

## XXVI SIMPÓSIO BRASILEIRO DE RECURSOS HIDRÍCOS

### **ESTIMATIVA DO CONSUMO DIÁRIO DE ÁGUA EM GALPÕES LOGÍSTICOS: UMA ABORDAGEM COM REGRESSÃO LINEAR MULTIPLA**

*Adriano Henrique Tognato<sup>1</sup>; Tiago Zenker Gireli<sup>2</sup>; Patrícia Dalsoglio Garcia<sup>3</sup>; Vinicius de Carvalho Neiva Pinheiro<sup>4</sup>; Gabriel Gonçalves Silveira<sup>5</sup> & Verena Corazza Vinha<sup>6</sup>*

**Abstract:** This study aims to estimate the daily water consumption in logistics distribution centers (DCs) using a Multiple Linear Regression (MLR) model, incorporating both seasonal and infrastructure variables. The analysis was based on field data collected over 12 consecutive months from nine DCs located in different regions of Brazil. The model considered variables such as average monthly maximum temperature, hydrostatic pressure, floating population, and the type and number of hydraulic devices. The results showed a strong correlation between the independent variables and water consumption, with an adjusted  $R^2$  of 0.768 and a Root Mean Square Error (RMSE) of 443.58 liters. The model demonstrated good predictive accuracy, achieving a Mean Absolute Percentage Error (MAPE) of 11.08%, making it a practical tool for hydraulic planning in logistics facilities. Traditional estimation methods based on per capita consumption were found to significantly overestimate actual consumption, with errors exceeding 600% in some cases. In contrast, the proposed model proved to be more aligned with the Brazilian context, offering more accurate predictions and thus contributing to more efficient water supply planning and better guidance for infrastructure investment.

**Resumo:** O presente trabalho tem como objetivo estimar o consumo diário de água em centros de distribuição logística (CD's) por meio de um modelo de Regressão Linear Múltipla (MLR), utilizando variáveis sazonais e de infraestrutura. O estudo foi desenvolvido com dados de campo obtidos em nove CD's localizados em diferentes regiões do Brasil, ao longo de 12 meses consecutivos. Foram consideradas variáveis como temperatura máxima média mensal, pressão hidrostática máxima, população flutuante, tipo e quantidade dos equipamentos hidráulicos. Os resultados indicaram forte correlação entre as variáveis independentes e o consumo, com  $R^2$  ajustado de 0,768 e RMSE de 443,58 litros. O modelo demonstrou capacidade preditiva satisfatória, com MAPE de 11,08%, destacando-se como ferramenta prática para o dimensionamento hídrico de empreendimentos logísticos. Verificou-se também que métodos tradicionais de estimativa baseados em consumo per capita superestimam significativamente o consumo real, com erros superiores a 600% em alguns casos. O modelo proposto, por sua vez, mostrou-se mais aderente à realidade brasileira, oferecendo estimativas mais precisas e por consequência, contribuindo para eficiência no planejamento de sistemas de abastecimento e orientação para investimentos em infraestrutura.

**Palavras-Chave:** consumo de água, centros de distribuição logística, regressão linear múltipla, modelagem matemática.

1) UNICAMP: Faculdade de Engenharia Civil, Arquitetura e Urbanismo – Departamento de Recursos Hídricos: a208330@dac.unicamp.br

2) UNICAMP: Faculdade de Engenharia Civil, Arquitetura e Urbanismo – Departamento de Recursos Hídricos: zenker@unicamp.br

3) UNICAMP: Faculdade de Engenharia Civil, Arquitetura e Urbanismo – Departamento de Recursos Hídricos: pdgargia@unicamp.br

4) UNICAMP: Faculdade de Engenharia Civil, Arquitetura e Urbanismo – Departamento de Recursos Hídricos: vpinheiro@unicamp.br

5) UNICAMP: Faculdade de Engenharia Civil, Arquitetura e Urbanismo – Departamento de Recursos Hídricos: g167943@dac.unicamp.br

6) UNICAMP: Faculdade de Engenharia Civil, Arquitetura e Urbanismo – Departamento de Recursos Hídricos: v225063@dac.unicamp.br

## INTRODUÇÃO:

Um centro de distribuição logística (CD) é uma instalação estrategicamente localizada e projetada para receber, armazenar, organizar e expedir mercadorias de forma ágil para atender às demandas de diversos pontos de entrega. Sua principal finalidade é otimizar o fluxo de produtos dentro da cadeia de suprimentos, reduzindo prazos de entrega, custos operacionais e aumentando a disponibilidade de itens para os consumidores finais (BALLOU; SRIVASTAVA, 2007).

Nos últimos anos, tem-se observado um crescimento expressivo na construção e expansão de CD's, impulsionado pelo avanço do comércio eletrônico e pela necessidade de maior capilaridade nas entregas. Esse cenário tem motivado empresas a realizarem investimentos significativos tanto em infraestrutura física, visando ampliar a capacidade logística e acomodar um número maior de colaboradores, quanto em tecnologias voltadas à automação, à gestão da informação e à escolha estratégica de localizações. Tais medidas buscam atender, com maior eficiência, às demandas dos mercados regionais e nacionais (DE FREITAS, 2025).

