

UNIVERSIDADE DO ESTADO DE SANTA CATARINA – UDESC
CENTRO DE CIÊNCIAS TECNOLÓGICAS – CCT
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL

ALLYSON BELLI BOGO

**GRÁFICOS DE CONTROLE ESTATÍSTICO DE PROCESSO PARAMÉTRICOS E
NÃO PARAMÉTRICOS APLICADOS AO MONITORAMENTO DO CONSUMO DE
ÁGUA**

JOINVILLE

2022

ALLYSON BELLI BOGO

**GRÁFICOS DE CONTROLE ESTATÍSTICO DE PROCESSO PARAMÉTRICOS E
NÃO PARAMÉTRICOS APLICADOS AO MONITORAMENTO DO CONSUMO DE
ÁGUA**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Civil da Universidade do Estado de Santa Catarina, como requisito parcial para obtenção do título de Mestre em Engenharia Civil.

Orientadora: Dra. Elisa Henning
Coorientadora: Dra. Andreza Kalbusch

JOINVILLE

2022

**Ficha catalográfica elaborada pelo programa de geração automática da
Biblioteca Setorial do CCT/UDESC,
com os dados fornecidos pelo(a) autor(a)**

Bogo, Allyson Belli

GRÁFICOS DE CONTROLE ESTATÍSTICO DE PROCESSO
PARAMÉTRICOS E NÃO PARAMÉTRICOS APLICADOS AO
MONITORAMENTO DO CONSUMO DE ÁGUA / Allyson Belli

Bogo. -- 2022.

109 p.

Orientadora: Elisa Henning

Coorientadora: Andreza Kalbusch

Dissertação (mestrado) -- Universidade do Estado de Santa
Catarina, Centro de Ciências Tecnológicas, Programa de
Pós-Graduação em Engenharia Civil, Joinville, 2022.

1. Consumo de água. 2. Gráficos de controle estatístico de processo.
3. Detecção de vazamentos. I. Henning, Elisa. II. Kalbusch, Andreza. III.
Universidade do Estado de Santa Catarina, Centro de Ciências
Tecnológicas, Programa de Pós-Graduação em Engenharia Civil. IV.
Título.

ALLYSON BELLI BOGO

**GRÁFICOS DE CONTROLE ESTATÍSTICO DE PROCESSO PARAMÉTRICOS E
NÃO PARAMÉTRICOS APLICADOS AO MONITORAMENTO DO CONSUMO DE
ÁGUA**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Civil da Universidade do Estado de Santa Catarina, como requisito parcial para obtenção do título de Mestre em Engenharia Civil.

Orientadora: Dra. Elisa Henning

Coorientadora: Dra. Andreza Kalbusch

BANCA EXAMINADORA

Dra. Elisa Henning

Universidade do Estado de Santa Catarina (UDESC)

Membros:

Dra. Luciana Rosa Leite

Universidade do Estado de Santa Catarina (UDESC)

Dr. Daniel Richard Sant'Ana

Universidade de Brasília (UnB)

Joinville, 25 de agosto de 2022.

“Imagine uma nova história para a sua vida, e acredite nela.” (Paulo Coelho)

RESUMO

A adoção de estratégias para o monitoramento do consumo de água é essencial para a manutenção e promoção da sustentabilidade dos recursos hídricos. Na escala urbana, as técnicas de monitoramento podem auxiliar na gestão dos sistemas de distribuição, e na escala edificada, contribuir para o desenvolvimento de práticas sustentáveis de consumo de água. A utilização de gráficos de controle para o monitoramento do consumo de água apresenta vantagens relacionadas à sua aplicação e interpretação, podendo fornecer uma ampla gama de benefícios, como a detecção de vazamentos de água e a identificação de horários de maior consumo. Diante disso, esta pesquisa realizou um estudo comparativo entre a aplicação de gráficos de controle paramétricos e não paramétricos para o monitoramento de séries temporais do consumo diário de água no município de Joinville, na região Sul do Brasil. O procedimento metodológico incluiu a investigação do consumo diário de água do município (em m³/dia), classificando-o conforme as categorias residencial, comercial, industrial e público, com a utilização dos gráficos de controle de Shewhart e EWMA, além da alternativa não paramétrica EWMA-SN. O consumo de água de dois conjuntos habitacionais também foi investigado para avaliar a eficácia dos gráficos de controle na detecção de vazamentos de água na escala edificada. Os conjuntos de dados, obtidos por meio de um sistema de medição de telemetria da concessionária de água do município, representam um período de 243 dias. Os resultados mostram que os gráficos de controle são uma ferramenta poderosa na identificação de mudanças nos padrões do consumo de água. Em relação à eficácia dos gráficos de controle na detecção de vazamentos em ambas as escalas municipal e edificada, os tradicionais gráficos de Shewhart e EWMA apresentaram o melhor desempenho, detectando simultaneamente os vazamentos de água. O gráfico EWMA apresentou os melhores resultados no monitoramento do padrão da demanda urbana de água, tornando mais clara a visualização de tendências no consumo diário de água. A metodologia utilizada nesta pesquisa pode auxiliar no desenvolvimento de ferramentas automatizadas para o monitoramento do consumo de água urbano ou de um edifício.

Palavras-chave: Consumo de água; Gráficos de controle estatístico de processo; Detecção de vazamentos.

ABSTRACT

The adoption of strategies for monitoring water consumption is essential for maintaining and promoting the sustainability of water resources. At the urban scale, monitoring techniques can help in the management of distribution systems, and at the built scale, contribute to the development of sustainable water consumption practices. The use of control charts for water consumption monitoring has advantages related to its application and interpretation, and can provide a wide range of benefits, such as the detection of water leaks and the identification of peak demand periods. That said, this paper conducts a comparative study between parametric and non-parametric control charts for the monitoring of daily water consumption time series in the city of Joinville, southern Brazil. The methodological procedure includes the investigation of the city's daily water consumption (in m^3/day), classifying it according to residential, commercial, industrial and public categories, using the Shewhart and the EWMA control charts in addition to the non-parametric alternative, the EWMA-SN. Water consumption of one housing compound was also investigated to assess the effectiveness of the control charts in detecting leaks at a built scale. The data sets, obtained through a telemetry metering system from the city's water utility, represent a period of 243 days. The results show that control charts are a powerful tool in identifying changes in water consumption patterns. Regarding the effectiveness of the control charts in detecting leaks at both municipal and residential scales, the traditional Shewhart and EWMA charts performed the best, simultaneously detecting the water leaks. The EWMA chart yielded the best results in monitoring the pattern of urban water consumption, making the visualization of trends in daily water consumption clearer. The methodology used in this research can help in the development of automated monitoring tools for the water consumption in a city or a building.

Keywords: Water consumption; Statistical process control charts; Leak detection.

LISTA DE ILUSTRAÇÕES

Figura 1 – Gráfico de controle típico	14
Figura 2 – Gráfico de controle $\bar{x}$ com comportamento não aleatório	24
Figura 3 – Regras da <i>Western Electric</i>	26
Figura 4 – Função de autocorrelação das amostras	29
Figura 5 – Quantidade e ano de publicação dos estudos	41
Figura 6 – Distribuição geográfica das publicações	42
Figura 7 – Método de obtenção dos dados por ano	43
Figura 8 – Obtenção dos dados por escala espacial.....	44
Figura 9 – Métodos utilizados nas pesquisas a cada ano	45
Figura 10 – Métodos utilizados por escala espacial	46
Figura 11 – Fluxograma dos procedimentos metodológicos conduzidos nesta pesquisa	51
Figura 12 – <i>Boxplots</i> dos consumos de água dos conjuntos habitacionais	60
Figura 13 – Série temporal do consumo de água do Condomínio A	61
Figura 14 – Série temporal do consumo de água do Condomínio B	62
Figura 15 – Funções de autocorrelação para os conjuntos habitacionais	63
Figura 16 – Gráficos de controle para o Condomínio A	65
Figura 17 – Gráficos de controle para o Condomínio B	67
Figura 18 – Quantidade de matrículas e dados de consumo de cada categoria.....	70
Figura 19 – <i>Boxplots</i> da demanda urbana de água	70
Figura 20 – Séries temporais da demanda urbana de água.....	72
Figura 21 – Funções de autocorrelação para a demanda urbana de água.....	74
Figura 22 – Gráficos de controle para o consumo residencial.....	75
Figura 23 – Gráficos de controle para o consumo comercial	77
Figura 24 – Gráficos de controle para o consumo industrial	78
Figura 25 – Gráficos de controle para o consumo público	79
Figura 26 – Gráficos de controle para o consumo total.....	81

LISTA DE TABELAS

Tabela 1 – Estatísticas descritivas dos conjuntos habitacionais	59
Tabela 2 – Resultados dos testes de Shapiro-Wilk para os conjuntos habitacionais	63
Tabela 3 – Estatísticas descritivas da demanda urbana de água.....	69
Tabela 4 – Resultados dos testes de Shapiro-Wilk para a demanda urbana de água.....	73

SUMÁRIO

1	INTRODUÇÃO.....	7
1.1	OBJETIVOS	12
1.1.1	Geral.....	12
1.1.2	Específicos.....	12
1.2	ESTRUTURA DO TRABALHO	12
2	REFERENCIAL TEÓRICO.....	14
2.1	FUNDAMENTOS DOS GRÁFICOS DE CONTROLE.....	14
2.1.1	Tamanho da amostra e frequência de amostragem	15
2.1.2	Fases I e II da aplicação dos gráficos de controle	15
2.1.3	Comprimento médio de sequência	16
2.2	GRÁFICOS DE CONTROLE DE SHEWHART.....	18
2.2.1	Base estatística dos gráficos de Shewhart	19
2.2.2	Gráficos de controle baseados em valores de referência	20
2.2.3	Gráficos de controle para $\bar{x}$ e R	21
2.2.4	Gráficos de controle para $\bar{x}$ e s	22
2.2.5	Gráficos de controle para medidas individuais	23
2.2.6	Análise de padrões dos gráficos de Shewhart.....	24
2.3	PREMISSAS PARA A APLICAÇÃO DOS GRÁFICOS DE SHEWHART	26
2.3.1	Aderência à distribuição normal	26
2.3.2	Independência das amostras	28
2.4	GRÁFICO DE CONTROLE EWMA.....	29
2.4.1	Base estatística do gráfico EWMA	30
2.4.2	Parâmetros estatísticos do gráfico EWMA.....	31
2.4.3	Cálculo do ARL para o gráfico EWMA	32
2.5	ALTERNATIVAS NÃO PARAMÉTRICAS AOS GRÁFICOS DE CONTROLE	33
2.6	GRÁFICO DE CONTROLE NÃO PARAMÉTRICO EWMA-SN.....	35
2.6.1	Base estatística do gráfico EWMA-SN.....	35
2.6.2	Parâmetros estatísticos do gráfico EWMA-SN	36
2.6.3	Cálculo do ARL para o gráfico EWMA-SN	37
2.7	TÉCNICAS DE MONITORAMENTO DE ÁGUA	38
2.7.1	Revisão de estudos sobre o monitoramento de água.....	38
2.7.2	Síntese dos estudos revisados	41

2.7.3	Vantagens e limitações identificadas nos estudos	46
3	METODOLOGIA	49
3.1	AQUISIÇÃO E FORMATAÇÃO DOS DADOS	50
3.2	ANÁLISE EXPLORATÓRIA DOS DADOS	53
3.3	APLICAÇÃO DOS GRÁFICOS DE CONTROLE	54
3.3.1	Avaliação das premissas para definição da fase I	54
3.3.2	Fase II de aplicação dos gráficos de controle	56
3.4	ANÁLISE COMPARATIVA	56
4	RESULTADOS E DISCUSSÃO	58
4.1	CONJUNTOS HABITACIONAIS	58
4.1.1	Análise das séries temporais	59
4.1.2	Avaliação das premissas	62
4.1.3	Aplicação dos gráficos de controle	64
4.1.3.1	Gráficos de controle para o Condomínio A	64
4.1.3.2	Gráficos de controle para o Condomínio B	66
4.1.4	Análise comparativa	67
4.2	DEMANDA URBANA DE ÁGUA	68
4.2.1	Análise das séries temporais	69
4.2.2	Avaliação das premissas	73
4.2.3	Aplicação dos gráficos de controle	74
4.2.3.1	Gráficos de controle para o consumo residencial	75
4.2.3.2	Gráficos de controle para o consumo comercial	76
4.2.3.3	Gráficos de controle para o consumo industrial	77
4.2.3.4	Gráficos de controle para o consumo público	79
4.2.3.5	Gráficos de controle para o consumo total	80
4.2.4	Análise comparativa	82
5	CONCLUSÃO	84
5.1	SUGESTÕES PARA TRABALHOS FUTUROS	86
	REFERÊNCIAS BIBLIOGRÁFICAS	87
	APÊNDICE A – LISTA DE ARTIGOS CONSIDERADOS NA REVISÃO	97
	ANEXO A – FATORES PARA A CONSTRUÇÃO DE GRÁFICOS DE CONTROLE PARA VARIÁVEIS	103

1 INTRODUÇÃO

A disponibilidade de recursos hídricos é um fator essencial tanto para o desenvolvimento socioeconômico como ambiental de uma nação (WEI *et al.*, 2018). Conforme os autores destacam, o mundo vem enfrentando problemas crescentes em relação à prestação de serviços de água, principalmente em países em desenvolvimento. Embora o crescimento populacional seja um dos principais fatores responsáveis pela escassez de água, a melhora nos padrões de vida, a urbanização, as condições climáticas e uma demanda maior e mais desigual também aumentam a pressão sobre este recurso e representam um desafio claro no fornecimento de serviços de água satisfatórios (BATES *et al.*, 2008).

Atualmente, os processos de urbanização e industrialização tornam mais evidente a discrepância entre o aumento do consumo e a consequente escassez de água (WEI *et al.*, 2018), de modo que alcançar o desenvolvimento sustentável dos recursos hídricos é uma questão de importância global, especialmente nos países em desenvolvimento, onde as questões relacionadas ao consumo de água são um desafio crescente em termos de sustentabilidade (WANG; WANG, 2020). Em 2010, a demanda mundial por água era de aproximadamente 4.600 km³ por ano, e estima-se que até o ano de 2050 esse consumo aumentará em 30%, atingindo o patamar de 6.000 km³ por ano (BUREK *et al.*, 2016), dificultando cada vez mais a manutenção do equilíbrio entre a demanda e a oferta de água (ALI; SHAFIEE; BERGLUND, 2017). Segundo dados da Agência Nacional de Águas (ANA), o Brasil já enfrenta uma situação crítica em termos de disponibilidade hídrica. O número de pessoas afetadas pela seca aumentou de 9,6 milhões em 2015 para 37,9 milhões em 2017 e um dos efeitos da seca é o racionamento de água em muitas partes do país (ANA, 2018). De acordo com a Associação Brasileira de Engenharia Sanitária e Ambiental (ABES, 2013), grande parte do problema do desperdício de água está concentrada nas perdas na distribuição. No Brasil, estima-se que este fator atinja 40% do total de água tratada (ABES, 2013).

As fontes primárias de perdas de água nos sistemas de distribuição estão associadas a vazamentos e rompimentos na rede e podem surgir de uma série de mecanismos, incluindo corrosão do material, fadiga associada a variações na pressão da água, carga superficial excessiva, falhas estruturais causadas por movimentos ou erosão do solo, entre muitos outros (ZAN *et al.*, 2014). Ainda de acordo com os autores, o vazamento nas tubulações de água impõe vários dilemas associados, como perda financeira, problemas de qualidade da água e consequentes danos à saúde e segurança pública. Entre esses problemas, os custos financeiros associados às perdas de água estão relacionados ao valor de produção da água não paga, custos

de reparo, custos associados à interrupção do serviço e custos intangíveis, como insatisfação do cliente e diminuição da qualidade da água e da segurança pública (ZAN *et al.*, 2014). Portanto, conforme os autores destacam, a detecção e localização de vazamentos tornam-se atividades imperativas para as autoridades competentes no fornecimento de serviços de água.

Em pesquisa realizada com aproximadamente 180 concessionárias de água distribuídas em 14 países, a necessidade de serviços de detecção de vazamentos foi reconhecida pela maioria das empresas, contudo, apenas 40% relataram possuir métodos de controle de vazamentos (MOSER; PAAL; SMITH, 2015). De acordo com os autores, a maioria das concessionárias reage aos vazamentos respondendo a rompimentos óbvios e reparando a infraestrutura conforme necessário, sendo que vazamentos menores ou não informados demoram mais tempo para serem corrigidos, evidenciando a necessidade da adoção de estratégias mais racionais e sistemáticas para o bom gerenciamento da infraestrutura.

As abordagens de detecção de vazamentos podem ser amplamente classificadas como baseadas em modelos ou em dados reais do processo (LI *et al.*, 2015). As técnicas baseadas em modelos são limitadas ao lidar com sistemas de grande escala, em virtude da complexidade característica dos processos (COLOMBO; LEE; KARNEY, 2009), e as abordagens baseadas em dados representam alternativas atraentes, podendo obter resultados mais precisos (ROMANO; KAPELAN; SAVIĆ, 2011; WU; LIU, 2017). Muitos métodos de inspeção, incluindo eletromagnético, acústico, ultrassônico, radiográfico e termográfico, foram propostos para detectar vazamentos, porém eles não possibilitam um monitoramento contínuo pelas empresas, visto que requerem o acompanhamento de uma equipe especializada para a operação dos equipamentos e interpretação das medições (LIU; KLEINER, 2013).

Diante deste cenário, a utilização de métodos computacionais para o monitoramento do nexos água-energia está recebendo cada vez mais atenção (CHINI; STILLWELL, 2019). Motivados pela transformação digital destas indústrias, a maioria dos estudos sobre o perfil da demanda de água foram inspirados por avanços anteriores no setor elétrico (DI MAURO *et al.*, 2021). Os autores relatam que, com um desenvolvimento mais avançado e consolidado de tecnologias de medição inteligente e *Internet das Coisas* (do inglês, *Internet of Things* - IoT) no setor elétrico, conjuntos de dados do consumo de eletricidade tornaram-se disponíveis mais cedo do que conjuntos de dados semelhantes no setor de água. Ainda de acordo com os autores, os desenvolvimentos de medidores inteligentes de água e eletricidade avançaram em velocidades diferentes, devido a presença de algumas diferenças tecnológicas como a necessidade de baterias nos hidrômetros inteligentes, que limitam sua vida operacional e

frequência de transmissão de dados, enquanto medidores de eletricidade são alimentados por uma fonte de energia por concepção (DI MAURO *et al.*, 2021).

Apesar destas disparidades, as técnicas de monitoramento nos dois setores apresentam algumas semelhanças: enquanto a modelagem de redes de distribuição de energia tem a finalidade de ajudar os serviços públicos e as autoridades na gestão da rede elétrica, a medição inteligente de eletricidade no ambiente construído visa melhorar a consciência dos usuários e promover comportamentos sustentáveis e possibilidades de economia de energia (BARAI; KRISHNAN; VENKATESH, 2015). Da mesma forma, no setor hídrico, o objetivo principal da modelagem de redes de distribuição de água é auxiliar na gestão das redes, e o da medição inteligente, é promover o uso racional de água (VOULIS; WARNIER; BRAZIER, 2018; WANG *et al.*, 2019).

Além disso, a escala temporal tem grande influência na etapa de coleta de dados nos dois setores (ALAHAKOON; YU, 2016). Conforme os autores destacam, geralmente são utilizados dados diários ou mensais para a modelagem de redes de distribuição, enquanto o consumo sub diário é adotado para a medição inteligente em edificações. Nessa escala fina, os dados de consumo de água e eletricidade são utilizados para aumentar a eficiência do comportamento dos consumidores, melhorar a previsão da demanda, promover oportunidades de economia, investigar diferentes segmentos de clientes e, potencialmente, auxiliar na gestão financeira das empresas do setor (HAN; ANG, 2018).

A maior parte das pesquisas de análise dos padrões de consumo de água são voltadas à escala doméstica ou de usos finais (FUENTES; MAURICIO, 2020; GAUTAM *et al.*, 2020; RIZZO *et al.*, 2021), uma vez que o acesso aos dados de centros urbanos é geralmente restrito e limitado pelas medidas de proteção da privacidade dos consumidores, seja pela anonimização dos dados, restrição de acesso ou existência de filtros de controle de acesso (DI MAURO *et al.*, 2021). Poucos estudos se concentram no processamento e análise de dados de consumo de água por telemetria para melhorar a operação e manutenção de redes de distribuição de água (LOUREIRO *et al.*, 2014).

Contudo, impulsionadas pelo recente avanço tecnológico em sistemas de telemetria e equipamentos medidores de água, pesquisas envolvendo o monitoramento das redes de distribuição de água vêm gradativamente ocupando mais espaço (ANTZOULATOS *et al.*, 2020; NARAYANAN; SANKARANARAYANAN, 2020; CASSIDY *et al.*, 2021). Esses estudos têm o objetivo de auxiliar as concessionárias de água na identificação de vazamentos e, conseqüentemente, reduzir as taxas de desperdício, dando um passo à frente no que tange ao controle de perdas de água em redes de distribuição.

Neste contexto, o uso de técnicas estatísticas tem ganhado destaque por apresentarem vantagens relacionadas à sua aplicação contínua e simplicidade na leitura e interpretação dos dados (JUNG; LANSEY, 2015; BORGES; JUNG; KIM, 2017; AHN; JUNG, 2019). Métodos estatísticos apresentam ferramentas importantes no diagnóstico e otimização da gestão e operação de vários sistemas, desde humanos até os mais complexos sistemas físicos (MONTGOMERY, 2016). Conforme o autor destaca, a aplicação dessas técnicas tem desempenhado um papel importante na solução de problemas tanto da indústria como relacionados a serviços de saúde, sendo empregado em bancos, contabilidade, marketing e diversos outros setores.

Os gráficos de controle são considerados uma ferramenta importante do Controle Estatístico de Processos (CEP), tendo sido desenvolvidos na década de 1920 por Walter A. Shewhart (MONTGOMERY, 2016). Esse método apresenta a vantagem de identificar situações em que um processo possa estar se comportando de maneira irregular, permitindo que uma investigação seja iniciada para descobrir a causa dessa mudança no processo e possam ser fornecidas ações corretivas (DEVORE, 2016). Até o final do século XX, o paradigma tradicional era o processo industrial, de maneira que os gráficos de controle eram direcionados exclusivamente para o monitoramento de linhas de produção, potencializando a produtividade e lucratividade de uma empresa, porém esta técnica já se estendeu a diversos outros campos de atuação (SAMOHYL, 2009).

Na área da saúde, o uso de CEP teve um crescimento expressivo recentemente, com a aplicação dos gráficos de controle nas mais diversas especialidades, como epidemiologia e controle de infecções (GOMES; OLIVEIRA; MINGOTI, 2011; MIHALOVITS; KEMÉNY, 2020; OKORO *et al.*, 2020; PERLA *et al.*, 2020; ROBERTS *et al.*, 2020). O uso de gráficos de controle também tem sido aplicado com sucesso no monitoramento ambiental, sendo amplamente utilizado para auxiliar no gerenciamento da qualidade da água para consumo, por exemplo (HENNING *et al.*, 2014a; HENNING *et al.*, 2014b; ELEVLI *et al.*, 2016; IGLESIAS *et al.*, 2016; AHMED *et al.*, 2020; SANUSI *et al.*, 2020).

No ambiente construído, o monitoramento do consumo de água utilizando técnicas de CEP apresenta um alto grau de consistência e desempenho, tendo potencial de estimular o uso racional e a gestão preventiva de redes de abastecimento de água (JUNG *et al.*, 2015; ROMANO; WOODWARD; KAPELAN, 2017; FREITAS *et al.*, 2019). Contudo, a aplicação de gráficos de controle nessa área é escassa e, nesta ótica, acredita-se ser possível o avanço.

O monitoramento do consumo de água em uma edificação por meio da utilização de gráficos de controle se mostrou vantajoso para associar a mudança dos padrões de usos com

certos eventos, como vazamentos, cortes no fornecimento de água e mudanças no cronograma de atividades (FREITAS *et al.*, 2019). Romano, Woodward e Kapelan (2017) apresentaram uma metodologia baseada em CEP para determinar a localização aproximada de vazamentos de água em redes de distribuição, possibilitando a determinação da área onde um vazamento ocorreu e, conseqüentemente, reduzindo substancialmente os esforços necessários e custos operacionais na identificação de um vazamento.

Por se apresentar como uma técnica eficaz para avaliar mudanças e identificar quando elas ocorreram, a aplicação de gráficos de controle para o monitoramento do consumo de água possibilita determinar se alterações na legislação ou em campanhas educacionais tiveram impacto na redução do desperdício de água, investigando também os efeitos da economia do uso de água, bem como identificar vazamentos em uma rede de distribuição. Nesse contexto, justifica-se o desenvolvimento de novas metodologias capazes de monitorar o consumo de água por meio do uso de gráficos de controle.

Em relação às características de aplicação dos gráficos de controle, existem duas possibilidades de categorização dos mesmos, podendo ser classificados como paramétricos ou não paramétricos. A principal característica da primeira classe está relacionada ao comportamento estatístico do processo sendo monitorado, sendo que somente quando os dados atendem as premissas de normalidade e independência das observações, é possível obter conclusões quanto ao estado estatístico do processo (MONTGOMERY, 2016). A segunda classe surgiu como uma alternativa à exigência dessas premissas, de modo que a aplicação dos gráficos de controle não paramétricos não depende de qualquer suposição específica sobre a distribuição dos dados (CHAKRABORTI; VAN DER LAAN; BAKIR, 2001).

A aplicação de gráficos de controle paramétricos para o monitoramento do consumo de água apresenta resultados positivos (JUNG *et al.*, 2015; FREITAS *et al.*, 2019). Contudo, a demanda de água está muito associada a variáveis meteorológicas e climáticas, de modo que mudanças no clima e no nível de precipitação influenciam significativamente o padrão do consumo de água ao longo do tempo (BALLING JR.; GOBER, 2007; BALLING JR.; GOBER; JONES, 2008; CHANG; PRASKIEVICZ; PARANDVASH, 2014). Dessa forma, devido à dispersão dos valores, muitas vezes os dados de consumo de água não apresentam distribuição normal, exigida para a aplicação dos gráficos de controle paramétricos, sendo necessária a utilização de técnicas específicas para trazer o conjunto de dados para a normalidade (GARCIA, 2020). Outra possibilidade seria a aplicação de gráficos de controle não paramétricos (CHAKRABORTI; VAN DER LAAN; BAKIR, 2001), ainda pouco utilizados para esse propósito.

Neste contexto, a proposta desta pesquisa consiste em investigar o desempenho de gráficos de controle não paramétricos para o monitoramento do consumo de água, comparando-os com o desempenho de gráficos de controle paramétricos. O diferencial desta pesquisa consiste no fato de existirem poucos estudos envolvendo a aplicação de gráficos de controle não paramétricos para este propósito, o que pode representar uma opção adicional para o monitoramento e gestão dos recursos hídricos.

1.1 OBJETIVOS

1.1.1 Geral

O objetivo principal desta pesquisa consiste em realizar um estudo comparativo entre a aplicação de gráficos de controle paramétricos e não paramétricos para o monitoramento de séries temporais do consumo diário de água no município de Joinville, em Santa Catarina.

1.1.2 Específicos

Os objetivos específicos foram elaborados como meio para alcançar o objetivo geral e fomentar o desenvolvimento do estudo, e são especificados a seguir:

- i. analisar séries temporais do consumo diário de água de dois condomínios residenciais e do município de Joinville, em Santa Catarina, dividindo-o nas categorias residencial, comercial, industrial e público;
- ii. aplicar os gráficos de controle às séries temporais e avaliar o desempenho destas técnicas na detecção de vazamentos e identificação de mudanças no processo de consumo de água;
- iii. comparar o desempenho dos gráficos de controle paramétricos e não paramétricos para o monitoramento das séries temporais, identificando qual mecanismo obteve melhores resultados em cada análise.

1.2 ESTRUTURA DO TRABALHO

Este trabalho teve seu desenvolvimento estruturado em cinco capítulos. O presente capítulo introduz o assunto, apresentando sua motivação, justificativa e importância, além de determinar os objetivos geral e específicos.

O segundo capítulo aborda o referencial teórico, em que é apresentada uma revisão sobre os fundamentos para construção dos gráficos de controle, com ênfase nos gráficos de controle de Shewhart e *Exponentially Weighted Moving Average* (EWMA, sigla em inglês para Média Móvel Exponencialmente Ponderada). Também é desenvolvida a temática sobre gráficos de controle não paramétricos, apresentando o gráfico de sinais EWMA-SN. A última seção deste capítulo apresenta uma revisão narrativa de pesquisas acerca do monitoramento de água, com o objetivo de sintetizar os equipamentos e métodos mais utilizados para este propósito.

O terceiro capítulo apresenta a metodologia utilizada nesta pesquisa, identificando os passos metodológicos conduzidos para a sua realização. São apresentadas a forma de aquisição dos dados e a aplicação destes nos gráficos de controle revisados no referencial teórico. É abordado o procedimento de construção dos gráficos de controle, detalhando como foram realizadas as fases de diagnóstico e monitoramento.

No quarto capítulo são discutidos os principais passos para o desenvolvimento do estudo e os resultados obtidos. A metodologia proposta foi aplicada aos conjuntos de dados, de modo a verificar e comparar a eficácia dos gráficos de controle paramétricos e não paramétricos para o alcance dos objetivos específicos.

Por fim, o quinto capítulo compreende a conclusão e considerações finais da pesquisa, apresentando o desempenho dos gráficos de controle em cada análise. Nesse capítulo são relatadas ainda as limitações da pesquisa e sugestões para trabalhos futuros.

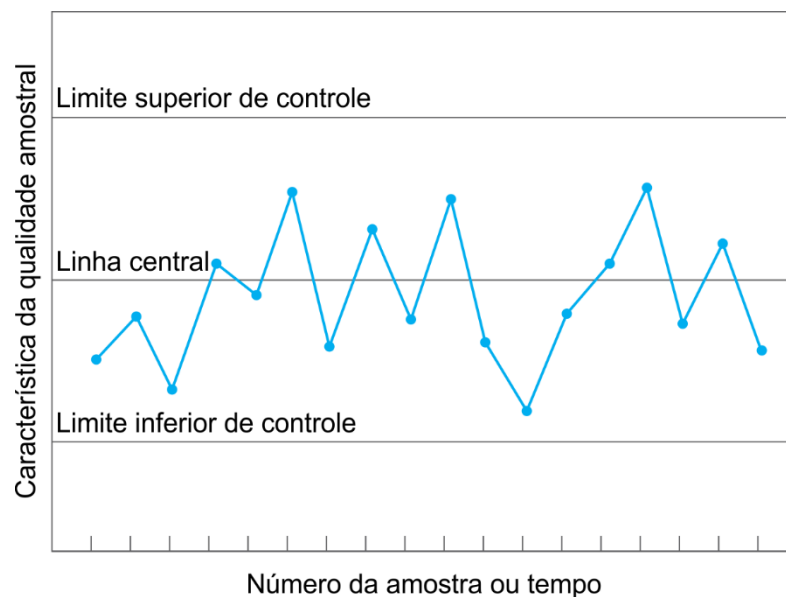
2 REFERENCIAL TEÓRICO

2.1 FUNDAMENTOS DOS GRÁFICOS DE CONTROLE

O gráfico de controle é um mecanismo poderoso dentro do CEP, o qual é composto por um conjunto de ferramentas de resolução de problemas cujo objetivo é medir, monitorar, controlar e melhorar a qualidade de um processo, examinando a variabilidade dos dados em séries temporais (MONTGOMERY; RUNGER, 2016).

Gráficos de controle possibilitam a visualização da presença de alterações no processo e consistem basicamente na representação de três linhas principais, denominadas de Linha Central (LC), que é comumente definida a partir do valor médio das amostras, Limite Superior de Controle (LSC) e Limite Inferior de Controle (LIC) (Figura 1). Quando há a presença de pontos além dos limites de controle, diz-se que o processo pode estar fora de controle, ou seja, é necessário investigar as causas responsáveis por este evento (SAMOHYL, 2009).

Figura 1 – Gráfico de controle típico



Fonte: Montgomery (2016)

Existe uma relação muito próxima entre os gráficos de controle e o teste de hipóteses. De certa maneira, o gráfico de controle testa a hipótese de que o processo esteja em um determinado estado de controle estatístico repetidamente. Pontos localizados dentro dos limites de controle equivalem a não rejeição da hipótese nula de que processo está sob controle, e pontos fora dos limites de controle indicam a rejeição dessa hipótese (MONTGOMERY, 2016).

Os limites são calculados de maneira a balancear dois importantes riscos associados ao uso dos gráficos de controle: o risco de declarar um ponto como fora de controle estatístico quando não houve mudança no processo, chamado de erro tipo I e representado por α , e o risco de não detectar uma mudança no processo quando ele está fora de controle, denominado erro tipo II e representado por β . Limites mais estreitos têm maior probabilidade de disparar alarmes falsos e aumentar o número de erros do tipo I, enquanto que limites mais amplos tendem a atrasar a detecção de situações que estejam realmente fora de controle estatístico, aumentando o número de erros tipo II (COSTA; EPPRECHT; CARPINETTI, 2016).

2.1.1 Tamanho da amostra e frequência de amostragem

No planejamento de um gráfico de controle, é necessário especificar tanto o tamanho da amostra a ser usada quanto a frequência de amostragem. Em geral, amostras maiores potencializarão a detecção de pequenas mudanças no processo. Na escolha do tamanho da amostra, deve-se ter em mente a magnitude da mudança que será detectada. Se a mudança no processo for relativamente grande, tamanhos de amostra menores terão melhores resultados, assim como o contrário (MONTGOMERY, 2016).

Em relação a frequência de amostragem e do ponto de vista de detecção de mudanças, a situação mais desejável seria a obtenção de grandes amostras muito frequentemente, no entanto, nem sempre esta opção é viável. O que ocorre normalmente é a coleta de pequenas amostras a intervalos curtos ou de amostras maiores a intervalos mais longos (ROSA, 2015).

Com o desenvolvimento da tecnologia de sensores e medidores automáticos, é possível aumentar a frequência de amostragem. Sistemas de medição automática e microcomputadores que aplicam programas de CEP são ferramentas cada vez mais eficientes para um monitoramento do processo online e em tempo real (DI MAURO *et al.*, 2021).

