



COMPARAÇÃO DO CONSUMO DE ÁGUA E ENERGIA ELÉTRICA E INDICADORES SOCIOECONÔMICOS EM CONJUNTOS HABITACIONAIS COM E SEM SISTEMA DE AQUECIMENTO SOLAR

Marcelo Hiroshi Tutia¹, Paulo André de Oliveira² & Carlos Roberto Padovani³

RESUMO. Atualmente fontes de energia renováveis são exploradas para o desenvolvimento da sociedade e se destacam por serem mais limpas e menos poluentes. Entre estas fontes de energia renováveis, a energia solar é usada em sistemas de aquecimento solar de água, que 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, o qual 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 o sistema de aquecimento solar de água. O objetivo desta pesquisa foi comparar o consumo de água e energia elétrica, e variáveis socioeconômicas entre dois conjuntos habitacionais, sem e com o sistema de aquecimento solar, na cidade de Ourinhos-SP. Os resultados permitiram concluir que existe diferença significativa entre os conjuntos de variáveis de consumo e as socioeconômicas.

PALAVRAS-CHAVE: T² de Hotelling, consumo, socioeconômico, sistema de aquecimento solar, habitações de interesse social.

WATER AND ELECTRICITY CONSUMPTION COMPARISON AND SOCIOECONOMIC INDICATORS ON HOUSING SETS WITH AND WITHOUT SOLAR HEATING SYSTEM.

ABSTRACT: Currently, renewable energy sources are used to the society development and stand out for being less polluting. Among the renewable energy sources, the solar energy is used in water heating systems, that has among its objectives, to provide savings of electricity heating water without the need for electrical systems such as the electric shower which is one of the appliances that consume more energy. To provide energy consumption economy, minimize environmental impacts, redirect investment in new sources of electricity, the Government of the State of Sao Paulo, together with the Development Company of Housing and Urban of the State of Sao Paulo - CDHU, installed theses systems in homes of social interest. The objective of this research was to compare the consumption of water and electricity, and socioeconomic variables between two housing projects with and without solar heating system in the city of Ourinhos-SP. The results showed a significant difference between consumption and socioeconomic variables.

1 INTRODUÇÃO

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.

Segundo Naspolini et al. (2008), Ijumba et al. (2009), Naspolini et al. (2010), Bodach e Hamhaber (2010), Silva et al. (2011), Naspolini e Rütther (2011), o sistema

de aquecimento solar de água instalado em habitações sociais de 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.

O termo habitações sociais, 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

¹ Faculdade de Tecnologia de Ourinhos. Email: marcelo.tutia@fatecourinhos.edu.br

² Doutorado em Agronomia e Professor da Faculdade de Tecnologia de Botucatu. Email: pauloandre68@gmail.com

³ Instituto de Biociências, Departamento de Bioestatística. Departamento De Bioestatística, UNESP/Botucatu

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).

A conversão de energia solar para térmica (aquecimento de água) é a mais conhecida no Brasil e sua utilização vem crescendo nos últimos anos, contribuindo para a diminuição do consumo de energia elétrica, segundo Napolini et al. (2008), Silva et al. (2011). Porém, em praticamente todas publicações o consumo de energia é uma variável analisada independente de outros fatores, o que torna uma forma simplista para analisar conclusões sobre a eficiência do consumo. Na presente pesquisa o consumo de energia elétrica será visto conjuntamente com outras variáveis, que podem contribuir ou não para a sua redução.

Coloca-se como hipótese a ser verificada que os consumos de energia elétrica e água, e as variáveis socioeconômicas não devem ser vistas isoladamente, mas analisadas conjuntamente para que forneçam resultados mais robustos e fidedignos sobre a instalação do sistema de aquecimento solar como política de habitação.

Deve-se 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. Neste sentido, a análise multivariada é uma das ferramentas estatísticas que proporciona o estudo conjunto de variáveis e, conseqüentemente, fornece resultados mais precisos que podem ser imprescindíveis para melhorar as tomadas de decisão.

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).

