



Correlação entre o consumo residencial de energia e fatores socioeconômicos na área de concessão da Cemig Distribuição: uma análise espacial

Tema: Planejamento da Expansão

Autores: Danilo de Deus Mota

Co-Autores: -

Empresa: Cemig Distribuição S. A.

Resumo

Este estudo explora a relação entre fatores socioeconômicos e o consumo residencial de energia na área de concessão da Cemig Distribuição, com o objetivo de compreender como variáveis como renda média e conectividade influenciam os padrões de consumo energético. Utilizando dados de consumo da Cemig, indicadores do IBGE e outras fontes, a análise incluiu mapas temáticos, modelos de regressão e técnicas de estatística espacial, como Índice de Moran e LISA Maps. Os resultados revelam uma forte correlação positiva entre renda e consumo energético, destacando padrões espaciais significativos, como clusters de alto e baixo consumo em regiões específicas. Modelos de regressão geograficamente ponderada (GWR) e autorregressivos espaciais (SAR) apresentaram maior precisão ao incorporar a dependência espacial. Aplicações práticas incluem o planejamento mais preciso da compra de energia e da expansão da infraestrutura, bem como a possibilidade de utilização do consumo residencial de energia como proxy para estimativas atualizadas de renda, contribuindo para estratégias mais eficientes em diversos setores.

1. Introdução

A Inteligência Geográfica pode contribuir significativamente para um controle e planejamento mais eficazes da geração, transmissão e distribuição de energia elétrica. No setor de distribuição, utilizando grandes volumes de dados geográficos das distribuidoras de energia, é possível analisar padrões de consumo energético em diferentes regiões, correlacionando-os com dados socioeconômicos, como os obtidos através do Censo do IBGE. A origem dos dados de consumo de energia de forma georreferenciada é a base de faturamento das próprias distribuidoras de energia que possuem as informações de consumo e localização por consumidor, além de outras informações.

Neste contexto, este trabalho tem como objetivo analisar a relação entre variáveis socioeconômicas e o consumo residencial de energia na área de concessão da Cemig Distribuição considerando o componente espacial. Buscou-se verificar a hipótese de que maiores níveis de renda média e outros indicadores socioeconômicos resultam em maior consumo de energia residencial. Utilizando dados de consumo energético da Cemig e diversos indicadores socioeconômicos públicos, como renda média, PIB per capita, e densidade de banda larga, a análise foi realizada através de mapas temáticos, Índice de Moran, LISA Maps, e regressões múltiplas e espaciais utilizando os softwares R e QGIS.

Utilizando informações da base de dados de consumo dos domicílios da AES Eletropaulo, FRANCISCO (2010) analisou em sua tese de doutorado o uso do consumo de energia elétrica como um indicador socioeconômico. Dentre outros objetivos, a pesquisa investiga indicadores de renda e consumo de energia elétrica agregados por áreas de ponderação do município de São Paulo, utilizando também dados do Censo Demográfico 2000. Os resultados mostram uma forte correlação espacial entre renda e consumo de energia elétrica.

No contexto atual, a disseminação de placas solares, carros elétricos e o aumento do uso de aparelhos de ar-condicionado têm causado alterações significativas no perfil de consumo de energia elétrica nas residências. A posse e o uso crescente desses equipamentos resultam em padrões de consumo mais elevados e variáveis, impactando diretamente a demanda energética. Conhecer o perfil de consumo de energia elétrica, considerando o perfil socioeconômico dos consumidores, é fundamental para que distribuidoras e outras entidades do setor elétrico possam planejar e suprir a energia de forma eficaz e eficiente.

Os resultados deste trabalho revelaram uma forte correlação positiva entre renda e consumo energético na área de concessão da Cemig Distribuição, destacando padrões espaciais significativos, como clusters de alto e baixo consumo em regiões específicas, o que contribui para um planejamento mais preciso da compra de energia, da expansão da infraestrutura e da implementação de políticas de eficiência energética, garantindo que a oferta de energia atenda adequadamente às necessidades da população. Foram utilizados dados de quatro fontes: Cemig Distribuição (Consumo de Energia Residencial e quantidade de consumidores), IBGE (PIB e Renda), Ministério dos Transportes (automóveis e total de veículos) e Anatel (densidade de celulares, densidade de banda larga e Índice Brasileiro de Conectividade - IBC). Nos mapas temáticos foram utilizadas tonalidades diferentes para as fontes. Um resumo das informações das variáveis utilizadas é apresentado na tabela 1.

