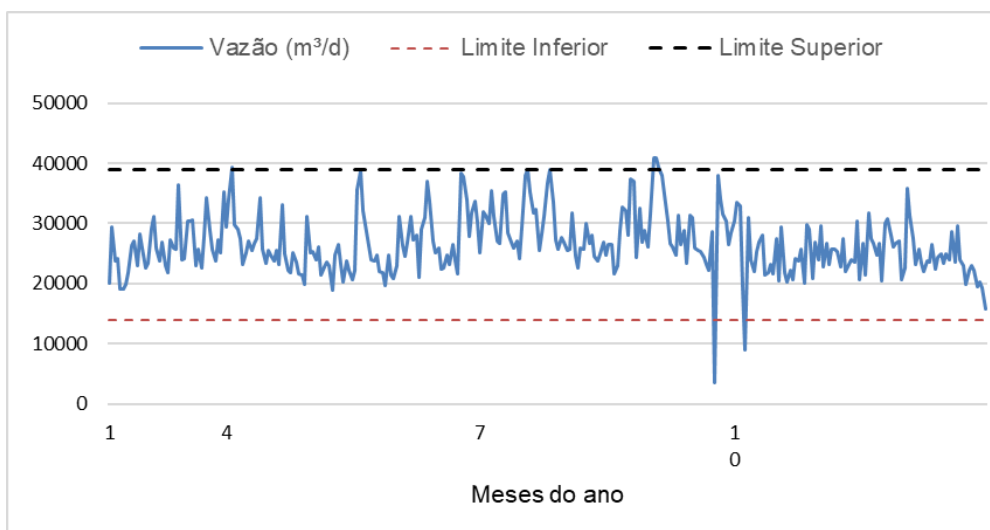


Figura 4 – Vazão de esgoto tratado monitorada pela ETE-SJNV em 2018, referente ao tratamento de esgoto sanitário com lodos ativados. Registros diários de vazão, exceto para janeiro (10 registros) e fevereiro (um registro).

Os registros de pH afluente não apresentaram *outliers*, mas foram identificados três *outliers* para pH efluente em 2018 (Tabela 1). A avaliação do descarte dos mesmos foi baseada na variação do pH ao longo do ano (Figura 5); sendo positiva quando o pH aumenta no efluente, negativa quando diminui e nula quando não há variação. Apenas os registros de dezembro não mostraram variação de pH. Na Figura 5 ainda estão indicados os *outliers* identificados nos registros dos meses de agosto (7,9), setembro (7,8) e novembro (7,8).

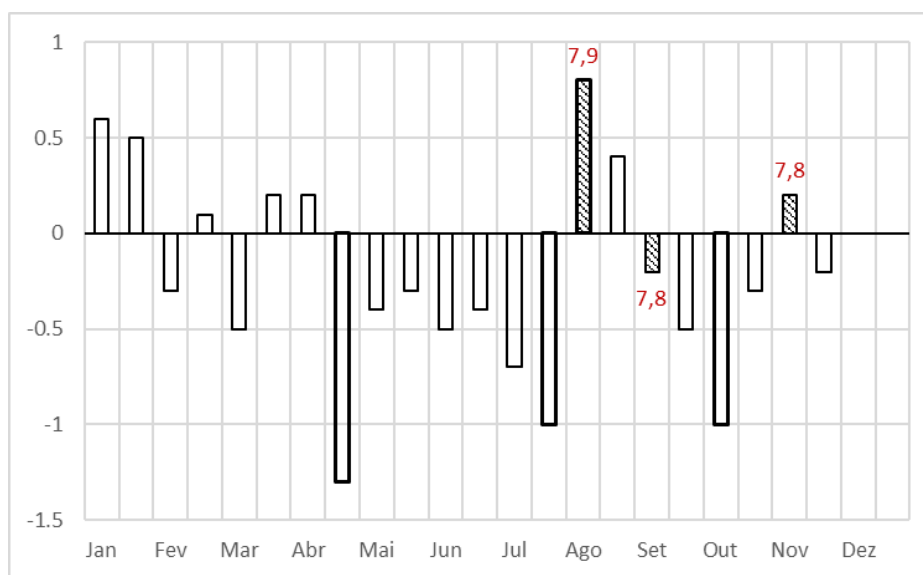


Figura 5 – Variação do pH (efluente - afluente) no tratamento de esgoto sanitário com lodos ativados na ETE-SJNV ao longo de 2018. *Outlier* identificado nos meses com caixa cheia. Valor do *outlier* no pH efluente em vermelho.

