UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"
FACULDADE DE CIÊNCIAS AGRONÔMICAS
CÂMPUS DE BOTUCATU

UTILIZAÇÃO DE PROCEDIMENTOS MULTIVARIADOS NO CONSUMO DE ÁGUA E ENERGIA ELÉTRICA EM HABITAÇÕES SOCIAIS COM SISTEMA DE AQUECIMENTO SOLAR

MARCELO HIROSHI TUTIA

Tese apresentada à Faculdade de Ciências Agronômicas da UNESP – Câmpus de Botucatu, para obtenção do título de Doutor em Agronomia (Energia na Agricultura).

BOTUCATU-SP
Junho - 2015

UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"
FACULDADE DE CIÊNCIAS AGRONÔMICAS
CÂMPUS DE BOTUCATU

UTILIZAÇÃO DE PROCEDIMENTOS MULTIVARIADOS NO CONSUMO DE ÁGUA E ENERGIA ELÉTRICA EM HABITAÇÕES SOCIAIS COM SISTEMA DE AQUECIMENTO SOLAR

MARCELO HIROSHI TUTIA

Orientador: Prof. Dr. Carlos Roberto Padovani

Tese apresentada à Faculdade de Ciências Agronômicas da UNESP – Câmpus de Botucatu, para obtenção do título de Doutor em Agronomia (Energia na Agricultura).

BOTUCATU-SP
Junho - 2015

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉCNICA DE AQUISIÇÃO E TRATAMENTO DA INFORMAÇÃO - DIRETORIA TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - UNESP - FCA - LAGEADO - BOTUCATU (SP)

Tutia, Marcelo Hiroshi, 1977-
T966u Utilização de procedimentos multivariados no consumo de água e energia elétrica em habitações sociais com sistema de aquecimento solar / Marcelo Hiroshi Tutia. - Botucatu : [s.n.], 2015
viii, 98 f. : grafs. color., tabs.

Tese (Doutorado) - Universidade Estadual Paulista, Faculdade de Ciências Agronômicas, Botucatu, 2015
Orientador: Carlos Roberto Padovani
Inclui bibliografia

1. Água - Aquecimento solar. 2. Análise multivariada. 3. Habitação popular. 4. Fatores socioeconômicos. I. Padovani, Carlos Roberto. II. Universidade Estadual Paulista "Júlio de Mesquita Filho" (Câmpus de Botucatu). Faculdade de Ciências Agronômicas. III. Título.

UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"

FACULDADE DE CIÊNCIAS AGRONÔMICAS

CAMPUS DE BOTUCATU

CERTIFICADO DE APROVAÇÃO

TÍTULO: **"UTILIZAÇÃO DE PROCEDIMENTOS MULTIVARIADOS NO CONSUMO DE ÁGUA E ENERGIA ELÉTRICA EM HABITAÇÕES SOCIAIS COM SISTEMA DE AQUECIMENTO SOLAR"**

ALUNO: MARCELO HIROSHI TUTIA

ORIENTADOR: PROF. DR. CARLOS ROBERTO PADOVANI

Aprovado pela Comissão Examinadora

PROF. DR. CARLOS ROBERTO PADOVANI

PROF. DR. OSMAR DE CARVALHO BUENO

PROF. DR. PAULO ANDRÉ DE OLIVEIRA

PROF. DR. ROGÉRIO ANTONIO DE OLIVEIRA

PROFA. DRA. MARIE OSHIIWA

Data da Realização: 06 de maio de 2015

DEDICATÓRIA

A minha esposa Evellyn.

A meus pais, Masakazu (*in memorian*) e Sumie, aos meus irmãos, Fábio e Flávia, e sobrinhos, Guilherme, Gabriel e Pedro.

AGRADECIMENTO

A Deus.

Ao Prof. Dr. Carlos Roberto Padovani, pela orientação, sabedoria e ensinamentos para realização deste trabalho.

Aos Professores, Dr. Osmar de Carvalho Bueno, Dr. Paulo André de Oliveira, Drª. Marie Oshiiwa, Dr. Rogério Antonio de Oliveira, Dr. Sergio Augusto Rodrigues, Drª. Nilza Regina da Silva, Dr. Paulo Isler, Dr. Sidney Carlos Ferrari, Dr. Mauri da Silva, Drª. Gislaine Cristina Batistela, pela troca de conhecimentos e sugestões.

A Superintendência de Água e Esgoto – SAE de Ourinhos-SP e a Companhia Paulista de Luz e Força – CPFL, pelos dados disponibilizados.

Ao Sr. Gustavo Gomes e ao Sr. Roberto Alves de Oliveira, pelas ideias, informações e sugestões.

A Prefeitura de Ourinhos pelas informações disponibilizadas.

A Faculdade de Tecnologia de Ourinhos, a Diretora, Profª. Drª. Lia Cupertino Duarte Albino, a Diretora Acadêmica, Renata Rosolem Góes, a Diretora de Serviços, Telma Menegazzo, aos Coordenadores de Curso, aos Professores e Funcionários.

Aos funcionários da Unesp – Lageado, biblioteca e pós-graduação.

A todos que contribuíram para o desenvolvimento deste trabalho, meus sinceros agradecimentos.

Sumário

RESUMO 1
SUMMARY 2
1 INTRODUÇÃO 3
2 REVISÃO DE LITERATURA 6
2.1 Energia solar 6
2.2 Sistema de aquecimento solar 9
2.3 Habitações de interesse social 12
2.4 Técnicas univariadas 13
2.5 Técnicas multivariadas 14
3 MATERIAIS E MÉTODOS 17
3.1 Materiais 17
3.1.1 Variáveis de consumo de água 22
3.1.2 Variáveis de consumo de energia elétrica 22
3.1.3 Variáveis socioeconômicas 23
3.2 Métodos 26
3.2.1 Análise exploratória 26
3.2.2 Análise multivariada 27
3.2.3 Teste de dois vetores de médias 29
3.2.2.1 Intervalos de confiança simultâneos de Hotelling 34
3.2.4 Análise de correlação canônica 34
3.2.4 Análise fatorial 38
4 RESULTADOS E DISCUSSÃO 42
4.1 Análise exploratória de dados 42
4.2 Teste T^2 de Hotelling para comparação de dois vetores de médias e respectivo intervalo de confiança. 55
4.2 Correlação canônica 60
4.3 Análise Fatorial 71
5 CONCLUSÃO 80
6 REFERÊNCIAS 84
ANEXO 1 – Ficha A 89
ANEXO 2 – Critério de classificação Brasil 91
APÊNDICE 1 – Questionário 96

LISTA DE TABELAS

Tabela 1: Representação tabular dos dados de consumos de água, energia elétrica e socioeconômicos. 28
Tabela 2: Distribuição da quantidade de residências pesquisadas amostradas, por bairro. 42
Tabela 3:Distribuição do número de moradores por residência. 43
Tabela 4:Distribuição do número de cômodos por residência. 43
Tabela 5: Distribuição da utilização de lâmpadas fluorescentes em grande parte dos cômodos. 44
Tabela 6: Distribuição da condição de ocupação das residências. 44
Tabela 7: Distribuição da faixa etária média por residência. 45
Tabela 8: Distribuição da quantidade de moradores do sexo feminino por residência. 45
Tabela 9: Distribuição da quantidade de moradores do sexo masculino por residência. 46
Tabela 10: Distribuição da classe econômica por residência no ano de 2013. 46
Tabela 11: Distribuição do grau de instrução do chefe da família por residência. 47
Tabela 12: Distribuição do consumo médio de água (m^3) por residência. 48
Tabela 13: Distribuição do consumo médio de energia elétrica (kWh) por residência. 48
Tabela 14: Distribuição da razão do número de moradores por número de cômodos por residência. 49
Tabela 15: Distribuição da razão da renda média familiar por número de moradores por residência. 49
Tabela 16: Distribuição do indicador da variação do sexo por residência. 50
Tabela 17: Distribuição da média dos anos de estudos por residência. 50
Tabela 18: Distribuição do indicador de idade dependente por residência. 51
Tabela 19: Distribuição do índice de maior permanência por residência. 51
Tabela 20:Distribuição da quantidade de itens que consomem energia elétrica por residência. 52
Tabela 21: Distribuição do índice de equipamentos com consumo alto por residência. 52
Tabela 22: Nível descritivo e resultado do teste de hipóteses da homogeneidade. 53
Tabela 23: Medidas descritivas das variáveis de consumo (água e energia elétrica). 53
Tabela 24: Medidas descritivas das variáveis socioeconômicas. 54
Tabela 25: Diferença entre as médias dos conjuntos habitacionais, limite inferior e limite superior do intervalo de confiança simultâneo de Hotelling. 58
Tabela 26: Correlação linear, variâncias e covariâncias das variáveis de consumo e as variáveis socioeconômicas, Conjunto Habitacional Sem Sistema. 60
Tabela 27: Correlação linear, variâncias e covariâncias das variáveis de consumo e as variáveis socioeconômicas, Conjunto Habitacional Com Sistema. 61
Tabela 28: Correlação canônica das variáveis de consumo e as variáveis socioeconômicas dos conjuntos habitacionais. 61
Tabela 29: Coeficientes canônicos e suas associações com as variáveis de consumo e as variáveis socioeconômicas no conjunto habitacional Sem Sistema. 63
Tabela 30: Coeficientes canônicos e suas associações com as variáveis de consumo e as variáveis socioeconômicas, no Conjunto Habitacional Com Sistema. 66
Tabela 31: Indicativo das características estudas para consumo baixo de água e energia nos conjuntos habitacionais. 69
Tabela 32: Correlação e nível descritivo das 11 variáveis consideradas na análise fatorial. 71
Tabela 33: Estatística Kaiser-Meyer-Olkin e Teste de Esferacidade de Bartlett. 72

Tabela 34: Variância total explicada pelos fatores. 73
Tabela 35: Medida de adequação da amostra, MSA, para cada variável. 74
Tabela 36: Comunalidades das variáveis consideradas na análise fatorial. 75
Tabela 37: Cargas fatoriais das variáveis para os cinco fatores extraídos (rotação Varimax). 75
Tabela 38: Mediana, mínimo, máximo, nível descritivo e resultado do teste não paramétrico Mann-Whitney. 78

LISTA DE FIGURAS

Figura 1: Q-Q plot para o conjunto de variáveis de consumo dos Conjuntos Habitacionais Com e Sem Sistema. 56
Figura 2: Q-Q plot para o conjunto de variáveis socioeconômicas dos Conjuntos Habitacionais Com e Sem Sistema. 57
Figura 3: Intervalos de confiança simultâneos de Hotelling para a diferença entre as médias dos conjuntos habitacionais, variáveis de consumo. 58
Figura 4: Intervalos de confiança simultâneos de Hotelling para a diferença entre as médias dos conjuntos habitacionais, variáveis socioeconômicas. 59
Figura 5: Gráfico de dispersão do primeiro par de variáveis canônicas dos Conjuntos Habitacionais Com Sistema e Sem Sistema. 62
Figura 6: Gráfico de dispersão da associação e associação cruzada de cada respectiva variável, Conjunto Habitacional Sem Sistema. 63
Figura 7: Gráfico das correlações das variáveis com os componentes canônicos, Conjunto Habitacional Sem Sistema. 64
Figura 8: Gráfico de dispersão da associação e associação cruzada de cada respectiva variável, Conjunto Habitacional Com Sistema. 66
Figura 9: Gráfico das correlações das variáveis com os componentes canônicos, Conjunto Habitacional Com Sistema. 67
Figura 10: Gráfico *Scree Plot* das variáveis. 73

RESUMO

Atualmente, fontes de energia renováveis são exploradas e usadas para o desenvolvimento da sociedade e se destacam por serem mais limpas e menos poluentes que as fontes de energia não renováveis. Entre estas fontes de energia renováveis, a energia solar é uma das mais aproveitadas, usada em sistemas fotovoltaicos para geração de energia elétrica e em sistemas de aquecimento solar de água, sendo a segunda utilização citada a mais conhecida, podendo ser vista em residências, comércios e indústrias. O sistema de aquecimento solar tem entre seus objetivos, o de proporcionar economia de energia elétrica, pois fornece água quente sem a necessidade de utilizar sistemas elétricos, como o chuveiro elétrico, que é um dos aparelhos que mais consome energia. Para proporcionar economia do consumo de energia elétrica, minimizar impactos ambientais, redirecionar investimentos em novas fontes geradoras de energia elétrica, o Governo do Estado de São Paulo, em conjunto com a Companhia de Desenvolvimento Habitacional e Urbano do Estado de São Paulo - CDHU, como política de habitação, instalaram, nas residências de interesse social construídas a partir de 2007, o sistema de aquecimento solar de água. Na cidade de Ourinhos – SP, o conjunto habitacional Helena Braz Vendramini foi contemplado pelo sistema de aquecimento solar de água nas residências, porém, na mesma cidade, existe outro conjunto habitacional, o Orlando Quagliato, mas, por ser mais antigo, não possui o sistema de aquecimento solar, fator que motivou o presente estudo quanto à busca de determinantes para a viabilização e recomendação de características fundamentais para a otimização do aproveitamento de fonte de energia alternativa. O objetivo desta pesquisa foi aplicar procedimentos e técnicas de análise multivariada na avaliação do impacto econômico e na busca dos identificadores socioeconômicos fundamentais para a descrição de uma política de consumo racional de água e energia em conjuntos habitacionais com e sem o sistema de aquecimento solar de água. Os resultados permitiram concluir que diferenças entre os conjuntos habitacionais são observadas para os conjuntos de variáveis de consumo e socioeconômicos, e a positividade da associação entre estes conjuntos de variáveis foi verificada, não existindo um comportamento único de associação nos dois conjuntos habitacionais.

Palavras-chave: Análise multivariada de dados, consumo, socioeconômico, sistema de aquecimento solar de água, habitações de interesse social.

MULTIVARIATE PROCEDURES FOR USE IN WATER CONSUMPTION AND ELECTRICITY IN SOCIAL HOUSING WITH SOLAR HEATING SYSTEM. Botucatu, 2015. 97p. Tese (Doutorado em Agronomia/Energia na Agricultura) - Faculdade de Ciências Agronômicas, Universidade Estadual Paulista.

Author: MARCELO HIROSHI TUTIA

Adviser: CARLOS ROBERTO PADOVANI

SUMMARY

Nowadays, renewable energy sources are exploited and used for the development of society and stand out for being cleaner and less polluting than non-renewable energy sources. These renewable energy sources, solar energy is one of the most exploited, used in photovoltaic systems for power generation and solar water heating systems, the second use cited the best-known and can be seen in homes, businesses and industries. The solar heating system has among its objectives, to provide energy saving because it provides hot water without the need for electrical systems, such as electric shower, which is one of the devices that consume more energy. To provide economy of energy consumption, minimize environmental impacts, redirect investments in new sources of electricity, the Government of the State of São Paulo, together with the Development Company Housing and Urban State of São Paulo - CDHU, as housing policy, installed in homes of social interest built from 2007, the solar heating system water. In Ourinhos - SP, Helena Braz Vendramini housing project was awarded the solar heating system water in homes, but in the same city, there is another housing project, the Orlando Quagliato, but being older, does not have the solar heating system, a factor that motivated this study and the search for determinants for the viability and recommendation of key features to optimize the alternative energy source to use. The objective of this research was to apply procedures and techniques of multivariate analysis in evaluating the economic impact and the pursuit of fundamental socioeconomic identifiers for the description of a rational use of water policy and energy in housing with and without solar water heating system. The results showed that differences between the housing are observed for combinations of consumption and socioeconomic variables and the association of positivity between these sets of variables was verified, there is no single association behavior in the two housing estates.

Keywords: multivariate data analysis, consumption, socioeconomic, solar water heating system, social housing.

1 INTRODUÇÃO

O progresso da humanidade sempre está associado à utilização de fontes de energia; sem estas não haveria a constante expansão e desenvolvimento, tanto social como tecnológico, que podem ser percebidos por todos ao longo do tempo.

Sem uma matriz energética, qualquer que seja ela, as tarefas mais básicas seriam prejudicadas ou até mesmo, com grande chance, não poderiam ser realizadas. Para muitas atividades do cotidiano, é necessária alguma fonte de energia para que estas sejam executadas.

Em diversas partes do mundo são exploradas diferentes fontes de energia, tais como as renováveis e as não renováveis. Entre as renováveis, a energia solar vem se destacando por sua ampla disponibilidade, menor degradação da natureza em sua exploração, por não proporcionar a emissão de gases poluentes quando utilizada (menos poluente em relação a outras fontes), por contribuir na economia de outras fontes energéticas quando empregadas simultaneamente, pelo desenvolvimento de novas tecnologias que melhoram seu aproveitamento e aplicação, enfim, por características que fazem com que ela seja uma grande promessa entre as fontes renováveis que são aproveitadas na atualidade.

São várias as vantagens e os benefícios da energia solar. Além de ser responsável, de forma direta ou indireta, pela origem de outras fontes energéticas

renováveis ou não, pode ser considerada inesgotável, contribui para o meio ambiente; entre outros benefícios.

No Brasil, seu uso é percebido com maior frequência nos sistemas de aquecimento solar de água, porém, sistemas fotovoltaicos para geração de energia elétrica vêm sendo implantados em locais de difícil acesso, onde as redes elétricas não foram instaladas.

Acredita-se que, pelas vantagens citadas e a crescente conscientização para a preservação do meio ambiente em um futuro não muito distante, exista um maior número de consumidores usando esta fonte de energia renovável. Pode-se observar que, hoje, é maior a quantidade de residências com o sistema de aquecimento solar de água, sendo este o mais utilizado e conhecido no aproveitamento da energia solar; empregado com maior frequência em residências de classes média e alta por apresentar um custo relativamente alto em sua instalação, mas que, ao longo do tempo, torna-se recuperável financeiramente com a economia da energia elétrica.

A economia de energia elétrica com a utilização do sistema de aquecimento solar de água é percebida em diversas pesquisas. Para contribuir na diminuição de custos com eletricidade, minimizar impactos ambientais (menor degradação dos recursos ambientais) e proporcionar a sustentabilidade em residências de interesse social, o Governo do Estado de São Paulo, em conjunto com a Companhia de Desenvolvimento Habitacional e Urbano do Estado de São Paulo - CDHU, instalaram, nas residências construídas desde 2007, o sistema de aquecimento solar de água.

Além da economia de energia elétrica e por consequência a economia financeira para as famílias beneficiadas, há também a conscientização ambiental, já que utilizam o sistema de aquecimento solar de água - como já dito, fonte renovável de energia que proporciona menor degradação ao meio ambiente - para gerar conforto aos moradores da residência. As fornecedoras de energia elétrica e o governo também são favorecidos, porque podem distribuir a energia que seria fornecida nestas residências para outros locais e redirecionarem os investimentos que seriam aplicados para geração de novas fontes de energia elétrica.

Na cidade de Ourinhos, estado de São Paulo, o Conjunto Habitacional Helena Braz Vendramini recebeu a instalação do sistema de aquecimento solar de água por ser mais recente, enquanto que o Conjunto Habitacional Orlando Quagliato, por se mais antigo, não foi contemplado. Por apresentarem características

semelhantes, salvo o sistema de aquecimento solar, os dois conjuntos habitacionais foram utilizados nesta pesquisa.

Na literatura especializada ao tema, encontram-se vários indicativos para a redução do consumo de energia elétrica em locais com diferentes sistemas de aquecimento solar de água. Utilizando-se metodologias univariadas e ferramentas específicas para estruturas de variabilidade simples, conseguiu-se verificar a redução do consumo de energia elétrica, mostrando que o sistema de aquecimento solar de água com diferentes configurações de instalação e meios de utilização cumpre o seu objetivo. No entanto, as técnicas univariadas são simplistas, pois não consideram o consumo de energia elétrica associado a outras variáveis, mas existem situações nas quais devem ser consideradas a análise conjunta de variáveis, sendo uma nova proposta na obtenção de resultados. Coloca-se como hipótese a ser verificada que os consumos de energia elétrica e de água associados às variáveis socioeconômicas não devem ser vistos isoladamente, mas analisados conjuntamente para que forneçam resultados mais robustos e fidedignos; e possibilitem identificar quais variáveis e fatores podem interferir ou diferenciar o consumo.

Neste sentido, a análise multivariada é uma das ferramentas estatísticas que proporciona o estudo conjunto de variáveis e, consequentemente, fornece resultados mais precisos que podem ser imprescindíveis para melhorar as tomadas de decisão. Complementando, o interesse do estudo deve considerar que, para as políticas públicas e utilização racional de energia, tanto na zona rural como na urbana, são necessárias informações mais consistentes e abrangentes quanto à distribuição de recursos financeiros e aproveitamento das fontes de energia disponíveis, levando assim a uma otimização no direcionamento das decisões.

Na presente pesquisa, foram consideradas algumas técnicas de estatística multivariada, para se analisar o conjunto de variáveis. Algumas destas coletadas nos respectivos conjuntos habitacionais, outras fornecidas pelos responsáveis pela distribuição de energia elétrica e água do município. O objetivo foi o de avaliar e aprofundar conceitos sobre o impacto econômico, consumo racional de energia e água e características socioeconômicas em habitações construídas com ou sem o sistema de aquecimento solar; além de construir, por meio de técnicas multivariadas, alguns indicadores como subsídios de entendimento e melhoria da qualidade de informação e implementação de políticas públicas.

2 REVISÃO DE LITERATURA

2.1 Energia solar

Fontes de energia provenientes da natureza ou geradas pelo homem são de fundamental importância para o desenvolvimento dos seres vivos e da humanidade.

O homem usa, desde os primórdios, diferentes fontes de energia que proporcionaram a evolução da sociedade. Desenvolveram-se, ao longo do tempo, diferentes tecnologias para explorar e aproveitar os recursos naturais, transformando-os em energia que pode ser utilizada para geração de produtos, bens, novas fontes de energia, entre outros.

Segundo Cavalcanti (2008, p. 18),

> o desenvolvimento da humanidade, em geral, e a melhoria da qualidade de vida do ser humano estão diretamente ligados à utilização de energia. Assim, a energia é um elemento imprescindível na vida do homem, tendo em vista que todas as suas atividades, doméstica, agropecuária, agroindustrial ou industrial dependem do uso de energia.

As diversas formas de energia utilizadas pelo homem podem ser agrupadas em duas categorias, uma composta pelas fontes de energias renováveis e a outra pelas não renováveis. Independente do seu consumo, as fontes de energia renováveis são aquelas que existem em uma quantidade considerada inesgotável na natureza. Pode-se

dizer que este tipo de energia sempre irá atender ao consumo da humanidade. Já a energia não renovável pode ser definida como sendo aquela que existe em quantidades limitadas na natureza (CAVALCANTI, 2008).

Segundo Foster et al. (2010), algumas das principais fontes de energia renováveis são provenientes do sol, vento, mares e biomassa. A energia solar é referida como renovável e ou sustentável, pois está disponível enquanto existir o sol, cujas estimativas de vida são de aproximadamente quatro ou cinco milhões de anos. A energia eólica, biomassa e biogás são indiretamente provenientes da energia solar, pois são derivadas do aquecimento não uniforme da superfície terrestre e dos efeitos por ela produzidos. São muitas as vantagens da energia renovável: sustentabilidade (não será esgotada ou será necessário um longo tempo até a extinção de suas atividades); onipresença (encontrada em todos os lugares do mundo, em contraste com os combustíveis fósseis e minerais) e essencialmente não poluente e livre de carbono (ou apresentando quantidades de poluentes e carbono significativamente inferiores em comparação às energias não renováveis). As desvantagens das energias renováveis, contudo, também existem: variabilidade da fonte de origem (a origem da energia não é constante, podendo sofrer influência do clima- nem sempre haverá vento; poderá haver períodos nebulosos, sem a predominância do sol; entre outros fatores que podem interferir na geração destas fontes); baixa densidade e, geralmente, os equipamentos de conversão apresentam custo inicial de produção mais elevado. Para diferentes formas de energia renováveis, outras desvantagens ou problemas aparecem: poluição visual; odor a partir de biomassa e necessidade de terrenos grandes para a conversão solar. A energia solar é a fonte que sustenta a vida na Terra, pois as plantas, os animais e as pessoas a utilizam direta ou indiretamente, satisfazendo as necessidades de fontes limpas e abundantes. A luz solar está disponível facilmente, segura de tensões geopolíticas, não representando uma ameaça para o nosso ambiente e clima global, sendo a fonte mais livre de carbono do planeta. A Terra é essencialmente um grande coletor de energia solar, recebendo grandes quantidades que se manifestam de várias formas; em uma hora a energia recebida no planeta pelo Sol, $4{,}3 \times 10^{20}$ J, é maior que toda a energia consumida em um ano, $4{,}1 \times 10^{20}$ J.

Mesmo com desvantagens, as fontes de energia renováveis apresentam vantagens superiores, principalmente ao meio ambiente.

Fontes de energia renovável podem ser exploradas e utilizadas com maior frequência para que haja uma queda da utilização de fontes não renováveis,

contribuindo tanto para a natureza (diminuindo a emissão de poluentes e menor degradação do meio ambiente) quanto para a economia das sociedades (proporcionando maior economia na utilização de fontes de energia mais dispendiosas).

Cavalcanti (2008) salienta que a energia total fornecida pelo Sol ao nosso planeta, durante um ano, equivale a aproximadamente 10.000 vezes a energia total consumida pela população mundial nesse mesmo período; superando muito o total fornecido pelo petróleo ou gerado em hidrelétricas. Entre as fontes de energia renováveis, a energia solar representa um elevado potencial de utilização por sua grande disponibilidade e as inúmeras vantagens que se encontram associadas a ela. O Sol funciona como um imenso reator à fusão, gerando energia constantemente e lançando-a no nosso planeta. A combinação da busca por recursos naturais e o elevado custo de manutenção da energia convencional vem tornando cada vez mais viável o uso da tecnologia que permite aproveitar a energia solar.

Pelos relatos da literatura, e na própria natureza, pode ser observado que a energia proveniente do Sol dá origem a muitas outras fontes de energia, sendo ela uma das mais importantes e abundante, ainda pouco aproveitada pelo homem, mas com grande tendência de crescimento de sua utilização.

O uso da energia solar é uma alternativa viável e promissora, está presente em todos os lugares, abunda, não polui, dentre outras vantagens. Com o desenvolvimento de tecnologias, ela pode ser explorada de forma que possa substituir, com eficiência, as fontes energéticas já utilizadas.

“Da perspectiva puramente técnica, os sistemas de conversão de energia solar são potencialmente capazes de produzir a maior parte da futura demanda de energia de todo o planeta” (PALZ, 2002, p. 107).

Em longo prazo, muitos cientistas acreditam que a energia solar se tornará a mais importante, pois a luz do Sol está em toda parte, podendo ser aproveitada até mesmo em regiões temperadas ou continuamente cobertas por nuvens; e coletada por pequenos aparelhos, não havendo a dependência de centrais de fornecimento de energia (WALISIEWICZ, 2008).

Cavalcanti (2008) relata que, além das transformações da energia solar, que ocorrem de forma espontânea na natureza, ela também pode ser transformada em energia elétrica ou térmica, como o efeito fotovoltaico, aquecimento de água e a calefação de ambientes, utilizando-se sistemas desenvolvidos pelo homem. A captação e a

transformação da energia solar em outras formas de energia utiliza tecnologia bastante simples, eficiente e relativamente barata; e vem sendo utilizada com maior frequência e com sucesso considerável, não somente no Brasil, mas principalmente em países desenvolvidos, como os Estados Unidos e os europeus.

Em alguns locais do Brasil, onde a energia elétrica não está disponível ou é de difícil acesso, sistemas fotovoltaicos são utilizados para fornecer eletricidade, proporcionando conforto para as pessoas. Nos grandes centros, onde a energia elétrica é de fácil acesso, a energia solar é transformada em energia térmica por meio dos sistemas de aquecimento solar de água.

2.2 Sistema de aquecimento solar

Os sistemas de aquecimento solar de água são mais conhecidos e divulgados hoje em dia, seus benefícios são relevantes e a disponibilidade de diversas marcas e modelos no mercado faz com que estes sejam mais acessíveis às diferentes classes sociais.

Para Walisiewicz (2008), atualmente a energia solar é mais utilizada para o fornecimento de água aquecida em residências e piscinas. Só nos Estados Unidos, mais de dois milhões de casas são equipadas com coletores de placas planas que usam a luz solar para aquecer a água com uma eficiência de aproximadamente 50%.

A conversão de energia solar para térmica (aquecimento de água) é a mais conhecida e sua utilização vem crescendo nos últimos anos no Brasil, contribuindo para a diminuição do consumo de energia elétrica, segundo Naspolini et al. (2008), Silva et al. (2011). Porém, em todos estes trabalhos o consumo de energia é uma variável analisada independente de outros fatores, mas na presente pesquisa o consumo de energia elétrica será visto com uma variável dependente, associada a outras variáveis que podem contribuir ou não para a sua redução.

O aquecimento solar de água é considerado um sistema caro por muitos, porém, tem ser tornado mais acessível com a divulgação e o incentivo do governo, e vem se difundindo nas residências - principalmente nas de classe baixa e média, em que se percebe a sua crescente utilização -, estabelecimentos comerciais e empresas.

“O custo inicial de instalação de coletores solares para aquecimento de água caiu consistentemente nos últimos vinte anos no Brasil, mas seu valor continua

alto se comparado com seu principal concorrente, os chuveiros" (RODRIGUES; MATAJS, 2004, p. 53).

> A tecnologia termossolar é uma possibilidade factível e competitiva de substituir a parte da eletricidade que é usada para aquecimento nos setores residencial, comercial, industrial e de serviços e, consequentemente, de minorar a preocupação ambiental à geração de energia elétrica. Por substituir hidroeletricidade e combustíveis fósseis, cada instalação de aquecedor solar reduz de uma vez e para sempre o dano ambiental associado a fontes convencionais, não produz emissão de gases tóxicos, que contribuem para poluição urbana, e não deixa lixo radiativo como uma herança perigosa para as gerações futuras (RODRIGUES; MATAJS, 2004, p. 40).

Varella, Fantinelli e Pereira (2004) relatam que a implantação de coletores solares de baixo custo, para aquecimento de água, contribuem para com a redução do consumo de energia elétrica, minimizando a dependência ao uso do chuveiro elétrico.

A grande vantagem do sistema de aquecimento solar de água é a redução do consumo de energia elétrica em horários que apresentam grande demanda por energia elétrica, além do conforto em sua utilização.

