

UNIVERSIDADE FEDERAL DE MATO GROSSO
INSTITUTO DE CIÊNCIAS EXATAS E DA TERRA
BACHARELADO EM ESTATÍSTICA

ILIAS DE MUSIS

**EFICIÊNCIA ENERGÉTICA DE EDIFÍCIOS RESIDENCIAIS POR ANÁLISE DE
CORRELAÇÃO CANÔNICA**

CUIABÁ – MT
2024

ILIAS DE MUSIS

**EFICIÊNCIA ENERGÉTICA DE EDIFÍCIOS RESIDENCIAIS POR ANÁLISE DE
CORRELAÇÃO CANÔNICA**

Trabalho de conclusão de curso
apresentado à banca examinadora do Curso
de Bacharelado em Estatística da
Universidade Federal de Mato Grosso
como requisito parcial para a obtenção do
título de Bacharel em Estatística.

Orientador: Prof. Dr. Kuang Hongyu

CUIABÁ – MT

2024

ILIAS DE MUSIS

**EFICIÊNCIA ENERGÉTICA DE EDIFÍCIOS RESIDENCIAIS POR ANÁLISE DE
CORRELAÇÃO CANÔNICA**

Trabalho de conclusão de curso
apresentado à banca examinadora do Curso
de Bacharelado em Estatística da
Universidade Federal de Mato Grosso
como requisito parcial para a obtenção do
título de Bacharel em Estatística.

APROVADA em 04 de abril de 2024.

BANCA EXAMINADORA

Prof. Dr. Kuang Hongyu

Orientador

Prof. Dra. Édila Cristina de Souza

Examinador

Prof. Dr. Rômulo Môra

Examinador

CUIABÁ – MT

2024

RESUMO

Nos contextos global e nacional, os edifícios são importantes fontes de emissões de CO₂ e consumo de energia. O setor de edificações, com destaque nos edifícios residenciais, é um dos principais consumidores de energia elétrica no Brasil. Diante desse desafio, estratégias como eficiência energética (EE) destacam-se como soluções. No contexto dos edifícios, a EE refere-se a práticas destinadas a reduzir o consumo de energia, começando desde a fase de construção e incluindo aspectos do design arquitetônico e conforto térmico. Este estudo buscou investigar a relação entre características físicas relacionadas ao formato do edifício (compactação relativa, área de superfície, área da parede, área do telhado, altura total, orientação, área envidraçada e distribuição da área envidraçada) de edifícios residenciais e sua eficiência energética, caracterizada por cargas térmicas (carga de aquecimento e resfriamento), por meio de análise de correlação canônica (ACC). Trata-se de uma técnica de estatística multivariada que visa quantificar a relação entre dois conjuntos de variáveis, criando combinações lineares que maximizam a correlação entre eles. Com ela, é possível verificar quais variáveis possuem maior influência na relação entre os grupos, entre outras informações. A análise da primeira função canônica revelou uma correlação canônica de 0,95, indicando uma forte relação linear entre as características dos edifícios e as cargas térmicas. Além disso, explicou 89,05% da variância das cargas térmicas e 33,12% da variância das características dos edifícios. Entre os parâmetros analisados, área do telhado e altura total destacaram-se como os mais influentes sobre as cargas térmicas. Uma maior altura total se mostrou associada a maiores cargas térmicas, enquanto maiores áreas de telhado mostraram uma relação inversa com as cargas térmicas. Esses achados sugerem que menores alturas e telhados mais extensos estão associados a um melhor desempenho energético em edificações residenciais.

Palavras-chave: Cargas térmicas; conforto térmico; formato do edifício; estatística multivariada.

ABSTRACT

In both global and national contexts, buildings are important sources of CO₂ emissions and energy consumption. The building sector, with a focus on residential buildings, is one of the main consumers of electricity in Brazil. Faced with this challenge, strategies such as energy efficiency (EE) stand out as solutions. In the context of buildings, EE refers to practices aimed at reducing energy consumption, starting from the construction phase and including aspects of architectural design and thermal comfort. This study aimed to investigate the relationship between physical characteristics related to building form (relative compactness, surface area, wall area, roof area, total height, orientation, glazed area, distribution of glazed area) of residential buildings and their energy efficiency, characterized by thermal loads (heating and cooling loads), through canonical correlation analysis (CCA). CCA is a multivariate statistical technique that aims to quantify the relationship between two sets of variables by creating linear combinations that maximize the correlation between them. With it, it is possible to verify which variables have the greatest influence on the relationship between the groups, among other information. The analysis of the first canonical function revealed a canonical correlation of 0.95, indicating a strong linear relationship between building traits and thermal loads. Moreover, it explained 89.05% of the variance of the thermal loads and 33.12% of the variance of the building characteristics. Among the parameters analyzed, roof area and total height stood out as the most influential on thermal loads. A greater total height was associated with higher thermal loads, while larger roof area showed an inverse relationship to thermal loads. These findings suggest that lower heights and more extensive roofs are associated with better energy performance in residential buildings.

Keywords: Thermal loads; thermal comfort; building form; multivariate statistics.

LISTA DE FIGURAS

Figura 1 – Diagrama de dispersão e modelo linear ajustado para a primeira função canônica.....	32
Figura 2 – Gráfico <i>biplot</i> da distribuição espacial das variáveis em relação às funções canônicas 1 e 2	34

LISTA DE TABELAS

Tabela 1 – Diretrizes gerais para a interpretação de coeficientes de correlação	21
Tabela 2 – Descrição das variáveis utilizadas no estudo.....	26
Tabela 3 – Variações de distribuição de área envidraçada.....	27
Tabela 4 – Correlações canônicas, porcentagens de variância explicada e teste estatístico Lambda de Wilks para as funções canônicas.....	30
Tabela 5 – Análise de redundância respectiva a cada conjunto de variáveis e função canônica	31
Tabela 6 – Pesos canônicos padronizados, cargas canônicas e cargas canônicas cruzadas referentes à primeira função canônica	33
Tabela 7 – Cargas canônicas quadradas, cargas canônicas cruzadas quadradas e índices de redundância referentes à primeira função canônica.....	35

LISTA DE ABREVIATURAS E SIGLAS

ACC – Análise de Correlação Canônica

ACP – Análise de Componentes Principais

ASHRAE – *American Society of Heating, Refrigerating and Air-Conditioning Engineers* (Associação Americana de Engenheiros de Aquecimento, Refrigeração e Ar-Condicionado)

AVAC – Aquecimento, Ventilação e Ar-Condicionado

CL – *Cooling load* (carga de resfriamento)

EE – Eficiência Energética

HL – *Heating load* (carga de aquecimento)

MANOVA – *Multivariate Analysis of Variance* (Análise Multivariada de Variância)

SUMÁRIO

1	INTRODUÇÃO	9
1.2	Objetivo Geral	11
1.3	Objetivo Específicos	11
2	REFERENCIAL TEÓRICO	12
2.1	Eficiência Energética.....	12
2.2	Cargas Térmicas.....	13
2.3	Análise Multivariada.....	13
2.4	Análise de Correlação Canônica	15
2.4.1	SUPOSIÇÕES	16
2.4.2	CÁLCULO DAS FUNÇÕES E CORRELAÇÕES CANÔNICAS	16
2.4.3	CRITÉRIOS PARA SELEÇÃO DAS FUNÇÕES CANÔNICAS	18
2.4.3.1	Testes de Significância.....	19
2.4.3.2	Magnitude da Correlação Canônica	20
2.4.3.3	Análise de Redundância	21
2.4.4	INTERPRETAÇÃO	23
2.4.4.1	Pesos canônicos padronizados.....	23
2.4.4.2	Cargas canônicas	23
2.4.4.3	Cargas canônicas cruzadas	24
3	MATERIAL E MÉTODOS.....	26
3.1	Características do conjunto de dados.....	26
3.2	Análise de Correlação Canônica	28
4	RESULTADOS E DISCUSSÕES	30
5	CONCLUSÃO	37
	REFERÊNCIAS	38

1 INTRODUÇÃO

O consumo energético é um tema de cada vez mais destaque devido à crescente demanda de energia elétrica e às consequências da não otimização dos sistemas envolvidos. Uma das soluções investigadas é a adoção de medidas de Eficiência Energética (EE). De acordo com o Plano Nacional de Eficiência Energética (Brasil, 2011), a EE consiste em uma série de ações com a finalidade da redução da energia necessária para atender as demandas da sociedade em termos de serviços de energia. Em outras palavras, seu objetivo é minimizar o consumo energético sem comprometer a produção.

Em 2021, as edificações foram responsáveis por cerca de 34% da demanda global de energia e, com a produção de materiais de construção, mais de 37% das emissões de CO₂ (United Nations Environment Programme, 2022). No Brasil, conforme dados divulgados no Balanço Energético Nacional de 2022, ano base 2021 (Empresa de Pesquisa Energética, 2022), os setores industrial, comercial e residencial, em conjunto, consomem 79% da energia elétrica disponibilizada no país. Esses dados destacam o setor de edificações como um dos principais consumidores de energia elétrica, evidenciando a necessidade de identificar e implementar soluções mais eficientes em termos energéticos.

A implementação de medidas de eficiência energética pode gerar economia de recursos financeiros, além de contribuir para a redução das emissões de gases de efeito estufa e a preservação do meio ambiente (Li; Li; Wang, 2022). O Programa Nacional de Eficiência Energética em Edificações estima que a adoção de medidas de eficiência energética nos sistemas de iluminação, ar-condicionado, bem como intervenções arquitetônicas reduziriam em 30% o consumo de energia elétrica em edificações existentes, e em até 50% em edificações novas, reforçando o potencial dessa abordagem (Programa Brasileiro de Etiquetagem de Edificações, 2020).

Conforme descrito pela Associação Americana de Engenheiros de Aquecimento, Refrigeração e Ar-Condicionado (ASHRAE, do inglês *American Society of Heating, Refrigerating and Air-Conditioning Engineers*), as cargas térmicas são definidas como a quantidade de calor que deve ser retirada (resfriamento) ou fornecida (aquecimento) a um sistema para a manutenção de condições ideais de temperatura do ambiente. Para isso, devem ser consideradas as dimensões e usos do espaço, entre outros fatores.

A ASHRAE (2021) destaca o cálculo das cargas de aquecimento e resfriamento como base para a determinação de maior parte dos sistemas de aquecimento e refrigeração, influenciando diretamente no dimensionamento de componentes como tubulações, ventiladores e difusores. Os requerimentos de energia de sistemas de aquecimento, ventilação e ar condicionado apresentam impacto direto nos custos operacionais de um edifício e, de forma mais ampla, no meio ambiente.

Além disso, a evolução tecnológica permitiu a incorporação de ferramentas de simulações computacionais avançadas no processo de cálculo das cargas térmicas, tornando seu uso fundamental para o design de edifícios energeticamente eficientes. Estas facilitam a realização de cálculos

complexos, considerando fatores como incidência solar e número de habitantes, e permitem a identificação, ajuste e experimentação de parâmetros do projeto antes de sua construção, evitando custos desnecessários. Além disso, oferecem resultados que refletem com considerável precisão os resultados reais (Yezioro; Dong; Leite, 2008).

Por intermédio da análise multivariada, pesquisadores podem discernir com maior precisão quais fatores exercem maior impacto na economia de energia, bem como empregar técnicas avançadas de modelagem e previsão. Diversos estudos exploram essa abordagem: através da utilização de diversas técnicas multivariadas. Chen, Wang e Steemers (2013) determinaram o efeito de características domiciliares e comportamento dos ocupantes em Hangzhou (China) sobre o consumo de energia, destacando idade e renda dos moradores como fatores significativos.

Hygh *et al.* (2012), por sua vez, buscaram apresentar uma nova abordagem de modelagem, incorporando simulações computacionais e modelos de regressão linear multivariados, com o objetivo de quantificar o desempenho energético de edifícios em seus estágios iniciais de desenvolvimento. Com base nas informações fornecidas, é observável a eficácia da estatística multivariada na análise de aspectos da eficiência energética e conforto térmico em edifícios. No entanto, para uma avaliação mais abrangente da relação de múltiplos fatores que afetam a eficiência energética, pode ser útil um método que possa analisar as relações entre dois conjuntos de características.

Conforme apontado por Tsanas e Xifara (2012), pesquisas sobre a eficiência energética de edificações frequentemente recorrem a pressupostos matemáticos simplificados e inadequados para aplicações mais complexas. Mesmo ao recorrer ao uso de técnicas mais heterodoxas, como o *machine learning*, é comum a falha em realizar uma análise detalhada, omitindo a identificação das variáveis mais significativas para o problema abordado, por exemplo.

Por esses motivos, a escolha da análise a ser empregada no presente estudo recaiu sobre a Análise de Correlação Canônica (ACC), uma técnica multivariada focada na quantificação da relação entre dois grupos de variáveis. Consiste na realização de combinações lineares de forma a maximizar a correlação entre os grupos e, assim, identificar as variáveis mais influentes nesse relacionamento, entre outras informações úteis. Ao analisar todas as variáveis de cada grupo simultaneamente, representa interações entre elas que poderiam ser ignoradas em análises univariadas ou multivariadas que não consideram esses aspectos (Hair Jr. *et al.*, 2018; Thompson, 2005).