No presente trabalho foi considerada a estatística T^2 de Hotelling, técnica de estatística multivariada para se comparar conjuntos de variáveis, complementada com a construção dos intervalos de confiança simultâneos, com o objetivo de avaliar e aprofundar a análise do relacionamento do consumo de energia elétrica e água e características socioeconômicas em habitações

construídas com ou sem o sistema de aquecimento solar; como subsídios de entendimento e melhoria da qualidade de informação e implementação de políticas públicas.

2 MATERIAL E MÉTODOS

O estudo envolve dados coletados nos Conjuntos Habitacionais (bairros) Sem Sistema e Com Sistema de Aquecimento; o primeiro com 523 residências, e o segundo com 596 residências, segundo informações da Prefeitura de Ourinhos.

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 (5%) percentuais 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.

Como instrumento para coleta de dados foi utilizado um questionário desenvolvido com base em outras duas ferramentas de levantamento de dados, a Ficha A e o Critério de Classificação Brasil. 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 (BRASIL, 2003). 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).

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 residência. Estas informações foram disponibilizadas pela Superintendência de Água e Esgoto – SAE e a Companhia Paulista de Luz e Força – CPFL, respectivamente.

Por meio do consumo mensal de água e energia elétrica foi possível gerar as seguintes variáveis que serão utilizadas na pesquisa:

- (X_1) SAE: média do consumo mensal de água por residência.
- (X_2) CPFL: média do consumo mensal de energia por residência.

Por meio das variáveis do questionário aplicado foram determinados alguns domínios e novas variáveis foram construídas a seguir:

- (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) IGEN: indicador da variação da quantidade de gênero por unidade amostral, [(n° de moradores do sexo feminino - n° de moradores do sexo masculino)/(n° de moradores)], variável que assume valores de -1 a 1; quanto mais próximo de -1, maior a quantidade de moradores do gênero masculino na unidade amostral; e quanto mais próximo de 1, maior quantidade de moradores do gênero feminino na unidade amostral. Caso o indicador atinja a unidade negativa, a residência é composta apenas de ocupantes do gênero masculino e, se atingir a unidade positiva, apenas ocupantes do gênero 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, [(n° de moradores de 0 a 14 anos) + (n° de moradores com idade igual e superior a 65 anos)]/[n° de moradores], 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, [n° 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, [n° de equipamentos que consomem maior quantidade de energia elétrica]/[quantidade de itens na residência que consomem energia], 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).

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, $k = 1, \dots, 11$, [$k = 1$ (consumo de água); $k = 2$ (consumo de energia); $k = 3, \dots, 11$ (socioeconômicas)], medida na j -ésima unidade amostral (residência), $j = 1, \dots, n = n_1 + n_2$ ($n = 287$ 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.

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

No estudo utilizou-se a estatística T^2 de Hotelling para comparação entre os conjuntos de variáveis de consumo (X_1 e X_2) e as variáveis socioeconômicas ($X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}$ e X_{11}), para os respectivos conjuntos habitacionais, Sem Sistema e Com Sistema. 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 representação matricial da Tabela 1 pode ser expressa em $X_{n \times p}$, com $n = 287$ e $p = 11$, como

$$X_{n \times p} = \begin{bmatrix} x_{111} & \dots & x_{11p} \\ \vdots & \ddots & \vdots \\ x_{1n_11} & \dots & x_{1n_1p} \\ x_{211} & \dots & x_{21p} \\ \vdots & \ddots & \vdots \\ x_{2n_21} & \dots & x_{2n_2p} \end{bmatrix}; \text{ com } n = n_1 + n_2.$$

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

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 Conj. Hab. Sem Sistema	1	x_{111}	x_{112}	x_{113}	...	x_{1111}	x'_{11}
	2	x_{121}	x_{122}	x_{123}	...	x_{1211}	x'_{12}
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	$n_1=135$	x_{1n_11}	x_{1n_12}	x_{1n_13}	...	x_{1n_111}	x'_{1n_1}
Bairro 2 Conj. Hab. Com Sistema	1	x_{211}	x_{212}	x_{213}	...	x_{2111}	x'_{21}
	2	x_{221}	x_{222}	x_{223}	...	x_{2211}	x'_{22}
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	$n_2=152$	x_{2n_21}	x_{2n_22}	x_{2n_23}	...	x_{2n_211}	x'_{2n_2}