Os benefícios proporcionados pela agregação de energia solar térmica, aliado à queda nos custos dos sistemas termossolares, ocorridos nos últimos anos, tornam a energia solar para aquecimento de água uma tendência a ser adotada em praticamente todas as camadas sociais (NASPOLINI et al., 2008).

Os sistemas de aquecedores solares de água são uma tendência, como pode ser observado nas considerações de diferentes autores. Os custos para implantação decresceram ao longo do tempo, tornando-se um item considerado nas diferentes classes sociais. Suas vantagens são percebidas tanto para o conforto dos usuários finais como para a economia destes e das concessionárias responsáveis pela distribuição de energia e força.

O sistema de aquecimento de água por energia solar permite alcançar temperaturas ideais para banho, tornando assim o uso do aquecimento elétrico apenas esporádico (BASSO, 2010).

Segundo Fantinelli (2006),

> a moradia construída e mantida com consumo energético mínimo não significa necessariamente o sacrifício do conforto dos usuários. A incorporação de soluções tecnológicas eficientes e adequadas para a redução do consumo de energia significa planejar com visão estratégica, tanto do ponto de vista econômico quanto social e ambiental. A análise conjunta dos fatores de eficiência e dos fatores econômicos passa a determinar quais os parâmetros desejáveis e compatíveis para uma determinada realidade social. A busca de soluções adequadas para o aquecimento doméstico de água com o uso da energia solar encontra respaldo quando se analisam as condições socioeconômicas de uma parcela da população brasileira que se situa na faixa de pobreza extrema. Essa realidade evidencia a má distribuição da renda e a exclusão social.

A utilização do sistema de aquecimento solar de água por moradores de baixa renda pode oferecer um duplo benefício; reduz o consumo de energia elétrica, podendo proporcionar consumo abaixo de 220 kWh/mês (limite máximo de consumo imposto pelo regime de tarifa subsidiada) e, com isto, consegue tarifas mais baixas. Em média, mais de 73% das residências utilizam chuveiros de resistência elétrica, em algumas regiões de clima mais temperado, no sul do país, onde a maioria da população está concentrada, chuveiros elétricos estão presentes em mais de 90% das construções residenciais (NASPOLINI; MILITÃO; RÜTHER, 2010).

Segundo Nasponili e Rüther (2011), da perspectiva do fornecedor de energia, o uso de chuveiro elétrico, especialmente por consumidores de baixa renda com tarifas subsidiadas, é considerado inadequado. Esse dispositivo é um investimento de baixo custo para o usuário final com consequência de alto custo às concessionárias de distribuição de energia e força, pois, a concessionária poderá deslocar essa energia, que seria fornecida para o consumidor residencial de baixa renda, para outras classes de consumo a um preço consideravelmente maior.

O uso de sistemas de aquecimento solar de água em populações de baixa renda pode ajudar a evitar custos de expansão do sistema e, ao mesmo tempo, permitir uma melhor gestão de cargas, aumentando a confiabilidade do sistema, especialmente em horários de pico. A utilização do sistema pode trazer benefício muito maior para a empresa concessionária do que para o consumidor residencial, e deve, portanto, ser estimulada pelo setor de energia devido ao potencial de demanda que pode oferecer (NASPOLINI; RÜTHER, 2011).

Para a CDHU e o Governo do Estado de São Paulo, a instalação do sistema de aquecimento solar de água em habitações de interesse social visa a diminuição do consumo de energia elétrica, a sustentabilidade das habitações, a redução dos

investimentos com outras fontes de geração de energia elétrica e, entre outros, contribuir para a preservação do meio ambiente (hidroelétricas – impedindo inundações e o impacto ambiental, termoelétricas – impedindo a queima de combustíveis fosseis, madeira e a redução de lixo radioativo). Já para as distribuidoras de energia elétrica, o sistema tem o intuito de reduzir o consumo de energia elétrica em horários de pico, pois, em residências de interesse social, o chuveiro elétrico é o equipamento de maior consumo, sistema de distribuição mais confiável. Desta forma, podem distribuir a energia que seria consumida pelas habitações de interesse social para outros fins.

Em habitações de interesse social, a instalação de sistemas de aquecimento solar de água de baixo custo é incentivada tanto pelo Governo Federal como Estadual, proporcionando aos usuários e às companhias de distribuição de energia as vantagens relatadas.

2.3 Habitações de interesse social

O termo habitações de interesse social, utilizado nesta pesquisa, indica as habitações destinadas à população de baixa renda; o mesmo nome é empregado em trabalhos de diferentes autores. Outras definições de habitações são fornecidas por Abiko (1995), habitação popular (moradia voltada para população de baixa renda); habitação de baixo custo (utilizado para designar moradia barata, sem que isto signifique necessariamente habitação para população de baixa renda) e/ou habitação para população de baixa renda (este termo necessita da renda máxima das famílias e indivíduos situados nesta faixa de atendimento social); esta última muito semelhante à utilizada neste estudo.

Promover condições dignas de moradia para a população de baixo poder aquisitivo consiste na principal atribuição da Secretaria de Estado da Habitação e de seus dois braços operacionais, a Companhia de Desenvolvimento Habitacional e Urbano – CDHU, e a Agência Paulista de Habitação Social (Casa Paulista), cujo foco de atuação está dirigido ao público-alvo situado na faixa de um a dez salários mínimos, com atendimento prioritário até cinco salários (CDHU, 2015).

Segundo Abiko e Coelho (2004), o mutirão, também conhecido como sistema de ajuda mútua, é a alternativa habitacional baseada no esforço coletivo e organizado da comunidade para a construção das moradias. Tem o objetivo de diminuir os

custos, melhorar a qualidade do conjunto habitacional e das unidades habitacionais e identificar o usuário com o produto de seu trabalho.

> Mais recentemente, com a promulgação do Estatuto das Cidades (Lei Federal Nº 10.257, de 10 de julho de 2001), que regulamenta a Constituição, foi ratificada a função social do solo urbano e a habitação assume efetivamente o caráter de direito básico da população. As políticas e estratégias habitacionais para a população de baixa renda passam a ser legalmente submetidas ao interesse da sociedade, sobretudo em nível local nos municípios, onde se dão os impactos de sua implantação (LARCHER, 2005).

Investimentos e incentivos do governo federal, estadual e municipal para construção de novas habitacionais de interesse social são notados. No estado de São Paulo, como em todo território nacional, existe um déficit habitacional que pode ser minimizado com a construção dessas habitações.

2.4 Técnicas univariadas

Diversos trabalhos, considerando o emprego do sistema de aquecimento solar de água em habitações de interesse social, foram publicados. Os estudados nesta pesquisa apresentam, em sua metodologia, para análise dos dados, abordagens considerando estatística descritiva, análise exploratória de dados, matemática financeira, modelos matemáticos básicos e triviais, e estudos de caso com entrevistas e relatos coletados.

Naspolini et al. (2008) realizam sua pesquisa em habitações de baixa renda com sistemas de aquecimento solar de água para banho na cidade de Florianópolis – SC. Constatam que a economia anual de energia elétrica proporcionada pela agregação da energia solar térmica ao banho é significativa por meio de critérios técnicos univariados, os quais utilizam a energia média mensal consumida, kWh, e em outro critério a demanda média na ponta, kW.

No trabalho realizado por Silva et. al (2011), adotam uma abordagem qualitativa e quantitativa, abrangendo levantamento documental e de campo. O levantamento de campo utilizou um questionário organizado em quatro tópicos: caracterização socioeconômica, adequação aos hábitos e costumes da população, economia financeira e conscientização ambiental. Os dados coletados foram analisados

estatisticamente pelo método de Análise Exploratória de Dados e pelo método de análise de conteúdo. O trabalho foi realizado no Programa Lares Habitação Popular da COHAB-MG em Itatiaiuçu, verificando que os principais impactos gerados com a incorporação do aquecedor solar na habitação de interesse social foram: a economia financeira e o aumento do conforto (sob aspectos como o aumento do tempo de banho).

Diakoulaki et al. (2001) analisam o custo beneficio, para o sistema de aquecimento solar de água considerando vários cenários. Com a utilização de técnicas de matemática financeira concluem que o sistema é superior em relação as tecnologias convencionais em locais nos quais utilizam eletricidade ou diesel.

O impacto do sistema de aquecimento solar de água nas perdas do sistema de energia foi estudado por Ijumba et al. (2009), com um programa que modela a radiação solar e outro que simula várias performances de um sistema de energia elétrica ao longo do tempo. A pesquisa deixou claro que o sistema de aquecimento solar de água tem o potencial de reduzir substancialmente tanto a demanda quanto as perdas na transmissão em um sistema de energia.

Nestes trabalhos destacados na revisão e em outros estudados para o desenvolvimento do projeto como bibliografia complementar para o entendimento do sistema de aquecimento solar de água em habitações de interesse social, observa-se que o tratamento estatístico oferecido aos dados dispensam o uso de técnicas de análise multivariada com rara exceções as quais ficam restritas ao consumo de energia. Todos os procedimentos, praticamente, envolvem o consumo de energia elétrica, demanda de energia e outras variáveis consideradas separadamente. Portanto, a pesquisa proposta nesta tese abrange uma nova dinâmica para a análise dos resultados, que levará em consideração a aplicação de técnicas de análise multivariada para o estudo dos consumos de energia elétrica e água, em habitações de interesse social, conjuntamente com outras variáveis abrangendo características socioeconômicas.

2.5 Técnicas multivariadas

A Análise Multivariada refere-se a um conjunto de métodos estatísticos que torna possível a análise simultânea de múltiplas medidas realizadas em cada indivíduo, objeto ou fenômeno observado. São procedimentos mais complexos do que os utilizados na análise univariada ou bivariada, mas, em muitos casos, não passam de

adaptações dessas análises em situações com mais de duas variáveis (CORRAR; PAULO; DIAS FILHO, 2009).

Hair Jr et al. (2009) relata que,

> para ser considerada verdadeiramente multivariada, todas as variáveis devem ser aleatórias e inter-relacionadas de tal maneira que seus diferentes efeitos não podem ser significativamente interpretados em separado. Alguns autores estabelecem que o objetivo da análise multivariada é medir, explicar e prever o grau de relação entre variáveis estatísticas.

A técnica tem interesse em dados que consistem em um conjunto de medições de um número de indivíduos ou objetos, cujas medições feitas sobre um único indivíduo podem ser reunidas em um vetor coluna (ANDERSON, 2003).

Os procedimentos multivariados permitem às organizações gerarem conhecimentos envolvendo toda a estrutura de variação dos dados, melhorando as tomadas de decisões, pois simultaneamente analisam múltiplas medidas sobre indivíduos ou objetos, assim, qualquer análise simultânea de mais do que duas variáveis respostas pode ser considerada, a princípio, como multivariada (HAIR JR et al., 2009).

Segundo Corrar, Paulo e Dias Filho (2009),

> desse modo, a análise multivariada pode ser definida como o conjunto de métodos que permitem a análise simultânea dos dados recolhidos para um ou mais conjuntos de indivíduos (populações ou amostras) caracterizados por mais de duas variáveis correlacionadas entre si, sendo que as variáveis podem ser quantitativas ou qualitativas. Somente as técnicas de estatística multivariada permitem que se explore a *performance* conjunta das variáveis e se determine a influência ou importância de cada uma, estando as demais presentes.

Atualmente, muita informação é coletada e armazenada em bancos de dados e está disponível para a realização de análises com fins de melhoria na tomada de decisões. Parte desta informação pode ser analisada e compreendida com estatística simples, mas uma grande proporção necessita de técnicas estatísticas multivariadas mais complexas para converter estes dados em conhecimento (HAIR JR et al., 2009).

A complexidade de muitos fenômenos requer uma investigação de observações coletadas referente a várias diferentes variáveis. Quando medições

simultâneas de muitas variáveis são coletas, a metodologia estatística mais adequada é a análise multivariada (JOHNSON; WICHERN, 2007).

Trabalhos como os de Cabral (2004) que usa a análise multivariada para avaliar os fatores de risco anteparto para ocorrência de cesárea, de Rigão (2009) o qual trabalha com a correlação canônica entre caracteres de tubérculos para seleção precoce de clones de batata, Barbosa (2010) estuda a associação entre qualidade da carne e características quantitativas de suínos por meio da correlação canônica, de Jaumot (2011) que usa análise multivariada nos estudos de equilíbrio químico, em Wisser (2011) utiliza a análise multivariada no estudo da resistência das doenças do milho, de Protásio (2012) o qual usa a análise de correlação canônica entre características da madeira e do carvão vegetal de *Eucalyptus*, Kavradin (2013) aplica análise multivariada para estudar a esperança em pessoas com câncer na Turquia, Nazaré (2013) estuda o questionário de confiança parental com o uso da análise fatorial confirmatória numa amostra comunitária de casais, Jangarelli (2014) emprega a abordagem multivariada para endogamia e valor fenotípico utilizando diferentes estratégias de cruzamento, em Menezes (2014) é verificada a correlação entre uso da terra e qualidade da água subterrânea, e Gok (2015) por meio de análise multivariada estuda a diferenciação das amostras de mel da Anatólia de diferentes origens botânicas.

Existem diversos trabalhos em diferentes áreas do conhecimento que utilizam técnicas de análise multivariada, entretanto houve dificuldade de encontrar estudos que envolvessem as técnicas multivariadas aplicadas no estudo de sistemas de aquecimento solar de água.

3 MATERIAIS E MÉTODOS

3.1 Materiais

Foram considerados dois bairros da cidade de Ourinhos, interior do estado de São Paulo, localizada a aproximadamente 370 km ao sudoeste da capital, situada na latitude -22^{o} 58’ 44”, longitude -49^{o} 52’ 14” e altitude de 483 metros. Segundo a classificação climática de Koppen, o clima é do tipo Am, que caracteriza o clima tropical chuvoso com inverno seco. O município pertence à microrregião de mesmo nome, mesorregião de Assis e região administrativa de Marília. Segundo censo demográfico de 2010, possui 103.035 habitantes, área da unidade territorial de 296,269 km^2, densidade demográfica de 347,78 habitantes/km^2, segundo IBGE (2012).

De acordo com SEADE (2013), no ano de 2000, o índice de desenvolvimento humano – IDH- do município foi de 0,813; valor semelhante ao do estado, com renda *per capita* (em Reais correntes) de R$ 724,40 e valor do rendimento médio mensal das pessoas responsáveis pelos domicílios particulares permanentes de R$ 1.553,64 para o ano de 2010.

O estado de São Paulo, constituído por 645 munícipios, no ano de 2010, apresentava uma população de 41.262.199 habitantes, área de 248.222,801 km^2, com densidade demográfica de 166,25 habitantes/km^2, IBGE (2013). No ano de 2010, a renda *per capita* (em Reais correntes) era de R$ 853,75, sendo de R$ 1870,49 o rendimento

médio mensal das pessoas responsáveis pelos domicílios particulares permanentes e índice de desenvolvimento humano, em 2000, de 0,814, segundo SEADE (2013).

O presente estudo envolve dados coletados nos Conjuntos Habitacionais (bairros) Orlando Quagliato e Helena Braz Vendramini; o primeiro com 523 residências, área total do conjunto de 105.494,04 m^2 e altitude de aproximadamente 421 metros. O segundo com 596 residências, área total do conjunto de 125.021,95 m^2 e altitude de aproximadamente 432 metros, segundo informações da Prefeitura de Ourinhos. As residências destes conjuntos habitacionais foram construídas no sistema de mutirão pelas famílias contempladas no sorteio realizado pela Companhia de Desenvolvimento Habitacional e Urbano do Estado de São Paulo – CDHU. São habitações de interesse social, consideradas de pequeno porte, com cinco cômodos (dois quartos, sala, cozinha e banheiro (contabilizado no total de cômodos)) e apresentavam área inicial construída de 36,47m^2 (Orlando Quagliato) e 46,12 m^2 (Helena Braz Vendramini).

Somente famílias que satisfizeram características pré-estabelecidas puderam fazer a inscrição para participar do referido sorteio, sendo que alguns dos pré-requisitos necessários para a inscrição foram residir na cidade de Ourinhos por mais de três anos, apresentar renda nas faixas salariais determinadas no edital, não possuir imóvel e não possuir financiamento habitacional.

Para o conjunto habitacional Orlando Quagliato, as inscrições e os sorteios foram realizados nos anos de 1997 a 1998. Para famílias inscritas com renda de até um salário mínimo, foram disponibilizadas para o sorteio 50% das residências, para famílias com renda entre um a dois salários mínimos, 25% das residências e para famílias com renda entre dois a três salários mínimos, 15% das residências. Famílias com rendas entre três a cinco salários mínimos, 5% das residências e com renda acima de cinco salários mínimos, 5% das residências.

No conjunto habitacional Helena Braz Vendramini, as inscrições e os sorteios foram realizados nos anos de 2003 a 2004. A proporção de residências sorteadas para as famílias inscritas nas determinadas faixas de renda é semelhante à citada anteriormente no conjunto habitacional Orlando Quagliato; além da existência de cotas para deficientes, aposentados e policiais militares em cada estrato de renda.

Atualmente, a característica das residências quanto à renda familiar e outras informações consideradas no ato da inscrição não são corroboradas caso sejam realizadas pesquisas. São vários os motivos destas mudanças, segundo funcionários da

prefeitura que acompanharam o processo de inscrição e no momento monitoram as famílias residentes nos conjuntos habitacionais, destacando-se tempo desde as inscrições das famílias, residências que voltaram novamente para sorteio por motivos de inadimplência, casos de venda da residência, aluguel e outras situações que não deveriam ocorrer de acordo com o contrato entre o mutuário e a CDHU.

Com a intenção de proporcionar a economia de energia elétrica, a CDHU e o Governo do estado de São Paulo instalaram, nas habitações sociais do Conjunto Habitacional Helena Braz Vendramini, coletores solares e reservatórios térmicos (sistema de aquecimento solar de água). Não há o sistema instalado nas residências do Conjunto Habitacional Orlando Quagliato, por terem sido construídas em período anterior a implantação deste benefício pela companhia.

Nesta pesquisa, os Conjuntos Habitacionais Orlando Quagliato e Helena Braz Vendramini serão denominados Conjuntos Habitacionais Sem Sistema e Com Sistema, respectivamente.

Segundo Soletrol (2013), o sistema de aquecimento solar de água instalado nas residências é o Aquecedor Solar Compacto Solarmax, composto por um reservatório térmico Solarmax de 200 litros, etiquetado pelo Instituto Nacional de Metrologia, Qualidade e Tecnologia - INMETRO e um coletor solar Soletrol modelo Max Alumínio com classificação A do INMETRO de 1,6 m^2 ou 2,0 m^2 , dependendo do modelo fornecido. Possui um reservatório de água fria para abastecimento direto da rede pública e não necessita estar interligado à caixa d'água da edificação. Esse reservatório é normalmente fornecido com 25 litros e, opcionalmente, pode ser fornecido com capacidade para 80 litros. Apresenta todas as interligações em mangueiras flexíveis de borracha EPDM para alta temperatura, que facilitam a instalação.

Este sistema de aquecimento solar de água não possui no reservatório térmico resistência elétrica para auxiliar no aquecimento da água quando o dia está nublado ou em dias de chuva; ou quando a utilização da água quente é excessiva. Quando se dá a ocorrência destas situações, é necessário que seja utilizado o chuveiro elétrico para complementar a temperatura da água. O equipamento não dispõe da resistência elétrica no reservatório térmico, pois o dispositivo aumentaria o consumo de energia elétrica e o custo do sistema, não sendo esta a intenção da instalação dos aquecedores.

O aquecedor solar tem capacidade limitada de água quente, o uso deve ser restrito ao banho de quatro ou cinco pessoas para evitar a falta de água, de acordo com as especificações da empresa fabricante.

Nas residências dos bairros mencionados anteriormente, com e sem o sistema de aquecimento solar de água, são observadas variáveis sociais (socioeconômicas) de consumo de energia e água, quantidade de determinados itens domésticos que consomem energia, características da residência e de seus moradores, presença de lâmpadas fluorescentes e de aquecimento solar de água; variáveis que serão estudadas conjuntamente para análise e comparação dos dois bairros.

Considerando o universo da pesquisa, composto por 1119 residências, nos dois bairros, foi calculado o tamanho da amostra em 287 residências (unidade amostral) considerando um nível de confiança de 95%, erro amostral de cinco pontos percentuais (5%) e probabilidade casual (50%) da residência sorteada participar da pesquisa (BOLFARINE; BUSSAB, 2005). O número de residências para a amostragem por bairro foi determinado em quantidades proporcionais ao número de residências existentes, indicando 47% das incluídas na amostra provenientes do Conjunto Habitacional Sem Sistema e 53% do Conjunto Habitacional Com Sistema.

Como critério de inclusão das residências na amostra, foi utilizada a técnica de amostragem sistemática com todos os logradouros dos conjuntos habitacionais percorridos, nos quais o lado do logradouro (direito ou esquerdo do ponto onde se encontrava o pesquisador) a ser inicialmente percorrido e a primeira residência para entrevista foram selecionados casualmente no momento. Após a seleção da primeira residência, as demais unidades amostrais foram sistematizadas de quatro em quatro domicílios, aproximadamente, ajustando-se em algumas situações, nas quais houve dificuldade de encontrar moradores nos endereços selecionados ou quando não estavam dispostos a responder o questionário. Para estabelecer o salto de quatro em quatro domicílios, considerou-se o seguinte cálculo:

$$Salto\ para\ seleção\ da\ amostra = \frac{Tamanho\ do\ Universo\ da\ Pesquisa}{Tamanho\ da\ Amostra}$$

Foram utilizadas duas pesquisadoras para coleta de dados, cada uma responsável por um conjunto habitacional, que foram treinadas em como proceder na seleção da unidade amostral, abordagem dos respondentes e aplicação do instrumento de coleta de dados. Para facilitar a aceitação e aplicação do instrumento, as pesquisadoras eram moradoras e conhecidas pela comunidade de cada conjunto habitacional. O período de coleta foi de aproximadamente um mês, com início em primeiro de abril de 2013.

As pesquisadoras foram orientadas a entrar em contato com o morador responsável pela residência selecionada para constituir a amostra, explicar o motivo do levantamento de dados e verificar a possibilidade da participação da residência na pesquisa. Caso não houvesse concordância em participar da pesquisa, o pesquisador selecionaria, de maneira casual, qual residência, da direita ou da esquerda do local que se encontrava, iria ser abordada para iniciar novamente o procedimento de contato.

Como instrumento para coleta de dados foi utilizado um questionário desenvolvido (Apêndice 1) com base em outras duas ferramentas de levantamento de dados, a Ficha A (Anexo 1) e o Critério de Classificação Brasil (Anexo 2). A Ficha A é utilizada pelas Secretarias Municipais de Saúde para realizar o acompanhamento dos residentes, suas respectivas doenças, características do domicílio visitado, condições sanitárias, entre outras variáveis. Os dados são coletados pelas agentes de saúde para alimentar o Sistema de Informação de Atenção Básica (SIAB) que irá fornecer informações para os gestores tomarem decisões sobre a situação de saúde na comunidade.

O Critério de Classificação Brasil é um instrumento de segmentação econômica que utiliza o levantamento de características domiciliares no qual são verificadas a presença e a quantidade de alguns itens domiciliares de conforto e o grau de escolaridade do chefe da família, fornecendo a classificação econômica (ABEP, 2012).

O questionário obtido com a mescla das duas ferramentas de coleta e o termo de consentimento livre e esclarecido, que informa sobre o objetivo da pesquisa e se apresenta em duas vias que devem ser assinadas (ficando uma para o respondente e outra para o responsável pela pesquisa), foi submetido ao Comitê de Ética em Pesquisa da Universidade Estadual Paulista – Faculdade de Medicina de Botucatu (Protocolo Nº: 4427-2013) e impetrou aprovação para ser utilizado na obtenção de dados nesta tese.

O instrumento de coleta de dados aplicado na amostra de residências obteve informações socioeconômicas de cada unidade amostral. Para o

desenvolvimento desta pesquisa, foi necessário também a utilização de informações sobre o consumo de água e o de energia elétrica de cada respectiva unidade. Estas informações foram disponibilizadas pela Superintendência de Água e Esgoto – SAE e a Companhia Paulista de Luz e Força – CPFL, respectivamente.

3.1.1 Variáveis de consumo de água

No banco de dados fornecido pela SAE, são apresentadas as variáveis: endereço, complemento e consumo de água (consumo de janeiro de 2008 a dezembro de 2011, em m^3). Para entendimento destas tem-se:

- Endereço: constam informações sobre o nome da rua, avenida, travessa, entre outras, número da residência e complemento.
- Consumo mensal: consumo de água em metros cúbicos – m^3.

Por meio do consumo mensal de água foi possível gerar a seguinte variável que será utilizada na pesquisa:

- (X_1) SAE: média do consumo mensal de água por residência.

3.1.2 Variáveis de consumo de energia elétrica

Pelo banco de dados fornecido pela CPFL, são apresentadas as variáveis: endereço, tensão, classe, valor de consumo de maio de 2010 a junho de 2011e consumo de energia no mesmo período (kW). Para entendimento destas tem-se:

- Endereço: constam informações sobre o nome da rua, avenida, travessa, entre outras, número da residência e complemento.
- Tensão: tensão fornecida para as unidades de consumo, baixa tensão ou alta tensão.
- Classe: classe de consumo, residencial, comercial, outros.
- Valor de consumo: valor do consumo no referido mês, em Reais.
- Consumo mensal: consumo mensal de energia em quilowatts – kW.

Com os valores do consumo mensal de energia a seguinte variável foi gerada:

- (X_2) CPFL: média do consumo mensal de energia por residência.

Há incompatibilidade de períodos nos dados fornecidos pelas distribuidoras de energia elétrica e água, devido à disponibilidade de dados consolidados e questões estratégicas.

3.1.3 Variáveis socioeconômicas

O banco de dados, com a tabulação dos questionários aplicados nas unidades amostrais dos respectivos conjuntos habitacionais, apresenta as seguintes variáveis: endereço, data de entrada na residência, número de moradores, número de cômodos, utilização de lâmpadas fluorescentes, utilização de sistemas de aquecimento solar de água, condição de ocupação, data de nascimento, idade, sexo, escolaridade, ocupação, renda (para as variáveis data de nascimento até renda são disponibilizadas informações de cada morador), quantidade de itens elétricos e eletrônicos na residência (televisão, rádio, entre outros), quantidade de banheiros, empregada mensalista, grau de instrução do chefe de família, entre outras variáveis. Para melhor esclarecimento destas, têm-se as seguintes descrições:

- Endereço: constam informações sobre o nome do logradouro, número da residência, complemento e bairro, para identificação da unidade amostral.
- Data de entrada na residência: dia, mês e ano da entrada na residência para posterior relação com os dados de consumo de energia e água.
- Número de moradores na residência: quantidade de moradores na residência.
- Número de cômodos: quantidade de cômodos e sua influência no consumo de recursos (energia e água).
- Utilização de lâmpadas fluorescentes em grande parte dos cômodos.
- Possui aquecimento solar de água: averiguar se possuem e usam o sistema de aquecimento solar de água na residência.
- Condição de moradia: dados sobre a condição da ocupação da residência, fornece informação se a moradia é própria, alugada ou cedida.
- Data de nascimento: data de nascimento de cada morador da residência.
- Idade: em anos.
- Sexo: sexo de cada morador da residência.
- Escolaridade: grau de instrução de cada morador da residência.

- Ocupação: profissão de cada morador da residência.
- Renda: renda mensal aproximada de cada morador da residência.
- Itens e Quantidade: coleta de informações sobre itens eletroeletrônicos e suas respectivas quantidades, para definição da classe social, e também se a presença e a quantidade destes influenciam no consumo de energia.
- Grau de instrução do chefe de família: a informação do grau de instrução do chefe de família foi utilizada para classificação da classe social.

As variáveis citadas anteriormente foram analisadas por meio da análise exploratória de dados e deram origem a alguns domínios e novas variáveis foram estabelecidas. Para a compreensão, são descritas a seguir:

- NQUEST: número associado a cada questionário aplicado nas unidades amostrais, indicando informações sobre o nome, logradouro, número e bairro. Identifica cada unidade amostral.
- (X_3) NMOR/NCOM: índice de ocupação da unidade amostral, número médio de moradores por cômodo- quanto maior o valor da variável, maior a quantidade de moradores na unidade amostral; e quanto menor o valor da variável, menor a quantidade de moradores na unidade amostral.
- (X_4) REND/NMOR: indicador de renda *per capita* da unidade amostral, renda média por morador.
- (X_5) IDADE: idade média dos moradores da residência.
- (X_6) ISEXO: indicador da variação da quantidade de sexo por unidade amostral, [(FEM - MASC)/(NMOR)], variável que assume valores de -1 a 1; quanto mais próximo de -1, maior de quantidade de moradores do sexo masculino na unidade amostral; e quanto mais próximo de 1, maior quantidade de moradores do sexo feminino na unidade amostral. Caso o indicador atinja a unidade negativa, a residência é composta apenas de ocupantes do sexo masculino e, se atingir a unidade positiva, apenas ocupantes do sexo feminino. Se o valor do indicador apresentar a nulidade, o número de moradores masculinos iguala-se ao feminino.
- (X_7) ESTUDO: média dos anos de estudo na residência.
- (X_8) IDEPEND: indicador de idade dependente na residência, número de moradores considerados inativos (0 a 14 anos e 65 anos e mais de idade) sobre o total de moradores da residência, [(Nº DE MORADORES DE 0 A 14 ANOS) + (Nº

DE MORADORES COM IDADE IGUAL E SUPERIOR A 65 ANOS)]/[NMOR], variável que assume valores de 0 a 1; quanto mais próximo de 0, menor a quantidade de moradores considerados inativos; e quanto mais próximo de 1, maior a quantidade de moradores considerados inativos na unidade amostral. Se o valor indicar a unidade, a residência é composta somente por inativos e, se zero, ausência total de inativos.