Este estudo explora a relação entre as cargas térmicas de aquecimento e resfriamento e um conjunto de atributos da envoltória de edificações, incluindo transparência, orientação, compactação e morfologia. Para isso, será realizada uma análise multivariada por meio da ACC sobre dados disponibilizados no *UCI Machine Learning Repository* (Kelly; Longjohn; Nottingham, 2023; Tsanas; Xifara, 2012), referentes a simulações no *software Ecotect* das variáveis descritas.

1.2 Objetivo Geral

Investigar as relações entre as características físicas de edifícios residenciais e sua eficiência energética por meio da análise de correlação canônica.

1.3 Objetivo Específicos

- Observar as relações das variáveis em seus respectivos grupos;
- Determinar a magnitude com a qual os atributos físicos de edifícios e suas cargas térmicas são correlacionados;
- Analisar a relação entre os grupos de variáveis, de forma a determinar quais características físicas de edifícios estão mais fortemente relacionadas à eficiência energética.

2 REFERENCIAL TEÓRICO

2.1 Eficiência Energética

A Eficiência Energética (EE) é definida pela otimização da relação entre a quantidade de energia empregada e a quantidade efetivamente utilizada em uma atividade. A ideia é obter o melhor resultado possível com o mínimo de energia consumida. Pode ser definida pela seguinte equação:

$$EE = \frac{\text{Energia utilizada}}{\text{Energia fornecida}},$$

com valores mais próximos de 1 implicando maior eficiência energética.

Enquanto alguns autores utilizam os termos “eficiência energética” e “conservação de energia” como sinônimos, outros fazem uma distinção entre eles. Schwarz (2018) afirma que a conservação consiste em economizar energia ao utilizar menos bens e serviços que consumam energia, enquanto a eficiência energética consiste em utilizar menos energia para obter um nível fixo de bens e serviços.

Para ilustrar esses conceitos, o autor exemplifica a conservação de energia como diminuir a temperatura de um termostato durante o inverno e colocar roupas mais quentes, de forma a economizar energia. Um exemplo de eficiência energética seria isolar termicamente a casa, de forma que se gaste menos energia para aquecê-la.

A adoção de medidas de eficiência energética oferece diversos benefícios, tanto para os indivíduos quanto para a sociedade como um todo. Esses benefícios incluem a redução de impactos ao meio ambiente, o estímulo de inovações tecnológicas e a promoção de um setor intensivo em geração de empregos qualificados. A EE se destaca como alternativa na redução de dependência de fontes de energia não renováveis, propiciando um uso mais eficiente dos recursos do planeta e reduzindo a emissão de gases de efeito estufa (Empresa de Pesquisa Energética, 2020; International Energy Agency, 2019).

No contexto de edificações, a EE se refere ao conjunto de práticas aplicadas para reduzir o consumo de energia em construções, mantendo o conforto e funcionalidade do espaço. Isso envolve aspectos como projeto arquitetônico, isolamento térmico, sistemas de ventilação, entre outros (Pérez-Lombard; Ortiz; Pout, 2008).

Estudos indicam que o sistema de condicionamento térmico é responsável por grande parte do consumo de energia em edificações, representando 59% do consumo elétrico no setor residencial (Baltar; Kaehler; Pereira, 2005). Isso evidencia a necessidade de se incorporar considerações sobre o conforto térmico no planejamento e construção de edificações, visando a eficiência energética.

2.2 Cargas Térmicas

As cargas térmicas se referem à quantidade de calor que deve ser adicionada ou removida de um ambiente para manter a temperatura desejada pelos usuários do recinto. Os ganhos de calor podem ser provenientes de fontes externas ou internas ao espaço condicionado (Silva, 2010).

Quantificando a dinâmica energética de determinado espaço, as cargas térmicas são parâmetros fundamentais no planejamento e dimensionamento de sistemas de Aquecimento, Ventilação e Ar-Condicionado (AVAC). Estes sistemas são projetados de forma a neutralizar as cargas calculadas, garantindo que as condições ambientais internas permaneçam dentro de limites predefinidos de temperatura e umidade. O cálculo de cargas térmicas e projeto adequado dos sistemas AVAC possuem impacto direto no conforto térmico, preservação de materiais, custo operacional e consumo de energia (ASHRAE, 2021).

Os cálculos de carga térmica de pico buscam determinar a taxa máxima de transferência de energia que possa ser necessária em qualquer ponto no tempo. Embora os processos de resfriamento e aquecimento sejam teoricamente idênticos, são utilizadas abordagens diferentes para calcular suas respectivas cargas (Burdick, 2011).

A carga de resfriamento (CL, do inglês *cooling load*) se refere à quantidade de calor que deve ser retirada de um ambiente para alcançar as condições ideais de temperatura. Para determiná-la, deve ser considerado o ganho de calor por fontes externas ou internas, levando em consideração fatores como superfícies opacas, superfícies transparentes, ventilação e ocupação do ambiente a ser resfriado (ASHRAE, 2021; Rutkowski, 2014).

Por sua vez, a carga de aquecimento (HL, do inglês *heating load*) diz respeito à quantidade de calor necessária para compensar as perdas de calor para o ambiente externo, geralmente durante os meses mais frios. Seu cálculo pode ser simplificado e envolve estimar o máximo possível de calor perdido em um ambiente a ser aquecido, ignorando ganhos de calor internos ou externos (ASHRAE, 2021; Matos, 2010). Conforme Wang, Yan e Xiao (2012), o uso de ferramentas de simulação é um dos métodos mais eficazes para a quantificação de energia.

Existem diversas variáveis que afetam as cargas térmicas de um edifício, tais como a área de superfície, a altura e a localização do prédio. Os autores Pérez-Lombard, Ortiz e Pout (2008) afirmam que unidades habitacionais menores demandam menos energia, tendo em vista a menor área a ser condicionada e superfície de transferência, bem como menor número de ocupantes.

2.3 Análise Multivariada

Quando um pesquisador conduz uma amostragem ou experimento, geralmente há interesse em mais de uma variável. Isso ocorre pois as variáveis medidas geralmente estão relacionadas entre si, e explorar essas relações pode levar a análises mais completas (Ferreira, 2008). A análise multivariada

consiste em um conjunto de técnicas estatísticas utilizadas para estudar múltiplas variáveis simultaneamente.

O avanço tecnológico é um dos principais fatores impulsionando a popularidade desse tipo de análise nas últimas décadas. A crescente disponibilidade de técnicas computacionais mais avançadas supera muitas das dificuldades técnicas que a análise multivariada apresentava, como a necessidade de processamento de grandes quantidades de dados e a complexidade dos cálculos envolvidos.

Anderson (2003) afirma que os problemas estatísticos de populações multivariadas são análogos àqueles surgidos em populações univariadas, com os métodos adequados para lidar com esses problemas sendo similarmente relacionados. O autor ilustra esse conceito exemplificando que, em situações nas quais, no caso univariado, se desejaria testar a hipótese de que a média de uma variável é zero, na análise multivariada seria testada a hipótese do vetor de médias de múltiplas variáveis ser o vetor nulo.

Explorando essa questão, Huberty e Morris (1989) apresentam uma comparação entre a análise multivariada e a aplicação de várias análises univariadas. Os autores pontuam que o uso de várias análises univariadas leva a um aumento do risco de erro do tipo I, resultando na rejeição indevida da hipótese nula. Nesse sentido, a análise multivariada fornece uma solução mais integrada, capaz de oferecer uma visão mais completa das relações entre as variáveis, enquanto exerce maior controle sobre o risco de tipo I. Com isso em mente, os autores destacam que ambas as abordagens buscam responder perguntas de pesquisa distintas, e são válidas em contextos diferentes.

Rencher (2002), por sua vez, destaca que a estatística multivariada é uma abordagem fundamental para dados complexos. A autor argumenta que, em frente a um banco de dados multivariado, não é mais justificável que sejam utilizados métodos de análise exclusivamente univariados, dada a abundância de técnicas e acessibilidade ao poder computacional para a execução facilitada de análises multivariadas.

As técnicas de análise multivariada abrangem uma ampla gama de métodos, cada um com objetivos específicos e adequados para diferentes tipos de perguntas de pesquisa. De acordo com Mingoti (2005), esses métodos são utilizados principalmente para simplificação estrutural e para realização de inferências estatísticas. Johnson e Wichern (2007) dividem os propósitos da análise de forma similar, incluindo agrupamento de dados, investigação de dependência entre variáveis, redução de dados e testes de hipóteses. As técnicas também podem ser combinadas para uma análise mais robusta.

Para os propósitos do presente trabalho, foi escolhido o método de Análise de Correlação Canônica. Essa escolha se deu pela sua capacidade de analisar dois conjuntos de variáveis, permitindo identificar relações significativas entre eles, o que é de interesse para o objetivo do trabalho. Seu detalhamento será fornecido no próximo tópico.

2.4 Análise de Correlação Canônica

A Análise de Correlação Canônica (ACC) é uma técnica de análise multivariada que busca identificar e compreender as relações entre dois conjuntos de variáveis. Proposta por Hotelling (1936), é frequentemente utilizada em análises multivariadas, tendo aplicações em diferentes áreas de estudo. Pode ser descrita como uma extensão da regressão múltipla e é útil quando se têm várias variáveis dependentes.

Tabachnick e Fidell (2013) descrevem a ACC como uma das técnicas mais gerais, pois engloba vários outros procedimentos comumente utilizados. As técnicas de regressão múltipla, análise discriminante e MANOVA (*Multivariate Analysis of Variance*, ou Análise Multivariada de Variância) podem ser vistos como casos especiais dela. Apesar de pouco utilizada, a flexibilidade da ACC pode ser aproveitada para explorar relações entre conjuntos de variáveis em diversas áreas.

Destacando a versatilidade da ACC, Everitt (2005) apresenta várias aplicações práticas. Na psicologia, um investigador pode desejar medir um conjunto de variáveis de aptidão e outro de variáveis de sucesso em uma amostra de alunos. No marketing, um exemplo similar poderia envolver um conjunto de índices de preços e um conjunto de índices de previsão. Em muitas situações, um desses conjuntos contém múltiplas variáveis dependentes e o outro, múltiplas variáveis independentes ou explicativas. A ACC oferece um método para a previsão dessas variáveis dependentes a partir de duas ou mais variáveis independentes.

Enquanto a análise de regressão múltipla busca uma combinação linear de um vetor de variáveis independentes x com maior correlação com uma única variável dependente y , a ACC relaciona x com um conjunto de variáveis y . Outra técnica multivariada que pode ser comparada à ACC é a análise de componentes principais (ACP). Esta também é utilizada para reduzir a dimensionalidade de um conjunto de dados por meio de combinações lineares, mas enquanto a ACP analisa relacionamentos encontrados dentro de um único conjunto de variáveis, a ACC foca nos relacionamentos entre conjuntos (Mardia; Kent; Bibby, 1979).

Apesar do uso dos termos “dependente” ou “independente” ao descrever variáveis poder ser útil para o tópico de pesquisa, é importante destacar que não há necessariamente relação de causa e efeito entre elas e a correlação canônica não faz essa distinção (Hair Jr. *et al.*, 2018). Para esse tipo de análise, são apenas dois grupos de variáveis cujas associações buscamos investigar.

A ideia central da ACC é determinar combinações lineares para cada vetor de variáveis sendo estudado de forma a maximizar a correlação entre eles. Em outras palavras, buscam-se pares de variáveis canônicas que tenham a correlação mais alta possível. Ao encontrar as combinações lineares mais correlacionadas entre os conjuntos de dados, a correlação canônica ajuda a identificar as relações mais importantes entre as variáveis dos grupos estudados.

Ferreira (2008) aponta que, ao estabelecer uma relação entre dois vetores de p e q variáveis respectivamente, é possível estimar correlações ou covariâncias individuais para cada par de variáveis

e analisá-los. No entanto, quando esses vetores contêm dimensões muito grandes, essa análise se torna inviável. O aspecto de maximização de correlações na ACC é uma tentativa de concentrar uma relação de alta dimensão entre dois conjuntos de variáveis em poucas variáveis canônicas (Johnson; Wichern, 2007).

2.4.1 SUPOSIÇÕES

A análise de correlação canônica, assim como outras técnicas estatísticas, requer algumas suposições para sua aplicação efetiva: premissas para a realização da análise e melhor interpretação dos resultados. Essas suposições incluem, entre outras, linearidade, multicolinearidade e homocedasticidade, a serem detalhadas a seguir.

A linearidade é uma premissa crucial para a correlação canônica. Apesar de ser um modelo multivariado geral, a ACC baseia-se na relação linear entre as funções. Se esse relacionamento não for linear, a análise poderá não ser feita da forma correta, já que as relações não lineares não serão bem refletidas na análise (Mardia; Kent; Bibby, 1979). Nesse caso, uma ou ambas variáveis devem ser transformadas (Ribeiro, 2004).