Tabela 1 - Informações das Variáveis

Variável	Granularidade	Origem	Data	Tonalidade nos Mapas
Consumo de Energia Residencial	Domicílio	Cemig Distribuição	2012 e 2023	Verde
Quantidade de Consumidores	Município	Cemig Distribuição	2012 e 2023	Verde
Produto Interno Bruto per capita	Município	IBGE	2021	Laranja
Renda Média	Município	Censo Demográfico IBGE	2010	Laranja
Automóveis	Município	Ministério dos Transportes	Setembro de 2024	Vermelho
Total de Veículos	Município	Ministério dos Transportes	Setembro de 2024	Vermelho
Densidade de Celulares	Município	Anatel	Agosto de 2024	Azul
Densidade de Banda Larga	Município	Anatel	Agosto de 2024	Azul
Índice Brasileiro de Conectividade (IBC)	Município	Anatel	2023	Azul

A granularidade original das informações de consumo de energia é o domicílio, mas as informações fornecidas pela Cemig já estavam agregadas por município. Constavam os dados de consumo residencial mensal total e quantidade de domicílios (unidades consumidoras) para os anos de 2012 e 2023. Optou-se por trabalhar com o consumo médio por consumidor (CPC), obtido pela divisão do consumo total pela quantidade de consumidores. Isso deve-se ao fato de que a renda do Censo do IBGE reflete a renda média da população. Demais dados foram padronizados considerando a população dos municípios ou já vieram assim originalmente.

A granularidade dos dados de renda e PIB per capita é o município e estão disponíveis no portal do IBGE. O PIB per capita é de 2021 e a renda do Censo Demográfico de 2010. Devido a essa defasagem das informações de renda, foram utilizadas outras informações como proxies: automóveis, total de veículos, densidade de celulares, densidade de banda larga e IBC.

Segundo informações da Anatel, o Índice Brasileiro de Conectividade (IBC) é uma métrica criada pela agência para medir e comparar os níveis de conectividade em diferentes municípios e estados brasileiros e varia de 0 a 100, onde valores mais altos indicam melhor conectividade. A granularidade é o município e referem-se ao ano de 2023. O IBC é baseado em dados do setor de telecomunicações e considera várias variáveis, inclusive os dados de densidade de acessos de celulares e de banda larga que também foram utilizados neste trabalho:

- Densidade de acessos móveis de telefonia móvel: número de acessos de telefonia móvel dividido pela população, ponderado pela tecnologia do acesso (2G, 3G, 4G e 5G).
- Densidade de acessos de banda larga fixa: número de acessos de banda larga fixa dividido pela população, ponderado pela faixa de velocidade máxima contratada.
- Percentual da população coberta por telefonia móvel: estimativa da porcentagem da população coberta por telefonia móvel.
- Adensamento de estações: quantidade de estações rádio base (ERB) por 10.000 habitantes.
- Existência de backhaul de fibra ótica: presença de infraestrutura de fibra ótica nas localidades.
- Grau de competitividade de banda larga fixa: medido pelo inverso do índice de Herfindahl-Hirschman (HHI).
- Grau de competitividade de telefonia móvel: medido pelo inverso do índice de Herfindahl-Hirschman (HHI).

A granularidade dos dados de densidade da Anatel também é o município e referem-se ao mês de agosto de 2024.

O Ministério dos Transportes do Brasil disponibiliza dados abertos sobre a frota de veículos automotores por município. Esses dados são extraídos do Sistema Nacional de Registro de Veículos Automotores (RENAVAM) e são atualizados mensalmente. Foram utilizadas as informações de automóveis e o total de veículos por município para o mês de setembro de 2024. Os dados não vieram com o código dos municípios, o que dificultou a integração das informações pois os nomes dos municípios divergem entre as bases de dados. Foram disponibilizados os totais de automotores por município e, utilizando as informações de população do IBGE, obteve-se as informações médias por habitante (per capita). As informações de população por município foram acessadas através do Sistema IBGE de Recuperação Automática (SIDRA).