Das 24 amostras, apenas quatro se destacaram em relação à variação do pH, em que a maior variação foi negativa e igual 1,3 (Figura 5). Com base nessas informações, optou-se por descartar apenas o *outlier* correspondente a uma alta variação de pH (+0,8), relativo ao registro de agosto de 2018 (pH = 7,9), sendo ainda o maior valor de pH efluente observado no ano. Os outros dois *outliers* (setembro e novembro) foram mantidos, uma vez que nessas amostras a variação do pH no processo de tratamento não superou meia unidade. E como seus valores são iguais (7,8), sua exclusão provocaria alta distorção nas medidas estatísticas para esta variável.

• **Registros de 2019**

Foram identificados 13 *outliers* nos registros do afluente em 2019, envolvendo todas as variáveis (exceto vazão que é analisada como efluente) (Tabela 1).

Em março e junho, ocorreu aumento simultâneo de DBO, DQO, amônia e SST (Figura 6a). Contudo, os valores de DQO e SST para estes casos foram extremamente altos: 3.430 mg/l e 2.890 mg/l (março) e 3.793 mg/l e 3.713 mg/l (junho), respectivamente. Estes *outliers* foram descartados por estarem fortemente isolados da faixa usual de concentração das referidas variáveis. Considerar estes *outliers* influenciaria de forma significativa as medidas de tendência central dos registros, o que não aparenta ser a melhor alternativa, visto que são apenas dois registros em 28 para DQO e em 21 para SST.

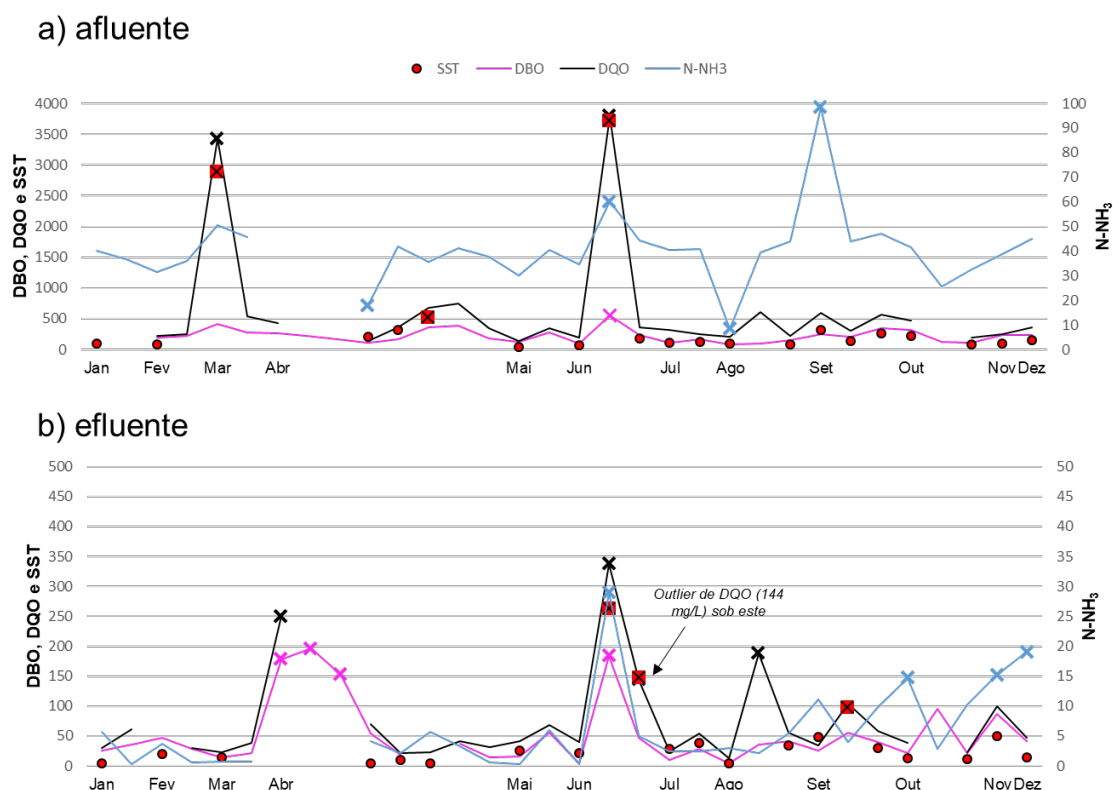


Figura 6 – Registros (mg/l) de (a) afluente e (b) efluente de parâmetros monitorados pela ETE-SJNV em 2019, referentes ao tratamento de esgoto sanitário com lodos ativados. *Outliers* identificados com o símbolo “X”. A descontinuação da linha representa ausência de registros da variável em determinada data.