A multicolinearidade é um problema comum, ocorrendo quando duas ou mais variáveis independentes apresentam alta correlação entre si. Com a presença de múltiplas variáveis explicando o mesmo evento, surge a dificuldade de isolar o impacto de cada uma, levando a resultados imprecisos ou inconsistentes (Hair Jr. *et al.*, 2018).

A presença de multicolinearidade pode inflacionar as estimativas de variância dos coeficientes, resultando em variáveis aparentemente não significativas, apesar de uma alta proporção da variância ser explicada pelo modelo, ou resultar em estimativas de coeficientes com valores distorcidos (Neisse; Hongyu, 2017).

A correlação canônica é mais eficiente na representação dos relacionamentos analisados quando as variáveis são homocedásticas. Isso é, quando as variáveis possuem variâncias homogêneas entre si (Meloun; Militký, 2011). Embora este não seja um requisito absoluto, a heterocedasticidade, seu oposto, pode levar a correlações menos significativas entre as variáveis, levando a resultados menos significativos e interpretações menos confiáveis (Hair Jr. *et al.*, 2018). Também pode ser necessária a transformação das variáveis para corrigir a heterocedasticidade.

2.4.2 CÁLCULO DAS FUNÇÕES E CORRELAÇÕES CANÔNICAS

Para a análise de correlação canônica, almeja-se encontrar combinações lineares das variáveis originais que capturem a maior correlação possível entre os grupos. Essas combinações, intituladas variáveis canônicas, são determinadas em pares e construídas de forma a maximizar a correlação entre elas. Cada par de variáveis canônicas corresponde a uma função canônica (Hair Jr. *et al.*, 2006). A

correlação entre cada par de variáveis canônicas é chamada correlação canônica. Assim, as correlações canônicas medem a intensidade da relação entre essas combinações lineares.

Sendo $\mathbf{x} = [x_1, x_2, \dots, x_p]$ e $\mathbf{y} = [y_1, y_2, \dots, y_q]$ os grupos de variáveis padronizadas que buscamos analisar e considerando U_1 e V_1 suas respectivas primeiras variáveis canônicas, definimos o par como:

$$U_1 = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_p x_p \quad (1)$$

e

$$V_1 = b_1 y_1 + b_2 y_2 + \dots + b_q y_q \quad (2)$$

em que $\boldsymbol{\alpha}' = [\alpha_1, \alpha_2, \dots, \alpha_p]$ e $\mathbf{b}' = [b_1, b_2, \dots, b_q]$ são os vetores de constantes que maximizam a correlação entre as variáveis U_1 e V_1 (Johnson; Wichern, 2007). Pode ser obtido ainda um segundo par de variáveis canônicas, não correlacionado com o primeiro, buscando maximizar a correlação entre as opções não contempladas pelo primeiro par.

Esse procedimento pode ser repetido até que toda a informação a ser explicada pelas funções canônicas seja exaurida, diminuindo a correlação com cada realização. Assim, o primeiro coeficiente de correlação canônica é sempre o de maior valor. Tendo dois vetores de variáveis com p e q dimensões respectivamente, o número de pares de variáveis canônicas que pode ser obtido é o valor mínimo entre p e q (Rencher, 2002). Supondo que $p \leq q$ e generalizando as equações (1) e (2) para todos os pares, temos (Ferreira, 2008):

$$\mathbf{U} = \boldsymbol{\alpha}' \mathbf{x}$$

e

$$\mathbf{V} = \mathbf{b}' \mathbf{y},$$

com $\mathbf{U}$ e $\mathbf{V}$ ambos com dimensões equivalentes ao valor de p .

Conforme descrito por Johnson e Wichern (2007), bem como Mingoti (2005), a correlação canônica entre $\mathbf{U}$ e $\mathbf{V}$ é dada por:

$$\rho = \text{Corr}(\mathbf{U}, \mathbf{V}) = \frac{\text{Cov}(\mathbf{U}, \mathbf{V})}{\sqrt{\text{Var}(\mathbf{U})\text{Var}(\mathbf{V})}} = \frac{\boldsymbol{\alpha}' \text{Cov}(\mathbf{x}, \mathbf{y}) \mathbf{b}}{\sqrt{\boldsymbol{\alpha}' \text{Var}(\mathbf{x}) \boldsymbol{\alpha}} \sqrt{\mathbf{b}' \text{Var}(\mathbf{y}) \mathbf{b}}}.$$

Assim, o valor máximo de correlação que pode ser obtido entre esses pares será dado por ρ_1^* , de forma que:

$$\rho_1^* = \max_{a,b} \text{Corr}(\mathbf{U}, \mathbf{V}) = \frac{\alpha_1' \text{Cov}(\mathbf{x}, \mathbf{y}) b_1}{\sqrt{\alpha_1' \text{Var}(\mathbf{x}) \alpha_1} \sqrt{b_1' \text{Var}(\mathbf{y}) b_1}}.$$

Surge então o problema de como identificar os valores de $\boldsymbol{\alpha}'$ e $\mathbf{b}'$ que maximizem essa correlação. A abordagem desse problema está ligada aos conceitos de autovalores e autovetores, elementos matemáticos utilizados no estudo de transformações lineares e matrizes (Afifi; Clark, 1999).

O autovetor de uma matriz $\mathbf{A}$ é o vetor não-nulo $\mathbf{u}$ que, quando multiplicado por essa matriz, não tem sua direção alterada, embora, em geral, alterem em magnitude. Portanto, deve ser solução de (Abdi, 2007; Higman, 1955):

$$\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$$

ou

$$(\mathbf{A} - \lambda)\mathbf{u} = 0,$$

na qual λ é um valor escalar conhecido como autovalor, associado a $\mathbf{u}$. Essas equações expressam que, ao aplicar a matriz $\mathbf{A}$ ao vetor $\mathbf{u}$, o resultado é simplesmente o vetor $\mathbf{u}$ escalado pelo autovalor λ .

Seja $\text{Corr}(\mathbf{x}) = \mathbf{R}_{xx}$, $\text{Corr}(\mathbf{y}) = \mathbf{R}_{yy}$, $\text{Corr}(\mathbf{x}, \mathbf{y}) = \mathbf{R}_{xy}$ e $\text{Corr}(\mathbf{y}, \mathbf{x}) = \mathbf{R}_{yx} = \mathbf{R}'_{xy}$, a matriz

de correlação entre os conjuntos de variáveis é dada por

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_{xx} & \mathbf{R}_{xy} \\ \mathbf{R}'_{xy} & \mathbf{R}_{yy} \end{pmatrix}.$$

Os vetores $\boldsymbol{\alpha}$ e $\mathbf{b}$ podem ser descritos como autovetores associados ao primeiro autovalor das matrizes $\mathbf{R}_{xx}^{-1}\mathbf{R}_{xy}\mathbf{R}_{yy}^{-1}\mathbf{R}_{yx}$ e $\mathbf{R}_{yy}^{-1}\mathbf{R}_{yx}\mathbf{R}_{xx}^{-1}\mathbf{R}_{xy}$, respectivamente, e são chamados de coeficientes canônicos, ou pesos canônicos (Johnson; Wichern, 2007; Neisse; Hongyu, 2017). Assim, conforme Mingoti (2005), os vetores $\boldsymbol{\alpha}'$ e $\mathbf{b}'$ podem ser encontrados como soluções do seguinte sistema de equações:

$$\begin{cases} (\mathbf{R}_{xy}\mathbf{R}_{yy}^{-1}\mathbf{R}_{yx} - \lambda_k\mathbf{R}_{xx})\boldsymbol{\alpha}_k = 0 \\ (\mathbf{R}_{yx}\mathbf{R}_{xx}^{-1}\mathbf{R}_{xy} - \lambda_k\mathbf{R}_{yy})\mathbf{b}_k = 0 \end{cases}$$

nos quais o autovalor λ_k , k -ésimo maior da matriz $(\mathbf{R}_{xx}^{-1}\mathbf{R}_{xy}\mathbf{R}_{yy}^{-1}\mathbf{R}_{yx})$, ou, equivalentemente, $\mathbf{R}_{yy}^{-1}\mathbf{R}_{yx}\mathbf{R}_{xx}^{-1}\mathbf{R}_{xy}$, satisfaz as seguintes equações:

$$\begin{cases} |\mathbf{R}_{xy}\mathbf{R}_{yy}^{-1}\mathbf{R}_{yx} - \lambda_k\mathbf{R}_{xx}| = 0 \\ |\mathbf{R}_{yx}\mathbf{R}_{xx}^{-1}\mathbf{R}_{xy} - \lambda_k\mathbf{R}_{yy}| = 0 \end{cases}.$$

O autovalor λ_k , associado a ambas as equações, recebe a denominação de raiz canônica, providenciando uma estimativa da variância compartilhada entre as variáveis canônicas do k -ésimo par (Hair Jr. *et al.*, 2018). Essa raiz canônica representa o quadrado da correlação máxima entre as k -ésimas variáveis canônicas, de forma que

$$\overline{\lambda}_1 = \text{Corr}(U_1, V_1) = \rho_1^*$$

corresponde à primeira correlação canônica. Por ser a primeira, trata-se também da maior correlação possível entre combinações lineares de $\mathbf{x}$ e $\mathbf{y}$, superando o valor de qualquer outra correlação entre essas variáveis (Rencher, 2002).

2.4.3 CRITÉRIOS PARA SELEÇÃO DAS FUNÇÕES CANÔNICAS

Após a obtenção das funções canônicas, é necessário verificar se o modelo canônico captura de forma suficiente a relação entre os conjuntos de variáveis preditoras e critério para justificar uma interpretação. Ou seja, deve-se verificar se existe uma relação relevante entre as variáveis em cada função canônica (Sherry; Henson, 2005).

Em seu trabalho, Hair Jr. *et al.* (2018) recomendam o uso de três critérios para a seleção de quais funções canônicas devem ser interpretadas: nível de significância, magnitude da correlação canônica e medida de redundância para a variância compartilhada entre os dois conjuntos de dados. Esses critérios são detalhados a seguir.

2.4.3.1 Testes de Significância

Existem diversas estatísticas de teste disponíveis para avaliar a significância estatística de uma hipótese nula de ausência de relação entre as variáveis explicativas e as variáveis preditoras. O nível de significância normalmente considerado é de 0,05 ou 5%, o que significa que há uma probabilidade de 5% de rejeitar a hipótese nula quando ela é verdadeira. Um nível de significância mais baixo (por exemplo, de 1%) torna o teste mais rigoroso (Rosenthal; Rubin, 1979).

Um teste aproximado para analisar a relação entre as variáveis x como um todo e as variáveis y como um todo foi proposto por Bartlett (1947), associado à distribuição qui-quadrado. Um valor significativamente grande indica que pelo menos uma das correlações canônicas r é significativa, enquanto um resultado não significativo sugere que a maior correlação canônica pode ser explicada apenas pela variação amostral (Manly; Navarro Alberto, 2017).

O teste mais utilizado, e normalmente fornecido por pacotes de computador, é a estatística F , baseada na aproximação de Rao (1948, 1973).

Ambos os testes são aplicáveis sob a suposição de que as variáveis x e y seguem conjuntamente uma distribuição normal multivariada. Um valor elevado de qui-quadrado ou na estatística F indica que ao menos uma das k correlações populacionais não é igual a zero (Afifi; Clark, 1999).

Além de testes separados de cada função canônica, testes multivariados incluindo todas as raízes canônicas também podem ser usadas para avaliar a importância dessas (Hair Jr. *et al.*, 2018). Cada teste oferece uma abordagem diferente para avaliar a variação entre e dentro das amostras, e como esta se relaciona com as médias populacionais (Manly; Navarro Alberto, 2017):

- Estatística Lambda de Wilks (Λ): Desenvolvida por Wilks (1932), compara a variação dentro dos grupos com a variação total (dentro e entre grupos). Um valor de Λ pequeno indica que a variação dentro de cada grupo é baixa comparada à variação total, indicando que há uma diferença significativa entre os grupos. Na ACC, indica que a correlação canônica é diferente de 0, e, portanto, estatisticamente significativa. Pode ser expressa da seguinte forma:

$$\Lambda = \prod_{i=1}^p \frac{1}{(1 + \lambda_i)}$$

- Maior Raiz de Roy (λ_1): Originalmente apresentada por Roy (1953), baseia-se no valor da maior raiz canônica obtida na análise (λ_1), podendo ser expressa diretamente como λ_1 ou ajustada para

$\frac{\lambda_1}{1-\lambda_1}$. Um valor elevado para a Maior Raiz de Roy indica que a maior correlação canônica é estatisticamente significativa, sugerindo uma forte associação entre os conjuntos de variáveis.