2.1.2 Fases I e II da aplicação dos gráficos de controle

Em geral, o processo de monitoramento estatístico envolve duas fases. Na fase I, é realizada uma análise retrospectiva e coletado um conjunto de dados do processo, são então construídos limites de controle de teste para determinar se o processo estava sob controle durante o período de coleta dos dados, verificando-se também se podem ser estabelecidos limites de controle confiáveis para o monitoramento do processo futuro (RAMOS; ALMEIDA; ARAÚJO, 2013).

A fase II se inicia após a obtenção de um conjunto de dados sob condições estáveis e representativas do processo sob controle. Nesta fase, o gráfico de controle é utilizado para monitoramento do processo, comparando a estatística amostral com cada observação subsequente obtida, à medida que ela é extraída do processo (MONTGOMERY, 2016).

Os gráficos de controle de Shewhart são extremamente úteis na fase I de implementação do CEP, quando o processo tem a possibilidade de estar fora de controle e suscetível a causas que resultem em parâmetros incorretos. Como estes parâmetros também podem fornecer orientação a respeito da causa atribuível à fuga do controle, este gráfico tem a vantagem de conseguir trazer um processo fora de controle para o controle estatístico (COSTA; EPPRECHT; CARPINETTI, 2016).

Contudo, uma vez que a decisão dos gráficos de controle de Shewhart se baseia na análise do último ponto plotado, a desconsideração de informações obtidas anteriormente resulta em uma grande desvantagem, tornando estes gráficos relativamente insensíveis a mudanças pequenas no processo (ROSA, 2015).

Diante deste cenário, duas alternativas ao gráfico de controle de Shewhart podem ser utilizadas para monitorar mudanças pequenas no processo: o gráfico de controle *Cumulative Sum* (CUSUM, sigla em inglês para Soma Acumulada) e o gráfico de controle EWMA. Em conjunto, estes gráficos são chamados de gráficos de controle ponderados pelo tempo e apresentam mais chances de serem eficazes no monitoramento de mudanças pequenas no processo (MONTGOMERY, 2016).

2.1.3 Comprimento médio de sequência

Quando se avalia a eficiência de um gráfico de controle, é desejável que ele detecte a presença de uma mudança no processo logo que ela tenha ocorrido e que a sinalização de alarmes falsos seja baixa. Este desempenho é medido pelo Comprimento Médio de Sequência (do inglês, *Average Run Length* - ARL) o qual representa o número médio de amostras que devem ser coletadas até que um ponto obrigatoriamente indique uma condição de fora de controle, condição conhecida como erro tipo I (ROSA, 2015).

Quando o processo está sob controle estatístico, o ARL_0 indica o número médio de pontos até que ocorra a sinalização do primeiro alarme falso no gráfico de controle. Já o ARL_1 representa o número de amostras necessárias para a detecção de uma alteração no processo quando este já está fora de controle (COSTA; EPPRECHT; CARPINETTI, 2016).

Assim, o ideal é que o ARL_0 seja o maior possível, indicando que serão emitidos poucos alarmes falsos quando o processo está sob controle estatístico, e que o ARL_1 seja o menor valor possível, já que este valor está diretamente relacionado ao tempo necessário para a detecção de uma causa especial (MONTGOMERY, 2016). Dessa forma, a escolha dos limites de controle nos gráficos deve ser feita com base no ARL desejado e no nível de alarmes falsos aceitável (MARGAVIO *et al.*, 1995).

Para um gráfico de controle de Shewhart, o ARL_0 pode ser expresso como o inverso da probabilidade de um ponto estar fora de controle e é dado pela Equação (1) (MONTGOMERY, 2016).

$$ARL_0 = \frac{1}{\alpha} \quad (1)$$

sendo ARL_0 o número médio de observações até que um falso alarme seja emitido e α é a probabilidade de se cometer um erro tipo I. Para o gráfico $\bar{x}$ de Shewhart com os limites 3σ , toma-se $\alpha = 0,0027$, portanto, o $ARL \cong 370$ indica que, mesmo o processo estando sob controle, um sinal de fora de controle será emitido, em média, a cada 370 amostras (MONTGOMERY, 2016).

O ARL_1 pode ser escrito em função da probabilidade de erro tipo II, que é a probabilidade de o gráfico não detectar uma alteração na estatística monitorada quando o processo estiver fora de controle estatístico e é calculado pela Equação (2) (MONTGOMERY, 2016).

$$ARL_1 = \frac{1}{1 - \beta} \quad (2)$$

em que ARL_1 é o número médio de observações até que seja sinalizada uma situação de fora de controle e β é a probabilidade de ocorrer um erro tipo II, o qual representa a probabilidade de não detecção de um processo fora de controle.

As equações para cálculo do ARL supõem independência das amostras e aderência à distribuição normal, portanto elas somente são aplicáveis aos gráficos de controle do tipo Shewhart e não podem ser aplicadas aos gráficos ponderados pelo tempo, pois nesses gráficos os pontos plotados não são independentes. Nesses casos, a determinação exata do ARL nem sempre é possível, pelo fato das estatísticas de controle serem variáveis aleatórias dependentes

(ROSA, 2015). No entanto, existem alguns métodos numéricos que permitem otimizar este parâmetro, como a aproximação por meio de uma cadeia de Markov, proposta por Brook e Evans (1972) e as aproximações propostas por Siegmund (1985) e Hawkins (1992).

2.2 GRÁFICOS DE CONTROLE DE SHEWHART

Essencialmente, um gráfico de controle testa a hipótese de que uma observação é igual a um parâmetro de interesse repetidamente, em pontos diferentes no tempo, dentro da região crítica estabelecida para o teste de hipóteses. De modo geral, pode-se estabelecer um modelo universal para o gráfico de controle. Seja w uma estatística amostral que mede um parâmetro de interesse, μ_w a média e σ_w o seu desvio padrão. Então, a linha central, o limite superior de controle e o limite inferior de controle podem ser estimados pelas Equações (3), (4) e (5) (MONTGOMERY, 2016).

$$LSC = \mu_w + L\sigma_w \quad (3)$$

$$LC = \mu_w \quad (4)$$

$$LIC = \mu_w - L\sigma_w \quad (5)$$

sendo o fator L a distância dos limites de controle até a linha central, expressa em unidades de desvio-padrão.

Essa teoria geral dos gráficos de controle foi originalmente proposta por Walter S. Shewhart, e os gráficos de controle desenvolvidos com base nesses princípios são geralmente chamados de gráficos de controle de Shewhart. Para a maioria dos gráficos de controle de Shewhart os limites de controle são calculados considerando-se $L = 3$, o que representa um $ARL_0 \cong 370$ (MONTGOMERY, 2016). Ainda de acordo com o autor, essa definição também representa que 99,7% das amostras estarão dentro dos limites de controle, se houver somente causas comuns atuando sobre o sistema e a distribuição for normal.

Os gráficos de controle de Shewhart podem ser classificados sob dois aspectos gerais. Quando a característica a ser monitorada pode ser quantificada, ela é chamada de variável, nesse caso, os gráficos de controle para tais medidas são chamados de gráficos de controle para variáveis (LOUZADA *et al.*, 2013). Conforme os autores destacam, quando a característica do processo não é medida em uma escala quantitativa, designa-se cada observação como conforme ou não conforme, os gráficos de controle para tais indicadores são chamados de gráficos de controle para atributos.

O controle estatístico de um processo que possui observações quantificáveis pode ser obtido por meio da análise de uma estatística de posição (média ou mediana) ou de uma medida de variabilidade (desvio padrão ou amplitude) calculadas a partir de amostras retiradas do processo (SAMOHYL, 2009). Segundo Paladini (2019), a utilização de gráficos de controle para o monitoramento da variabilidade se justifica pelo fato de que nenhum processo produtivo pode garantir que dois produtos sejam exatamente iguais, ou seja, há variabilidade entre eles.

Para os gráficos de controle de variáveis, podem existir processos que analisam amostras contendo uma ou mais observações, nesse último caso, pode-se mensurar a média das amostras pelo gráfico $\bar{x}$ e a variabilidade amostral por meio dos gráficos para amplitude (R) ou desvio padrão (s). Quando o tamanho de amostra é igual a um, os gráficos que analisam o comportamento de tal processo são chamados de gráficos de controle para medidas individuais (ROSA, 2015).

2.2.1 Base estatística dos gráficos de Shewhart

Assumindo a suposição de que um conjunto de dados seja normalmente distribuído e tenha seus valores de média μ e desvio-padrão σ conhecidos, uma amostra x de tamanho n tem sua média amostral dada pela Equação (6) (MONTGOMERY, 2016).

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (6)$$

quando $\bar{x}$ apresenta uma distribuição normal e desvio padrão $\sigma_{\bar{x}} = \sigma / \sqrt{n}$, existe uma probabilidade de $1 - \alpha$ de a média das amostras estar entre os intervalos $\mu + Z_{\alpha/2} \sigma_{\bar{x}}$ e $\mu - Z_{\alpha/2} \sigma_{\bar{x}}$.

Dessa forma, quando μ e σ são conhecidos, estes intervalos podem ser usados como limites de controle para as médias amostrais. No entanto, geralmente os padrões μ e σ não são conhecidos. Logo, eles devem ser estimados a partir de amostras preliminares retiradas de quando o processo estava sob controle. Tais estimativas devem se basear em pelo menos 20 ou 25 amostras, cada uma com n observações do parâmetro de monitoramento (RAMOS; ALMEIDA; ARAÚJO, 2013).

2.2.2 Gráficos de controle baseados em valores de referência

Quando existem disponíveis valores de referência para a média μ e desvio-padrão σ do processo, tais valores podem ser utilizados na construção dos gráficos $\bar{x}$, R e s sem que seja necessário recorrer à análise de dados passados do processo sob controle. Assim, a linha central e os limites superior e inferior de controle do gráfico $\bar{x}$ podem ser estimados pelas Equações (7), (8) e (9) (MONTGOMERY, 2016).

$$LSC = \mu + A\sigma \quad (7)$$

$$LC = \mu \quad (8)$$

$$LIC = \mu - A\sigma \quad (9)$$

em que a constante A é calculada a partir do número de observações das amostras e encontra-se listada no ANEXO A (MONTGOMERY, 2016).

O gráfico R , que monitora a variabilidade do processo pela análise das amplitudes amostrais R , possui os parâmetros calculados pelas Equações (10), (11) e (12) (MONTGOMERY, 2016).

$$LSC = D_2\sigma \quad (10)$$

$$LC = d_2\sigma \quad (11)$$

$$LIC = D_1\sigma \quad (12)$$

Para o monitoramento do desvio padrão s pode ser aplicado o gráfico s , o qual possui seus parâmetros quantificados pelas Equações (13), (14) e (15) (MONTGOMERY, 2016).

$$LSC = B_6\sigma \quad (13)$$

$$LC = c_4\sigma \quad (14)$$

$$LIC = B_5\sigma \quad (15)$$

As constantes d_2 , D_1 , D_2 , c_4 , B_5 e B_6 também dependem da quantidade de n observações das amostras e encontram-se listadas no ANEXO A (MONTGOMERY, 2016).

2.2.3 Gráficos de controle para $\bar{x}$ e R

Quando o processo sendo analisado não possui valores de referência, se faz necessário estimar os valores de média μ e desvio padrão σ a partir de amostras ou subgrupos preliminares do processo sob controle. Considerando-se um conjunto de m amostras, cada uma com n observações da característica sendo monitorada, o melhor estimador de μ é a média geral, representada pela Equação (16) (MONTGOMERY, 2016).

$$\bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \cdots + \bar{x}_m}{m} \quad (16)$$

em que $\bar{x}_1 + \bar{x}_2 + \cdots + \bar{x}_m$ representam a média de cada uma das m amostras e $\bar{\bar{x}}$, representando a média geral, deve ser utilizada como a linha central no gráfico $\bar{x}$.

A amplitude R das amostras quantifica a diferença entre a maior e a menor observação dentro do subgrupo e é calculada por $R = x_{max} - x_{min}$. Seja uma amostra de tamanho m , sua amplitude média é dada pela Equação (17) (MONTGOMERY, 2016).

$$\bar{R} = \frac{R_1 + R_2 + \cdots + R_m}{m} \quad (17)$$

Assim, a linha central e os limites superior e inferior de controle do gráfico $\bar{x}$ considerando-se a amplitude R amostral podem ser estimados pelas Equações (18), (19) e (20) (MONTGOMERY, 2016).

$$LSC = \bar{\bar{x}} + A_2 \bar{R} \quad (18)$$

$$LC = \bar{\bar{x}} \quad (19)$$

$$LIC = \bar{\bar{x}} - A_2 \bar{R} \quad (20)$$

O gráfico R monitora a variabilidade do processo pela análise das amplitudes amostrais R . Nesse caso, a linha central e os limites de controle são calculados pelas Equações (21), (22) e (23), em que R é utilizado como estimativa da amplitude média e seu desvio padrão é dado pelas constantes D_3 e D_4 (MONTGOMERY, 2016).

$$LSC = D_4 \bar{R} \quad (21)$$

$$LC = \bar{R} \quad (22)$$

$$LIC = D_3 \bar{R} \quad (23)$$

As constantes A_2 , D_3 e D_4 são estimadores obtidos em função do tamanho n de observações de cada amostra e encontram-se listadas no ANEXO A (MONTGOMERY, 2016).

2.2.4 Gráficos de controle para $\bar{x}$ e s

A construção e operação dos gráficos de controle $\bar{x}$ e s requerem mais ou menos a mesma sequência de etapas que os gráficos $\bar{x}$ e R , porém neste caso são consideradas a média amostral $\bar{x}$ e o desvio padrão amostral s dos dados preliminares do processo sob controle.

Seja um conjunto de m amostras, cada uma com n observações, o desvio padrão de cada amostra é calculado pela Equação (24) (MONTGOMERY, 2016).

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (24)$$

em que s representa o desvio padrão amostral. Dessa forma, o desvio padrão médio de um conjunto de m amostras é dado pela Equação (25) (MONTGOMERY, 2016).

$$\bar{s} = \frac{1}{m} \sum_{i=1}^m s_i \quad (25)$$

O gráfico $\bar{x}$ para monitoramento da média com base no desvio padrão amostral possui a linha central e os limites superior e inferior de controle calculados pelas Equações (26), (27) e (28) (MONTGOMERY, 2016).

$$LSC = \bar{\bar{x}} + A_3 \sigma \quad (26)$$

$$LC = \bar{\bar{x}} \quad (27)$$

$$LIC = \bar{\bar{x}} - A_3 \sigma \quad (28)$$

Para monitoramento da variabilidade do processo por meio do desvio padrão s é utilizado o gráfico s . Nesse caso, a linha central e os limites de controle são obtidos pelas Equações (29), (30) e (31) (MONTGOMERY, 2016).

$$LSC = B_4 \bar{s} \quad (29)$$

$$LC = \bar{s} \quad (30)$$

$$LIC = B_3 \bar{s} \quad (31)$$

Os valores das constantes A_3 , B_3 e B_4 para a construção dos limites de controle dependem do valor de n observações das amostras e encontram-se listadas no ANEXO A (MONTGOMERY, 2016).

2.2.5 Gráficos de controle para medidas individuais

Em muitos processos o número de observações da amostra é $n = 1$, isso significa que a amostra consiste em uma única unidade individual. Nestes casos são empregados os gráficos de controle para medidas individuais. Geralmente, esses gráficos utilizam a amplitude móvel de duas observações consecutivas como base para estimar a variabilidade do processo. A amplitude móvel é definida como MR e é calculada por $MR_i = |x_i - x_{i-1}|$, sendo i um número inteiro que representa o ponto sendo observado (MONTGOMERY, 2016).

O gráfico de controle para medidas individuais para monitoramento da média possui a linha central e os limites superior e inferior de controle calculados pelas Equações (32), (33) e (34) (MONTGOMERY, 2016).

$$LSC = \bar{\bar{x}} + \frac{3\overline{MR}}{d_2} \quad (32)$$

$$LC = \bar{\bar{x}} \quad (33)$$

$$LIC = \bar{\bar{x}} - \frac{3\overline{MR}}{d_2} \quad (34)$$

Também é possível realizar o monitoramento da amplitude móvel para medições individuais. Nesse caso, a linha central e os limites de controle são obtidos pelas Equações (35), (36) e (37) (MONTGOMERY, 2016).

$$LSC = D_4 \overline{MR} \quad (35)$$

$$LC = \overline{MR} \quad (36)$$

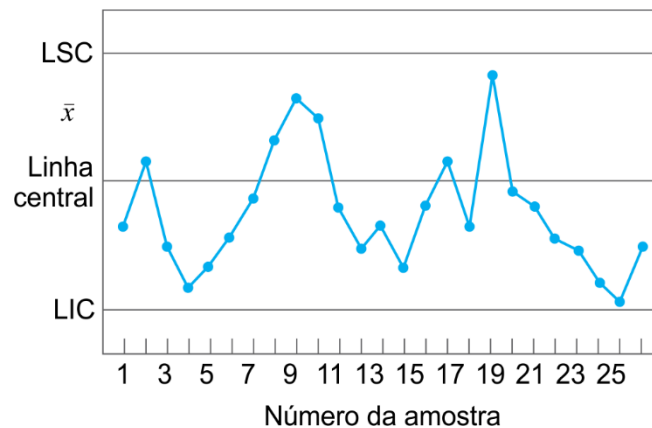
$$LIC = D_3 \overline{MR} \quad (37)$$

Quando o número de observações é $n = 2$, tem-se que $d_2 = 1,128$, $D_3 = 0$ e $D_4 = 3,267$, constantes que dependem do tamanho de n observações e encontram-se listadas no ANEXO A. Como o gráfico de controle para medidas individuais utiliza como base a análise de duas observações consecutivas, tais valores podem ser utilizados na construção dos limites de controle (MONTGOMERY, 2016). Embora seja comum a aplicação em conjunto com um gráfico para amplitude móvel (MR), para o monitoramento da variabilidade, alguns autores não o consideram efetivo para este fim (SMETI *et al.*, 2007). Dessa forma, este gráfico não é apresentado neste trabalho.

2.2.6 Análise de padrões dos gráficos de Shewhart

Um gráfico de controle pode indicar que determinado processo se encontra fora de controle quando um ou mais pontos estiver além dos limites de controle ou exibir algum padrão não aleatório de comportamento (MONTGOMERY, 2016), ilustrado na Figura 2. Especificamente para a figura ilustrada, 19 dos 25 pontos encontram-se abaixo da linha central, se a distribuição dos pontos fosse aleatória, seria esperado uma plotagem mais uniforme deles acima e abaixo da linha central (MONTGOMERY; RUNGER, 2016).

Figura 2 – Gráfico de controle $\bar{x}$ com comportamento não aleatório



Fonte: Montgomery (2016)

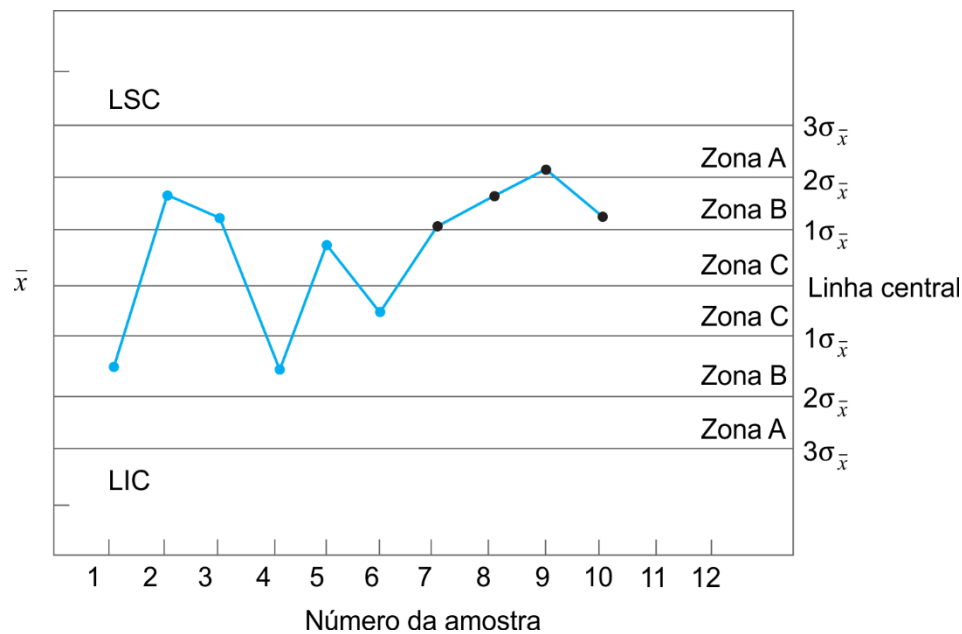
Outros fatores também são utilizados para essa análise, como por exemplo sequências de pontos simultâneos subindo ou descendo, ou a distribuição exibindo comportamento cíclico, mesmo que todos os pontos estejam dentro dos limites (MONTGOMERY, 2016). Para auxiliar na detecção de padrões não aleatórios de comportamento nos gráficos de controle, um conjunto de regras foi sugerido pela *Western Electric* (MONTGOMERY, 2016). O objetivo destas regras é detectar pequenas mudanças na média, uma vez que as mesmas poderiam resultar em pontos mais afastados do centro, porém ainda dentro dos limites de controle, não sendo detectados (WADSWORTH; STEPHENS; GODFREY, 2001). De acordo com essa metodologia, o processo é classificado como fora de controle se atender a um dos seguintes pontos:

- a) um ponto estiver fora dos limites de controle 3σ ;
- b) dois, em três pontos consecutivos, estiverem além dos limites de 2σ ;
- c) quatro, em cinco pontos consecutivos, estiverem a uma distância de 1σ ou além da central;
- d) oito pontos consecutivos estiverem de um mesmo lado da linha central.

Essas regras se aplicam a um lado da linha central de cada vez e são geralmente utilizadas para acentuar a sensibilidade dos gráficos de controle. Com o uso dessas regras pode-se permitir que mudanças menores no processo sejam detectadas mais rapidamente do que quando o único critério de monitoramento é a usual violação dos limites de controle 3σ , contudo a sua aplicação implica em um aumento na sinalização de causas especiais quando elas não existem, podendo chegar a 5% (VAUGHN, 1990).

Além destes princípios, a plotagem dos pontos de maneira cíclica também pode ser um sinal de um processo fora de controle. Nesse caso, a redução de algumas fontes de variação pode eliminar este comportamento, sendo que a interpretação destes comportamentos exige um conhecimento técnico tanto dos princípios estatísticos quanto do processo. Contudo, o uso destas regras aumenta a complexidade do uso e dificulta a interpretação dos gráficos de controle de Shewhart (MONTGOMERY, 2016).

A Figura 3 exemplifica a aplicação do critério *Western Electric* ao gráfico de controle de um determinado processo. Considerando esse critério, o processo está fora de controle, uma vez que os últimos quatro pontos consecutivos excedem o limite de 1σ , indicando um padrão não aleatório.

Figura 3 – Regras da *Western Electric*

Fonte: Montgomery (2016)

2.3 PREMISSAS PARA A APLICAÇÃO DOS GRÁFICOS DE SHEWHART

Os gráficos de controle convencionais exigem que os dados do processo atendam às suposições de normalidade e independência das amostras, sendo que somente quando estas características forem satisfeitas é possível aplicar os gráficos de controle e inferir conclusões relacionadas ao estado estatístico do processo. Dados que fogem à normalidade não seguem o comportamento estatístico nos quais estão baseados os princípios de funcionamento dos gráficos de controle. A independência das observações também apresenta extrema importância, sendo que os gráficos de controle convencionais podem apresentar níveis muito elevados de alarmes falsos mesmo para níveis pequenos de autocorrelação. Dessa forma, quando o processo apresenta evidências de um desvio da normalidade ou independência das observações, os limites de controle especificados anteriormente podem ser totalmente inapropriados para o monitoramento do processo na fase II (MONTGOMERY, 2016).

2.3.1 Aderência à distribuição normal

A avaliação da suposição de normalidade é exigida pela maioria dos procedimentos estatísticos, quando a suposição de normalidade é violada, a interpretação e a inferência das análises podem não ser confiáveis ou válidas. Portanto, é importante verificar essa suposição

antes de prosseguir com quaisquer procedimentos estatísticos. Basicamente, a maneira mais fácil de se verificar a suposição de normalidade é utilizando recursos gráficos, dentre os quais o gráfico de probabilidade normal é a ferramenta de diagnóstico mais comumente empregada. Outros métodos gráficos comuns que também podem ser utilizados incluem o histograma, o *boxplot* e o diagrama de ramo e folha (RAZALI; WAH, 2011).

Mesmo os métodos gráficos servindo como uma ferramenta útil na verificação da normalidade de uma amostra de n observações independentes, eles ainda não são suficientes para fornecer evidências conclusivas de que a suposição normal é válida. Portanto, para apoiar os métodos gráficos, existem métodos mais formais, especificamente os métodos numéricos e os testes de normalidade, que devem ser realizados antes de se assumir qualquer conclusão sobre a normalidade dos dados (RAZALI; WAH, 2011).

Os métodos numéricos incluem os coeficientes de assimetria e curtose (medida de forma que caracteriza o achatamento da curva da função de distribuição de probabilidade), enquanto que os testes de normalidade são um procedimento mais formal que envolvem testar a hipótese nula de que um conjunto de dados segue uma distribuição normal (RAZALI; WAH, 2011). De acordo com os autores, há uma quantidade expressiva de testes de normalidade disponíveis na literatura, no entanto, os procedimentos mais comuns disponíveis em *softwares* estatísticos são o teste de Shapiro-Wilk, o teste de Kolmogorov-Smirnov, o teste de Anderson-Darling e o teste de Lilliefors. Diferentes testes de normalidade frequentemente produzem resultados distintos, ou seja, alguns testes rejeitam a hipótese nula de normalidade, enquanto que outros falham em rejeitá-la. Portanto, a escolha do teste de normalidade a ser utilizado deve receber grande atenção (RAZALI; WAH, 2011).

Os testes de Anderson-Darling e Kolmogorov-Smirnov são baseados na função de distribuição empírica (FDE) dos dados. O teste de Kolmogorov-Smirnov trabalha com a maior diferença entre a distribuição empírica e a hipotética. O teste de Lilliefors é utilizado quando os parâmetros da distribuição não são conhecidos (RAZALI; WAH, 2011). O teste Anderson-Darling pertence à classe quadrática de estatísticas baseadas na FDE, trabalhando com as diferenças quadráticas entre a distribuição empírica e a hipotética. Por sua vez, o teste de Shapiro-Wilk se baseia nos valores amostrais ordenados elevados ao quadrado e tem sido o teste de normalidade mais utilizado por se mostrar o mais poderoso teste entre os disponíveis (LEOTTI; BIRCK; RIBOLDI, 2005).

Para a aplicação dos gráficos de controle quando os dados não apresentam aderência à distribuição normal, podem ser utilizadas técnicas de transformação com o objetivo de trazer os dados para a normalidade. Transformar os dados é um procedimento padrão muitas vezes

utilizado quando os dados apresentam fuga da normalidade. As técnicas de transformação dos dados mais comumente utilizadas são a logarítmica, raiz quadrada, raiz cúbica, angular, percentual, hiperbólica de primeiro ou segundo grau e a transformação em valores de z (MONTGOMERY, 2016).

2.3.2 Independência das amostras

A verificação da independência dos dados pode ser obtida por meio dos testes de autocorrelação e de autocorrelação parcial. A autocorrelação é definida como a medida de dependência de uma observação num dado tempo t em relação àquela mesma observação num instante $t - k$, em que k é a distância considerada entre duas observações, denominada de defasagem (*lag*) (MONTGOMERY, 2016).

Ainda de acordo com o autor, para mensurar a autocorrelação ao longo de uma série de observações distribuídas no tempo tem-se uma característica longitudinal, ou seja, uma série temporal, cujo coeficiente linear é chamado de função de autocorrelação e é dado pela Equação (38) (MONTGOMERY, 2016).

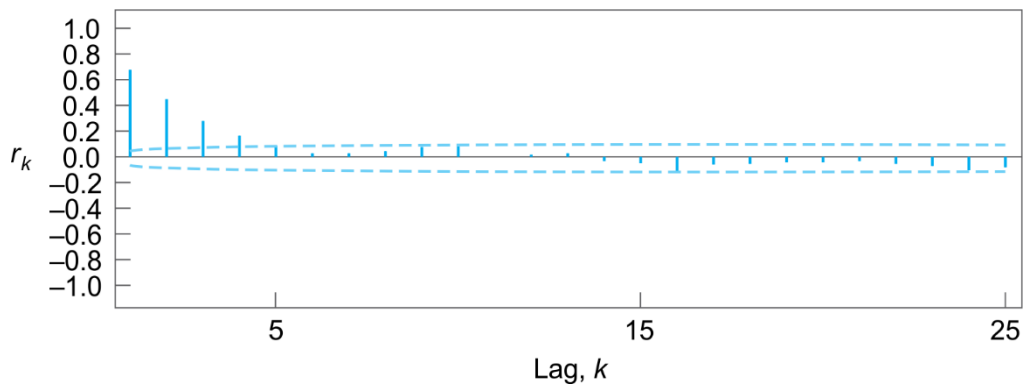
$$\rho_k = \frac{cov(x_t, x_{t-k})}{V(x_t)} \quad (38)$$

em que $cov(x_t, x_{t-k})$ é a covariância de observações separadas por k períodos de tempo e $V(x_t)$ representa a variância constante. Em linhas gerais, os valores de ρ_k são estimados pela função de autocorrelação, apresentada pela Equação (39) (MONTGOMERY, 2016).

$$r_k = \frac{\sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t-k} - \bar{x})}{\sum_{t=1}^{n-k} (x_t - \bar{x})^2} \quad (39)$$

sendo x_t a observação individual ou média dos subgrupos no tempo t , $\bar{x}$ a média das amostras e k um número inteiro que representa o *lag*. A série temporal é plotada em um diagrama (Figura 4) com limites superior e inferior de dois desvios padrão, representados pelas linhas tracejadas. Quando as observações extrapolam as linhas tracejadas, significa que os dados apresentam autocorrelação (MONTGOMERY, 2016).

Figura 4 – Função de autocorrelação das amostras



Fonte: Montgomery (2016)

A autocorrelação dos dados é um problema que pode ocasionar em um número maior de alarmes falsos, sendo associada à frequência de amostragem, como intervalos pequenos entre as observações, erros na coleta ou a presença de variáveis cíclicas. Uma solução para o problema de autocorrelação é a utilização de um modelo autorregressivo de série temporal, para a eliminação da autocorrelação nos resíduos, a outra seria a construção de gráficos de controle com limites mais largos (MIRANDA, 2005).

2.4 GRÁFICO DE CONTROLE EWMA

O gráfico de controle EWMA é uma alternativa ao gráfico de controle de Shewhart quando se pretende detectar pequenas mudanças no processo, sendo tipicamente mais utilizado para medições individuais, ou seja, $n = 1$ (MONTGOMERY, 2016). Apresentado inicialmente por Roberts em 1959, foi introduzido por Wortham e Heinrich (1972) para aplicações em processos em que amostras unitárias por período de tempo estivessem disponíveis. A observação unitária pode ser uma média, um valor individual, uma fração ou uma observação similar (WADSWORTH; STEPHENS; GODFREY, 2001), ou seja, a estatística EWMA pode ser utilizada tanto para medições individuais quanto para subgrupos racionais.

Devido sua semelhança com os gráficos de Shewhart, o gráfico EWMA possui um bom desempenho e facilidade na compreensão dos dados pelo usuário (WOODALL; MARAGAH, 1990). Segundo Hunter (1986), as principais diferenças entre os gráficos de Shewhart e EWMA estão relacionadas à maneira com que cada um dos métodos utiliza as informações coletadas no processo. De acordo com o autor, enquanto o gráfico de Shewhart depende unicamente do último ponto coletado, o gráfico EWMA atribui um peso maior às amostras mais recentes e menor para as amostras anteriores, decrescendo de forma geométrica da mais recente até a mais

distante. Nos gráficos EWMA, médias móveis passadas e atuais são calculadas recebendo pesos que decrescem exponencialmente com a distribuição dos dados, dessa forma, a detecção de comportamentos irregulares nesse gráfico é mais influenciada pelo desempenho recente do processo (MONTGOMERY, 2016). Esta ponderação é realizada por meio da constante de suavização λ , a qual é um dos parâmetros de projeto do gráfico. Como o gráfico EWMA pode ser considerado uma média ponderada de todas as observações do processo, ele é menos sensível que o gráfico de Shewhart à suposição de normalidade dos dados (ZAGO, 2009).

Em estudo comparativo de gráficos para medidas individuais realizado por Borror, Montgomery e Runger (1999), os gráficos de Shewhart apresentaram um grande aumento no número de alarmes falsos quando os dados apresentam uma distribuição t ou gama, o qual pode ocasionar ajustes desnecessários no processo, falta de confiança na ferramenta e perda de produtividade. Embora ambos os gráficos se mostraram capazes de uma detecção rápida das mudanças no processo, o EWMA foi capaz de manter o índice de ARL tanto para as distribuições normais quanto para as não normais, mostrando-se mais robusto que o gráfico Shewhart em relação a não normalidade.

Ao comparar os procedimentos de tomada de decisão em relação ao estado do processo, o gráfico EWMA é semelhante ao gráfico de Shewhart, especialmente em termos de limites de controle, de maneira que se algum ponto plotado exceder os limites de controle do gráfico EWMA, o processo é considerado fora de controle estatístico. Como nenhuma outra regra de decisão é necessária, esta característica torna o gráfico EWMA uma ferramenta bastante simples (ZAGO, 2009).

2.4.1 Base estatística do gráfico EWMA

O gráfico EWMA é definido pela Equação (40) (MONTGOMERY, 2016).

$$z_i = \lambda x_i + (1 - \lambda)z_{i-1} \quad (40)$$

sendo z_i o valor atual do peso da média móvel, x_i o valor atual da amostra, i um número inteiro que representa o índice atual e $0 < \lambda < 1$ a constante de suavização. O valor de z_0 é o alvo do processo e deve ser estimado para a primeira amostra, podendo assumir o valor da média do processo.

A estatística EWMA pode ser definida de duas formas equivalentes. A primeira utiliza os dados na forma original, aplicando-os diretamente na equação acima, a segunda envolve uma padronização dos dados, utilizando-se a Equação (41) (MONTGOMERY, 2016).

$$y_i = \frac{(x_i - \mu_0)\sqrt{n}}{\sigma} \quad (41)$$

Dessa forma, a estatística EWMA padronizada pode ser definida pela Equação (42) (MONTGOMERY, 2016), em que λ representa a memória do gráfico e é definida pelo parâmetro que atribui um peso às observações mais recentes.

$$z_i = \lambda y_i + (1 - \lambda)z_{i-1} \quad (42)$$

Em suma, ambas as equações representam estatísticas EWMA funcionalmente semelhantes, pois indicam matematicamente a mesma informação, sendo que usualmente utiliza-se $z_0 = 0$ para a estatística padronizada.