Ainda em junho, amônia e DBO afluente também aumentaram e constituíram *outliers* (Amônia 60,15 mg/l; DBO 557 mg/l) (Figura 6a). Estes valores também foram descartados do banco de dados, em concordância com a exclusão dos *outliers* de DQO e SST do mesmo mês. O *outlier* de DBO também é o único para a variável em todo o ano.

Amônia teve outro aumento em setembro, indicando seu maior *outlier* em 2019 (98,34 mg/l) (Figura 6a), que também foi excluído pela sua distinção de valor, evitando grandes distorções da tendência central do composto químico nos registros de 2019. Registros de amônia ainda compõem os menores *outliers* do afluente do ano, sendo detectados em abril (17,98 mg/l) e agosto (8,46 mg/l). Estes dois pequenos valores foram descartados, uma vez que não foi notada relação dessas reduções com o comportamento das outras variáveis que justificasse mantê-los. De qualquer forma, ainda restaram quatro registros de amônia em abril e dois em agosto, sendo suficiente para representar os meses na distribuição anual.

Por fim, optou-se por manter o menor valor de *outlier* de SST, que foi registrado no final de abril (520 mg/l) (Figura 6a). Este valor indica um aumento de SST acompanhado de aumento com intensidade semelhante nas concentrações de DQO, DBO e amônia, o que justifica mantê-lo.

Tratando-se dos registros de concentração no efluente do tratamento de esgoto durante o ano de 2019, foram determinados 15 *outliers* (Figura 6b). A maior parte dos registros de DBO, DQO e SST situaram-se em torno de 100 mg/l ou inferior. Optou-se por descartar os três *outliers* de DBO e um de DQO em abril, porque não existem registros de SST e amônia nas mesmas datas para interpretar possíveis aumentos simultâneos das variáveis. No afluente em abril também não houve aumento de DBO e DQO, que explicasse o consequente aumento da concentração efluente dessas variáveis.

Os *outliers* de junho também foram excluídos, em concordância com a exclusão de *outliers* realizada nos registros de afluente do mesmo mês (Figura 6a). O *outlier* de DQO em agosto também foi descartado, levando em consideração a baixa eficiência de remoção obtida (69%) neste dia para DQO. Embora o *outlier* de SST em setembro representa um aumento da variável junto com DQO e DBO e o valor do *outlier* está abaixo de 100 mg/l, a eficiência de remoção desse registro foi de apenas 31%, bem abaixo do normal. Entende-se que esta baixa eficiência não é comum, portanto, este *outlier* foi descartado. Da mesma forma, para os *outliers* de amônia em outubro (14,71 mg/l), novembro (15,23 mg/l) e dezembro (19,03 mg/l), as eficiências de remoção dos registros são em torno de 60%, valor abaixo do normal também. Estes *outliers* foram descartados.

A maior parte dos registros de vazão em 2019 está compreendida entre 13.230 m³/d e 40.080 m³/d, sendo que nenhum valor foi menor que o primeiro (Figura 7). Em contrapartida, um registro no final de abril supera o limite superior da faixa usual de observações, indicando a maior vazão do ano (50.576 m³/d). Chuvas foram registradas pela estação uma semana antes e depois da data deste pico de vazão, mas não exatamente no mesmo dia. Além disso, os valores dos demais *outliers* de vazão identificados em 2019 não foram tão diferentes da faixa usual. Assim como em 2018, todos os *outliers* representam uma pequena parte dos dados, cerca de 1,6% dos 365 registros diários de vazão no referido ano. Por essas razões, optou-se por descartar todos os seis *outliers* identificados nos registros de vazão em 2019.