Segundo Vianna e Depexe (2013), investimentos em infraestrutura necessitam de indicadores e métodos de previsão de demanda eficazes, a fim de evitar que os projetos sejam subdimensionados ou superdimensionados. Logo, a previsão da demanda de consumo de água é fundamental para o planejamento e a gestão eficiente dos recursos hídricos, especialmente em instalações urbanas e industriais, como é o caso dos centros de distribuição logística.

Ao estimar as necessidades de consumo, torna-se possível otimizar a operação dos sistemas de abastecimento, identificar e reduzir perdas, prevenir situações de escassez e orientar adequadamente os investimentos em infraestrutura. Além disso, a previsão contribui para a formulação de políticas públicas e privadas mais eficazes e sustentáveis, promovendo o uso racional da água frente aos cenários de crescimento populacional e mudanças climáticas (TOMAZ, 2000).

No contexto brasileiro, é comum que o dimensionamento de Sistemas de Abastecimento de Água (SAA), sejam eles públicos ou privados, baseie-se unicamente no consumo per capita, expresso em litros por pessoa por dia (L/pessoa.dia). Embora a quantidade de pessoas seja o principal fator de influência na demanda, esse critério isolado não abrange toda a complexidade do consumo de água. Diversos outros fatores também exercem influência significativa e devem ser considerados no processo de previsão, como as condições climáticas, os fatores socioeconômicos, os tipos de uso, a pressão disponível no sistema de abastecimento e as características construtivas das edificações, incluindo o tipo e a quantidade dos equipamentos hidráulicos instalados (FEIL; HAETINGER, 2014).

Neste sentido, a modelagem matemática é uma ferramenta essencial na gestão hídrica, pois viabiliza a obtenção de indicadores e ferramentas de dimensionamento que fundamentam a expansão e a regulação do fornecimento de água. De acordo com Trautwein Jr (2024), a utilização de modelos matemáticos para o estudo do consumo de água na maioria dos casos possui dois objetivos: Determinar os melhores métodos de previsão e as variáveis que apresentam melhor correlação com o consumo de água.

Bassanezi (2002) define a modelagem matemática como a arte de transformar uma situação real e atual em um problema matemático, com o objetivo de encontrar soluções por meio de uma linguagem acessível. Devido à complexidade das funções envolvidas, é comum o uso de softwares nesse processo. Atualmente, a maioria dos estudos sobre modelagem matemática voltados à previsão de consumo de água concentra-se em grandes consumidores, como cidades (TRAUTWEIN JR, 2004; VIANNA; DEPEXE, 2013; FEIL; HAETINGER, 2014; SILVA et al., 2016) ou países (ZHOU et al., 2000; SCHLEICH; HILLENBRAND, 2009; YASAR et al., 2012). Assim, observa-se uma escassez

de modelos preditivos voltados a segmentos específicos da indústria, o que evidencia uma lacuna na literatura.

Uma conclusão geral que pode ser extraída da literatura citada é que o modelo de Regressão Linear Múltipla (MLR) se destaca como o mais utilizado nesse tipo de estudo. Isso se deve ao fato de que o consumo de água apresenta relações lineares com as variáveis analisadas. Além disso, a ampla adoção do método também se justifica pela sua aplicação relativamente simples, especialmente por meio do uso de ferramentas computacionais.

Prioritariamente, a Regressão Linear Múltipla (MLR) é uma técnica estatística amplamente utilizada para prever e modelar relações entre variáveis de entrada e saída. De acordo com Frost (2019), a MLR estuda a dependência de uma variável dependente em relação a múltiplas variáveis independentes, com o objetivo de estimar o valor médio de uma série de dados. Seu fundamento baseia-se na aplicação do método dos mínimos quadrados e na formulação de equações lineares que descrevem a relação entre as variáveis dependente e independentes. Maiores detalhes sobre o método podem ser encontrados na obra de Frost (2019).

Diante do exposto, este trabalho tem como objetivo analisar o consumo de água em centros de distribuição logística no Brasil por meio da aplicação da Regressão Linear Múltipla (MLR) e de ferramentas estatísticas associadas. Considerando a crescente demanda por infraestrutura logística e a importância de um dimensionamento hídrico preciso, busca-se desenvolver um modelo preditivo que supere as limitações das estimativas baseadas exclusivamente no consumo per capita. Ao incorporar múltiplas variáveis explicativas, o estudo visa oferecer uma abordagem mais realista e robusta para estimar o consumo diário de água, contribuindo para a gestão eficiente dos recursos hídricos em um setor estratégico e em expansão.

## DELINAMENTO EXPERIMENTAL:

Previsão é um método que usa séries temporais, ou seja, dados coletados ao longo do tempo para combinar informações estatisticamente e gerar uma estimativa do que pode acontecer no futuro. Neste trabalho, a previsão almejada é o consumo diário em litros (L/dia) para centros de distribuição logísticos – CD's, que dentro do modelo de relação linear múltipla proposto foi definido como a variável dependente ( $Y$ ). As variáveis independentes ( $X_1, X_2, X_n$ ) foram classificadas em duas categorias: variáveis sazonais e variáveis de infraestrutura. A Tabela 1 apresenta a listagem das variáveis que são utilizadas na Equação 1.