- Traço de Pillai (V): Obtido a partir dos autovalores λ_1 a λ_p , o Traço de Pillai oferece uma visão geral da importância das raízes canônicas combinadas. Valores grandes de V indicam que, em conjunto, as correlações canônicas são significativas. Pode ser calculado da seguinte forma (Pillai, 1955):

$$V = \sum_{i=1}^p \frac{\lambda_i}{1 + \lambda_i}.$$

Cada uma das estatísticas dos testes multivariados expostos pode ser aproximada à estatística mais familiar F e então avaliada. Em geral, os três testes tendem a fornecer níveis de significância similares (Manly; Navarro Alberto, 2017). Dentre eles, o Lambda de Wilks é o mais comumente utilizado na análise de correlação canônica, por sua maior aplicabilidade geral (Sherry; Henson, 2005).

2.4.3.2 Magnitude da Correlação Canônica

A significância prática de cada função canônica também pode ser medida pela magnitude de sua correlação canônica correspondente. A correlação canônica é definida pelo coeficiente de correlação de Pearson entre as duas variáveis canônicas em uma determinada função canônica. Como resultado da padronização dos coeficientes usados, esse valor deve ser positivo e varia de 0 a 1 (Sherry; Henson, 2005).

A magnitude da correlação canônica fornece uma medida de associação linear entre os dois conjuntos de variáveis. Magnitudes mais altas indicam uma relação mais forte, enquanto magnitudes mais baixas sugerem uma associação mais fraca ou potencialmente insignificante. As funções canônicas com maiores correlações geralmente representam padrões subjacentes mais relevantes para a ACC (Johnson; Wichern, 2007).

Como descrito por Anderson (2003), em casos onde os conjuntos são muito extensos, pode-se optar por considerar apenas algumas combinações lineares de cada conjunto, focando naquelas combinações mais fortemente correlacionadas. Rencher (2002) destaca as seguintes propriedades das correlações canônicas:

- A correlação canônica é invariante a mudanças de escala, tanto nas variáveis **y** quanto nas **x**. Se a escala de medida, por exemplo, mudar de polegadas para centímetros, as correlações canônicas não se alterarão, apesar de mudanças nos coeficientes canônicos.
- A primeira correlação canônica é a correlação máxima entre funções lineares de **y** e **x**. Portanto, essa excede o valor absoluto da correlação simples entre qualquer **y** e qualquer **x** ou da correlação múltipla entre qualquer **y** e todos os **x** ou entre qualquer **x** e todos os **y**.

Conforme descrito por Hair Jr. *et al.* (2018), não existem diretrizes universalmente aceitas sobre os tamanhos adequados para as correlações canônicas, se tratando, então, de uma decisão baseada no impacto prático das correlações canônicas dentro do contexto específico da investigação.

Apesar da diversidade dos critérios a serem utilizados, existem faixas de valores comumente aceitos e utilizados para a interpretação de um coeficiente de correlação r . Essas são definidas como (Hinkle; Wiersma; Jurs, 1988; Mukaka, 2012):

Tabela 1 – Diretrizes gerais para a interpretação de coeficientes de correlação

Coeficiente de correlação	Interpretação
$0,9 < r \leq 1$	Muito forte
$0,7 < r \leq 0,9$	Forte
$0,5 < r \leq 0,7$	Moderada
$0,3 < r \leq 0,5$	Fraca
$0 < r \leq 0,3$	Negligível

Fonte: Autoria própria, com base em Mukaka (2012) e Callegari-Jacques (2007).

Essa abordagem fornece diretrizes gerais para a interpretação da magnitude de uma correlação, por meio de coeficientes de Pearson. No entanto, é importante considerar também a significância estatística do coeficiente, o contexto da pesquisa e a relevância prática da correlação encontrada.

Ainda que a maioria dos pesquisadores concorde que um coeficiente menor que 0,1 representa uma relação negligível e, um maior que 0,9, uma relação muito forte, a classificação dos valores intermediários é discutível. Um coeficiente de correlação de 0,65, por exemplo, pode ser considerado "bom" ou "moderado", conforme o critério adotado (Schober; Boer; Schwarte, 2018).

2.4.3.3 Análise de Redundância

As raízes canônicas, calculadas por meio do quadrado das correlações canônicas, apresentam uma estimativa simples da variância compartilhada entre as variáveis canônicas. Apesar disso, cada função canônica explica apenas uma parcela da variância das variáveis que as compõem. Além disso, embora determinado par de variáveis canônicas possa apresentar correlação alta, isso não implica que cada uma das variáveis seja explicada nessa mesma proporção, visto que as cargas das variáveis individuais não são levadas em consideração (Hair Jr. *et al.*, 2018; Van Den Wollenberg, 1977).

Como apontado por Stewart e Love (1968), é relevante reconhecer que duas funções lineares podem apresentar uma forte correlação canônica entre si sem necessariamente extrair quantidades significativas de variância de seus respectivos conjuntos. Portanto, para a obtenção da variância explicada para cada variável canônica, deve ser levada em consideração não apenas sua raiz canônica,

mas também suas cargas canônicas, ou coeficientes estruturais (Thompson, 2005; Van Den Wollenberg, 1977).

Como forma de contornar o viés inerente na utilização de raízes canônicas como medida de variância compartilhada, Stewart e Love (1968) propuseram um método para a obtenção de uma estimativa assimétrica de redundância entre conjuntos de dados. Servindo como medida da variância compartilhada, de forma similar ao R^2 da regressão múltipla (Hair Jr. *et al.*, 2018). Definido pela variância extraída por determinada variável canônica, multiplicada por sua respectiva raiz canônica, o coeficiente de redundância (Rd) é um índice da proporção média da variância nas variáveis em um conjunto que pode ser explicada pela variável canônica do outro conjunto (Miller; Farr, 1971; Thompson, 2005).

Analogamente ao papel da variância extraída pela variável canônica, a estatística R^2 em uma regressão representa a correlação entre a variável dependente e a independente. No contexto canônico, buscamos a correlação entre a variável canônica e cada uma de suas variáveis constituintes (Hair Jr. *et al.*, 2018).

As cargas canônicas, que indicam a correlação de cada variável com sua variável canônica correspondente, são elevadas ao quadrado para quantificar a variação em cada variável explicada pela variável canônica. Assim, a proporção da variância do conjunto de variáveis extraída por sua variável canônica pode ser definida da seguinte forma (Hair Jr. *et al.*, 2018; Stewart; Love, 1968):

$$VC_{xn} = \frac{R_{xn1}^2 + \dots + R_{xni}^2}{i} \text{ e } VC_{yn} = \frac{R_{yn1}^2 + \dots + R_{ynj}^2}{j},$$

em que

VC_{xn} : variância compartilhada da n -ésima variável canônica do conjunto independente (X);

VC_{yn} : variância compartilhada da n -ésima variável canônica do conjunto dependente (Y);

R_{xni} : carga canônica da i -ésima variável do conjunto independente (X);

R_{ynj} : carga canônica da j -ésima variável do conjunto dependente (Y).

Valores altos de VC_n indicam que a n -ésima variável canônica representa bem o conjunto original de variáveis, extraindo uma alta proporção da variância total.

A estimativa da variância compartilhada entre as variáveis canônicas dependentes e independentes, necessária para o cálculo do índice de redundância, é dada pela raiz canônica, definida pelo quadrado da correlação canônica. Dessa forma, o índice de redundância é descrito como (Hair Jr. *et al.*, 2018; Stewart; Love, 1968; Tabachnick; Fidell, 2013):

$$Rd_{Cn} = VC_{Cn} \times \lambda_n,$$

Sendo Rd_{cn} o índice de redundância respectivo à n -ésima variável canônica do conjunto C e λ_n a n -ésima raiz canônica. Valores altos de Rd_{cn} indicam que a n -ésima variável canônica referente ao conjunto C é eficaz em explicar as variáveis do conjunto oposto.

2.4.4 INTERPRETAÇÃO

2.4.4.1 Pesos canônicos padronizados

Os coeficientes α e b , associados às variáveis canônicas U e V , respectivamente, refletem diferenças na contribuição das variáveis para a correlação canônica, bem como diferenças de escalonamento entre as variáveis. Sua padronização por meio da multiplicação pelo desvio padrão elimina o efeito de escala, de modo que representem a contribuição relativa de cada variável em sua respectiva variável canônica (Rencher, 2002).

A análise da magnitude dos pesos canônicos associados às variáveis é um dos métodos de interpretação de sua importância na função canônica. Variáveis com maior peso contribuem mais com a função canônica, e portanto tem papel mais importante na correlação dos conjuntos analisados, e variáveis com peso positivo contrastam com aquelas de peso negativo (Ferreira, 2008; Hair Jr. *et al.*, 2018).

Os coeficientes canônicos são interpretados de maneira análoga à interpretação dos coeficientes de regressão. Por exemplo, imaginamos, em uma análise, determinada variável “Sexo” possui coeficiente -0.5 na primeira função canônica. Nesse caso específico, determinado indivíduo ter sexo feminino leva a um decréscimo de 0.5 na variável canônica correspondente quando os outros preditores são mantidos constantes (UCLA, 2023).

Alguns dos obstáculos que podem ser encontrados na análise por pesos canônicos são: a instabilidade de resultados entre amostras, sensibilidade à seleção de variáveis e dificuldade de interpretação de valores em conjunto de dados com multicolinearidade (Hair Jr. *et al.*, 2018). Por esses motivos, técnicas adicionais de análise são recomendadas.

2.4.4.2 Cargas canônicas

As cargas canônicas (*canonical loadings*), também chamadas de correlações estruturais, se tratam do coeficiente de correlação de Pearson entre a variável original do estudo e a sua correspondente variável canônica. De acordo com Mingoti (2005), são definidas como:

$$R_{U_k X}^* = R_{xx} \alpha_k$$

e

$$R_{V_k Y}^* = R_{yy} b_k.$$

Essas cargas podem variar de -1 a +1. Elas contribuem na definição da estrutura da função canônica, quantificando a importância de cada variável original na obtenção de sua respectiva variável canônica e, portanto, sua utilidade no modelo (Sherry; Henson, 2005). Dessa forma, cada variável independente possui sua respectiva carga canônica.

O quadrado de uma carga canônica representa a proporção de variância compartilhada linearmente por uma variável com o seu respectivo composto canônico (Thompson, 2005). Essa estatística é análoga a outras medidas de tamanho do efeito de tipo R^2 .

Caso os coeficientes canônicos e as cargas canônicas mostrem diferenças significativas, apresentando sinais opostos, ambos devem ser analisados cuidadosamente para a interpretação dos resultados, conforme destacado por Afifi e Clark (1999). Os problemas de interpretação tendem a ser relacionados às variáveis canônicas subsequentes, em relação às primeiras, visto que a exigência de que combinações lineares subsequentes das variáveis sejam independentes das já obtidas impõe limitações aos resultados que podem dificultar sua interpretação.

Afifi e Clark (1999) também apontam que, quando o conjunto de variáveis usadas em uma variável canônica não estão correlacionados entre si, as cargas canônicas serão iguais aos pesos canônicos padronizados. Quando algumas das variáveis originais são altamente correlacionadas, as cargas e os coeficientes podem variar significativamente. Meredith (1964) acrescenta que, se as variáveis dentro de cada conjunto são moderadamente interrelacionadas, a possibilidade de interpretar as variáveis canônicas pela inspeção dos pesos canônicos é praticamente nula.

Nesses casos, é preferível a interpretação das cargas canônica no lugar dos pesos canônicos. Para ilustrar esse conceito, suponha que existam duas variáveis positivamente correlacionadas entre si, e que cada uma também está positivamente correlacionada com a variável canônica. Nesse exemplo, é possível que um peso respectivo na variável canônica seja positivo e o outro negativo. Enquanto isso, as cargas das variáveis canônicas seriam ambas positivas, refletindo o resultado esperado (Afifi; Clark, 1999).

Apesar de manter alguns dos mesmos desafios, como instabilidade amostral, Hair Jr. *et al.* (2018) constatam que a abordagem por cargas canônicas é considerada mais confiável que a interpretação por pesos canônicos. Os autores recomendam, ainda, o uso de cargas canônicas cruzadas a fim de evitar os problemas associados aos outros tipos de interpretação.

2.4.4.3 Cargas canônicas cruzadas

As cargas canônicas cruzadas (*canonical cross-loadings*) são definidas pela correlação entre a variável original de um conjunto com a variável canônica do outro conjunto. Podem ser representadas por (Mingoti, 2005):

$$R_{U_k Y}^* = R_{YX} \alpha_k$$

e

$$R_{V_k X}^* = R_{XY} b_k.$$

e revelam de forma mais direta os relacionamentos entre os grupos de variáveis. Como as cargas cruzadas tradicionais, se tratam de coeficientes de correlação de Pearson, e, por consequência, variam de -1 a +1.

As cargas cruzadas oferecem a vantagem de expressarem a relação de cada variável preditora original separadamente com a variável canônica de resposta, sem a interferência de outras variáveis predictoras. Sendo mais conservadoras e menos propensas à superestimação do que as cargas canônicas tradicionais, as cargas cruzadas estabelecem uma base mais confiável para a interpretação (Kuylen; Verhallen, 1981; Lambert; Durand, 1975).

