

INVESTIGAÇÃO DE FATORES RELACIONADOS AO CONSUMO DE ÁGUA EM EDIFÍCIOS PÚBLICOS DE ESCRITÓRIOS EM JOINVILLE

, Ihor Parcias Olijnyk¹, Andreza Kalbush², Elisa Henning³

¹ Acadêmico do Curso de Engenharia Civil - CCT- PROBIC/UDESC.

² Orientadora, Departamento de Engenharia Civil, CCT- andreza.kalbush@udesc.br

³ Professora Participante do Departamento de Matemática – CCT.

Palavras-chave: modelos, regressão múltipla, consumo de água.

Na gestão do uso de água é fundamental obter-se informações confiáveis sobre como se dá o seu consumo a fim de permitir tomadas de decisões precisas no desenvolvimento de projetos que busquem eficiência e proporcionem economia e segurança hídrica ao longo prazo. Este projeto teve como objetivo construir um modelo matemático que auxiliasse no entendimento da influência das características físicas, sociodemográficas, operativas e dos sistemas prediais hidrossanitários no consumo de água dos edifícios de escritórios públicos administrativos em Joinville. Os edifícios selecionados e analisados tinham, obrigatoriamente, medidores individuais do consumo de água e características exclusivamente administrativas, evitando-se escritórios que possuíssem atributos demasiadamente peculiares e ocasionassem distorção nos dados. Dessa forma, edifícios públicos da área de saúde, educação e segurança não entraram na amostra para a criação do modelo. O modelo matemático tem o formato de regressão linear múltipla, o qual busca associar variáveis explicativas, relacionadas às características dos escritórios e de seus usuários, com uma variável independente resposta: o consumo de água. Os dados do consumo, compreendidos de Janeiro de 2016 a Dezembro de 2017, foram adquiridos diretamente com a Companhia de Águas de Joinville (CAJ), a exploradora direta dos serviços públicos de abastecimento de água da cidade. Foram avaliados dois modelos, um para o consumo bruto, em metros cúbicos por mês - CB ($m^3/mês$), e outro para o Indicador de Consumo - IC, dado em litros por funcionário por dia. As informações coletadas nos escritórios separam-se em 3 grupos principais: (i) características físicas dos edifícios, incluindo o tamanho do terreno, a área construída, a idade dos edifícios; (ii) características sociodemográficas e operacionais, tais como o número de trabalhadores do sexo feminino e masculino, a jornada de trabalho, o horário de funcionamento, o número de atendimentos, o tempo médio gasto em um atendimento; e (iii) as características de consumo de água e dos sistemas prediais hidrossanitários: número de cozinhas, uso de galões de água mineral, número de banheiros. As informações foram coletadas por meio de questionários, tanto para as características físicas quanto para as características sociodemográficas e operacionais. Os dados relacionados a características de consumo de água e dos sistemas prediais hidrossanitários foram coletados *in-loco*. As informações obtidas foram compiladas e tratadas com auxílio do software livre R Studio e de seus pacotes de análise estatísticas, entre os quais destacam-se: car, leaps, gvlma, MASS e olsrr. Com o intuito de proceder análises paramétricas, buscou-se encontrar indícios de normalidade na distribuição dos dados obtidos e, caso contrário, alguma transformação que tornasse-os dessa forma. Consequentemente observou-se que as distribuições dos dados mostravam normalidade quando se aplicava transformação logarítmica natural na variável dependente CB ($\ln CB$) e também nas

variáveis independentes. Como exceção, alguns dados transformados em percentuais, tais como o percentual de torneiras com arejador e o percentual de bacias sanitárias com acionamento duplo, não atenderam aos critérios de normalidade. Isso pode ser explicado pelo fato de que para esses tipos de variáveis geralmente haviam valores extremos: em geral a maioria ou quase nenhum dos componentes tinham a característica considerada. Para definir a importância inicial de cada variável explicativa o exame de correlação bivariada de *Pearson* foi aplicado entre todas as variáveis preditivas e as variáveis de desfecho LnCB e IC. As que indicaram correlação significativa foram então avaliadas com metodologias dentro do R que utilizam regras de decisões distintas para conferir a importância das variáveis na criação dos modelos: *Stepwise* (critério AIC), *Leaps* (critério do R-Ajustado), *Mallows* (critério estatístico de *Mallows*) e multicolineariedade (critério de correlação bivariadas de *Pearsons* entre as variáveis independentes). Assim, quatro modelos com variáveis independentes distintas foram obtidos e analisados quanto à distribuição normal de erros, independência dos erros (estatística de autocorreção de *Durbin Watson*), homocedasticidade (variância de erro constante - NCV), multicolineariedade (fator de inflação da variância - VIF), apresentação de *outliers* (pontos fora da curva) e de dados demasiadamente influentes (distância de *Cook* maior do que 1). Foram, assim, obtidos dois modelos: um para consumo bruto e outro para o indicador de consumo de água. Observou-se que o tamanho do terreno (LnLA) e o número de torneiras (LnTap) tem influência positiva no consumo, enquanto a presença de bacias sanitárias de 6 litros por fluxo (PerLF) e com acionamento duplo (PerDF) tem correlação negativa, ou seja, fazem com que o consumo diminua.

Fig. 1 Resumo do modelo IC gerado pelo software R

Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	22.47199	13.29757	1.690	0.1168
LnLA	3.43261	1.49850	2.291	0.0409 *
PerLF	-0.16192	0.06542	-2.475	0.0292 *
PerDF	-0.18732	0.07558	-2.478	0.0290 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
Residual standard error: 9.969 on 12 degrees of freedom				
Multiple R-squared: 0.6851, Adjusted R-squared: 0.6063				
F-statistic: 8.701 on 3 and 12 DF, p-value: 0.002445				

Fig. 2 Resumo do modelo LnCB gerado pelo software R

Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.715872	0.588856	-1.216	0.24955
LnLA	0.121449	0.060315	2.014	0.06917 .
LnTap	1.486679	0.209689	7.090	2.02e-05 ***
PerLF	-0.010850	0.002496	-4.348	0.00116 **
PerDF	-0.006942	0.002833	-2.450	0.03223 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
Residual standard error: 0.3691 on 11 degrees of freedom				
Multiple R-squared: 0.9379, Adjusted R-squared: 0.9153				
F-statistic: 41.52 on 4 and 11 DF, p-value: 1.42e-06				