Tabela 1 – Variáveis do modelo.

Variáveis	Tipo	Termo	Classificação	Unidade
Consumo Diário Médio Mensal	Dependente	$Y$	Objetivo	Litros / Dia
Temperatura Máxima Média Mensal	Independente	$X_1$	Sazonal	° C
Carga Mensal Distribuída	Independente	$X_2$	Sazonal	m <sup>3</sup>
População Atendida	Independente	$X_3$	Sazonal	Pessoas
Máxima Pressão Hidrostática Disponível	Independente	$X_4$	Infraestrutura	mca
Quantidade de Caixas Acopladas	Independente	$X_5$	Infraestrutura	Unidade
Quantidade de Válvulas de Descarga	Independente	$X_6$	Infraestrutura	Unidade
Quantidade de Chuveiros e Duchas	Independente	$X_7$	Infraestrutura	Unidade
Quantidade de Torneiras	Independente	$X_8$	Infraestrutura	Unidade

$$Y = a + m_1X_1 + m_2X_2 + m_3X_3 + m_4X_4 + m_5X_5 + m_6X_6 + m_7X_7 + m_8X_8 \quad (1)$$

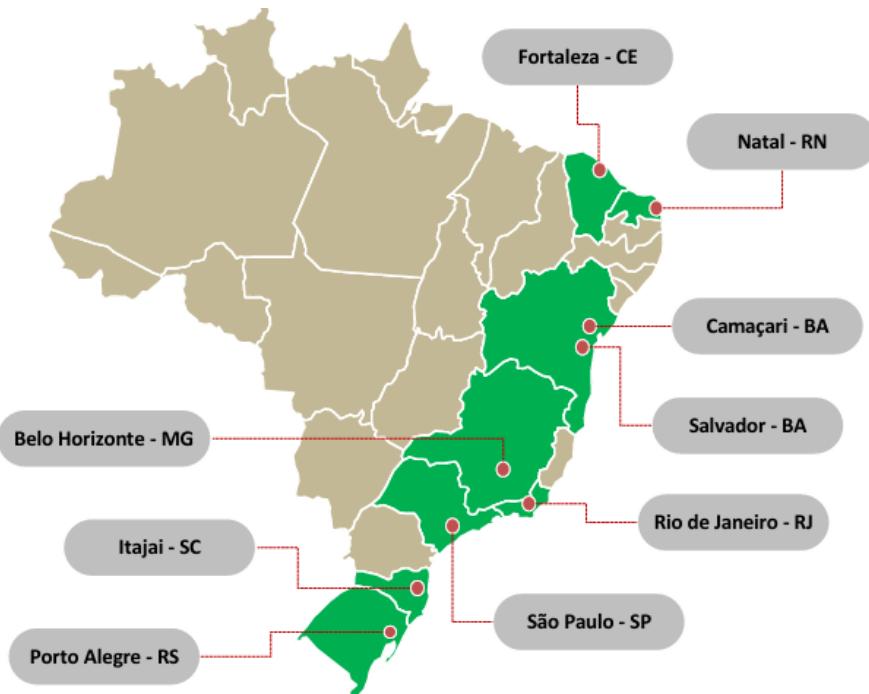
A constante ( $a$ ) denota o intercepto em ( $Y$ ). Os valores dos coeficientes angulares ( $m_1, m_2, m_n$ ) foram calculados usando a função PROJ.LIN disponível no Microsoft Excel. Ao empregar uma ferramenta acessível e amplamente disponível, o modelo proposto torna-se de fácil reprodução, mesmo por profissionais da área que não possuam experiência prévia com softwares estatísticos.

Este trabalho parte da hipótese de que o consumo diário de água em centros de distribuição logística não é determinado apenas pelo número de pessoas que frequentam o local. Acredita-se que outros fatores exerçam influência significativa sobre esse consumo, como as variáveis sazonais previamente mencionadas e diversos parâmetros relacionados à infraestrutura dos galpões. Entre esses parâmetros, destacam-se a quantidade e o tipo de equipamentos instalados, além da pressão hidrostática máxima disponível no sistema de distribuição hidráulico.

Diante disso, para que fosse possível construir um modelo matemático de previsão que seja representativo e confiável, tornou-se necessário realizar o levantamento de dados em múltiplos galpões logísticos com localizações geográficas, características operacionais e estruturais distintas. Essa diversidade de contextos visa garantir que o modelo capture a variabilidade do fenômeno estudado, evitando vieses localizados e aumentando sua capacidade de generalização e aplicação prática em diferentes cenários.