O quadrado de uma carga canônica cruzada representa a proporção de variância compartilhada linearmente entre uma variável e o composto canônico oposto. Essa medida é particularmente relevante ao determinar a eficácia com que as variáveis independentes explicam as variáveis dependentes (Encabo, 2011). Adicionalmente, pode ser utilizada para o cálculo do índice de redundância, um indicador de sobreposição de informações entre os conjuntos de variáveis (Kuylen; Verhallen, 1981; Stewart; Love, 1968).

Embora a computação de cargas cruzadas não seja tão amplamente acessível em *softwares* estatísticos quanto as outras alternativas de interpretação, Hair Jr. *et al.* (2018) destacam essa abordagem como preferida sempre que disponível. Permitindo que se observe a correlação entre variáveis atreladas a conjuntos diferentes, essas não estariam tão sujeitas à multicolinearidade. Assim, o uso de cargas canônicas cruzadas serve para construir um modelo mais estável.

É importante notar que a utilização de cargas canônicas, tradicionais ou cruzadas, não substitui a análise de outros elementos da correlação canônica. Rencher (2002) argumenta que, por trazerem informações univariadas, devem ser interpretadas com cuidado. Hair Jr. *et al.* (2018) sublinham que, enquanto os coeficientes canônicos podem ser mais adequados para predições, as cargas canônicas são mais eficazes para explicar os construtos subjacentes, ressaltando a utilidade da consideração de vários métodos.

3 MATERIAL E MÉTODOS

3.1 Características do conjunto de dados

Os dados utilizados no estudo são provenientes do site *UCI Machine Learning Repository* (Kelly; Longjohn; Nottingham, 2023). Nomeado de “*Energy efficiency Data Set*”, esse conjunto foi criado e processado por Tsanas e Xifara (2012) e consiste em informações sobre uma análise energética de 12 formatos de edifício distintos, gerados por meio de simulações no *software Ecotect*. O detalhamento da metodologia das simulações foi disponibilizado pelos pesquisadores.

A partir de um cubo elementar de dimensões $3,5 \times 3,5 \times 3,5\text{m}$, foram gerados 12 modelos de edifícios, cada um composto de 18 cubos elementares. Apesar de todos possuírem o mesmo volume, $771,75\text{ m}^3$, apresentam áreas de superfície e dimensões distintas. A simulação assume que os edifícios estão localizados na cidade de Athena, na Grécia, e residem 7 pessoas, de estilo de vida sedentário. Definiu-se como ideal o intervalo de temperatura interna de 19 a 24°C .

O conjunto de dados contém 10 atributos, dos quais 8 são variáveis explicativas relacionadas às características físicas das edificações, incluindo dados sobre transparência, orientação cardinal e morfologia, e 2 são variáveis de resposta referentes ao consumo de energia para aquecimento e refrigeração. As variáveis incluídas na análise foram organizadas e codificadas como indicado na Tabela 2.

Tabela 2 – Descrição das variáveis utilizadas no estudo

Código	Construto	Número de valores possíveis	Grupo
X1	Compactação relativa	12	Características dos edifícios (X)
X2	Área de superfície	12	
X3	Área da parede	7	
X4	Área do telhado	4	
X5	Altura total	2	
X6	Orientação	4	
X7	Área envidraçada	4	
X8	Distribuição da área envidraçada	6	Cargas térmicas dos edifícios (Y)
Y1	Carga de aquecimento (HL)	586	
Y2	Carga de resfriamento (CL)	636	

Fonte: Autoria própria, com base em Tsanas e Xifara (2012).

A compactação relativa de um edifício é definida de forma que a razão entre seu volume e sua superfície é comparada com aquela do formato mais compacto possível, possuindo o mesmo volume. Em contraste com outros indicadores de compactação, que dependem do tamanho da forma, a compactação relativa é puramente dependente do formato do edifício (Mahdavi; Gurtekin, 2002; Pessenlehner; Mahdavi, 2003). Devido às suposições realizadas no estudo, constata-se que as variáveis de compactação relativa e área de superfície possuem uma relação inversamente proporcional.

Para a quantidade e distribuição de janelas, foram utilizadas quatro alternativas para a dimensão total da área envidraçada, expressas em porcentagens da área de superfície do piso: 0%, 10%, 25% e 40%. Além disso, foram simulados cinco cenários diferentes de distribuição para cada alternativa de área total envidraçada maior que 0%, conforme a Tabela 3. Todos os modelos foram rotacionados em intervalos de 90° de forma a se voltar aos quatro pontos cardeais.

Tabela 3 – Variações de distribuição de área envidraçada

Variação	Porcentagem da área envidraçada voltada à direção			
	Norte	Sul	Leste	Oeste
Uniforme	25	25	25	25
Norte	55	15	15	15
Sul	15	55	15	15
Leste	15	15	55	15
Oeste	15	15	15	55

Fonte: Autoria própria, com base em Tsanas e Xifara (2012) e Pessenlehner e Mahdavi (2003).

Dados os 12 formatos, 4 opções de área envidraçada, 5 variações de distribuição da área envidraçada e 4 orientações, foram obtidas 768 configurações únicas de edifícios para a amostra final. Esse resultado é calculado pela soma $(12 \times 3 \times 5 \times 4) + (12 \times 4)$. A primeira parte do cálculo aborda as combinações com área envidraçada de 10%, 25% e 40%, multiplicadas pelas cinco variações de distribuição e pelas quatro orientações. A adição final de 12×4 incorpora os modelos sem área envidraçada, multiplicando os 12 formatos de edifícios pelas 4 orientações.

Para cada um dos 768 edifícios, foram registradas as cargas de aquecimento e resfriamento. Esses parâmetros foram classificados como variáveis de resposta e representados por Y_1 e Y_2 , respectivamente, formando o grupo de cargas térmicas (Tabela 2).

3.2 Análise de Correlação Canônica

A metodologia empregada neste estudo visou explorar as relações entre as características de edifícios residenciais, designadas como variáveis independentes ou explicativas (X), e suas cargas térmicas, designadas como variáveis dependentes ou de resposta (Y), grupos de variáveis detalhados na Tabela 2. Para tanto, foi realizada uma análise de correlação canônica (ACC), uma técnica estatística multivariada que permite examinar as relações existentes entre dois conjuntos de variáveis quantitativas (Mingoti, 2005; Rencher, 2002).

Segundo metodologia proposta por Hotelling (1975), foram determinadas as funções e correlações canônicas, sendo as funções canônicas definidas como pares de combinações lineares das variáveis estudadas de forma a maximizar a correlação entre os conjuntos X e Y . O número máximo de funções canônicas que podem ser obtidas é determinado pelo tamanho do menor conjunto de variáveis entre os dois analisados (Ferreira, 2008; Johnson; Wichern, 2007). Dessa forma, foi possível a determinação de 2 funções canônicas ou 2 pares de variáveis canônicas.

A avaliação e seleção das funções canônicas obtidas baseou-se na análise de sua significância estatística, magnitude das correlações canônicas e medidas de redundância, conforme indicado por Hair Jr. *et al.* (2018). Para determinação do nível de significância das correlações canônicas, foi empregado o teste multivariado Lambda de Wilks, escolhido por sua flexibilidade e aplicabilidade geral (Sherry; Henson, 2005), considerando um nível de significância de 1%.

As medidas de redundância para a avaliação da variância extraída consideradas foram a variância compartilhada, a raiz canônica e o índice de redundância, conforme definido por Stewart e Love (1968). As raízes canônicas refletem a variância compartilhada entre as variáveis canônicas de um par, enquanto a medida de variância compartilhada representa a proporção da variância extraída pela variável canônica de cada conjunto de dados.

Finalmente, o coeficiente de redundância (R_d), calculado pela multiplicação da raiz canônica pela variância compartilhada, indica a proporção média da variância nas variáveis de um conjunto que pode ser explicada pela variável canônica do outro conjunto.

Após a seleção de funções canônicas a serem interpretadas, procedeu-se ao cálculo e análise dos pesos canônicos padronizados, cargas canônicas e cargas canônicas cruzadas. Dentre essas medidas, as cargas canônicas cruzadas foram priorizadas na interpretação por fornecerem informações mais diretas sobre o relacionamento entre as características de interesse e as cargas térmicas, além de sua maior estabilidade (Hair Jr. *et al.*, 2018; Kuylen; Verhallen, 1981; Lambert; Durand, 1975).

O quadrado das cargas e cargas cruzadas oferecem medidas adicionais de variância compartilhada, providenciando mais detalhes sobre a contribuição e capacidade explanatória de cada variável individual (Encabo, 2011; Thompson, 2005).

Para a visualização dos resultados da ACC, foram empregados diagramas de dispersão e gráficos *biplot*. Os escores canônicos, utilizados em ambas representações gráficas, são calculados pela

multiplicação das variáveis originais pelos seus respectivos pesos canônicos e representam a posição de cada observação nas dimensões canônicas respectivas (Mardia; Kent; Bibby, 1979; Thompson, 2005).

Por meio da diagramação da dispersão dos escores canônicos, junto da exibição da reta do modelo linear ajustado de determinada função canônica, a correlação linear entre os conjuntos de variáveis pode ser avaliada. Esse tipo de gráfico também é útil para acessar a presença de observações atípicas no conjunto de dados (Afifi; Clark, 1999; Johnson; Wichern, 2007). O gráfico *biplot*, por sua vez, auxilia na interpretação das relações multivariadas, destacando a distribuição espacial das variáveis em relação a ambas funções canônicas identificadas (Gabriel, 1971; Park; Huh, 1996).

Toda a análise estatística e cálculos envolvidos foram realizados por meio do *software* R (2023), versão 4.3.2. O principal pacote estatístico utilizado para as análises foi o *candisc* (Friendly; Fox, 2021), versão 0.8-6. Adicionalmente, foi utilizado o pacote CCA (González; Déjean, 2023), versão 1.2.2, para a obtenção do gráfico *biplot*.

4 RESULTADOS E DISCUSSÕES

A Análise de Correlação Canônica (ACC) realizada neste estudo visou explorar as relações entre os conjuntos de características físicas de edifícios residenciais simulados, aqui denominadas variáveis explicativas ou independentes (X), e suas respectivas cargas térmicas, aqui denominadas variáveis resposta ou dependentes (Y), conforme os procedimentos metodológicos detalhados na seção anterior.

A Tabela 4 detalha as principais estatísticas das funções canônicas obtidas, incluindo as correlações canônicas, a porcentagem da variância explicada por cada função e a aplicação do teste de significância multivariado Lambda de Wilks, com os níveis de significância associados. Nesse caso, foram obtidas 2 funções canônicas, sendo esse o menor número de variáveis entre os conjuntos estudados, conforme delineado por Hotelling (1936).

Ambas as funções canônicas foram consideradas estatisticamente significativas pelo teste de Wilks (p -valor $< 0,01$), indicando que as funções distinguem de forma significativa os conjuntos analisados. Contudo, deve ser considerado que os p -valores associados a uma ACC são altamente influenciados pelo tamanho de amostra, podendo apresentar significância para efeitos de magnitude mínima quando uma amostra é suficientemente grande (Sherry; Henson, 2005).

Diante dessas observações, a seleção das funções canônicas para análise detalhada foi orientada pelos critérios estabelecidos por Hair Jr. *et al.* (2018): (1) nível de significância, (2) magnitude da correlação canônica e (3) medida de redundância para a variância compartilhada entre os dois conjuntos de dados.

Tabela 4 – Correlações canônicas, porcentagens de variância explicada e teste estatístico Lambda de Wilks para as funções canônicas

Função canônica	Correlação canônica	Raiz canônica (R^2)	Estatística do teste	Aproximação à distribuição F	p-valor
1	0,95	0,90	0,08	157,77	$<0,01$
2	0,35	0,12	0,87	10,28	$<0,01$

A análise revela uma correlação canônica muito elevada associada à primeira função canônica (0,95), bem como uma raiz canônica substancial (0,90), sugerindo uma forte associação linear e considerável proporção da variância compartilhada entre as características físicas dos edifícios e suas cargas térmicas.

Apesar de um resultado significativo segundo a aplicação do teste Lambda de Wilks, a segunda função canônica apresentou um valor de correlação canônica consideravelmente baixo (0,35), com sua raiz canônica indicando uma proporção de variância compartilhada entre as variáveis canônicas de apenas 0,12.

A Tabela 5 apresenta detalhes adicionais relacionados à análise de variância compartilhada entre as variáveis canônicas específicas, conforme indicado por Stewart e Love (1968). Notavelmente, a primeira função canônica revela uma variância compartilhada de 98,03% para o conjunto de variáveis dependentes (Y) e de 36,46% para o conjunto de variáveis independentes (X).

Em contraste, a segunda função canônica apresenta uma variância compartilhada de 1,97% para o conjunto Y e de 14,15% para o conjunto X . Esses valores são muito menores, indicando que as variáveis canônicas dessa função não foram capazes de explicar tanto dos conjuntos correspondentes nas variáveis originais.