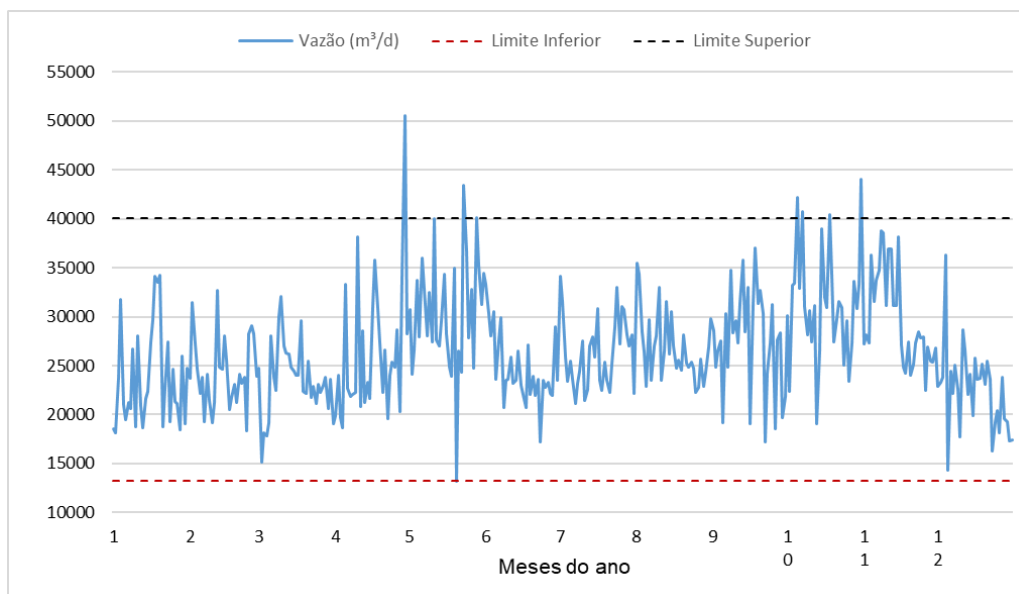


Figura 7 – Vazão de esgoto tratado monitorada pela ETE-SJNV em 2019, referente ao tratamento de esgoto sanitário com lodos ativados. Registros diários de vazão.

Os *outliers* identificados no pH resultaram em um número um pouco maior para 2019 (5), em relação a 2018 (3). Em 2019, entretanto, foram obtidos *outliers* nas duas fases do esgoto, afluente (3) e efluente (2). Outra diferença mais interessante consiste no fato de que em 2018 as ocorrências de *outliers* no pH, além de serem apenas em amostras do efluente, foram mais espaçadas entre as amostras do ano. Já em 2019, os *outliers* se concentraram nas amostras coletadas em agosto e setembro (Figura 8).

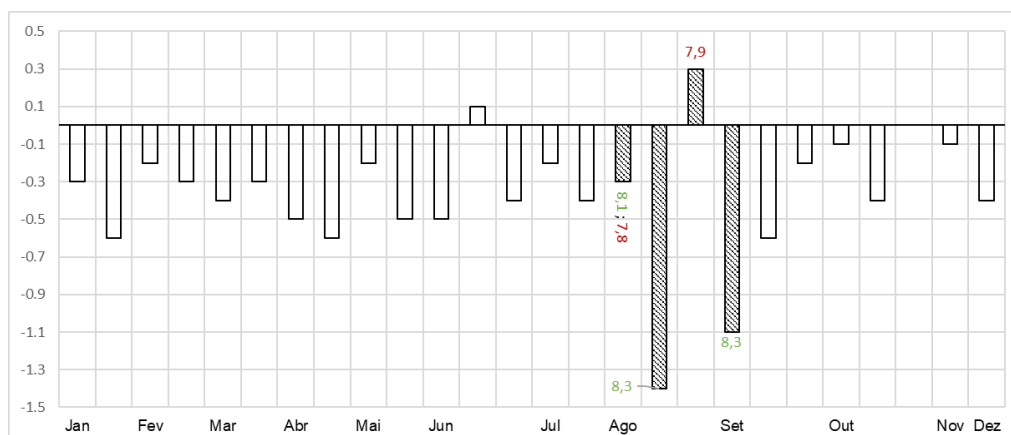


Figura 8 – Variação do pH (efluente - afluente) no tratamento de esgoto sanitário com lodos ativados na ETE-SJNV ao longo de 2019. *Outlier* identificado nos meses com caixa cheia. Valor do *outlier* no: pH afluente em verde; pH efluente em vermelho.

Da mesma forma que foi feito para os registros de pH em 2018, optou-se por descartar os *outliers* identificados em amostras com as maiores variações no pH (-1,4 e -1,1), correspondente aos *outliers* do pH afluente (pH = 8,3) de agosto e setembro de 2019 (Figura 8). Estes valores são ainda os maiores *outliers* do monitoramento do pH em 2019 e se originam das únicas amostras com variação superior a uma unidade de pH.

Estatísticas descritivas dos registros

A Tabela 2 apresenta as estatísticas que descrevem as concentrações de DBO, DQO, amônia e SST. As estatísticas das variáveis temperatura, pH e vazão estão dispostas na Tabela 3.

Tabela 2 – Estatísticas descritivas (mg/l) dos parâmetros que apresentaram maiores coeficientes de distorção (CD). (CV: Coeficiente de variação).