Considerando μ_1 e μ_2 os vetores de médias populacionais dos bairros 1 e 2, respectivamente, as hipóteses estatísticas a serem testadas são

$$\begin{cases} H_0: \mu_1 - \mu_2 = \mathbf{0} \\ H_1: \mu_1 - \mu_2 \neq \mathbf{0} \end{cases}$$

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 $\bar{x}'_i = (\bar{x}_{i \cdot 1} \dots \bar{x}_{i \cdot k} \dots \bar{x}_{i \cdot p})$, em que

$$\bar{x}_i = \begin{bmatrix} \bar{x}_{i \cdot 1} \\ \bar{x}_{i \cdot 2} \\ \bar{x}_{i \cdot 3} \\ \vdots \\ \bar{x}_{i \cdot p} \end{bmatrix} = \begin{bmatrix} \bar{x}_{i2x1}^{\{c\}} \\ \bar{x}_{i9x1}^{\{s\}} \end{bmatrix},$$

$$S_{i \times p} = \begin{bmatrix} S_{11}^{(i)} & S_{12}^{(i)} & | & S_{13}^{(i)} & \dots & S_{1k}^{(i)} & \dots & S_{3p}^{(i)} \\ S_{21}^{(i)} & S_{22}^{(i)} & | & S_{23}^{(i)} & \dots & S_{2k}^{(i)} & \dots & S_{2p}^{(i)} \\ \hline S_{31}^{(i)} & S_{32}^{(i)} & | & S_{33}^{(i)} & \dots & S_{3k}^{(i)} & \dots & S_{3p}^{(i)} \\ \vdots & \vdots & | & \vdots & \ddots & \vdots & \ddots & \vdots \\ S_{k1}^{(i)} & S_{k2}^{(i)} & | & S_{k3}^{(i)} & \dots & S_{kk}^{(i)} & \dots & S_{kp}^{(i)} \\ \vdots & \vdots & | & \vdots & \ddots & \vdots & \ddots & \vdots \\ S_{p1}^{(i)} & S_{p2}^{(i)} & | & S_{p3}^{(i)} & \dots & S_{pk}^{(i)} & \dots & S_{pp}^{(i)} \end{bmatrix} = \begin{bmatrix} S_{i2x2}^{\{c\}} & | & S_{i2x9}^{\{cs\}} \\ \hline S_{i9x2}^{\{sc\}} & | & S_{i9x9}^{\{s\}} \end{bmatrix},$$

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

O índice sobrescrito no vetor $\bar{x}_i^{(l)}$ e na matriz $S_i^{(l)}$, será denominado por l , sendo $l = c$ para consumo e $l = s$ para socioeconômico. Neste sentido, os vetores $\bar{x}_{i2x1}^{\{c\}}$ e $\bar{x}_{i9x1}^{\{s\}}$ representam as médias de consumo (água e energia elétrica) e socioeconômicas (número médio de moradores por cômodo, renda *per capita* da residência, 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, $S_{i_{2x2}}^{[c]}$, representa a matriz de variâncias e covariâncias dos consumos, $S_{i_{9x9}}^{[s]}$, a matriz de variâncias e covariâncias dos socioeconômicos e, por fim, $S_{i_{2x9}}^{[cs]}$ e $S_{i_{9x2}}^{[sc]}$, as matrizes das covariâncias entre consumo e socioeconômico. A matriz $S_{i_{2x9}}^{[cs]}$ terá como transposta $S_{i_{9x2}}^{[sc]}$.

2.1 T² de Hotelling

Para a comparação dos dois conjuntos habitacionais foram realizados, separadamente, o teste T² de Hotelling para as variáveis de consumo ($k = 1, 2$) e, posteriormente, para as variáveis socioeconômicas ($k = 3, \dots, 11$).