O índice de redundância, por sua vez, mede a capacidade de explicação de uma função canônica em relação ao conjunto de variáveis oposto. Esse também foi baixo na segunda função canônica: ela foi capaz de explicar 1,78% das variáveis independentes (X) com base nas variáveis dependentes (Y), e apenas 0,24% das variáveis dependentes (Y) com base nas independentes (X).

Tabela 5 – Análise de redundância respectiva a cada conjunto de variáveis e função canônica

Função canônica	Variância compartilhada (%)	Raiz canônica (R^2)	Índice de redundância (%)
Variáveis independentes (X)			
1	36,46	0,90	33,12
2	14,15	0,12	1,78
Variáveis dependentes (Y)			
1	98,03	0,90	89,05
2	1,97	0,12	0,24

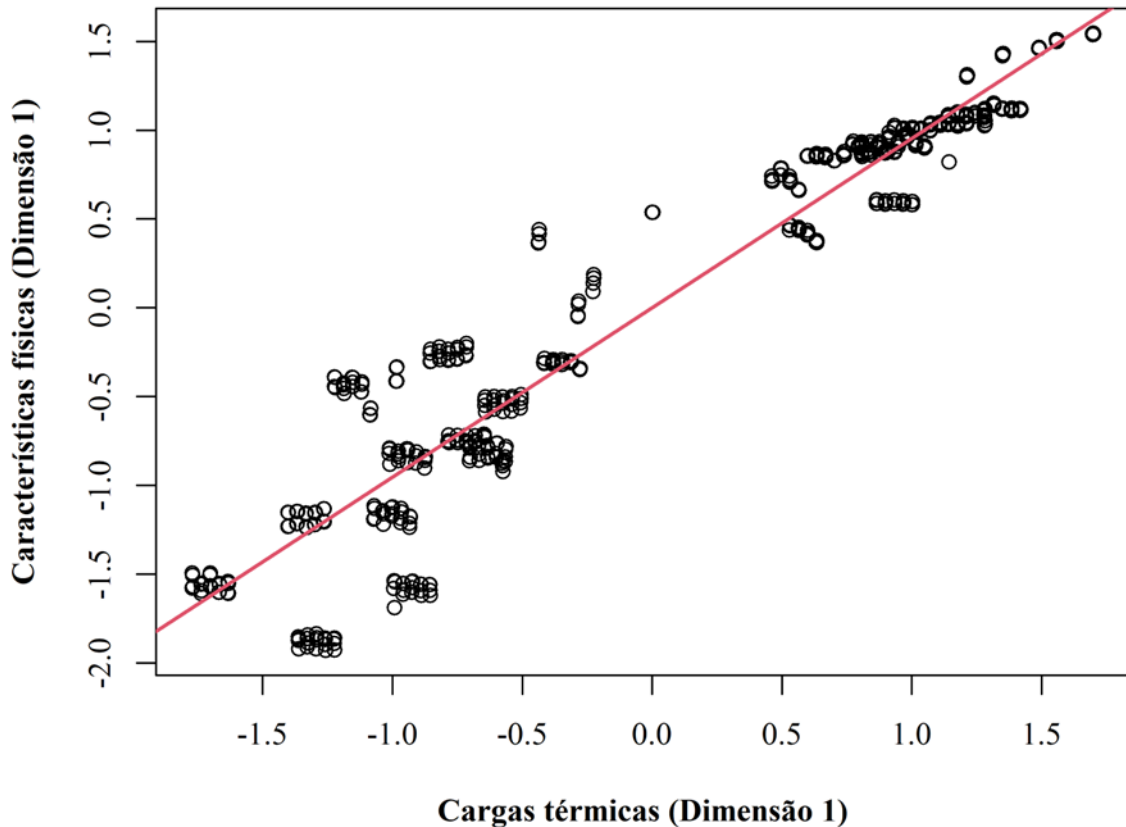
Esses resultados indicam que, embora seja estatisticamente significativa, a capacidade explanatória da função canônica 2 é limitada, principalmente em comparação com a função anterior. Por esses motivos, apenas a primeira função canônica será considerada para análise detalhada.

A Figura 1 apresenta a dispersão dos escores canônicos associados aos conjuntos de variáveis das características físicas dos edifícios e das cargas térmicas, ambos referentes à primeira função canônica. A linha vermelha indica uma regressão entre esses, sugerindo a tendência da relação entre as duas variáveis canônicas. Esse tipo de gráfico pode ser útil para verificar a existência de *outliers* no conjunto de dados (Afifi; Clark, 1999).

A reta ilustrada pelo ponto médio de ambos os conjuntos de escores canônicos, e sua inclinação indica a força e a direção da correlação canônica (Thompson, 2005). Dessa forma, o gráfico oferece uma representação visual da alta correlação canônica apresentada na Tabela 4, com a inclinação da reta sendo

equivalente a 0,95. A disposição aproximada dos pontos ao longo da linha de regressão e ausência de *outliers* notáveis corrobora a correlação linear significativa entre as variáveis analisadas.

Figura 1 – Diagrama de dispersão e modelo linear ajustado para a primeira função canônica



Adicionalmente, foi realizada a padronização dos coeficientes associados à função canônica escolhida, de forma a torna-los comparáveis entre variáveis de escalas diferentes (Rencher, 2002). A análise da primeira função canônica, com base nos pesos canônicos padronizados, cargas canônicas e cargas canônicas cruzadas, é sumarizada na Tabela 6.

Observa-se que, nos três métodos de interpretação analisados, um destaque consistente foi dado às variáveis X_1 , X_2 , X_4 e X_5 no grupo de variáveis independentes. Nota-se que, apresentando valores elevados tanto nos coeficientes canônicos quanto nas cargas canônicas, sugere-se que, além de contribuírem significativamente para a variável canônica, a variável original é bem representada por essa. Dessas, X_1 e X_4 demonstraram uma inversão de sinal ao comparar os pesos canônicos e as cargas canônicas, indicando uma relação de proporcionalidade inversa entre essas variáveis e as variáveis canônicas.

Tabela 6 – Pesos canônicos padronizados, cargas canônicas e cargas canônicas cruzadas referentes à primeira função canônica

Variáveis	Pesos canônicos	Cargas canônicas	Cargas canônicas cruzadas
Características físicas (<i>X</i>)			
<i>X</i> 1	0,75	-0,64	-0,61
<i>X</i> 2	6,19	0,68	0,65
<i>X</i> 3	-2,97	-0,49	-0,47
<i>X</i> 4	-5,61	0,90	0,86
<i>X</i> 5	-0,84	-0,94	-0,89
<i>X</i> 6	0,00	0,00	0,00
<i>X</i> 7	-0,21	-0,27	-0,25
<i>X</i> 8	-0,05	-0,16	-0,15
Cargas térmicas (<i>Y</i>)			
<i>Y</i> 1	-0,85	-1,00	-0,95
<i>Y</i> 2	-0,15	-0,98	-0,93

*Valores absolutos acima de 0,5 destacados em negrito.

Dentro do conjunto de cargas térmicas, tanto a carga de aquecimento (*Y*1) quanto a carga de resfriamento (*Y*2) apresentaram altos valores negativos, indicando alta proporção de variância compartilhada com sua respectiva variável canônica e com a oposta, e um relacionamento direto entre as cargas de aquecimento (HL) e resfriamento (CL).

Apesar das variáveis *X*2 e *X*3 destacarem-se em magnitude na análise dos pesos canônicos, essa característica não foi observada nas cargas canônicas, tradicionais ou cruzadas. Isso indica que, embora essas variáveis possuam um peso considerável na composição da função canônica, sua influência direta nas relações canônicas não é tão pronunciada quanto inicialmente sugerido.

Devido às cargas canônicas cruzadas fornecerem uma visão mais direta da relação entre os grupos de variáveis, bem como sua maior consistência em diferentes iterações e seleções de variáveis (Hair Jr. *et al.*, 2018; Kuylen; Verhallen, 1981), a interpretação subsequente terá foco nessas medidas. Conforme os dados da Tabela 6, essas cargas, definindo a correlação entre variáveis de um grupo com as canônicas do grupo oposto, exibiram valores consistentes com as cargas canônicas tradicionais, tanto em sinais quanto em magnitude. Essa congruência sugere que as variáveis significativas dentro de seu próprio grupo possuem uma relação equivalente e direta com as variáveis canônica do outro grupo.

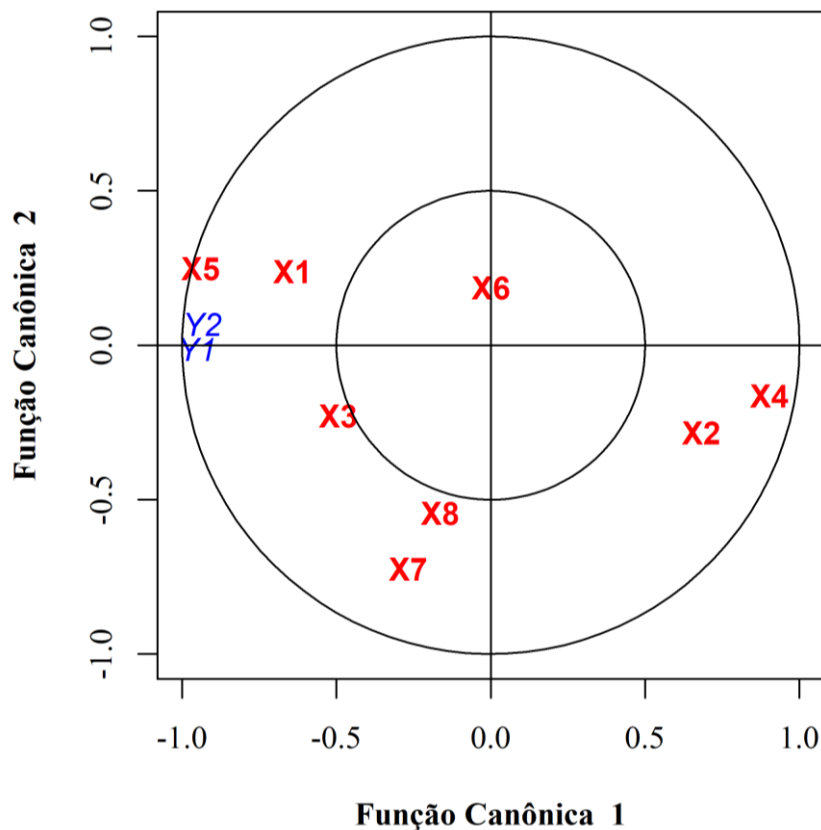
Na análise das cargas canônicas cruzadas, observa-se um maior destaque nas variáveis *X*5 (-0,89), *X*4 (0,86), *X*2 (0,65) e *X*1 (-0,61), em ordem de magnitude, indicando uma relação mais forte

com a variável canônica do conjunto Y . Especificamente, $X1$ e $X5$ apresentaram as maiores cargas canônicas cruzadas positivas, sugerindo uma associação inversa com as cargas térmicas, de valores negativos.

Por outro lado, $X2$ e $X4$ mostraram cargas canônicas cruzadas negativas, apontando para uma relação direta com as cargas térmicas. Isso sugere que valores maiores de área de superfície ($X2$) e área de telhado ($X4$) estão associados a uma diminuição das cargas térmicas, enquanto maiores valores de compactação relativa ($X1$) e altura total ($X5$) estão associados a um aumento das cargas térmicas.

A disposição espacial das variáveis analisadas também pode ser observada em uma análise bidimensional das funções canônicas 1 e 2, conforme Figura 2. As variáveis $Y1$ e $Y2$, correspondentes às cargas térmicas HL e CL, se encontram distanciadas da origem no eixo referente à função canônica 1, e próximas dessa naquele associado à função canônica 2. Esse posicionamento indica que as variáveis dependentes apresentam uma menor relevância da segunda função canônica, o que se alinha com sua menor capacidade de explicação.

Figura 2 – Gráfico *biplot* da distribuição espacial das variáveis em relação às funções canônicas 1 e 2



Ao analisar a distribuição bidimensional, a proximidade entre as variáveis de compactação relativa ($X1$) e altura total ($X5$) em relação a HL ($Y1$) e CL ($Y2$) sugere uma relação de proporcionalidade direta entre elas. Por sua vez, as variáveis de área de superfície ($X2$) e área de telhado ($X4$) se encontram

na direção oposta do gráfico, indicando proporcionalidade inversa com as cargas térmicas. Esses resultados corroboram as interpretações derivadas da análise de cargas canônicas e cargas canônicas cruzadas.

Tabela 7 – Cargas canônicas quadradas, cargas canônicas cruzadas quadradas e índices de redundância referentes à primeira função canônica

Variáveis	Cargas canônicas quadradas	Cargas canônicas cruzadas quadradas	Índice de redundância
Características físicas (X)			
X1	0,41	0,37	33,12%
X2	0,47	0,42	
X3	0,24	0,22	
X4	0,82	0,74	
X5	0,88	0,80	
X6	0,00	0,00	
X7	0,07	0,06	
X8	0,03	0,02	
Cargas térmicas (Y)			
Y1	1,00	0,91	89,05%
Y2	0,96	0,87	

*Valores absolutos acima de 0,5 destacados em negrito.