	DBO		DQO		N-NH ₃		SST	
	2018	2019	2018	2019	2018	2019	2018	2019
Afluente								
<i>n</i>	23	29	21	26	24	25	11	19
Mínimo	31	85	76	133	15,36	25,74	46	49
Média	189	206	297	359	45,49	39,45	119	170
Mediana	196	199	288	329	39,95	40,61	110	124
Máximo	444	415	835	743	137,17	50,58	186	520
Desvio Padrão	85	88	149	171	24,21	5,70	46	118
CV (%)	45	43	50	48	53,21	14,44	39	69
CD (adimensional)	1,4	1,4	1,6	1,5	1,7	0,4	1,2	2,4
Efluente								
<i>n</i>	23	26	20	24	23	25	11	18
Mínimo	6	4	16	13	0,19	0,29	5	5
Média	37	34	52	45	4,39	3,7	30	21
Mediana	33	29	50	40	2,33	3	22	17
Máximo	95	96	98	104	15,50	11,09	71	50
Desvio Padrão	23	22	21	23	4,25	3,12	22	15
CV (%)	62	66	41	52	96,77	84,15	74	69
CD (adimensional)	2,1	2,3	1,3	1,7	3,8	3,1	2,6	2,4

Remoção média (%)	81	84	83	88	90,36	90,61	75	88
-------------------	----	----	----	----	-------	-------	----	----

Tabela 3 – Estatísticas descritivas dos parâmetros que apresentaram menores coeficientes de distorção (CD). (CV: Coeficiente de variação).

	Temperatura (°C)		pH		Vazão (m³/d)	
	2018	2019	2018	2019	2018	2019
Afluente						
<i>n</i>	24	28	24	24	-	-
Mínimo	16	19	6,2	7	-	-
Média	22,6	23,8	7,2	7,2	-	-
Mediana	22,8	23,3	7,2	7,2	-	-
Máximo	27,5	28,5	8,1	8,1	-	-
Desvio Padrão	3,7	3	0,5	0,3	-	-
CV (%)	16,4	12,6	6,4	3,7	-	-
CD (adimensional)	0,5	0,4	0,2	0,1	-	-
Efluente						
<i>n</i>	24	28	23	26	307	359
Mínimo	17	19	6,3	6,4	15.882	13.230
Média	22,9	24,1	6,9	7	26.479	26.147
Mediana	22,5	24	6,9	6,9	25.800	25.382
Máximo	28,5	29	7,8	7,9	38.786	40.080
Desvio Padrão	4,1	3,1	0,4	0,3	4.390	5.017
CV (%)	17,9	12,8	5,8	4,9	17	19
CD (adimensional)	0,5	0,4	0,2	0,1	0,5	0,6
Variação média (efluente - afluente)	0,3	0,3	-0,3	-0,2	-	-

O número de observações foi igual para 2018 e 2019 apenas nas amostras de pH afluente. No restante, as amostras de 2019 apresentaram mais registros que 2018, destacando-se as amostras de SST em 2019 que apresentaram no afluente oito registros a mais que 2018, e no efluente sete. Da mesma forma, foram obtidos mais registros de vazão em 2019, embora o número de observações soma mais de 300 registros nos dois anos. Por outro lado, as amostras de SST foram as menores de todas as amostras consideradas.

Em relação à variação dos dados (CV), de forma geral, o pH foi a variável que menos sofreu variação em seus valores (CV<6,5%) (Tabela 3). Os registros de temperatura em 2019 foram os que apresentaram a segunda menor variação tanto no afluente (CV=12,6%) como no efluente (CV=12,8%). A temperatura em 2018 e a vazão nos dois anos mostraram variação entre 16 e 19%. Na outra ponta se encontram as amostras de amônia efluente, as quais apresentaram maior variação de registros nos dois anos (CV>84%) (Tabela 2).

Comparando a variação das amostras de cada variável entre os anos, no afluente, DBO e DQO apresentaram variação semelhante nos dois anos, enquanto amônia e SST diferiram em mais de 30% entre 2018 e 2019. No efluente, as diferenças entre os anos para cada variável foram mais baixas, sendo que as maiores distinções ocorreram para as amostras de amônia (13%) e DQO (11%) entre 2018 e 2019 (Tabela 2).