Inicialmente, buscou-se identificar empresas responsáveis pela gestão de galpões logísticos que estivessem dispostas a disponibilizar, por um período mínimo 12 meses, o histórico de consumo de água, o número de colaboradores e o volume de carga distribuída. Além disso, era imprescindível que autorizassem a realização de medições diretas do consumo diário por meio de hidrômetros, ao longo de 1 mês de operação. O objetivo foi desenvolver o modelo preditivo tendo como variável dependente ( $Y$ ) o consumo diário médio extraído do histórico mensal e, a partir disso, avaliar a qualidade das estimativas e a validade da hipótese por meio da comparação com os dados de campo e análise estatística. A Figura 1 apresenta a localização dos 9 CD's que atenderam os requisitos propostos.

Figura 1: Localização dos CD's.



A Tabela 2 exibe os valores levantados em campo para as variáveis relacionadas à infraestrutura, ou seja, aquelas que dependem exclusivamente das características construtivas de cada CD. Optou-se por distinguir dois tipos de dispositivos de descarga sanitária: caixa acoplada e válvula de descarga. Essa separação se justifica pelo fato de que as válvulas de descarga, em geral, consomem significativamente mais água do que as caixas acopladas. Não considerar essa diferença poderia comprometer a precisão e a integridade do modelo proposto.

Tabela 2 – Valores das variáveis de infraestrutura por CD.

Localização	X4 Pressão (mca)	X5 Cx. Acopladas	X6 Válv. de Descarga	X7 Chuveiros	X8 Torneiras
Belo Horizonte - MG	14,00	7	0	6	19
Camaçari - BA	16,00	12	8	8	29
Fortaleza - CE	6,00	18	0	6	21
Itajaí - SC	10,00	8	0	3	12
Natal - RN	9,00	10	0	8	21
Porto Alegre - RS	12,00	14	1	5	17
Rio de Janeiro - RJ	5,00	3	12	7	23
Salvador - BA	5,00	12	2	12	21
São Paulo - SP	10,00	2	16	8	33

A Tabela 3 exibe, além dos valores de consumo médio diário mensal ( $Y$ ), os dados referentes às variáveis sazonais, as quais variam conforme a localidade e o mês. Ressalta-se que, para fins de modelagem, foi implementada a temperatura máxima média mensal, definida como a média das temperaturas máximas diárias ao longo de cada mês. Os dados de temperatura foram obtidos a partir dos “Normais Climatológicos” do Instituto Nacional de Meteorologia (INMET, 2025).

O volume distribuído de cargas corresponde à quantidade total de mercadorias movimentadas e efetivamente expedidas a partir do CD. Essa métrica, expressa em metros cúbicos ( $m^3$ ), reflete diretamente a intensidade das operações logísticas realizadas no CD em determinado intervalo de tempo. A variação sazonal desse volume pode ser influenciada por fatores climáticos, culturais e, sobretudo, por dinâmicas de mercado que são, em grande medida, determinadas pelas características das mercadorias armazenadas, uma vez que diferentes categorias de produtos apresentam demandas específicas ao longo do ano.

A população mensal de um centro de distribuição (CD) é composta por dois grupos: a população fixa, formada pelos colaboradores que atuam continuamente no local, e a população flutuante, que inclui os profissionais externos envolvidos nas operações de entrega e coleta de cargas. Mesmo antes da aplicação da análise estatística, foi possível perceber uma correlação direta entre o volume de cargas movimentadas e a variação na população flutuante.

Tabela 3 – Valores das variáveis sazonais por mês e CD.

Mês	CD:	Y	X1	X2	X3	Mês	CD:	Y	X1	X2	X3
jan/24	Belo Horizonte	3233,33	28,7	4541,29	137	jul/24	Belo Horizonte	4733,33	24,9	2390,15	112
fev/24	Belo Horizonte	3353,33	29,1	4780,3	140	ago/24	Belo Horizonte	3033,33	26,3	2151,14	121
mar/24	Belo Horizonte	3266,67	28,4	4302,27	135	set/24	Belo Horizonte	3766,67	27,9	3346,21	132
abr/24	Belo Horizonte	2966,67	27,6	3107,2	130	out/24	Belo Horizonte	3066,67	28,7	3824,24	137
mai/24	Belo Horizonte	3666,67	25,7	2629,17	117	nov/24	Belo Horizonte	3966,67	27,7	3585,23	130
jun/24	Belo Horizonte	2966,67	24,9	2868,18	112	dez/24	Belo Horizonte	4366,67	28,2	4063,26	134



Mês	CD:	Y	X1	X2	X3	Mês	CD:	Y	X1	X2	X3
jan/24	Camaçari	7933,33	31	3248,53	139	jul/24	Camaçari	5433,33	26,6	1709,75	111
fev/24	Camaçari	8066,67	31,1	3419,51	139	ago/24	Camaçari	5333,33	26,7	1538,78	112
mar/24	Camaçari	7233,33	30,9	3077,56	138	set/24	Camaçari	5633,33	27,7	1880,73	118
abr/24	Camaçari	7333,33	29,6	2906,58	130	out/24	Camaçari	6033,33	29,1	2222,68	127
mai/24	Camaçari	7500	28,1	2564,63	121	nov/24	Camaçari	6633,33	29,8	2393,65	132
jun/24	Camaçari	5133,33	27,1	2051,7	114	dez/24	Camaçari	6566,67	30,6	2735,6	137