Detalhados na Tabela 7, os quadrados das cargas canônicas e das cargas canônicas cruzadas quantificam a proporção de variância compartilhada linearmente com a variável canônica de seu próprio grupo e com a variável canônica do grupo oposto, respectivamente. Os altos valores desses observados no conjunto *Y* indicam alta proporção de variância compartilhada entre as cargas térmicas e a variável canônica derivada do conjunto de características físicas (*X*), bem como sua própria variável canônica.

No conjunto de variáveis *X*, as variáveis equivalentes a área do telhado (*X*4) e altura total (*X*5) sobressaem-se, apresentando valores elevados tanto nas cargas canônicas quadradas (0,82 e 0,88) quanto nas cargas canônicas cruzadas ao quadrado (0,74 e 0,80). Esses resultados sugerem que essas variáveis não apenas compartilham uma significativa proporção de sua variância com a variável canônica de seu próprio grupo, indicando um bom ajuste e alta contribuição para essa, como à do grupo oposto, indicando que elas também têm um papel importante no relacionamento com as cargas térmicas.

Diversas investigações têm dedicado-se a estudar a relação de medidas de compactação do edifício com sua eficiência energética, em particular no que diz respeito às cargas térmicas, indicando a relevância desse parâmetro (D'Amico; Pomponi, 2019; Li; Yao, 2021; Ourghi; Al-Anzi; Krarti, 2007).

Devido aos cálculos específicos envolvidos no desenvolvimento do banco de dados, constata-se que as variáveis de compactação relativa e área de superfície possuem uma relação inversamente proporcional.

Conforme Kahraman e Köymen (2023), é sugerido que a envoltória de edifícios sejam mantidas o mais compactas o possível, aproximando-se a formas retangulares básicas para minimizar a perda de calor. Notavelmente, sua conceitualização envolve as medidas de área de superfície e volume, estando este último relacionado à altura, variáveis dadas como significativas na ACC.

Park *et al.* (2021) avaliaram como materiais de mudança de fase melhoram a eficiência energética de edifícios residenciais, considerando fatores como orientação e altura dos edifícios. Por meio dessa análise, os pesquisadores constataram que a economia de energia diminui conforme o aumento da altura do edifício, sendo influenciado também pelas condições climáticas específicas de cada região. Essa tendência é confirmada por outras pesquisas, evidenciando os desafios de manter edifícios mais altos energeticamente eficientes (Liu *et al.*, 2019; Lotfabadi; Hançer, 2019).

De forma complementar, Tsanas e Xifara (2012) empregaram técnicas de aprendizado de máquina para avaliar o efeito de diferentes variáveis nas cargas térmicas de edifícios. A compactação relativa e a área de superfície foram identificadas como características importantes para a determinação de cargas térmicas, de maneira similar aos resultados do presente estudo.

Embora os autores destaquem a relevância da área envidraçada em suas análises, essa foi pouco relevante na ACC. Outros pesquisadores, incluindo Hee *et al.* (2015) e Alwetaishi (2019), também enfatizam a relevância dessa característica na eficiência energética de edifícios. Inversamente, a altura total do edifício, pouco impactante na pesquisa de Tsanas e Xifara, foi significativa na análise de correlação canônica. Isso pode implicar que diferentes métodos têm diferentes sensibilidades à estrutura de dados, com a ACC sendo mais sensível a relações lineares.

Os índices de redundância obtidos mostram que 89,05% da variância das cargas térmicas foi explicada pela variável canônica das características físicas dos edifícios. De maneira inversa, a variância das características físicas (X) que é explicada pelas cargas térmicas (Y) é de 33,12%. Conforme Hair Jr. *et al.* (2018), quando há essa delimitação clara entre variáveis dependente e independente, é comum obter uma redundância mais baixa para as variáveis independentes, e isso não representa necessariamente um problema.

Nesse caso, a redundância mais baixa das variáveis dependentes é resultante da variância compartilhada relativamente baixa entre essas e sua própria variável canônica, não da raiz canônica, como pode ser verificado na Tabela 5. Embora não existam diretrizes específicas para a interpretação dos índices de redundância, 89,05% pode ser considerado notavelmente alto. Em comparação, Hair Jr. *et al.* (2018) obtiveram valores de 0,484 (descrito como substancial) e 0,204 (descrito como baixo) para o índice de redundância, enquanto Fialho, Jayme Jr. e Hermeto (2016) registraram medidas de redundância entre 0,04 e 0,41.

5 CONCLUSÃO

Os resultados do estudo indicam que a análise de correlação canônica foi eficiente para a avaliação das relações entre os parâmetros de construção e as cargas térmicas. A análise da primeira função canônica obtida revelou uma correlação de 0,95, refletindo uma forte correlação linear entre os conjuntos de variáveis estudados.

Adicionalmente, essa função canônica foi responsável por explicar 89,05% da variância das cargas térmicas e 33,12% da variância das características dos edifícios, demonstrando a sua eficácia. Dentre as variáveis independentes analisadas, a compactação relativa, área de superfície, área do telhado e altura total fornecem as contribuições mais substantivas para as variáveis canônicas e maior importância na determinação das cargas térmicas, com destaque para área de telhado e altura total.

Entre essas, a altura total demonstrou relacionamento direto com as cargas térmicas, indicando que maiores valores dessa característica podem ser desfavoráveis para a eficiência energética de um edifício. Em contrapartida, a área do telhado se apresentou inversamente correlacionada com as cargas térmicas, indicando que uma maior extensão de telhado pode levar a uma maior economia de energia.

REFERÊNCIAS

- ABDI, H. The eigen-decomposition: Eigenvalues and eigenvectors. In: SALKIND, N. J. (ed.). **Encyclopedia of measurement and statistics**. California: Sage Thousand Oaks, 2007. p. 304–308. Disponível em: <https://personal.utdallas.edu/~herve/Abdi-EVD2007-pretty.pdf>. Acesso em: 6 mar. 2024.
- AFIFI, A. A.; CLARK, V. **Computer-aided multivariate analysis**. 1. ed. Boca Raton: Chapman & Hall/CRC, 1999.
- ALWETAISHI, M. Impact of glazing to wall ratio in various climatic regions: A case study. **Journal of King Saud University - Engineering Sciences**, Riyadh, Saudi Arabia, v. 31, n. 1, p. 6–18, 2019. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1018363916300381>. Acesso em: 7 mar. 2024.
- ANDERSON, T. W. **An introduction to multivariate statistical analysis**. 3. ed. Hoboken, N.J: Wiley-Interscience, 2003. (Wiley series in probability and statistics).
- ASHRAE. **2021 ASHRAE handbook: Fundamentals**. Peachtree Corners, GA: ASHRAE, 2021.
- BALTAR, M. G.; KAEHLER, J. W. M.; PEREIRA, L. A. Indústria da construção civil e eficiência energética. In: CONFERÊNCIA ENGENHARIA'2005: INOVAÇÃO E DESENVOLVIMENTO, 3., 2005, Covilhã, Portugal. **Anais [...]**. Covilhã, Portugal: Universidade da Beira Interior, 2005. Disponível em: <https://repositorio.pucrs.br/dspace/bitstream/10923/3210/5/000384715-Texto%2bCompleto%2bAnexo%2bE-5.pdf>. Acesso em: 21 out. 2023.
- BARTLETT, M. S. The General Canonical Correlation Distribution. **The Annals of Mathematical Statistics**, United States, v. 18, n. 1, p. 1–17, 1947. Disponível em: <http://projecteuclid.org/euclid.aoms/1177730488>. Acesso em: 14 fev. 2024.
- BRASIL, MINISTÉRIO DE MINAS E ENERGIA. **Plano Nacional de Eficiência Energética**. Brasília, 2011. Disponível em: <http://bibliotecadigital.economia.gov.br/handle/123456789/914>. Acesso em: 3 abr. 2023.
- BURDICK, A. **Strategy Guideline: Accurate Heating and Cooling Load Calculations**. Pittsburgh, Pennsylvania: U.S. Department of Energy, 2011. Disponível em: <http://www.osti.gov/servlets/purl/1018100/>. Acesso em: 5 fev. 2024.
- CALLEGARI-JACQUES, S. M. **Bioestatística: Princípios E Aplicações**. Porto Alegre: Artmed, 2007.
- CHEN, J.; WANG, X.; STEEMERS, K. A statistical analysis of a residential energy consumption survey study in Hangzhou, China. **Energy and Buildings**, [s. l.], v. 66, p. 193–202, 2013. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0378778813004362>. Acesso em: 24 maio 2023.
- D'AMICO, B.; POMPONI, F. A compactness measure of sustainable building forms. **Royal Society Open Science**, United Kingdom, v. 6, n. 6, 2019. Disponível em: <https://royalsocietypublishing.org/doi/10.1098/rsos.181265>. Acesso em: 23 out. 2023.
- EMPRESA DE PESQUISA ENERGÉTICA - EPE. **Balanco Energético Nacional 2022: Ano base 2021**. Rio de Janeiro, 2022. Disponível em: <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/balanco-energetico-nacional-2022>. Acesso em: 2 abr. 2023.

EMPRESA DE PESQUISA ENERGÉTICA - EPE. **Ações para promoção da eficiência energética nas edificações brasileiras**: No caminho da transição energética. Rio de Janeiro, 2020. Nota técnica. Disponível em: <https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/Documents/NT%20DEA-SEE-007-2020.pdf>. Acesso em: 2 abr. 2023.

ENCABO, H. C. Canonical Correlation Analysis of Student Perception on Instructional Quality and Satisfaction. **JPAIR Multidisciplinary Research**, Philippines, v. 6, n. 1, 2011.

EVERITT, B. S. **An R and S-PLUS® Companion to Multivariate Analysis**. London: Springer, 2005. (Springer Texts in Statistics). Disponível em: <http://link.springer.com/10.1007/b138954>. Acesso em: 10 fev. 2024.

FERREIRA, D. F. **Estatística multivariada**. Lavras: Editora UFLA, 2008.

FIALHO, T. M. M.; JAYME JR., F. G.; HERMETO, A. M. Desenvolvimento do sistema financeiro e pobreza no Brasil: uma análise multivariada. **Economia e Sociedade**, Campinas, v. 25, n. 1, p. 247–278, 2016. Disponível em: <https://www.scielo.br/j/ecos/a/PLJwWq5y8LGhFLvySjhyCFq/>. Acesso em: 6 mar. 2024.

FRIENDLY, M.; FOX, J. **Candisc**: Visualizing Generalized Canonical Discriminant and Canonical Correlation Analysis. Versão 0.8-6. [S. l.], 2021. Disponível em: <https://CRAN.R-project.org/package=candisc>. Acesso em: 18 abr. 2023.

GABRIEL, K. R. The biplot graphic display of matrices with application to principal component analysis. **Biometrika**, Great Britain, v. 58, n. 3, p. 453–467, 1971. Disponível em: <https://www.jstor.org/stable/2334381>. Acesso em: 6 mar. 2024.

GONZÁLEZ, I.; DÉJEAN, S. **CCA**: Canonical Correlation Analysis. Versão 1.2.2. [S. l.], 2023. Disponível em: <https://CRAN.R-project.org/package=CCA>. Acesso em: 19 maio 2023.

HAIR JR., J. F. *et al.* Canonical Correlation. In: HAIR JR., J. F. *et al.* **Multivariate data analysis**. 8. ed. Hampshire, United Kingdom: Cengage, 2018. Disponível em: <https://mvstats.com/supplements/>. Acesso em: 21 out. 2023. Capítulo suplementar.

HAIR JR., J. F. *et al.* **Multivariate data analysis**. 6. ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2006.

HEE, W. J. *et al.* The role of window glazing on daylighting and energy saving in buildings. **Renewable and Sustainable Energy Reviews**, United Kingdom, v. 42, p. 323–343, 2015. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S136403211400793X>. Acesso em: 7 mar. 2024.

HIGMAN, B. **Applied group-theoretic and matrix methods**. Oxford: Clarendon press, 1955.

HINKLE, D. E.; WIERSMA, W.; JURS, S. G. **Applied statistics for the behavioral sciences**. 2. ed. Boston: Houghton Mifflin, 1988.

HOTELLING, H. Relations Between Two Sets of Variates. **Biometrika**, Oxford, v. 28, 1936.

HUBERTY, C. J.; MORRIS, J. D. Multivariate analysis versus multiple univariate analyses. **Psychological Bulletin**, United States, v. 105, n. 2, p. 302–308, 1989. Disponível em: <https://psycnet.apa.org/record/1989-21163-001>. Acesso em: 4 abr. 2024.