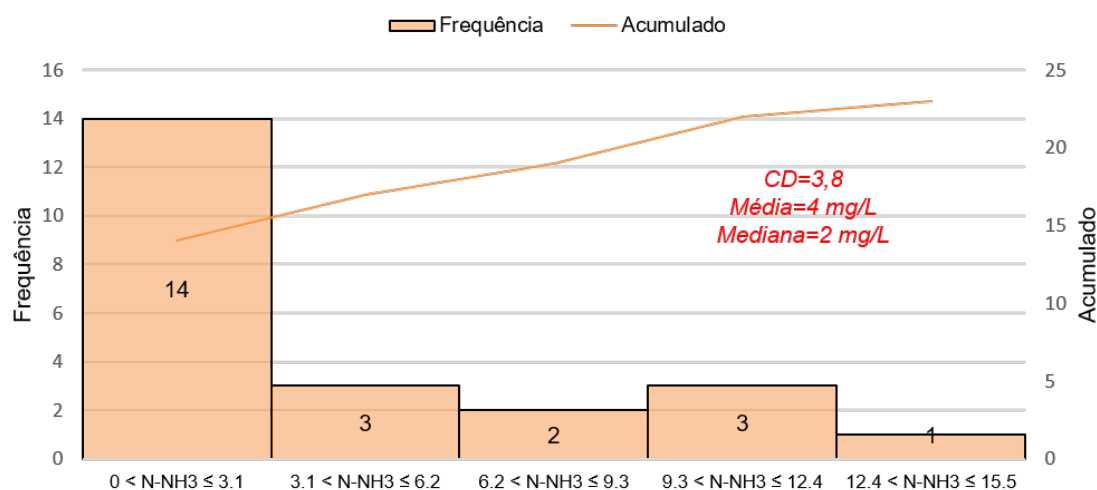
Alguns autores relatam que a composição típica de esgoto doméstico bruto considerado fraco e com pequenas contribuições de efluente industrial pode apresentar as seguintes concentrações (VOLCKE ET AL., 2020): DQO de 500 mg/l, DBO de 230 mg/l, amônia de 20 mg/l e SST de 250 mg/l. A composição do esgoto afluente à ETE-SJNV é inferior a estes valores para DQO, DBO e SST nos dois anos de monitoramento adotados no presente estudo (Tabela 2). Apenas a variável amônia apresentou concentração superior (2018: 45 mg/l; 2019: 39 mg/l), se aproximando da condição de esgoto considerado médio pelos autores (45 mg/l).

As médias das concentrações no afluente registradas em 2019 foram superiores aos registros de 2018, com exceção da variável amônia. As maiores diferenças no afluente entre os dois anos aconteceram para DQO (62 mg/l) e SST (51 mg/l). Nos registros do efluente foi observado o contrário, sendo que todas as quatro variáveis de controle apresentaram concentrações maiores em 2018. As diferenças no efluente entre os anos não ultrapassaram 10 mg/l para as variáveis.

Distribuição dos registros

Os parâmetros da Tabela 2 apresentaram maiores coeficientes de distorção (CD) e, por isso, se destacaram em relação à assimetria. Por outro lado, as variáveis da Tabela 3 apresentaram CD próximo a zero, ou seja, se mostraram com distribuição simétrica. A forte assimetria da amostra de amônia efluente em 2018 se destacou (CD=3,8), enquanto amônia afluente em 2019 foi a única exceção dos altos CDs da Tabela 2 (CD=0,4) (Figura 9).