- (X_9) IMP: indicador de maior permanência na residência; número de pessoas que permanecem maior tempo na residência (do lar, estudante, aposentado, desempregado) sobre o total de moradores da residência, [Nº DE MORADORES QUE PERMANECEM MAIOR TEMPO NA RESIDÊNCIA]/[NMOR], variável que assume valores de 0 a 1; quanto mais próximo de 0, menor a quantidade de moradores que permanecem maior tempo na unidade amostral; e quanto mais próximo de 1, maior a quantidade de moradores que permanecem maior tempo na unidade amostral. A ocorrência do valor zero indica que não há moradores que permaneçam maior tempo na residência e, se for um,todos os moradores permanecem maior tempo na residência.
- (X_{10}) QUANTELE: quantidade de itens na residência que consomem energia (todos os itens, sendo os que consomem menor ou maior quantidade de energia elétrica).
- (X_{11}) ICA: indicador de equipamentos que consomem maior quantidade de energia elétrica; número de equipamentos que consomem mais energia (geladeira, freezer, chuveiro, fogão elétrico, forno elétrico, churrasqueira elétrica, panela elétrica) sobre a quantidade de equipamentos que consomem energia, [Nº DE EQUIPAMENTOS QUE CONSOMEM MAIOR QUANTIDADE DE ENERGIA ELÉTRICA]/[QUANTELE], variável que assume valores de 0 a 1. Quando o indicador receber o valor zero, indica ausência de equipamentos que consomem mais energia; porém, se o valor for a unidade, todos os equipamentos eletrônicos da residência são de maiores consumo de energia. Os equipamentos que consomem maior quantidade de energia elétrica foram definidos com base no Programa Nacional de Conservação de Energia Elétrica (PROCEL, 2014).

Com os dados coletados na amostra de residências nos dois bairros e os fornecidos pela CPFL e SAE foi construído um novo banco de dados para possibilitar o desenvolvimento e realização desta pesquisa. Neste novo banco de dados, são

consideradas 11 variáveis, como descritas anteriormente, que foram utilizadas na análise multivariada de dados.

3.2 Métodos

Os dados coletados e armazenados em planilha computacional foram submetidos a técnicas de análises exploratórias e de estatística multivariada, como mencionado, para a geração de informações, comparações entre os bairros e construção de indicadores que tenham em vista melhorar e contribuir para futuras políticas públicas e, ainda, a utilização desses resultados no uso racional de energia elétrica e água.

A utilização de técnicas específicas e mais apropriadas para gerar estas informações tornam-se de fundamental importância para contribuir de maneira efetiva e relevante na construção de indicadores multivariados para entendimento da situação. O uso de procedimentos que envolvem técnicas de estatística multivariada, por considerarem simultaneamente as informações, certamente contribuiu para melhorar as respostas que são fornecidas e, consequentemente, possibilitam obter parâmetros mais fidedignos e relevantes para decisões e análises dos responsáveis.

3.2.1 Análise exploratória

A análise exploratória de dados tem por objetivo procurar interpretações sobre as respostas encontradas nas características da amostra selecionada, explicando melhor os resultados e obtendo descrições detalhadas.

Segundo Tukey (1993), a Análise Exploratória de Dados não se trata de técnicas novas, mas técnicas mais organizadas, que utilizam dispositivos gráficos melhores e não familiares. A novidade está em:

- Organização de um conjunto de ferramentas e abordagens.
- Técnicas novas ou desconhecidas.
- Destaque em visualizar os resultados por meio de gráficos ou numericamente.
- Procedimentos legitimados, mostrando que eles foram utilizados e trabalhados e não são derivados de um modelo.

- Muitas técnicas podem ser utilizadas sem qualquer menção de probabilidade.
- Observar situações que haviam de fato acontecido por acaso ao lado de situações não casuais.
- Não exigir conclusões únicas (para significância ou confiança);
- Evidenciar dado = ajuste + resíduo, no qual o ajuste é reconhecidamente incompleto e não apenas incerto.
- Extrair as descrições da estrutura das respostas, passo a passo.

Nesta pesquisa, a Análise Exploratória de Dados foi utilizada por meio das distribuições de frequência e medidas estatísticas descritivas, para obter informações e explorar as características das unidades amostrais selecionadas nos conjuntos habitacionais, sem e com o sistema de aquecimento solar de água.

É necessário que as respostas sejam analisadas em algum tipo de contexto, pois, segundo Tukey (1993), o objetivo da Análise Exploratória de Dados é descobrir padrões nos dados. De uma maneira metafórica, este autor muitas vezes compara esta técnica como o trabalho de um detetive; o papel do analista de dados é "ouvir" os dados de várias maneiras possíveis até que uma "história" plausível seja evidente, mesmo que esta descrição não seja confirmada em amostras subsequentes (BEHRENS, 1997).

A técnica auxiliará a compreender melhor as características elementares dos conjuntos habitacionais e, posteriormente, contribuir nas análises e discussões de resultados obtidos por meio de técnicas confirmatórias.

3.2.2 Análise multivariada

A estrutura geral de um banco de dados multivariado correspondente a g grupos avaliados em p variáveis contendo n_i unidades amostrais por grupo, pode ser expressa tendo como elemento genérico X_{ijk} (i = 1, ..., g; j = 1, ..., n_i e k = 1, ..., p).

No presente estudo, o banco de dados com todas as informações de consumo de água, energia elétrica e dados socioeconômicos agregados apresenta o elemento genérico X_{ijk}, o qual representará a resposta da k-ésima variável, k = 1, ..., 11, [k = 1 (consumo de água); k = 2 (consumo de energia); k = 3, ..., 11 (socioeconômicas)],

medida na j-ésima unidade amostral (residência), j = 1, ..., n = n_1 + n_2 (n = 287 amostras, n_1 = 135, número de unidades amostrais do Conjunto Habitacional Sem Sistema, n_2 = 152, número de unidades amostrais do Conjunto Habitacional Com Sistema), do i-ésimo bairro, considerando i = 1, 2 (i =1: Conjunto Habitacional Sem Sistema e i = 2: Conjunto Habitacional Com Sistema).

A estrutura tabular geral dos dados das variáveis utilizadas nesta pesquisa pode ser definida conforme Tabela 1.

Tabela 1: Representação tabular dos dados de consumos de água, energia elétrica e socioeconômicos.

Bairro	Residência	Variáveis					Vetor de Resposta
		Consumo {c}		Socioeconômicas {s}			
		X_1	X_2	X_3	...	X_{11}	
Bairro 1	1	$x_{1\,1\,1}$	$x_{1\,1\,2}$	$x_{1\,1\,3}$	...	$x_{1\,1\,11}$	$\boldsymbol{x}'_{1\,1}$
Conj. Hab.	2	$x_{1\,2\,1}$	$x_{1\,2\,2}$	$x_{1\,2\,3}$	...	$x_{1\,2\,11}$	$\boldsymbol{x}'_{1\,2}$
Sem Sistema	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	n_1=135	$x_{1\,n_1\,1}$	$x_{1\,n_1\,2}$	$x_{1\,n_1\,3}$	...	$x_{1\,n_1\,11}$	$\boldsymbol{x}'_{1\,n_1}$
Bairro 2	1	$x_{2\,1\,1}$	$x_{2\,1\,2}$	$x_{2\,1\,3}$	...	$x_{2\,1\,11}$	$\boldsymbol{x}'_{2\,1}$
Conj. Hab.	2	$x_{2\,2\,1}$	$x_{2\,2\,2}$	$x_{2\,2\,3}$	...	$x_{2\,2\,11}$	$\boldsymbol{x}'_{2\,2}$
Com Sistema	⋮	⋮	⋮	⋮	⋮	⋮	
	n_2=152	$x_{2\,n_2\,1}$	$x_{2\,n_2\,2}$	$x_{2\,n_2\,3}$	...	$x_{2\,n_2\,11}$	$\boldsymbol{x}'_{2\,n_2}$

O vetor linha $\boldsymbol{x}'_{i\,j}$ representa as respostas das variáveis X_1, ..., X_{11}, da j-ésima unidade amostral (residência) do i-ésimo bairro (vetor reposta de dimensão 11).

A representação matricial da Tabela 1 pode ser expressa em $\mathbb{X}_{(nxp)}$, com n = 287 e p = 11, como

$$\mathbb{X}_{nxp} = \begin{bmatrix} x_{1\,1\,1} & \cdots & x_{1\,1\,p} \\ \vdots & \ddots & \vdots \\ x_{1\,n_1\,1} & \cdots & x_{1\,n_1 p} \\ x_{2\,1\,1} & \cdots & x_{2\,1\,p} \\ \vdots & \ddots & \vdots \\ x_{2\,n_2\,1} & \cdots & x_{2\,n_2\,p} \end{bmatrix}; \text{ com } n = n_1 + n_2.$$

*

3.2.3 Teste de dois vetores de médias

Para realizar a comparação entre os dois conjuntos habitacionais, considerando as características de consumo de água e energia e as socioeconômicas, foi utilizado o procedimento T^2 de Hotelling. A estatística T^2 de Hotelling é utilizada para testar a igualdade de dois vetores de médias de populações normais multivariadas e, portanto, apropriada para comparar respostas de um conjunto de dados de uma população (população 1) com respostas independentes de um conjunto de dados de outra população (população 2) (JOHNSON; WICHERN, 2007).

A estrutura geral dos dados para esta técnica pode ser definida conforme Tabela 1, tomando-se a seguinte partição da matriz geral $\mathbb{X}_{nxp} = [\mathbb{X}_{nxp1} \mid \mathbb{X}_{nxp2}]$, com *p1* = $p_{\{c\}}$ = 2 (consumo) e *p2* = $p_{\{s\}}$ = 9 (socioeconômico).

Considerando $\boldsymbol{\mu}_1$ e $\boldsymbol{\mu}_2$ os vetores de médias populacionais dos bairros 1 e 2, respectivamente, as hipóteses estatísticas a serem testadas são

$$\begin{cases} H_0: \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \mathbf{0} \\ H_1: \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 \neq \mathbf{0} \end{cases}$$

Sob a partição adotada, considerando amostras aleatórias independentes do vetor com p (p = 11) variáveis para os dois bairros, o vetor de médias do bairro i pode ser estimado por $\overline{\boldsymbol{x}}_i' = (\bar{x}_{i\cdot 1} \ \ldots \ \bar{x}_{i\cdot k} \ \ldots \ \bar{x}_{i\cdot p})$**,** em que

$$\overline{\boldsymbol{x}}_i = \begin{bmatrix} \bar{x}_{i\cdot 1} \\ \bar{x}_{i\cdot 2} \\ --- \\ \bar{x}_{i\cdot 3} \\ \vdots \\ \bar{x}_{i\cdot p} \end{bmatrix} = \begin{bmatrix} \overline{\boldsymbol{x}}_{i_{2x1}}^{\{c\}} \\ --- \\ \overline{\boldsymbol{x}}_{i_{9x1}}^{\{s\}} \end{bmatrix},$$

$$\mathbb{S}_{i_{pxp}} = \begin{bmatrix} s_{1\,1}^{(i)} & s_{1\,2}^{(i)} & | & s_{1\,3}^{(i)} & \cdots & s_{1\,k}^{(i)} & \cdots & s_{3\,p}^{(i)} \\ s_{2\,1}^{(i)} & s_{2\,2}^{(i)} & | & s_{2\,3}^{(i)} & \cdots & s_{2\,k}^{(i)} & \cdots & s_{2\,p}^{(i)} \\ - - & - - & - - & - - & - - & - - & - - & - - \\ s_{3\,1}^{(i)} & s_{3\,2}^{(i)} & | & s_{3\,3}^{(i)} & \cdots & s_{3\,k}^{(i)} & \cdots & s_{3\,p}^{(i)} \\ \vdots & \vdots & | & \vdots & \ddots & \vdots & \ddots & \vdots \\ s_{k\,1}^{(i)} & s_{k\,2}^{(i)} & | & s_{k\,3}^{(i)} & \cdots & s_{k\,k}^{(i)} & \cdots & s_{k\,p}^{(i)} \\ \vdots & \vdots & | & \vdots & \ddots & \vdots & \ddots & \vdots \\ s_{p\,1}^{(i)} & s_{p\,2}^{(i)} & | & s_{p\,3}^{(i)} & \cdots & s_{p\,k}^{(i)} & \cdots & s_{p\,p}^{(i)} \end{bmatrix} = \begin{bmatrix} \mathbb{S}_{i_{2x2}}^{\{c\}} & | & \mathbb{S}_{i_{2x9}}^{\{cs\}} \\ - - - & | & - - - \\ \mathbb{S}_{i_{9x2}}^{\{sc\}} & | & \mathbb{S}_{i_{9x9}}^{\{s\}} \end{bmatrix},$$

com

$\bar{x}_{i\,\cdot\,k} = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{i\,j\,k}$, para $i = 1, 2$, $k = 1, \ldots, 11$, $c =$ consumo e $s =$ socioeconômico.

O índice sobrescrito no vetor $\overline{\boldsymbol{x}}_i^{\{l\}}$ e na matriz $\mathbb{S}_i^{\{l\}}$, serão denominados por *l*, sendo *l* = *c* para consumo e *l* = *s* para socioeconômico. Neste sentido, os vetores $\overline{\boldsymbol{x}}_{i_{2x1}}^{\{c\}}$ e $\overline{\boldsymbol{x}}_{i_{9x1}}^{\{s\}}$ representam as médias de consumo (água e energia elétrica) e socioeconômicas (número médio de moradores por cômodo, renda *per capita* da unidade amostral, idade média dos moradores, indicador da variação da quantidade de sexo, médias dos anos de estudo na residência, indicador de idade dependente, indicador de maior permanência, quantidade de itens na residência que consomem energia, indicador de equipamentos que consomem maior quantidade de energia elétrica), respectivamente. Da mesma forma, $\mathbb{S}_{i_{2x2}}^{\{c\}}$, representa a matriz de variâncias e covariâncias dos consumos, $\mathbb{S}_{i_{9x9}}^{\{s\}}$, a matriz de variâncias e covariâncias dos socioeconômicos e, por fim, $\mathbb{S}_{i_{2x9}}^{\{cs\}}$e $\mathbb{S}_{i_{9x2}}^{\{sc\}}$, as matrizes das covariâncias entre consumo e socioeconômico. A matriz $\mathbb{S}_{i_{2x9}}^{\{cs\}}$ terá como transposta $\mathbb{S}_{i_{9x2}}^{\{sc\}}$.

Para a comparação dos dois conjuntos habitacionais foram realizados, separadamente, o teste T^2 de Hotelling para as variáveis de consumo (k = 1, 2) e, posteriormente, para as variáveis socioeconômicas (k = 3, ..., 11).

Para o teste de consumo, as hipóteses estatísticas a serem testadas são dadas por

$$\begin{cases} H_0: \boldsymbol{\mu}_1^{\{c\}} - \boldsymbol{\mu}_2^{\{c\}} = \mathbf{0} \\ H_1: \boldsymbol{\mu}_1^{\{c\}} - \boldsymbol{\mu}_2^{\{c\}} \neq \mathbf{0} \end{cases};$$

e o vetor de médias e a matriz de variância-covariância do bairro i podem ser estimados por

$$\overline{\boldsymbol{x}}_{i_{2x1}}^{\{c\}} = \begin{bmatrix} \bar{x}_{i\cdot 1} \\ \bar{x}_{i\cdot 2} \end{bmatrix} \text{ e } \mathbb{S}_{i_{2x2}}^{\{c\}} = \begin{bmatrix} s_{11}^{(i)} & s_{12}^{(i)} \\ s_{21}^{(i)} & s_{22}^{(i)} \end{bmatrix}, \text{ com}$$

$$s_{kk}^{(i)} = \frac{\sum_{j=1}^{n_i} x_{ijk}^2 - n_i\, \overline{x}_{i\cdot k}^2}{n_i - 1} \text{ e } s_{kk'}^{(i)} = \frac{\sum_{j=1}^{n_i} x_{ijk}\, x_{ijk'} - n_i\, \bar{x}_{i\cdot k}\, \bar{x}_{i\cdot k'}}{n_i - 1 \cdot},$$

$k \neq k', k = 1,2$ e $i = 1{,}2$.

Para o teste socioeconômico, as hipóteses estatísticas a serem testadas são dadas por

$$\begin{cases} H_0: \boldsymbol{\mu}_1^{\{s\}} - \boldsymbol{\mu}_2^{\{s\}} = \mathbf{0} \\ H_1: \boldsymbol{\mu}_1^{\{s\}} - \boldsymbol{\mu}_2^{\{s\}} \neq \mathbf{0} \end{cases};$$

O vetor de médias e a matriz de variância-covariância do bairro i, para o conjunto de variáveis socioeconômicas, podem ser estimados por

$$\overline{\boldsymbol{x}}_{i_{9x1}}^{\{s\}} = \begin{bmatrix} \bar{x}_{i\cdot 3} \\ \vdots \\ \bar{x}_{i\cdot 11} \end{bmatrix} \text{ e } \mathbb{S}_{i_{9x9}}^{\{s\}} = \begin{bmatrix} s_{3\,3}^{(i)} & s_{3\,4}^{(i)} & \cdots & s_{3\,11}^{(i)} \\ s_{4\,3}^{(i)} & s_{4\,4}^{(i)} & \cdots & s_{4\,11}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ s_{11\,3}^{(i)} & s_{11\,4}^{(i)} & \cdots & s_{11\,11}^{(i)} \end{bmatrix}, \text{ com}$$

$$s_{kk}^{(i)} = \frac{\sum_{j=1}^{n_i} x_{ijk}^2 - n_i\, \overline{x}_{i\cdot k}^2}{n_i - 1} \text{ e } s_{kk'}^{(i)} = \frac{\sum_{j=1}^{n_i} x_{ijk}\, x_{ijk'} - n_i\, \bar{x}_{i\cdot k}\, \bar{x}_{i\cdot k'}}{n_i - 1},$$

$k \neq k', k = 3, \ldots, 11$ e $i = 1, 2$.

Antes da realização do teste, é necessário que as pressuposições quanto à normalidade multivariada dos dados e sua homogeneidade sejam verificadas.

Por meio do gráfico *Q-Q plot* pode-se inspecionar a normalidade multivariada dos dados para a análise estatística. Este procedimento gráfico utiliza no eixo das abscissas a distância generalizada de Mahalanobis expressa na amostra por:

$\delta_{ij}^{2\{l\}} = \left(\boldsymbol{x}_{ij}^{\{l\}} - \overline{\boldsymbol{x}}_{i}^{\{l\}}\right)' \mathbb{S}_{i}^{\{l\}} \left(\boldsymbol{x}_{ij}^{\{l\}} - \overline{\boldsymbol{x}}_{i}^{\{l\}}\right)$, i = 1, 2, j = 1, ..., n_i ($n_1 = 135$ e $n_2 = 157$), l = c (consumo) ou s (socioeconômico);

no qual $\boldsymbol{x}_{ij}^{\{l\}}$ são os vetores $p_{\{l\}}$-dimensionais dos valores observados, $\overline{\boldsymbol{x}}_{i}^{\{l\}}$ o vetor de médias e $\mathbb{S}_{i}^{\{l\}}$ a matriz de variâncias e covariâncias.

Para construção do gráfico *Q-Q plot*, é necessário calcular a distância de Mahalanobis, $\delta_{ij}^{2\{l\}}$, para todos os elementos da amostra e, posteriormente, dispor os valores calculados em ordem crescente, $\delta_{(1)}^{2\{l\}} \leq \delta_{(2)}^{2\{l\}} \leq \cdots \leq \delta_{(n_i)}^{2\{l\}}$, sendo $\delta_{(j)}^{2\{l\}}$ a j-ésima estatística de ordem. Após, deve-se construir o gráfico dos pares $\left[\delta_{(j)}^{2\{l\}}; \chi^2_{\left(p_{\{l\}};\ \left(\left(j-\frac{1}{2}\right)/n_i\right)\right)}\right]$, nas quais $\chi^2_{\left(p_{\{l\}};\left(\left(j-\frac{1}{2}\right)/n_i\right)\right)}$ representa o quantil de ordem $100\left(\left(j-\frac{1}{2}\right)/n_i\right)$% da distribuição qui-quadrado com $p_{\{l\}}$ graus de liberdade, isto é,

$$P\left[\chi^2_{p_{\{l\}}} \leq \left[\chi^2_{\left(p_{\{l\}};\left(\left(j-\frac{1}{2}\right)/n_i\right)\right)}\right]\right] = \left(j-\frac{1}{2}\right)/n_i.$$

O gráfico resultante é dado por uma nuvem de pontos distribuídos na proximidade de uma reta, quando a normalidade multivariada se ajusta bem aos dados amostrais.

Além da distribuição multinormal dos dados, para realizar o teste T^2 de Hotelling, também é necessário que seja efetuado o teste multivariado de homogeneidade de matrizes de variâncias-covariâncias. O teste de Box (JOHNSON; WICHERN, 2007) pode ser utilizado para verificar a homogeneidade referida e existem vários softwares estatísticos que processam este teste.

$$\begin{cases} H_0: \Sigma_1^{\{l\}} = \Sigma_2^{\{l\}} \\ H_1: \Sigma_1^{\{l\}} \neq \Sigma_2^{\{l\}} \end{cases}; \quad l = \text{c (consumo) ou s (socioeconômico).}$$

A estatística do teste de homogeneidade, para este trabalho, é dada por,

$$M_{\{l\}} = (n-2)ln\left|\mathbb{S}^{\{l\}}\right| - \sum_{i=1}^{2}(n_i - 1)ln\left|\mathbb{S}_i^{\{l\}}\right|$$

$$C_{\{l\}}^{-1} = 1 - \frac{\left(2p_{\{l\}}^2 + 3p_{\{l\}} - 1\right)}{6\left(p_{\{l\}} + 1\right)}\left(\sum_{i=1}^{2}\frac{1}{n_i - 1} - \frac{1}{(n-2)}\right),$$

com

$\mathbb{S}^{\{l\}} = \dfrac{(n_1 - 1)\mathbb{S}_1^{\{l\}} + (n_2 - 1)\mathbb{S}_2^{\{l\}}}{n_1 + n_2 - 2}$, $\quad l =$ c (consumo) ou s (socioeconômico) .

A estatística do teste é dada por $M_{\{l\}}C_{\{l\}}^{-1}$ com distribuição $\chi^2_{\left(\frac{1}{2}p_{\{l\}}(p_{\{l\}}+1)\right)}$, sob a veracidade de H_0; com a regra de decisão habitual.

Duas circunstâncias poderão ser observadas quanto à estrutura das matrizes de variâncias e covariâncias, iguais ou diferentes, ou seja, não rejeição ou rejeição da hipótese de nulidade, respectivamente.

Considerando a situação na qual as matrizes de variâncias-covariâncias são iguais (não rejeição de H_0), a estatística do teste de igualde dos vetores de médias, sob a veracidade de H_0, é expressa por:

$$T^{2^{\{l\}}} = \left[\left(\overline{\boldsymbol{x}}_1^{\{l\}} - \overline{\boldsymbol{x}}_2^{\{l\}}\right)\right]' \left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\mathbb{S}^{\{l\}}\right]^{-1} \left[\left(\overline{\boldsymbol{x}}_1^{\{l\}} - \overline{\boldsymbol{x}}_2^{\{l\}}\right)\right],$$

com $\mathbb{S}^{\{l\}}$conforme definição anterior.

Se o valor da estatística $T^{2^{\{l\}}}$ for maior do que $\frac{(n_1+n_2-2)\,p_{\{l\}}}{n_1+n_2-p_{\{l\}}-1}F_{(\alpha,p_{\{l\}},n_1+n_2-p_{\{l\}}-1)}$, rejeita-se a hipótese nula, H_0, com nível de significância α, indicando que os vetores das médias populacionais são diferentes quanto ao conjunto de características considerados.

Quando se tem matrizes de variâncias-covariâncias diferentes, a estatística do teste é expressa por:

$$T^{2^{\{l\}}} = \left[\left(\overline{\boldsymbol{x}}_1^{\{l\}} - \overline{\boldsymbol{x}}_2^{\{l\}}\right)\right]' \left[\frac{1}{n_1}\mathbb{S}_1^{\{l\}} + \frac{1}{n_2}\mathbb{S}_2^{\{l\}}\right]^{-1} \left[\left(\overline{\boldsymbol{x}}_1^{\{l\}} - \overline{\boldsymbol{x}}_2^{\{l\}}\right)\right].$$

A regra de decisão neste caso é rejeitar a hipótese nula, H_0, ao nível de significância α, quando $T^{2^{\{l\}}}$ for maior que o valor dado por $\chi^2_{(\alpha,p_{\{l\}})}$, e a mesma conclusão se faz presente quanto aos vetores de médias.

3.2.2.1 Intervalos de confiança simultâneos de Hotelling

Para complementar o teste T^2 de Hotelling, intervalos de confiança simultâneos de Hotelling foram construídos para o contraste entre pares de médias. Para cada variável o contraste entre os pares de médias dos dois conjuntos habitacionais será indicado como significativo se o valor zero não situar dentro do intervalo, e a não existência de diferença significante, quando o valor zero situar-se dentro do intervalo.

Quando as matrizes de variâncias-covariâncias são iguais, os limites dos Intervalos de Confiança Simultâneos de Hotelling para as diferenças entre os elementos dos vetores de médias com nível de 100 (1-α)% de confiança são dados por:

$$ICS\left[\mathbf{a}'(\boldsymbol{\mu}_1^{\{l\}} - \boldsymbol{\mu}_2^{\{l\}})\right]: \quad \mathbf{a}'(\overline{\boldsymbol{x}}_1^{\{l\}} - \overline{\boldsymbol{x}}_2^{\{l\}}) \pm \sqrt{\frac{(n_1+n_2-2)p_{\{l\}}}{n_1+n_2-p_{\{l\}}-1} F_{(\alpha;\, p_{\{l\}};\, n_1+n_2-p_{\{l\}}-1)} \frac{n_1+n_2}{n_1\, n_2} \mathbf{a}'\mathbb{S}^{\{l\}}\, \mathbf{a}},$$

sendo $\boldsymbol{a}' = [a_1,\ a_2, \ldots,\ a_p]$, o qual no caso da comparação de pares de médias consiste em um vetor formado por zeros e um, sendo este responsável por indicar a diferença de médias da variável desejada para a comparação estatística.

Para o caso onde as matrizes de variâncias-covariâncias são diferentes, os limites dos Intervalos de Confiança Simultâneos de Hotelling são dados por:

$$ICS\left[\mathbf{a}'(\boldsymbol{\mu}_1^{\{l\}} - \boldsymbol{\mu}_2^{\{l\}})\right]: \quad \mathbf{a}'(\overline{\boldsymbol{x}}_1^{\{l\}} - \overline{\boldsymbol{x}}_2^{\{l\}}) \pm \sqrt{\chi^2_{(\alpha,p_{\{l\}})}\, \mathbf{a}'\left(\frac{1}{n_1}\mathbb{S}_1^{\{l\}} + \frac{1}{n_2}\mathbb{S}_2^{\{l\}}\right) \mathbf{a}}.$$

3.2.4 Análise de correlação canônica

A Análise de Correlação Canônica avalia o grau de associação entre dois conjuntos de variáveis aleatórias. A técnica consiste em identificar e quantificar

as associações máximas entre dois conjuntos de variáveis. Baseia-se na correlação linear simples, entre uma combinação linear das variáveis do primeiro conjunto e uma combinação linear das variáveis do segundo conjunto, cujos pares de combinações lineares são chamados de variáveis canônicas e a correlação existente entre as variáveis canônicas resume-se na chamada correlação canônica (JOHNSON; WICHERN, 2007).

Nesta pesquisa, a técnica da correlação canônica foi utilizada para analisar a associação entre os conjuntos de variáveis de consumo (água e energia elétrica) e as variáveis socioeconômicas, para os respectivos conjuntos habitacionais, Sem Sistema e Com Sistema.

A estrutura geral dos dados das variáveis utilizadas nesta técnica pode ser representada conforme Tabela 1, descrita posteriormente.

O vetor linha resposta $\boldsymbol{x'_{ij}}$ representa as medidas da j-ésima unidade amostral (residência) relativas às p variáveis no i-ésimo conjunto habitacional.

A representação genérica apresentada pode ser expressa em uma matriz, fracionada em duas, sendo a primeira relativa ao bairro Orlando Quagliato e a segunda ao bairro Helena Braz Vendramini, nas quais $n_1 = 135$, $n_2 = 152$ e $p = 11$.

Logo:

$$\mathbb{X}_{nxp} = \begin{bmatrix} \mathbb{X}_{1\,(n_1 xp)} \\ --- \\ \mathbb{X}_{2\,(n_2 xp)} \end{bmatrix}, i = 1, 2, j = 1, \dots, n_i, n = n_1 + n_2 \text{ e k} = 1, \dots, \text{p; sendo}$$

$$\mathbb{X}_{1\,(n_1 xp)} = \begin{bmatrix} x_{1\,1\,1} & \cdots & x_{1\,1\,p} \\ \vdots & \vdots & \vdots \\ x_{1\,n_1\,1} & \cdots & x_{1\,n_1\,p} \end{bmatrix}; \qquad \mathbb{X}_{2\,(n_2 xp)} = \begin{bmatrix} x_{2\,1\,1} & \cdots & x_{2\,1\,p} \\ \vdots & \vdots & \vdots \\ x_{2\,n_2\,1} & \cdots & x_{2\,n_2\,p} \end{bmatrix}.$$

A técnica de correlação canônica realizada para cada bairro envolve a associação entre o conjunto de variáveis de consumo (água e energia elétrica) $\boldsymbol{X}_i^{\{c\}} = (\boldsymbol{X}_{i\,1}, \boldsymbol{X}_{i\,2})$ e o conjunto de variáveis socioeconômicas $\boldsymbol{X}_i^{\{s\}} = (\boldsymbol{X}_{i\,3}, \boldsymbol{X}_{i\,4}, \dots, \boldsymbol{X}_{i\,11})$.

O vetor de resposta dos conjuntos envolvendo as 11 variáveis e as respectivas medidas de posição e variabilidade são as mesmas consideradas anteriormente.

Pelo fato das variáveis originais $(\boldsymbol{X}_k)$ não apresentarem as mesmas unidades de medidas, serão usadas nesta técnica variáveis padronizadas $(\boldsymbol{Z}_k)$.

A partir dos conjuntos dados relativos às variáveis padronizadas são desenvolvidas combinações lineares, o par destas combinações lineares é denominado de variáveis canônicas e suas associações de correlações canônicas. Ou seja, as variáveis canônicas são representadas pelas combinações lineares das variáveis padronizadas de consumo de água e energia elétrica (u_{it}) e pelas variáveis padronizadas socioeconômicas (v_{it}). O número t de variáveis canônicas é determinado por, t = min ($p_{\{c\}}$ = quantidade de variáveis de consumo, $p_{\{s\}}$ = quantidade de variáveis socioeconômicas), quantidade possível de pares entre os dois conjuntos de variáveis.