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

$$\begin{cases} H_0: \mu_1^{[c]} - \mu_2^{[c]} = \mathbf{0} \\ H_1: \mu_1^{[c]} - \mu_2^{[c]} \neq \mathbf{0} \end{cases};$$

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

$$\bar{x}_{i_{2x1}}^{[c]} = \begin{bmatrix} \bar{x}_{i \cdot 1} \\ \bar{x}_{i \cdot 2} \end{bmatrix} \text{ e } S_{i_{2x2}}^{[c]} = \begin{bmatrix} S_{11}^{(i)} & S_{12}^{(i)} \\ S_{21}^{(i)} & S_{22}^{(i)} \end{bmatrix}, \text{ com}$$

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

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

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

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

$$\begin{cases} H_0: \mu_1^{[s]} - \mu_2^{[s]} = \mathbf{0} \\ H_1: \mu_1^{[s]} - \mu_2^{[s]} \neq \mathbf{0} \end{cases};$$

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

$$\bar{x}_{i_{9x1}}^{[s]} = \begin{bmatrix} \bar{x}_{i \cdot 3} \\ \vdots \\ \bar{x}_{i \cdot 11} \end{bmatrix} \text{ e } S_{i_{9x9}}^{[s]} = \begin{bmatrix} S_{33}^{(i)} & S_{34}^{(i)} & \dots & S_{311}^{(i)} \\ S_{43}^{(i)} & S_{44}^{(i)} & \dots & S_{411}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ S_{113}^{(i)} & S_{114}^{(i)} & \dots & S_{1111}^{(i)} \end{bmatrix},$$

com

$$S_{kk}^{(i)} = \frac{\sum_{j=1}^{n_i} x_{ijk}^2 - n_i \bar{x}_{i \cdot k}^2}{n_i - 1} \quad (3) \text{ e}$$

$$S_{kk'}^{(i)} = \frac{\sum_{j=1}^{n_i} x_{ijk} x_{ijk'} - n_i \bar{x}_{i \cdot k} \bar{x}_{i \cdot k'}}{n_i - 1} \quad (4),$$

$k \neq k', k = 3, \dots, 11 \text{ e } i = 1, 2.$

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:

$$\delta_{ij}^{2(l)} = (\mathbf{x}_{ij}^{(l)} - \bar{\mathbf{x}}_i^{(l)})' S_i^{(l)} (\mathbf{x}_{ij}^{(l)} - \bar{\mathbf{x}}_i^{(l)}) \quad (5),$$

$i = 1, 2, j = 1, \dots, n_i$ ($n_1 = 135$ e $n_2 = 157$), $l = c$ (consumo) ou s (socioeconômico);

no qual $\mathbf{x}_{ij}^{(l)}$ são os vetores $p_{(l)}$ -dimensionais dos valores observados, $\bar{\mathbf{x}}_i^{(l)}$ o vetor de médias e $S_i^{(l)}$ 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, $\delta_{(1)}^{2(l)} \leq \delta_{(2)}^{2(l)} \leq \dots \leq \delta_{(n_i)}^{2(l)}$, sendo $\delta_{(j)}^{2(l)}$ a j -ésima estatística de ordem. Após, deve-se construir o gráfico dos pares $\left[\delta_{(j)}^{2(l)}; \chi_{(p_{(l)}; ((j-\frac{1}{2})/n_i))}^2 \right]$, nas quais $\chi_{(p_{(l)}; ((j-\frac{1}{2})/n_i))}^2$ representa o quantil de ordem $100 \left((j - \frac{1}{2}) / n_i \right) \%$ da distribuição qui-quadrado com $p_{(l)}$ graus de liberdade, isto é,

$$P \left[\chi_{p_{(l)}}^2 \leq \left[\chi_{(p_{(l)}; ((j-\frac{1}{2})/n_i))}^2 \right] \right] = (j - \frac{1}{2}) / n_i. \quad (6)$$

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 T² 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.

$$\begin{cases} H_0: \Sigma_1^{[l]} = \Sigma_2^{[l]} \\ H_1: \Sigma_1^{[l]} \neq \Sigma_2^{[l]} \end{cases};$$

$l = c$ (consumo) ou s (socioeconômico).