a) Amônia efluente (mg/L) - 2018



b) Amônia afluente (mg/L) - 2019

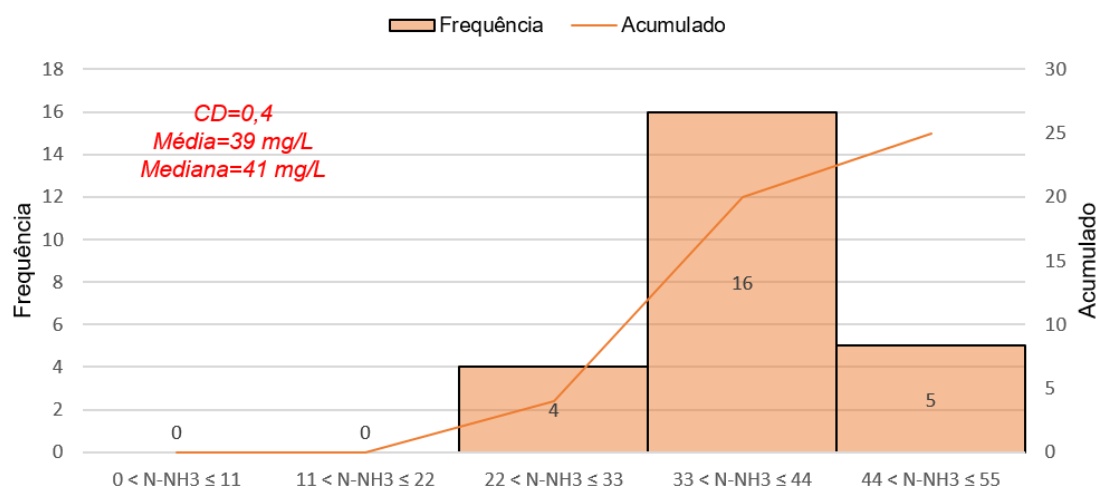


Figura 9 – Distribuição de frequência de registros de amônia sinalizando (a) assimetria e (b) simetria da distribuição.

É comum que valores de concentração (afluente e efluente) de variáveis do tratamento de esgoto apresentem distribuição assimétrica, por serem dados com limite mínimo (não existe concentração negativa) e sem limites

máximos (embora o efluente deve atender os limites legais). Por isso, a distribuição log-normal se adequa a esse tipo de dado (VON SPERLING; VERBYLA; OLIVEIRA, 2020).

Os resultados evidenciam essa suposição, em que o CD da distribuição log-normal indicou que a maior parte das amostras é assimétrica ($CD > 1$) (Tabela 2), destacando-se a amônia efluente em 2018 (Figura 9a). Portanto, a mediana pode ser adequada como medida de tendência central para esses dados (VON SPERLING; VERBYLA; OLIVEIRA, 2020). Por outro lado, a amônia afluente em 2019 apresentou baixo CD e se mostrou simétrica (Figura 9b). Além disso, a média e mediana desta amostra são quase iguais. Esses fatores indicam que a distribuição log-normal destes registros é similar à distribuição normal e, pela simetria, tanto a média quanto a mediana pode ser adotada como medida de tendência central dos registros (OTT, 1995; VON SPERLING; VERBYLA; OLIVEIRA, 2020).

Portanto, é importante conhecer a distribuição de um conjunto de dados, uma vez que adotar a média aritmética para uma amostra sem confirmação da simetria de sua distribuição pode levar a resultados incorretos e não confiáveis (MODARRES ET AL., 2005).

Comparação entre os registros de 2018 e 2019

As concentrações (afluente e efluente) e eficiências de remoção das variáveis não diferiram significativamente entre 2018 e 2019. Diante das estatísticas descritivas do afluente (Tabela 2), esperava-se que a diferença entre as medianas de 2018 e 2019 fossem significativas para DQO e SST. O resultado contrário gerado pelo teste de hipóteses pode estar relacionado ao tamanho das amostras em dois aspectos: tamanhos diferentes entre os anos; e as maiores amostras ($n \approx 30$) podem ser consideradas pequenas, dificultando a rejeição da hipótese nula – igualdade das medianas (EPA, 2006; VON SPERLING; VERBYLA; OLIVEIRA, 2020). A mesma explicação cabe ao resultado da análise das eficiências de remoção, que não foram diferentes entre 2018 e 2019, embora a maior diferença de remoção média absoluta foi de 13% para SST (Tabela 2).

CONCLUSÕES

Medidas estatísticas foram usadas para avaliação das variáveis de controle do processo de lodos ativados, usando dados de monitoramento padrão que representam ETE de grande porte em escala real. A variável SST apresentou maior número de registros ausentes e censurados. A avaliação conjunta dos dados de todas as variáveis possibilitou visualizar relações entre os *outliers*, indicando possíveis alterações das concentrações que justificavam o não descarte de alguns valores.

O coeficiente de distorção da maior parte das amostras de DBO, DQO, N-NH₃ e SST indicou assimetria da distribuição dos registros das amostras. Portanto, é possível reforçar que dados do tratamento de esgoto geralmente se distribuem de forma não simétrica e não se adequam à distribuição normal. Esta análise de simetria, por sua vez, aponta que a mediana das concentrações pode ser a medida de tendência central mais adequada para o banco de dados investigado, considerando os parâmetros de monitoramento mencionados. Por outro lado, as amostras de temperatura, pH e vazão se mostraram simétricas, ou seja, tanto a média como a mediana dessas amostras podem representar a tendência central dos registros.

As concentrações afluente e efluente das respectivas variáveis, bem como suas eficiências de remoção não diferiram entre os períodos estudados (2018 e 2019). O tamanho das amostras disponíveis pode ter limitado a análise estatística, sendo que quantidades maiores e mais homogêneas de registros podem contribuir para a realização de uma análise mais representativa das variáveis.