As variáveis canônicas expressas nos vetores $\boldsymbol{a_t}$ e $\boldsymbol{b_t}$ são as soluções do sistema linear homogêneo

$$\begin{cases} \left(\mathbb{R}_i^{\{cs\}} \left(\mathbb{R}_i^{\{s\}}\right)^{-1} \mathbb{R}_{\boldsymbol{i}}^{\{sc\}} - \boldsymbol{\lambda}_{it}\mathbb{R}_i^{\{c\}}\right)\boldsymbol{a}_{it} = \mathbf{0} \\ \left(\mathbb{R}_i^{\{sc\}} \left(\mathbb{R}_i^{\{c\}}\right)^{-1} \mathbb{R}_{\boldsymbol{i}}^{\{cs\}} - \boldsymbol{\lambda}_{it}\mathbb{R}_i^{\{s\}}\right)\boldsymbol{b}_{it} = \mathbf{0} \end{cases}$$

onde $\boldsymbol{\lambda}_{it}$ é o t-ésimo autovalor da matriz $\mathbb{R}_i^{\{cs\}} \left(\mathbb{R}_i^{\{ss\}}\right)^{-1} \mathbb{R}_{\boldsymbol{i}}^{\{sc\}}$ ou da matriz $\mathbb{R}_i^{\{sc\}} \left(\mathbb{R}_i^{\{cc\}}\right)^{-1} \mathbb{R}_{\boldsymbol{i}}^{\{cs\}}$ e para cada autovalor são determinados os respectivos autovetores padronizados ($\boldsymbol{a}'_{it}$ e $\boldsymbol{b}'_{it}$). As matrizes de correlação $\mathbb{R}_i^{\{c\}}, \mathbb{R}_i^{\{s\}}$ e $\mathbb{R}_{\boldsymbol{i}}^{\{cs\}}$ correspondentes às matrizes de variâncias e covariâncias $\mathbb{S}_i^{\{c\}}, \mathbb{S}_i^{\{s\}}$ e $\mathbb{S}_i^{\{cs\}}$ das variáveis originais. Observar que a matriz de correlação da variável original é idêntica à matriz de variâncias e covariâncias das respectivas variáveis padronizadas.

Neste sentido, tem-se:

$$u_{it} = \boldsymbol{a}'_t\mathbf{Z}_i^{\{c\}} = a_{it\,1}\boldsymbol{Z}_1 + a_{it\,2}\boldsymbol{Z}_2$$

$$v_{it} = \boldsymbol{b}'_t\mathbf{Z}_{\boldsymbol{i}}^{\{\boldsymbol{s}\}} = b_{it\,1}\boldsymbol{Z}_3 + b_{it\,2}\boldsymbol{Z}_4 + b_{it\,3}\boldsymbol{Z}_5 + b_{it\,4}\boldsymbol{Z}_6 + b_{it\,5}\boldsymbol{Z}_7 + b_{it\,6}\boldsymbol{Z}_8 + b_{it\,7}\boldsymbol{Z}_9 + \\ + b_{it\,8}\boldsymbol{Z}_{\mathbf{10}} + b_{it\,9}\boldsymbol{Z}_{11},$$

com, $i = 1, 2; t = min(p_{\{c\}}, p_{\{s\}})$

O coeficiente de correlação canônico para o par (u_{it}, v_{it}) é dado por

$$r_{u_{it},v_{it}} = \frac{\boldsymbol{a}'_{it} Cov\left(\mathbf{Z}_i^{\{c\}}, \mathbf{Z}_i^{\{s\}}\right) \boldsymbol{b}_{it}}{\sqrt{\left(\boldsymbol{a}'_{it} Var\left(\mathbf{Z}_i^{\{c\}}\right) \boldsymbol{a}_{it}\right)\left(\boldsymbol{b}'_{it} Var(\mathbf{Z}_i^{\{s\}}) \boldsymbol{b}_{it}\right)}} = \frac{\boldsymbol{a}'_{it}\, \mathbb{R}_i^{\{cs\}}\, \boldsymbol{b}_{it}}{\sqrt{\left(\boldsymbol{a}'_{it}\, \mathbb{R}_i^{\{c\}}\, \boldsymbol{a}_{it}\right)\left(\boldsymbol{b}'_{it}\, \mathbb{R}_i^{\{s\}}\, \boldsymbol{b}_{it}\right)}}.$$

Na sequência, será estabelecido o teste de hipótese relativo apenas à primeira correlação canônica. Esta conduta metodológica deve-se ao fato da primeira correlação canônica reter a maior quantidade informação da variabilidade dos dados contida neste par de variáveis canônicas.

$$\begin{cases} H_0: \rho_{u_{i1} v_{i1}} = 0 \\ H_1: \rho_{u_{i1} v_{i1}} \neq 0 \end{cases}$$

A estatística do teste da primeira correlação canônica é dada por

$$\chi_i^2 = -\left(n_i - \frac{p_{\{c\}} + p_{\{s\}} + 3}{2}\right) ln\Lambda_i,$$

com $\Lambda_i = (1 - \lambda_{i1})(1 - \lambda_{i2})$.

Se o valor da estatística χ_i^2 for maior que $\chi^2_{[p_{\{c\}} p_{\{s\}}; \alpha]}$ rejeita-se a hipótese nula no nível de 100 α % de significância.

Complementa-se o estudo com algumas medidas que permitem auxiliar na interpretação e na análise da qualidade das variáveis canônicas. Essas medidas referem-se às associações entre as variáveis canônicas e suas respectivas variáveis componentes da combinação linear (associação direta), e àquelas que envolvem a variável canônica u_{i1} com as variáveis componentes de v_{i1} (associação cruzada) e vice-versa.

As associações diretas são determinadas por:

$$\mathbb{R}^*_{u_{i1} X_i^{\{c\}}} = \mathbb{R}^{\{c\}}_{u_{i1}} \text{ e } \mathbb{R}^*_{v_{i1} X_i^{\{s\}}} = \mathbb{R}^{\{s\}}_{v_{i1}};$$

enquanto as cruzadas por:

$$\mathbb{R}^*_{u_{i1} X_i^{\{s\}}} = \mathbb{R}^{\{sc\}}_{u_{i1}^{\{c\}}} \text{ e } \mathbb{R}^*_{v_{i1} X_i^{\{c\}}} = \mathbb{R}^{\{cs\}}_{v_{i1}^{\{s\}}}.$$

As dimensões dessas matrizes são dadas por:

$$\mathbb{R}^*_{u_{i1} X_i^{\{c\}}} \Rightarrow \dim(2x2),\ \mathbb{R}^*_{v_{i1} X_i^{\{s\}}} \Rightarrow \dim(9x9),\ \mathbb{R}^*_{u_{i1} X_i^{\{s\}}} \Rightarrow \dim(9x2), \mathbb{R}^*_{v_{i1} X_i^{\{c\}}} \Rightarrow \dim(2x9).$$

3.2.4 Análise fatorial

-

A Análise fatorial tem o objetivo de explicar as variáveis originais estudadas, X_k, por meio do agrupamento destas em subconjuntos de variáveis, denominados fatores (variáveis não observadas, variáveis latentes). Portanto, a técnica trata de um tipo de análise exploratória de dados que consiste em reduzir a quantidade de variáveis originais em um número menor de fatores independentes, de tal maneira que estes fatores contêm subconjuntos de variáveis que apresentam maior semelhança (MINGOTI, 2005).

Para Johnson e Wichern (2007), a finalidade essencial da análise fatorial é descrever, se possível, as relações de covariância entre muitas variáveis em termos de algumas variáveis subjacentes, não observáveis, chamadas fatores aleatórios. Basicamente, a técnica é motivada pelo seguinte argumento: as variáveis podem ser agrupadas por suas correlações.

Para Corrar, Paulo e Dias Filho (2009), a técnica avalia a correlação existente entre um grande número de variáveis e identifica a possibilidade de essas variáveis serem agrupadas em um número menor de variáveis latentes e de que, obviamente, se possa identificar o significado dos agrupamentos realizados; bem como avalia a possibilidade de agrupar k variáveis em um número menor de m fatores. A análise fatorial é uma técnica de interdependência na qual o objetivo é identificar uma estrutura de relacionamentos que permita a explicação das variações ocorridas nas variáveis analisadas.

O propósito geral de técnicas de análise fatorial é encontrar um modo de condensar, resumir, a informação contida em diversas variáveis originais em um conjunto menor de novas dimensões compostas ou variáveis estatísticas (fatores) com uma perda mínima de informação, ou seja, buscar e definir os construtores fundamentais ou dimensões assumidas como inerentes às variáveis originais (Hair Jr et al., 2009).

Neste estudo, a análise fatorial foi utilizada para determinar fatores que pudessem reunir as variáveis originais com maior associação. O entendimento das variáveis e seus resultados podem ser explicados de maneira mais simples, pelo agrupamento das variáveis originais em um número menor de fatores.

A estrutura geral do modelo de análise fatorial dos dados, composta a partir da matriz de correlação das variáveis estudadas, pode ser definida conforme a representação matricial expressa por,

$$\boldsymbol{Z} = \mathbb{L}\boldsymbol{F} + \boldsymbol{\varepsilon},$$

considerando o vetor aleatório $\boldsymbol{X}' = [X_1, \dots, X_p]$; $\boldsymbol{Z}' = [Z_1, \dots, Z_p]$, (p = 11), sendo

$$Z_k = \frac{X_k - \mu_k}{\sigma_k}, k = 1, \dots, p,$$

com μ_k e σ_k estimados pela média e desvio-padrão de X_k, e

$$\mathbb{L}_{pxm} = \begin{bmatrix} l_{1\,1} & l_{1\,2} & \cdots & l_{1\,m} \\ l_{2\,1} & l_{2\,2} & \cdots & l_{2\,m} \\ \vdots & \vdots & \vdots & \vdots \\ l_{p\,1} & l_{p\,2} & \cdots & l_{p\,m} \end{bmatrix};\ \boldsymbol{F}_{mx1} = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{bmatrix};\ \boldsymbol{\varepsilon}_{px1} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{bmatrix}.$$

O vetor $\mathbf{F}_{mx1}$ é aleatório contendo m fatores não observáveis, também chamados de variáveis latentes, que descrevem os elementos da população em estudo, $1 \leq m \leq p = 11$. O vetor $\boldsymbol{\varepsilon}_{px1}$ é o vetor de erros aleatórios correspondente aos erros de medida e à variação de Z_k, que não é explicada pelos fatores comuns F_r, com r = 1, ..., m, abrangidos no modelo. A matriz $\mathbb{L}_{pxm}$ armazena os coeficientes, l_{kr} ("loading" – fator de carga), da k-ésima variável padronizada, Z_k, no r-ésimo fator F_r e representa o grau de relacionamento linear existente entre Z_k e F_r, que são parâmetros constantes a ser estimados por meio do método das componentes principais, (MINGOTI, 2005).

Para a construção do modelo fatorial, torna-se necessário a utilização de alguns indicadores estatísticos antes da obtenção dos fatores que são: a estatística de *Kaiser-Meyer-Olkin* (KMO), o teste esfericidade de *Bartlett*, as comunalidades, os autovalores foram calculados e o gráfico *scree plot* confeccionado a fim de auxiliar na interpretação dos resultados da análise fatorial.

Segundo Hair Jr et al. (2009):

- A estatística *KMO* (medida de adequação da amostra) quantifica o grau de intercorrelações entre as variáveis e a adequação da análise fatorial, o índice varia de 0 a 1 (1 quando cada variável é perfeitamente prevista sem erro pelas outras variáveis);

- O teste de esfericidade de Bartlett, para a presença de associações entre as variáveis, fornece a significância estatística de que a matriz de correlação tem correlações significantes entre pelo menos algumas das variáveis;
- Semelhante à estatística *KMO* a estatística *MSA*, consiste na medida de adequação da amostra de cada variável.
- As comunalidades representam a variância de Z_k explicada pelos fatores incluídos no modelo fatorial;
- Os autovalores representam a variância total explicada por cada fator;
- O gráfico *scree plot* é utilizado para identificar o número ótimo de fatores que podem ser extraídos. O gráfico é feito com os autovalores em relação ao número de fatores em sua ordem de extração e o ponto no qual o gráfico começa a ficar horizontal é considerado indicativo do número máximo de fatores a serem extraídos. Outro critério para determinar o número de fatores é o do autovalor; apenas os fatores com autovalores acima de um (1,00) são considerados. O autovalor corresponde a quanto o fator consegue explicar da variância, ou seja, quanto da variância total dos dados pode ser associado ao fator. Pelo fato dos dados serem padronizados, cada variável tem média zero e variância igual a um, logo, autovalores abaixo de um implicam em fatores menos significativos do que uma variável original (CORRAR, PAULO, DIAS FILHO, 2009).

Verificadas as condições básicas para a construção do modelo fatorial, deve-se escolher um método de estimação para a obtenção dos elementos da matriz $\mathbb{L}$ e das variâncias associadas aos erros (Ψ_k: variância especifica ou inicidade).

No presente estudo optou-se por utilizar o método de estimação denominado por componentes principais.

O método dos componentes principais consiste na decomposição espectral da matriz $\mathbb{R}$ de correlações, considerando os autovalores $\hat{\lambda}_k$de $\mathbb{R}$ juntamente com seus respectivos autovetores normalizados $\hat{e}_k$ expressos na seguinte aproximação:

$$\mathbb{R} \approx \widehat{\mathbb{L}}\widehat{\mathbb{L}}' + \widehat{\Psi},$$

sendo $\widehat{\mathbb{L}}\widehat{\mathbb{L}}' = \sum_{k=1}^{p} \hat{\lambda}_k \, \hat{e}_k \hat{e}_k'$ e $\widehat{\Psi} = diag(\widehat{\Psi}_1, \dots, \widehat{\Psi}_p)$ com $\widehat{\Psi}_k = \widehat{Var}(\varepsilon_k)$.

As variâncias especificas estimadas $\widehat{\Psi}_k$ são os complementos das comunalidades (h_k^2), isto é, $\widehat{\Psi}_k + h_k^2 = 1$, portanto indicam a parte da variabilidade de Z_k associada apenas ao erro aleatório ε_k; sendo o seu complemento a comunalidade, a parte da variabilidade de Z_k explicada pelos m fatores incluídos no modelo fatorial.

A interpretação dos fatores originais pode não ser fácil, devido a coeficientes com valores semelhantes em vários fatores diferentes. Nesta situação, a suposição de ortogonalidade dos fatores é violada e a partição das variáveis originais em diferentes fatores não fica clara, mas difícil de ser justificada. Pode-se utilizar o recurso da transformação ortogonal dos fatores originais na tentativa de se alcançar uma estrutura mais simples, rotação fatorial (MINGOTI, 2005). O método da rotação tem o objetivo de simplificar e melhorar os resultados da técnica da análise fatorial.

Nesta pesquisa, será utilizada a rotação Varimax, um método de rotação fatorial ortogonal, considerado superior a outros métodos de rotação fatorial ortogonal para conseguir uma estrutura fatorial simplificada (HAIR JR et al., 2009).

O critério de rotação Varimax busca a matriz $\mathbb{T}$ tendo como base a tentativa de encontrar para um fator fixo um grupo de variáveis Z_k altamente correlacionadas com o fator e um outro grupo de variáveis que tenham correlação desprezível ou moderada com o fator. Para cada fator fixo, a solução é obtida por meio da maximização da variação dos quadrados das cargas fatoriais originais da matriz $\widehat{\mathbb{L}}$ (MINGOTI, 2005). O critério seleciona $\tilde{l}_{kr}$ que maximizam a quantidade $V = \frac{1}{p}\sum_{r=1}^{m}\left[\sum_{k=1}^{p}\tilde{l}_{kr}^4 - \frac{1}{p}\left(\sum_{k=1}^{p}\tilde{l}_{kr}^2\right)^2\right]$, com $\tilde{l}_{kr} = \frac{\hat{l}_{kr}}{\hat{h}_k}$. Os coeficientes da matriz transformada $\widehat{\mathbb{L}}^* = \widehat{\mathbb{L}}\mathbb{T}$ são expressos por $\hat{l}_{kr}^*$, k = 1, ..., p, são obtidos pelo produto de $\tilde{l}_{kr}$ por $\hat{h}_k$ para que as comunalidades originais sejam preservadas.

Os programas computacionais estatísticos utilizados para trabalhar com o banco de dados e gerar os resultados das técnicas estatísticas neste trabalho foram SPSS Statistics 17.0 e R 3.0.2.

4 RESULTADOS E DISCUSSÃO

4.1 Análise exploratória de dados

Os resultados apresentados a seguir descrevem algumas características dos moradores, das residências e dos conjuntos habitacionais, fornecidos pela análise exploratória de dados.

Tabela 2: Distribuição da quantidade de residências pesquisadas amostradas, por bairro.

Conjunto Habitacional	Quantidade	Percentual
Sem Sistema	135	47,04
Com Sistema	152	52,96
Total Geral	287	100,00

Os percentuais apresentados na Tabela 2 indicam as proporcionalidades das residências dos dois bairros, conjuntos habitacionais, conforme plano amostral estabelecido.

Tabela 3:Distribuição do número de moradores por residência.

Número de Moradores	Quantidade		Percentual		Quantidade Total	Percentual Total
	Sem Sistema	Com Sistema	Sem Sistema	Com Sistema		
1	2	2	1,48	1,32	4	1,39
2	15	18	11,11	11,84	33	11,50
3	28	37	20,74	24,34	65	22,65
4	43	41	31,85	26,97	84	29,27
5	21	32	15,56	21,05	53	18,47
6	15	9	11,11	5,92	24	8,36
7	9	9	6,67	5,92	18	6,27
8	0	3	0,00	1,97	3	1,05
9	2	1	1,48	0,66	3	1,05
Total Geral	135	152	100,00	100,00	287	100,00

Aproximadamente 68,15% das residências do Conjunto Habitacional Sem Sistema apresentam de três a cinco moradores, no Conjunto Habitacional Com Sistema, para esta mesma quantidade de moradores, a proporção é de 72,36%.

Tabela 4:Distribuição do número de cômodos por residência.

Número de Cômodos	Quantidade		Percentual		Quantidade Total	Percentual Total
	Sem Sistema	Com Sistema	Sem Sistema	Com Sistema		
4	36	0	26,67	0,00	36	12,54
5	41	123	30,37	80,92	164	57,14
6	27	18	20,00	11,84	45	15,68
7	20	6	14,81	3,95	26	9,06
8	4	4	2,96	2,63	8	2,79
9	4	1	2,96	0,66	5	1,74
10	2	0	1,48	0,00	2	0,70
11	1	0	0,74	0,00	1	0,35
Total Geral	135	152	100,00	100,00	287	100,00

Por informações fornecidas pela prefeitura, as residências apresentavam cinco cômodos no momento da sua entrega (dois quartos, sala, cozinha e banheiro - considerado como cômodo). Entretanto, em algumas residências do Conjunto Habitacional Sem Sistema, informaram quatro cômodos, talvez por não considerarem o banheiro ou por terem feito alguma reforma. O Conjunto Habitacional Com Sistema foi entregue anos depois, reformas para ampliação não ocorreram em 80,92% das residências.

No Sem Sistema, 65,18% das residências apresentam cinco, seis ou sete cômodos e 26,67% relataram quatro cômodos na residência.

Tabela 5: Distribuição da utilização de lâmpadas fluorescentes em grande parte dos cômodos.

	Quantidade		Percentual			
Lâmpada Fluorescente	Sem Sistema	Com Sistema	Sem Sistema	Com Sistema	Quantidade Total	Percentual Total
Sim	81	82	60,00	53,95	163	56,79
Não	54	70	40,00	46,05	124	43,21
Total Geral	135	152	100,00	100,00	287	100,00

É observado no Conjunto Com Sistema que um pouco mais da metade, 53,95% das residências, utilizam lâmpadas fluorescentes. No Conjunto Habitacional Sem Sistema, este percentual é de 60%.

Na Tabela 6, são apresentadas as distribuições da condição de ocupação das residências.

Tabela 6: Distribuição da condição de ocupação das residências.

	Quantidade		Percentual			
Condição de Ocupação	Sem Sistema	Com Sistema	Sem Sistema	Com Sistema	Quantidade Total	Percentual Total
Alugada	4	7	2,96	4,61	11	3,83
Própria	125	143	92,59	94,08	268	93,38
Cedida	6	2	4,44	1,32	8	2,79
Total Geral	135	152	100,00	100,00	287	100,00

Como relatado anteriormente, as residências destes conjuntos habitacionais não podem ser comercializadas, vendidas ou alugadas. Quando questionado a condição de ocupação das residências, grande parte delas, tanto no Com Sistema quanto no Sem Sistema, relataram que são “Próprias”, 94,08% e 92,59%; respectivamente. Porém, algumas unidades amostrais apresentam condição de ocupação “Alugada” ou “Cedida”, o que não poderia ocorrer nestes tipos de residência.

Tabela 7: Distribuição da faixa etária média por residência.

Faixa Etária Média	Quantidade Sem Sistema	Quantidade Com Sistema	Percentual Sem Sistema	Percentual Com Sistema	Quantidade Total	Percentual Total
9 ⊢ 22	29	61	21,48	40,13	90	31,36
22 ⊢ 35	61	64	45,19	42,11	125	43,55
35 ⊢ 48	29	17	21,48	11,18	46	16,03
48 ⊢ 61	9	7	6,67	4,61	16	5,57
61 ⊢ 74	7	3	5,19	1,97	10	3,48
Total Geral	135	152	100,00	100,00	287	100,00

As idades médias dos moradores das unidades amostrais foram organizadas na Tabela 7. No Conjunto Habitacional Sem Sistema, 33,34% apresentam médias de idades maiores ou iguais a 35 anos ou menores do que 74 anos. Nesta mesma faixa etária média, são 17,76% residências do Conjunto Habitacional Com Sistema. No Conjunto Habitacional Com Sistema de aquecimento solar existe maior percentual de residências, 82,24%, nas quais as idades médias são maiores ou iguais a nove anos ou menores do que 35. Já para o conjunto habitacional Sem Sistema, nesta mesma faixa, o percentual de residências é de 66,67%.

Na Tabela 8, estão organizadas as quantidades de moradores do sexo feminino por residência.

Tabela 8: Distribuição da quantidade de moradores do sexo feminino por residência.

Quantidade Feminino	Quantidade Sem Sistema	Quantidade Com Sistema	Percentual Sem Sistema	Percentual Com Sistema	Quantidade Total	Percentual Total
0 ⊢ 3	85	104	62,96	68,42	189	65,85
3 ⊢ 7	50	48	37,04	31,58	98	34,15
Total Geral	135	152	100,00	100,00	287	100,00

Aproximadamente 63% das residências do Conjunto Habitacional Sem Sistema apresentam no máximo dois moradores do sexo feminino. Para esta mesma classe de moradores, no Conjunto Habitacional Com Sistema, o percentual é de 68,42%. Para a quantidade de 3 a 6 moradores do sexo feminino, no Conjunto Habitacional Sem Sistema e no Com Sistema, são de 37,04% e 31,58%, respectivamente.

Tabela 9: Distribuição da quantidade de moradores do sexo masculino por residência.

Quantidade Masculino	Quantidade		Percentual		Quantidade Total	Percentual Total
	Sem Sistema	Com Sistema	Sem Sistema	Com Sistema		
0 ⊢ 3	103	116	76,30	76,32	219	76,31
3 ⊢ 8	32	36	23,70	23,68	68	23,69
Total Geral	135	152	100,00	100,00	287	100,00

No Conjunto Habitacional Sem Sistema, 76,3% das residências apresentam no máximo dois moradores do sexo masculino e 23,7% possuem de 3 a 7 moradores do sexo masculino. Para o Com Sistema são observados percentuais semelhantes, 76,32% e 23,68% para as respectivas quantidades de moradores do sexo masculino.

Tabela 10: Distribuição da classe econômica por residência no ano de 2013.

Classe Econômica	Renda Média Bruta Familiar por mês (R$)	Quantidade		Percentual		Quantidade Total	Percentual Total
		Sem Sistema	Com Sistema	Sem Sistema	Com Sistema		
D	714,00	21	10	15,56	6,58	31	10,80
C2	1.024,00	51	43	37,78	28,29	94	32,75
C1	1.541,00	45	56	33,33	36,84	101	35,19
B2	2.565,00	16	39	11,85	25,66	55	19,16
B1	4.418,00	2	4	1,48	2,63	6	2,09
Total Geral		135	152	100,00	100,00	287	100,00

A classe econômica e a renda média bruta familiar por mês foram determinadas pelo Critério de Classificação Econômica Brasil da Associação Brasileira de Empresas de Pesquisa – ABEP.

Como mencionado anteriormente, grande parte das residências foram sorteadas para famílias de baixa renda, nos anos em que ocorreram os sorteios. Os dados coletados nesta pesquisa mostram que 86,67% das residências do Sem Sistema foram classificadas nas classes econômicas de D a C1; para estas mesmas classes econômicas, tem-se 71,71% das residências do Com Sistema. O percentual complementar para 100% são residências classificadas nas classes econômicas B2 e B1, nenhuma residência foi classificada nas classes A2 e A1, nos respectivos conjuntos habitacionais.

Tabela 11: Distribuição do grau de instrução do chefe da família por residência.

Grau de Instrução	Quantidade		Percentual		Quantidade Total	Percentual Total
	Sem Sistema	Com Sistema	Sem Sistema	Com Sistema		
1	48	23	35,56	15,13	71	24,74
2	44	62	32,59	40,79	106	36,93
3	26	40	19,26	26,32	66	23,00
4	17	26	12,59	17,11	43	14,98
5	0	1	0,00	0,66	1	0,35
Total Geral	135	152	100,00	100,00	287	100,00

	Grau de Instrução do Chefe de Família	
	Nomenclatura Antiga	Nomenclatura Atual
1	Analfabeto/Primário Incompleto	Analfabeto/Funtamental 1 Incompleto
2	Primário Completo/Ginásio Incompleto	Fundamental 1 Completo/Fundamental 2 Incompleto
3	Ginásio Completo/Colegial Incompleto	Fundamental 2 Completo/Médio Incompleto
4	Colegial Completo/Superior Incompleto	Médio Completo/Superior Incompleto
5	Superior Completo	Superior Completo

Das residências dos Conjuntos Habitacionais Sem Sistema e Com Sistema, 12,59% e 17,77%, respectivamente, o grau de instrução do chefe de família foi informado como “Colegial Completo/Superior Incompleto” – “Médio Completo/Superior Incompleto” ou “Superior Completo”. O percentual para completar 100% para cada um dos conjuntos habitacionais é de residências onde informaram o grau de instrução do chefe de família de “Analfabeto/Primário Incompleto” – “Analfabeto/Fundamental 1 Incompleto” a “Ginásio Completo/Colegial Incompleto” – “Fundamental 2 Completo/Médio Incompleto”.

Com o questionário aplicado nas unidades amostrais foram levantados dados de diversas variáveis socioeconômicas, que forneceram informações sobre a residência e seus moradores, apresentadas anteriormente. Estas variáveis deram origem a outras mais interessantes e informativas para o estudo, as quais foram utilizadas no plano analítico da pesquisa, Tabelas 12 a 21.

Tabela 12: Distribuição do consumo médio de água (m^3) por residência.

Consumo Médio Água	Quantidade		Percentual		Quantidade Total	Percentual Total
	Sem Sistema	Com Sistema	Sem Sistema	Com Sistema		
3,20 ⊢ 9,35	9	39	6,67%	25,66%	48	16,72%
9,35 ⊢ 15,50	51	62	37,78%	40,79%	113	39,37%
15,50 ⊢ 21,65	43	40	31,85%	26,32%	83	28,92%
21,65 ⊢ 27,80	22	10	16,30%	6,58%	32	11,15%
27,80 ⊢ 33,95	10	1	7,41%	0,66%	11	3,83%
Total Geral	135	152	100,00	100,00	287	100,00

Na Tabela 12, são apresentados os consumos médios de água (m^3), nos respectivos conjuntos habitacionais. Observa-se no Conjunto Habitacional Sem Sistema maior percentual (55,56%) de residências nas quais o consumo médio de água é mais elevado. No Conjunto Habitacional Com Sistema, o consumo médio de água é menor, ou seja, maior percentual de residências (66,45%) nas quais o consumo médio é menor.

Tabela 13: Distribuição do consumo médio de energia elétrica (kWh) por residência.

Consumo Médio Energia Elétrica	Quantidade		Percentual		Quantidade Total	Percentual Total
	Sem Sistema	Com Sistema	Sem Sistema	Com Sistema		
40,90 ⊢ 103,05	19	56	14,07%	36,84%	75	26,13%
103,05 ⊢ 165,20	54	67	40,00%	44,08%	121	42,16%
165,20 ⊢ 227,35	41	24	30,37%	15,79%	65	22,65%
227,35 ⊢ 289,50	17	2	12,59%	1,32%	19	6,62%
289,50 ⊢ 351,65	4	3	2,96%	1,97%	7	2,44%
Total Geral	135	152	100,00	100,00	287	100,00

O consumo médio de energia elétrica (kWh) por residência foi apresentado na Tabela 13, respectivamente para cada Conjunto Habitacional. Como na distribuição do consumo médio de água, o Conjunto Habitacional Com Sistema apresenta maior percentual de residências (80,92%) nas quais o consumo médio de energia elétrica é menor. Já para o Conjunto Habitacional Sem Sistema é observado maior percentual de residências com consumo médio de energia maior.

Tabela 14: Distribuição da razão do número de moradores por número de cômodos por residência.

Número de Moradores/ Número de Cômodos	Quantidade		Percentual		Quantidade Total	Percentual Total
	Sem Sistema	Com Sistema	Sem Sistema	Com Sistema		
0,20 ⊢ 0,62	48	59	35,56%	38,82%	107	37,28%
0,62 ⊢ 1,04	70	77	51,85%	50,66%	147	51,22%
1,04 ⊢ 1,46	12	13	8,89%	8,55%	25	8,71%
1,46 ⊢ 1,88	4	3	2,96%	1,97%	7	2,44%
1,88 ⊢ 2,30	1	0	0,74%	0,00%	1	0,35%
Total Geral	135	152	100,00	100,00	287	100,00

A razão do número de moradores por número de cômodos, Tabela 14, indica a quantidade média de moradores por cômodo, ou seja, é um indicador da densidade de moradores por unidade amostral. É observado equilíbrio entre os dois conjuntos habitacionais, os percentuais de residências nos intervalos estão balanceados.