2.4.2 Parâmetros estatísticos do gráfico EWMA

Se as observações x_i são variáveis aleatórias independentes com variância σ^2 , então a variância de z_i é dada pela Equação (43) (MONTGOMERY, 2016).

$$\sigma_{z_i}^2 = \sigma^2 \left(\frac{\lambda}{2 - \lambda} \right) [1 - (1 - \lambda)^{2i}] \quad (43)$$

Dessa forma, os limites de controle do gráfico EWMA são obtidos com base no valor assintótico da variância de z_i . Portanto, estimado o desvio padrão do processo, os limites de controle são representados pelas Equações (44), (45) e (46) (MONTGOMERY, 2016).

$$LSC = \mu_0 + L\sigma \sqrt{\frac{\lambda}{(2 - \lambda)} [1 - (1 - \lambda)^{2i}]} \quad (44)$$

$$LC = \mu_0 \quad (45)$$

$$LIC = \mu_0 - L\sigma \sqrt{\frac{\lambda}{(2 - \lambda)} [1 - (1 - \lambda)^{2i}]} \quad (46)$$

sendo o fator L a extensão dos limites de controle, ou seja, o número de múltiplos de desvio padrão que determina a distância dos limites de controle até a linha central.

À medida que o índice i cresce, o termo $[1 - (1 - \lambda)^{2i}]$ se aproxima de 1 (MONTGOMERY, 2016). Isso significa que, depois de o gráfico EWMA ter percorrido diversos períodos de tempo, os limites de controle se aproximarão dos valores de estado estacionário, dados pelas Equações (47), (48) e (49) (MONTGOMERY, 2016).

$$LSC = \mu_0 + L\sigma \sqrt{\frac{\lambda}{(2 - \lambda)}} \quad (47)$$

$$LC = \mu_0 \quad (48)$$

$$LIC = \mu_0 - L\sigma \sqrt{\frac{\lambda}{(2 - \lambda)}} \quad (49)$$

Montgomery (2016) destaca que para pequenos valores de i , o desempenho do gráfico de controle EWMA aumentará consideravelmente utilizando-se as equações primitivas para o cálculo dos limites de controle.

2.4.3 Cálculo do ARL para o gráfico EWMA

O procedimento de planejamento ótimo do gráfico EWMA consiste da especificação do ARL e da magnitude de mudança do processo, para então proceder-se à escolha da combinação de λ e L que forneça o ARL desejado. Em geral, valores de $0,05 \leq \lambda \leq 0,25$ têm uma boa performance, sendo $\lambda = 0,05$; $0,10$ e $0,20$ escolhas populares. Uma boa regra empírica é o uso de valores menores de λ para detecção de mudanças menores. Por exemplo, para valores de λ próximos a 0, a observação anterior tem maior peso, e para $\lambda = 1$, a observação anterior é anulada, e o EWMA se transforma em um gráfico de Shewhart (CROWDER, 1989).

Crowder (1989) apontou que os erros tipo I e tipo II são caracterizados pela definição dos ARLs, de modo que a escolha da combinação λ e L deve levar em consideração que a probabilidade de um erro tipo I influenciará na possibilidade de um erro tipo II acontecer. Diversos estudos foram feitos com relação ao ARL dos gráficos EWMA e entre eles podem ser citadas as contribuições de Crowder (1987b, 1989), Roberts (1959), Robinson e Ho (1978) e Lucas e Saccucci (1990). Crowder (1987a, 1989) apresentou três possíveis metodologias para

aplicação do gráfico EWMA, incluindo valores tabelados de ARL, métodos gráficos ou métodos computacionais.

De maneira geral, o gráfico EWMA tem melhor desempenho para mudanças pequenas, e em contrapartida, o gráfico de Shewhart consegue detectar mais rapidamente alterações maiores. Para melhorar a sensibilidade dos gráficos a mudanças maiores sem diminuir a habilidade em detectar mudanças pequenas, é comum a combinação dos gráficos de Shewhart e EWMA para o monitoramento do processo (MONTGOMERY, 2016). O gráfico EWMA também pode ser utilizado para o monitoramento de processos em que o número de amostras seja maior que um, nesse caso, são construídos os gráficos para subgrupos racionais. Em virtude das análises realizadas nesta pesquisa considerarem exclusivamente o monitoramento de medições individuais do consumo diário de água, este gráfico não é apresentado neste trabalho.

2.5 ALTERNATIVAS NÃO PARAMÉTRICAS AOS GRÁFICOS DE CONTROLE

Em diversas situações o processo não segue uma distribuição normal e, como resultado, as propriedades estatísticas dos gráficos de controle convencionais podem ser altamente afetadas (CHAKRABORTI; VAN DER LAAN; BAKIR, 2001). De acordo com os autores, nesses casos podem ser utilizados gráficos de controle não paramétricos, os quais não dependem da normalidade ou de qualquer outra suposição específica sobre a distribuição dos dados. A maioria das técnicas não paramétricas de CEP depende de postos que representam a posição das observações em relação a um parâmetro predefinido e foram desenvolvidas como alternativa aos usuais gráficos de controle de Shewhart, CUSUM e EWMA (MONTGOMERY, 2016).

Uma definição formal de um gráfico de controle não paramétrico, também chamado de livre de distribuição, é dada em termos do seu ARL, se a estimação do ARL for igual ao equivalente gráfico paramétrico mesmo quando os dados não apresentam distribuição normal, o gráfico é chamado de livre de distribuição (CHAKRABORTI; VAN DER LAAN; BAKIR, 2001). Conforme os autores destacam, a principal vantagem desses gráficos é a flexibilidade obtida pela não necessidade de se presumir uma distribuição de probabilidade paramétrica para o processo, pelo menos no que diz respeito ao estabelecimento e implementação do gráfico. Comumente, os gráficos de controle não paramétricos tendem a compartilhar as propriedades de robustez de testes não paramétricos e intervalos de confiança e, portanto, são menos sensíveis à presença de *outliers*.

Em suma, as vantagens dos gráficos de controle não paramétricos incluem a sua simplicidade, falta de necessidade de se presumir uma distribuição normal, um ARL equivalente mesmo quando os dados não apresentam distribuição normal, maior eficiência na detecção de mudanças no processo quando a distribuição do conjunto de dados foge à normalidade e, por fim, a não necessidade de estimação da variância para construção dos gráficos (CHAKRABORTI; VAN DER LAAN; BAKIR, 2001).

O ponto de partida para projetar um gráfico de controle é geralmente uma estatística de controle que representa um estimador do parâmetro de interesse. A estatística de controle tradicional para a média é a média da amostra ($\bar{x}$), enquanto que para a variação do processo utiliza-se o desvio padrão da amostra (s) ou a amplitude amostral (R). Ao construir gráficos de controle não paramétricos, ocorre a substituição das estatísticas de controle convencionais por equivalentes livres de distribuição. Com isso, permite-se a comparação das propriedades do ARL entre os gráficos (CHAKRABORTI; VAN DER LAAN; BAKIR, 2001).

Reconhecendo o potencial dos gráficos de controle não paramétricos, Chakraborti, Van der Laan e Bakir (2001) forneceram um relato completo da literatura disponível sobre a aplicação dessas técnicas até o final do ano de 2000, a qual foi atualizada por Chakraborti e Graham (2008) e Chakraborti e Graham (2019). Em estudo recente, Chakraborti e Graham (2019) apontam que as pesquisas nessa área crescem em um ritmo rápido e muitas novas contribuições têm sido adicionadas nos últimos anos.

Conforme Chakraborti e Graham (2019) destacam, existem disponíveis na literatura diversas técnicas não paramétricas de CEP, desenvolvidas a partir da modificação dos tradicionais gráficos de Shewhart, CUSUM e EWMA, podendo ser utilizadas para o monitoramento de processos univariados ou multivariados. Estudos sugerem que o gráfico EWMA para medidas individuais proporciona vantagens claras no monitoramento da demanda de água (JUNG *et al.*, 2015; FREITAS *et al.*, 2019). No entanto, estes gráficos são sensíveis à suposição de normalidade dos dados, o que pode se refletir no aumento da sinalização de alarmes falsos (AMIN; SEARCY, 1991). Dessa forma, a técnica não paramétrica escolhida para a realização do monitoramento das séries temporais do consumo de água neste trabalho é o gráfico de controle EWMA-SN, desenvolvido a partir do gráfico de controle EWMA.

2.6 GRÁFICO DE CONTROLE NÃO PARAMÉTRICO EWMA-SN

Elaborado por Graham, Chakraborti e Human (2011), o gráfico EWMA-SN utiliza a estatística de sinais como parâmetro de análise aplicado a um gráfico de controle EWMA. Esse gráfico pode ser utilizado tanto para medições individuais ou subgrupos e apresenta a vantagem de não exigir o valor numérico das observações, apenas a indicação de se elas são maiores ou menores que um parâmetro de interesse estabelecido. Essa técnica exige que os parâmetros de interesse do processo sob controle sejam conhecidos, classificando-se como gráfico de controle de fase II. Algumas das vantagens da aplicação do gráfico não paramétrico EWMA-SN estão listadas a seguir:

- a) a suposição de simetria não é exigida;
- b) pode ser utilizado para monitorar qualquer parâmetro de interesse conhecido da distribuição, incluindo a mediana;
- c) pode ser utilizado tanto para amostras de medições individuais ($n = 1$) quanto para subgrupos ($n > 1$);
- d) é um gráfico eficiente quando as amostras apresentam distribuição assimétrica;
- e) não requer os valores numéricos reais das observações, a única informação necessária é se a observação é maior ou menor do que o parâmetro de interesse especificado. Dessa forma, o gráfico não paramétrico EWMA-SN somente pode ser aplicado para variáveis dicotomizadas, ou seja, quando é possível afirmar que a única informação importante do processo é se o dado é "alto" ou "baixo";
- f) a variância do processo não precisa ser especificada;

2.6.1 Base estatística do gráfico EWMA-SN

Seja X_i a medição individual de um processo de tamanho i com função de distribuição cumulativa F e mediana θ sendo monitorada, define-se a estatística de sinais pela Equação (50) (GRAHAM; CHAKRABORTI; HUMAN, 2011).

$$SN_i = \text{sign}(x_i - \theta_0) \quad (50)$$

em que i é um número inteiro que representa o tamanho do conjunto de dados e θ é o valor especificado da mediana a ser monitorada. A variável sign é um sinal atribuído à observação

atual em relação a um parâmetro pré-definido. É obtido pela Equação (51), assumindo um valor de 1, 0 ou -1 para cada observação (GRAHAM; CHAKRABORTI; HUMAN, 2011).

$$sign = \begin{cases} 1 & \text{se } x > 0 \\ 0 & \text{se } x = 0 \\ -1 & \text{se } x < 0 \end{cases} \quad (51)$$

A estatística de plotagem do gráfico não paramétrico EWMA-SN é obtida acumulando-se sequencialmente a estatística de sinal SN_i e é definida pela Equação (52) (GRAHAM; CHAKRABORTI; HUMAN, 2011).

$$Z_i = \lambda SN_i + (1 - \lambda)Z_{i-1} \quad (52)$$

em que $0 < \lambda < 1$ é uma constante de suavização que será apresentada adiante e $Z_0 = 0$.

Graham, Chakraborti e Human (2011) destacam que, embora o gráfico EWMA-SN possa ser considerado para monitorar qualquer percentil de uma distribuição contínua, suas análises se concentraram no 50º percentil, ou seja, a mediana, de modo que θ denota o valor conhecido ou especificado da mediana do processo sob controle. Os autores apontam que as razões para focar na mediana envolvem a sua melhor representação do valor central e sua maior robustez em relação à média, além de se igualar à mesma quando a distribuição é simétrica.

2.6.2 Parâmetros estatísticos do gráfico EWMA-SN

Quando a observação sendo plotada é igual ao parâmetro definido, o sinal atribuído é nulo, assim, a média é $E(Z_i) = 0$. O desvio padrão da estatística de plotagem Z_i é obtido pela Equação (53) (GRAHAM; CHAKRABORTI; HUMAN, 2011).

$$\sigma_{Z_i} = \sqrt{\frac{\lambda}{2 - \lambda} (1 - (1 - \lambda)^{2i})} \quad (53)$$

Dessa forma, a linha central e os limites de controle superior e inferior do gráfico de controle não paramétrico EWMA-SN para a mediana são dados pelas Equações (54), (55) e (56) (GRAHAM; CHAKRABORTI; HUMAN, 2011).

$$LSC = L \sqrt{\frac{\lambda}{2 - \lambda} (1 - (1 - \lambda)^{2i})} \quad (54)$$

$$LC = 0 \quad (55)$$

$$LIC = -L \sqrt{\frac{\lambda}{2 - \lambda} (1 - (1 - \lambda)^{2i})} \quad (56)$$

em que L é a distância dos limites de controle até a linha central.

Basicamente, o gráfico não paramétrico EWMA-SN monitora a estatística de sinal Z_i acompanhada da linha central e limites de controle no eixo vertical em relação ao número de observações i no eixo horizontal. Quando qualquer observação Z_i plotar além dos limites de controle, o processo é declarado como fora de controle estatístico e uma busca por causas atribuíveis pode ser iniciada. Caso as observações sejam plotadas entre os limites de controle, o processo é considerado sob controle, não sendo necessário nenhum tipo de intervenção (GRAHAM; CHAKRABORTI; HUMAN, 2011).

2.6.3 Cálculo do ARL para o gráfico EWMA-SN

O ARL do gráfico de controle EWMA-SN é definido determinando-se λ e L específicos. A constante λ é o parâmetro de suavização e é selecionada dependendo da magnitude da mudança a ser detectada, quando $\lambda = 1$ obtém-se um gráfico de controle de Shewhart, ou seja, não se considera a observação anterior. A constante L deve ser escolhida em combinação com o valor do parâmetro de suavização λ e representa a distância da linha central aos limites de controle, quanto maior seu valor, mais largos os limites de controle (GRAHAM; CHAKRABORTI; HUMAN, 2011).

A distribuição do ARL do gráfico não paramétrico EWMA-SN é derivada a partir da utilização de uma abordagem da cadeia de Markov. Esta abordagem implica que a estatística de plotagem seja vista como seguindo uma cadeia de Markov, caracterizada por um espaço de estado S e uma matriz de probabilidade de transição M (BROOK; EVANS, 1972). De acordo com os autores, para uma estatística ponderada pelo tempo, tem-se uma cadeia de Markov de estado contínuo com uma matriz de probabilidade de transição de estado infinito. Dessa forma, os valores da estatística de sinais utilizada no gráfico EWMA-SN são discretizados em uma cadeia de Markov de estado finito para que os resultados aproximados de ARL possam ser obtidos.

O espaço de estado consiste em dois tipos de estados: um estado de absorção, quando a estatística Z_i é plotada além dos limites de controle e o gráfico emite um sinal de alerta, e r estados transientes ou não absorventes. De modo que, no total, existem $r + 1$ estados. Os r estados transitórios correspondem a r subintervalos de comprimento igual obtidos dividindo o intervalo entre os limites de controle superior e inferior. O valor de r afeta diretamente a precisão dos resultados, sendo que quanto maior o seu valor, mais precisos serão os resultados aproximados (BROOK; EVANS, 1972).

A matriz de probabilidade de transição é desenhada de forma a conter todas as probabilidades de um estado transiente ir para um estado de absorção ou para outro estado transiente e também o oposto, ou seja, um estado de absorção ir para um estado transiente ou para outro estado de absorção. O valor do ARL do gráfico de controle não paramétrico EWMA-SN é quantificado pelo tempo de espera para a cadeia de Markov entrar no estado de absorção pela primeira vez (BROOK; EVANS, 1972).

2.7 TÉCNICAS DE MONITORAMENTO DE ÁGUA

Nas últimas décadas, a gestão da demanda de água surgiu como uma estratégia fundamental para a busca da sustentabilidade na distribuição de água e o aprimoramento dos serviços de água (DI MAURO *et al.*, 2021). Conforme os autores relatam, o desenvolvimento de tecnologias para o gerenciamento e medição do consumo de água promoveu a coleta de dados em escalas espaciais cada vez maiores e intervalos de tempo gradativamente menores. Ainda de acordo com os autores, a disponibilidade de dados de consumo de água em escalas residenciais ou de usos finais com uma amostragem de tempo de segundos ou minutos abriu oportunidades sem precedentes para a compreensão dos comportamentos do consumo de água e a modelagem da demanda de água.

2.7.1 Revisão de estudos sobre o monitoramento de água

Para aprimorar o conhecimento dos estudos envolvendo o monitoramento de água e a compreensão das técnicas utilizadas para este fim, este capítulo apresenta uma revisão narrativa de estudos abrangendo a temática do monitoramento de água. O objetivo desta revisão é apresentar uma síntese geral dos equipamentos e métodos mais utilizados para a obtenção e posterior monitoramento dos dados de consumo de água.

Em um artigo de revisão elaborado por Di Mauro *et al.* (2021), os autores contribuíram com a disponibilização de uma lista abrangente e atualizada de publicações abordando a temática do consumo de água (DI MAURO *et al.*, 2020). Com a finalidade de identificar as ferramentas utilizadas nas pesquisas, esta lista foi atualizada com estudos publicados posteriormente à data da sua publicação e pesquisas não contempladas na revisão, sendo incluída também uma classificação quanto aos equipamentos e técnicas utilizados para o monitoramento da demanda de água.

A estratégia de pesquisa utilizou a combinação das seguintes palavras-chave nas plataformas *Web of Science* e *Google Scholar*: *water demand* (demanda de água), *water consumption* (consumo de água), *water monitoring* (monitoramento de água), *pipe bursts* (vazamentos em tubulações), *residential water consumption monitoring* (monitoramento do consumo de água residencial), *urban water consumption monitoring* (monitoramento do consumo de água urbano), *residential water demand forecast* (previsão da demanda de água residencial), *urban water demand forecast* (previsão da demanda de água urbana).

Não houve limitação no período de busca, de modo a serem consideradas todas as publicações desde o início das plataformas. Após a compilação de todos os estudos catalogados no artigo de revisão (DI MAURO *et al.*, 2020) e das publicações obtidas utilizando a estratégia de pesquisa definida, os artigos foram revisados e classificados de acordo com três critérios: escala, obtenção dos dados e método.

O critério referente à escala classifica os estudos conforme a escala espacial adotada e é categorizado da seguinte forma:

- a) urbana: se refere a uma cidade, centro urbano ou distrito. A escala cidade ou centro urbano inclui toda a rede de distribuição de água urbana sendo medida pela concessionária de água. A escala distrital se refere a um grupo de edifícios cujo consumo de água é medido pela empresa fornecedora de serviços de água;
- b) residencial: se refere a uma única habitação ou edificação residencial ligada a um medidor de água. Dependendo da disponibilidade dos dados, pode ter sido considerado o uso interno, externo ou total.

Foram identificados quatro equipamentos principais para a obtenção dos dados de demanda de água, os quais foram classificados conforme as definições a seguir:

- a) hidrômetro: instrumento de medição volumétrica de água individual, utilizado em larga escala pelas concessionárias de água para o faturamento das unidades consumidoras (RATHNAYAKA *et al.*, 2017). Em alguns estudos, os dados de consumo foram obtidos por meio da leitura das faturas emitidas a partir das medições dos hidrômetros;

- b) medidor de consumo com *data logger*: dispositivo eletrônico que monitora e registra medições em tempo real em intervalos de tempo constantes, permitindo determinar o comportamento de uma variável ao longo do tempo. Os dados são coletados após as medições, em consulta ao aparelho (MEYER; JACOBS; ILEMOBADE, 2021);
- c) medidor de vazão: equipamento de medição volumétrica de água de grandes dimensões, podendo ser instalado para medir a demanda de água em bairros ou conjuntos de edificações (BORGES; JUNG; KIM, 2017);
- d) *smart meter*: dispositivo eletrônico que registra medições em tempo real em intervalos de tempo constantes e tem a característica de retornar simultaneamente informações a respeito do padrão de consumo. Se difere por permitir a comunicação bidirecional entre o medidor e o consumidor, retornando os dados em tempo real (COMINOLA *et al.*, 2018).

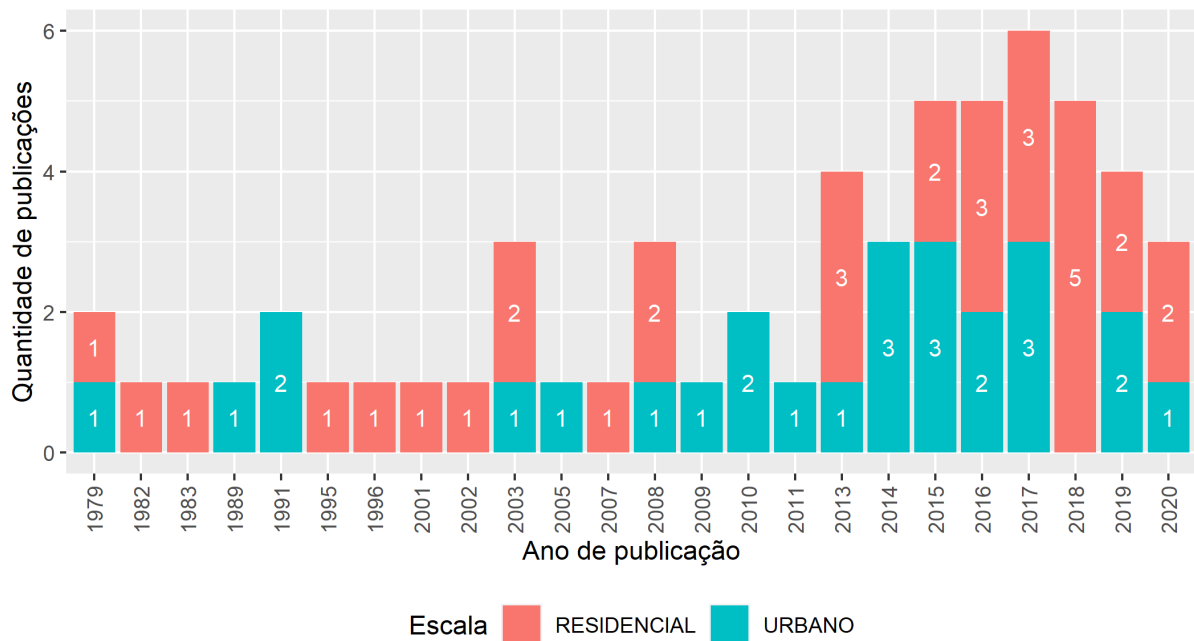
Quanto aos métodos de monitoramento de água abordados nos estudos, foram identificadas oito principais técnicas, definidas a seguir:

- a) descritivo: se refere à utilização de medidas numéricas, como a comparação de médias, medianas e desvios padrão, ou de técnicas visuais, por meio da análise de histogramas, gráficos *boxplot* ou gráficos de séries temporais;
- b) gráficos de controle: contempla a aplicação de gráficos de controle, podendo ter sido utilizados gráficos univariados ou multivariados de Shewhart, CUSUM, EWMA ou a combinação destes;
- c) inferencial: quando o padrão de consumo de água é comparado estatisticamente por meio de testes de hipótese;
- d) *machine learning*: se refere à prática da utilização de algoritmos para que uma máquina possa fazer uma determinação ou predição sobre o consumo de água;
- e) processo estocástico: a modelagem estocástica considera a probabilidade de o consumo de água seguir determinado padrão, caracterizado em termos do tipo de distribuição dos dados e dos parâmetros definidos;
- f) regressão: se refere a modelagem estatística dos dados a partir de variáveis independentes, tendo como objetivo prever o consumo de água com base nas informações de outros fatores;
- g) teste de detecção de mudanças: abrange a identificação de pontos em que a distribuição do consumo de água sofreu uma mudança significativa por meio da utilização de testes de detecção de mudanças ou de detecção de comportamentos anômalos.

2.7.2 Síntese dos estudos revisados

Ao todo foram identificadas 58 pesquisas abordando a temática de monitoramento de água dentro dos critérios estabelecidos, listadas no APÊNDICE A. Estas publicações estão divididas em 26 estudos de escala urbana e 32 estudos de escala residencial. Uma visão geral da distribuição temporal das publicações (Figura 5), sugere que houve um crescimento na quantidade de pesquisas publicadas a partir do ano de 2013. Isso provavelmente se deve tanto pelo crescente interesse no desenvolvimento de estratégias para a gestão da demanda de água quanto pelo surgimento de tecnologias de medição inteligente da água durante o período de 2011-2015 (COMINOLA *et al.*, 2015).

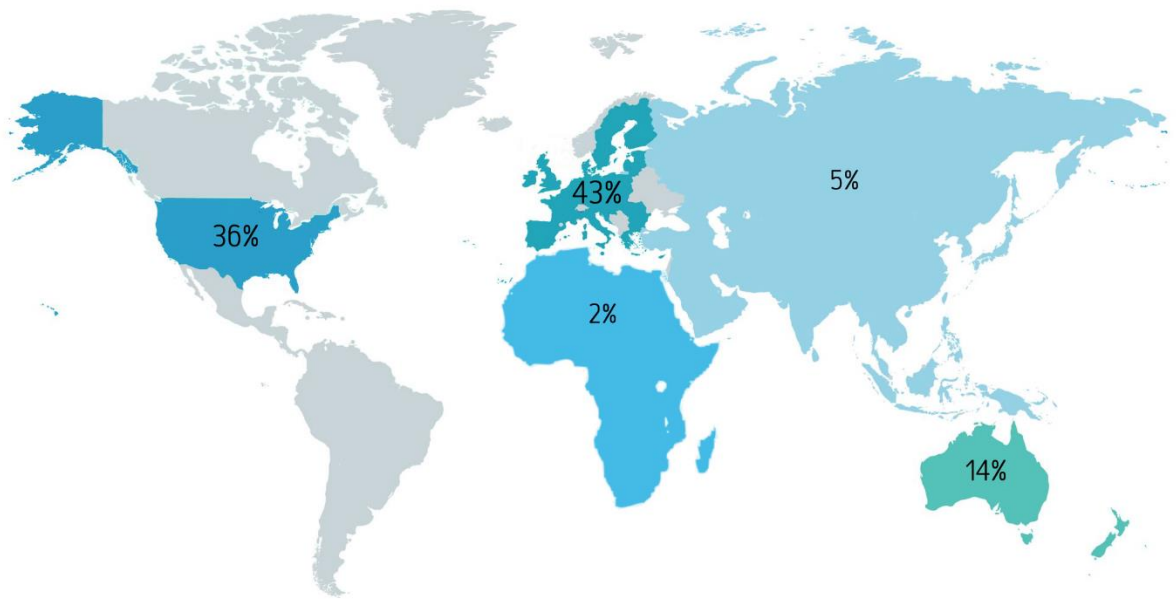
Figura 5 – Quantidade e ano de publicação dos estudos



Fonte: Elaborado pelo autor (2022)

A disposição geográfica das publicações revisadas (Figura 6) mostra uma distribuição espacial desigual, com mais de 75% (46) dos estudos revisados oriundos dos EUA ou do continente europeu, de modo que 43% (25) são de origem europeia, 36% (21) são estadunidenses, 14% (8) da Austrália, 5% (3) do continente asiático e, por fim, 2% (1) da África do Sul.

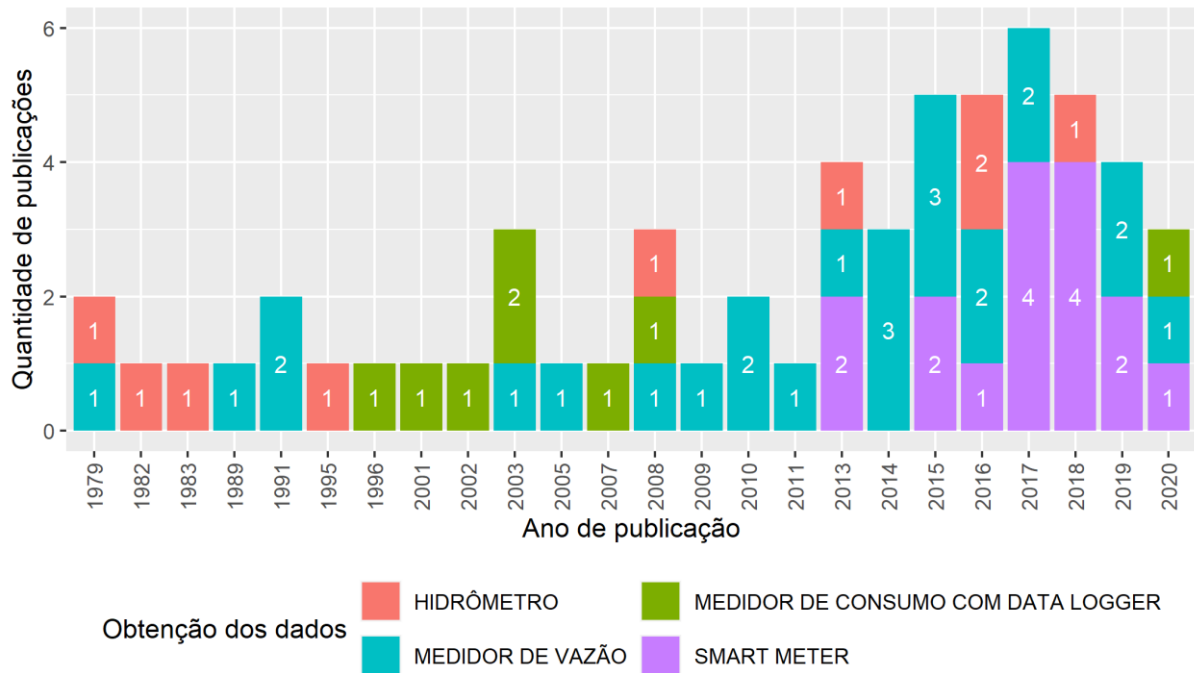
Figura 6 – Distribuição geográfica das publicações



Fonte: Elaborado pelo autor (2022)

A Figura 7 ilustra o efeito do surgimento de novas tecnologias de medição de água no desenvolvimento de pesquisas na área, de modo que após o ano de 2011 o número de estudos compreendendo o consumo de água teve um crescimento expressivo, associado principalmente à utilização de *smart meters*. A maior disponibilidade de tecnologias de obtenção dos dados de água estimulou uma busca cada vez maior pela compreensão dos padrões de consumo de água (DI MAURO *et al.*, 2021). Conforme os autores citam, este processo promoveu um aumento no número de pesquisas tanto em escala urbana quanto residencial, com diferentes escalas temporais e métodos de interpretação dos dados.

Figura 7 – Método de obtenção dos dados por ano

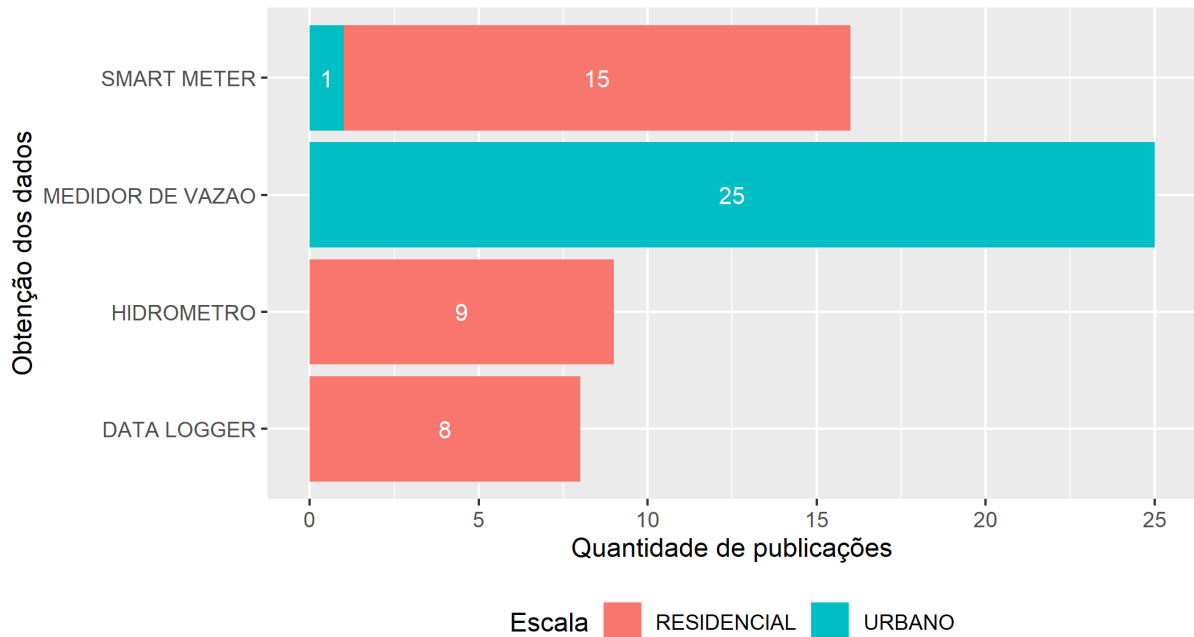


Fonte: Elaborado pelo autor (2022)

Os dados de demanda de água coletados em escala urbana são usados principalmente para monitorar padrões de demanda de água na rede ou para fornecer informações de entrada para modelos de simulação de sistemas de distribuição de água (ARAL *et al.*, 2006). Esses conjuntos de dados são quase que em sua totalidade obtidos por meio de relatórios de faturamento ou observações periódicas dos medidores de vazão (Figura 8), apresentando uma escala temporal de resolução diária ou mensal.

A Figura 8 também aponta que os *smart meters* são majoritariamente utilizados para pesquisas em escala residencial, indicando que o desenvolvimento desses equipamentos intensificou o estudo de padrões de consumo de água em nível domiciliar. As pesquisas abrangendo o consumo de água em escala residencial são de grande interesse para estudos comportamentais e fornecem informações importantes para promoção da conservação de água, projeção de tarifas de água, promoção de usos mais sustentáveis de recursos, caracterização da demanda de água durante horários de interesse e melhoria das capacidades de previsão e gestão da demanda de água (WILLIS *et al.*, 2013).