A análise estatística realizada neste estudo tem aplicabilidade em situações similares, em que os registros de monitoramento padrão de estações de tratamento de esgoto podem subsidiar uma ampla avaliação do desempenho do processo de tratamento. Associação das medidas estatísticas com as especificações da legislação ambiental constituem alternativas para outros estudos. Análises envolvendo outras medidas de distribuição log-normal, como média geométrica, também podem auxiliar na interpretação do comportamento central de variáveis de controle de processos de tratamento de esgoto.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ATLAS ESGOTOS. *Despoluição das Bacias Hidrográficas*. Agência Nacional de Águas, Secretaria Nacional de Saneamento Ambiental. 2017. Disponível em: <atlasesgotos.ana.gov.br/>. Acesso em 9 dez. 2020.
2. BAPTISTA, M. T. *Avaliação dos modelos matemáticos ASM1 e ASM3 calibrados com dados de monitoramento padrão de um sistema de lodos ativados*. 2020. Dissertação (Mestrado em Recursos Hídricos e Saneamento Ambiental) – Instituto de Pesquisas Hidráulicas, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2020. 96 f.
3. BRASIL. *Resolução CONAMA No 430/2011* - Dispõe sobre condições e padrões de lançamento de efluentes, complementa e altera a Resolução no 357, de 17 de março de 2005, do Conselho Nacional do Meio Ambiente - CONAMA. Brasília, Ministério do Meio Ambiente - Conselho Nacional de Meio Ambiente, 2011. Disponível em: <<http://www2.mma.gov.br/port/conama/legiabre.cfm?codlegi=646>>. Acesso em 20 out. 2020.
4. EPA - UNITED STATES ENVIRONMENTAL PROTECTION AGENCY. *Data Quality Assessment: Statistical Methods for Practitioners*. Washington, DC. 2006.
5. MODARRES, Reza; GASTWIRTH, Joseph L.; EWENS, Warren. A cautionary note on the use of non-parametric tests in the Analysis of Environmental Data. *Environmetrics*, [s. l.], v. 16, n. 4, p. 319–326, 2005. <http://doi.wiley.com/10.1002/env.695>
6. OTT, W. R. *Environmental Statistics and Data Analysis*. Boca Raton, Florida: CRC Press LLC, 1995.
7. RIO GRANDE DO SUL. *Resolução CONSEMA no 355/2017* - "Dispõe sobre os critérios e padrões de emissão de efluentes líquidos para as fontes geradoras que lancem seus efluentes em águas superficiais no Estado do Rio Grande do Sul. Porto Alegre, Secretaria do Ambiente e Desenvolvimento Sustentável - Conselho Estadual do Meio Ambiente, 2017. Disponível em: <<https://www.sema.rs.gov.br/upload/arquivos/201707/19110149-355-2017-criterios-e-padres-de-emissao-de-efluentes-liquidos.pdf>>. Acesso em 20 out. 2020.
8. TCHOBANOGLIOUS, G.; STENSEL, H. D.; TSUCHIHASHI, R.; BURTON, F.; ABU-ORF, M.; BOWDEN, G.; PFRANG, W. *Tratamento de Efluentes e Recuperação de Recursos*. Tradução de Ivanildo Hespanhol, José Carlos Mierzwa. Porto Alegre: AMGH, 2016. Ed. 5, 1984 p.
9. VOLCKE, E. I. P.; SOLON, K.; COMEAU, Y.; HENZE, M. Wastewater characteristics. In: CHEN, G.; VAN LOOSDRECHT, M.C.M.; EKAMA, G.A.; BRDJANOVIC, D. (Eds.). *Biological Wastewater Treatment: Principles, Modeling and Design*. Londres: IWA Publishing, 2020. ed. 2, cap. 3, p. 77–110.
10. VON SPERLING, Marcos. Activated Sludge and Aerobic Biofilm Reactors. In: *Biological Wastewater Treatment Series*. Londres: IWA Publishing, 2007.
11. _____. *Urban wastewater treatment in Brazil*. Water and Sanitation Division – Technical Note N° IDB-TN-970. IDB, 2016. Disponível em: <<https://publications.iadb.org/en/urban-wastewater-treatment-brazil>>. Acesso em: 22 jun. 2021.
12. VON SPERLING, Marcos; VERBYLA, Matthew E.; OLIVEIRA, Sílvia M. A. C. *Assessment of Treatment Plant Performance and Water Quality Data: A Guide for Students, Researchers and Practitioners*. Londres: IWA Publishing, 2020.