Tabela 15: Distribuição da razão da renda média familiar por número de moradores por residência.

Renda Média Familiar/ Número de Moradores	Quantidade		Percentual		Quantidade Total	Percentual Total
	Sem Sistema	Com Sistema	Sem Sistema	Com Sistema		
102 ⊢ 390	97	90	71,85%	59,21%	187	65,16%
390 ⊢ 678	30	35	22,22%	23,03%	65	22,65%
678 ⊢ 966	8	18	5,93%	11,84%	26	9,06%
966 ⊢ 1254	0	1	0,00%	0,66%	1	0,35%
1254 ⊢ 1542	0	8	0,00%	5,26%	8	2,79%
Total Geral	135	152	100,00	100,00	287	100,00

Observa-se, na Tabela 15, a distribuição das rendas *per capita* por residência. No Conjunto Habitacional Sem Sistema, o intervalo com valores de rendas maiores ou iguais a R$ 102,00 ou menores que R$ 390,00 possui maior percentual de residências, 71,85% em relação ao Conjunto Habitacional Com Sistema, ou seja, existe maior percentual de residências com menor renda *per capita* no Conjunto Habitacional Sem Sistema. No Conjunto Habitacional Com Sistema, é observado percentual maior de residências (5,92%) com valores de renda *per capita* em intervalos maiores em relação ao Conjunto Habitacional Sem Sistema, logo, no Conjunto Habitacional Com Sistema existe maior percentual de residências com renda *per capita* maior.

Tabela 16: Distribuição do indicador da variação do sexo por residência.

Indicador Sexo	Quantidade		Percentual		Quantidade Total	Percentual Total
	Sem Sistema	Com Sistema	Sem Sistema	Com Sistema		
-1,00 ⊢ -0,60	1	4	0,74%	2,63%	5	1,74%
-0,60 ⊢ -0,20	27	34	20,00%	22,37%	61	21,25%
-0,20 ⊢ 0,20	61	70	45,19%	46,05%	131	45,64%
0,20 ⊢ 0,60	36	38	26,67%	25,00%	74	25,78%
0,60 ⊢| 1,00	10	6	7,41%	3,95%	16	5,57%
Total Geral	135	152	100,00	100,00	287	100,00

No indicador da variação do sexo por unidade amostral, os percentuais indicam equilíbrio entre os dois conjuntos habitacionais.

Tabela 17: Distribuição da média dos anos de estudos por residência.

Média de Anos de Estudos	Quantidade		Percentual		Quantidade Total	Percentual Total
	Sem Sistema	Com Sistema	Sem Sistema	Com Sistema		
0,00 ⊢ 2,60	4	3	2,96%	1,97%	7	2,44%
2,60 ⊢ 5,20	30	19	22,22%	12,50%	49	17,07%
5,20 ⊢ 7,80	58	70	42,96%	46,05%	128	44,60%
7,80 ⊢ 10,40	35	48	25,93%	31,58%	83	28,92%
10,40 ⊢I 13,00	8	12	5,93%	7,89%	20	6,97%
Total Geral	135	152	100,00	100,00	287	100,00

Para a média dos anos de estudo por residência, Tabela 17, os percentuais são mais altos, 85,52% para média de anos de estudos maiores iguais a 5,2 anos e menores ou iguais a 13 anos, no Conjunto Habitacional Com Sistema, em relação ao Sem Sistema. Maior percentual de residências nas quais a média de anos de estudos são maiores para o Conjunto Habitacional Com Sistema e para o Conjunto Habitacional Sem Sistema há maior percentual de residências nas quais a média de anos de estudos é menor.

Tabela 18: Distribuição do indicador de idade dependente por residência.

Indicador Dependência	Quantidade Sem Sistema	Quantidade Com Sistema	Percentual Sem Sistema	Percentual Com Sistema	Quantidade Total	Percentual Total
0,00 ⊢ 0,20	35	36	25,93%	23,68%	71	24,74%
0,20 ⊢ 0,40	49	49	36,30%	32,24%	98	34,15%
0,40 ⊢ 0,60	41	53	30,37%	34,87%	94	32,75%
0,60 ⊢ 0,80	6	12	4,44%	7,89%	18	6,27%
0,80 ⊢I 1,00	4	2	2,96%	1,32%	6	2,09%
Total Geral	135	152	100,00	100,00	287	100,00

Para o Conjunto Habitacional Sem Sistema, o percentual de unidades amostrais nas quais o indicador é maior ou igual a 0 ou menor do que 0,40 é de 62,23%. No Conjunto Habitacional Com Sistema, o percentual de residências neste intervalo é de 55,92%, ou seja, existe maior percentual de residências no Conjunto Habitacional Com Sistema nas quais o indicador é maior ou igual a 0,40 e menor ou igual a 1,00.

Tabela 19: Distribuição do índice de maior permanência por residência.

Índice de Maior Permanência	Quantidade Sem Sistema	Quantidade Com Sistema	Percentual Sem Sistema	Percentual Com Sistema	Quantidade Total	Percentual Total
0,00 ⊢ 0,20	3	9	2,22%	5,92%	12	4,18%
0,20 ⊢ 0,40	22	18	16,30%	11,84%	40	13,94%
0,40 ⊢ 0,60	32	43	23,70%	28,29%	75	26,13%
0,60 ⊢ 0,80	58	60	42,96%	39,47%	118	41,11%
0,80 ⊢I 1,00	20	22	14,81%	14,47%	42	14,63%
Total Geral	135	152	100,00	100,00	287	100,00

A Tabela 19 apresenta o percentual de residências semelhantes, distribuídos nos intervalos do indicador.

Tabela 20:Distribuição da quantidade de itens que consomem energia elétrica por residência.

Quantidade de Itens	Quantidade		Percentual		Quantidade Total	Percentual Total
	Sem Sistema	Com Sistema	Sem Sistema	Com Sistema		
3 ⊢ 8	44	36	32,59%	23,68%	80	27,87%
8 ⊢ 13	70	76	51,85%	50,00%	146	50,87%
13 ⊢ 18	13	32	9,63%	21,05%	45	15,68%
18 ⊢ 23	7	7	5,19%	4,61%	14	4,88%
23 ⊢I 28	1	1	0,74%	0,66%	2	0,70%
Total Geral	135	152	100,00	100,00	287	100,00

No Conjunto Habitacional Com Sistema, existe maior percentual de residências com quantidade de itens maior ou igual a 13 ou menor ou igual a 28, 26,32%, em comparação ao outro conjunto. Para o Conjunto Habitacional Sem Sistema, o percentual de residências com quantidades de itens maior ou igual a 3 ou menor do que 13 é maior em relação ao Conjunto Habitacional Com Sistema, 84,44%, 73,68%, respectivamente.

Tabela 21: Distribuição do índice de equipamentos com consumo alto por residência.

Indicador Consumo Alto	Quantidade		Percentual		Quantidade Total	Percentual Total
	Sem Sistema	Com Sistema	Sem Sistema	Com Sistema		
0,17 ⊢ 0,27	42	54	31,11%	35,53%	96	33,45%
0,27 ⊢ 0,37	70	80	51,85%	52,63%	150	52,26%
0,37 ⊢ 0,47	23	13	17,04%	8,55%	36	12,54%
0,47 ⊢ 0,57	0	3	0,00%	1,97%	3	1,05%
0,57 ⊢ 0,67	0	2	0,00%	1,32%	2	0,70%
Total Geral	135	152	100,00	100,00	287	100,00

Observa-se equilíbrio entre os percentuais de residências com valores de indicador maiores ou iguais a 0,17 e menores do que 0,37. Para valores do indicador maiores ou iguais a 0,47 ou menores do que 0,67, o Conjunto Habitacional Com Sistema apresenta percentual de 3,29% e para o Conjunto Habitacional Sem Sistema não existem residências neste intervalo.

Para avaliar a significância da distribuição de frequências de algumas das variáveis de interesse apresentadas nas tabelas anteriores e verificar a existência de características comuns ou distintas, foi realizado o teste estatístico para

comparar os resultados obtidos nos dois conjuntos habitacionais. A homogeneidade da distribuição das frequências nos conjuntos habitacionais foi verificada pelo teste Qui-Quadrado. Os resultados são apresentados na Tabela 22.

Tabela 22: Nível descritivo e resultado do teste de hipóteses da homogeneidade.

Variável	Valor p	Resultado
LAMPF	0,302	Homogeneidade entre os Conjuntos Habitacionais
CONDOCUP	0,220	Homogeneidade entre os Conjuntos Habitacionais
%FEM	0,569	Homogeneidade entre os Conjuntos Habitacionais
%MAS	0,569	Homogeneidade entre os Conjuntos Habitacionais
CLAECON	0,004	Não há homogeneidade entre os Conjuntos Habitacionais
GRAUINST	0,002	Não há homogeneidade entre os Conjuntos Habitacionais

LAMPF: lâmpada fluorescente em grande parte dos cômodos; CONDOCUP: condição de ocupação; %FEM: porcentagem de moradores do sexo feminino; %MAS: porcentagem de moradores do sexo masculino; CLAECON: classe econômica determinada pelo critério de classificação ABEP; GRAUINST: grau de instrução do chefe da família.

As variáveis CLAECON (classe econômica) e GRAUINST (grau de instrução do chefe da família) apresentaram resultados significativos considerando nível de significância de 5%, ou seja, não existem evidências de homogeneidade ou não existem evidências de igualdade entre os conjuntos habitacionais para as respectivas variáveis.

As informações fornecidas pelas tabelas anteriores acumulam elementos básicos, que irão auxiliar no entendimento dos achados da pesquisa e na interpretação dos resultados que serão gerados por análises estatísticas, envolvendo Inferência Estatística e Análise Multivariada.

Os resultados a seguir são referentes a algumas medidas descritivas das variáveis quantitativas coletadas nas unidades amostrais, das variáveis geradas e das variáveis fornecidas pelos órgãos responsáveis pela distribuição de água e energia elétrica.

Tabela 23: Medidas descritivas das variáveis de consumo (água e energia elétrica).

Variável	Conjunto Habitacional	Média	Desvio Padrão	Mínimo	Máximo
SAE (X_1)	Geral	15,2310	6,1920	3,2667	33,8958
	Sem Sistema	17,4830	6,3605	4,7500	33,8958
	Com Sistema	13,2309	5,3064	3,2667	31,1333
CPFL (X_2)	Geral	145,6501	57,5546	40,9167	351,4286
	Sem Sistema	166,5853	59,3886	45,8333	351,4286
	Com Sistema	127,0563	49,0334	40,9167	334,9286

Na Tabela 23, foram apresentadas algumas medidas descritivas para se ter uma análise primária das informações das variáveis de consumo. Diferenças entre os consumos de água e energia elétrica comparando os dois conjuntos habitacionais podem ser observadas e serão corroboradas com a utilização de um teste estatístico aplicado para verificar a existência de diferenças significativas entre o conjunto de variáveis de consumo (água e energia elétrica) dos dois conjuntos habitacionais e entre o conjunto de variáveis socioeconômicos. Os resultados serão apresentados no subcapítulo 4.2.

Tabela 24: Medidas descritivas das variáveis socioeconômicas.

Variável	Conjunto Habitacional	Média	Desvio Padrão	Mínimo	Máximo
NMOR/NCOM (X_3)	Geral	0,7782	0,2992	0,2000	2,2500
	Sem Sistema	0,7790	0,3171	0,2000	2,2500
	Com Sistema	0,7775	0,2835	0,2000	1,6000
REND/NMOR (X_4)	Geral	421,3377	253,6533	102,00	1541,00
	Sem Sistema	362,6099	169,9104	102,00	855,00
	Com Sistema	473,4973	300,7009	128,00	1541,00
IDADE (X_5)	Geral	29,5111	12,4035	9,1429	73,5000
	Sem Sistema	32,0413	13,2031	9,1429	73,5000
	Com Sistema	27,2639	11,2200	10,2857	69,5000
ISEXO (X_6)	Geral	0,0601	0,3876	-1,0000	1,0000
	Sem Sistema	0,0942	0,3878	-0,6667	1,0000
	Com Sistema	0,0298	0,3862	-1,0000	1,0000
ESTUDO (X_7)	Geral	7,0394	2,2011	0,0000	13,0000
	Sem Sistema	6,6951	2,3014	0,0000	12,0000
	Com Sistema	7,3451	2,0679	2,0000	13,0000
IDEPEND (X_8)	Geral	0,3522	0,2324	0,0000	1,0000
	Sem Sistema	0,3386	0,2295	0,0000	1,0000
	Com Sistema	0,3642	0,2351	0,0000	1,0000
IMP (X_9)	Geral	0,6211	0,2369	0,0000	1,0000
	Sem Sistema	0,6322	0,2293	0,0000	1,0000
	Com Sistema	0,6113	0,2437	0,0000	1,0000
QUANTELE (X_{10})	Geral	11,0836	3,9991	3,0000	28,0000
	Sem Sistema	10,7556	4,1076	3,0000	28,0000
	Com Sistema	11,3750	3,8906	4,0000	26,0000
ICA (X_{11})	Geral	0,3004	0,0664	0,1786	0,6667
	Sem Sistema	0,3025	0,0588	0,1786	0,4167
	Com Sistema	0,2985	0,0727	0,1923	0,6667

Na Tabela 24, observam-se as medidas descritivas das características socioeconômicas, mas existe a necessidade de teste estatístico para

confirmar ou não a existência de diferenças entre os dois conjuntos habitacionais, que terão os resultados apresentados no subcapítulo 4.2.

A seguir, são apresentados resultados da aplicação de técnicas estatísticas multivariadas no conjunto de variáveis apresentadas anteriormente, X_1 a X_{11}.

4.2 Teste T^2 de Hotelling para comparação de dois vetores de médias e respectivo intervalo de confiança.

Para verificar a existência de diferenças entre os vetores de respostas dos dois conjuntos habitacionais, foi realizado o teste T^2 de Hotelling, tanto para o conjunto de variáveis de consumo quanto para as variáveis socioeconômicas.

As médias das variáveis foram as apresentadas nas Tabelas 23 e 24, para cada respectivo conjunto habitacional.

As matrizes de variâncias e covariâncias amostrais dos conjuntos habitacionais (Conjunto Habitacional Sem Sistema = S_1 e Conjunto Habitacional Com Sistema = S_2), do conjunto de variáveis de consumo e do conjunto de variáveis socioeconômicas são dadas por:

1) Consumo (Sem Sistema)

$$\begin{bmatrix} 40{,}4556 & 213{,}5549 \\ 213{,}5549 & 3527{,}0109 \end{bmatrix};$$

2) Socioeconômica (Sem Sistema)

$$\begin{bmatrix} 0{,}1006 & -29{,}4030 & -2{,}1266 & -0{,}0083 & -0{,}0485 & 0{,}0110 & 0{,}0050 & -0{,}1504 & 0{,}0005 \\ -29{,}4030 & 28869{,}5319 & 492{,}6858 & 1{,}2569 & 103{,}9269 & -6{,}0308 & -5{,}2477 & 280{,}3293 & -0{,}9280 \\ -2{,}1266 & 492{,}6858 & 174{,}3221 & -0{,}1508 & -11{,}2368 & -0{,}1479 & 0{,}0656 & -9{,}8008 & 0{,}0905 \\ -0{,}0083 & 1{,}2569 & -0{,}1508 & 0{,}1504 & -0{,}1127 & 0{,}0058 & 0{,}0189 & 0{,}0366 & -0{,}0004 \\ -0{,}0485 & 103{,}9269 & -11{,}2368 & -0{,}1127 & 5{,}2964 & -0{,}1804 & -0{,}2048 & 3{,}0080 & -0{,}0171 \\ 0{,}0110 & -6{,}0308 & -0{,}1479 & 0{,}0058 & -0{,}1804 & 0{,}0527 & 0{,}0277 & -0{,}0909 & 0{,}0002 \\ 0{,}0050 & -5{,}2477 & 0{,}0656 & 0{,}0189 & -0{,}2048 & 0{,}0277 & 0{,}0526 & -0{,}1687 & 0{,}0018 \\ -0{,}1504 & 280{,}3293 & -9{,}8008 & 0{,}0366 & 3{,}0080 & -0{,}0909 & -0{,}1687 & 16{,}8726 & -0{,}0391 \\ 0{,}0005 & -0{,}9280 & 0{,}0905 & -0{,}0004 & -0{,}0171 & 0{,}0002 & 0{,}0018 & -0{,}0391 & 0{,}0035 \end{bmatrix}.$$

3) Consumo (Com Sistema)

$$\begin{bmatrix} 28{,}1578 & 155{,}6414 \\ 155{,}6414 & 2404{,}2770 \end{bmatrix};$$

4) Socioeconômico (Com Sistema)

$$\begin{bmatrix} 0{,}0803 & -50{,}3569 & -1{,}6294 & 0{,}0033 & -0{,}1784 & 0{,}0283 & 0{,}0178 & 0{,}0309 & -0{,}0013 \\ -50{,}3569 & 90421{,}0469 & 685{,}3589 & -11{,}0253 & 301{,}7233 & -25{,}4898 & -19{,}9140 & 402{,}2701 & -2{,}8796 \\ -1{,}6294 & 685{,}3589 & 125{,}8880 & -0{,}2869 & -1{,}4913 & -0{,}8637 & -0{,}5091 & -6{,}1177 & 0{,}0941 \\ 0{,}0033 & -11{,}0253 & -0{,}2869 & 0{,}1491 & 0{,}0955 & 0{,}0085 & 0{,}0028 & -0{,}0844 & 0{,}0021 \\ -0{,}1784 & 301{,}7233 & -1{,}4913 & 0{,}0955 & 4{,}2763 & -0{,}2370 & -0{,}1804 & 2{,}1785 & -0{,}0169 \\ 0{,}0283 & -25{,}4898 & -0{,}8637 & 0{,}0085 & -0{,}2370 & 0{,}0553 & 0{,}0281 & -0{,}1146 & 0{,}0006 \\ 0{,}0178 & -19{,}9140 & -0{,}5091 & 0{,}0028 & -0{,}1804 & 0{,}0281 & 0{,}0594 & -0{,}0969 & -0{,}0015 \\ 0{,}0309 & 402{,}2701 & -6{,}1177 & -0{,}0844 & 2{,}1785 & -0{,}1146 & -0{,}0969 & 15{,}1366 & -0{,}0742 \\ -0{,}0013 & -2{,}8796 & 0{,}0941 & 0{,}0021 & -0{,}0169 & 0{,}0006 & -0{,}0015 & -0{,}0742 & 0{,}0053 \end{bmatrix}.$$

Antes da realização do teste T^2 de Hotelling, foi feito o gráfico *Q-Q plot* e o teste de Box.

Os gráficos *Q-Q plot* são apresentados a seguir para o conjunto de variáveis, que são considerados nesta pesquisa para cada respectivo conjunto habitacional. O gráfico *Q-Q plot* auxiliará na verificação do ajuste da normalidade multivariada aos dados amostrais.

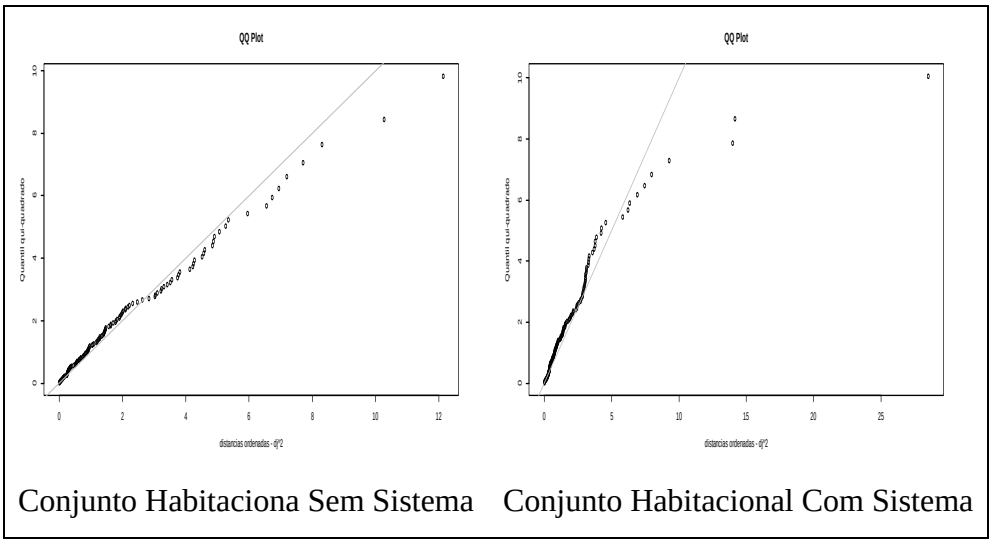
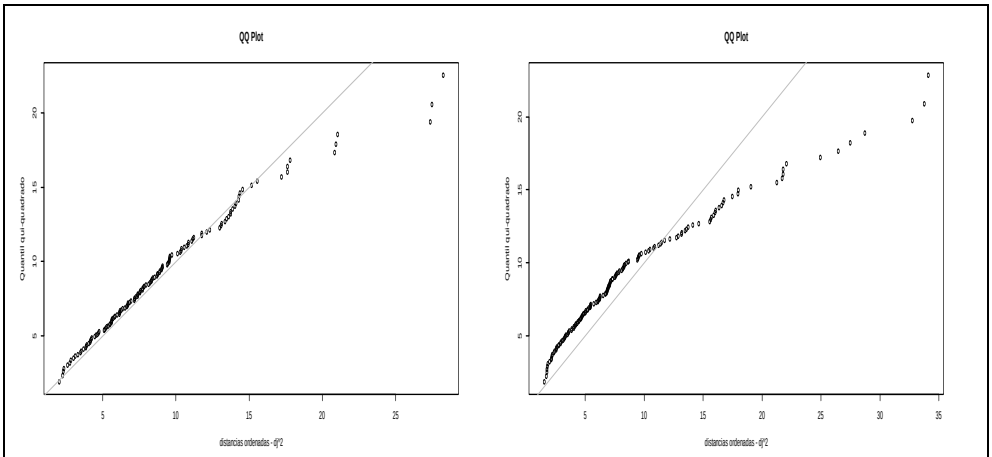


Figura 1: Q-Q plot para o conjunto de variáveis de consumo dos Conjuntos Habitacionais Com e Sem Sistema.



Conjunto Habitaciona Sem Sistema Conjunto Habitacional Com Sistema

Figura 2: Q-Q plot para o conjunto de variáveis socioeconômicas dos Conjuntos Habitacionais Com e Sem Sistema.

Os gráficos nas Figuras 1 e 2 apresentam nuvens de pontos distribuídos próximos da reta. Observa-se que o ajuste não foi totalmente sucedido, pois existem alguns pontos distantes da reta, tanto na Figura 1 como na Figura 2. Frente à quantidade alta de unidades amostrais, esta fuga moderada de alguns valores pode-se sobrestar o ajuste à normalidade dos dados (MARDIA, 1974).

O teste de Box (JOHNSON; WICHERN, 2007) foi utilizado para testar a homogeneidade das matrizes de variâncias e covariâncias. Nos dois casos a serem testados (conjunto de variáveis de consumo e conjunto de variáveis socioeconômicas), os resultados (valor $p < 0,01$, variáveis de consumo: $p = 0,009$, variáveis socioeconômicas: $p = 0,000$) indicam a rejeição da hipótese de homogeneidade considerando um nível de significância de 5%, nas duas situações consideradas.

Portanto, o teste de comparação dos vetores de médias foi utilizado empregando a estatística T^2 de Hotelling para matrizes de variâncias e covariâncias diferentes, cujos resultados dos níveis descritivos obtidos são inferiores a 0,001, ou seja, a hipótese da igualdade dos dois vetores de médias é rejeitada nos dois casos, variáveis de consumo e variáveis socioeconômicas. Existe diferença significativa entre os vetores de médias do conjunto de variáveis de consumo dos Conjuntos Habitacionais Sem Sistema e Com Sistema, o mesmo ocorrendo para os vetores de médias do conjunto de variáveis socioeconômicas.

Segundo Naspolini et al. (2008), Ijumba et al. (2009), Naspolini et al. (2010), Bodach e Hamhaber (2010), Silva et al. (2011), Naspolini e Rüther (2011), o sistema de aquecimento solar de água instalado em habitações sociais, baixa renda,

proporciona, nestas residências, a redução do consumo de energia elétrica (principalmente no horário de pico), economia financeira e, para as distribuidoras de energia, redução de investimentos de expansão, confiabilidade do sistema, otimização da distribuição de energia e também a eficiência energética na distribuição.

Complementou-se a análise com a construção dos intervalos de confiança simultâneos de Hotelling para as diferenças entre todos os pares de elementos dos vetores de médias considerando o nível de 95% de confiança para cada uma das variáveis da pesquisa, Tabela 25.

Tabela 25: Diferença entre as médias dos conjuntos habitacionais, limite inferior e limite superior do intervalo de confiança simultâneo de Hotelling.

Resultado do Teste Multivariado	Variável	Sem Sistema – Com Sistema $\left(\bar{x}_1^{\{l\}} - \bar{x}_2^{\{l\}}\right)$	Limite Inferior	Limite Superior
Consumo ($l = c$)	SAE (X_1)	4,2521	2,5476	5,9567
	CPFL (X_2)	39,5290	23,6765	55,3816
Socioeconômica ($l = s$)	NMOR/NCOM (X_3)	0,0015	-0,1453	0,1483
	REND/NMOR (X_4) x 10^2	-1,1089	-2,2786	0,0609
	IDADE (X_5)	4,7774	-1,2109	10,7657
	ISEXO (X_6)	0,0644	-0,1239	0,2527
	ESTUDO (X_7)	-0,6500	-1,7176	0,4176
	IDEPEND (X_8)	-0,0255	-0,1384	0,0874
	IMP (X_9)	0,0209	-0,0940	0,1358
	QUANTELE (X_{10})	-0,6194	-2,5687	1,3298
	ICA (X_{11})	0,0040	-0,0280	0,0359

Nas Figuras 3 e 4, são apresentados os intervalos de confiança simultâneos de Hotelling, para os respectivos conjuntos de variáveis.

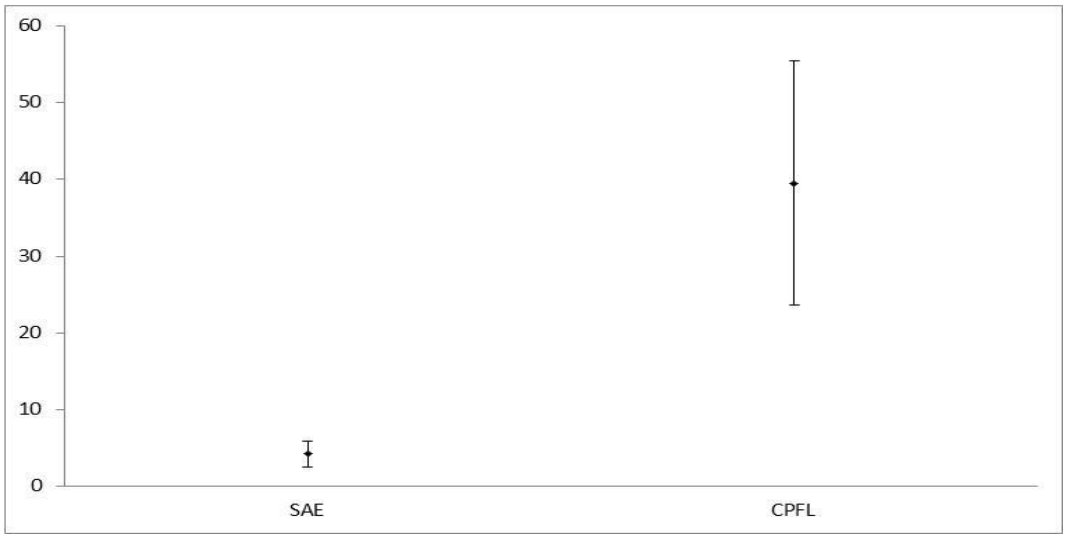


Figura 3: Intervalos de confiança simultâneos de Hotelling para a diferença entre as médias dos conjuntos habitacionais, variáveis de consumo.

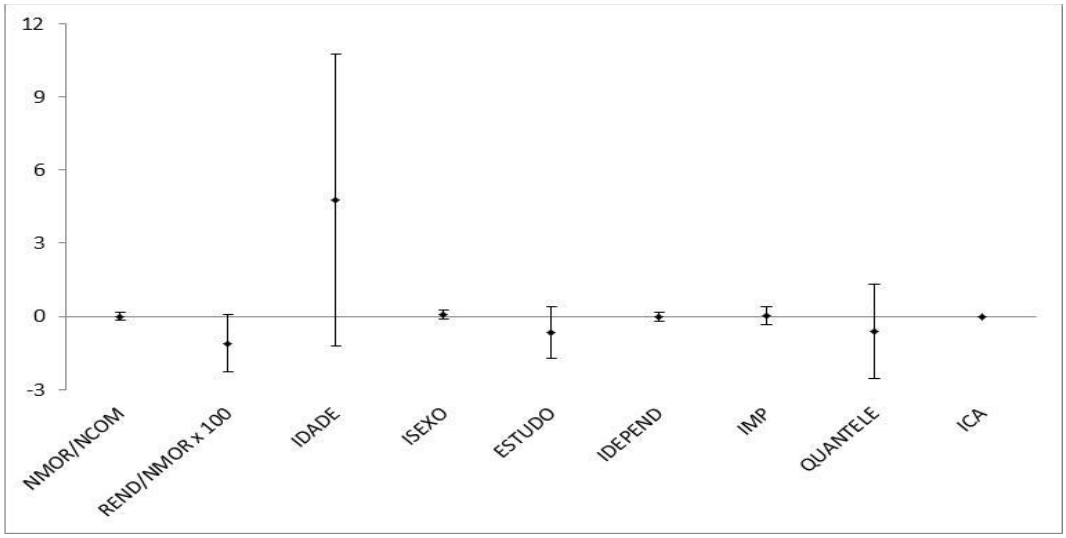


Figura 4: Intervalos de confiança simultâneos de Hotelling para a diferença entre as médias dos conjuntos habitacionais, variáveis socioeconômicas.