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

$$M_{\{l\}} = (n - 2) \ln |S^{[l]}| - \sum_{i=1}^2 (n_i - 1) \ln |S_i^{[l]}| \quad (7)$$

$$C_{\{l\}}^{-1} = 1 - \frac{(2p_{\{l\}}^2 + 3p_{\{l\}} - 1) \left(\sum_{i=1}^2 \frac{1}{n_i - 1} - \frac{1}{(n-2)} \right)}{6(p_{\{l\}} + 1)} \quad (8),$$

com

$$S_{\{l\}} = \frac{(n_1 - 1)S_1^{\{l\}} + (n_2 - 1)S_2^{\{l\}}}{n_1 + n_2 - 2} \quad (9),$$

$l = c$ (consumo) ou s (socioeconômico).

A estatística do teste é dada por $M_{\{l\}}C_{\{l\}}^{-1}$ com distribuição $\chi_{\left(\frac{1}{2}p_{\{l\}}(p_{\{l\}}+1)\right)}^2$, 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 igualdade dos vetores de médias, sob a veracidade de H_0 , é expressa por:

$$T^{2\{l\}} = \left[(\bar{x}_1^{\{l\}} - \bar{x}_2^{\{l\}}) \right]' \left[\left(\frac{1}{n_1} + \frac{1}{n_2} \right) S_{\{l\}} \right]^{-1} \left[(\bar{x}_1^{\{l\}} - \bar{x}_2^{\{l\}}) \right] \quad (10),$$

com $S_{\{l\}}$ 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 um determinado 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:

$$T^{2\{l\}} = \left[(\bar{x}_1^{\{l\}} - \bar{x}_2^{\{l\}}) \right]' \left[\frac{1}{n_1} S_1^{\{l\}} + \frac{1}{n_2} S_2^{\{l\}} \right]^{-1} \left[(\bar{x}_1^{\{l\}} - \bar{x}_2^{\{l\}}) \right]. \quad (11)$$

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 $\chi_{(\alpha, p_{\{l\}})}^2$, e a mesma conclusão se faz presente quanto aos vetores de médias.

2.2 Intervalos de confiança simultâneos de Hotelling

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

$$ICS \left[\mathbf{a}'(\boldsymbol{\mu}_1^{\{l\}} - \boldsymbol{\mu}_2^{\{l\}}) \right]: \mathbf{a}'(\bar{x}_1^{\{l\}} - \bar{x}_2^{\{l\}}) \pm$$

$$\sqrt{\frac{(n_1+n_2-2)p_{\{l\}}}{n_1+n_2-p_{\{l\}}-1} F_{(\alpha, p_{\{l\}}, n_1+n_2-p_{\{l\}}-1)} \frac{n_1+n_2}{n_1 n_2} \mathbf{a}' S_{\{l\}} \mathbf{a}} \quad (12),$$

sendo $\mathbf{a}' = [a_1, a_2, \dots, a_p]$, 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 sejam diferentes, os limites dos Intervalos de Confiança Simultâneos de Hotelling são dados por:

$$ICS \left[\mathbf{a}'(\boldsymbol{\mu}_1^{\{l\}} - \boldsymbol{\mu}_2^{\{l\}}) \right]: \mathbf{a}'(\bar{x}_1^{\{l\}} - \bar{x}_2^{\{l\}}) \pm \sqrt{\chi_{(\alpha, p_{\{l\}})}^2 \mathbf{a}' \left(\frac{1}{n_1} S_1^{\{l\}} + \frac{1}{n_2} S_2^{\{l\}} \right) \mathbf{a}} \quad (13).$$

Os resultados do teste T^2 de Hotelling, intervalos de confiança Simultâneos e os gráficos que serão apresentados, foram calculados e gerados por meio do *software* livre R-Gui (THE COMPREHENSIVE, 2012).