Figura 8 – Obtenção dos dados por escala espacial

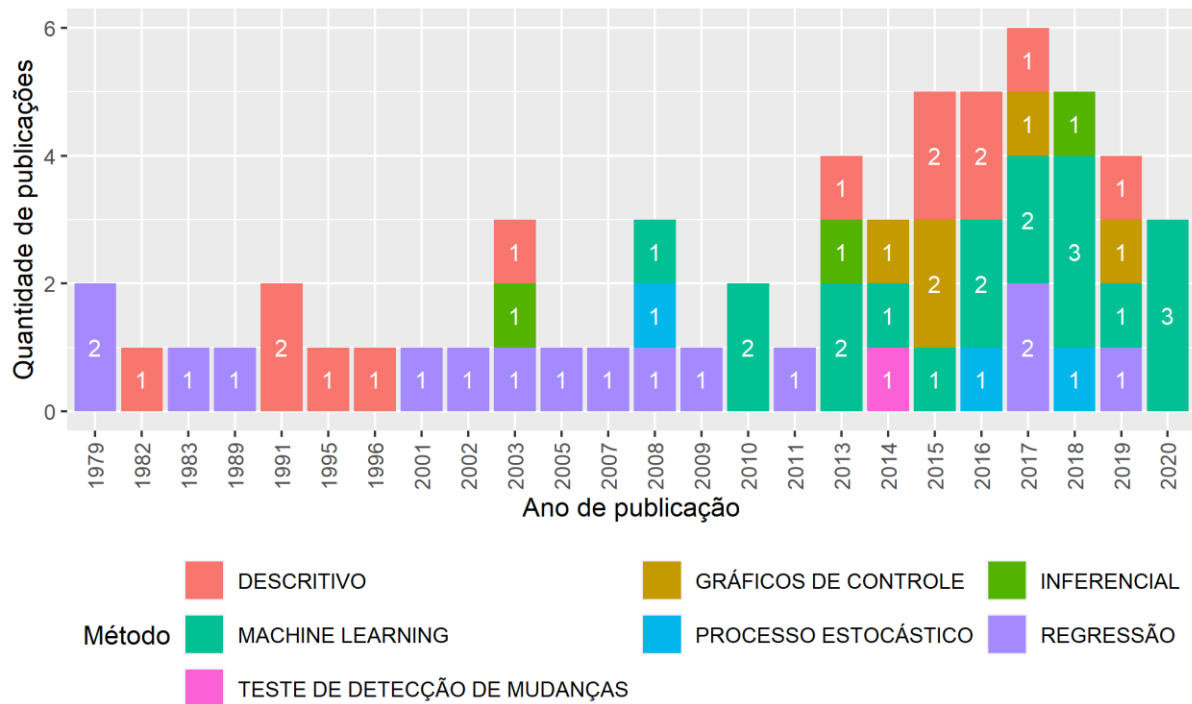


Fonte: Elaborado pelo autor (2022)

Pesquisas envolvendo a previsão da demanda de água vêm sendo realizadas há algumas décadas, com diferentes métodos de modelagem e maior ênfase a partir do ano de 2000 (Figura 9). Inicialmente, a maioria das pesquisas utilizava modelos de regressão para modelagem dos dados ou métodos descritivos para auxiliar na compreensão dos padrões do consumo de água. A partir do ano de 2010, o surgimento de medidores inteligentes de água potencializou o estudo dos padrões do consumo de água. Os *smart meters* proporcionam uma série de possibilidades no que tange ao estudo dos padrões do consumo de água, principalmente por apresentarem a vantagem de serem capazes de medir e armazenar os dados de uso de água com resolução subdiária, de até mesmo intervalos de segundos (GARGANO *et al.*, 2016).

Ainda neste contexto, o desenvolvimento de novos métodos de modelagem transformou o cenário do monitoramento hídrico, em que aproximadamente 45% das pesquisas utilizaram técnicas de *machine learning* como ferramenta para a modelagem dos conjuntos de dados.

Figura 9 – Métodos utilizados nas pesquisas a cada ano

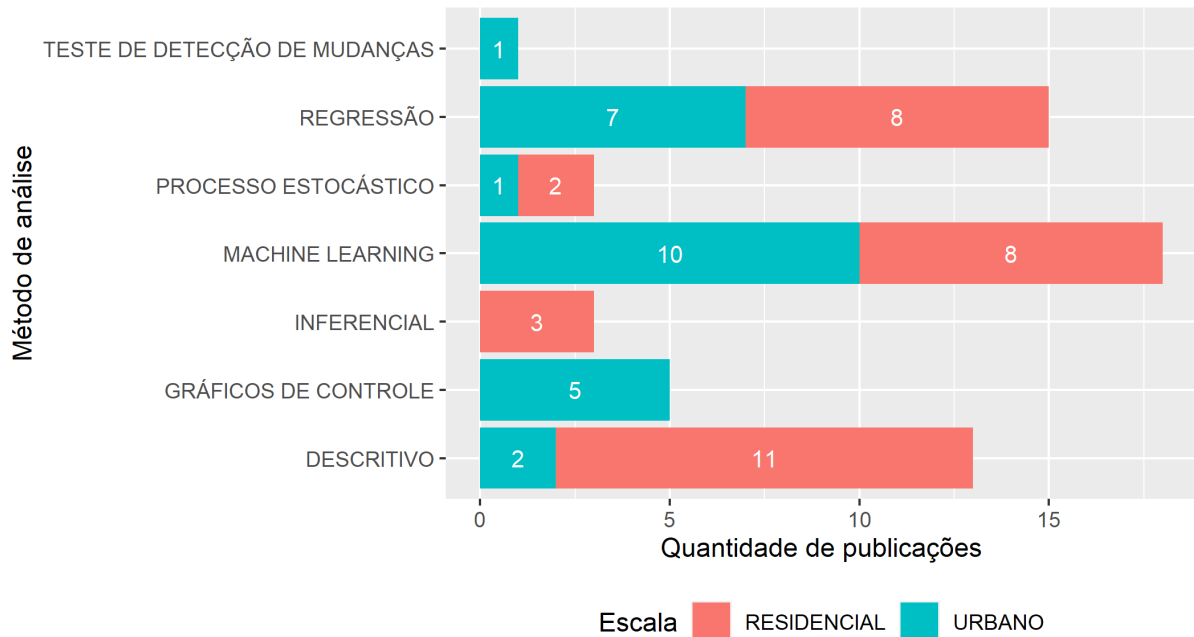


Fonte: Elaborado pelo autor (2022)

Em relação aos métodos de modelagem dos dados, existe uma maior concentração na utilização de técnicas de regressão e *machine learning*, divididas uniformemente entre as escalas urbana e residencial e representando mais de 50% dos artigos revisados (Figura 10). Isso mostra que os gráficos de controle, apesar de serem ferramentas já consolidadas no setor industrial, ainda são pouco utilizados para o monitoramento de água, sua aplicação no setor hídrico se dá principalmente em escala urbana e a partir do ano de 2014 (BAKKER *et al.*, 2014).

No que diz respeito à aplicação das técnicas de estatística inferencial, processos estocásticos e testes de detecção de mudanças, a principal característica consiste na possibilidade de se retirar conclusões sobre determinado modelo com base nas observações coletadas, apresentando forte relação com a aplicação de testes de hipóteses (GUTTORG, 1995). Assis, Sousa e Linhares (2020) definem o teste de hipóteses como sendo um processo estatístico utilizado com o objetivo de verificação da igualdade ou desigualdade de uma ou mais suposições, a partir da análise de dois ou mais conjuntos de dados.

Figura 10 – Métodos utilizados por escala espacial



Fonte: Elaborado pelo autor (2022)

2.7.3 Vantagens e limitações identificadas nos estudos

Algumas características e tendências puderam ser identificadas a partir das informações coletadas dos estudos. A resolução temporal da amostragem representa um fator importante para o desenvolvimento das pesquisas, de modo que dados de consumo de água coletados em intervalos de tempo menores permitem uma compreensão maior e mais precisa da amostra, podendo ser usados para detectar vazamentos de água mais rapidamente e fornecendo estimativas precisas dos horários de maior consumo (COMINOLA *et al.*, 2018).

Considerando-se os 58 artigos incluídos nesta revisão, observa-se uma grande variedade na definição das escalas temporais adotadas, a depender do equipamento sendo utilizado para a medição dos dados de consumo de água. Os conjuntos de dados coletados em escala urbana incluem principalmente dados registrados com resolução diária ou mensal, consistindo em medidas obtidas por observações periódicas de medidores ou relatórios de faturamento. Essa característica é consistente com as necessidades dos estudos, em que o objetivo é geralmente estimar a demanda de água para a rede ou a otimização da gestão e operação da rede de distribuição de água. Apenas algumas exceções incluem dados com uma resolução de tempo em intervalos de minutos, o que pode estar atribuído à dimensão dos estudos, à extensão do período sendo estudado e aos custos de implantação de *smart meters* em larga escala. Por sua vez, os conjuntos de dados de escala residencial apresentam mais estudos com dados coletados

em intervalos de tempo menores. A classificação desses conjuntos de dados com base em sua resolução temporal revela que 62,5% das pesquisas contém dados coletados em intervalos de tempo de horas, minutos ou até mesmo segundos.

A amostragem também influencia nos métodos que podem ser utilizados para a realização do monitoramento do consumo de água. A análise descritiva apresenta a vantagem de ser uma síntese numérica, podendo ser aplicada a qualquer escala temporal e tendo como objetivo básico resumir uma série de valores por meio de um conjunto de ferramentas e técnicas, como tabelas, gráficos ou medidas de variabilidade e tendência, que ajudam na obtenção de uma visualização global dos dados (CRESCO, 2020). Desse modo, esse método se caracteriza pela capacidade de analisar praticamente quaisquer conjuntos de dados, fornecendo informações importantes de valores máximos, mínimos, atípicos e inclusive dados faltantes, representando uma etapa fundamental para qualquer estudo (MANCUSO *et al.*, 2019).

No entanto, a estatística descritiva não permite testar hipóteses, como comparar o uso de água em residências unifamiliares ou edifícios multifamiliares, ou verificar mudanças na demanda de água ao longo do tempo. Nesse contexto, os estudos de estatística inferencial, incluindo os métodos de processo estocástico e os testes de detecção de mudança, são utilizados como um método de averiguação sobre a veracidade de uma afirmação (HIRAKATA; BRANCO MANCUSO; DE JESUS CASTRO, 2019). Ainda segundo os autores, a estatística inferencial utiliza testes de hipóteses como regra de decisão para aceitar ou rejeitar uma hipótese, com base nos dados coletados em uma amostra, sendo uma ferramenta poderosa para a tomada de decisões.

Em outros estudos, o objetivo é compreender o comportamento do consumo de água em relação a outras variáveis, identificando se a alteração destas pode causar mudanças no padrão do consumo de água. Conforme afirmado por Leotti *et al.* (2020), a modelagem estatística apresenta técnicas capazes de cumprir esse objetivo, sendo a análise de regressão muito utilizada para supor relações de dependência entre o consumo de água e outras variáveis, ajustando a melhor função matemática capaz de prever o comportamento do consumo de água. A identificação de fatores que alteram o consumo de água traz vantagens claras para o desenvolvimento de práticas sustentáveis quanto ao uso da água, podendo orientar a criação de políticas públicas de conservação e campanhas de conscientização (GARCIA *et al.*, 2019).

Assim como nos modelos de regressão, a capacidade de gerar modelos capazes de prever ou simular a demanda de água, tanto em nível urbano quanto residencial, também é objeto de estudo em pesquisas utilizando técnicas de *machine learning*. Nesse cenário, são utilizados algoritmos para a coleta e interpretação de dados, seguindo parâmetros definidos pelos

pesquisadores, a fim de auxiliar uma máquina a reconhecer padrões e tomar decisões (RODRIGUES, 2021). A utilização de algoritmos computacionais proporciona uma série de vantagens para o estudo, dado que a máquina aprende manipulando uma grande quantidade de informações em um curto período de tempo, auxiliando a compreensão das variações da demanda de água (OYEBODE; IGHRAVWE, 2019).

No contexto do monitoramento do consumo de água, os gráficos de controle apresentam uma vantagem relacionada à sua versatilidade, podendo ser aplicados em conjuntos de dados de escala urbana ou residencial e com diferentes escalas temporais. Dessa forma, a aplicação de gráficos de controle para a gestão da água se revela como uma ferramenta atual e relevante, podendo representar um avanço no que tange à automatização de processos de monitoramento de grandes volumes de água (DURMUSOGLU, 2018).

Em suma, esforços de pesquisas diferenciados permitem um maior entendimento dos conjuntos de dados de consumo de água e contribuem para aprimorar os conhecimentos acerca da gestão hídrica. A utilização de gráficos de controle para o monitoramento de água pode ser útil a empresas concessionárias de água, auxiliando na identificação de vazamentos ou horários de maior consumo, além de auxiliar na análise do uso de água doméstico, contribuindo para a promoção de práticas sustentáveis do consumo de água.

3 METODOLOGIA

As abordagens de pesquisa podem ser amplamente classificadas como quantitativa ou qualitativa. Na abordagem quantitativa, uma das principais preocupações é a mensurabilidade, em que um conjunto de variáveis passível de ser mensurado é utilizado para testar hipóteses do modelo. Na abordagem qualitativa, a realidade subjetiva e as interpretações individuais do pesquisador são consideradas para o desenvolvimento da pesquisa (MARTINS, 2012). Dessa forma, a abordagem desta pesquisa se define como quantitativa, por estudar um conjunto de dados reais do processo.

Segundo Bertrand e Fransoo (2002), observam-se na literatura duas classes distintas de pesquisas quantitativas: a pesquisa axiomática e a pesquisa empírica. A primeira classe utiliza estratégias mais racionalistas no desenvolvimento do estudo. Nesse modelo de pesquisa, o foco do estudo é obter soluções para um problema em específico, garantindo que tais soluções consigam esclarecer a estrutura do problema e compreender o comportamento de variáveis específicas do modelo, possibilitando também a manipulação de certas variáveis. As pesquisas empíricas têm características mais interpretativas e são baseadas em observações e experiências da realidade, estando primariamente voltadas a criar modelos que se adequem a processos reais. Em outras palavras, essa linha de pesquisa investiga a aplicação de soluções desenvolvidas por meio de pesquisas teóricas em problemas reais (MORABITO NETO; PUREZA, 2012). Uma vez que o objetivo deste trabalho é desenvolver uma metodologia para a aplicação de gráficos de controle paramétricos e não paramétricos para o monitoramento do consumo diário de água, comparando o desempenho das técnicas e indicando qual obteve melhores resultados, o método adotado se enquadra na categoria de pesquisa axiomática.

A pesquisa axiomática pode ainda ser classificada como normativa ou descritiva (MORABITO NETO; PUREZA, 2012). De acordo com os autores, a pesquisa axiomática normativa visa desenvolver estratégias a fim de encontrar uma solução ótima para um problema ou comparar a aplicação de diferentes estratégias a um mesmo problema. Em geral, esse modelo de pesquisa se baseia em modelos de programação matemática. Enquanto que o objetivo da pesquisa axiomática descritiva, por sua vez, é entender o processo modelado ou explicar suas características. Sua diferença em relação à pesquisa normativa está no desenvolvimento de modelos que descrevam o comportamento do sistema, permitindo uma melhor compreensão dos relacionamentos funcionais entre as variáveis (MORABITO NETO; PUREZA, 2012).

Com base nas características descritas, este trabalho se enquadra na classe normativa da pesquisa axiomática quantitativa, pois pretende comparar estratégias de monitoramento do

consumo de água na cidade de Joinville, utilizando um conjunto de dados reais do processo e empregando os gráficos de controle de Shewhart, EWMA e EWMA-SN.

As técnicas de monitoramento por meio de gráficos de controle se caracterizam como uma pesquisa longitudinal e não experimental (GIL, 2002). Estudos longitudinais têm como característica a coleta dos dados em diferentes períodos de tempo, com objetivo de realizar inferências sobre a evolução do processo. O desenho de pesquisa não experimental não possui atribuições aleatórias, manipulação de variáveis ou utilização de grupos de controle. Os pesquisadores observam o que aconteceu naturalmente sem qualquer interferência. Existem muitas razões para a realização de pesquisas não experimentais, em primeiro lugar, quando as variáveis não são afetadas pela manipulação experimental ou randomização, ou no caso do impedimento da manipulação de certas variáveis, por razões éticas. Em alguns outros casos, também, o evento causador da variável independente já ocorreu e não pode ser controlado (SILVEIRA; CÓRDOVA, 2009).

3.1 AQUISIÇÃO E FORMATAÇÃO DOS DADOS

O objetivo principal desta pesquisa consiste na realização de um estudo comparativo entre a aplicação de gráficos de controle paramétricos e não paramétricos para o monitoramento de séries temporais do consumo diário de água na cidade de Joinville, em Santa Catarina. Os métodos utilizados foram os gráficos de controle de Shewhart, EWMA e a alternativa não paramétrica EWMA-SN. Foi investigado o consumo diário de água total medido por sistema de telemetria no município, classificando-o conforme as categorias de uso predefinidas pelo sistema de medição implantado no município, e o consumo de água de dois conjuntos habitacionais individualmente, a fim de analisar a eficácia dos gráficos de controle na detecção de vazamentos na escala edificada. Embora as análises realizadas nesta pesquisa estejam considerando os dados na escala municipal como a demanda total do município de Joinville, vale ressaltar que o sistema de telemetria não atende a todas os consumidores da cidade, de modo que na categoria residencial são computados dados de 424 unidades assistidas pelo sistema, na categoria comercial são 166 unidades e nas categorias industrial e público, 42 e 70 unidades, respectivamente. A categoria público contempla edificações que fornecem serviços destinados ao público. Incluem-se nesta categoria edifícios administrados por órgãos públicos, escolas, hospitais, entre outros.

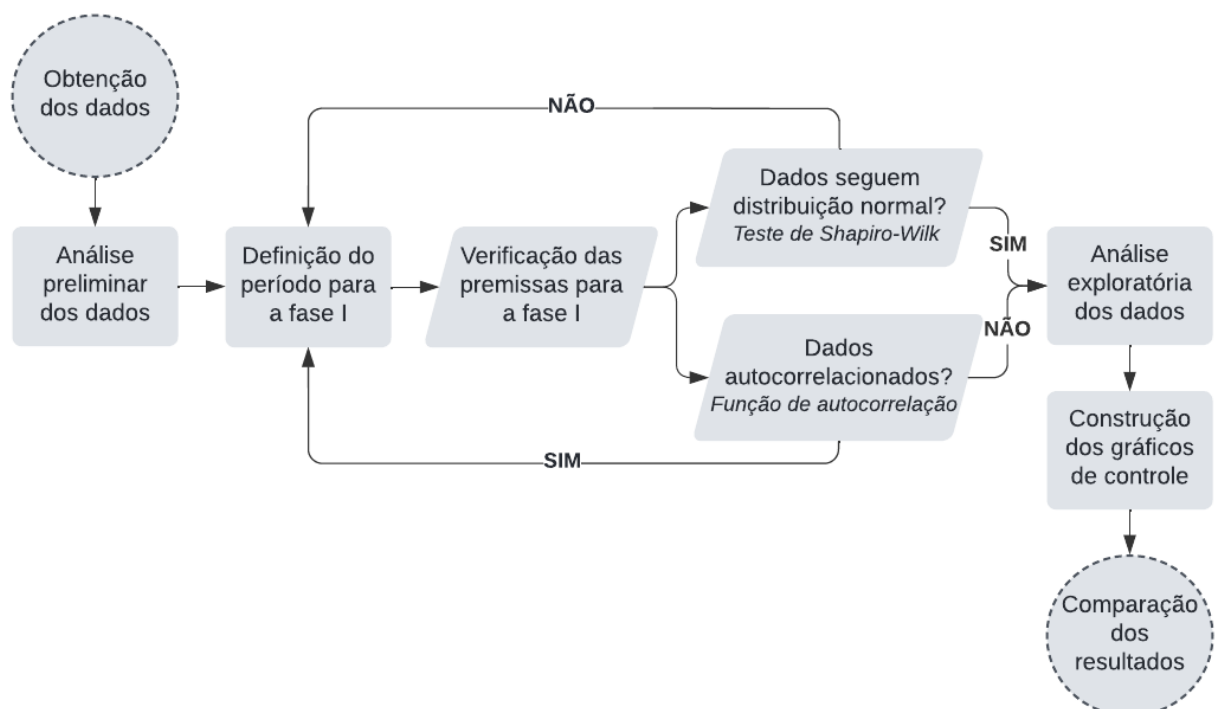
A obtenção dos dados de consumo de água utilizados neste estudo foi realizada junto à companhia responsável pelo abastecimento de água do município de Joinville. Esses dados são

decorrentes de medição por sistema de telemetria, o qual realiza a captação dos dados do consumo de água em intervalos subdiários de cada unidade consumidora atendida pelo sistema e envia a um banco de dados por meio da *internet* local. Como estudo retrospectivo, foram analisados dados previamente medidos pelo sistema, os quais foram agrupados em uma série temporal representando o consumo diário de água total medido por sistema de telemetria.

Os dados de consumo de escala municipal foram analisados conforme os usos residencial, comercial, industrial ou público, previamente especificados pela companhia de abastecimento de água do município de acordo com a atividade principal que contribui para o consumo de água de cada edificação. Essa divisão permitiu investigar o comportamento do padrão do consumo de água por segmento, possibilitando a validação de estudos dentro dessa temática e fornecendo suporte para novas pesquisas.

Os procedimentos metodológicos utilizados para condução desta pesquisa foram divididos em etapas, conforme mostra a Figura 11. Dessa maneira, é possível visualizar todas as etapas precedentes ao estudo comparativo entre os gráficos de controle paramétricos e não paramétricos. Esses procedimentos estão detalhados neste capítulo e, em linhas gerais, incluem a análise preliminar dos dados, seleção de um conjunto de dados para estimação dos parâmetros dos gráficos de controle, análise exploratória da série temporal, construção dos gráficos de controle e, por fim, comparação dos resultados obtidos em cada análise.

Figura 11 – Fluxograma dos procedimentos metodológicos conduzidos nesta pesquisa



Fonte: Elaborado pelo autor (2022)

Após a aquisição e classificação dos dados, realizou-se a análise preliminar e filtragem dos conjuntos de dados. Nessa etapa, após análise minuciosa das planilhas, houve a necessidade de supressão de dados em que o consumo apresentava comportamento discrepante, com a indicação de valores nulos ou expressivamente superiores aos demais, possivelmente causados por uma manutenção no sistema de telemetria, em que a captação foi paralisada e a medição subsequente acumulou o consumo de dias ou até mesmo semanas anteriores. A imputação dos valores faltantes foi realizada a partir de interpolação linear, por meio de função específica do pacote *imputeTS* (MORITZ; BARTZ-BEIELSTEIN, 2017), com auxílio do *software R* (R CORE TEAM, 2020).

Em seguida, com base na série retrospectiva de dados, foi investigada a definição dos períodos para as fases I e II de construção dos gráficos de controle. Conforme Montgomery (2016) aponta, para a delimitação da fase I de aplicação dos gráficos de controle paramétricos, é exigido que o conjunto de dados atenda a certas premissas específicas, do contrário, não é possível inferir conclusões relacionadas ao estado estatístico do processo. Dessa forma, iniciou-se a etapa de investigação de um conjunto de dados compreendido na série temporal que atendesse a estas suposições. Essa etapa somente foi finalizada após a obtenção de um período sob condições estáveis e representativas do processo sob controle para todos os conjuntos de dados, tanto para os dados de consumo dos conjuntos habitacionais como para os dados de consumo em escala municipal.

Após realizada a investigação, análise e teste de todas as séries temporais, identificou-se um período que atendesse simultaneamente a estas suposições para todos os conjuntos de dados. O restante da série foi utilizado para a fase II, denominada monitoramento. O detalhamento das análises realizadas nesta etapa está discutido e explicado na seção 3.3.1. Ao final dessa etapa, a série temporal definida representava os dados do consumo diário de água compreendidos no período entre 01/04/2021 e 29/11/2021, totalizando 243 dias, com a fase I sendo delimitada pelos primeiros 30 dias da série, representando o mês de abril.

A etapa seguinte à definição das séries temporais envolveu a análise exploratória dos dados, visando a obtenção de informações detalhadas a respeito do consumo de água nas escalas municipal e edificada. Foram estimadas medidas descritivas relativas ao consumo de água para todas as séries temporais, verificando suas peculiaridades por meio do cálculo de medidas numéricas e da construção de gráficos descritivos (DEVORE, 2016).

Definidos os períodos para as fases I e II e realizada a análise exploratória dos dados, a próxima etapa consistiu na comparação dos resultados obtidos a partir da aplicação dos gráficos de controle paramétricos e não paramétricos nas diferentes séries temporais estabelecidas,

identificando qual técnica melhor se adequou ao monitoramento do consumo de água, seja na escala municipal ou edificada, verificando a velocidade de detecção de mudanças no padrão do consumo de água, sinalização de alarmes falsos e o efeito da sazonalidade na aplicação dos gráficos de controle.

Todas as análises e aplicações foram realizadas com auxílio do *software* R (R CORE TEAM, 2020) e dos pacotes *qcc* (SCRUCCA, 2004), *ggplot2* (WICKHAM, 2016), *forecast* (HYNDMAN; KHANDAKAR, 2008) e *imputeTS* (MORITZ; BARTZ-BEIELSTEIN, 2017).

3.2 ANÁLISE EXPLORATÓRIA DOS DADOS

A análise exploratória de dados se concentra principalmente na verificação dos pressupostos exigidos para o ajuste de modelos e o teste de hipóteses, além de detectar comportamentos anômalos do processo, manusear valores faltantes e fazer transformações de variáveis quando necessário (CHATFIELD, 1995). Neste estudo, essa etapa concentrou-se na aplicação de técnicas estatísticas capazes de descrever e sumarizar os dados do consumo de água em ambas escalas municipal e edificada. Foram avaliados os padrões do consumo de água para todas as séries temporais estabelecidas, a partir do cálculo de medidas numéricas, como médias, medianas e desvios padrão, e do uso de técnicas visuais, em que foram utilizados os gráficos *boxplot* e os gráficos de séries temporais.

O gráfico *boxplot* é uma ferramenta gráfica utilizada para representar a variação nos dados por meio de quartis (NAVIDI, 2014). Conforme o autor destaca, o *boxplot* apresenta uma reta, indicando a mediana do conjunto de dados, e os quartis superior e inferior, além de indicar a presença de *outliers*, calculados multiplicando-se o primeiro e terceiros quartis por 1,5. Em linhas gerais, o *boxplot* identifica o intervalo em que estão localizados 50% dos valores do conjunto de dados, a mediana e os valores atípicos (ROSS, 2009; NAVIDI, 2014; MANN, 2015).

O gráfico de séries temporais é utilizado para visualizar uma tendência nos dados em intervalos de tempo, auxiliando na visualização de padrões ao longo do tempo, como a existência de tendências e padrões sazonais (MANN, 2015). De acordo com o autor, gráficos de séries temporais exigem que os dados sejam coletados em intervalos regulares, como uma vez por dia ou uma vez por mês. Dados coletados em intervalos irregulares não conseguem representar graficamente o comportamento do processo (MANN, 2015).

A análise exploratória dos dados possibilitou explorar e investigar estatisticamente o comportamento dos padrões do consumo de água nas escalas urbana e edificada.

Oportunamente, a classificação do consumo de água municipal por categoria de uso proporcionou a investigação individual de cada segmento, facilitando a visualização e comparação dos resultados das aplicações dos gráficos de controle para o monitoramento do consumo de água por categoria.

3.3 APLICAÇÃO DOS GRÁFICOS DE CONTROLE

3.3.1 Avaliação das premissas para definição da fase I

Para a aplicação dos gráficos de controle paramétricos foi confirmada a suposição de aderência à distribuição normal do conjunto de dados, exigida para o correto comportamento dos gráficos de controle EWMA e Shewhart, esse último exigindo também a verificação da independência dos dados, a qual foi realizada por meio do teste de autocorrelação.

A verificação da normalidade foi realizada a partir do teste de hipóteses de Shapiro-Wilk, por meio da aceitação ou rejeição da hipótese nula de que os dados seguem uma distribuição normal. Esse é um dos testes de normalidade mais populares existentes, baseado na função de distribuição empírica (WADSWORTH, 2001). Em estudo realizado por Seier (2002), comparando diversos testes de normalidade disponíveis, o teste de Shapiro-Wilk apresentou os melhores resultados na maioria das situações.

Quando os dados não apresentam aderência à distribuição normal, podem ser utilizadas duas técnicas com a finalidade de trazer os dados para a normalidade: a primeira envolve a utilização de técnicas de transformação, dentre as quais, as mais comumente utilizadas são a logarítmica, a raiz quadrada e a raiz cúbica (MONTGOMERY, 2016). A segunda técnica é a aplicação de um gráfico de controle não paramétrico, o qual não exige a pressuposição de normalidade (CHAKRABORTI; VAN DER LAAN; BAKIR, 2001).

No presente estudo, investigou-se o desempenho de gráficos de controle não paramétricos para o monitoramento do consumo de água a partir da aplicação do gráfico proposto por Graham, Chakraborti e Human (2011), o qual envolve a estatística de sinais aplicada a um gráfico de controle EWMA, denominado EWMA-SN. Essa técnica apresenta a vantagem de não exigir o valor numérico das observações, apenas a indicação de se elas são maiores ou menores que um parâmetro de interesse estabelecido. Dessa forma, para que seja possível a aplicação dessa técnica, é exigido que um valor de controle de referência seja conhecido ou especificado (GRAHAM; CHAKRABORTI; HUMAN, 2011).

Diferentemente dos gráficos de controle de Shewhart e EWMA, não existe disponível uma linguagem de programação desenvolvida para aplicação do gráfico de controle não paramétrico EWMA-SN. Dessa forma, foi adaptada uma rotina em ambiente R capaz de calcular sua base estatística e limites de controle a partir do código já existente no pacote qcc (SCRUCCA, 2014) para aplicação do gráfico de controle EWMA. As modificações realizadas no código estão expostas no Quadro 1. Em suma, as modificações envolvem a mudança do cálculo dos limites de controle para o padrão específico do gráfico de controle EWMA-SN, quantificados a partir da definição do valor de λ , e a inclusão da condicional referente ao cálculo da estatística de plotagem do gráfico, obtida acumulando-se sequencialmente a informação do sinal da observação atual em relação ao parâmetro de interesse sendo monitorado.

Quadro 1 – Rotina adaptada para aplicação do gráfico de controle EWMA-SN

```
ewmaSmoothSN <- function (x, y, w, lambda = 0.2, start, ...)
{
  ord <- order(x)
  x <- x[ord]
  y <- y[ord]
  n <- length(y)
  med <- median(y[w])
  if (missing(start))
    start <- y[1]
  z <- c(start, y)
  for (i in 2:(n + 1)) z[i] <- (ifelse(y[i-1] < med, -lambda,
    ifelse(y[i-1] > med, lambda, 0))) + (1 - lambda) * z[i - 1]
}

ewmasn <- function (data, sizes, center, std.dev, lambda = 0.2,
  nsigmas = 3, parameter, data.name, labels, newdata,
  newsizes, newlabels, plot=TRUE, ...)
{
  size <- 1:length(data)
  ewma <- ewmaSmoothSN(indices, statistics, size, lambda = lambda,
    start = center)
  sigma2 <- ((lambda/(2 - lambda)) * (1 - (1 - lambda)^(2 * (1:n))))
  ucl <- nsigmas * sqrt(sigma2)
  lcl <- - nsigmas * sqrt(sigma2)
}
```

Fonte: Elaborado pelo autor (2022)

O valor de controle de referência adotado neste estudo foi a mediana, a qual apresenta uma melhor representação do valor central e uma maior robustez em relação à média, além de se igualar à mesma quando a distribuição é simétrica (GRAHAM; CHAKRABORTI; HUMAN, 2011).

3.3.2 Fase II de aplicação dos gráficos de controle

Após a fase de diagnóstico, foi iniciada a fase correspondente ao período de monitoramento. Os gráficos de controle foram aplicados às séries temporais do consumo diário de água nas escalas municipal e edificada, analisando o comportamento dos processos nas diferentes aplicações. A construção dos limites de controle foi verificada quanto ao atendimento às tendências de sazonalidade no padrão do consumo de água, observando as variações semanais na demanda de água.

Destaca-se que apesar de o gráfico de controle não paramétrico EWMA-SN não exigir o atendimento do conjunto de dados a quaisquer suposições de simetria ou independência da amostra, possibilitando a escolha de qualquer período da série temporal para a fase I, optou-se por utilizar os mesmos períodos para as fases I e II em todas as análises gráficas, com o intuito de auxiliar na comparação dos resultados e facilitar a visualização do comportamento dos gráficos de controle quando aplicados às diferentes escalas e categorias de consumo de água.

Isto posto, a etapa final consistiu na comparação dos resultados obtidos a partir das aplicações dos gráficos de controle de Shewhart, EWMA e EWMA-SN. Nesta etapa, os desempenhos dos gráficos de controle foram comparados verificando-se qual técnica detectou mais rapidamente uma situação de alteração no processo e apresentou melhores índices de sinalização de alarmes falsos.

3.4 ANÁLISE COMPARATIVA

Albarracín (2014) destaca que quando se tem como objetivo avaliar o desempenho de gráficos de controle e comparar diferentes técnicas, pode-se considerar os valores da probabilidade de ocorrência dos erros tipo I (α) e II (β). No entanto, conforme o autor ressalta, outro parâmetro comumente utilizado é a verificação do número de observações necessárias até que um alarme falso seja sinalizado, denominado ARL.

Neste estudo, a análise do desempenho dos gráficos empregados foi feita a partir da comparação dos ARL_0 , ou seja, a frequência de alarmes falsos. Dessa forma, os valores de L , que representam a distância dos limites de controle até a linha central, e de λ , que representam a memória dos gráficos de controle EWMA e EWMA-SN, foram determinados de modo a se obter um mesmo $ARL_0 \cong 370$ para todas as análises. Essa definição é amplamente utilizada na área industrial e facilita a comparação entre os gráficos de controle (SAMOHYL, 2009). Visto

que cada observação representa o consumo total de água por dia, esta definição consiste na sinalização de um alarme falso a cada 370 dias, ou seja, aproximadamente um ano.

Para obtenção do valor alvo do ARL_0 , utilizou-se o valor de $L = 3$ para o gráfico de controle de Shewhart (MONTGOMERY, 2016) e as combinações de $\lambda = 0,2$ e $L = 2,86$ para o gráfico EWMA (BORROR; MONTGOMERY; RUNGER, 1999) e $\lambda = 0,10$ e $L = 2,585$ para o gráfico de controle não paramétrico EWMA-SN (GRAHAM; CHAKRABORTI; HUMAN, 2011).