HYGH, J. S. *et al.* Multivariate regression as an energy assessment tool in early building design. **Building and Environment**, [s. l.], v. 57, p. 165–175, 2012. Disponível em:

<https://www.sciencedirect.com/science/article/abs/pii/S0360132312001400>. Acesso em: 24 maio 2023.

INTERNATIONAL ENERGY AGENCY - IEA. **Energy Efficiency 2019**. Paris, France: IEA Publications, 2019.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 6. ed. Upper Saddle River, N.J: Pearson Prentice Hall, 2007.

KAHRAMAN, Ö.; KÖYMEN, E. The Influence of Building Form Compactness on Energy Efficiency in Accommodation Structures: The Case of Türkiye. **Sakarya University Journal of Science**, Sakarya, Türkiye, v. 27, n. 5, p. 1055–1078, 2023. Disponível em: <https://dergipark.org.tr/en/pub/saufenbilder/issue/80257/1253136>. Acesso em: 7 mar. 2024.

KELLY, M.; LONGJOHN, R.; NOTTINGHAM, K. **The UCI Machine Learning Repository**. Irvine, California, 2023. Disponível em: <https://archive.ics.uci.edu/>. Acesso em: 1 out. 2023.

KUYLEN, A. A. A.; VERHALLEN, T. M. M. The use of canonical analysis. **Journal of Economic Psychology**, [s. l.], v. 1, n. 3, p. 217–237, 1981. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/0167487081900398>. Acesso em: 12 fev. 2024.

LAMBERT, Z. V.; DURAND, R. M. Some Precautions in Using Canonical Analysis. **Journal of Marketing Research**, United States, v. 12, n. 4, p. 468–475, 1975. Disponível em: <http://journals.sagepub.com/doi/10.1177/002224377501200411>. Acesso em: 13 fev. 2024.

LI, R.; LI, L.; WANG, Q. The impact of energy efficiency on carbon emissions: Evidence from the transportation sector in Chinese 30 provinces. **Sustainable Cities and Society**, [Basel, Switzerland], v. 82, 2022. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S2210670722002062>. Acesso em: 21 out. 2023.

LI, X.; YAO, R. Modelling heating and cooling energy demand for building stock using a hybrid approach. **Energy and Buildings**, [s. l.], v. 235, 2021. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0378778821000244>. Acesso em: 23 out. 2023.

LIU, C. *et al.* Analysis and optimization of load matching in photovoltaic systems for zero energy buildings in different climate zones of China. **Journal of Cleaner Production**, [s. l.], v. 238, p. 117914, 2019. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0959652619327842>. Acesso em: 22 out. 2023.

LOTFABADI, P.; HANÇER, P. A Comparative Study of Traditional and Contemporary Building Envelope Construction Techniques in terms of Thermal Comfort and Energy Efficiency in Hot and Humid Climates. **Sustainability**, [Basel, Switzerland], v. 11, n. 13, p. 3582, 2019. Disponível em: <https://www.mdpi.com/2071-1050/11/13/3582>. Acesso em: 22 out. 2023.

MAHDAVI, A.; GURTEKIN, B. Shapes, Numbers, Perception: Aspects and Dimensions of the Design-Performance Space. In: DESIGN AND DECISION SUPPORT SYSTEMS IN ARCHITECTURE, 2002, Eindhoven, The Netherlands. **Sixth Design and Decision Support Systems in Architecture and Urban Planning - Part one: Architecture Proceedings**. Eindhoven, The Netherlands: Eindhoven University of Technology, 2002. Disponível em: <https://papers.cumincad.org/data/works/att/ddssar0222.content.pdf>. Acesso em: 13 fev. 2024.

MANLY, B. F. J.; NAVARRO ALBERTO, J. A. **Multivariate statistical methods: a primer**. 4. ed. Boca Raton: CRC Press, Taylor & Francis Group, 2017.

MARDIA, K. V.; KENT, J. T.; BIBBY, J. M. **Multivariate analysis**. 1. ed. London: Academic Press, 1979. (Probability and mathematical statistics).

MATOS, R. S. **Climatização**. Apostila da disciplina TM182 - Refrigeração e Climatização. Paraná: UFPR, 2010.

MELOUN, M.; MILITKÝ, J. Statistical analysis of multivariate data. In: MELOUN, M.; MILITKÝ, J. **Statistical Data Analysis: A Practical Guide**. India: Woodhead Publishing, 2011. p. 151–403.

MEREDITH, W. Canonical correlations with fallible data. **Psychometrika**, United States, v. 29, n. 1, p. 55–65, 1964. Disponível em: <http://link.springer.com/10.1007/BF02289567>. Acesso em: 13 fev. 2024.

MILLER, J. K.; FARR, S. D. Bimultivariate Redundancy: A Comprehensive Measure of Interbattery Relationship. **Multivariate Behavioral Research**, United States, v. 6, n. 3, p. 313–324, 1971. Disponível em: http://www.tandfonline.com/doi/abs/10.1207/s15327906mbr0603_4. Acesso em: 28 fev. 2024.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: Editora UFMG, 2005.

MUKAKA, M. M. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. **Malawi Medical Journal**, Malawi, v. 24, n. 3, p. 69–71, 2012. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3576830/>. Acesso em: 3 mar. 2024.

NEISSE, A. C.; HONGYU, K. Variáveis Psicológicas e Desempenho Acadêmico: Uma Análise Da Existência de Correlação Canônica. **E&S Engineering and Science**, Mato Grosso, v. 6, n. 2, p. 76–86, 2017. Disponível em: <https://periodicoscientificos.ufmt.br/ojs/index.php/eng/article/view/4577>. Acesso em: 23 abr. 2023.

OURGHI, R.; AL-ANZI, A.; KRARTI, M. A simplified analysis method to predict the impact of shape on annual energy use for office buildings. **Energy Conversion and Management**, [s. l.], v. 48, n. 1, p. 300–305, 2007. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0196890406001397>. Acesso em: 23 out. 2023.

PARK, J. H. *et al.* Energy retrofit of PCM-applied apartment buildings considering building orientation and height. **Energy**, [s. l.], v. 222, 2021. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0360544221001262>. Acesso em: 22 out. 2023.

PARK, M.-R.; HUH, M.-H. Canonical correlation biplot. **Communications for Statistical Applications and Methods**, Korea, v. 3, n. 1, p. 11–19, 1996. Disponível em: <https://koreascience.kr/article/JAKO199611920523151.page>. Acesso em: 6 mar. 2024.

PÉREZ-LOMBARD, L.; ORTIZ, J.; POUT, C. A review on buildings energy consumption information. **Energy and Buildings**, [s. l.], v. 40, n. 3, p. 394–398, 2008. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0378778807001016>. Acesso em: 1 abr. 2023.

PESSSENLEHNER, W.; MAHDAVI, A. Building morphology, transparency, and energy performance. In: EIGHTH INTERNATIONAL IBPSA CONFERENCE, 2003, Eindhoven, Netherlands. **Anais [...]**. Eindhoven, Netherlands: Building Simulation, 2003. Disponível em: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=8e354f38264892070c7037d9203e68f0219b858f>. Acesso em: 13 dez. 2024.

PILLAI, K. C. S. Some New Test Criteria in Multivariate Analysis. **The Annals of Mathematical Statistics**, United States, v. 26, n. 1, p. 117–121, 1955. Disponível em: <http://projecteuclid.org/euclid.aoms/1177728599>. Acesso em: 14 fev. 2024.

PROGRAMA BRASILEIRO DE ETIQUETAGEM DE EDIFICAÇÕES. **Sobre o PBE Edifica**. Brasil, 2020. Disponível em: <https://www.pbeedifica.com.br/sobre>. Acesso em: 12 fev. 2024.

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Versão 4.2.3. Vienna, Austria: R Foundation for Statistical Computing, 2023. Disponível em: <https://www.R-project.org/>. Acesso em: 18 abr. 2023.

RAO, C. R. **Linear Statistical Inference and its Applications**. 2. ed. [S. l.]: Wiley, 1973. (Wiley Series in Probability and Statistics). Disponível em: <https://onlinelibrary.wiley.com/doi/book/10.1002/9780470316436>. Acesso em: 14 fev. 2024.

RAO, C. R. Tests of Significance in Multivariate Analysis. **Biometrika**, Oxford, v. 35, n. 1/2, p. 58, 1948. Disponível em: <https://www.jstor.org/stable/2332629?origin=crossref>. Acesso em: 14 fev. 2024.

RENCHE, A. C. **Methods of multivariate analysis**. 2. ed. New York: J. Wiley, 2002.

RIBEIRO, M. M. **Análise de correlação canônica: estruturação teórica e aplicações em estatística ambiental**. 2004. Trabalho de Conclusão de Curso (Bacharelado em Estatística) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2004.

ROSENTHAL, R.; RUBIN, D. B. Comparing significance levels of independent studies. **Psychological Bulletin**, [s. l.], v. 86, n. 5, p. 1165–1168, 1979. Disponível em: <https://doi.apa.org/doi/10.1037/0033-2909.86.5.1165>. Acesso em: 14 fev. 2024.

ROY, S. N. On a Heuristic Method of Test Construction and its use in Multivariate Analysis. **The Annals of Mathematical Statistics**, United States, v. 24, n. 2, p. 220–238, 1953. Disponível em: <http://projecteuclid.org/euclid.aoms/1177729029>. Acesso em: 14 fev. 2024.

RUTKOWSKI, H. **Residential Load Calculation Manual J**. 7. ed. Arlington, VA: Air Conditioning Contractors of America, 2014.

SCHÖBER, P.; BOER, C.; SCHWARTE, L. A. Correlation Coefficients: Appropriate Use and Interpretation. **Anesthesia & Analgesia**, [s. l.], v. 126, n. 5, p. 1763–1768, 2018. Disponível em: <https://journals.lww.com/00005339-201805000-00050>. Acesso em: 28 fev. 2024.

SCHWARZ, P. M. **Energy economics**. New York: Routledge, Taylor & Francis Group, 2018. (Routledge Textbooks in Environmental and Agricultural Economics, v. 9).

SHERRY, A.; HENSON, R. K. Conducting and Interpreting Canonical Correlation Analysis in Personality Research: A User-Friendly Primer. **Journal of Personality Assessment**, [s. l.], v. 84, n. 1, p. 37–48, 2005. Disponível em: http://www.tandfonline.com/doi/abs/10.1207/s15327752jpa8401_09. Acesso em: 16 maio 2023.

SILVA, J. G. da. **Introdução à tecnologia da refrigeração e da climatização**. São Paulo: Artliber, 2010.

STEWART, D.; LOVE, W. A general canonical correlation index. **Psychological Bulletin**, Florida, v. 70, n. 3, Pt.1, p. 160–163, 1968. Disponível em: <https://psycnet.apa.org/record/1968-18070-001>. Acesso em: 19 maio 2023.

TABACHNICK, B. G.; FIDELL, L. S. **Using multivariate statistics**. 6. ed. Boston: Pearson, 2013.

THOMPSON, B. Canonical Correlation Analysis. In: EVERITT, B. S.; HOWELL, D. C. (org.). **Encyclopedia of Statistics in Behavioral Science**. 1. ed. Hoboken, N.J: John Wiley & Sons, 2005. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/0470013192.bsa068>. Acesso em: 5 nov. 2023.

TSANAS, A.; XIFARA, A. Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. **Energy and Buildings**, [s. l.], v. 49, p. 560–567,

2012. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S037877881200151X>. Acesso em: 9 abr. 2023.

UCLA: STATISTICAL CONSULTING GROUP. **Canonical Correlation Analysis**. UCLA: Los Angeles, 2023. Disponível em: <https://stats.oarc.ucla.edu/r/dae/canonical-correlation-analysis/>. Acesso em: 11 fev. 2024.

UNITED NATIONS ENVIRONMENT PROGRAMME. **2022 Global Status Report for Buildings and Construction**: Towards a Zero-emission, Efficient and Resilient Buildings and Construction Sector. Nairobi: GlobalABC, 2022.

VAN DEN WOLLENBERG, A. L. Redundancy analysis an alternative for canonical correlation analysis. **Psychometrika**, Nijmegen, v. 42, n. 2, p. 207–219, 1977. Disponível em: <http://link.springer.com/10.1007/BF02294050>. Acesso em: 19 maio 2023.

WANG, S.; YAN, C.; XIAO, F. Quantitative energy performance assessment methods for existing buildings. **Energy and Buildings**, [s. l.], v. 55, p. 873–888, 2012. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0378778812004410>. Acesso em: 29 fev. 2024.

WILKS, S. S. Certain Generalizations in the Analysis of Variance. **Biometrika**, Oxford, v. 24, n. 3/4, p. 471, 1932. Disponível em: <https://www.jstor.org/stable/2331979?origin=crossref>. Acesso em: 14 fev. 2024.

YEZIORO, A.; DONG, B.; LEITE, F. An applied artificial intelligence approach towards assessing building performance simulation tools. **Energy and Buildings**, [s. l.], v. 40, n. 4, 2008. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0378778807001430>. Acesso em: 4 abr. 2023.