Mês	CD:	Y	X1	X2	X3	Mês	CD:	Y	X1	X2	X3
jan/24	Fortaleza	3333,33	31,2	2673,49	110	jul/24	Fortaleza	3800	30,6	1688,52	96
fev/24	Fortaleza	3566,67	31,1	2814,20	107	ago/24	Fortaleza	3933,33	31,1	1407,1	107
mar/24	Fortaleza	2900,00	30,8	2532,78	100	set/24	Fortaleza	4533,33	31,4	1266,39	113
abr/24	Fortaleza	4300,00	30,6	2110,65	96	out/24	Fortaleza	3733,33	31,6	1547,81	118
mai/24	Fortaleza	2800,00	30,8	1969,94	100	nov/24	Fortaleza	4133,33	31,7	1829,23	120
jun/24	Fortaleza	4233,33	30,5	2251,36	94	dez/24	Fortaleza	3066,67	31,7	2392,07	120

Mês	CD:	Y	X1	X2	X3	Mês	CD:	Y	X1	X2	X3
jan/24	Itajaí	700,00	31,4	1706,99	35	jul/24	Itajaí	800,00	21,6	1167,94	29
fev/24	Itajaí	1500,00	31,6	1796,83	35	ago/24	Itajaí	700,00	23,5	808,57	30
mar/24	Itajaí	1266,67	30,4	1617,15	35	set/24	Itajaí	600,00	24,0	898,42	30
abr/24	Itajaí	900,00	27,9	1078,10	33	out/24	Itajaí	1166,67	26,1	1527,31	32
mai/24	Itajaí	533,33	24,3	988,26	31	nov/24	Itajaí	1366,67	28,3	1437,47	33
jun/24	Itajaí	600,00	22,1	1257,78	29	dez/24	Itajaí	833,33	30,5	1347,62	35

Mês	CD:	Y	X1	X2	X3	Mês	CD:	Y	X1	X2	X3
jan/24	Natal	2400,00	30,4	1637,89	68	jul/24	Natal	2233,33	28,4	1034,46	56
fev/24	Natal	2866,67	30,7	1724,10	70	ago/24	Natal	2333,33	28,6	862,05	57
mar/24	Natal	2666,67	30,7	1551,69	70	set/24	Natal	1500,00	29,0	775,84	60
abr/24	Natal	2766,67	30,3	1293,07	68	out/24	Natal	1766,67	29,7	948,25	64
mai/24	Natal	2233,33	29,8	1206,87	65	nov/24	Natal	2033,33	30,0	1120,66	66
jun/24	Natal	2933,33	28,8	1379,28	59	dez/24	Natal	2766,67	30,4	1465,48	68

Mês	CD:	Y	X1	X2	X3	Mês	CD:	Y	X1	X2	X3
jan/24	Porto Alegre	5700,00	31,0	2332,38	110	jul/24	Porto Alegre	3733,33	19,7	1595,84	88
fev/24	Porto Alegre	5066,67	30,5	2455,13	109	ago/24	Porto Alegre	3066,67	21,8	1104,81	92
mar/24	Porto Alegre	4566,67	29,2	2209,62	106	set/24	Porto Alegre	2033,33	22,8	1227,57	94
abr/24	Porto Alegre	2533,33	26,4	1473,08	101	out/24	Porto Alegre	1633,33	25,2	2086,86	98
mai/24	Porto Alegre	1800,00	22,6	1350,32	94	nov/24	Porto Alegre	2366,67	27,7	1964,11	103
jun/24	Porto Alegre	2700,00	20,3	1718,59	89	dez/24	Porto Alegre	2433,33	30,0	1841,35	108

Mês	CD:	Y	X1	X2	X3	Mês	CD:	Y	X1	X2	X3
jan/24	Rio de Janeiro	2333,33	30,5	1547,86	98	jul/24	Rio de Janeiro	2133,33	25,0	814,66	80
fev/24	Rio de Janeiro	5033,33	31,1	1629,33	100	ago/24	Rio de Janeiro	2300,00	26,1	1221,99	84
mar/24	Rio de Janeiro	2200,00	29,9	1466,39	96	set/24	Rio de Janeiro	2666,67	26,9	1059,06	86
abr/24	Rio de Janeiro	2166,67	28,4	1140,53	91	out/24	Rio de Janeiro	3866,67	28,1	1303,46	90
mai/24	Rio de Janeiro	2466,67	25,8	977,60	83	nov/24	Rio de Janeiro	2600,00	28,4	1221,99	91
jun/24	Rio de Janeiro	2500,00	25,2	896,13	81	dez/24	Rio de Janeiro	5266,67	29,9	1384,93	96