Nota-se, na Tabela 25 e na Figura 3, para as variáveis de consumo, que os intervalos de confiança simultâneos de Hotelling não apresentam o valor zero dentro dos limites, ou seja, indicação da diferença significativa entre os conjuntos habitacionais para os consumos de energia elétrica e água ($p < 0,05$). Os valores das médias das variáveis do Conjunto Habitacional Sem Sistema são maiores em relação ao Conjunto Habitacional Com Sistema, mostrando maior consumo no conjunto Sem Sistema em ambas variáveis.

Para as variáveis socioeconômicas, observa-se, na Tabela 25 e Figura 4, que embora o teste multivariado indique diferença significativa entre os conjuntos, os intervalos de confiança simultâneos de Hotelling apresentam o valor zero dentro dos limites dos intervalos, indicando que não foi possível encontrar diferença significativa entre os conjuntos habitacionais (ausência de diferença, $p > 0,05$). Todavia, alguns limites dos intervalos apresentam assimetria muito forte, indicando, quanto ao entendimento prático, a importância dessas variáveis na significância do teste T^2 de Hotelling. Destacam-se, neste contexto, as variáveis REND/NMOR (X_4), IDADE (X_5) e ESTUDO (X_7), sendo que os valores apresentados para o Conjunto Habitacional Sem Sistema são menores em relação ao Com Sistema, na primeira e terceira variáveis mencionadas (assimetria negativa), e os valores da segunda variável são maiores para o Conjunto Habitacional Sem Sistema em relação ao Com Sistema (assimetria positiva).

A ausência de diferença entre os conjuntos habitacionais dos intervalos de confiança não acompanha o encontrado na análise global para todas as

variáveis, justificando-se pelo fato do teste T^2 de Hotelling ter maior poder estatístico que os intervalos de confiança simultâneos. Além disso, o fato de o coeficiente de variação ser alto para as variáveis socioeconômicas (REND/NMOR (X_4), ISEXO (X_6) e IDEPEND (X_8)) pode ter contribuído.

4.2 Correlação canônica

São apresentados, nesta seção, os resultados da correlação canônica entre as variáveis de consumo (água e energia elétrica) e as variáveis socioeconômicas.

Antes dos resultados das correlações canônicas, serão expostas as correlações simples (correlação linear de Pearson) entre as variáveis de cada conjunto e as respectivas variâncias e covariâncias para cada um dos conjuntos habitacionais.

Na Tabela 26, são descritas as correlações entre as variáveis do conjunto de consumo e do conjunto socioeconômicas do Conjunto Habitacional Sem Sistema, bem como suas variâncias e covariâncias. Na diagonal, são apresentados os valores das variâncias das variáveis; abaixo da diagonal, os valores das covariâncias e, acima da diagonal, os valores da correlação linear de Pearson, com destaque em negrito para os resultados significativos.

Tabela 26: Correlação linear, variâncias e covariâncias das variáveis de consumo e as variáveis socioeconômicas, Conjunto Habitacional Sem Sistema.

	(X_1)	(X_2)	(X_3)	(X_4)	(X_5)	(X_6)	(X_7)	(X_8)	(X_9)	(X_{10})	(X_{11})
(X_1)	40,46	**0,57**	**0,39**	**-0,23**	**-0,22**	-0,09	-0,03	-0,05	0,09	0,16	0,15
(X_2)	213,55	3527,01	**0,30**	-0,13	**-0,27**	-0,06	0,14	-0,04	-0,12	**0,40**	**0,19**
(X_3)	0,78	5,74	0,10	**-0,55**	**-0,51**	-0,07	-0,07	0,15	0,07	-0,12	0,03
(X_4)	-251,02	-1326,02	-29,40	28869,53	**0,22**	0,02	**0,27**	-0,15	-0,13	**0,40**	-0,09
(X_5)	-18,72	-210,95	-2,13	492,69	174,32	-0,03	**-0,37**	-0,05	0,02	**-0,18**	0,12
(X_6)	-0,21	-1,32	-0,01	1,26	-0,15	0,15	-0,13	0,07	**0,21**	0,02	-0,02
(X_7)	-0,40	18,63	-0,05	103,93	-11,24	-0,11	5,30	**-0,34**	**-0,39**	**0,32**	-0,13
(X_8)	-0,07	-0,48	0,01	-6,03	-0,15	0,01	-0,18	0,05	**0,53**	-0,10	0,02
(X_9)	0,13	-1,58	0,00	-5,25	0,07	0,02	-0,20	0,03	0,05	**-0,18**	0,13
(X_{10})	4,31	97,03	-0,15	280,33	-9,80	0,04	3,01	-0,10	-0,17	16,87	-0,16
(X_{11})	0,06	0,65	0,00	-0,93	0,09	-0,00	-0,02	0,00	0,00	-0,04	0,00

X_1 = SAE; X_2 = CPFL; X_3 = NMOR/NCOM; X_4 = REND/NMOR; X_5 = IDADE; X_6 = ISEXO; X_7 = ESTUDO; X_8 = IDEPEND; X_9 = IMP; X_{10} = QUANTELE; X_{11} = ICA.

As estatísticas encontradas para o Conjunto Habitacional Sem Sistema, apresentadas na Tabela 26, também foram calculadas para o Conjunto Habitacional Com Sistema e estão apresentadas na Tabela 27.

Tabela 27: Correlação linear, variâncias e covariâncias das variáveis de consumo e as variáveis socioeconômicas, Conjunto Habitacional Com Sistema.

	(X_1)	(X_2)	(X_3)	(X_4)	(X_5)	(X_6)	(X_7)	(X_8)	(X_9)	(X_{10})	(X_{11})
(X_1)	28,16	**0,60**	**0,37**	**-0,17**	**-0,19**	-0,11	-0,06	**0,17**	0,09	**0,17**	-0,10
(X_2)	155,64	2404,28	**0,34**	**-0,27**	-0,16	-0,09	-0,01	0,02	0,04	**0,17**	-0,06
(X_3)	0,55	4,74	0,08	**-0,59**	**-0,51**	0,03	**-0,30**	**0,42**	**0,26**	0,03	-0,06
(X_4)	-265,10	-3969,07	-50,36	90421,05	**0,20**	-0,09	**0,49**	**-0,36**	**-0,27**	**0,34**	-0,13
(X_5)	-11,52	-86,79	-1,63	685,36	125,89	-0,07	-0,06	**-0,33**	**-0,19**	-0,14	0,12
(X_6)	-0,24	-1,62	0,00	-11,03	-0,29	0,15	0,12	0,09	0,03	-0,06	0,07
(X_7)	-0,66	-0,93	-0,18	301,72	-1,49	0,10	4,28	**-0,49**	**-0,36**	**0,27**	-0,11
(X_8)	0,21	0,28	0,03	-25,49	-0,86	0,01	-0,24	0,06	**0,49**	-0,13	0,03
(X_9)	0,12	0,48	0,02	-19,91	-0,51	0,00	-0,18	0,03	0,06	-0,10	-0,08
(X_{10})	3,53	31,84	0,03	402,27	-6,12	-0,08	2,18	-0,11	-0,10	15,14	**-0,26**
(X_{11})	-0,04	-0,21	-0,00	-2,88	0,09	0,00	-0,02	0,00	-0.00	-0,07	0,01

X_1 = SAE; X_2 = CPFL; X_3 = NMOR/NCOM; X_4 = REND/NMOR; X_5 = IDADE; X_6 = ISEXO; X_7 = ESTUDO; X_8 = IDEPEND; X_9 = IMP; X_{10} = QUANTELE; X_{11} = ICA.

As correlações canônicas entre as variáveis de consumo e socioeconômicas são apresentadas na Tabela 28.

Tabela 28: Correlação canônica das variáveis de consumo e as variáveis socioeconômicas dos conjuntos habitacionais.

Bairro	Variável Canônica	Autovalor	Correlação Canônica	Estatística do Teste	gl	Valor p
Bairro 1 Conj. Hab. Sem Sistema	(U_1,V_1)	0,3940	0,6277	82,3164	18	0,00000
Bairro 2 Conj. Hab. Com Sistema	(U_1,V_1)	0,2258	0,4752	47,4542	18	0,00018

Os dois conjuntos habitacionais mostram resultados de significância para a primeira correlação canônica (Tabela 28). A existência de associação significativa ($p < 0,0001$) entre o conjunto de variáveis de consumo, água e energia elétrica, e o conjunto de variáveis socioeconômicas nos respectivos conjuntos habitacionais, indica que quanto maior apresentaram-se os valores de consumo, maior também os valores das variáveis socioeconômicas.

A Figura 5 mostra o diagrama de dispersão para o primeiro par de variáveis canônicas de cada conjunto habitacional, indicando a associação linear entre eles apontada pela correlação canônica existente entre os conjuntos de variáveis estudadas.

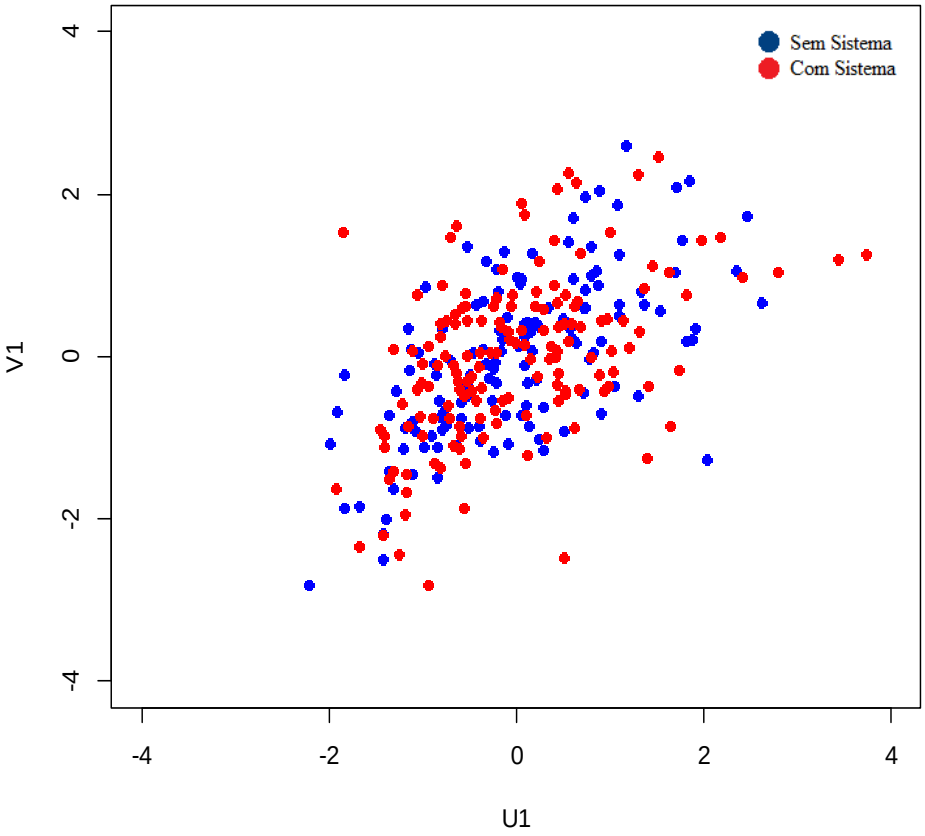


Figura 5: Gráfico de dispersão do primeiro par de variáveis canônicas dos Conjuntos Habitacionais Com Sistema e Sem Sistema.

São apresentados posteriormente os resultados dos coeficientes, associações e associações cruzadas dos conjuntos de variáveis de cada respectivo conjunto habitacional.

Tabela 29: Coeficientes canônicos e suas associações com as variáveis de consumo e as variáveis socioeconômicas no conjunto habitacional Sem Sistema.

Conjunto de Variáveis	Variável	Coeficiente	Associação	Associação Cruzada
Consumo	SAE	-0,3561	0,7539	0,4732
	CPFL	-0,9344	0,9681	0,6077
Socioeconômica	NMOR/NCOM	-0,4357	0,5736	0,3600
	REND/NMOR	0,2922	-0,2791	-0,1752
	IDADE	0,0457	-0,4492	-0,2819
	ISEXO	0,0730	-0,1143	-0,0717
	ESTUDO	0,0024	0,1598	0,1003
	IDEPEND	0,0978	-0,0675	-0,0424
	IMP	0,0155	-0,1056	-0,0663
	QUANTELE	-0,7514	0,5845	0,3669
	ICA	-0,3780	0,3105	0,1949

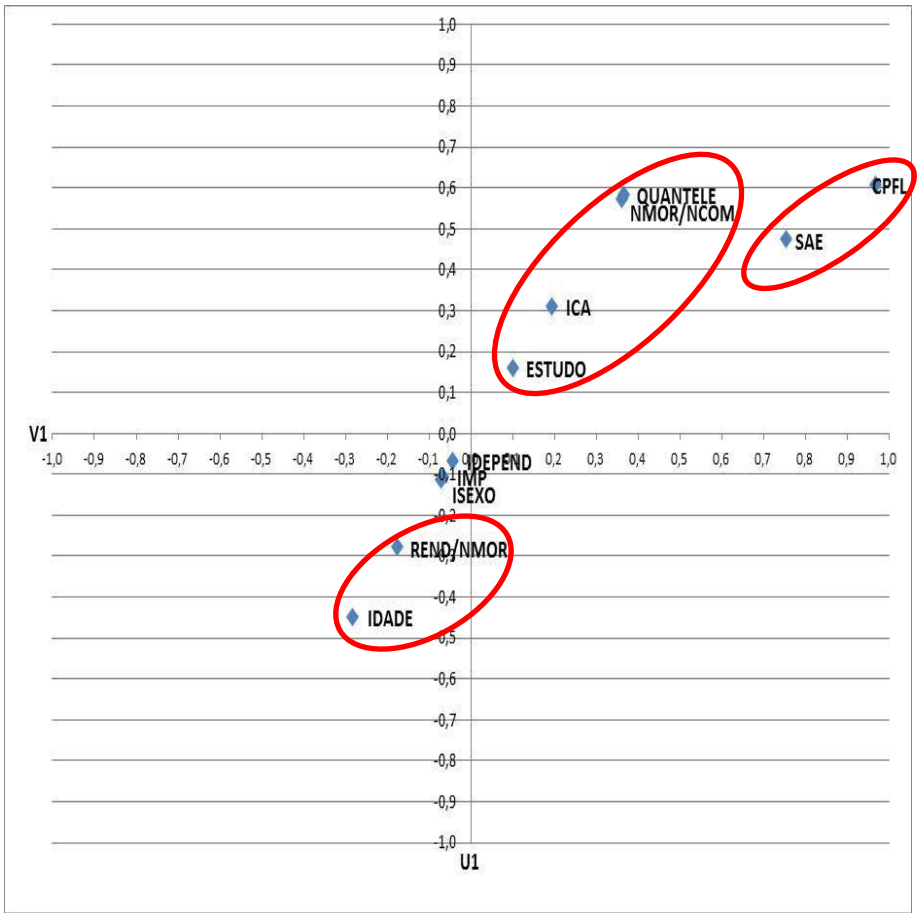


Figura 6: Gráfico de dispersão da associação e associação cruzada de cada respectiva variável, Conjunto Habitacional Sem Sistema.

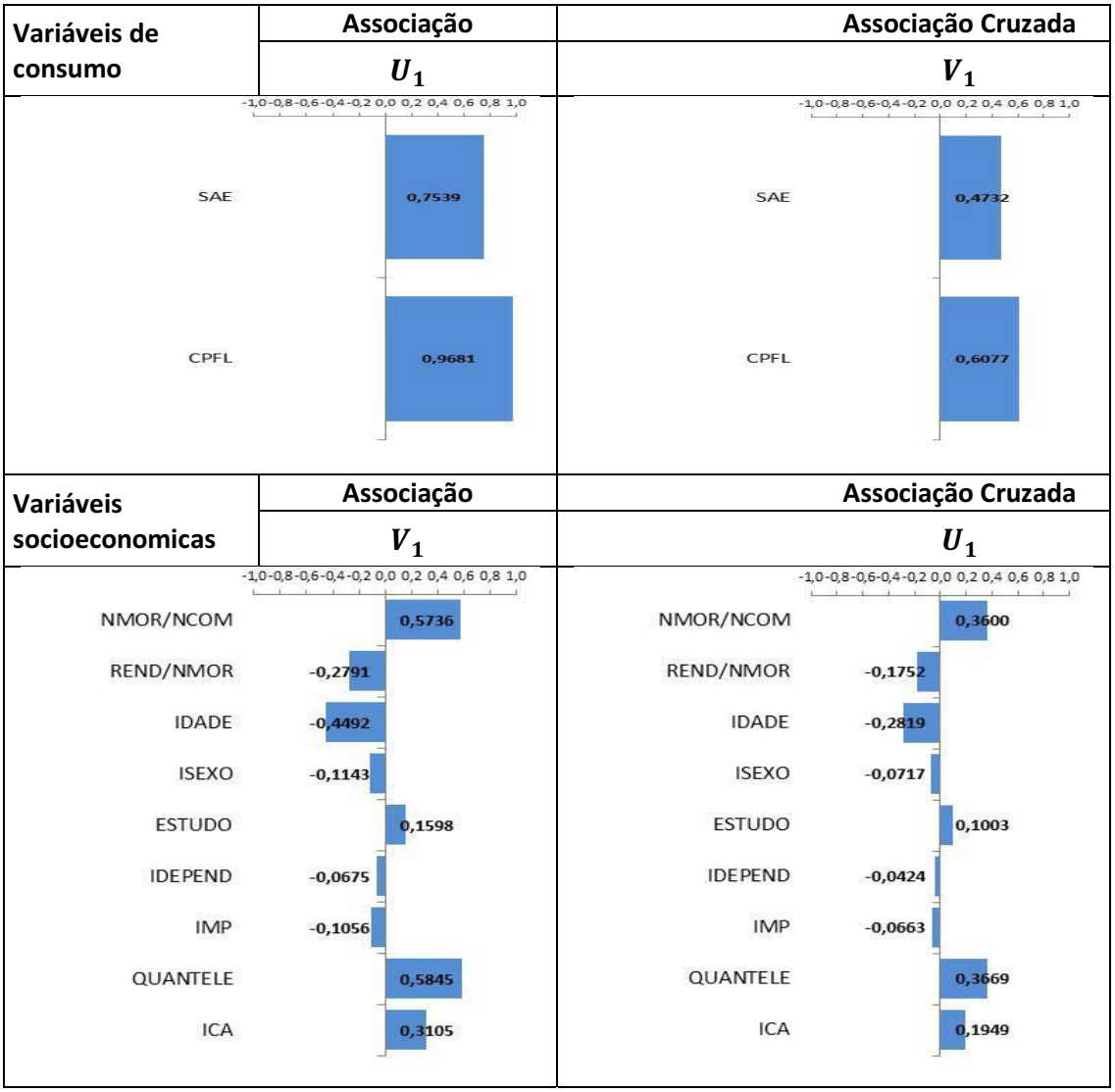


Figura 7: Gráfico das correlações das variáveis com os componentes canônicos, Conjunto Habitacional Sem Sistema.

Para o conjunto habitacional sem o sistema de aquecimento solar, as variáveis de consumo (água e energia elétrica) se somam na contribuição do componente consumo, U_1. Já no componente social, V_1, a contribuição se dá pelo contraste da soma das variáveis média de moradores por cômodo (NMOR/NCOM), média dos anos de estudo (ESTUDO), quantidade de itens que consomem energia (QUANTELE) e o indicador de equipamentos que consomem maior quantidade de energia (ICA) em relação à renda *per capita* (REND/NMOR) e idade média dos moradores (IDADE), como pode ser observado na Tabela 29.

As associações entre as variáveis de consumo e o componente canônico U_1 (consumo), mostram que valores maiores destas variáveis associam-se com

valores maiores do componente consumo. O componente consumo, U_1, apresenta valores elevados associados com valores maiores tanto para o consumo de água (SAE) quanto para o consumo de energia elétrica (CPFL).

As associações entre as variáveis sociais e o componente canônico V_1 (social) informam que valores elevados das variáveis NMOR/NCOM, ESTUDO, QUANTELE e ICA e valores menores das variáveis REND/NMOR e IDADE associam-se com valores maiores do componente social. O componente social, V_1, apresenta valor elevado associado com valores maiores para as variáveis NMOR/NCOM, ESTUDO, QUANTELE e ICA e menores valores para as variáveis REND/NMOR e IDADE, ou seja, o componente social assume valores altos associados ao maior contraste entre estes dois conjuntos de variáveis.

As associações cruzadas, entre as variáveis de consumo e o componente canônico V_1 (social), mostram-se positivamente associadas, ou seja, valores altos das variáveis de consumo associam-se com valores altos do componente social. Para que o componente social apresente valores elevados, a média de moradores por cômodo, média dos anos de estudo, quantidade de itens que consomem energia e o indicador de equipamentos que consomem maior quantidade de energia devem assumir valores altos e valores menores devem ser apresentados nas variáveis renda *per capita* e idade média dos moradores; destacando-se a variável de consumo de energia elétrica, CPFL, que apresenta maior associação com o componente social.

As associações cruzadas, entre as variáveis sociais e o componente canônico U_1 (consumo), indicam que o componente de consumo está associado diretamente (correlação positiva) com NMOR/NCOM, QUANTELE e ICA e inversamente (correlação negativa) com REND/NMOR e IDADE, ou seja, a associação cruzada sugere que valores elevados para as variáveis média de moradores por cômodo, quantidade de itens que consomem energia e o indicador de equipamentos que consomem maior quantidade de energia; e valores menores para renda *per capita* e idade média dos moradores associam-se com valores maiores do componente consumo. Para o componente consumo assumir valores altos existe a necessidade de valores altos do consumo de água e do consumo de energia elétrica.

Tabela 30: Coeficientes canônicos e suas associações com as variáveis de consumo e as variáveis socioeconômicas, no Conjunto Habitacional Com Sistema.

Conjunto de Variáveis	Variável	Coeficiente	Associação	Associação Cruzada
Consumo	SAE	-0,5578	0,8458	0,4019
	CPFL	-0,8300	0,9335	0,4436
Socioeconômica	NMOR/NCOM	-0,6479	0,8213	0,3903
	REND/NMOR	0,4706	-0,5336	-0,2536
	IDADE	-0,0987	-0,4033	-0,1916
	ISEXO	0,2815	-0,2277	-0,1082
	ESTUDO	-0,2721	-0,0694	-0,0330
	IDEPEND	0,0232	0,1916	0,0911
	IMP	-0,0230	0,1454	0,0691
	QUANTELE	-0,4398	0,3947	0,1876
	ICA	0,0368	-0,1789	-0,0850

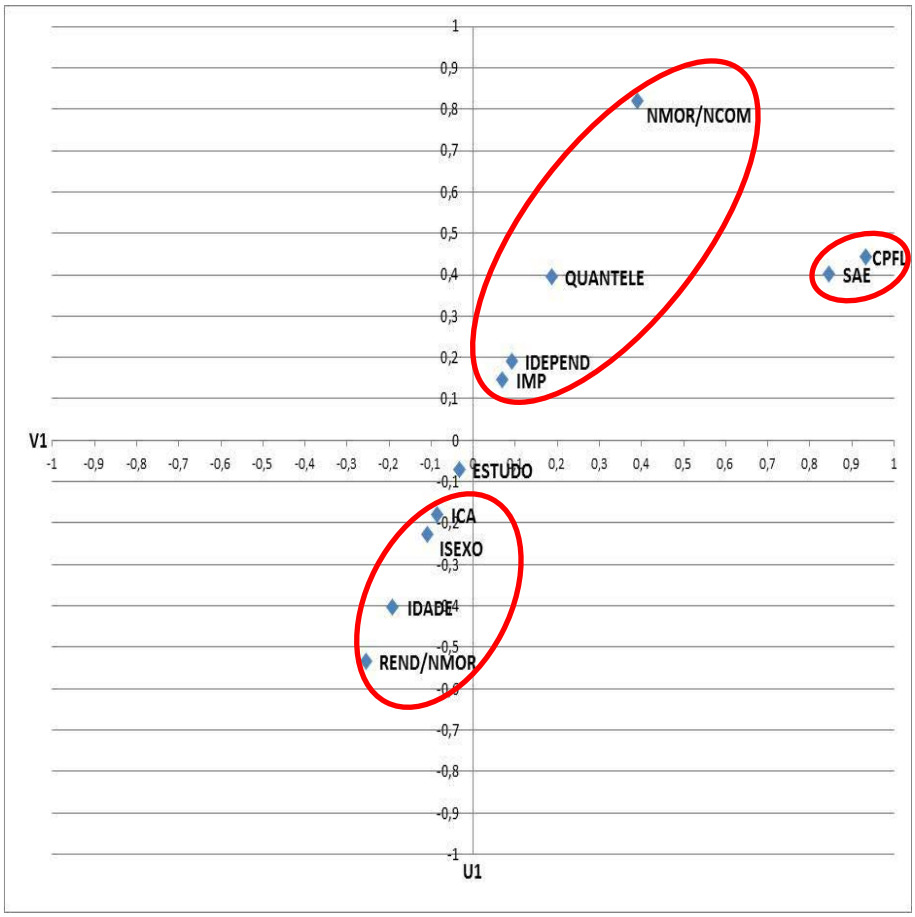


Figura 8: Gráfico de dispersão da associação e associação cruzada de cada respectiva variável, Conjunto Habitacional Com Sistema.

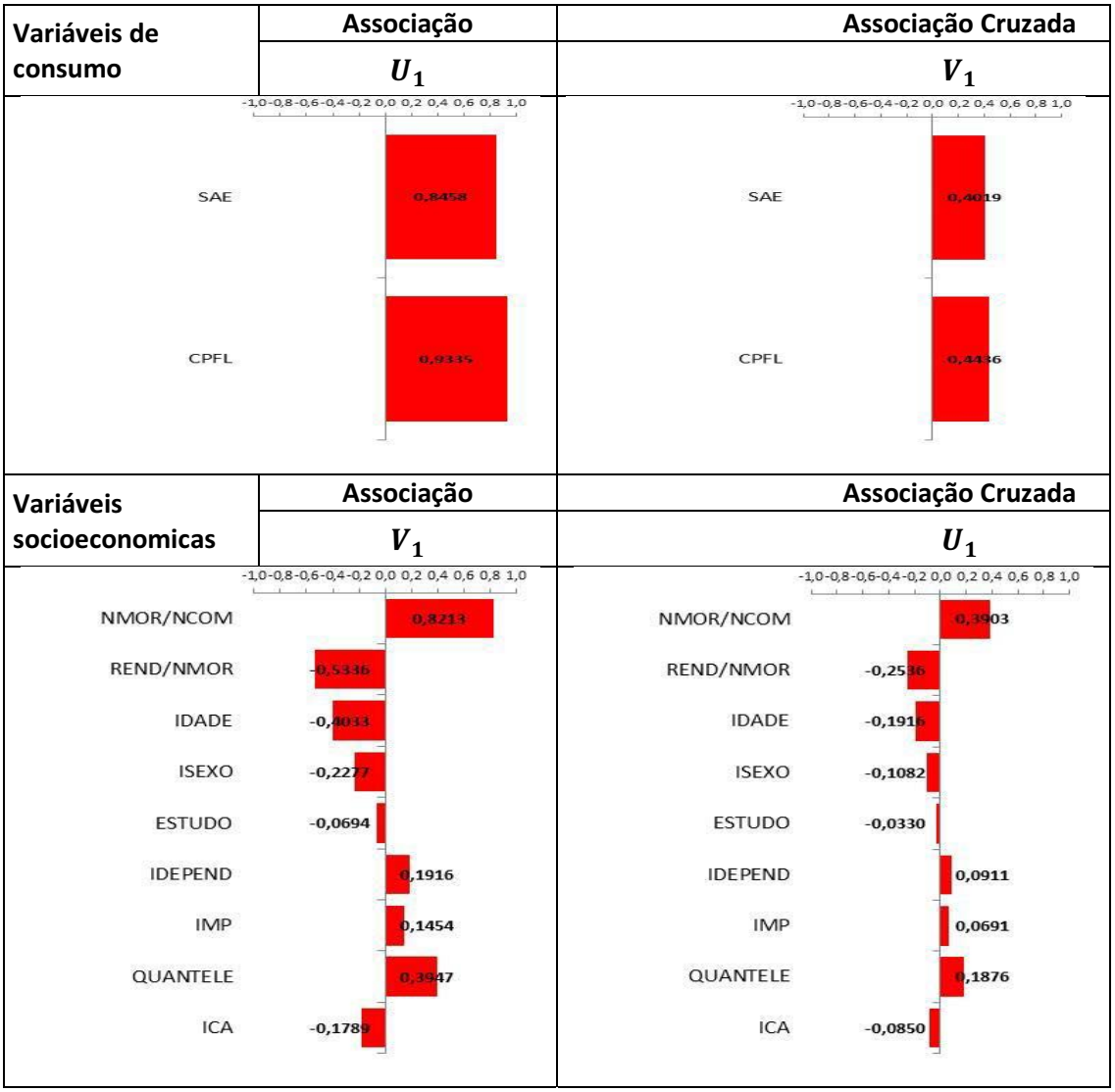


Figura 9: Gráfico das correlações das variáveis com os componentes canônicos, Conjunto Habitacional Com Sistema.

No conjunto habitacional com o sistema de aquecimento solar, as variáveis de consumo (água e energia elétrica) se somam para contribuir no componente consumo U_1. No componente social V_1, a contribuição se dá pelo contraste da soma das variáveis número médio de moradores por de cômodo (NMOR/NCOM), indicador de dependência na residência (IDEPEND), indicador de maior permanência na residência (IMP) e quantidade de itens que consomem energia (QUANTELE) em relação às variáveis renda *per capita* (REND/NMOR), idade média dos moradores (IDADE), indicador da variação da quantidade de sexo por unidade amostral (ISEXO) e indicador de equipamentos que consomem maior quantidade de energia elétrica (ICA), Tabela 30, Figuras 9 e 10.

Pode ser observado que as associações entre as variáveis de consumo e o componente canônico U_1 (consumo) indicam a existência de uma associação positiva. Valores maiores, tanto para o consumo de água (SAE) quanto para o consumo de energia elétrica (CPFL), estão associados com valores elevados do componente consumo, U_1.