4 RESULTADOS E DISCUSSÃO

Neste capítulo são apresentados e discutidos os resultados obtidos a partir da aplicação da metodologia desenvolvida nesta pesquisa aos conjuntos de dados especificados. Seguindo os procedimentos propostos, foram estudados os dados do consumo de água de dois conjuntos habitacionais individualmente, assim como os dados do consumo diário de água total do município de Joinville medido por sistema de telemetria, analisando ainda as categorias de consumo residencial, comercial, industrial e público.

Conforme comentado no capítulo 3.1, após avaliação das planilhas, houve a necessidade de supressão e interpolação de dados em que o consumo apresentava valores discrepantes, desde valores nulos a valores expressivamente superiores aos demais. Após a filtragem e investigação dos conjuntos de dados em relação ao atendimento às premissas para aplicação dos gráficos de controle paramétricos, a série temporal definida representava os dados do consumo diário de água delimitados entre os dias 01/04/2021 e 29/11/2021, totalizando 243 dias. Os primeiros 30 dias da série temporal foram definidos para a fase I de aplicação dos gráficos de controle de Shewhart, EWMA e EWMA-SN, visto que atendiam às premissas exigidas para aplicação dos gráficos de controle convencionais. Essa definição foi aplicada tanto para os conjuntos de dados de escala municipal quanto para os conjuntos de dados dos condomínios residenciais.

4.1 CONJUNTOS HABITACIONAIS

Esta seção apresenta os resultados obtidos a partir da análise exploratória dos dados e da aplicação da metodologia proposta às séries temporais do consumo de água dos dois conjuntos habitacionais. A fim de garantir a privacidade e segurança das informações pessoais e particulares de cada unidade, os conjuntos habitacionais foram denominados como Condomínio A e Condomínio B. Ambos foram escolhidos junto à empresa concessionária de água do município por terem apresentado vazamentos de água previamente detectados pelo sistema de telemetria. O vazamento no Condomínio A ocorreu entre os meses de junho e agosto de 2021 e no Condomínio B, entre os meses de setembro e novembro de 2021.

O Condomínio A representa um condomínio residencial localizado em região central do município, possui área de 15.566 m² e contém 11 residências de alto padrão. O Condomínio B corresponde a um conjunto habitacional de interesse social localizado em região periférica do município, possui 16.989,01 m² de área construída, distribuídos em 20 blocos de 4 pavimentos e 16 apartamentos cada, totalizando 320 apartamentos no empreendimento.

Inicialmente, os dados de consumo dos conjuntos habitacionais apresentavam escala temporal de medição em intervalos irregulares de horas. Com o objetivo de suprimir valores de medição errôneos e padronizar o intervalo de leitura, essas medições foram agrupadas em intervalos de 24 horas. Dessa forma, foi considerado o consumo diário de água para todas as análises, em m³/dia.

4.1.1 Análise das séries temporais

Os padrões de consumo diário de água dos dois conjuntos habitacionais foram inicialmente investigados por meio da análise exploratória global dos dados, contemplando integralmente a série temporal. A Tabela 1 traz as informações das medidas numéricas de posição e dispersão do consumo diário de água dos dois condomínios para os 243 dias analisados nesta pesquisa.

Tabela 1 – Estatísticas descritivas dos conjuntos habitacionais

Descrição	Média (m ³ /dia)	Mediana (m ³ /dia)	Mínimo (m ³ /dia)	Máximo (m ³ /dia)	Desvio padrão (m ³ /dia)	Coefficiente de variação
Condomínio A	20,19	20,90	4,10	44,00	7,87	38,96%
Condomínio B	153,60	141,10	83,30	330,70	36,13	23,52%

Fonte: Elaborado pelo autor (2022)

Os resultados obtidos a partir das análises descritivas mostram que a amplitude dos dados, em ambos os condomínios, sugere a existência de alterações no padrão do consumo de água ao longo da série temporal estudada. Os valores máximos correspondem aos períodos em que ocorreram os vazamentos de água e os valores mínimos representam períodos em que o consumo de água é reduzido, como finais de semanas, feriados, recessos trabalhistas e férias escolares.

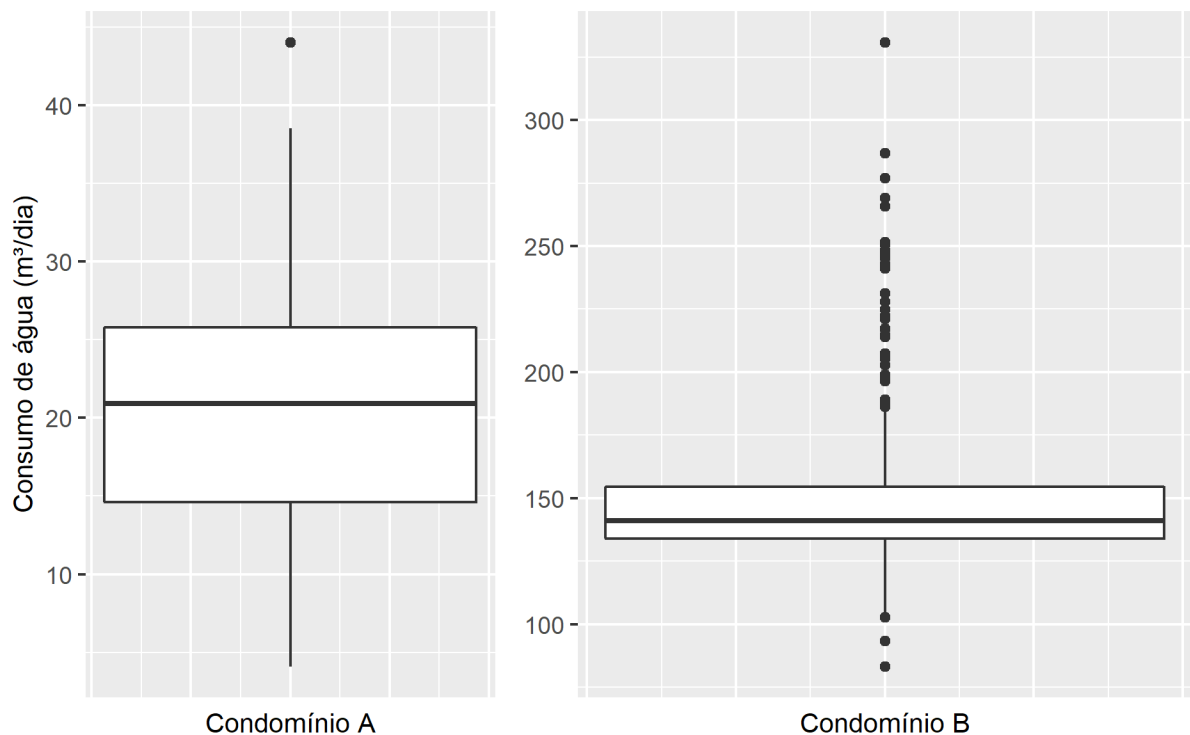
Em relação às medidas de posição para o Condomínio A, a proximidade dos valores de média e mediana poderiam indicar uma simetria na distribuição. Contudo, essa suposição é contrariada ao se analisar a dispersão dos dados, observando-se os valores de desvio padrão e coeficiente de variação em torno da média.

Para o Condomínio B, a média do consumo diário de água apresenta indícios de ser superior à mediana, possivelmente influenciada pela ocorrência do vazamento de água na série temporal. Comparado ao Condomínio A, os resultados revelam uma menor variabilidade dos

dados em torno da média, porém ainda expressiva, apresentando um coeficiente de variação igual a 23,52%.

Com o auxílio dos gráficos *boxplot* (Figura 12), é possível observar que 50% dos valores de consumo para o Condomínio A estão entre 14,60 e 25,80 m³/dia, referentes ao primeiro e terceiro quartis, respectivamente. O Condomínio B apresenta 50% dos valores de consumo no intervalo entre 133,9 e 154,5 m³/dia. Ambos os condomínios apresentam *outliers* nos gráficos *boxplot*, a presença destes *outliers* pode indicar que o conjunto de dados apresenta um comportamento discrepante ao longo da série temporal. Esse comportamento será monitorado pelos gráficos de controle, a aplicação dessa técnica poderá definir se esse comportamento está dentro do controle estatístico ou se os valores atípicos de consumo podem ser classificados como fora do controle estatístico.

Figura 12 – *Boxplots* dos consumos de água dos conjuntos habitacionais

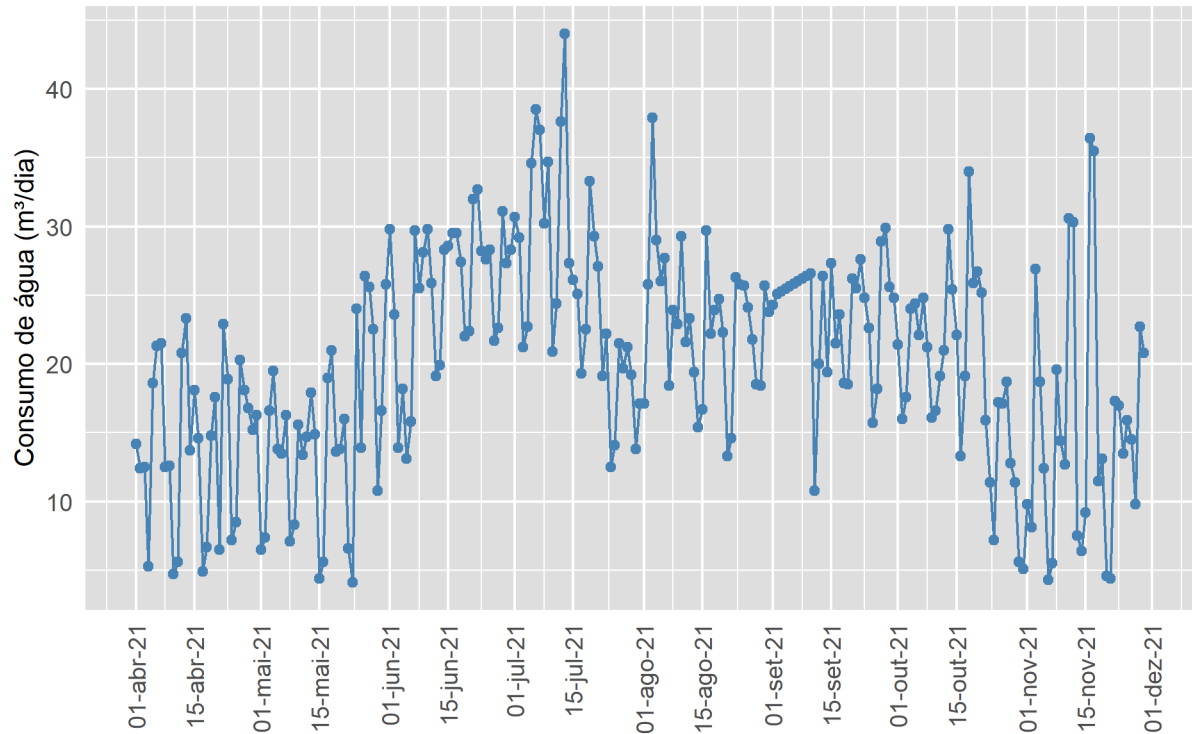


Fonte: Elaborado pelo autor (2022)

A Figura 13 ilustra o gráfico de séries temporais para o consumo diário de água do Condomínio A, no qual é possível visualizar a ocorrência do vazamento de água entre os meses de junho e agosto de 2021. Percebe-se também que o padrão do consumo de água do

Condomínio A apresenta uma estabilidade nos intervalos semanais, mantendo-se consistente durante os dias úteis, com uma queda aos finais de semana.

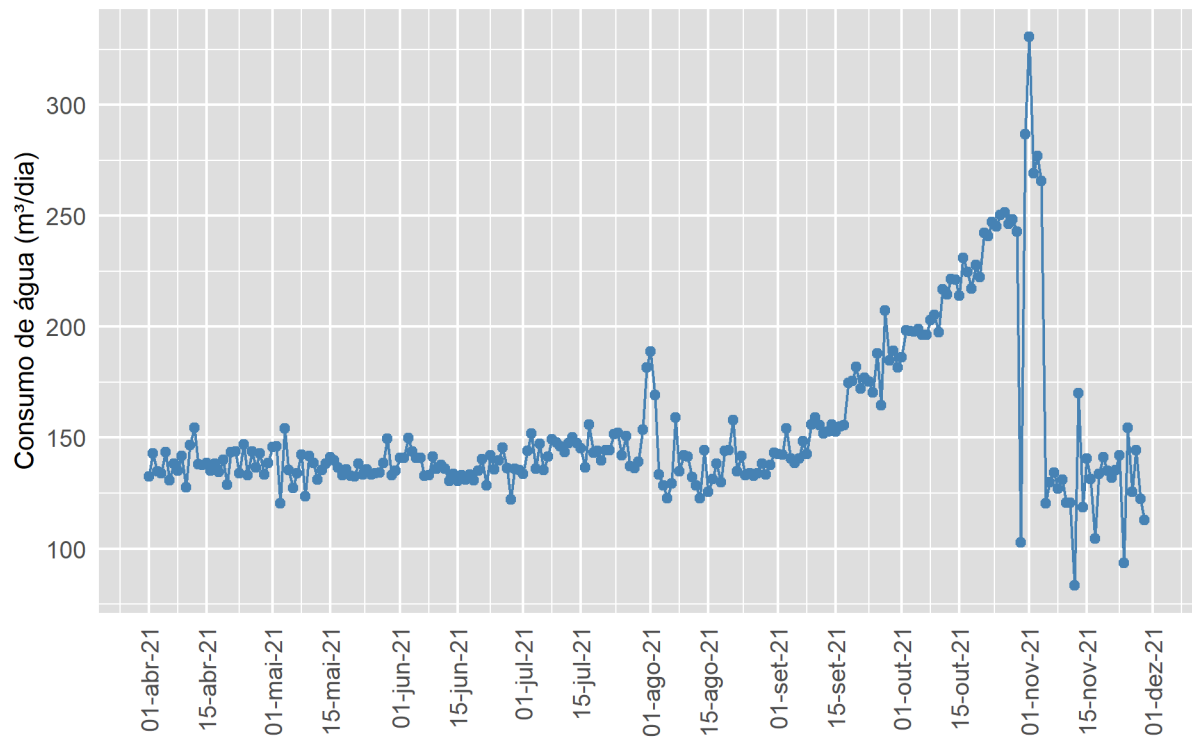
Figura 13 – Série temporal do consumo de água do Condomínio A



Fonte: Elaborado pelo autor (2022)

O período de ocorrência do vazamento de água no Condomínio B é caracterizado pelo aumento expressivo no valor das observações entre os meses de setembro e novembro de 2021, conforme pode ser observado no gráfico de séries temporais (Figura 14). A partir da segunda quinzena de outubro, os valores baixos de consumo seguidos de uma observação alta podem representar uma manutenção no sistema de telemetria, em que a medição subsequente acumulou um valor de consumo não mensurado no dia anterior.

Figura 14 – Série temporal do consumo de água do Condomínio B



Fonte: Elaborado pelo autor (2022)

4.1.2 Avaliação das premissas

O período definido para a fase I, de estimação dos parâmetros e construção dos limites de controle, foi delimitado pelos primeiros 30 dias da série temporal. Esse período representa o mês de abril de 2021 e atende satisfatoriamente às suposições de normalidade e independência dos dados.

Atenta-se que, quando o processo sendo monitorado apresenta características sazonais, a definição de um período específico para a fase I pode resultar em um comportamento tendencioso dos gráficos de controle, uma vez que esse conjunto de dados pode não estar considerando o comportamento do processo em outros períodos da série temporal. Especificamente em relação aos dados sendo analisados nesta pesquisa, apesar da utilização do mês de abril para construção dos limites de controle considerar as variações semanais, essa definição pode não representar o comportamento do consumo de água em outros meses, o que pode influenciar a sinalização de mudanças significativas nos gráficos de controle na fase de monitoramento.

Para verificação da normalidade dos dados aplicou-se o teste de hipóteses de Shapiro-Wilk (Tabela 2), em que ambos os conjuntos de dados retornaram um $p\text{-valor} > 0,05$, aceitando a hipótese nula de que os dados seguem uma distribuição normal. Dessa forma, as premissas exigidas para aplicação dos gráficos de controle EWMA e Shewhart foram satisfeitas, esse último exigindo ainda a verificação da independência dos dados.

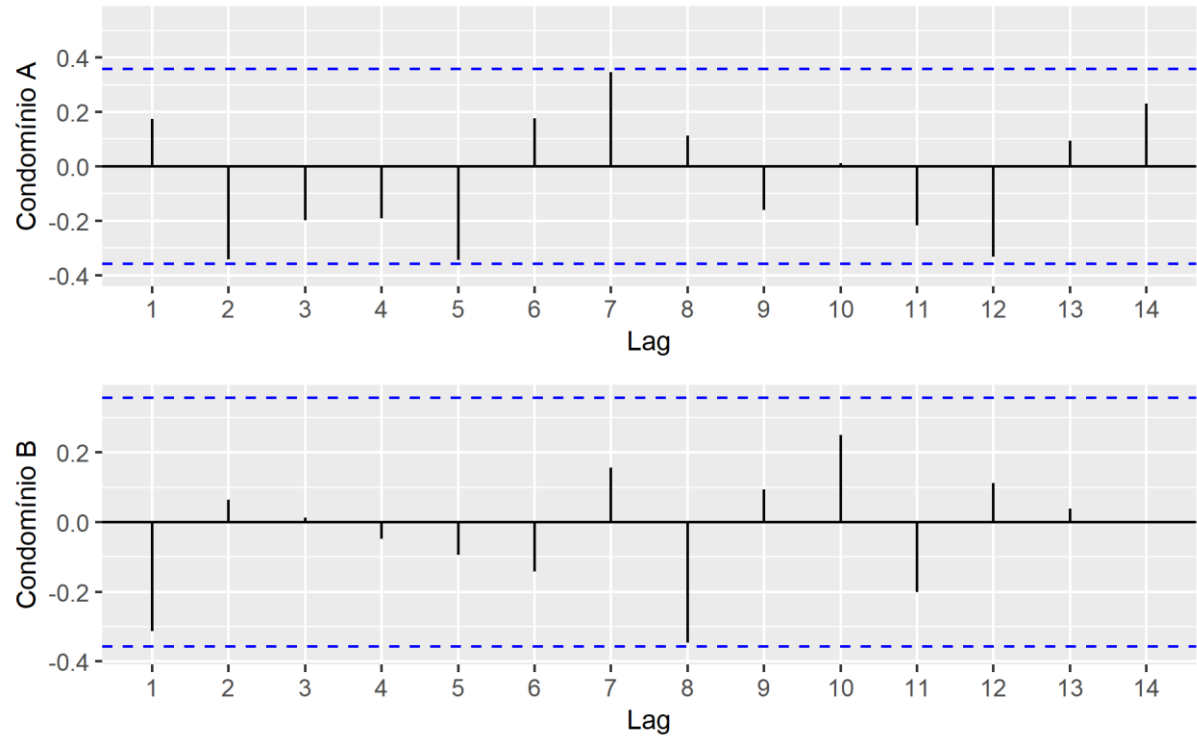
Tabela 2 – Resultados dos testes de Shapiro-Wilk para os conjuntos habitacionais

Descrição	p-valor
Condomínio A	0,071
Condomínio B	0,534

Fonte: Elaborado pelo autor (2022)

A independência dos dados foi aferida aplicando-se a função de autocorrelação aos conjuntos de dados (Figura 15). Uma vez que as observações não extrapolam o intervalo de confiança, é confirmada a independência dos dados, permitindo a aplicação do gráfico de controle de Shewhart aos conjuntos de dados.

Figura 15 – Funções de autocorrelação para os conjuntos habitacionais



Fonte: Elaborado pelo autor (2022)

4.1.3 Aplicação dos gráficos de controle

Esta seção aborda o desempenho dos gráficos de controle de Shewhart, EWMA e EWMA-SN para o monitoramento do consumo diário de água dos dois conjuntos habitacionais. Os resultados obtidos são discutidos individualmente por condomínio e o desempenho dos gráficos foi comparado com base nos valores do ARL_0 de cada técnica, ou seja, a frequência de alarmes falsos.

4.1.3.1 Gráficos de controle para o Condomínio A

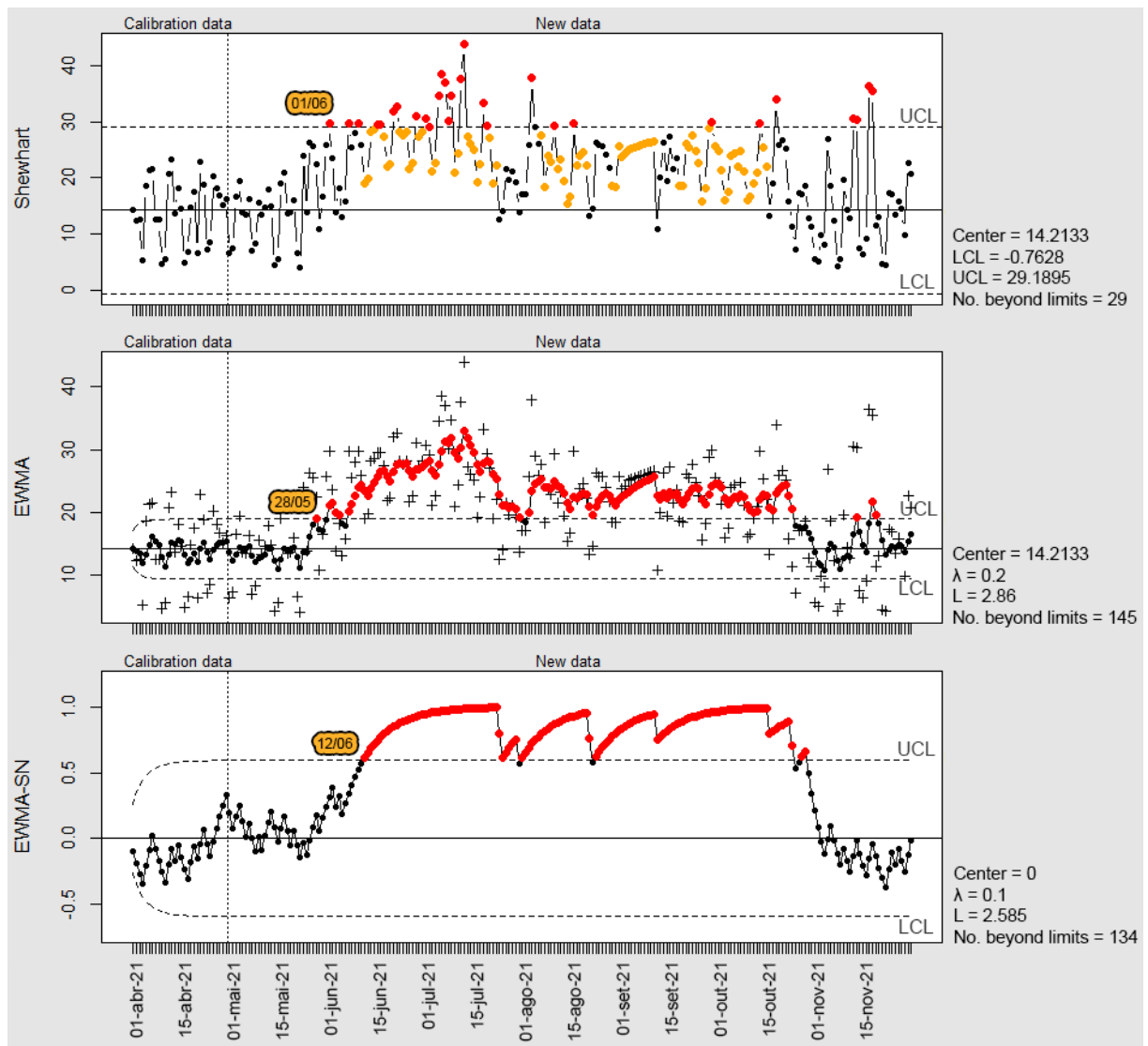
A construção do gráfico de controle de Shewhart para o Condomínio A resultou em 29 observações plotadas além dos limites de controle, indicando que o consumo de água estava fora do controle estatístico nesses pontos (Figura 16). Além da análise convencional dos pontos excedendo os limites de controle, o gráfico também sinalizou períodos em que o processo violou o conjunto de regras da *Western Electric*, principalmente referente à plotagem de pontos do mesmo lado da linha central ou a uma distância de pelo menos 1σ .

O gráfico de Shewhart sinalizou uma mudança no processo a partir do dia 01/06/21, correspondendo ao período de ocorrência do vazamento de água informado pela empresa concessionária de água. A partir do mês de setembro, os pontos plotados além do LSC correspondem a datas isoladas e podem estar relacionados a mudanças causadas por variáveis meteorológicas e climáticas, que influenciam o padrão do consumo de água, ou até mesmo alarmes falsos (BALLING JR.; GOBER, 2007; BALLING JR.; GOBER; JONES, 2008; CHANG; PRASKIEVICZ; PARANDVASH, 2014). Estas observações apresentam menor magnitude e recorrência em relação ao período de ocorrência do vazamento de água. Dessa forma, uma vez que a fuga do controle estatístico é discreta e passageira, a suposição de ocorrência de um vazamento poderia ser descartada, reiterando a eficácia do gráfico de controle de Shewhart como técnica para o monitoramento do consumo de água.

Em relação ao gráfico de controle EWMA, sua aplicação resultou em 145 pontos sinalizados além do LSC, concentrados principalmente no intervalo de ocorrência do vazamento de água no condomínio, porém se estendendo até o mês de outubro, motivados principalmente pelo acúmulo do peso de observações antigas na estatística de plotagem do gráfico (Figura 16). Esta característica também contribuiu para um aumento na sinalização de alarmes falsos, conforme pode ser observado a partir do mês de setembro. Esse período não corresponde à ocorrência de um vazamento de água, porém foram sinalizadas mudanças

menores no consumo de água. Conforme Montgomery (2016) aponta, essa é uma característica própria do gráfico de controle EWMA, sendo muito utilizado para a detecção de mudanças pequenas no processo. O gráfico EWMA sinalizou uma fuga do controle estatístico mais rapidamente que o gráfico de Shewhart, a partir do dia 28/05/21.

Figura 16 – Gráficos de controle para o Condomínio A



Fonte: Elaborado pelo autor (2022)

O gráfico não paramétrico EWMA-SN apresentou comportamento análogo ao gráfico EWMA na sinalização de alarmes falsos, sinalizando 134 observações além dos limites de controle (Figura 16 **Erro! Fonte de referência não encontrada.**). Apesar de essa técnica apresentar a vantagem de não exigir o valor numérico das observações, apenas a indicação de se elas são maiores ou menores que um parâmetro de interesse estabelecido, essa característica

resultou em um atraso na sinalização de uma mudança no processo, uma vez que a magnitude da observação sendo plotada não é considerada pelo gráfico. O gráfico EWMA-SN sinalizou a fuga do controle estatístico no dia 12/06/21, representando um atraso de aproximadamente duas semanas em relação às outras técnicas.

4.1.3.2 Gráficos de controle para o Condomínio B

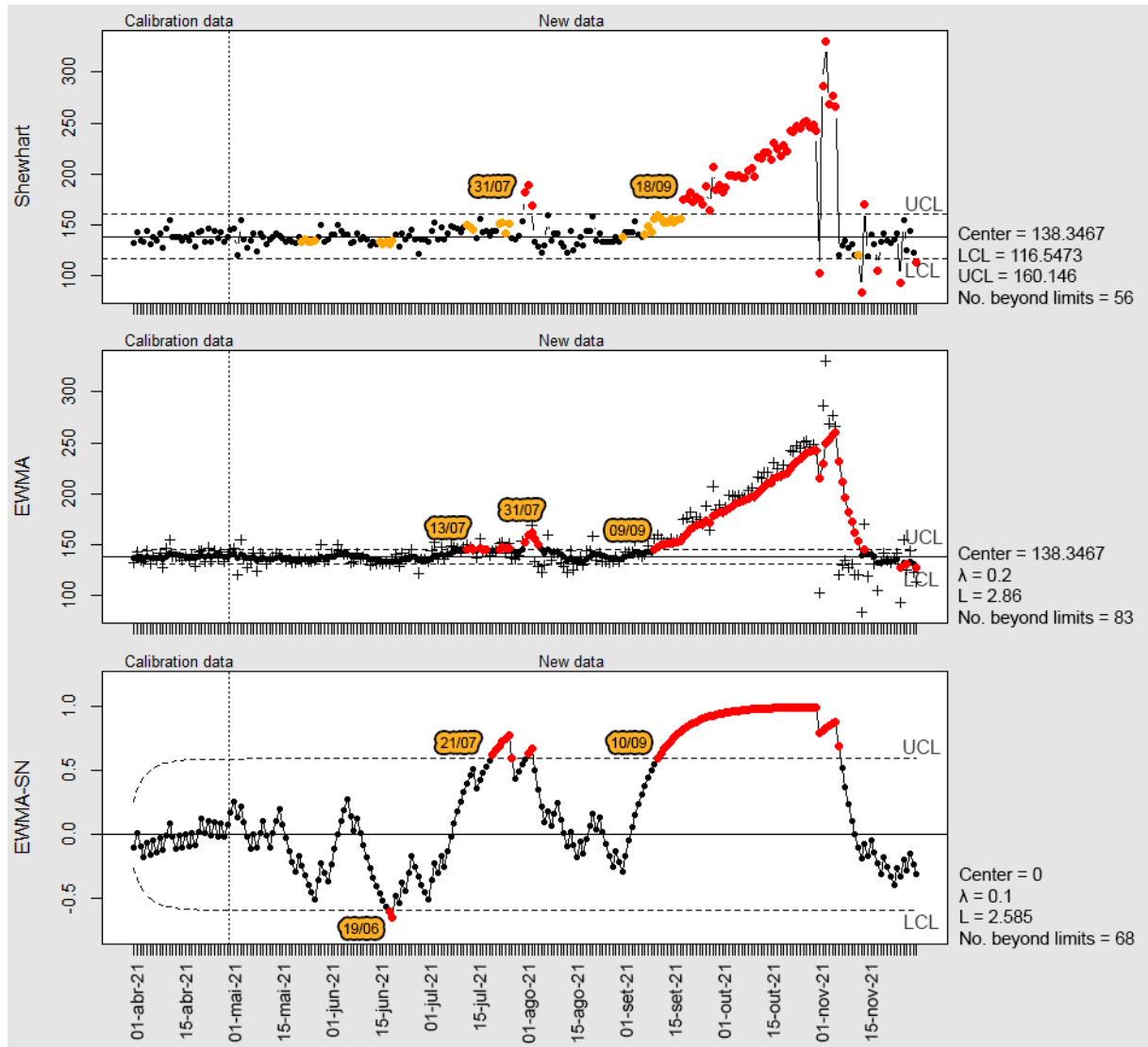
Para o Condomínio B, o gráfico de controle de Shewhart sinalizou 56 pontos fora do controle estatístico, sendo 3 observações no mês de agosto e o restante distribuído no período de ocorrência do vazamento de água, além de 74 observações violando o conjunto de regras da *Western Electric* (Figura 17). Assim como nos resultados obtidos para o Condomínio A, as primeiras observações plotadas além do LSC podem estar relacionadas a variáveis meteorológicas e climáticas (BALLING JR.; GOBER, 2007; BALLING JR.; GOBER; JONES, 2008; CHANG; PRASKIEVICZ; PARANDVASH, 2014) e, a partir da segunda quinzena de outubro, as observações plotadas além dos limites inferior e superior de controle podem representar uma manutenção no aparelho de medição motivada pelo vazamento de água, em que a captação foi paralisada e a medição subsequente acumulou o consumo do dia anterior. O vazamento foi sinalizado no gráfico no dia 18/09/21, mas ao se olhar as regras suplementares, verificam-se indícios de aumento na média de consumo a partir do dia 06/09/21.

O gráfico de controle EWMA para o Condomínio B sinalizou 83 pontos além dos limites de controle, motivados principalmente pelo comportamento não aleatório do consumo de água, conforme pode ser observado a partir do dia 13/07/21 (Figura 17 **Erro! Fonte de referência não encontrada.**). O gráfico EWMA sinalizou o vazamento de água no dia 09/09/21, detectando mais rapidamente que o gráfico de Shewhart, sem a aplicação das regras suplementares. A magnitude das observações referentes ao vazamento de água também causou a sinalização de alarmes falsos após o período de ocorrência do evento, em virtude de o gráfico continuar atribuindo um peso às observações recentes.

Em comparação ao gráfico EWMA, o gráfico não paramétrico EWMA-SN obteve um índice de sinalização de alarmes falsos menor, totalizando 68 pontos (Figura 17 **Erro! Fonte de referência não encontrada.**). A indicação do vazamento de água no gráfico EWMA-SN ocorreu no dia 10/09/21, mesmo período em que houve a sinalização nas outras técnicas. O diferencial dos resultados obtidos a partir do gráfico não paramétrico está na sinalização de duas observações abaixo do LIC no mês de junho, assim como a maior estabilidade do gráfico

após a ocorrência do vazamento de água, não sendo afetado pela possível manutenção no sistema de medição.

Figura 17 – Gráficos de controle para o Condomínio B



Fonte: Elaborado pelo autor (2022)

4.1.4 Análise comparativa

Para o consumo em escala edificada, em ambos os condomínios o gráfico EWMA obteve os melhores resultados de velocidade de detecção dos vazamentos de água. Apesar do gráfico de Shewhart ter detectado o vazamento mais rapidamente considerando-se o conjunto de regras da *Western Electric*, a utilização destas regras implica na alteração do ARL_0 , refletindo em uma maior sinalização de alarmes falsos.

O gráfico de controle EWMA apresentou a maior quantidade de sinalizações de alarmes em ambas as análises, motivadas pelo acúmulo consecutivo do peso de observações com comportamento não aleatório do consumo de água. A utilização da mediana como valor de referência para o gráfico EWMA-SN resultou em uma maior robustez do gráfico a mudanças repentinas no processo. Esse aspecto diminuiu a sinalização de alarmes, atrasando a identificação do vazamento de água no Condomínio A. No Condomínio B não foi observado esse comportamento, uma vez que a sinalização do vazamento de água ocorreu após um período de observações sucessivas acima da LC. Em ambos os condomínios os vazamentos de água foram precedidos por uma sequência de observações acima da LC, e não uma mudança repentina no processo, sugerindo que mudanças pequenas no consumo de água podem representar vazamentos (PATABENDIGE *et al.*, 2018).