3 RESULTADOS E DISCUSSÃO

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 (conjunto de variáveis de consumo e conjunto de variáveis socioeconômicas) os resultados (valor $p < 0,01$, variáveis de consumo: $p = 0,009$, variáveis socioeconômicas: $p < 0,001$) indicam a rejeição da hipótese de homogeneidade considerando um nível de significância de 5%.

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. Portanto, 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.

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.

Nota-se, na Tabela 2, 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 de consumo do Conjunto Habitacional Sem Sistema são maiores em relação ao Conjunto Habitacional Com Sistema, em ambas variáveis.

Para as variáveis socioeconômicas, observa-se, na Tabela 2, que embora o teste multivariado indique diferença significativa entre os conjuntos, os intervalos de confiança simultâneos de Hotelling apresentam o valor zero dentro de seus limites, indicando que não foi suficientemente poderoso para localizar de forma isolada as variáveis que contribuem para rejeição. 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.

Tabela 2 - Diferença entre as médias dos conjuntos habitacionais e limites do intervalo de confiança simultâneo.

Resultado do Teste Multivariado	Variável	Sem – Com ($\bar{x}_1^{(l)} - \bar{x}_2^{(l)}$)	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

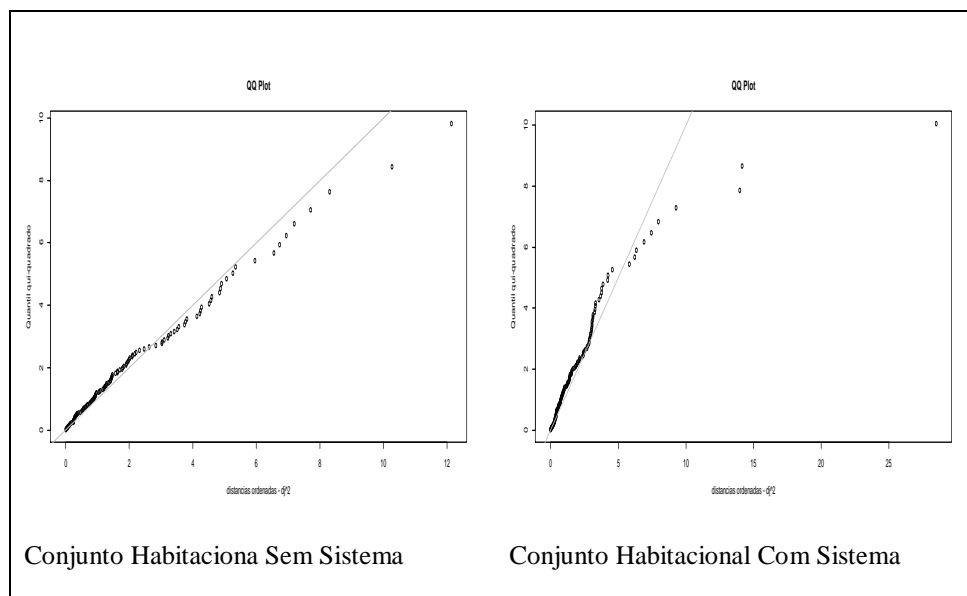


Figura 1 - Q-Q plot para o conjunto de variáveis de consumo dos Conjuntos Habitacionais Sem e Com Sistema.