23 a 28 de novembro de 2025 - Vitória - ES

Mês	CD:	Y	X1	X2	X3	Mês	CD:	Y	X1	X2	X3
jan/24	Salvador	4200,00	31,0	3255,31	139	jul/24	Salvador	2700,00	26,6	1713,32	112
fev/24	Salvador	4800,00	31,1	3426,64	140	ago/24	Salvador	5133,33	26,7	1541,99	112
mar/24	Salvador	3733,33	30,9	3083,97	139	set/24	Salvador	3866,67	27,7	1884,65	119
abr/24	Salvador	4000,00	29,6	2912,64	130	out/24	Salvador	3300,00	29,1	2227,31	127
mai/24	Salvador	1700,00	28,1	2569,98	121	nov/24	Salvador	3000,00	29,8	2398,65	132
jun/24	Salvador	4800,00	27,1	2055,98	115	dez/24	Salvador	5966,67	30,6	2741,31	137

Mês	CD:	Y	X1	X2	X3	Mês	CD:	Y	X1	X2	X3
jan/24	São Paulo	6766,67	28,6	3942,35	136	jul/24	São Paulo	3633,33	22,9	2074,92	110
fev/24	São Paulo	6466,67	29,0	4149,84	137	ago/24	São Paulo	3266,67	24,5	3112,38	118
mar/24	São Paulo	4433,33	28,0	3734,86	133	set/24	São Paulo	3300,00	25,2	2697,40	121
abr/24	São Paulo	4500,00	26,6	2904,89	127	out/24	São Paulo	4133,33	26,5	3319,87	126
mai/24	São Paulo	4600,00	23,4	2489,91	113	nov/24	São Paulo	5000,00	26,9	3112,38	128
jun/24	São Paulo	4066,67	22,9	2282,41	111	dez/24	São Paulo	5333,33	28,3	3527,37	134

## ANÁLISES E DISCUSSÕES:

A série de dados utilizada na Regressão Linear Múltipla apresentada possui 108 cenários, oriundos da mensuração de 9 localidades distintas durante 12 meses completos. A equação obtida é expressa abaixo:

$$Y = -7188,05 + 138,87X_1 + 0X_2 + 21,84X_3 + 180,97X_4 + 134,11X_5 + 158,56X_6 + 129,44X_7 + 8,53X_8 \quad (2)$$

Como o valor do coeficiente angular ( $m_2$ ), que representa a influência do volume de carga distribuída sobre o consumo diário, resultou em zero, confirma-se a correlação anteriormente identificada entre a população flutuante e o volume de carga distribuída. Isso indica que a alternância da variável dependente ( $Y$ ) não é explicada pelo volume de carga distribuída ( $X_2$ ), de forma independente. Portanto, essa variável e seu respectivo coeficiente podem ser removidos da equação do modelo, sem prejuízo à capacidade explicativa da regressão.

A Tabela 4 exibe as principais estatísticas utilizadas na avaliação de modelos de regressão segundo Frost (2019). O R múltiplo representa a correlação linear entre os valores observados e os valores estimados pelo modelo, refletindo o grau de associação entre eles. O  $R^2$  indica a proporção da variabilidade da variável dependente que é explicada pelo conjunto de variáveis independentes. O  $R^2$  ajustado, por sua vez, corrige o  $R^2$  com base no número de variáveis independentes e no tamanho da amostra, evitando a superestimação da capacidade explicativa do modelo. Por fim, o Erro Padrão mede a dispersão dos resíduos em torno da linha de regressão, atuando como um indicador de precisão das estimativas geradas pelo modelo.

Tabela 4 – Principais estatísticas do modelo.

R Múltiplo	0,886	R <sup>2</sup> ajustado	0,768
R <sup>2</sup>	0,785	Erro Padrão	830,13

Para avaliar a significância estatística dos coeficientes angulares estimados foi utilizado o teste de  $-p$ . Este teste avalia se uma variável independente exerce influência significativa sobre a variável dependente. Para cada coeficiente, testa-se a hipótese de que seu valor é igual a zero, ou seja, de que não há efeito da variável independente. Valores de  $-p$  inferiores ao nível de significância adotado (recomendado 0,05) indicam que a hipótese é nula e pode ser rejeitada, sugerindo que o coeficiente é estatisticamente significativo. Por outro lado, valores de  $-p$  superiores a este limite indicam ausência

de evidência suficiente para afirmar que a variável contribui de forma relevante para o modelo (Tomaz, 2000). A Tabela 5 apresenta os valores de -p calculados para cada variável independente.

Tabela 5 – Teste de -p.

Coeficiente:	Valor -p
Interseção ( $a$ )	1,18704E-10
População Flutuante ( $m_1$ )	0,000145709
Temperatura Máxima Média ( $m_2$ )	8,86407E-05
Volume de Carga Distribuída ( $m_3$ )	0,551140538
Pressão Hidrostática Máxima ( $m_4$ )	7,5538E-08
Quantidade de Caixas Acopladas ( $m_5$ )	7,63169E-05
Quantidade de Válvulas de Descarga ( $m_6$ )	0,000856205
Quantidade de Chuveiros ( $m_7$ )	0,038024768
Quantidade de Torneiras ( $m_8$ )	0,835354858

Os indicadores apresentados na Tabela 4 evidenciam a boa qualidade do modelo de regressão. O R múltiplo de 0,8860 aponta forte correlação linear entre os valores observados e estimados, enquanto o R<sup>2</sup> de 0,7850 indica que 78,5% da variabilidade da variável dependente é explicada pelas variáveis independentes. O R<sup>2</sup> ajustado, com valor de 0,7676, reforça a robustez do modelo ao corrigir o R<sup>2</sup> em função do número de variáveis independentes, neste caso 8, e do tamanho da amostra, evitando a superestimação da capacidade explicativa. O Erro Padrão de 830,13, embora dependa da escala da variável resposta, reforça a adequação do ajuste, indicando baixa dispersão residual.