4.2 DEMANDA URBANA DE ÁGUA

Esta seção apresenta os resultados obtidos a partir da análise exploratória dos dados e da aplicação da metodologia proposta ao conjunto de dados da demanda diária de água total medida por sistema de telemetria do município, em m³/dia. Os dados de consumo também foram analisados conforme os usos residencial, comercial, industrial e público, especificados de acordo com a atividade principal que contribui para o consumo de água de cada edificação. Conforme apresentado anteriormente, o consumo de água da categoria residencial representa aproximadamente 60% (424) do total de 702 matrículas atendidas pelo sistema de telemetria do município. Dessa forma, visto que neste estudo a demanda total do município é calculada somando-se o consumo de todas as categorias, as análises realizadas nesta pesquisa são influenciadas expressivamente pelo segmento residencial.

A classificação do consumo municipal de água por categoria proporcionou a investigação da demanda de água característica a cada segmento, explorando suas singularidades e fornecendo suporte para novas pesquisas acerca da demanda urbana de água. Esta categorização também facilitou a visualização e comparação dos resultados obtidos a partir das aplicações dos gráficos de controle para o monitoramento do consumo de água em cada segmento.

A investigação da eficácia dos gráficos de controle para a detecção de vazamentos em escala urbana foi realizada com base nas informações de interrupção no abastecimento de água no município, disponíveis na página eletrônica da concessionária de água. Essas publicações permitem investigar a ocorrência de vazamentos de água na série temporal da demanda total do

município, porém não possibilitam a classificação dos vazamentos por categoria de uso tampouco informam a data e local específico de ocorrência dos eventos, existindo a possibilidade de não terem sido captados pelo sistema de telemetria do município, e não estarem contemplados no conjunto de dados analisado nesta pesquisa.

4.2.1 Análise das séries temporais

Inicialmente, os padrões dos consumos de água por categoria e do total medido pelo sistema de telemetria do município foram investigados por meio da análise exploratória dos dados, contemplando os 243 dias da série temporal estudada nesta pesquisa (Tabela 3). A amplitude dos dados e os coeficientes de variação indicam a alta variabilidade dos dados e sugerem a existência de alterações no padrão do consumo de água ao longo da série temporal estudada, as quais podem estar associadas a mudanças no clima e no nível de precipitação. As categorias residencial, comercial e público apresentaram médias de consumo superiores às medianas, indicando que nestes segmentos existiu uma maior concentração de valores acima da média no período estudado.

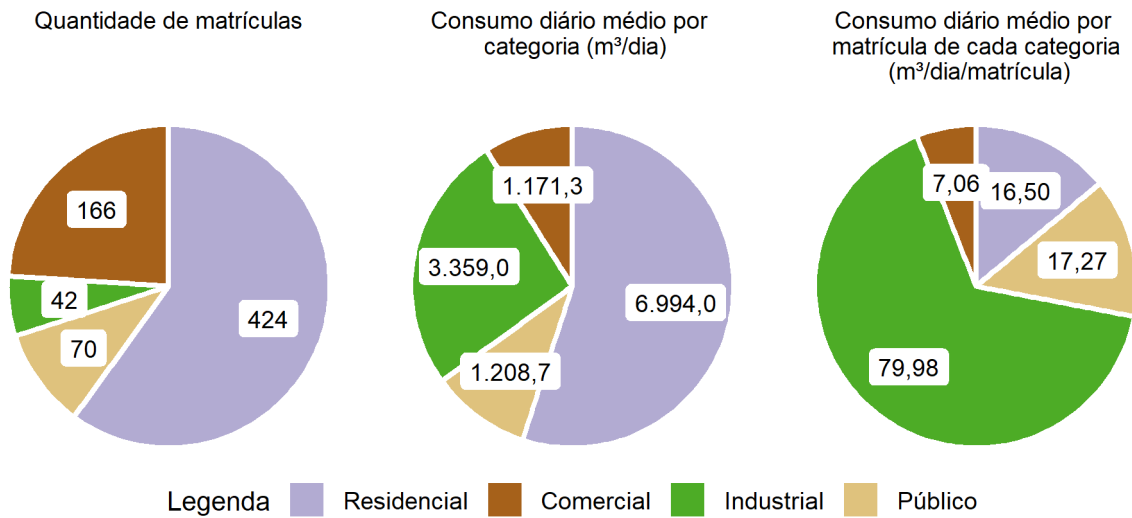
Tabela 3 – Estatísticas descritivas da demanda urbana de água

Descrição	Média (m ³ /dia)	Mediana (m ³ /dia)	Mínimo (m ³ /dia)	Máximo (m ³ /dia)	Desvio padrão (m ³ /dia)	Coeficiente de variação
Residencial	6.994,0	6.861,0	4.474,0	10.912,0	950,89	13,59%
Comercial	1.171,3	1.159,9	620,1	2.523,7	267,60	22,85%
Industrial	3.359,0	3.421,0	1.295,0	7.846,0	930,41	27,70%
Público	1.208,7	1.192,8	444,1	3.202,0	298,94	24,73%
Total	12.733,0	12.760,0	8.560,0	18.107,0	1.662,46	13,06%

Fonte: Elaborado pelo autor (2022)

Ao se investigar o consumo diário de água por matrícula de cada categoria, constata-se que o segmento industrial apresenta um consumo médio por matrícula igual a 79,98 m³/dia/matrícula, enquanto que nas categorias residencial, comercial e público, esse valor corresponde a 16,5 m³/dia/matrícula, 7,06 m³/dia/matrícula e 17,27 m³/dia/matrícula, respectivamente. Cabe ressaltar que uma matrícula pode ter mais de uma unidade consumidora, como é o caso de condomínios de casas, edifícios de apartamentos, edifícios comerciais, parques industriais, entre outros. A Figura 18 auxilia na comparação da quantidade de matrículas, consumo diário médio e consumo diário por matrícula de cada categoria.

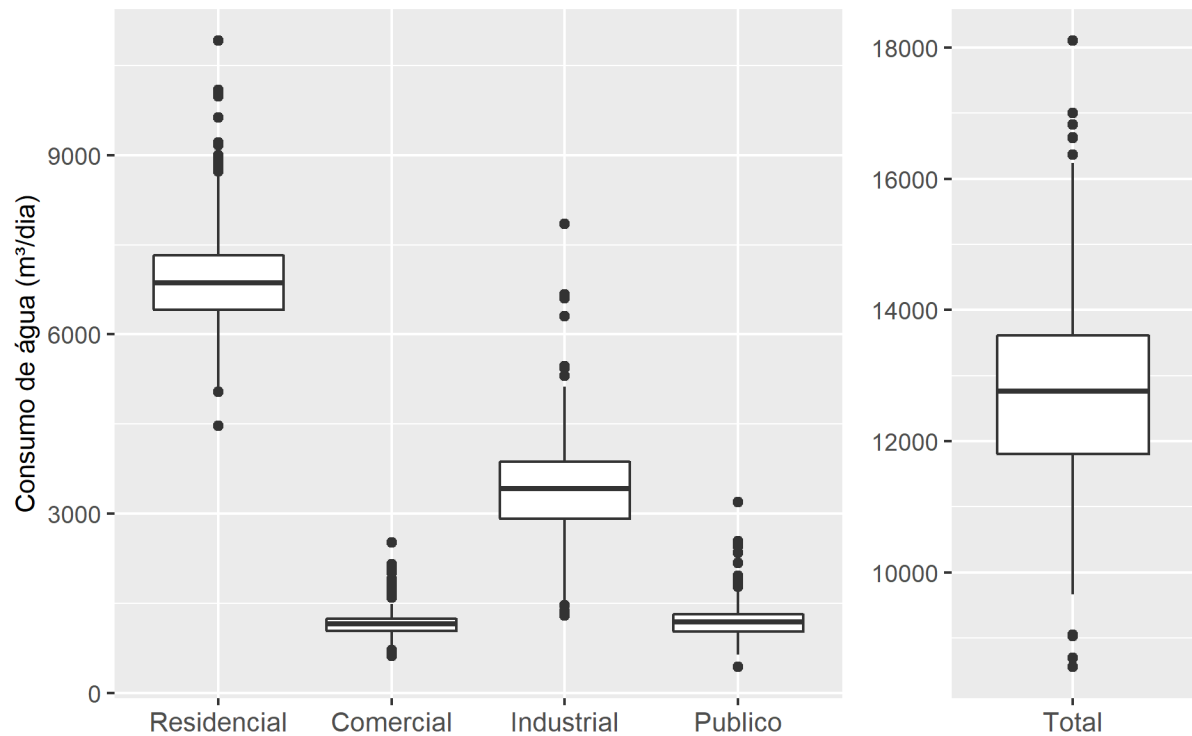
Figura 18 – Quantidade de matrículas e dados de consumo de cada categoria



Fonte: Elaborado pelo autor (2022)

Com o auxílio dos gráficos *boxplot* (Figura 19), é possível observar que todas as categorias apresentam possíveis *outliers*, ou seja, valores atípicos de consumo de água. Conforme mencionado anteriormente, a aplicação dos gráficos de controle poderá definir se esse comportamento está dentro do controle estatístico ou se a presença desses *outliers* pode ser classificada como fora do controle estatístico.

Figura 19 – Boxplots da demanda urbana de água

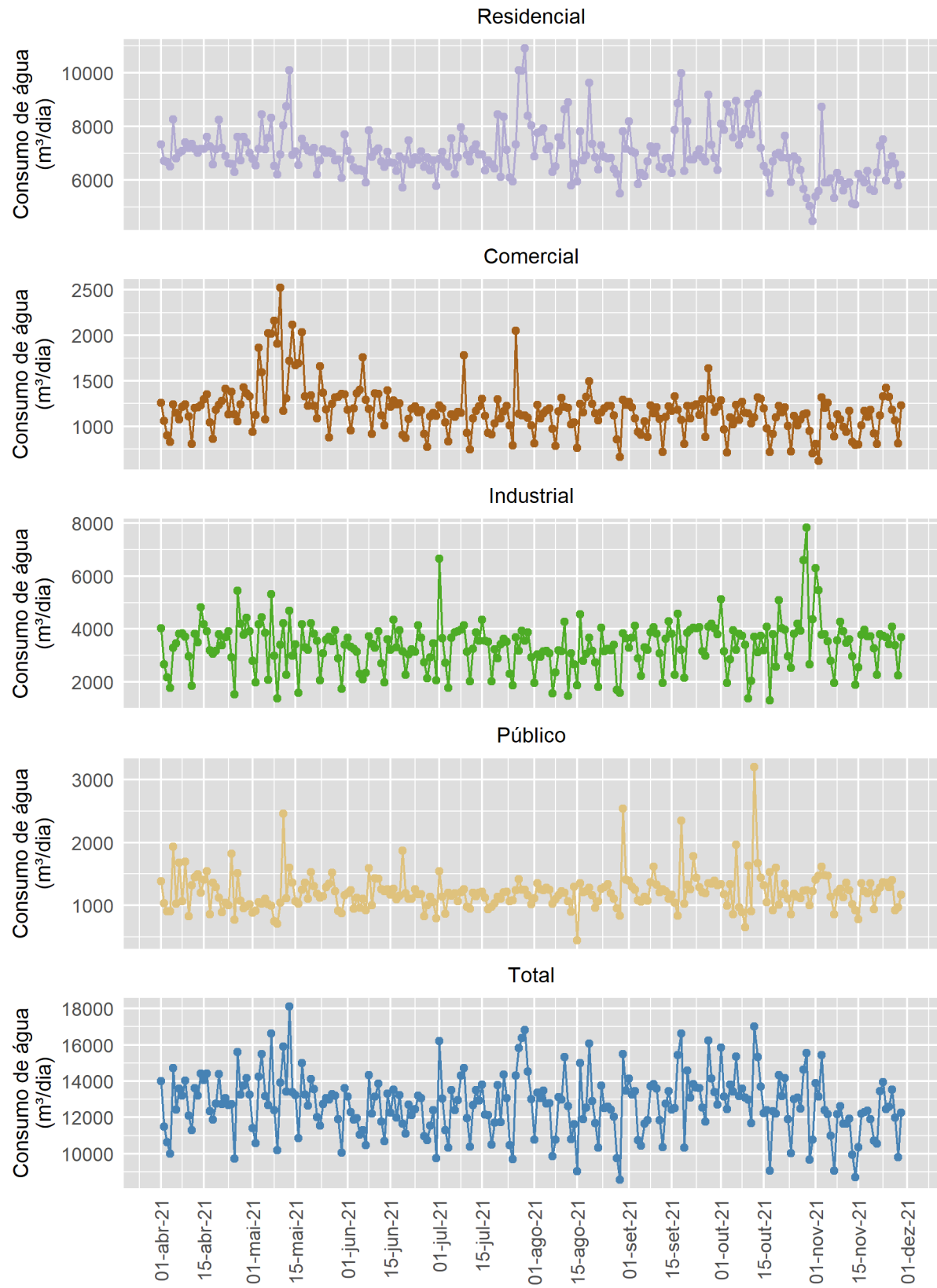


Fonte: Elaborado pelo autor (2022)

O gráfico de séries temporais (Figura 20) permite visualizar o comportamento sazonal da demanda de água no município de Joinville. Nas categorias comercial, industrial e público, o consumo de água semanal apresenta indícios de manter um comportamento estável durante os dias úteis, com uma queda aos finais de semana. Para a categoria residencial, os intervalos semanais de consumo não apresentam um padrão definido ao longo da série temporal estudada, de modo que os picos de consumo ocorrem alternadamente entre os dias úteis e finais de semana.

À exceção dos dias em que a demanda de água apresenta valores visivelmente superiores ao restante da série temporal, os quais podem indicar a ocorrência de vazamentos de água, não é possível visualizar mudanças robustas no comportamento da demanda urbana de água ao longo do período estudado. Em particular, notam-se diferenças mais expressivas no consumo de água na categoria comercial no mês de maio e no consumo residencial a partir do mês de novembro, que pode estar relacionado às férias escolares e recessos trabalhistas.

Figura 20 – Séries temporais da demanda urbana de água



Fonte: Elaborado pelo autor (2022)

4.2.2 Avaliação das premissas

Os limites de controle e parâmetros dos gráficos de controle foram estimados com base no período delimitado pelos primeiros 30 dias da série temporal, correspondendo ao mês de abril de 2021. Após investigação dos conjuntos de dados em relação ao atendimento às premissas exigidas para aplicação dos gráficos de controle convencionais, verificou-se que esse período atendia satisfatoriamente às suposições de normalidade e independência dos dados para todas as categorias de consumo.

A definição do mês de abril para estimação dos parâmetros pode enviesar a sinalização de comportamentos irregulares pelos gráficos de controle na fase de monitoramento, uma vez que as alterações no consumo de água relacionadas a variáveis meteorológicas ou climáticas podem não ter sido consideradas. Além disso, é importante ressaltar outros dois pontos que podem influenciar as análises realizadas nesta pesquisa. O primeiro ponto está relacionado à análise preliminar dos dados, visto que a exclusão de observações discrepantes pode descaracterizar o comportamento do processo. O segundo ponto é referente aos valores obtidos no teste de Shapiro-Wilk para as categorias comercial e público, os quais estão muito próximos do limite de 5% para aceitação da hipótese de normalidade dos dados e podem representar uma desvantagem para os gráficos de controle paramétricos.

A normalidade dos dados foi verificada a partir da aplicação do teste de hipóteses de Shapiro-Wilk (Tabela 4), em que todos os conjuntos de dados retornaram um p-valor $> 0,05$, aceitando a hipótese nula de que os dados seguem uma distribuição normal. Dessa forma, foram atendidas as premissas exigidas para aplicação dos gráficos de controle EWMA e Shewhart, esse último exigindo ainda a verificação da independência dos dados.

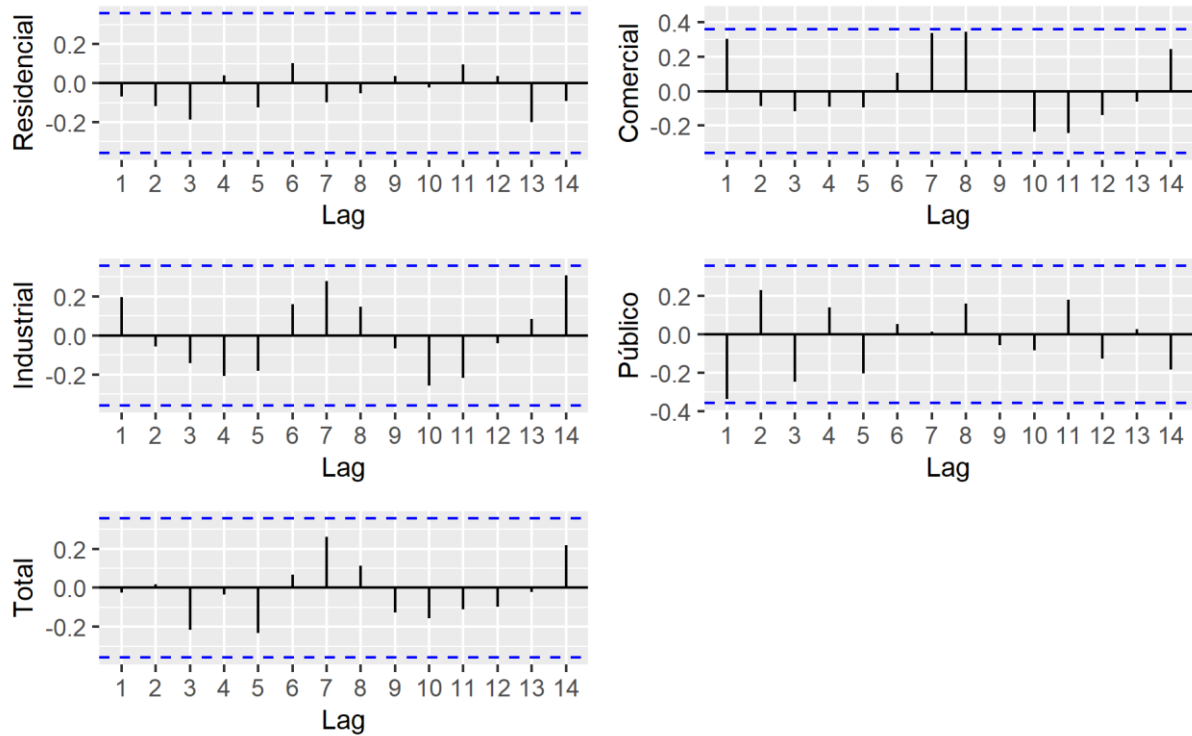
Tabela 4 – Resultados dos testes de Shapiro-Wilk para a demanda urbana de água

Descrição	p-valor
Residencial	0,16
Comercial	0,08
Industrial	0,17
Público	0,07
Total	0,27

Fonte: Elaborado pelo autor (2022)

Para a verificação da independência dos dados, aplicou-se a função de autocorrelação aos conjuntos de dados (Figura 21). Visto que as observações não extrapolam o intervalo de confiança, confirma-se a independência dos dados, permitindo a aplicação do gráfico de controle de Shewhart aos conjuntos de dados.

Figura 21 – Funções de autocorrelação para a demanda urbana de água



Fonte: Elaborado pelo autor (2022)

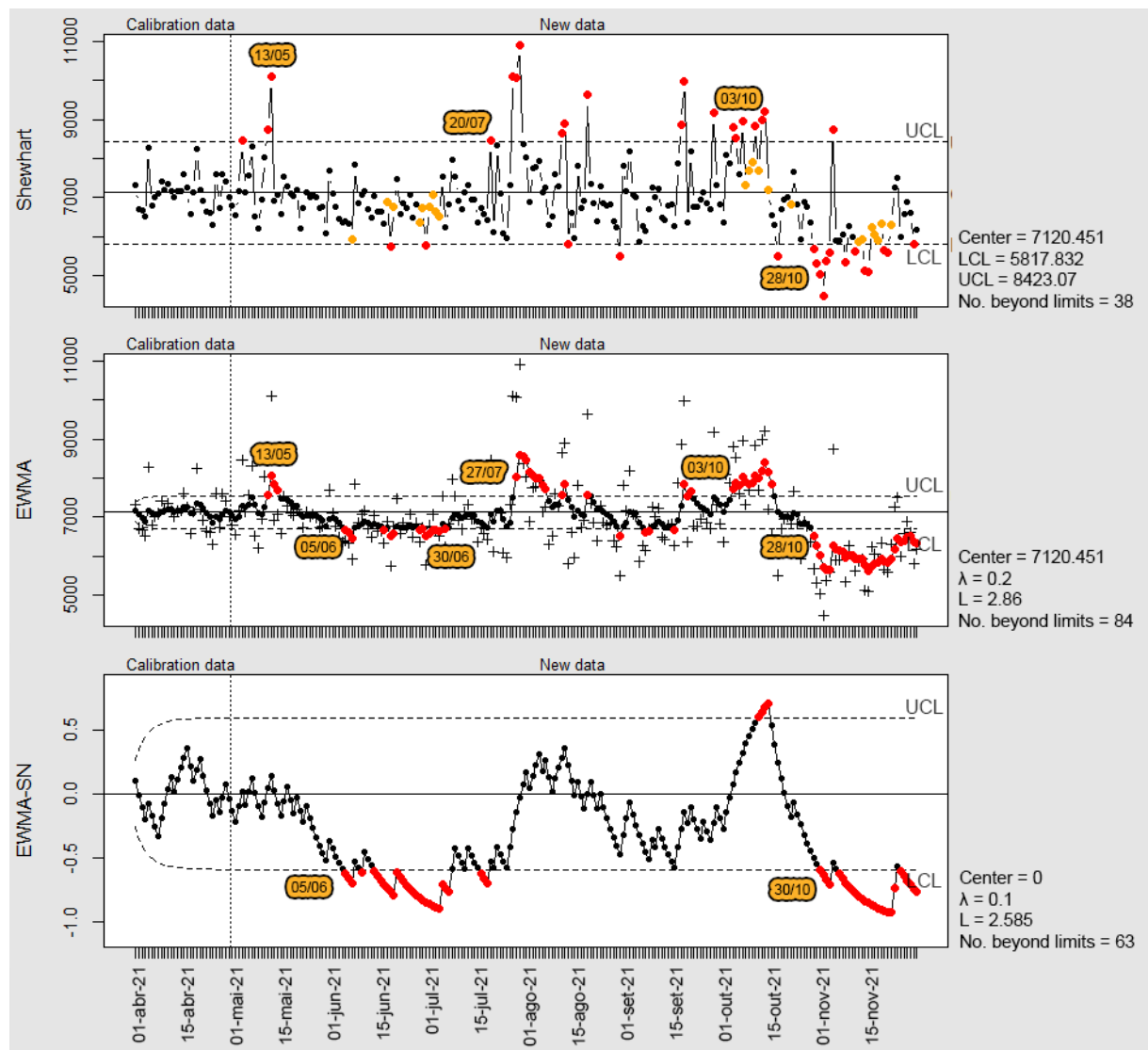
4.2.3 Aplicação dos gráficos de controle

Esta seção aborda o desempenho dos gráficos de controle de Shewhart, EWMA e EWMA-SN para o monitoramento da demanda urbana de água medida pelo sistema de telemetria do município de Joinville. Os resultados são apresentados e discutidos por categoria de uso, nomeadamente residencial, comercial, industrial e público, previamente especificados pela companhia de abastecimento de água do município. O desempenho dos gráficos foi comparado com base nos valores do ARL_0 de cada técnica, o qual representa a sinalização de alarmes falsos. Como não havia a disponibilidade de informações de vazamentos de água por categoria de consumo, nestas análises os gráficos de controle foram comparados em relação à velocidade de detecção de mudanças no processo, as quais podem estar relacionadas a alterações em políticas públicas, variáveis meteorológicas ou recessos trabalhistas e escolares.

4.2.3.1 Gráficos de controle para o consumo residencial

Para a categoria residencial, o gráfico de controle de Shewhart sinalizou proporcionalmente pontos acima e abaixo dos limites de controle (Figura 22). Em relação às observações plotadas acima do LSC, 85% correspondem a dias úteis, enquanto que para os pontos plotados abaixo do LIC, a maioria das observações ocorreu no mês de novembro, período caracterizado pelas férias escolares e recessos trabalhistas. Em ordem cronológica, os períodos mais prolongados de observações plotadas além dos limites de controle ocorreram a partir dos dias 20/07/21, 03/10/21 e 28/10/21.

Figura 22 – Gráficos de controle para o consumo residencial



Fonte: Elaborado pelo autor (2022)

Comparado ao gráfico de Shewhart, o gráfico de controle EWMA sinalizou um maior número de pontos além dos limites de controle, os quais foram detectados em períodos similares (Figura 22**Erro! Fonte de referência não encontrada.**). Nesta técnica, as sinalizações ocorreram a partir dos dias 27/07/21, 03/10/21 e 28/10/21, obtendo resultados equivalentes ao gráfico de controle de Shewhart.

A alternativa não paramétrica EWMA-SN apresentou um comportamento distinto em relação à detecção de mudanças no processo, sinalizando majoritariamente pontos abaixo do LIC, a partir dos dias 05/06/21 e 30/10/21 (Figura 22**Erro! Fonte de referência não encontrada.**).

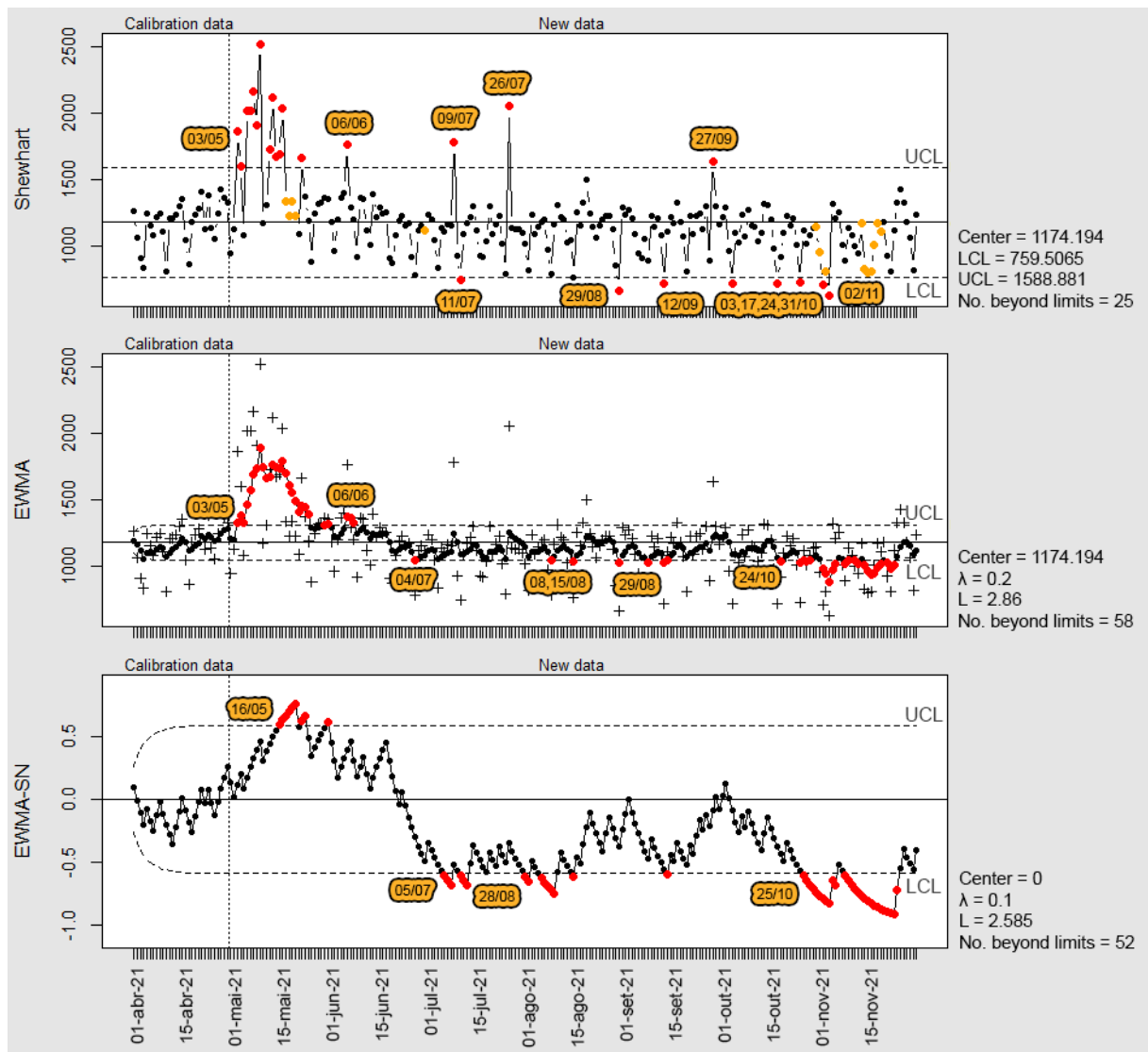
4.2.3.2 Gráficos de controle para o consumo comercial

A construção do gráfico de controle de Shewhart para a categoria comercial resultou em 25 observações além dos limites de controle (Figura 23). Os pontos plotados acima do LSC correspondem majoritariamente ao mês de maio, sinalizados a partir do dia 03/05/21, e podem estar relacionados ao feriado do dia das mães e ao avanço do processo de vacinação contra a COVID-19. Dentre os 8 pontos plotados abaixo do LIC, 7 ocorreram em domingos e somente 1 em uma terça-feira, referente ao feriado do dia 02/11/21.

Assim como no gráfico de Shewhart, o gráfico EWMA detectou uma mudança no processo a partir do dia 03/05/21, prolongando-se pelo mês de maio (Figura 23). Contudo, foram sinalizados pontos abaixo do LIC a partir do dia 24/10/21, os quais não foram considerados como fora do controle estatístico no gráfico de Shewhart. Este período é caracterizado pela ocorrência de feriados e antecede os recessos trabalhistas, o que pode causar uma redução no consumo de água nesta categoria devido à possibilidade dos munícipes viajarem a cidades próximas.

A desconsideração do valor numérico das observações no gráfico EWMA-SN resultou em um atraso na detecção de mudanças no processo no mês de maio, sinalizando a fuga do controle somente no dia 16/05/21 (Figura 23). Os pontos abaixo do LIC no mês de outubro foram sinalizados a partir do dia 25/10/21, de maneira similar ao gráfico EWMA.

Figura 23 – Gráficos de controle para o consumo comercial



Fonte: Elaborado pelo autor (2022)

4.2.3.3 Gráficos de controle para o consumo industrial

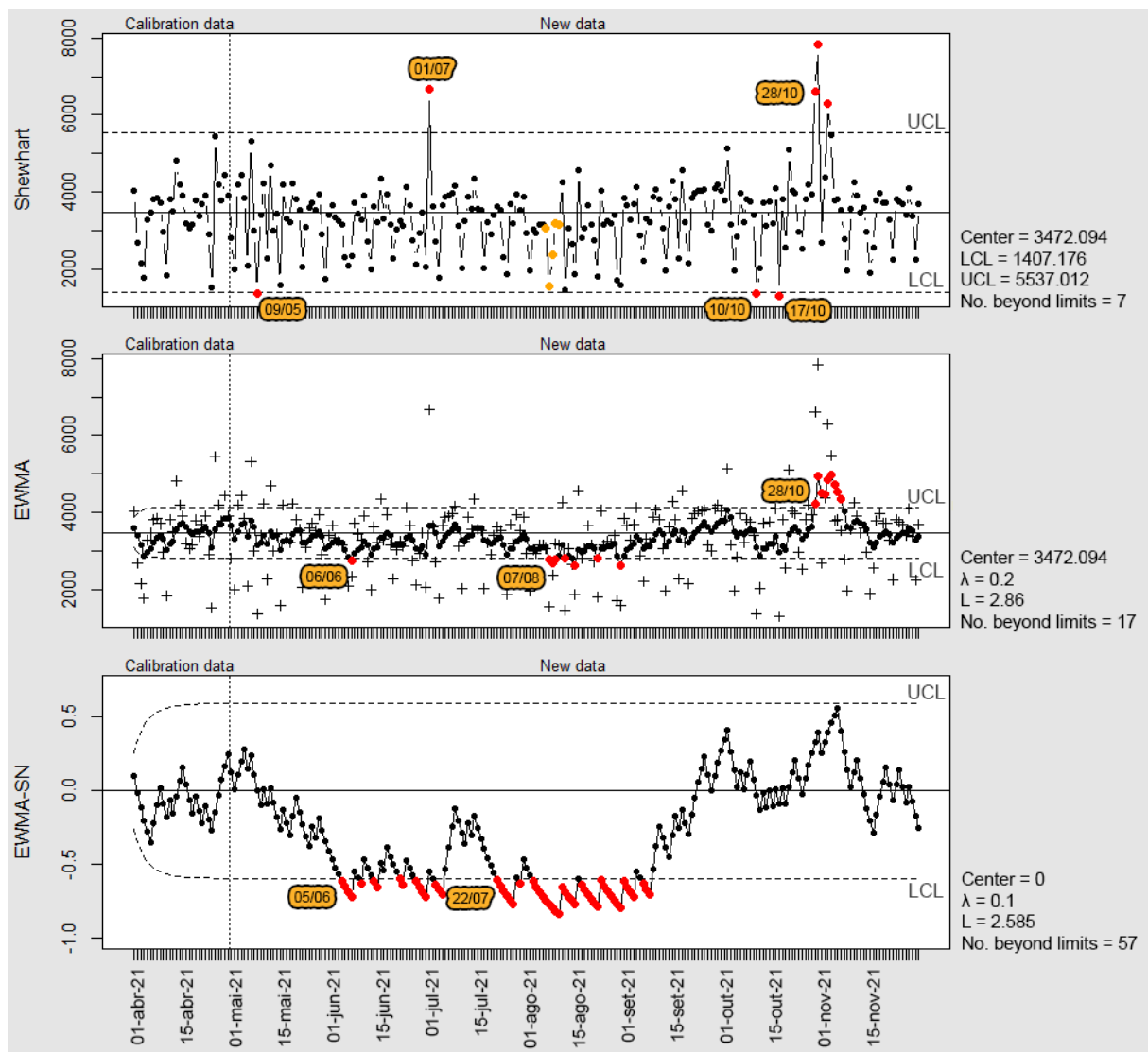
Para a categoria industrial, o gráfico de Shewhart sinalizou 7 observações além dos limites de controle (Figura 24). As observações indicando um consumo abaixo do LIC são referentes a finais de semana, quando as indústrias da cidade operam em menor escala e, consequentemente, o consumo de água desta categoria é reduzido. Os pontos acima do LSC ocorreram nos dias 01/07/21 e a partir do dia 28/10/21, ambos referentes a quintas-feiras.