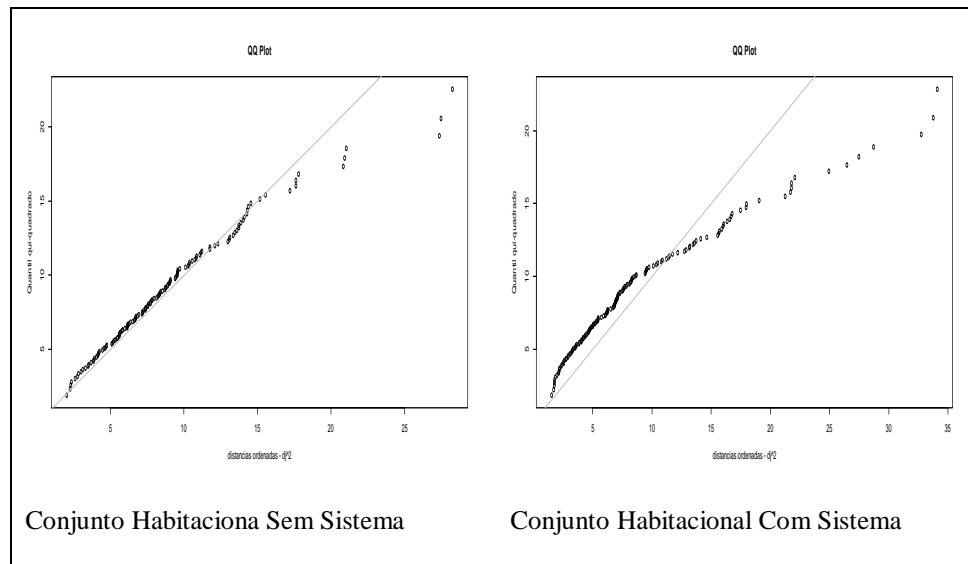


Figura 2 - Q-Q plot para o conjunto de variáveis socioeconômicas dos Conjuntos Habitacionais Sem e Com Sistema.

4 CONCLUSÃO

De acordo com o procedimento multivariado aplicado, T^2 de Hotelling, foi observado diferença significativa tanto no conjunto de variáveis de consumo (água e energia elétrica) como para as variáveis socioeconômicas, indicando que os conjuntos habitacionais são distintos.

O conjunto de variáveis de consumo indicam que o Conjunto Habitacional Sem Sistema apresenta valores maiores, consumo 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 também apresentou diferenças significativas pelo teste estatístico multivariado, todavia os intervalos não confirmam o teste, justificando-se pelos motivos mencionados. Pelo fato do teste T^2 apresentar diferença significativa, existem indícios que características socioeconômicas possam estar associadas, diretamente ou inversamente, ao consumo nos conjuntos habitacionais, com ou sem sistema de aquecimento.

5 REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE EMPRESAS DE PESQUISA – ABEP. Dados com base no levantamento socio econômico de 2011. São Paulo, SP, 2012.

ABIKO, A. K. **Introdução à gestão habitacional**. São Paulo: Escola Politécnica da USP, 1995. 31 p. (Série Texto Técnico, 12).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/ Brasília, DF: Ministério da Saúde, 2003.

COMPANHIA DE DESENVOLVIMENTO HABITACIONAL E URBANO DO ESTADO DE SÃO PAULO - CDHU. **Provisão de Moradias**. São Paulo, 2014. Disponível em: <http://www.cdhu.sp.gov.br/programas_habitacionais/pr_ovisao_de_moradias/parceria_com_municipios.asp>. Acesso em: 10 jan. 2015.

HAIR JR, J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E.; TATHAN, R. L. **Análise multivariada de dados**. 6. ed. Porto Alegre: Bookman, 2009.

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

MARDIA, K. V. Assessment of multinormality and the robustness of Hotelling T^2 test. **Applied Statistics**, Cidade, v. 24, p. 163-171, 1974.

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: Congresso Brasileiro de Energia Solar, 2.; Conferência Latinoamericana da ISES, 3., 2008, Florianópolis. **Anais**. São Paulo: ABENS, 2008. p. 745-754.

NASPOLINI, H. F.; MILITAO, H.; RUTHER, R. 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.

PROGRAMA NACIONAL DE CONSERVAÇÃO DE ENERGIA ELÉTRICA - PROCEL **Dicas de economia de energia**. Brasília, 2006. Disponível em: <<http://www.procelinfo.com.br/main.asp?View={E6BC2A5F-E787-48AF-B485-439862B17000}>>. Acesso em: 20 jan. 2014.

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 Edifício. **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 Edifício**, p. 698-707, Rio de Janeiro, 2011.

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