As associações entre as variáveis socioeconômicas e o componente canônico V_1 (social) mostram que valores maiores para as variáveis média de moradores por cômodo (NMOR/NCOM), indicador de dependência na residência (IDEPEND), indicador de maior permanência na residência (IMP) e quantidade de itens na residência que consomem energia (QUANTELE) e valores menores das variáveis renda média por morador (REND/NMOR), idade média dos moradores (IDADE), indicador da variação da quantidade de sexo por unidade amostral (ISEXO) e indicador de equipamentos que consomem maior quantidade de energia elétrica (ICA), estão associados com valores elevados do componente social. O componente social, V_1, apresenta valores elevados associados à ocorrência de valores altos para as variáveis NMOR/NCOM, IDEPEND, IMP e QUANTELE e valores menores para as variáveis REND/NMOR, IDADE, ISEXO e ICA, ou seja, o valor do componente social será maior quando o contraste entre estes conjuntos de variáveis apresentar maior diferença.

As associações cruzadas entre as variáveis de consumo e o componente canônico V_1 (social) informam que estão associadas diretamente, valores maiores das variáveis de consumo associam-se com valores maiores do componente social. Pelos valores da associação cruzada serem semelhantes, indicam equilíbrio na correlação das variáveis com o componente. Para que o componente social, V_1, apresente valores altos, os valores das variáveis NMOR/NCOM, IDEPEND, IMP e QUANTELE devem ser altos e os valores das variáveis REND/NMOR, IDADE, ISEXO e ICA devem ser baixos.

As associações cruzadas entre as variáveis sociais e a variável canônica U_1 (consumo) mostram que o componente consumo está associado diretamente com NMOR/NCOM e QUANTELE e inversamente com REND/NMOR e IDADE, ou seja, as associações sugerem que valores altos para o componente consumo associam-se de maneira positiva com as variáveis média de moradores por cômodo (NMOR/NCOM) e quantidade de itens que consomem energia (QUANTELE) e de maneira negativa com os valores das variáveis renda *per capita* por unidade amostral (REND/NMOR) e idade média

dos moradores da residência (IDADE). Valores maiores do componente consumo associam-se com elevados consumos de água e energia elétrica.

Os resultados encontrados indicam que o consumo de água e energia elétrica apresentarão valores baixos se houver associação positiva ou negativa com algumas variáveis estudadas. As características ou variáveis do componente canônico V_1, socioeconômico, que contribuem para o consumo baixo de água e energia elétrica são apresentadas na Tabela 31 para os respectivos conjuntos habitacionais.

As setas vermelhas e verdes indicam a contribuição das respectivas variáveis para o componente consumo, U_1. Valores baixos para o componente consumo estão associados tanto a valores baixos (seta vermelha) quanto a valores altos (seta verde) de algumas variáveis.

Tabela 31: Indicativo das características estudas para consumo baixo de água e energia nos conjuntos habitacionais.

	Conjunto Habitacional		
Variável	Sem Sistema	Com Sistema	Influência
NMOR/NCOM	↓	↓	Ambos
REND/NMOR	↑	↑	Ambos
IDADE	↑	↑	Ambos
ISEXO		↑	Com Sistema
ESTUDO	↓		Sem Sistema
IDEPEND		↓	Com Sistema
IMP		↓	Com Sistema
QUANTELE	↓	↓	Ambos
ICA	↓	↑	Ambos Inversamente

Consumo de água e energia elétrica estão associados positiva ou negativamente às variáveis socioeconômicas da pesquisa. Pelos resultados apresentados, observam-se que existem associações comuns aos dois conjuntos habitacionais. Em ambos os conjuntos, valores baixos de consumo (água e energia elétrica) estão associados a menor quantidade de moradores, menor quantidade de equipamentos que consomem energia, maior renda *per capita* e maior idade média.

Características distintas quanto às associações também são observadas nos respectivos conjuntos habitacionais. Para o Conjunto Habitacional Sem Sistema, valores baixos de consumo estão associados a menor média de estudos e menor quantidade de equipamentos que consomem maior quantidade de energia. Já para o

Conjunto Habitacional Com Sistema, valores baixos de consumo estão associados a menor indicador de dependência na residência, menor indicador de maior permanência na residência e maior indicador da variação da quantidade de sexo por unidade amostral.

Segundo Naspolini et al. (2008), a economia anual de energia elétrica proporcionada pela agregação da energia solar térmica ao banho é significativa e contribui para manter o consumidor na faixa de baixa renda, desonerando significativamente parcela tão carente da população. A significativa redução da demanda na ponta, proporcionada pela agregação da energia solar térmica ao banho da população de baixa renda, é um forte indicador técnico que permite vislumbrar elevados custos evitados com sistemas de transmissão e de distribuição de energia elétrica no país.

Na presente pesquisa, foram analisados conjuntamente, por meio de técnicas multivariadas, variáveis de consumo e socioeconômicas para obter resultados e informações mais detalhados sobre a redução de consumo de energia elétrica e a associação de variáveis que contribuem para este fato. Silva et al. (2011) relatou que os principais impactos gerados com a incorporação do aquecedor solar nestas habitações foram, primeiramente, a economia financeira, que interfere diretamente na qualidade de vida da população residente, causando um segundo impacto, que é o aumento do conforto sob aspectos como o aumento do tempo de banho. Ressalta-se que o uso do aquecimento solar não pode aumentar o consumo de água, sugerindo-se que, em trabalhos futuros, a análise do consumo de energia seja sempre feita em conjunto com a do consumo de água.

Os autores estudados não mencionam outras variáveis que influenciam na economia de energia elétrica em habitações de interesse social com sistema de aquecimento solar de água, a economia é atribuída somente ao sistema instalado.

A pesquisa realizada deixou claro que o aquecimento solar de água tem o potencial de reduzir substancialmente a procura e a perda na transmissão em um sistema de energia; sendo argumentos para a utilização de sistemas de aquecimento solar de água, a fim de reduzir a perda na transmissão de energia, especialmente durante os períodos de demanda de pico. Portanto, recomenda-se, como um meio de reduzir esses fenômenos, especialmente durante os períodos de demanda de pico, permitindo que o sistema de distribuição de energia trabalhe de maneira mais eficiente. A redução da tensão no sistema também pode prolongar a vida útil de equipamentos de transmissão e adiar a necessidade de investimentos para novas geradoras de energia e infraestrutura de transmissão (IJUMBA et al., 2009).

Segundo Bodach e Hamhaber (2010), existe um grande potencial para melhorar a eficiência energética nas residências sociais no Brasil. O estudo mostra que medidas energéticas eficientes são necessárias para melhorar o conforto térmico e reduzir o consumo de energia. A implementação de propostas de medidas energéticas eficientes contribuiria substancialmente para os objetivos do país, diminuindo o crescimento rápido do consumo de energia e, assim, reduzindo as necessidades de investimento no setor de geração de eletricidade. A análise econômica deste estudo mostra que a energia mais eficiente nas residências sociais melhoraria a renda da população mais pobre, devido à redução nas despesas de energia. O aprimoramento da eficiência energética em residências sociais poderia controlar o mercado em transformação por meio de mais sustentabilidade no setor de construção.

4.3 Análise Fatorial

São apresentados, neste subcapítulo, os resultados da análise fatorial com as variáveis originais de consumo e socioeconômicas das 287 unidades amostrais a partir da matriz de correlação dos dados, sem distinção quanto ao conjunto habitacional.

Na Tabela 32, encontram-se, na diagonal e acima da diagonal, os valores das correlações e, abaixo da diagonal, seus respectivos níveis descritivos (valores p).

Tabela 32: Correlação e nível descritivo das 11 variáveis consideradas na análise fatorial.

Variável	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}
X_1	**1.000**	0.629	0.355	-0.239	-0.127	-0.065	-0.090	0.035	0.099	0.130	0.025
X_2	0,000	**1.000**	0.302	-0.261	-0.137	-0.037	0.014	-0.024	-0.021	0.244	0.061
X_3	0,000	0,000	**1.000**	-0.532	-0.500	-0.018	-0.178	0.289	0.166	-0.045	-0.023
X_4	0,000	0,000	0,000	**1.000**	0.146	-0.071	0.405	-0.264	-0.226	0.356	-0.123
X_5	0,016	0,010	0,000	0,007	**1.000**	-0.030	-0.250	-0.193	-0.073	-0.173	0.117
X_6	0.138	0.264	0.379	0.114	0.305	**1.000**	-0.015	0.075	0.116	-0.024	0.037
X_7	0.064	0.405	0.001	0.000	0.000	0.400	**1.000**	-0.402	-0.373	0.302	-0.120
X_8	0.276	0.340	0.000	0.000	0.001	0.102	0.000	**1.000**	0.503	-0.107	0.026
X_9	0,046	0.364	0,002	0,000	0.110	0,025	0,000	0,000	**1.000**	-0.141	0.004
X_{10}	0,014	0,000	0.222	0,000	0,002	0.342	0,000	0,036	0,008	**1.000**	-0.219
X_{11}	0.336	0.153	0.351	0,019	0,023	0.265	0,021	0.333	0.470	0,000	**1.000**

X_1 = SAE; X_2 = CPFL; X_3 = NMOR/NCOM; X_4 = REND/NMOR; X_5 = IDADE; X_6 = ISEXO; X_7 = ESTUDO; X_8 = IDEPEND; X_9 = IMP; X_{10} = QUANTELE; X_{11} = ICA.

Algumas correlações apresentam o nível descritivo menor do que 0,05; indicando associação linear significativa entre as variáveis. Outras apresentam nível descritivo maior do que 0,05 (valor p em vermelho), indicando que não há associação linear significativa entre as variáveis (correlações em vermelho). Nota-se que a variável X_6 = ISEXO apresenta associação positiva significativa somente com a variável X_9 = IMP.

O tamanho amostral considerado possibilita que, apesar de algumas correlações apresentarem magnitudes expressas em valores baixos, estas associações são significativas, como observado na estrutura triangular abaixo da diagonal principal.

A estatística *KMO* e o Teste de Esfericidade de Bartlett são apresentados na Tabela 33. A estatística *KMO*, com valor 0,631, avalia se as variáveis originais confirmam a utilização da técnica de análise fatorial e indica o grau de explicação dos dados a partir dos fatores encontrados na técnica. Como o valor da estatística foi maior que 0,50; significa que os fatores encontrados na análise fatorial conseguem descrever satisfatoriamente as variações dos dados das variáveis originais. O resultado do teste de esfericidade de Bartlett ($p < 0,001$) indicou a associação significativa entre as variáveis para aplicação da análise fatorial (CORRAR; PAULO; DIAS FILHO, 2009).

Tabela 33: Estatística Kaiser-Meyer-Olkin e Teste de Esferacidade de Bartlett.

Estatística *Kaiser-Meyer-Olkin* de adequação da amostra		0,631
Teste de Esfericidade de Bartlett	Aprox. Qui-Quadrado	786,131
	Graus de Liberdade	55
	Valor p	0,000

Em síntese, os resultados 0,631 da estatística *KMO* considerado moderado e o nível descritivo α do teste de esfericidade de Bartlett ($p<0,001$) sugerem a associação significativa entre as variáveis, corroborando com a possibilidade da utilização da técnica.

Os autovalores, os percentuais da variância total explicada e os percentuais acumulados são apresentados na Tabela 34.

Tabela 34: Variância total explicada pelos fatores.

	Autovalores			Cargas da Soma de Quadrados da Extração			Cargas da Soma de Quadrados da Rotação		
Fatores	Total	% de Variância	% Acumulada	Total	% de Variância	% Acumulada	Total	% de Variância	% Acumulada
1	2,684	24,399	24,399	2,684	24,399	24,399	1,970	17,907	17,907
2	2,005	18,228	42,627	2,005	18,228	42,627	1,838	16,706	34,612
3	1,369	12,442	55,069	1,369	12,442	55,069	1,696	15,416	50,029
4	1,079	9,805	64,875	1,079	9,805	64,875	1,616	14,694	64,723
5	1,005	9,137	74,011	1,005	9,137	74,011	1,022	9,289	74,011
6	0,834	7,578	81,589						
7	0,584	5,308	86,897						
8	0,436	3,963	90,861						
9	0,412	3,750	94,611						
10	0,311	2,825	97,435						
11	0,282	2,565	100,000						

Os cinco fatores extraídos relativos aos autovalores maiores que a unidade explicam 74,011% da variância total, ou seja, os fatores conseguem explicar mais de 74% da variância total dos dados originais.

No Figura 11 é apresentado o *Scree Plot*, um dos critérios para definir o número de fatores que serão utilizados na técnica.

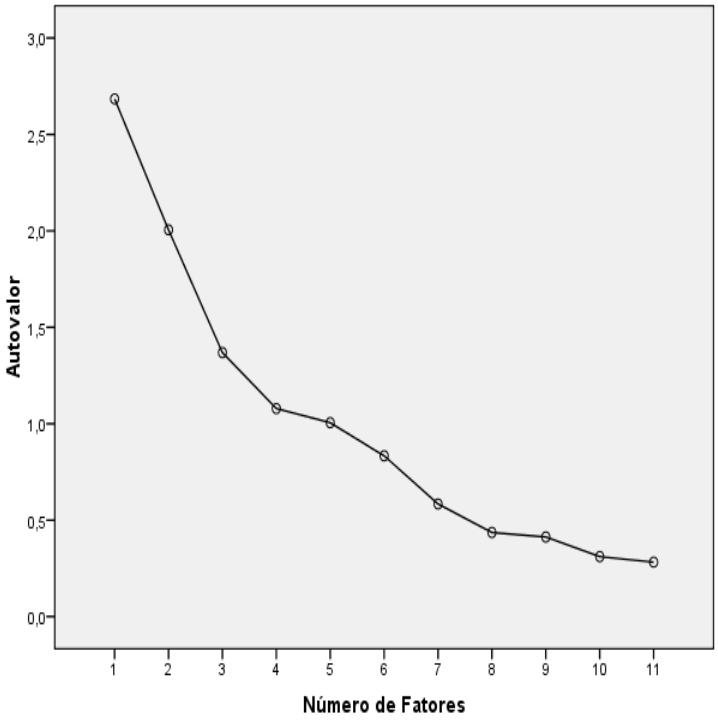


Figura 10: Gráfico *Scree Plot* das variáveis.

A Figura 11, *Scree Plot*, indica, por meio de inspeção visual, a utilização de cinco fatores, pois, a partir deste número de fatores a queda no gráfico começa a mostrar estabilidade, proporcionando pouca contribuição para a informação global e admitindo uma relação custo-benefício baixa se existir a entrada de novos fatores.

Na Tabela 35, são apresentados os resultados da *MSA* para cada variável; os valores indicam a explicação de cada variável, a partir dos fatores determinados na técnica, estes devem ser maiores do que 0,50 para que a variável não seja excluída, todas as variáveis consideradas apresentam valores acima de 0,50.

Tabela 35: Medida de adequação da amostra, MSA, para cada variável.

Variável	*MSA*
SAE (X_1)	0,628
CPFL (X_2)	0,571
NMOR/NCOM (X_3)	0,671
RENMED/NMOR (X_4)	0,672
IDADE (X_5)	0,505
ISEXO (X_6)	0,508
ESTUDO (X_7)	0,639
IDEPEND (X_8)	0,686
IMP (X_9)	0,704
QUANTELE (X_{10})	0,592
ICA (X_{11})	0,625

Na Tabela 36, são apresentadas as comunalidades de cada variável, ou seja, o poder de explicação da variável considerando todos os fatores extraídos.

Tabela 36: Comunalidades das variáveis consideradas na análise fatorial.

Variável	Valor Inicial	Valor após Extração dos Fatores
SAE (X_1)	1,000	0,765
CPFL (X_2)	1,000	0,812
NMOR/NCOM (X_3)	1,000	0,789
RENMED/NMOR (X_4)	1,000	0,724
IDADE (X_5)	1,000	0,826
ISEXO (X_6)	1,000	0,899
ESTUDO (X_7)	1,000	0,726
IDEPEND (X_8)	1,000	0,685
IMP (X_9)	1,000	0,685
QUANTELE (X_{10})	1,000	0,722
ICA (X_{11})	1,000	0,509

Os valores das comunalidades para as respectivas variáveis apresentadas na Tabela 36 são maiores do que 0,50; indicando considerável proporção da variância explicada pelos fatores para cada variável.

As cargas fatoriais das 11 variáveis para os cinco fatores extraídos na análise fatorial, utilizando a rotação Varimax, são apresentados na Tabela 37. Os resultados sem a rotação não são apresentados, pois não são conclusivos (Fator 4 não incorporou variáveis).

Tabela 37: Cargas fatoriais das variáveis para os cinco fatores extraídos (rotação Varimax).

	Fatores				
Variável	1	2	3	4	5
SAE (X_1)	0,088	**0,861**	0,108	-0,003	-0,067
CPFL (X_2)	-0,064	**0,891**	0,119	0,015	0,021
NMOR/NCOM (X_3)	0,214	0,315	**0,769**	-0,204	-0,102
RENMED/NMOR (X_4)	-0,308	-0,264	-0,425	**0,614**	0,044
IDADE (X_5)	0,038	-0,018	**-0,875**	-0,220	-0,098
ISEXO (X_6)	0,103	-0,047	0,024	-0,056	**0,939**
ESTUDO (X_7)	**-0,695**	-0,099	0,196	0,414	0,151
IDEPEND (X_8)	**0,786**	-0,054	0,247	0,001	0,055
IMP (X_9)	**0,816**	0,020	0,041	0,030	0,127
QUANTELE (X_{10})	-0,147	0,319	-0,007	**0,767**	0,103
ICA (X_{11})	-0,128	0,130	-0,172	**-0,621**	0,245

Observa-se, na Tabela 37, quais variáveis estão associadas a cada um dos fatores extraídos. O Fator 1 está associado as variáveis média de anos de estudos (ESTUDO (X_7)), indicador de idade dependente (IDEPEND (X_8)) e indicador de maior permanência (IMP (X_9)). O Fator 2 reuni as variáveis consumo médio de água (SAE (X_1)) e consumo médio de energia (CPFL (X_2)). No Fator 3, estão agrupadas as variáveis número médio de moradores por cômodo (NMOR/NCOM (X_3)) e média de idade (IDADE (X_5)). As variáveis renda *per capita* (RENMED/NMOR (X_4)), quantidade de itens na residência que consomem energia (QUANTELE (X_{10})) e o indicador de equipamentos que consomem maior quantidade de energia elétrica (ICA(X_{11})) associam-se ao Fator 4. A variável indicador da variação da quantidade de sexo por unidade amostral (ISEXO(X_6)) é associada ao Fator 5.

Pelo fato de existir cinco autovalores com valores acima de um e o gráfico *Scree Plot* indicar cinco fatores como melhor alternativa, definiu-se que o modelo da análise fatorial será composto por cinco fatores que acumulam 74,011% da variância total explicada no modelo fatorial, apresentados na Tabela 37.

No contexto da interpretação dos fatores estes serão denominados como:

- Fator 1: Permanência na Residência

  Fator responsável por 17,907% da variância explicada. É representado pelas variáveis média de anos de estudos (ESTUDO (X_7)), indicador de idade dependente (IDEPEND (X_8)) e indicador de maior permanência (IMP (X_9)); uma variável que representa a escolaridade média e duas variáveis associadas à permanência na residência. A primeira variável referente à permanência na residência, envolve moradores em faixas etárias que tendem a ficar maior tempo na residência e a segunda variável está relacionada ao número de moradores aos quais as ocupações informadas levam a crer que permanecem maior tempo na residência também. Valores altos de X_8 e X_9 e valores baixos de X_7 estão associados a valores altos do Fator 1, ou seja, valores altos para o Fator 1 estão associados à maior quantidade de moradores inativos, à maior quantidade de moradores que permanecem maior tempo na residência e a menor média de anos de estudos.

- Fator 2: Consumo

  Fator responsável por 16,706% da variância explicada. É representado pelas variáveis consumo médio de água (SAE (X_1)) e consumo médio de energia (CPFL (X_2)); as duas variáveis estão associadas ao consumo, a primeira informa sobre o consumo médio de água em m^3 e a segunda variável sobre o consumo médio de energia elétrica em kWh. Valores altos de X_1 e X_2 estão associados a valores altos para o Fator 2, ou seja, valores altos de consumo de água e de consumo de energia elétrica estão associados positivamente ao Fator 2.

- Fator 3: Estrutura Familiar

  Fator responsável por 15,416% da variância explicada. É representado pelas variáveis número médio de moradores por cômodo (NMOR/NCOM (X_3)) e idade média dos moradores (IDADE (X_5)); as duas variáveis estão associadas a informações familiares, a primeira variável indica o número médio de moradores por cômodo na residência (densidade por residência), valores altos indicam maior quantidade de moradores; a segunda variável informa sobre a idade média por morador na residência. Valores altos de X_3 e valores baixos de X_5 estão associados a valores altos para o Fator 3, ou seja, maior densidade e menor idade média por morador estão associados a maiores valores do Fator Estrutura Familiar.

- Fator 4: Poder de Compra

  Fator responsável por 14,694% da variância explicada. É representado pelas variáveis renda *per capita* (RENMED/NMOR (X_4)), quantidade de itens na residência que consomem energia (QUANTELE (X_{10})) e o indicador de equipamentos que consomem maior quantidade de energia elétrica (ICA(X_{11})); as três variáveis estão associadas ao poder de compra, a primeira variável indica a renda média por morador, a segunda informa sobre a quantidade de itens que consomem energia na residência e a última representa a quantidade de equipamentos que consomem maior quantidade de energia elétrica. Valores altos de X_4 e X_{10} e valores baixos de X_{11} estão associados a valores altos para o Fator 4, ou seja, o Fator Poder de Compra associa-se positivamente com maior renda *per capita* e maior quantidade de itens que consomem energia e inversamente ao indicador de equipamentos que consomem maior quantidade de energia elétrica.

- Fator 5: Sexo

  Fator responsável por 9,289% da variância explicada. É representado pela variável indicador da variação da quantidade de sexo por unidade amostral (ISEXO(X_6)). O Fator Sexo associa-se positivamente com valores do indicador da variação da quantidade de sexo por residência, ou seja, valores altos do indicador da variação da quantidade de sexo (valores altos indicam maior quantidade do sexo feminino) estão associados positivamente a valores altos para o Fator Sexo.

Com o intuito de avaliar o valor dos fatores encontrados na significância da comparação dos dois conjuntos habitacionais, foi realizado um teste estatístico para comparar os escores calculados usando as cargas fatoriais. Para a comparação foi utilizado o teste não paramétrico Mann-Whitney, pois não foi verificada normalidade dos escores em todos os fatores extraídos.

Tabela 38: Mediana, mínimo, máximo, nível descritivo e resultado do teste não paramétrico Mann-Whitney.

Fator	Conjunto Habitacional	Escores			Mann-Whitney	
		mediana	mínimo	máximo	Valor p	Resultado
Permanência na Residência	Sem Sistema	0,0597	-2,4949	3,2909	0,508	Não há evidência da diferença
	Com Sistema	-0,0046	-2,7284	2,9701		
Consumo	Sem Sistema	0,3228	-1,6463	2,9385	0,000	Há evidência da diferença
	Com Sistema	-0,5079	-1,9273	2,5342		
Estrutura Familiar	Sem Sistema	-0,0463	-2,9467	2,3827	0,004	Há evidência da diferença
	Com Sistema	0,3170	-2,7941	2,1094		
Poder de Compra	Sem Sistema	-0,2161	-1,8264	3,6089	0,002	Há evidência da diferença
	Com Sistema	0,1188	-3,5128	3,5069		
Sexo	Sem Sistema	-0,0333	-2,1990	2,6786	0,619	Não há evidência da diferença
	Com Sistema	-0,1044	-2,6205	2,8757		

Os resultados na Tabela 38 indicam que não foi possível verificar diferença significativa entre os conjuntos habitacionais para os Fatores Permanência na Residência e Sexo. Os Fatores Consumo, Estrutura Familiar e Poder de Compra são significativamente diferentes ($p < 0,01$), para os quais os escores sugerem no Fator Consumo que o conjunto habitacional sem sistema apresenta valores maiores do que no conjunto habitacional com sistema. Considerando os Fatores Estrutura Familiar e Poder de Compra, os escores indicam que o conjunto habitacional sem sistema apresenta valores menores do que no com sistema.

Os escores indicam que no conjunto habitacional com sistema, valores baixos no Fator Consumo estão associados ao menor consumo de energia e água;

para o Fator Estrutura Familiar, valores altos neste fator estão associados ao maior número de moradores por cômodo e menor idade média por residência; e valores altos do Fator Poder de Compra estão associados a uma maior renda per capita, maior quantidade de equipamentos que consomem energia e menor indicador de equipamentos que consomem maior quantidade de energia elétrica.

5 CONCLUSÃO

Neste trabalho, por meio das técnicas estatísticas aplicadas, observa-se a existência de diferença significativa entre os conjuntos habitacionais, tanto no conjunto de variáveis de consumo (água e energia elétrica) como para as variáveis socioeconômicas. O conjunto de variáveis de consumo (água e energia elétrica) indica que o Conjunto Habitacional Sem Sistema apresenta valores maiores para os consumos de água e energia elétrica em relação ao Conjunto Habitacional Com Sistema, corroborado pelos intervalos de confiança simultâneos de Hotelling.

Para o conjunto de variáveis socioeconômicas, a comparação dos vetores médios do Conjunto Habitacional Sem Sistema em relação ao Conjunto Habitacional Com Sistema de aquecimento apresentou diferenças significativas pelo teste estatístico multivariado, todavia os intervalos simultâneos não confirmam o teste, justificando-se pelos motivos mencionados. Pelo fato do teste T^2 de Hotelling apresentar diferença, ou seja, o conjunto de variáveis socioeconômicas é distinto entre os conjuntos habitacionais, existem indícios que características socioeconômicas possam estar associadas, diretamente ou inversamente, ao consumo de água e energia elétrica nos respectivos conjuntos habitacionais.

Constata-se que existe positividade das correlações canônicas, ou seja, componente de consumo (conjunto de variáveis de consumo) associado com

componente socioeconômico (conjunto de variáveis socioeconômicas). Porém a interpretação dessa positividade é diferenciada nos conjuntos habitacionais, não existe um comportamento único de associação. Para o Conjunto Habitacional Sem Sistema, valores altos do componente de consumo associam-se com valores altos do componente socioeconômico e valores altos do componente socioeconômico associam-se positivamente com as variáveis número médio de moradores por cômodo (NMOR/NCOM), média dos anos de estudo (ESTUDO), quantidade de itens que consomem energia (QUANTELE) e indicador de equipamentos que consomem maior quantidade de energia (ICA), contrastando com renda média por morador (REND/NMOR) e idade média (IDADE). Ou seja, o consumo baixo nas residências sem sistema de aquecimento solar está associado ao baixo número de moradores por cômodo, pequena média dos anos de estudo, pequena quantidade de itens que consomem energia elétrica, baixo índice de aparelhos que consomem maior quantidade de energia elétrica, maior renda per capita e maior idade média.

Para o Conjunto Habitacional Com Sistema, valores altos do componente de consumo associam-se com valores altos do componente socioeconômico e valores altos do componente socioeconômico associam-se positivamente com as variáveis número médio de moradores por cômodo (NMOR/NCOM), indicador de idade dependente (IDEPEND), indicador de maior permanência (IMP) e quantidade de itens que consomem energia (QUANTELE), contrastando com renda *per capita* (REND/NCOM), idade média dos moradores (IDADE), indicador da variação da quantidade de sexo (ISEXO) e indicador de equipamentos que consomem maior quantidade de energia (ICA). Ou seja, o consumo baixo nas residências com sistema de aquecimento solar está associado ao baixo número de moradores por cômodo, baixo índice de idade dependente, baixo índice de maior permanência, pequena quantidade de itens que consomem energia elétrica, maior renda per capita, maior idade média, maior indicador da variação da quantidade de sexo e maior índice de aparelhos que consomem maior quantidade de energia elétrica.

O fato de um conjunto habitacional apresentar o sistema de aquecimento solar de água e outro não, não modifica a associação entre o conjunto de variáveis de consumo e as variáveis socioeconômicas. Caso o conjunto habitacional possua ou não o sistema de aquecimento solar, o componente de consumo está associado diretamente ao componente socioeconômico, mas as variáveis socioeconômicas que constituem a associação se modificam.

Logo, nota-se que as características para entendimento do componente de consumo associado com componente socioeconômico nos conjuntos habitacionais são diferentes. No Conjunto Habitacional Sem Sistema, os fatores distintos que contribuem para o componente de consumo são a média de anos de estudo e o índice de equipamentos que consomem maior quantidade de energia. No Conjunto Habitacional Com Sistema, os fatores distintos que contribuem para o componente de consumo são indicador da variação da quantidade de sexo, o número de moradores em idade dependente, o número de moradores que permanecem maior tempo na residência e o índice de equipamentos que consomem maior quantidade de energia. Deste modo, mostra-se a existência de associações do consumo com características socioeconômicas diferentes nos respectivos conjuntos habitacionais.

Nas futuras habitações sociais, caso as famílias apresentem características como baixo número de moradores por cômodo, maior renda *per capita*, maior idade média, menor média dos anos de estudo, pequena quantidade de equipamentos que consomem energia e baixo índice de equipamentos que consomem maior quantidade de energia, as quais estão associadas ao baixo consumo de energia elétrica e água, torna-se praticamente sem eficiência instalar o sistema de aquecimento solar de água, ou seja, o sistema não produzirá eficiência energética para o consumo.

Se as famílias apresentarem características como baixo número de moradores por cômodo, maior renda *per capita*, maior idade média, maior indicador de variação de sexo, pequeno índice de idade dependente, baixo índice de maior permanência na residência, pequena quantidade de equipamentos que consomem energia e maior índice de equipamentos que consomem maior quantidade de energia, as quais estão associadas ao baixo consumo de energia elétrica e água, haverá a necessidade de instalar o sistema de aquecimento solar de água, pois o resultado será altamente eficiente.

O sistema de aquecimento solar de água é uma tecnologia que auxilia na redução reduz o consumo de energia elétrica, portanto sua instalação em residências de interesse social é recomendada para proporcionar benefícios aos usuários, com a economia de gastos com o consumo de energia elétrica. Além de proporcionar benefícios ao sistema, pois podem distribuir a energia que seria consumida pelas residências para outros usuários, favorecendo todo o processo de distribuição de energia. Mas a sua instalação ou não em habitações sociais podem ser associadas as características das famílias para que haja maior efetividade no seu uso.