O gráfico de controle EWMA obteve um comportamento similar ao gráfico de Shewhart, sinalizando uma fuga do controle estatístico do processo a partir do dia 07/08/21, assim como a violação das regras suplementares no gráfico de Shewhart, porém não foram detectadas algumas mudanças repentinas no processo, como as observações referentes aos dias

09/05/21, 01/07/21, 10/10/21 e 17/10/21 (Figura 24 **Erro! Fonte de referência não encontrada.**). Em relação aos pontos acima do LSC no mês de outubro, a sinalização no gráfico EWMA ocorreu no mesmo dia em que foi detectada a fuga do controle no gráfico de Shewhart.

O gráfico não paramétrico EWMA-SN sinalizou apenas pontos abaixo do LIC, a partir dos dias 05/06/21 e 22/07/21, os quais não foram considerados como fora do controle estatístico nos gráficos convencionais (Figura 24 **Erro! Fonte de referência não encontrada.**). A aplicação desta técnica também não detectou uma mudança no processo a partir do dia 28/10/21, sinalizada nos outros gráficos.

Figura 24 – Gráficos de controle para o consumo industrial

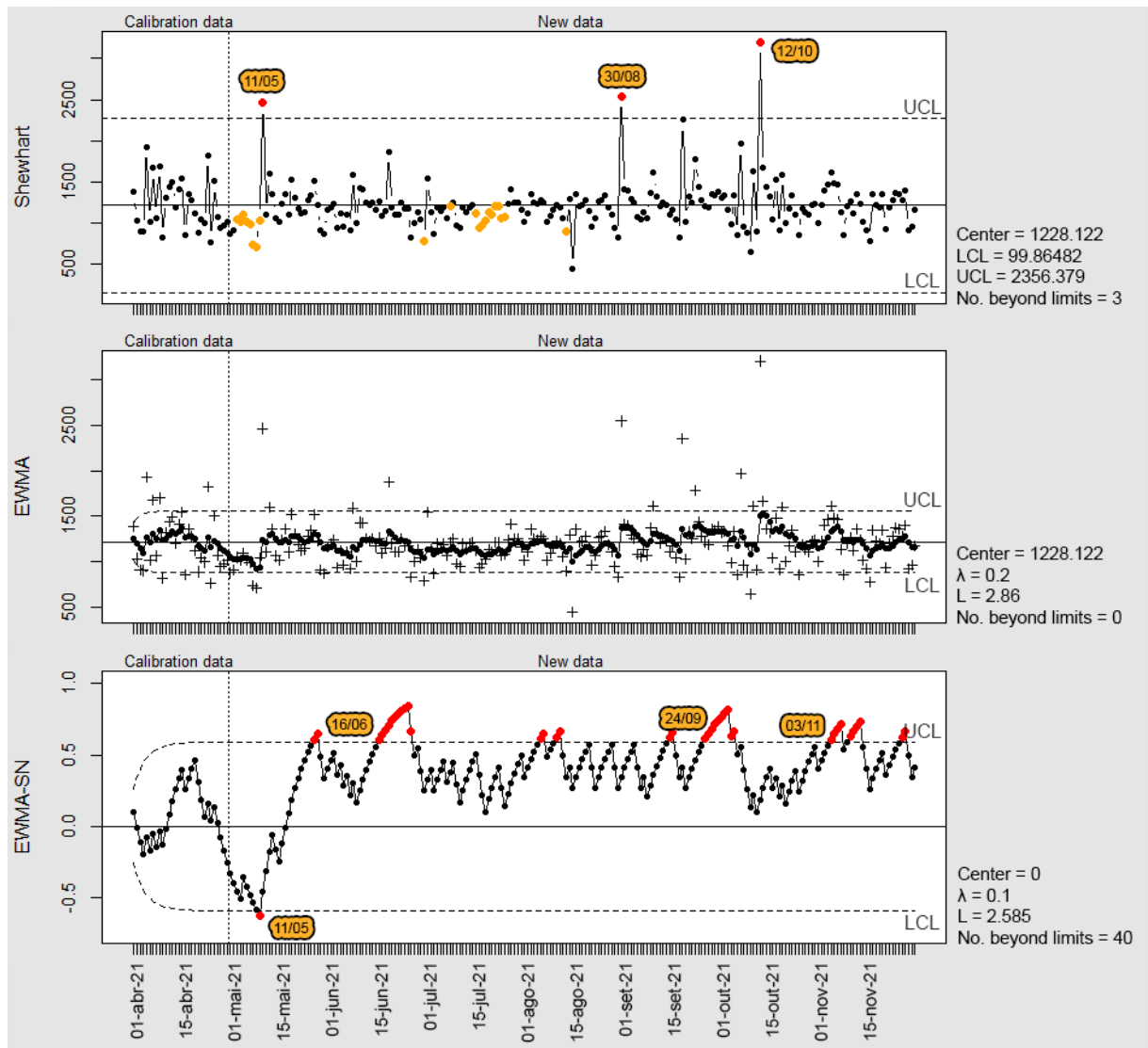


Fonte: Elaborado pelo autor (2022)

4.2.3.4 Gráficos de controle para o consumo público

A construção do gráfico de controle de Shewhart para a categoria público resultou em apenas 3 pontos além dos limites de controle, todos acima do LSC e referentes aos dias 11/05/21, 30/08/21 e 12/10/21 (Figura 25). Além da análise convencional dos pontos excedendo os limites de controle, o gráfico de Shewhart também sinalizou 21 observações violando o conjunto de regras suplementares em 5 períodos distintos, nos dias 03/05/21, 30/06/21, 08/07/21, 16/07/21 e 13/08/21.

Figura 25 – Gráficos de controle para o consumo público



Fonte: Elaborado pelo autor (2022)

Conforme pode ser observado na Figura 25**Erro! Fonte de referência não encontrada.**, o gráfico de controle EWMA não sinalizou quaisquer observações além dos limites de controle, nem mesmo quando o consumo de água apresentou um comportamento não aleatório, como os períodos em que foram observados valores consecutivos de consumo abaixo da LC, os quais foram sinalizados no gráfico de Shewhart considerando-se o conjunto de regras da *Western Electric*.

Comparado às outras técnicas, o gráfico EWMA-SN sinalizou uma quantidade expressiva de pontos acima do LSC, os quais ocorreram em períodos distintos das sinalizações no gráfico de Shewhart (Figura 25**Erro! Fonte de referência não encontrada.**). O ponto plotado abaixo do LIC no gráfico EWMA-SN corresponde à mesma observação sinalizada acima do LSC no dia 11/05/21 no gráfico de Shewhart, destacando a diferença entre essas técnicas na sinalização de mudanças no processo.

4.2.3.5 Gráficos de controle para o consumo total

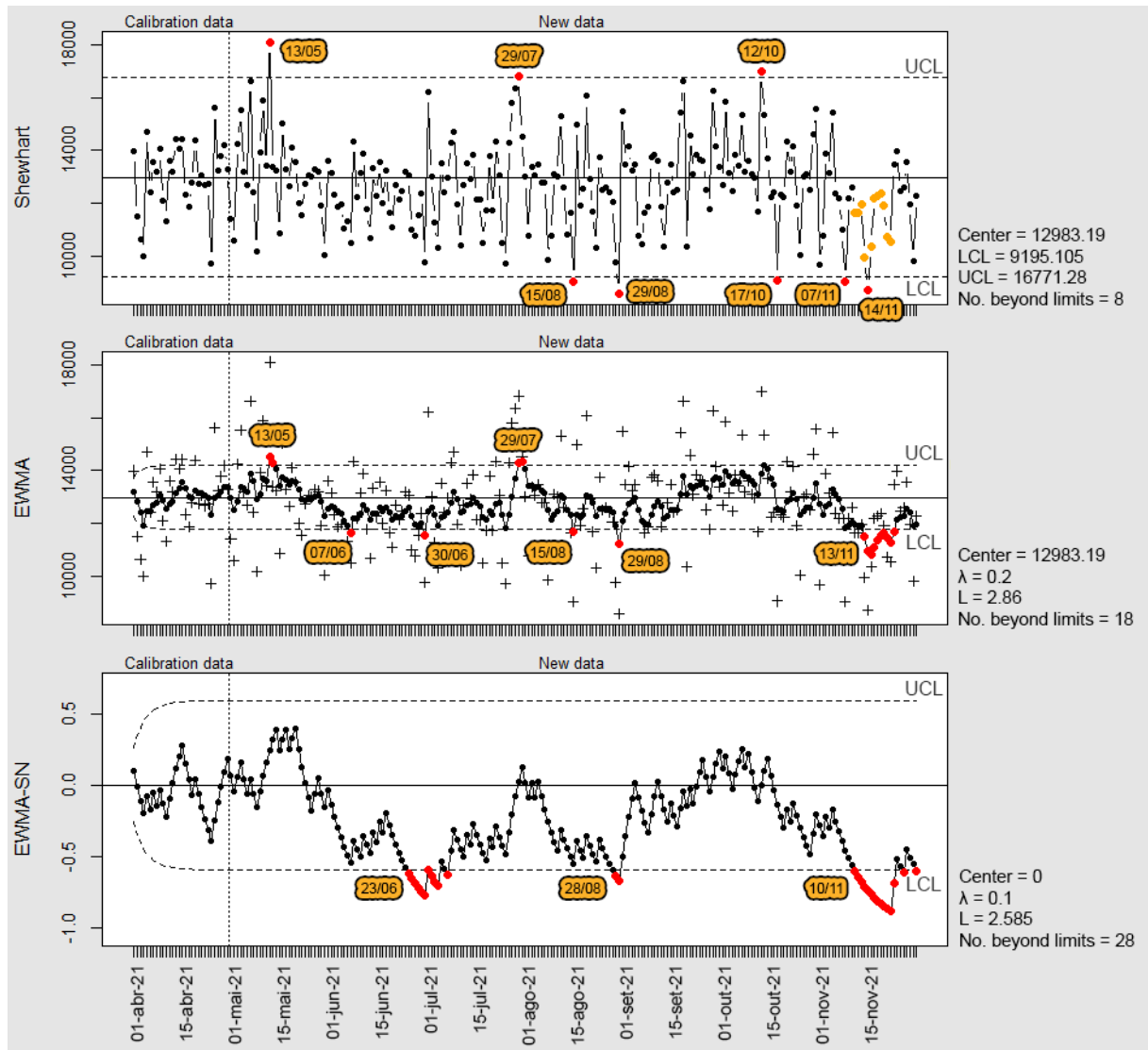
Para o monitoramento do consumo total de água medido por sistema de telemetria no município, o gráfico de Shewhart sinalizou 8 observações além dos limites de controle (Figura 26). As observações abaixo do LIC são todas referentes a domingos e evidenciam a redução na demanda urbana de água em virtude da interrupção no funcionamento de indústrias e comércios da cidade, impactando principalmente os consumos residencial, comercial e industrial, conforme comentado nas análises anteriores. As sinalizações acima do LSC são referentes aos dias 13/05/21, 29/07/21 e 12/10/21, e correspondem a duas quintas-feiras e uma terça-feira, respectivamente. Em consulta à página eletrônica da empresa concessionária de água, foram informados vazamentos de água em três bairros diferentes nos dias 12/05/21 e 13/05/21, os quais podem ter motivado a sinalização acima do LSC no gráfico de Shewhart, apesar de existir a possibilidade de estes eventos não terem sido captados pelo sistema de telemetria. Também é sinalizada uma mudança no comportamento do processo por meio das regras da *Western Electric* a partir dos dias 10/11/21 e 15/11/21.

O gráfico EWMA não detectou algumas mudanças repentinas no processo, referentes aos dias 12/10/21, 17/10/21 e 07/11/21 (Figura 26). Contudo, a aplicação dessa técnica facilitou a visualização de tendências sazonais na demanda de água no município de Joinville, a qual aparenta ser maior nos meses de maio, setembro e outubro, e menor nos meses de junho, julho e agosto, os quais correspondem aos meses de inverno e podem causar uma redução no consumo residencial. Também é possível visualizar uma diminuição no consumo de água a partir do mês

de novembro, em que os feriados prolongados e recessos podem reduzir o consumo no município. Assim como no gráfico de Shewhart, o gráfico EWMA detectou uma mudança no processo a partir do dia 13/05/21, a qual pode estar relacionada ao vazamento de água informado na página eletrônica da concessionária mencionado anteriormente.

Comparado às outras técnicas, o gráfico não paramétrico EWMA-SN apresentou uma quantidade maior de pontos além dos limites de controle, destacando que o comportamento não aleatório do processo, como sequências de pontos consecutivamente acima ou abaixo da LC, pode refletir em uma maior sinalização de alarmes no gráfico (Figura 26). Devido ao fato de esta técnica não considerar a magnitude das observações, sua aplicação não detectou algumas mudanças repentinas no processo, como a fuga do controle no dia 13/05/21, referente ao vazamento de água que pode ter sido captado pelo sistema de telemetria, detectado em ambos os gráficos de Shewhart e EWMA.

Figura 26 – Gráficos de controle para o consumo total



Fonte: Elaborado pelo autor (2022)

4.2.4 Análise comparativa

Referente aos dados de consumo em escala municipal, os gráficos paramétricos de Shewhart e EWMA apresentaram resultados similares em relação à sinalização de comportamentos irregulares na série temporal estudada, detectando o vazamento de água possivelmente contemplado neste conjunto de dados simultaneamente, enquanto que o gráfico EWMA-SN não sinalizou a ocorrência deste evento. De modo geral, o gráfico de Shewhart apresentou menor sinalização de pontos além dos limites de controle.

A aplicação do gráfico de controle não paramétrico EWMA-SN não detectou diversas mudanças repentinas no processo, uma vez que a estatística de plotagem desta técnica não considera o valor numérico das observações, apenas a indicação de se elas são maiores ou

menores que um parâmetro de interesse estabelecido. Este resultado está de acordo com a literatura, pois alternativas não paramétricas são menos poderosas do que os métodos paramétricos ao serem aplicados a dados com distribuição normal. De acordo com Ross (2009), métodos não paramétricos podem ser utilizados sem que sejam feitas quaisquer suposições sobre a distribuição dos dados, no entanto, não serão tão eficazes nos casos em que as distribuições são normais, em comparação aos testes que começam assumindo a suposição de normalidade dos dados.

Em relação ao comportamento do consumo de água de cada categoria, os segmentos residencial e comercial apresentaram as maiores mudanças ao longo da série temporal estudada, as quais podem estar relacionadas a alterações no clima ou períodos de férias e recessos (MINI; HOGUE; PINCETL, 2014; DZIMIŃSKA *et al.*, 2021). O padrão de queda observado ao longo da série temporal na categoria comercial está de acordo com estudo realizado por Batista *et al.* (2022) acerca do consumo de água em edifícios comerciais, em que os autores constataram haver uma correlação entre o consumo de água desse setor e variáveis climáticas. No mês de maio, o aumento no consumo de água observado nessa categoria pode estar relacionado ao avanço do processo de vacinação contra a COVID-19, o que flexibilizou as medidas de prevenção e combate ao coronavírus e pode ter causado o aumento no consumo de água comercial.

Na categoria residencial, os maiores valores de consumo foram observados nos meses de maio, setembro e outubro, e os menores, nos meses de junho, julho e agosto, que correspondem aos meses de inverno e podem representar uma diminuição no consumo de água nessa categoria (GATO-TRINIDAD; JAYASURIYA; ROBERTS, 2011; CARDELL-OLIVER, 2013). Também é possível visualizar uma diminuição no consumo de água a partir do mês de novembro, em que a ocorrência de feriados e recessos pode causar mudanças no comportamento do consumo de água doméstico (DZIMIŃSKA *et al.*, 2021).

Considerando-se as variações semanais, as categorias comercial, industrial e público apresentaram um padrão de consumo estável nos dias úteis e reduzido aos finais de semana. Na categoria residencial, não foi identificado um padrão de consumo ao longo da semana, de modo que os picos de consumo ocorrem alternadamente entre os dias úteis e finais de semana. De acordo com Bougadis, Adamowski e Diduch (2005), essas variações são um aspecto importante a ser considerado nas análises de previsão de demanda de água, podendo produzir resultados melhores em termos de precisão dos modelos.

As análises realizadas nesta pesquisa foram limitadas em virtude dos dados estudados representarem uma série temporal retrospectiva do consumo de água do município, em que não estavam disponíveis informações detalhadas das unidades consumidoras de cada segmento.

Desse modo, não foi possível identificar se as alterações no consumo aconteceram em uma matrícula em específico, o que poderia representar a ocorrência de um vazamento de água na unidade em questão, ou se as mudanças sinalizadas nos gráficos representam alterações no padrão do consumo de água da categoria como um todo. Ainda nesse contexto, a investigação do consumo de água por matrícula pode não refletir precisamente o consumo de água por unidade consumidora ou por pessoa, uma vez que uma matrícula pode ter mais de uma unidade consumidora.

5 CONCLUSÃO

Esta pesquisa buscou comparar o desempenho de gráficos de controle paramétricos e não paramétricos para o monitoramento do consumo de água nas escalas edificada e municipal, com o uso dos tradicionais gráficos de controle de Shewhart, EWMA e a alternativa não paramétrica EWMA-SN. Foi analisado o consumo de água total medido por sistema de telemetria no município de Joinville, em Santa Catarina, classificando-o conforme as categorias de uso residencial, comercial, industrial e público. Na escala edificada, os gráficos de controle foram aplicados a dois conjuntos habitacionais que apresentaram vazamentos de água no período estudado, a fim de analisar a eficácia dos gráficos na detecção desses eventos.

No tocante às análises realizadas nos conjuntos habitacionais, sem a aplicação das regras suplementares no gráfico de Shewhart, o gráfico EWMA obteve os melhores resultados, sinalizando primeiramente a ocorrência do vazamento de água. Apesar de o gráfico de Shewhart ter detectado mais rapidamente o vazamento de água considerando-se o conjunto de regras da *Western Electric*, o uso dessas regras altera o valor do ARL_0 , podendo aumentar a probabilidade de sinalização de alarmes falsos. O método não paramétrico EWMA-SN apresentou comportamento similar ao gráfico EWMA, devido às semelhanças no cálculo das estatísticas de plotagem dos mesmos, apresentando ainda a vantagem de não exigir o atendimento às suposições de independência ou normalidade dos dados.

Os resultados obtidos nas análises da demanda urbana de água mostram que os tradicionais gráficos de Shewhart e EWMA apresentaram desempenho similar, detectando a mudança que pode estar relacionada ao vazamento informado pela concessionária de água simultaneamente, ao passo que o gráfico não paramétrico EWMA-SN não sinalizou a ocorrência desse evento. A utilização da mediana como valor de referência para o gráfico EWMA-SN resultou em uma maior robustez a mudanças repentinas no processo, não detectando observações sinalizadas nos outros gráficos. A eficiência do gráfico não paramétrico está relacionada à normalidade do conjunto de dados estudado, uma vez que técnicas paramétricas são mais robustas ao serem aplicadas a dados com distribuição normal. As sinalizações que não correspondem a vazamentos de água apontam a eficácia dos gráficos para o monitoramento do padrão da demanda urbana de água, o qual pode identificar se políticas públicas ou campanhas de conscientização provocam mudanças no comportamento da demanda de água. Neste contexto, o gráfico EWMA obteve os melhores resultados, facilitando a visualização de tendências no consumo diário de água.

No que diz respeito à eficácia dos gráficos de controle na identificação de vazamentos de água em escala urbana, as análises realizadas nesta pesquisa podem ter sido prejudicadas em virtude dos dados representarem apenas o consumo de água de matrículas atendidas por sistema de telemetria, os quais podem não ter contemplado os vazamentos de água informados pela concessionária de água em comunicados em sua página eletrônica. Todavia, ainda é válido ressaltar que a magnitude dos vazamentos em relação ao consumo total sendo monitorado pode dificultar a detecção dos mesmos, de modo que eventos de menor magnitude podem não ser detectados pelos gráficos de controle. Essa desvantagem pode ser minimizada dividindo-se a aplicação dos gráficos de controle aos centros de distribuição de água do município. Conforme dados da concessionária de água, o município conta com 13 reservatórios que realizam o fornecimento de água a todos os domicílios. Dessa forma, ao realizar o monitoramento de maneira setorizada, os vazamentos podem representar uma parcela maior do consumo de água sendo monitorado, aumentando a eficácia dos gráficos de controle na detecção destes eventos.

Em linhas gerais, visto que o consumo de água apresenta tendências sazonais, a definição de um período específico para a estimação dos parâmetros pode resultar em um comportamento enviesado dos gráficos de controle, uma vez que o período definido pode não estar abrangendo mudanças no padrão do consumo de água relacionadas a variáveis externas, como o clima e o nível de precipitação. Dessa forma, sugere-se a consideração de um período mais abrangente do processo na etapa de construção dos gráficos de controle. Os parâmetros estatísticos podem ser definidos considerando-se os dados referentes a um semestre ou ano, o que consistiria em limites de controle mais largos e que poderiam compreender as variações sazonais no consumo de água. Em virtude da dispersão dos dados, outra sugestão seria a investigação do monitoramento do consumo de água em intervalos semanais, a partir da construção de gráficos de controle para a média semanal ou o desvio padrão, ao invés de medidas individuais.

Isto posto, os resultados obtidos neste estudo podem preencher importantes lacunas de conhecimento sobre o uso de gráficos de controle para o monitoramento do consumo de água. A metodologia utilizada nesta pesquisa pode auxiliar no desenvolvimento de ferramentas automatizadas para o monitoramento da demanda urbana de água ou de edificações. Os gráficos de controle são ferramentas consolidadas para o monitoramento de processos e permitem a identificação e geração de alarmes caso seja detectado um comportamento irregular. Dessa forma, a aplicação dessas técnicas pode agilizar a detecção de vazamentos de água, otimizando o tempo de intervenção e reduzindo perdas. Os resultados obtidos mostram que os gráficos de controle podem ser utilizados para o monitoramento do consumo de água de forma contínua e

remota, via IoT, sinalizando aos gestores caso sejam detectadas mudanças no processo, reduzindo os esforços relacionados à identificação e correção desses eventos. Os gráficos de controle também podem ser utilizados para a avaliação do consumo de água ao longo do tempo, identificando horários e períodos de maior consumo e auxiliando no desenvolvimento de políticas públicas voltadas à conservação da água nestes intervalos.

5.1 SUGESTÕES PARA TRABALHOS FUTUROS

A fim de contornar limitações encontradas no desenvolvimento desta pesquisa e complementar a metodologia aplicada, são propostas as seguintes estratégias para trabalhos futuros:

- a) setorizar o monitoramento da demanda urbana de água a centros de distribuição, de modo que comportamentos irregulares possam representar uma parcela maior do consumo de água sendo monitorado;
- b) utilizar um período mais abrangente do processo na fase de definição dos parâmetros estatísticos dos gráficos de controle, na ordem de um semestre ou ano, com o intuito de compreender as variações sazonais presentes no consumo de água;
- c) considerar o monitoramento da variabilidade semanal do processo, a partir da construção de gráficos de controle para a média, desvio padrão e amplitude amostrais;
- d) avaliar o desempenho da combinação de gráficos de controle para o monitoramento do consumo de água, com o intuito de diminuir a sinalização de alarmes falsos.

REFERÊNCIAS BIBLIOGRÁFICAS

ABES – Associação Brasileira de Engenharia Sanitária e Ambiental. **Perdas em sistemas de abastecimento de água: diagnóstico, potencial de ganhos com sua redução e propostas de medidas para o efetivo combate**. 2013. Disponível em: <<http://www.abes-sp.org.br/arquivos/perdas.pdf>>. Acesso em: 08 de nov. de 2020.

AHMED, U.; MUMTAZ, R.; ANWAR, H.; MUMTAZ, S.; QAMAR, A. M. Water quality monitoring: from conventional to emerging technologies. **Water Supply**, v. 20, n. 1, p. 28-45, 2020.

AHN, J.; JUNG, D. Hybrid Statistical Process Control Method for Water Distribution Pipe Burst Detection. **Journal of Water Resources Planning and Management**, v. 145, n. 9, p. 06019008, 2019.

ALAHAKOON, D.; YU, X. Smart Electricity Meter Data Intelligence for Future Energy Systems: A Survey. **IEEE Transactions on Industrial Informatics**, v. 12, n. 1, p. 425-436, 2016.

ALBARRACÍN, O. Y. E. **Monitoramento de séries de contagem por meio de gráficos de controle**. Dissertação (Mestrado em Ciências) – Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2014.

ALI, A. M.; SHAFIEE, M. E.; BERGLUND, E. Z. Agent-based modeling to simulate the dynamics of urban water supply: Climate, population growth, and water shortages. **Sustainable Cities and Society**, v. 28, p. 420-434, 2017.

AMIN, R. W., SEARCY, A. J. A nonparametric exponentially weighted moving average control scheme. **Communications in Statistics: Simulation and Computation**, v. 20, n. 4, p. 1049-1072, 1991.

ANA – Agência Nacional de Águas. **Atlas Brasil: abastecimento urbano de água: panorama nacional**. Brasília: Agência Nacional de Águas (ANA), 2010. 72 p.

ANA – Agência Nacional de Águas. **Conjuntura dos recursos hídricos no Brasil 2018: relatório pleno**. Brasília: Agência Nacional de Águas (ANA), 2018. 72 p.

ANA – Agência Nacional de Águas. **ODS 6 no Brasil: Visão da ANA sobre os indicadores**. Brasília: Agência Nacional de Águas (ANA), 2019. 100 p.

ANTZOULATOS, G.; MOURTZIOS, C.; STOURNARA, P.; KOULOGLOU, I.; PAPADIMITRIOU, N.; SPYROU, D.; MENTES, A.; NIKOLAIDIS, E.; KARAKOSTAS, A.; KOURTESIS, D.; VROCHIDIS, S.; KOMPATSIARIS, I. Making urban water smart: the SMART-WATER solution. **Water Science & Technology**, v. 82, n. 12, p. 2691-2710, 2020.

ARAL, M. M.; GUAN, J.; MASLIA, M. L.; GRAYMAN, W. M. Optimization model and algorithms for design of water sensor placement in water distribution systems. **Proceedings of the 8th Annual Water Distribution Systems Analysis Symposium**, Cincinnati, p. 27-30, ago. 2006.

ASSIS, J. P. de; SOUSA, R. P. de; LINHARES, P. C. F. **Testes de hipóteses estatísticas**. Mossoró: EdUFERSA, 2020. 182 p.

BALLING JR., R. C.; GOBER, P. Climate Variability and Residential Water Use in the City of Phoenix, Arizona. **Journal of Applied Meteorology and Climatology**, v. 46, p. 1130-1137, 2007. <https://doi.org/10.1175/JAM2518.1>

BALLING JR., R. C.; GOBER, P.; JONES, N. Sensitivity of residential water consumption to variations in climate: An intraurban analysis of Phoenix, Arizona. **Water Resources Research**, v. 44, W10401, 2008. <https://doi.org/10.1029/2007WR006722>

BAKKER, M.; JUNG, D.; VREEBURG, J.; VAN DE ROER, M.; LANSEY, K.; RIETVELD, L. Detecting pipe bursts using Heuristic and CUSUM methods. **Procedia Engineering**, v. 70, p. 85-92, 2014. <https://doi.org/10.1016/j.proeng.2014.02.011>

BARAI, G. R.; KRISHNAN, S.; VENKATESH, B. Smart metering and functionalities of smart meters in smart grid – A review. **Proceedings of the 2015 IEEE Electrical Power and Energy Conference: Smarter Resilient Power Systems (EPEC)**, Londres, p. 26-28, out. 2015.

BATES, B.; KUNDZEWICZ, Z. W.; WU, S.; PALUTIKOF, J. **Climate Change and Water**. Technical Paper of the Intergovernmental Panel on Climate Change. Geneva: IPCC Secretariat, 2008. 210 p.

BATISTA, L. T.; FRANCO, J. R. Q.; FAKURY, R. H.; PORTO, M. F.; BRAGA, C. M. P. Methodology for Determining Sustainable Water Consumption Indicators for Buildings. **Sustainability**, v. 14, n. 9, p. 5695, 2022.

BERTRAND, J. W. M.; FRANSOO, J. C. Operations management research methodologies using quantitative modelling. **International Journal of Operations and Production Management**, v. 22, n. 2, p. 241-264, 2002. <https://doi.org/10.1108/01443570210414338>

BOUGADIS, J.; ADAMOWSKI, K.; DIDUCH, R. Short-term municipal water demand forecasting. **Hydrological Processes**, v. 19, n. 1, p. 137-148, 2005. <https://doi.org/10.1002/hyp.5763>

BORGES, A. P. A.; JUNG, D.; KIM, J. H. Smart WDS management: Pipe burst detection using real-time monitoring data. **IEEE SmartWorld**, p. 1-4, 2017.

BORROR, C. M.; MONTGOMERY, D. C.; RUNGER, G. C. Robustness of the EWMA Control Chart to Non-Normality. **Journal of Quality Technology**, v. 31, n. 3, p. 309-316, 1999. <https://doi.org/10.1080/00224065.1999.11979929>

BROOK, D.; EVANS, D. A. An approach to the probability distribution of CUSUM run length. **Biometrika**, v. 59, n. 3, p. 539-549, 1972. <https://doi.org/10.2307/2334805>

BUREK, P.; SATOH, Y.; FISCHER, G.; KAHIL, T.; JIMENEZ, L. N.; SCHERZER, A.; TRAMBEREND, S.; WADA, Y.; EISNER, S.; FLÖRKE, M.; HANASAKI, N.; MAGNUSZIEWSKI, P.; COSGROVE, W.; WIBERG, D. **Water futures and solution: fast track initiative**. Áustria: IIASA, 2016. 88 p.

CARDELL-OLIVER, R. Water use signature patterns for analyzing household consumption using medium resolution meter data. **Water Resources Research**, v. 49, p. 8589-8599, 2013. <https://doi.org/10.1002/2013WR014458>

CASSIDY, J.; BARBOSA, B.; DAMIÃO, M.; RAMALHO, P.; GANHÃO, A.; SANTOS, A.; FELICIANO, J. Taking water efficiency to the next level: digital tools to reduce non-revenue water. **Journal of Hydroinformatics**, v. 23, n. 3, p. 453-465, 2021.

CHAKRABORTI, S.; GRAHAM, M. A. Nonparametric control charts. In: RUGGERI, F.; KENETT, R. S.; FALTIN, F. W. **Encyclopedia of Quality and Reliability**. Nova Iorque: John Wiley & Sons, 2008. <https://doi.org/10.1002/9780470061572.eqr262>

CHAKRABORTI, S.; GRAHAM, M. A. Nonparametric (distribution-free) control charts: An updated overview and some results. **Quality Engineering**, v. 31, n. 4, p. 523-544, 2019. <https://doi.org/10.1080/08982112.2018.1549330>

CHAKRABORTI, S.; VAN DER LAAN, P.; BAKIR, S. T. Nonparametric control charts: An overview and some results. **Journal of Quality Technology**, v. 33, n. 3, p. 304-15, 2001.

CHANG, H.; PRASKIEVICZ, S.; PARANDVASH, H. Sensitivity of Urban Water Consumption to Weather and Climate Variability at Multiple Temporal Scales: The Case of Portland, Oregon. **International Journal of Geospatial and Environmental Research**, v. 1, n. 1, 2014.

CHATFIELD, C. **Problem solving: a statistician's guide**. 2 ed. Londres: Chapman & Hall, 1995.

CHINI, C. M.; STILLWELL, A. S. The metabolism of U.S. cities 2.0. **Journal of Industrial Ecology**, v. 23, p. 1353-1362, 2019.

COLOMBO, A.; LEE, P.; KARNEY, B. A selective literature review of transient based leak detection methods. **Journal of Hydro-Environment Research**, v. 2, p. 212-227, 2009. <https://doi.org/10.1016/j.jher.2009.02.003>

COMINOLA, A.; GIULIANI, M.; PIGA, D.; CASTELLETTI, A.; RIZZOLI, A. E. Benefits and challenges of using smart meters for advancing residential water demand modeling and management: A review. **Environmental Modelling & Software**, v. 72, p. 198-214, 2015.

COMINOLA, A.; GIULIANI, M.; CASTELLETTI, A.; ROSENBERG, D.E.; ABDALLAH, A.M. Implications of data sampling resolution on water use simulation, end-use disaggregation, and demand management. **Environmental Modelling & Software**, v. 102, p. 199-212, 2018. <https://doi.org/10.1016/j.envsoft.2017.11.022>

COSTA, A. F. B.; EPPRECHT, E. K.; CARPINETTI, L. C. R. **Controle Estatístico de Qualidade**. 2 ed. São Paulo: Atlas, 2016.

CRESPO, A. A. **Estatística**. 20 ed. São Paulo: Saraiva, 2020.

CROWDER, S. V. A Simple Method for Studying Run Length Distributions of Exponentially Weighted Moving Average Charts. **Technometrics**, v. 29, p. 393-407, 1987a.

CROWDER, S. V. Computation of ARL for combined individual measurements and moving range charts. **Journal of Quality Technology**, v. 19, p. 98-102, 1987b.

CROWDER, S. V. Design of Exponentially Weighted Moving Average Schemes. **Journal of Quality Technology**, v. 21, p. 155-162, 1989.

DEVORE, J. L. **Probabilidade e Estatística para Engenharia e Ciências**. 8 ed. São Paulo: Cengage Learning, 2016.

DI MAURO, A.; COMINOLA, A.; CASTELLETTI, A.; DI NARDO, A. [Banco de dados] Urban Water Consumption at Multiple Spatial and Temporal Scales. A Review of Existing Datasets. Zenodo. 2020. <https://doi.org/10.5281/zenodo.4390460>

DI MAURO, A.; COMINOLA, A.; CASTELLETTI, A.; DI NARDO, A. Urban Water Consumption at Multiple Spatial and Temporal Scales. A Review of Existing Datasets. **Water**, v. 13, n. 36, 2021. <https://doi.org/10.3390/w13010036>

DURMUSOGLU, A. Updating technology forecasting models using statistical control charts. **Kybernetes**, v. 47, n. 4, p. 672-688, 2018. <https://doi.org/10.1108/K-04-2017-0144>

DZIMIŃSKA, P.; DRZEWIECKI, S.; RUMAN, M.; KOSEK, K.; MIKOŁAJEWSKI, K.; LICZNAR, P. The Use of Cluster Analysis to Evaluate the Impact of COVID-19 Pandemic on Daily Water Demand Patterns. **Sustainability**, v. 13, n. 11, p. 5772, 2021. <https://doi.org/10.3390/su13115772>

ELEVLI, S.; UZGÖREN, N.; BINGÖL, D.; ELEVLI, B. Drinking water quality control: control charts for turbidity and pH. **Journal of Water, Sanitation and Hygiene for Development**, v. 6, n. 4, p. 511-518, 2016. <https://doi.org/10.2166/washdev.2016.016>

FREITAS, L. L. G., HENNING, E.; KALBUSCH, A.; KONRATH, A. C.; WALTER, O. M. F. C. Analysis of water consumption in toilets employing Shewhart, EWMA, and Shewhart-EWMA combined control charts. **Journal of Cleaner Production**, v. 233, p. 1146-1157, 2019. <https://doi.org/10.1016/j.jclepro.2019.06.114>

FUENTES, H.; MAURICIO, D. Smart water consumption measurement system for houses using IoT and cloud computing. **Environmental Monitoring and Assessment**, v. 192, n. 602, 2020.