Verificou-se que os valores de -p para os coeficientes ( $m_2$ ) e ( $m_8$ ) ficaram acima de 0,05, indicando ausência de significância estatística. Esse resultado já era esperado para ( $m_2$ ), devido à correlação discutida anteriormente. Como novidade, o teste revelou que a quantidade de torneiras não exerce influência significativa sobre o consumo diário de água, enquanto as demais variáveis independentes demonstraram forte significância. Considerando que o modelo foi construído com dados de diversos CD's em períodos distintos, é razoável inferir que esse comportamento não se restringe a um contexto específico, mas representa uma característica geral do consumo de água em CDs. Literalmente, isso sugere que o uso de água por torneiras é pouco representativo quando comparado ao impacto de outros equipamentos consumidores ou à pressão hidrostática máxima.

A Tabela 6 exibe a comparação entre os valores de consumo médio diário medidos ao longo de 1 mês em cada CD e o valores previstos pelo modelo.

Tabela 6 – Análise de dados (consumo medido x previsto).

Localização	Mês Medição de Consumo	Consumo Diário Medido (l)	Consumo Diário Previsto (l)	Erro (l)	Erro (%)
Belo Horizonte - MG	Fevereiro	4172,80	4166,66	6,14	0,15%
Camaçari - BA	Abri	6927,27	6577,04	350,23	5,06%
Fortaleza - CE	Abri	3733,91	3439,25	294,66	7,89%
Itajaí - SC	Março	1552,73	1070,65	482,08	31,05%
Natal - RN	Abri	2790,10	2513,58	276,52	9,91%
Porto Alegre - RS	Abri	3051,80	3543,95	492,15	16,13%
Rio de Janeiro - RJ	Fevereiro	3018,25	3435,79	417,54	13,83%
Salvador - BA	Abri	4430,21	4152,77	277,44	6,26%
São Paulo - SP	Fevereiro	6359,69	5488,51	871,18	13,70%
					<b>RMSE (l) = 443,58</b>
					<b>MAPE (%) = 11,08</b>

Também são exibidos os valores de RMSE (Raiz do Erro Quadrático Médio), que indica o erro médio entre os valores previstos e observados, e de MAPE (Erro Percentual Absoluto Médio), que expressa esse erro em termos percentuais, facilitando a interpretação relativa do desempenho do modelo. Os resultados experimentais demonstram que o modelo preditivo é capaz de estimar com boa precisão o consumo diário de água em centros de distribuição logística, considerando as diferentes variáveis sazonais e variáveis de infraestrutura propostas. Os valores calculados de RMSE e MAPE foram, respectivamente, 443,58 litros e 11,08%. Observa-se que a porcentagem de erro tende a ser maior em cenários com menor consumo de água, uma vez que, nesses casos, pequenas imprecisões no modelo resultam em diferenças relativamente mais expressivas entre os valores medidos e os estimados.

Por fim, realizou-se a comparação entre a estimativa de consumo baseada em valores médios per capita extraídos da literatura e a previsão gerada pelo modelo. Embora sejam escassos os dados específicos para centros de distribuição, é comum a utilização de referências voltadas a setores industriais ou comerciais. Nesse contexto, o estudo de Tomaz (2000) apresenta uma abrangente compilação de dados nacionais e internacionais, na qual foi possível identificar apenas um valor de referência aplicável a galpões destinados ao transporte, armazenamento e distribuição de cargas, como é o caso dos CD's, correspondente a 238 litros por pessoa por dia. Contudo, esse dado foi publicado pela *American Water Works Association (AWWA)* em 1992 e não reflete com a realidade atual encontrada no Brasil, conforme exibido na Tabela 7.

Tabela 7 – Análise de dados (consumo medido x estimativa pelo per capita da AWWA).

Localização	Mês de Consumo	População Atendida	Consumo Diário Medido (l)	Previsto com Per Capita (l)	Erro	Erro (%)
Belo Horizonte - MG	Fevereiro	140	4172,80	31.920,00	27.747,20	664,95%
Camaçari - BA	Abril	130	6927,27	29.640,00	22.712,73	327,87%
Fortaleza - CE	Abril	96	3733,91	21.888,00	18.154,09	486,20%
Itajaí - SC	Março	35	1552,73	7.980,00	6.427,27	413,93%
Natal - RN	Abril	68	2790,10	15.504,00	12.713,90	455,68%
Porto Alegre - RS	Abril	101	3051,80	23.028,00	19.976,20	654,57%
Rio de Janeiro - RJ	Fevereiro	100	3018,25	22.800,00	19.781,75	655,40%
Salvador - BA	Abril	130	4430,21	29.640,00	25.209,79	569,04%
São Paulo - SP	Fevereiro	137	6359,69	31.236,00	24.876,31	391,16%
						<b>RMSE = 20.717,76</b>
						<b>MAPE = 87,28</b>

A análise de dados demonstra que as estimativas baseadas no consumo per capita indicado pela AWWA superestimam significativamente os valores medidos nos CD's investigados. Os erros percentuais ultrapassam 300% em todos os casos, com destaque para Belo Horizonte – MG, que atingiu 664,95%. O RMSE de 20.717,76 litros e o MAPE de 87,28% evidenciam a baixa acurácia do método, indicando que o consumo per capita da AWWA não é adequado para estimativas de consumo diário em CDs no contexto brasileiro.