As características dos moradores indicadas nesta pesquisa podem ser utilizadas como parâmetros para inscrição nos programas de habitação e políticas públicas. Desta forma, se terá uma perspectiva de como será o consumo de água e energia elétrica, prevendo a demanda de consumo. Além da possibilidade de verificar se a instalação do sistema de aquecimento solar de água será valido ou não para a respectiva família que ocupará a residência.

6 REFERÊNCIAS

ABEP – Associação Brasileira de Empresas de Pesquisa: Dados com base no levantamento socioeconômico de 2011. ABEP, 2012.

ABIKO, A. K. Introdução à gestão habitacional. São Paulo: Escola Politécnica da USP, 1995. 31 p. (Série Texto Técnico, TT/PCC/12).

ABIKO, A. K.; COELHO, L. O. Procedimentos de gestão de mutirão habitacional para população de baixa renda. **Ambiente Construído**, Porto Alegre, v. 4, n. 1, p. 7-20, jan./mar., 2004.

ANDERSON, T. W. **An introduction to multivariate statistical analysis**. 3. ed. New Jersey: John Wiley, 2003.

ARCHER, M. D.; BARBER, J. Photosynthesis and Photoconversion. In.: ARCHER, M.D.; Barber, J. **Molecular to Global Photosynthesis**. Londres: Imperial College Press, 2004. Disponível em < http://www.worldscibooks.com/etextbook/p218/p218_chap1_1.pdf>. Acesso em: 21/02/2010.

BAKKE, H. A.; LEITE, A. S. M.; SILVA, L. B. Estatística multivariada: Aplicação da análise fatorial na engenharia de produção. **Revista Gestão Industrial**, Ponta Grossa, v. 04, n. 04, p. 01-14, 2008.-

BARBOSA, L. T. et al. Associação entre qualidade da carne e características quantitativas de suínos por meio de correlação canônica. **Revista Brasileira de Saúde e Produção Animal**, v. 11, n. 4, p. 1150-1162, 2010.

BASSO, L. H. et al. Análise de um sistema de aquecimento de água para residências rurais, utilizando energia solar. **Engenharia Agrícola**, Jaboticabal, v. 30, n. 1, p. 14–21, 2010.

BEHRENS, J. T. Principles and procedures of exploratory data analysis. **Psychological Methods**, v. 2, n. 2, p. 131-160, 1997.

BODACH, S.; HAMHABER, J. Energy efficiency in social housing: Opportunities and barriers from a case study in Brazil. **Energy Policy**, v. 38, p. 7898-7910, 2010.

BOLFARINE, H.; BUSSAB, W. O. **Elementos de amostragem**. São Paulo: Editora Blucher, 2005.

BRASIL. Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de Atenção Básica. **SIAB: manual do sistema de Informação de Atenção Básica/ Ministério da Saúde**, Secretaria de Atenção à Saúde, Departamento de Atenção Básica. – 1. ed., 4.ª reimpr. – Brasília: Ministério da Saúde, 2003.

CABRAL, S. A. L. C. S. Análise multivariada dos fatores de risco anteparto para ocorrência de cesárea. **Revista Brasileira de Ginecologia e Obstetrícia**, v. 26, nº 3, p. 256-256, 2004.

CAVALCANTI, E. S. C.; LOPES, J. D. S.; LIMA, F. Z. **Energia solar para aquecimento de água.** Viçosa: Centro de Produções Técnicas, 2008.

CDHU – Companhia de Desenvolvimento Habitacional e Urbano do Estado de São Paulo. **Provisão de Moradias**. Disponível em: < http://www.cdhu.sp.gov.br/programas_habitacionais/provisao_de_moradias/parceria_com_municipios.asp>. Acesso em: 10 de janeiro de 2015.

COMETTA, E. **Energia solar**. São Paulo: Hemus Editora Limitada, 1982.

CORRAR, L. J.; PAULO, E.; DIAS FILHO, J. M. **Análise multivariada**: para os cursos de administração, ciências contáveis e economia. 1. ed. São Paulo: Atlas, 2009.

DIAKOULAKI, D.; ZERVOS, A.; SARAFIDIS, J.; MIRASGEDIS, S. Cost benefit analysis for solar water heting systems. **Energy Conversion and Management**, v. 42, p. 1727-1739, 2001.

FANTINELLI, J. T. **Análise da evolução de ações na difusão do aquecimento solar de água para habitações populares:** estudo de caso em Contagem-MG. 2006. 316 f. Tese (Doutorado em Planejamento de Sistemas Energéticos) – Faculdade de Engenharia Mecânica Universidade Estadual de Campinas, Campinas-SP, 2006.

FOSTER, R.; GHASSEMI, M.; COTA, A. **Solar energy**: renewable energy and the environment. Boca Raton: Taylor & Francis Group, 2010.

GOK, S. et al. Differentiation of Anatolian honey samples from different botanical origins by ATR-FTIR spectroscopy using multivariate analysis. **Food Chemistry**, v. 170, n. 1, p. 234-240, 2015.

HAIR JR, J. F.; et al. **Análise multivariada de dados**. Trad. Adonai Schlup Sant'Anna. 6. ed. Porto Alegre: Bookman, 2009.

IBGE – Instituto Brasileiro de Geografia e Estatística. **Censo demográfico 2010 – Cidades@**. Disponível em: <http://www.ibge.gov.br/cidadesat/topwindow.htm?1>. Acesso em: 29 de junho de 2012.

IJUMBA, K. P.; SEBITOSI, A. B.; PILLAY, P.; FOLLY, K. Impact of extensive residential solar water heating on power system losses. **Energy for Sustainable Development**, v. 13, p. 85-95, 2009.

JANGARELLI, M. Abordagem multivariada para endogamia e valor fenotípico utilizando diferentes estratégias de cruzamento. **Pesquisa Agropecuária Tropical**, v. 44, n. 1, p. 79-87, 2014.

JAUMOT, J.; ERITJA, R.; GARGALLO, R. Chemical equilibria studies using multivariate analysis methods. **Analytical and Bioanalytical Chemistry**, v. 399, n. 6, p. 1983-1997, 2011.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 6. ed. New Jersey: Prentice-Hall, 2007. 794 p.

KAVRADIN, S. T.; ÖZER, Z. C.; BOZCUK, H. Hope in people with cancer: a multivariate analysis from Turkey. **Journal of Advanced Nursing**, v. 69, n. 5, p. 1183-1196, 2013.

LARCHER, J. V. M. **Diretrizes visando a melhoria de projetos e soluções construtivas na expansão de habitações de interesse social**. 189f. Dissertação (Mestrado em Construção Civil) – Universidade Federal do Paraná, Curitiba-PR, 2005.

LUIZ, A. M. **Como aproveitar a energia solar**. São Paulo: Edgard Blücher Ltda, 1985.

MARDIA, K. V. Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. **Sankhyã B**, v. 36, 115-128.

MENEZES, J. P. C. et al. Correlação entre uso da terra e qualidade da água subterrânea. **Engenharia Sanitária e Ambiental**, v. 19, n. 2, p. 173-186, 2014.

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

NASPOLINI, H. F. ; MILITAO, H. ; RUTHER, R. . Avaliação do Potencial da Agregação da Energia Solar Térmica para Fins de Aquecimento da Água para o Banho Humano em Núcleos Habitacionais de Baixa Renda. In: II CBENS-Congresso Brasileiro de Energia Solar e III ISES-CLA-Conferência Latinoamericana da ISES, 2008, Florianópolis. **Anais do II Congresso Brasileiro de Energia Solar e III Conferência Latinoamericana da ISES**, p. 745-754, 2008.

_____. The role and benefits of solar water heating in the energy demands of low-income dwellings Brazil. **Energy Conversion and Management**, v. 51, p. 2835–2845, 2010.

NASPOLINI, H. F.; RÜTHER, R. The impacts of solar water heating in low-income households on the distribution utility's active, reactive and apparent power demands. **Solar Energy**, v. 85, p. 2023-2032, 2011.

NAZARÉ, B.; FONSECA, A.; CANAVARRO, M. C. Questionário de confiança parental: análise fatorial confirmatória numa amostra comunitária de casais. **Psicologia, Saúde & Doenças**, v. 14 (1), p. 23-37, 2013.

OLIVEIRA, L. F. C. et al. Potencial de redução do consumo de energia elétrica pela utilização de aquecedores solares no Estado de Goiás. **Engenharia Agrícola**, Jaboticabal, v. 28, n. 3, p. 406-416, 2008.

PALZ, W. **Energia solar e fontes alternativas**. Curitiba: Hemus SA, 2002.

PROCEL – Programa Nacional de Conservação de Energia Elétrica. **Dicas de economia de energia**. Disponível em: <http://www.procelinfo.com.br/main.asp?View={E6BC2A5F-E787-48AF-B485-439862B17000}>. Acesso em: 20 de janeiro de 2014.

PROTÁSIO, T. P. et al. Análise de correlação canônica entre características da madeira e do carvão vegetal de Eucalyptus. **Scientia Forestalis**, v. 40, n. 95, p. 317-326, Piracicaba, 2012.

RIGÃO, M. H. et al. Correlação canônica entre caracteres de tubérculos para seleção precoce de clones de batata. **Ciência Rural**, Santa Maria, v. 39, n. 8, p. 2347-2353, 2009.

RODRIGUES, D.; MATAJS, R. **Um banho de sol para o Brasil**: o que os aquecedores solares podem fazer pelo meio ambiente e sociedade. São Lourenço da Serra: Vitae Civilis, 2004.

SILVA, I. C. C.; CARVALHO, A. W. B.; TIBÚRCIO, T. M. S.; TIBIRIÇÁ, A. C. G. Uso do sistema de aquecimento solar de água na habitação de interesse social: estudo de caso do Programa Lares Habitação Popular da COHAB-MG em Itatiaiuçu. 2º Simpósio Brasileiro de Qualidade do Projeto no Ambiente Construído – SBQP (2011) e X Workshop Brasileiro de Gestão do Processo de Projeto na Construção de Edificio. **Anais do 2º Simpósio Brasileiro de Qualidade do Projeto no Ambiente Construído – SBQP (2011) e X Workshop Brasileiro de Gestão do Processo de Projeto na Construção de Edificio**, p. 698-707, Rio de Janeiro, 2011.

SEADE – Fundação Sistema Estadual de Análise de Dados. **Informações dos munícipios paulistas – IMP**. Disponível em: <http://www.seade.gov.br/produtos/imp/index.php?page=consulta&action=new&tema=1&tabs=1&aba=tabela1>. Acesso em: 25 de janeiro de 2013.

SEN, Z. Solar energy in progress and future research trends. **Progress in Energy and Combustion Science**, v. 30, p. 367–416, 2004.

SOLETROL, SOLARMAX – Aquecedor Solar de Água Compacto 200 litros – Manual do Usuário: Instalação e Utilização, 2013.

SPSS 17.0 for Windows [computer program] [cited 2008 Dez 25]. Statistical Package for Social Science (SPSS). Release Version 17.0.1. Chicago (IL): SPSS Incorporation; 2008.

THE COMPREHENSIVE R ArchiveNetwork. Disponível em: < http://cran.r-project.org/>. Acesso em: 10 mar. 2012.

TOLMASQUIM, M. T. (org.) **Fontes renováveis de energia no Brasil**. Rio de Janeiro: Interciência: CENERGIA, 2003.

TUKEY, J. W. **Exploratory Data Analysis: Past, Present, and Future**. Princenton: Princeton University, 1993.

VARELLA, F. K. O. M.; FANTINELLI, J. T.; PEREIRA, J. T. V. Coletores solares alternativos: tecnologia de baixo custo que contribui para redução do consumo de energia elétrica. In: IV Congresso Brasileiro de Planejamento Energético, 2004, Itajubá/MG. **Planejamento Energético**, 2004.

WALISIEWICZ, M. **Energia alternativa**: solar, eólica, hidrelétrica e de biocombustíveis; Trad. Elvira Serapicos. São Paulo: Publifolha, 2008.

WISSER, R. J. et al. Multivariate analysis of maize disease resistances suggests a pleiotropic genetic basis and implicates a *GST* gene. **Proceedings of the National Academy of Science of the United States of America**, v. 108 (3), p. 7339-7344, 2011.

ANEXO 1 – Ficha A

Ficha A - frente - modelo

FICHA A	SECRETARIA MUNICIPAL DE SAÚDE SISTEMA DE INFORMAÇÃO DA ATENÇÃO BÁSICA	UF __

ENDEREÇO	NÚMERO	BAIRRO	CEP
	___		_____ - ____

MUNICÍPIO	SEGMENTO	ÁREA	MICROÁREA	FAMÍLIA	DATA
_____	___	___	__	___	_____

CADASTRO DA FAMÍLIA

PESSOAS COM 15 ANOS OU MAIS NOME	DATA NASC.	IDADE	SEXO	ALFABETIZADO		OCUPAÇÃO	DOENÇA OU CONDIÇÃO REFERIDA (sigla)
				sim	não		

PESSOAS DE 0 A 14 ANOS NOME	DATA NASC.	IDADE	SEXO	FREQÜENTA A ESCOLA		OCUPAÇÃO	DOENÇA OU CONDIÇÃO REFERIDA (sigla)
				sim	não		

Fonte: (BRASIL, 2003).

Ficha A, *verso* - modelo

SITUAÇÃO DA MORADIA E SANEAMENTO

TIPO DE CASA	
Tijolo/Adobe	
Taipa revestida	
Taipa não revestida	
Madeira	
Material aproveitado	
Outro - Especificar:	
Número de cômodos / peças	
Energia elétrica	
DESTINO DO LIXO	
Coletado	
Queimado / Enterrado	
Céu aberto	

TRATAMENTO DA ÁGUA NO DOMICÍLIO	
Filtração	
Fervura	
Cloração	
Sem tratamento	
ÁBASTECIMENTO DE ÁGUA	
Rede geral	
Poço ou nascente	
Outros	
DESTINO DE FEZES E URINA	
Sistema de esgoto (rede geral)	
Fossa	
Céu aberto	

OUTRAS INFORMAÇÕES

Alguém da família possui Plano de Saúde?		Número de pessoas cobertas pelo Plano de Saúde	
Nome do Plano de Saúde			

EM CASO DE DOENÇA PROCURA	
Hospital	
Unidade de Saúde	
Benzedeira	
Farmácia	
Outros - Especificar:	
MEIOS DE COMUNICAÇÃO QUE MAIS UTILIZA	
Rádio	
Televisão	
Outros - Especificar:	

PARTICIPA DE GRUPOS COMUNITÁRIOS	
Cooperativa	
Grupo religioso	
Associações	
Outros - Especificar:	
MEIOS DE TRANSPORTE QUE MAIS UTILIZA	
Ônibus	
Caminhão	
Carro	
Carroça	
Outros - Especificar	

OBSERVAÇÕES

Fonte: (BRASIL, 2013)

ANEXO 2 – Critério de classificação Brasil

CRITÉRIO DE CLASSIFICAÇÃO ECONÔMICA BRASIL

ABEP
associação brasileira de empresas de pesquisa

Alterações na aplicação do Critério Brasil, válidas a partir de 01/01/2013

A dinâmica da economia brasileira, com variações importantes nos níveis de renda e na posse de bens nos domicílios, representa um desafio importante para a estabilidade temporal dos critérios de classificação socioeconômica. Em relação ao CCEB, os usuários têm apresentado dificuldades na manutenção de amostras em painel para estudos longitudinais. As dificuldades são maiores na amostragem dos estratos de pontuação mais baixa.

A ABEP vem trabalhando intensamente na avaliação e construção de um critério que seja fruto da nova realidade do país. Porém, para que os estudos produzidos pelos usuários do Critério Brasil continuem sendo úteis ao mercado e mantenham o rigor metodológico necessário, as seguintes recomendações são propostas às empresas que tenham estudos contínuos, com amostras em painel:

- A reclassificação de domicílios entre as classe C2 e D deve respeitar uma região de tolerância de 1 ponto, conforme descrito abaixo:
  - Domicílios classificados, no momento inicial do estudo, como classe D --> são reclassificados como C2, apenas no momento em que atingirem 15 pontos;
  - Domicílios classificados, no momento inicial do estudo, como classe C2 --> são reclassificados como D, apenas no momento em que atingirem 12 pontos;
  - O momento inicial de estudos desenvolvidos a partir de amostra mestra é o da realização da amostra mestra;
  - O momento inicial de estudos desenvolvidos sem amostra mestra é o da primeira medição (onda) do estudo.

IMPORTANTE: As alterações descritas acima são apenas para os estudos que usem amostras contínuas em painéis. Estudos *ad hoc* e estudos contínuos, com amostras independentes, devem continuar a aplicar o Critério Brasil regularmente.

Outra mudança importante no CCEB é válida para todos os estudos que utilizem o Critério Brasil. As classes D e E devem ser unidas para a estimativa e construção de amostras. A justificativa para esta decisão é o tamanho reduzido da classe E, que inviabiliza a leitura de resultados obtidos através de amostras probabilísticas ou por cotas, que respeitem os tamanhos dos estratos.

A partir de 2013 a ABEP deixa de divulgar os tamanhos separados destes dois estratos.

Finalmente, em função do tamanho reduzido da Classe A1 a renda média deste estrato deixa de ser divulgada. Assim, a estimativa de renda média é feita para o conjunto da Classe A.

O Critério de Classificação Econômica Brasil, enfatiza sua função de estimar o poder de compra das pessoas e famílias urbanas, abandonando a pretensão de classificar a população em termos de "classes sociais". A divisão de mercado definida abaixo é de **classes econômicas.**

SISTEMA DE PONTOS

Posse de itens

	Quantidade de Itens				
	0	1	2	3	4 ou +
Televisão em cores	0	1	2	3	4
Rádio	0	1	2	3	4
Banheiro	0	4	5	6	7
Automóvel	0	4	7	9	9
Empregada mensalista	0	3	4	4	4
Máquina de lavar	0	2	2	2	2
Videocassete e/ou DVD	0	2	2	2	2
Geladeira	0	4	4	4	4
Freezer (aparelho independente ou parte da geladeira duplex)	0	2	2	2	2

Grau de Instrução do chefe de família

Nomenclatura Antiga	Nomenclatura Atual	
Analfabeto/ Primário incompleto	Analfabeto/ Fundamental 1 Incompleto	0
Primário completo/ Ginasial incompleto	Fundamental 1 Completo / Fundamental 2 Incompleto	1
Ginasial completo/ Colegial incompleto	Fundamental 2 Completo/ Médio Incompleto	2
Colegial completo/ Superior incompleto	Médio Completo/ Superior Incompleto	4
Superior completo	Superior Completo	8

CORTES DO CRITÉRIO BRASIL

Classe	Pontos
A1	42 - 46
A2	35 - 41
B1	29 - 34
B2	23 - 28
C1	18 - 22
C2	14 - 17
D	8 - 13
E	0 - 7

PROCEDIMENTO NA COLETA DOS ITENS

É importante e necessário que o critério seja aplicado de forma uniforme e precisa. Para tanto, é fundamental atender integralmente as definições e procedimentos citados a seguir.

Para aparelhos domésticos em geral devemos:

Considerar os seguintes casos
Bem alugado em caráter permanente
Bem emprestado de outro domicílio há mais de 6 meses
Bem quebrado há menos de 6 meses

Não considerar os seguintes casos
Bem emprestado para outro domicílio há mais de 6 meses
Bem quebrado há mais de 6 meses
Bem alugado em caráter eventual
Bem de propriedade de empregados ou pensionistas

Televisores

Considerar apenas os televisores em cores. Televisores de uso de empregados domésticos (declaração espontânea) só devem ser considerados caso tenha(m) sido adquirido(s) pela família empregadora.

Rádio

Considerar qualquer tipo de rádio no domicílio, mesmo que esteja incorporado a outro equipamento de som ou televisor. Rádios tipo walkman, conjunto 3 em 1 ou microsystems devem ser considerados, desde que possam sintonizar as emissoras de rádio convencionais. Não pode ser considerado o rádio de automóvel.

Banheiro

O que define o banheiro é a existência de vaso sanitário. Considerar todos os banheiros e lavabos com vaso sanitário, incluindo os de empregada, os localizados fora de casa e os da(s) suite(s). Para ser considerado, o banheiro tem que ser privativo do domicílio. Banheiros coletivos (que servem a mais de uma habitação) não devem ser considerados.

Automóvel

Não considerar táxis, vans ou pick-ups usados para fretes, ou qualquer veículo usado para atividades profissionais. Veículos de uso misto (lazer e profissional) não devem ser considerados.

Empregado doméstico

Considerar apenas os empregados mensalistas, isto é, aqueles que trabalham pelo menos 5 dias por semana, durmam ou não no emprego. Não esquecer de incluir babás, motoristas, cozinheiras, copeiras, arrumadeiras, considerando sempre os mensalistas. Note bem: o termo empregados mensalistas se refere aos empregados que trabalham no domicílio de forma permanente e/ou continua, pelo menos 5 dias por semana, e não ao regime de pagamento do salário.

Máquina de Lavar

Considerar máquina de lavar roupa, somente as máquinas automáticas e/ou semiautomática
O tanquinho NÃO deve ser considerado.

Videocassete e/ou DVD

Verificar presença de qualquer tipo de vídeo cassete ou aparelho de DVD.

Geladeira e Freezer

No quadro de pontuação há duas linhas independentes para assinalar a posse de geladeira e freezer respectivamente. A pontuação será aplicada de forma independente:
Havendo geladeira no domicílio, independente da quantidade, serão atribuídos os pontos (4) correspondentes a posse de geladeira;
Se a geladeira tiver um freezer incorporado – 2^a. porta – ou houver no domicílio um freezer independente serão atribuídos os pontos (2) correspondentes ao freezer.

As possibilidades são:

Não possui geladeira nem freezer	0 pt
Possui geladeira simples (não duplex) e não possui freezer	4 pts
Possui geladeira de duas portas e não possui freezer	6 pts
Possui geladeira de duas portas e freezer	6 pts
Possui freezer mas não geladeira (caso raro mas aceitável)	2 pt

OBSERVAÇÕES IMPORTANTES

Este critério foi construído para definir grandes classes que atendam às necessidades de segmentação (por poder aquisitivo) da grande maioria das empresas. Não pode, entretanto, como qualquer outro critério, satisfazer todos os usuários em todas as circunstâncias. Certamente há muitos casos em que o universo a ser pesquisado é de pessoas, digamos, com renda pessoal mensal acima de US$ 30.000. Em casos como esse, o pesquisador deve procurar outros critérios de seleção que não o CCEB.

A outra observação é que o CCEB, como os seus antecessores, foi construído com a utilização de técnicas estatísticas que, como se sabe, sempre se baseiam em coletivos. Em uma determinada amostra, de determinado tamanho, temos uma determinada probabilidade de classificação correta, (que, esperamos, seja alta) e uma probabilidade de erro de classificação (que, esperamos, seja baixa). O que esperamos é que os casos incorretamente classificados sejam pouco numerosos, de modo a não distorcer significativamente os resultados de nossa investigação.

Nenhum critério, entretanto, tem validade sob uma análise individual. Afirmações freqüentes do tipo *"... conheço um sujeito que é obviamente classe D, mas pelo critério é classe B..."* não invalidam o critério que é feito para funcionar estatisticamente. Servem porém, para nos alertar, quando trabalhamos na análise individual, ou quase individual, de comportamentos e atitudes (entrevistas em profundidade e discussões em grupo respectivamente). Numa discussão em grupo um único caso de má classificação pode pôr a perder todo o grupo. No caso de entrevista em profundidade os prejuízos são ainda mais óbvios. Além disso, numa pesquisa qualitativa, raramente uma definição de classe exclusivamente econômica será satisfatória.

Portanto, é de fundamental importância que todo o mercado tenha ciência de que o CCEB, ou qualquer outro critério econômico, não é suficiente para uma boa classificação em pesquisas qualitativas. Nesses casos deve-se obter além do CCEB, o máximo de informações (possível, viável, razoável) sobre os respondentes, incluindo então seus comportamentos de compra, preferências e interesses, lazer e hobbies e até características de personalidade.

Uma comprovação adicional da conveniência do Critério de Classificação Econômica Brasil é sua discriminação efetiva do poder de compra entre as diversas regiões brasileiras, revelando importantes diferenças entre elas

Informações referentes ao LSE 2011
9 RMs – IBOPE Mídia

Classes	Renda média bruta familiar no mês em R$
Classe A	9.263
Classe B1	5.241
Classe B2	2.654
Classe C1	1.685
Classe C2	1.147
Classe DE	776

Classes	Gde. FORT	Gde. REC	Gde. SALV	Gde. BH	Gde. RJ	Gde. SP	Gde. CUR	Gde. POA	DF	9 Grandes Áreas
Classe A1	0,6	0,4	0,4	0,4	0,1	0,7	0,6	0,9	0,8	0,5
Classe A2	3,5	2,8	1,6	3,6	3,2	4,0	7,2	6,3	7,7	4,0
Classe B1	4,5	6,5	6,1	9,6	10,4	10,7	14,6	10,4	15,7	10,0
Classe B2	9,5	13,0	12,5	21,9	20,0	26,2	26,8	25,9	24,9	21,8
Classe C1	17,0	20,6	21,9	26,7	28,3	28,4	24,0	28,4	24,9	26,3
Classe C2	30,6	28,1	31,6	23,5	23,8	19,6	17,0	19,4	16,3	22,5
Classe DE	34,3	28,6	25,9	14,3	14,2	10,4	9,8	8,7	9,7	14,9
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

ABEP - Associação Brasileira de Empresas de Pesquisa – 2012 – www.abep.org – abep@abep.org
Dados com base no Levantamento Sócio Econômico 2011 - IBOPE

5

Fonte: ABEP (2012).

APÊNDICE 1 – Questionário

Data: ___/___/_____ Questionário: _____

Universidade Estadual Paulista "Júlio de Mesquita Filho" – Câmpus de Botucatu
Faculdade de Ciências Agronômicas – Programa de Pós Graduação em Agronomia – Energia na Agricultura
Faculdade de Tecnologia de Ourinhos

Endereço: _____ nº: _____ Complemento: _____
Bairro: _____

Data de entrada na residência: ___/___/_____ Número de moradores na residência: ____

Número de cômodos: ____

Utiliza lâmpadas fluorescentes em grande parte dos cômodos: ( ) sim ( ) não

Possui aquecimento solar de água: ( ) sim ( ) não

Condição de moradia: ( ) alugada ( ) própria ( ) cedida ( ) outro: _____

Nº	Data de Nascimento	Idade	Sexo	Escolaridade	Ocupação	Renda (R$)	Observação
01							
02							
03							
04							
05							
06							
07							
08							
09							
10							

Itens	Quantidade
Televisão em cores ( ) led ( ) lcd ( ) plasma ( ) tubo	
Rádio	
Banheiro	
Automóvel	
Empregada mensalista	
Máquina de lavar	
Videocassete e/ou DVD	
Geladeira	
Freezer (aparelho independente ou parte da geladeira duplex)	
Chuveiro elétrico	
Secador de cabelo	
Micro-ondas	
Cafeteira e/ou torradeira	
Fogão e/ou forno e/ou churrasqueira e/ou panela elétrica	
Computador e/ou notebook	
Ar-condicionado	
Aspirador de pó	

Ferro elétrico (ferro de passar roupa)	
Ventilador de mesa e/ou de teto	

Grau de Instrução do Chefe de Família		
Nomenclatura Antiga	Nomenclatura Atual	X
Analfabeto/Primário Incompleto	Analfabeto/Funtamental 1 Incompleto	
Primário Completo/Ginásio Incompleto	Fundamental 1 Completo/Fundamental 2 Incompleto	
Ginásio Completo/Colegial Incompleto	Fundamental 2 Completo/Médio Incompleto	
Colegial Completo/Superior Incompleto	Médio Completo/Superior Incompleto	
Superior Completo	Superior Completo	

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Convido o (a) senhor (a) a participar como voluntário desta pesquisa.

Este questionário, com duração de aproximadamente 10 minutos, pretende levantar características de sua residência para, posteriormente, as informações serem utilizadas no desenvolvimento da tese de Marcelo Hiroshi Tutia do Programa de Pós Graduação em Agronomia – Energia na Agricultura – Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP) – Campus de Botucatu. O autor desta pesquisa é aluno regularmente matriculado no programa de Pós-Graduação da referida Universidade e professor da Faculdade de Tecnologia de Ourinhos (FATEC – Ourinhos), cujo trabalho tem como título "Utilização de procedimentos multivariados no consumo de água e energia elétrica em habitações sociais com sistema de aquecimento solar".

Os questionários serão aplicados em uma amostra aleatória de residências, sem qualquer risco e dano para o respondente e com sigilo sobre as informações pessoais.

A participação será voluntária com consentimento livre e esclarecido e as informações sobre as características pessoais coletadas serão mantidas em sigilo. Os resultados obtidos nas análises dos dados serão divulgados exclusivamente com fins científicos envolvendo apresentações em eventos ou publicações em revistas especializadas, assegurando sigilo sobre sua participação.

A metodologia da coleta de dados, desenvolvimento do trabalho e confidencialidade das informações disponibilizadas é de responsabilidade de Marcelo Hiroshi Tutia, estando disponível para esclarecimentos no telefone (14) 9796-6386 ou pelo e-mail marcelotutia@yahoo.com.br.

Este documento é elaborado em duas vias, sendo uma via entregue ao respondente e outra via arquivada pelo pesquisador.

Estando de acordo e ciente sobre as informações fornecidas pelo pesquisador, por favor, assinar e datar este termo.

Qualquer dúvida adicional, você poderá entrar em contato com o Comitê de Ética em Pesquisa, através dos fones: (14) 3880-1608/3880-1609.

Ourinhos – SP, ___ de _____ de 2013.

_____ _____

Nome e assinatura do respondente **Marcelo Hiroshi Tutia – Responsável pela Pesquisa**

Pesquisador: Marcelo Hiroshi Tutia, Avenida Vitalina Marcusso, 1400, 19910-206, Ourinhos – SP, (14) 9796-6386/3326-1486, marcelotutia@yahoo.com.br.

Orientador: Prof. Dr. Carlos Roberto Padovani, UNESP – Campus de Botucatu, Programa de Pós Graduação em Agronomia – Energia na Agricultura, 18618-970, Botucatu – SP, (14) 3880-0070.