GARCIA, J.; SALFER, L. R.; KALBUSCH, A.; HENNING, E. Identifying the Drivers of Water Consumption in Single-Family Households in Joinville, Southern Brazil. **Water**, v. 11, 2019. <https://doi.org/10.3390/w11101990>

GARCIA, J. **Determinação dos fatores influentes no consumo de água em residências unifamiliares de Joinville**. Dissertação (Mestrado em Engenharia Civil) – Centro de Ciências Tecnológicas, Universidade do Estado de Santa Catarina, Joinville, 2020.

GARGANO, R.; DI PALMA, F.; DE MARINIS, G.; GRANATA, F.; GRECO, R. A stochastic approach for the water demand of residential end users. **Urban Water Journal**, v. 13, p. 569-582, 2016.

GATO-TRINIDAD, S.; JAYASURIYA, N.; ROBERTS, P. Understanding urban residential end uses of water. **Water Science and Technology**, v. 64, n. 1, p. 36-42, 2011. <https://doi.org/10.2166/wst.2011.436>

GAUTAM, J.; CHAKRABARTI, A.; AGARWAL, S.; SINGH, A.; GUPTA, S.; SINGH, J. Monitoring and forecasting water consumption and detecting leakage using an IoT system. **Water Supply**, v. 20, n. 3, p. 1103-1113, 2020.

GIL, A. C.. **Como elaborar projetos de pesquisa**. 4 ed. São Paulo: Atlas, 2002.

GOMES, I. C.; OLIVEIRA, C. L.; MINGOTI, S. A. A novel experience in the use of control charts for the detection of nosocomial infection outbreaks. **Clinics**, v. 66, n. 10, p.1681-1689, 2011.

GRAHAM, M. A.; CHAKRABORTI, S.; HUMAN, S. W. A Nonparametric EWMA Sign Chart for Location Based on Individual Measurements. **Quality Engineering**, v. 23, p. 227–241, 2011. <https://doi.org/10.1080/08982112.2011.575745>

GUTTORP, P. **Stochastic Modelling of Scientific Data**. Londres: Chapman & Hall, 1995. 384 p.

HAN, W.; ANG, L. Analysis of the use of smart meter data by energy companies in various countries. **IOP Conference Series: Materials Science and Engineering**, Bristol, v. 439, 2018.

HAWKINS, D. M. A Fast Accurate Approximation for Average Run Lengths of CUSUM Control Charts. **Journal of Quality Technology**, v. 24, p. 31-43, 1992. <https://doi.org/10.1080/00224065.1992.11979372>

HENNING, E.; WALTER, O. M. F. C.; SOUZA, N. S. DE; SAMOHYL, R. W. Um Estudo para a Aplicação de Gráficos de Controle Estatístico de Processo em Indicadores de Qualidade da Água Potável. **Sistemas & Gestão**, v. 9, n. 1, p. 2-13, 2014a. <https://doi.org/10.7177/sg.2014.V9.N1.A1>

HENNING, E.; KONRATH, A. C.; WALTER, O. M. F. C.; ALVES, C. C.; SAMOHYL, R. W. Aplicação de um gráfico cusum binomial no monitoramento de um indicador de água potável. **Produção Online**, v. 14, n. 1, p. 84-114, 2014b. <https://doi.org/10.14488/1676-1901.v14.i1.1452>

HIRAKATA, V. N.; BRANCO MANCUSO, A. C.; DE JEZUS CASTRO, S. M. Teste de Hipóteses: Perguntas que você sempre quis fazer, mas nunca teve coragem. **Clinical & Biomedical Research**, v. 39, n. 2, p. 181-185, 2019. <https://doi.org/10.4322/2357-9730.93649>

HUNTER, J. S. The exponentially weighted moving average. **Journal of Quality Technology**, v. 18, p. 203-210, 1986.

HYNDMAN, R. J.; KHANDAKAR, Y. Automatic time series forecasting: the forecast package for R. **Journal of Statistical Software**, v. 26, n. 3, p. 1-22, 2008.

IGLESIAS, C.; SANCHO, J.; PIÑEIRO, J. I.; MARTÍNEZ, J.; PASTOR, J. J.; TABOADA, J. Shewhart-type control charts and functional data analysis for water quality analysis based on a

global indicator. **Desalination and Water Treatment**, v. 57, n. 6, p. 2669-2684, 2016. <https://doi.org/10.1080/19443994.2015.1029533>

JUNG, D.; KANG, D.; LIU, J.; LANSEY, K. Improving the rapidity of responses to pipe burst in water distribution systems: a comparison of statistical process control methods. **Journal of Hydroinformatics**, v. 17, n. 2, p. 307-328, 2015.

JUNG, D.; LANSEY, K. Water Distribution System Burst Detection Using a Nonlinear Kalman Filter. **Journal of Water Resources Planning and Management**, v. 141, n. 5, p. 04014070-2, 2015.

LEOTTI, V. B.; BIRCK, A. R.; RIBOLDI, J. Comparação dos Testes de Aderência à Normalidade Kolmogorov-Smirnov, Anderson-Darling, Cramer-Von Mises e Shapiro-Wilk por Simulação. In: SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRONÔMICA, 11, 2005, Londrina. **Anais**. Florianópolis: UFSC; 2005.

LEOTTI, V. B.; MANCUSO, A. C. B.; BORGES, R. B.; DE JEZUS CASTRO, S. M.; HIRAKATA, V. N.; CAMEY, S. A. Modelagem estatística: Perguntas que você sempre quis fazer, mas nunca teve coragem. **Clinical & Biomedical Research**, v. 39, n. 4, p. 356-363, 2020. <https://doi.org/10.22491/2357-9730.98944>

LI, R.; HUANG, H.; XIN, K.; TAO, T. A review of methods for burst/leakage detection and location in water distribution systems. **Water Science & Technology: Water Supply**, v. 15, n. 3, p. 429-441, 2015.

LIU, Z.; KLEINER, Y. State of the art review of inspection technologies for condition assessment of water pipes. **Measurement**, v. 46, p. 1-15, 2013.

LOUREIRO, D.; ALEGRE, H.; COELHO, S. T.; MARTINS, A.; MAMADE, A. A new approach to improve water loss control using smart metering data. **Water Supply**, v. 14, n. 4, p. 618-625, 2014.

LOUZADA, F.; DINIZ, C.; FERREIRA, P.; FERREIRA, E. **Controle Estatístico de Processos**. Rio de Janeiro: LTC, 2013.

LUCAS, J. M.; SACCUCCI, M. S. Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements. **Technometrics**, v. 32, n. 1, p. 1-12, 1990. <https://doi.org/10.2307/1269835>

MANCUSO, A. C. B.; CASTRO, S. M. DE J.; GUIMARÃES, L. S. P.; LEOTTI, V. B.; HIRAKATA, V. N.; CAMEY, S. A. Estatística Descritiva: Perguntas que você sempre quis fazer, mas nunca teve coragem. **Clinical & Biomedical Research**, v. 38, n. 4, p. 414-418, 2019. <https://doi.org/10.4322/2357-9730.89242>

MANN, P. S. **Introdução à Estatística**. 8 ed. Rio de Janeiro: LTC, 2015.

MARGAVIO, T. M.; CONERLY, M. D; WOODALL, W. H.; DRAKE, L. G. Alarm rates for quality control charts. **Statistics & Probability Letters**, v. 24, p. 219-224, 1995.

MARTINS, R. A. Abordagens quantitativa e qualitativa. In: MIGUEL, P. A. C. **Metodologia de pesquisa em engenharia de produção e gestão de operações**. 2. ed. Rio de Janeiro: Elsevier, 2012, cap. 3, p. 47-63.

MEYER, B. E.; JACOBS, H. E.; ILEMOBADE, A. Classifying household water use into indoor and outdoor use from a rudimentary data set: a case study in Johannesburg, South Africa. **Journal of Water, Sanitation and Hygiene for Development**, v. 11, n. 3, p. 423-431, 2021. <https://doi.org/10.2166/washdev.2021.229>

MIHALOVITS, M.; KEMÉNY, S. Regression control chart with unknown parameters for detection of out-of-trend results in pharmaceutical on-going stability studies. **Journal of Pharmaceutical and Biomedical Analysis**, v. 188, 2020. <https://doi.org/10.1016/j.jpba.2020.113375>

MINI, C.; HOGUE, T. S.; PINCETL, S. Patterns and controlling factors of residential water use in Los Angeles, California. **Water Policy**, v. 16, n. 6, p. 1054-1069, 2014. <https://doi.org/10.2166/wp.2014.029>

MIRANDA, R. G. **Um modelo para a análise da capacidade de processos com ênfase na transformação de dados**. Dissertação (Mestrado em Engenharia de Produção) – Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis, 2005.

MONTGOMERY, D. C. **Introdução ao Controle Estatístico da Qualidade**. 7 ed. Rio de Janeiro: LTC, 2016.

MONTGOMERY, D. C.; RUNGER, G. C. **Estatística Aplicada e Probabilidade Para Engenheiros**. 6 ed. Rio de Janeiro: LTC, 2016.

MORABITO NETO, R.; PUREZA, V. Modelagem e simulação. In: MIGUEL, P. A. C. **Metodologia de pesquisa em engenharia de produção e gestão de operações**. 2. ed. Rio de Janeiro: Elsevier, 2012, cap. 8, p. 169-198.

MORITZ, S.; BARTZ-BEIELSTEIN, T. imputeTS: Time Series Missing Value Imputation in R. **The R Journal**, v. 9, n. 1, p. 207-218, 2017. <https://doi.org/10.32614/RJ-2017-009>

MOSER, G.; PAAL, S. G.; SMITH, I. F. C. Performance comparison of reduced models for leak detection in water distribution networks. **Advanced Engineering Informatics**, v. 29, n. 3, p. 714-726, 2015. <https://doi.org/10.1016/j.aei.2015.07.003>

NARAYANAN, L. K.; SANKARANARAYANAN, S. IoT-based water demand forecasting and distribution design for smart city. **Journal of Water and Climate Change**, v. 11, n. 4, p. 1411-1428, 2020.

NAVIDI, W. C. **Statistics for Engineers and Scientists**. 4 ed. Nova Iorque: McGraw-Hill Science/Engineering/Math, 2014.

OKORO, C.; HUANG, H.; CANNON, S.; LOW, D.; LISTON, D. E.; RICHARDS, M. J.; LENDVAY, T. S. The pudendal nerve block for ambulatory urology: What's old is new again. A quality improvement project. **Journal of Pediatric Urology**, v. 16, n. 5, 2020. <https://doi.org/10.1016/j.jpuro.2020.07.025>

OYEBODE, O.; IGHRAVWE, D. E. Urban Water Demand Forecasting: a comparative evaluation of conventional and soft computing techniques. **Resources**, v. 8, n. 3, p. 156, 2019. <https://doi.org/10.3390/resources8030156>

PALADINI, E. P. **Gestão da Qualidade: Teoria e Prática**. 4 ed. São Paulo: Atlas, 2019.

PATABENDIGE, S.; CARDELL-OLIVER, R.; WANG, R.; LIU, W. Detection and interpretation of anomalous water use for non-residential customers. **Environ. Modell. Software**, v. 100, p. 291-301, 2018. <https://doi.org/10.1016/j.envsoft.2017.11.028>

PERLA, R. J.; PROVOST, S. M.; PARRY, G. J.; LITTLE, K.; PROVOST, L. P. Understanding variation in covid-19 reported deaths with a novel Shewhart chart application. **International Journal for Quality in Health Care**, v. 33, n. 1, p. 1-8, 2020. <https://doi.org/10.1093/intqhc/mzaa069>

R CORE TEAM. **R: A language and environment for statistical computing**. Viena: R Foundation for Statistical Computing, 2020.

RAMOS, E. M. L. S.; ALMEIDA, S. S.; ARAÚJO, A. R. **Controle Estatístico da Qualidade**. Porto Alegre: Bookman, 2013.

RATHNAYAKA, K.; MALANO, H.; ARORA, M.; GEORGE, B.; MAHEEPALA, S.; NAWARATHNA, B. Prediction of urban residential end-use water demands by integrating known and unknown water demand drivers at multiple scales I: Model development. **Resources, Conservation and Recycling**, v. 117, p. 85-92, 2017. <https://doi.org/10.1016/j.resconrec.2016.11.014>

RAZALI, N. M.; WAH, Y. B. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. **Journal of Statistical Modeling and Analytics**, v. 2, n. 1, p. 21-33, 2011.

RIZZO, E. B.; COUSIN, F. A.; LUCCA, R. M.; LAUTENSCHLAGER, S. R. Autonomous metering system for monitoring water consumption. **Journal of Water Supply: Research and Technology-Aqua**, 2021.

ROBERTS, S. W. Control chart tests based on Geometric Moving Averages. **Technometrics** v. 1, 1959.

ROBERTS, N. A.; ALEXANDER, K.; WYLD, D.; JANDA, M. Statistical Process Control assessed implementation fidelity of Patient Reported Outcome Measures (PROMs) in routine care. **Journal of Clinical Epidemiology**, v. 127, p. 76-86, 2020. <https://doi.org/10.1016/j.jclinepi.2020.06.022>

ROBINSON, P. B.; HO, T. Y. Average Run Length of Geometric Moving Average Charts by Numerical Methods. **Technometrics**, v. 20, p. 85-93, 1978.

RODRIGUES, V. **Como máquinas aprendem: Volume II... Algoritmos clássicos de machine learning**. Florença: Publicação Independente, 2021.

ROMANO, M.; KAPELAN, Z.; SAVIĆ, D. A. Real-time leak detection in water distribution systems. **Proceedings of the 12th Annual Conference on Water Distribution Systems Analysis (WDSA)**, Tucson, p. 1074-1082, 2011.

ROMANO, M.; WOODWARD, K.; KAPELAN, Z. Statistical Process Control Based System for Approximate Location of Pipe Bursts and Leaks in Water Distribution Systems. **Procedia Engineering**, v. 186, p. 236-243, 2017.

ROSA, L. C. **Introdução ao Controle Estatístico de Processos**. 2 ed. Porto Alegre: Editora UFSM, 2015.

ROSS, S. **Introduction to Probability and Statistics for Engineers and Scientists**. 4 ed. Cambridge: Elsevier/Academic Press, 2009.

SAMOHYL, R. W. **Controle estatístico de qualidade**. Rio de Janeiro: Elsevier, 2009.

SANUSI, R. A.; CHONG, Z. L.; MUKHERJEE, A.; XIE, M. Distribution-free hybrid schemes for process surveillance with application in monitoring chlorine content of water. **Chemometrics and Intelligent Laboratory Systems**, v. 206, 2020. <https://doi.org/10.1016/j.chemolab.2020.104099>

SCRUCCA, L. qcc: an R package for quality control charting and statistical process control. **R News**, v. 4, n. 1, p. 11-17, 2004.

SEIER, E. **Comparison of tests for univariate normality**. InterStat. 2002. Disponível em: <https://www.researchgate.net/publication/228965974_Comparison_of_tests_of_univariate_normality>. Acesso em: 23 de jun. de 2021.

SIEGMUND, D. **Sequential Analysis: Tests and Conference Intervals**. Nova Iorque: Springer-Verlag, 1985. <https://doi.org/10.1007/978-1-4757-1862-1>

SILVEIRA, D. T.; CÓRDOVA, F. P. A Pesquisa Científica. In: GERHARDT, T. E.; SILVEIRA, D. T. **Métodos de Pesquisa**. Porto Alegre: Editora da UFRGS, 2009, p. 33-44.

SMETI, E. M.; THANASOULIAS, N. C.; KOUSOURIS, L. P.; TZOUMERKAS, P. C. An approach for the application of statistical process control techniques for quality improvement of treated water. **Desalination**, v. 213, n. 1-3, p. 273-281, 2007.

VAUGHN, R. C. **Quality assurance**. Iowa: Iowa State University Press, 1990.

VOULIS, N.; WARNIER, M.; BRAZIER, F. M. Understanding spatio-temporal electricity demand at different urban scales: A data-driven approach. **Applied Energy**, v. 230, p. 1157-1171, 2018.

WADSWORTH, H. M.; STEPHENS, K. S.; GODFREY, A. B. **Modern methods for quality control and improvement**. 2 ed. Nova Iorque: John Wiley & Sons, 2001.

WANG, Y.; CHEN, Q.; HONG, T.; KANG, C. Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges. **IEEE Transactions on Smart Grid**, v. 10, n. 3, p. 3125-3148, 2019.

WANG, Q.; WANG, X. Moving to economic growth without water demand growth – a decomposition analysis of decoupling from economic growth and water use in 31 provinces of China. **Science of the Total Environment**, v. 726, 2020.

WEI, Y.; WANG, Z.; WANG, H.; YAO, T.; LI, Y. Promoting inclusive water governance and forecasting the structure of water consumption based on compositional data: a case study of Beijing. **Science of the Total Environment**, v. 634, p. 407-416, 2018.

WICKHAM, H. **ggplot2: Elegant Graphics for Data Analysis**. Nova Iorque: Springer-Verlag, 2016.

WILLIS, R. M.; STEWART, R. A.; GIURCO, D. P.; TALEBPOUR, M. R.; MOUSAVINEJAD, A. End use water consumption in households: Impact of sociodemographic factors and efficient devices. **Journal of Cleaner Production**, v. 60, p. 107-115, 2013.

WOODALL, W. H., MARAGAH, H. D. Control charts using exponential smoothing techniques. **Technometrics**, v. 32, p. 1-29, 1990.

WORTHAM, A. W.; HEINRICH, G. F. Control charts using exponential smoothing techniques. **Annual Conference Transactions**, v. 26, Milwaukee, p. 451-458, 1972.

WU, Y.; LIU, S. A review of data-driven approaches for burst detection in water distribution systems. **Urban Water Journal**, v. 14, p. 972-983, 2017.

ZAGO, V. A. **Avaliação da aplicação de gráficos de controle com memória em uma indústria de papel e embalagens**. Dissertação (Mestrado em Engenharia de Produção) – Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis, 2009.

ZAN, T. T. T.; LIM, H. B.; WONG, K.; WHITTLE, A. J.; LEE, B. Event Detection and Localization in Urban Water Distribution Network. **IEEE Sensors Journal**, v. 14, n. 12, p. 4134-4142, 2014. <https://doi.org/10.1109/JSEN.2014.2358842>

APÊNDICE A – LISTA DE ARTIGOS CONSIDERADOS NA REVISÃO

Ano	Periódico	Título	Escala de consumo	Escala temporal	Obtenção dos dados	Método de análise	Origem
1979	Journal Water Resources Research	An analysis of residential demand for water using micro time-series data	Residencial	Diária	Hidrômetro	Regressão	Estados Unidos
1979	J. Am. Water Resour. Assoc.	Effect of price on the residential demand for water within an agency	Urbano	Mensal	Medidor de vazão	Regressão	Estados Unidos
1982	Journal Evaluation Review	Response To Water Conservation Campaigns: An Exploratory Look	Residencial	Mensal	Hidrômetro	Descritivo	Estados Unidos
1983	Sociol. Perspect.	Saving water: A Causal Model of Household Conservation	Residencial	Mensal	Hidrômetro	Regressão	Estados Unidos
1989	Journal / American Water Works Association	Demand management factors in residential water use: The southern Arizona experience	Urbano	Mensal	Medidor de vazão	Regressão	Estados Unidos
1991	Water and Environment Journal	Insights into Domestic Demand from a Metering Trial	Urbano	Diária	Medidor de vazão	Descritivo	Reino Unido
1991	Journal Water Resources Management	Peak coefficients of household potable water supply	Urbano	Subdiária	Medidor de vazão	Descritivo	Itália
1995	Water and Environment Journal	A Methodology for Surveying Domestic Water Consumption	Residencial	Subdiária	Hidrômetro	Descritivo	Reino Unido
1996	Journal of Water Resources Planning and Management	Intensity, duration, and frequency of residential water demands	Residencial	Subdiária	Medidor de consumo com data logger	Descritivo	Estados Unidos
2001	Conference Progress in Water Resources	Instantaneous residential water demand as stochastic point process	Residencial	Subdiária	Medidor de consumo com data logger	Regressão	Itália

Ano	Periódico	Título	Escala de consumo	Escala temporal	Obtenção dos dados	Método de análise	Origem
2002	Water Studies	Dynamic simulation of water distribution systems with instantaneous demands	Residencial	Subdiária	Medidor de consumo com data logger	Regressão	Estados Unidos
2003	Journal Hydro 2000: Interactive Hydrology	Domestic Water Use Study. In Perth, Western Australia 1998-2001	Residencial	Diária	Medidor de consumo com data logger	Descritivo	Australia
2003	American Water Works Association Research Foundation	Random demands, travel times and water quality in deadends	Residencial	Subdiária	Medidor de consumo com data logger	Inferencial	Estados Unidos
2003	Journal Water Resources Management	A stochastic model for representing drinking water demand at residential level	Urbano	Subdiária	Medidor de vazão	Regressão	Itália
2005	Australasian Journal of Water Resources	A simple time series approach to modelling urban water demand. Australas	Urbano	Diária	Medidor de vazão	Regressão	Australia
2007	8th Annual Water Distribution Systems Analysis Symposium	Effect of time step and data aggregation on cross correlation of residential demands	Residencial	Subdiária	Medidor de consumo com data logger	Regressão	Estados Unidos
2008	Journal of the American Water Resources Association	Residential water demand management: Lessons from Aurora, Colorado	Residencial	Mensal	Hidrômetro	Regressão	Estados Unidos
2008	J. Water Resour. Plan. Manag.	Spatial and temporal scaling properties of water demand	Residencial	Subdiária	Medidor de consumo com data logger	Processo estocástico	Itália
2008	Journal of Water Resources Planning and Management	Urban Water Demand Forecasting with a Dynamic Artificial Neural Network Model	Urbano	Diária	Medidor de vazão	Machine learning	Estados Unidos
2009	Water Policy	Residential water demand modeling in Queensland, Australia: A comparative panel data approach	Urbano	Mensal	Medidor de vazão	Regressão	Austrália

Ano	Periódico	Título	Escala de consumo	Escala temporal	Obtenção dos dados	Método de análise	Origem
2010	Water Sci. Technol. Water Supply	Implementation of an on-line artificial intelligence district meter area flow meter data analysis system for abnormality detection: A case study	Urbano	Subdiária	Medidor de vazão	Machine learning	Reino Unido
2010	Journal of Hydrology	Predictive models for forecasting hourly urban water demand	Urbano	Subdiária	Medidor de vazão	Machine learning	Espanha
2011	Conference WIT Transactions on the Built Environment	Characterizing maximum residential water demand	Urbano	Subdiária	Medidor de vazão	Regressão	Austrália
2013	Urban Water Journal	Smart meter enabled disaggregation of urban peak water demand: Precursor to effective urban water planning	Residencial	Diária	Smart meter	Descritivo	Australia
2013	World Environmental and Water Resources Congress	Residential water demand analysis due to water meter installation in California	Residencial	Mensal	Hidrômetro	Inferencial	Estados Unidos
2013	8th Intelligent Sensor Networks and Information Processing	Discovering water use activities for smart metering	Residencial	Subdiária	Smart meter	Machine learning	Australia
2013	Environ. Model. Softw.	A fully adaptive forecasting model for short-term drinking water demand.	Urbano	Subdiária	Medidor de vazão	Machine learning	Holanda
2014	Institute of Electrical and Electronics Engineers Inc.	Exploiting self-similarity for change detection	Urbano	Subdiária	Medidor de vazão	Teste de detecção de mudanças	Espanha
2014	26th Chinese Control and Decision Conference	Urban water demand forecasting by LS-SVM with tuning based on elitist teaching-learning-based optimization	Urbano	Subdiária	Medidor de vazão	Machine learning	China
2014	Procedia Engineering	Detecting pipe bursts using Heuristic and CUSUM methods	Urbano	Subdiária	Medidor de vazão	Gráficos de controle	Holanda

Ano	Periódico	Título	Escala de consumo	Escala temporal	Obtenção dos dados	Método de análise	Origem
2015	Journal Desalination and Water Treatment	Field application of waterworks automated meter reading systems and analysis of household water consumption	Residencial	Subdiária	Smart meter	Descritivo	Coreia
2015	Water Sci. Technol. Water Supply	Linking water consumption smart metering with census data to improve demand management	Residencial	Diária	Smart meter	Descritivo	Portugal
2015	Water Resour. Res.	Water consumption patterns as a basis for water demand modeling.	Urbano	Mensal	Medidor de vazão	Machine learning	Israel
2015	J. Water Resour. Plann. Manage.	Water Distribution System Burst Detection Using a Nonlinear Kalman Filter	Urbano	Subdiária	Medidor de vazão	Gráficos de controle	Estados Unidos
2015	Journal of Hydroinformatics	Improving the rapidity of responses to pipe burst in water distribution systems: a comparison of statistical process control methods	Urbano	Subdiária	Medidor de vazão	Gráficos de controle	EUA / Holanda / Itália
2016	Journal of Water Sanitation and Hygiene for Development	Assessing the impact of property size on residential water use for selected neighbourhoods in Lilongwe, Malawi	Residencial	Mensal	Hidrômetro	Descritivo	África do Sul
2016	Journal Water Science and Technology: Water Supply	Application of machine learning techniques to predict anomalies in water supply networks	Urbano	Subdiária	Medidor de vazão	Machine learning	Holanda
2016	Journal Water Science and Technology: Water Supply	A stochastic model for daily residential water demand	Urbano	Subdiária	Medidor de vazão	Processo estocástico	Itália
2016	Journal Water and Environment Journal	Residential water demand analysis of a Low-Income Rate Assistance Program in California, United States	Residencial	Mensal	Hidrômetro	Descritivo	Estados Unidos
2016	8th Int. Congr. Environ. Model. Softw.	Profiling residential water users' routines by eigenbehavior modelling	Residencial	Subdiária	Smart meter	Machine learning	Reino Unido / Suíça

Ano	Periódico	Título	Escala de consumo	Escala temporal	Obtenção dos dados	Método de análise	Origem
2017	Journal article Sustainability	Household smart water metering in Spain: Insights from the experience of remote meter reading in alicant	Residencial	Subdiária	Smart meter	Descritivo	Espanha
2017	Computing and Control for the Water Industry	Pattern Recognition in Residential End Uses of Water Using Artificial Neural Networks and Other Machine Learning Techniques	Residencial	Subdiária	Smart meter	Machine learning	Espanha
2017	International Conference on Data Science and Advanced Analytics	Predictive classification of water consumption time series using non-homogeneous markov model	Urbano	Subdiária	Smart meter	Machine learning	França
2017	Sci. Adv. 3.	Changes in water consumption linked to heavy news media coverage of extreme climatic events.	Urbano	Mensal	Medidor de vazão	Regressão	Estados Unidos
2017	J. Water Resour. Plan. Manag.	Models for Generating Household Water Demand Pulses: Literature Review and Comparison	Residencial	Subdiária	Smart meter	Regressão	Itália
2017	IEEE SmartWorld	Smart WDS management: Pipe burst detection using real-time monitoring data	Urbano	Subdiária	Medidor de vazão	Gráficos de controle	Estados Unidos
2018	Journal of Cleaner Production	Segmentation analysis of residential water-electricity demand for customized demand-side management programs	Residencial	Subdiária	Smart meter	Inferencial	Estados Unidos
2018	Water (Switzerland)	Exploring the statistical and distributional properties of residential water demand at fine time scales	Residencial	Subdiária	Smart meter	Processo estocástico	Grécia
2018	Environ. Model. Softw.	Forecasting urban household water demand with statistical and machine learning methods using large space-time data: A Comparative study	Residencial	Mensal	Hidrômetro	Machine learning	Estados Unidos
2018	EasyChair	Smart Water Demand Forecasting: Learning from the Data	Residencial	Subdiária	Smart meter	Machine learning	Reino Unido

Ano	Periódico	Título	Escala de consumo	Escala temporal	Obtenção dos dados	Método de análise	Origem
2018	Environ. Model. Softw	Re-engineering traditional urban water management practices with smart metering and informatics	Residencial	Subdiária	Smart meter	Machine learning	Australia
2019	Journal of Water Resources Planning and Management	Long-Term Projections of Domestic Water Demand: A Case Study of London and the Thames Valley	Residencial	Diária	Smart meter	Regressão	Reino Unido
2019	Lecture Notes in Computer Science	Calibration of a water distribution network with limited field measures: The case study of Castellammare di Stabia (Naples, Italy)	Urbano	Diária	Medidor de vazão	Machine learning	Itália
2019	Water e-Journal	Smart water metering technology for water management in urban areas	Residencial	Subdiária	Smart meter	Descritivo	Australia
2019	J. Water Resour. Plann. Manage.	Hybrid Statistical Process Control Method for Water Distribution Pipe Burst Detection	Urbano	Subdiária	Medidor de vazão	Gráficos de controle	Estados Unidos
2020	Urban Water J.	Applying human mobility and water consumption data for short-term water demand forecasting using classical and machine learning models	Urbano	Subdiária	Medidor de vazão	Machine learning	Polonia
2020	Environ. Model. Softw.	Smart meters data for modeling and forecasting water demand at the user-level	Residencial	Subdiária	Smart meter	Machine learning	Estados Unidos
2020	Sensors (Switzerland)	A low-cost, open source monitoring system for collecting high temporal resolution water use data on magnetically driven residential water meters	Residencial	Subdiária	Medidor de consumo com data logger	Machine learning	Estados Unidos

ANEXO A – FATORES PARA A CONSTRUÇÃO DE GRÁFICOS DE CONTROLE PARA VARIÁVEIS

n	Gráfico para média				Gráfico para desvio padrão						Gráfico para amplitude			
	Limites de controle			Linha central	Limites de controle						Linha central	Limites de controle		
	A	A ₂	A ₃	c ₄	B ₃	B ₄	B ₅	B ₆	d ₂	D ₁	D ₂	D ₃	D ₄	
2	2,121	1,88	2,659	0,7979	0	3,267	0	2,606	1,128	0	3,686	0	3,267	
3	1,732	1,023	1,954	0,8862	0	2,568	0	2,276	1,693	0	4,358	0	2,574	
4	1,5	0,729	1,628	0,9213	0	2,266	0	2,088	2,059	0	4,698	0	2,282	
5	1,342	0,577	1,427	0,94	0	2,089	0	1,964	2,326	0	4,918	0	2,114	
6	1,225	0,483	1,287	0,9515	0,03	1,97	0,029	1,874	2,534	0	5,078	0	2,004	
7	1,134	0,419	1,182	0,9594	0,118	1,882	0,113	1,806	2,704	0,204	5,204	0,076	1,924	
8	1,061	0,373	1,099	0,965	0,185	1,815	0,179	1,751	2,847	0,388	5,306	0,136	1,864	
9	1	0,337	1,032	0,9693	0,239	1,761	0,232	1,707	2,97	0,547	5,393	0,184	1,816	
10	0,949	0,308	0,975	0,9727	0,284	1,716	0,276	1,669	3,078	0,687	5,469	0,223	1,777	
11	0,905	0,285	0,927	0,9754	0,321	1,679	0,313	1,637	3,173	0,811	5,535	0,256	1,744	
12	0,866	0,266	0,886	0,9776	0,354	1,646	0,346	1,61	3,258	0,922	5,594	0,283	1,717	
13	0,832	0,249	0,85	0,9794	0,382	1,618	0,374	1,585	3,336	1,025	5,647	0,307	1,693	
14	0,802	0,235	0,817	0,981	0,406	1,594	0,399	1,563	3,407	1,118	5,696	0,328	1,672	
15	0,775	0,223	0,789	0,9823	0,428	1,572	0,421	1,544	3,472	1,203	5,741	0,347	1,653	
16	0,75	0,212	0,763	0,9835	0,448	1,552	0,44	1,526	3,532	1,282	5,782	0,363	1,637	
17	0,728	0,203	0,739	0,9845	0,466	1,534	0,458	1,511	3,588	1,356	5,82	0,378	1,622	
18	0,707	0,194	0,718	0,9854	0,482	1,518	0,475	1,496	3,64	1,424	5,856	0,391	1,608	
19	0,688	0,187	0,698	0,9862	0,497	1,503	0,49	1,483	3,689	1,487	5,891	0,403	1,597	
20	0,671	0,18	0,68	0,9869	0,51	1,49	0,504	1,47	3,735	1,549	5,921	0,415	1,585	
21	0,655	0,173	0,663	0,9876	0,523	1,477	0,516	1,459	3,778	1,605	5,951	0,425	1,575	
22	0,64	0,167	0,647	0,9882	0,534	1,466	0,528	1,448	3,819	1,659	5,979	0,434	1,566	
23	0,626	0,162	0,633	0,9887	0,545	1,455	0,539	1,438	3,858	1,71	6,006	0,443	1,557	
24	0,612	0,157	0,619	0,9892	0,555	1,445	0,549	1,429	3,895	1,759	6,031	0,451	1,548	
25	0,6	0,153	0,606	0,9896	0,565	1,435	0,559	1,42	3,931	1,806	6,056	0,459	1,541	

Fonte: Adaptado de Montgomery (2016)

Quando $n > 25$, utilizam-se as Equações (57), (58) e (59) para o gráfico para média, e as Equações (60), (61), (62) e (63) para o gráfico para desvio padrão.

$$A = \frac{3}{\sqrt{n}} \quad (57)$$

$$A_3 = \frac{3}{c_4 \sqrt{n}} \quad (58)$$

$$c_4 = \frac{4(n-1)}{4n-3} \quad (59)$$

$$B_3 = 1 - \frac{3}{c_4 \sqrt{2(n-1)}} \quad (60)$$

$$B_4 = 1 + \frac{3}{c_4 \sqrt{2(n-1)}} \quad (61)$$

$$B_5 = c_4 - \frac{3}{\sqrt{2(n-1)}} \quad (62)$$

$$B_6 = c_4 + \frac{3}{\sqrt{2(n-1)}} \quad (63)$$