Para fins comparativos, também foi verificado o consumo de 55 litros por pessoa por dia, recomendado por muitas concessionárias de abastecimento de água que atuam no Brasil, para edificações industriais com uso de água exclusivo para fins sanitários. Embora esta recomendação apresente desempenho superior à estimativa da AWWA, os erros ainda são significativos. O RMSE foi reduzido para 2.031,03 litros e o MAPE para 35,47%, indicando melhora considerável na acurácia geral. No entanto, persistem desvios expressivos. Esses resultados mostram que, apesar de mais aderente à realidade brasileira, essa abordagem ainda carece de ajustes específicos para a previsão precisa do consumo diário em centros de distribuição.

## CONCLUSÕES:

Os resultados experimentais demonstram que o modelo preditivo desenvolvido é capaz de estimar com boa precisão o consumo diário de água em centros de distribuição logística ao considerar conjuntamente variáveis sazonais e de infraestrutura. Com valores de RMSE e MAPE de 443,58 litros e 11,08%, respectivamente, o modelo apresentou desempenho satisfatório na maioria dos cenários avaliados, evidenciando seu potencial como ferramenta de apoio ao planejamento e dimensionamento hídrico em empreendimentos logísticos. Notou-se que a porcentagem de erro tende a ser maior em instalações com menor consumo absoluto, dado que, nesses casos, pequenas variações resultam em desvios percentualmente mais relevantes, um comportamento esperado em modelos baseados em dados empíricos.

Os testes estatísticos confirmaram a relevância de variáveis como população, temperatura, pressão disponível e número de dispositivos sanitários, enquanto indicaram que o volume de carga distribuída e o número de torneiras têm baixa influência no consumo. Essas observações revelam aspectos importantes do uso de água em CDs e reforçam a inadequação de estimativas per capita genéricas, que superestimaram em até 665% os valores medidos. Assim, o modelo proposto representa um avanço, por ser de fácil aplicação, alinhado à realidade brasileira e mais sensível às características operacionais dos CDs, contribuindo para o uso racional da água e um planejamento mais eficiente de sistemas de abastecimento em um setor estratégico em crescimento.

## REFERÊNCIAS:

- BALLOU, Ronald H.; SRIVASTAVA, Samir K. **Business logistics/supply chain management: planning, organizing, and controlling the supply chain.** [S.l.]: Pearson Education India, 2007.
- DE FREITAS, Thiago Fernandes. O CRESCIMENTO DO E-COMMERCE E O IMPACTO DA LOGÍSTICA INTELIGENTE. **LUMEN ET VIRTUS**, v. 16, n. 44, 2025.
- FEIL, Alexandre André; HAETINGER, Claus. Previsão de consumo de água via modelagem matemática de sistema de abastecimento de água. **Revista DAE**, v. 195, n. 1, p. 32–46, 2014.
- SCHLEICH, Joachim; HILLENBRAND, Thomas. Determinants of residential water demand in Germany. **Ecological economics**, v. 68, n. 6, p. 1756–1769, 2009.
- SILVA, W. T. P.; CAMPOS, M. M.; SANTOS, A. A. Previsão consumo diário de água: um estudo de caso de Nobres (Brasil). **Ingeniería del agua**, v. 20, n. 2, p. 73, 29 abr. 2016.
- SLAVÍKOVÁ, Lenka *et al.* Impacts of climate variables on residential water consumption in the Czech Republic. **Water Resources Management**, v. 27, p. 365–379, 2013.
- TOMAZ, Plínio. **Previsão de Consumo de Água**. São Paulo: Navegar, 2000.
- TRAUTWEIN JR, Breno. Avaliação de métodos para previsão de consumo de água para curtíssimo prazo: um estudo de caso em empresa de saneamento. 2004.
- VIANNA, Vinicius; DEPEXE, Marcelo Dacul. Modelagem de dados por regressão múltipla para previsão do consumo de água em Umuarama. **Exacta**, v. 11, n. 1, p. 33–46, 28 maio 2013.
- YASAR, Abdulkadir; BILGILI, Mehmet; SIMSEK, Erdogan. Water demand forecasting based on stepwise multiple nonlinear regression analysis. **Arabian journal for science and engineering**, v. 37, p. 2333–2341, 2012.
- ZHOU, Shuang Lin *et al.* Forecasting daily urban water demand: a case study of Melbourne. **Journal of hydrology**, v. 236, n. 3–4, p. 153–164, 2000.