2. Desenvolvimento

Foram realizados vários mapas temáticos para avaliar a distribuição espacial da variável resposta e das variáveis independentes para identificar qualitativamente autocorrelações espaciais. Na comparação entre consumo total e consumo médio por consumidor, é possível perceber uma autocorrelação espacial com mais nitidez no mapa do consumo médio do que no mapa do consumo total devido à influência da densidade populacional (figura 1). Desta forma, optou-se pelo consumo médio residencial mensal, até porque a renda e demais variáveis também são per capita ou por domicílio.

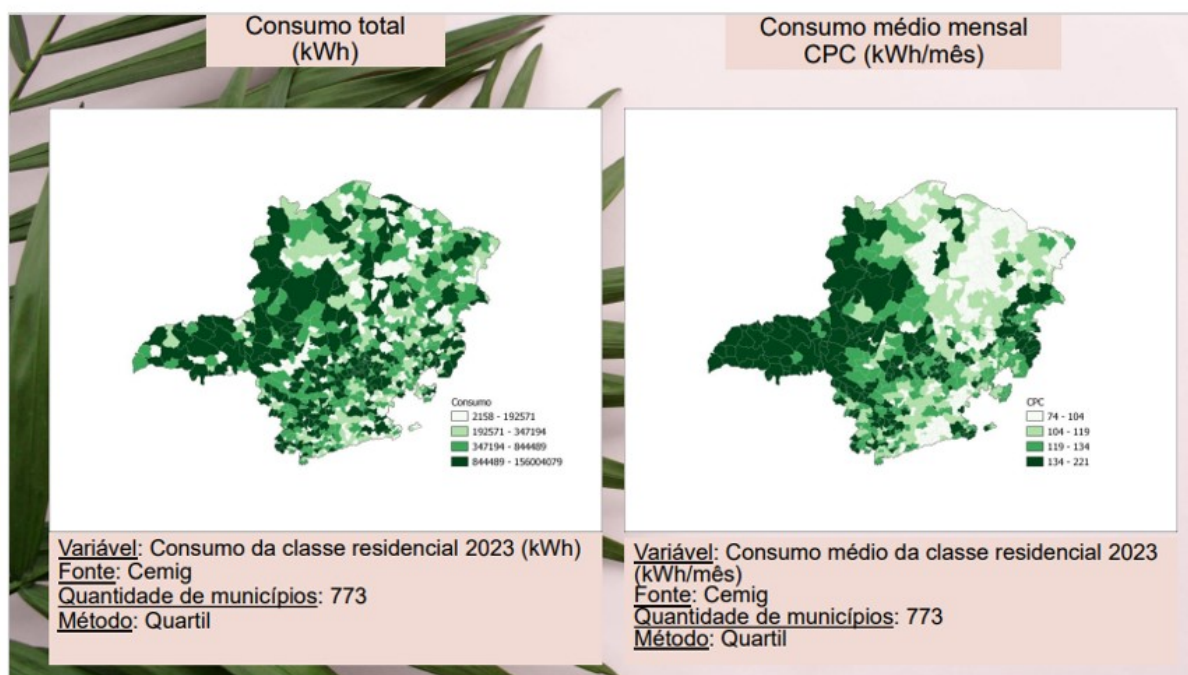


Figura 1: Comparando a distribuição espacial do consumo total de energia residencial e consumo médio residencial por consumidor

Uma análise qualitativa dos mapas de consumo médio mostra um padrão de dependência espacial entre as regiões do estado de Minas Gerais, com uma tendência decrescente do oeste para o leste. A autocorrelação espacial é bem nítida, com consumo médio maior em mesorregiões como Triângulo Mineiro/Alto Paranaíba, Noroeste de Minas e Sul/Sudoeste de Minas que fazem divisa com a região Centro-oeste e com o estado de São Paulo. O consumo médio vai reduzindo em direção ao leste de Minas Gerais, com menor consumo médio em regiões como Norte de Minas e Jequitinhonha.

Na comparação com as variáveis de Renda Média (figura 2) e PIB per capita (figura 3), o mesmo padrão de dependência espacial do consumo de energia é observado, com a mesma tendência decrescente do oeste para o leste, sugerindo uma forte correlação espacial, o que será analisado quantitativamente neste trabalho.

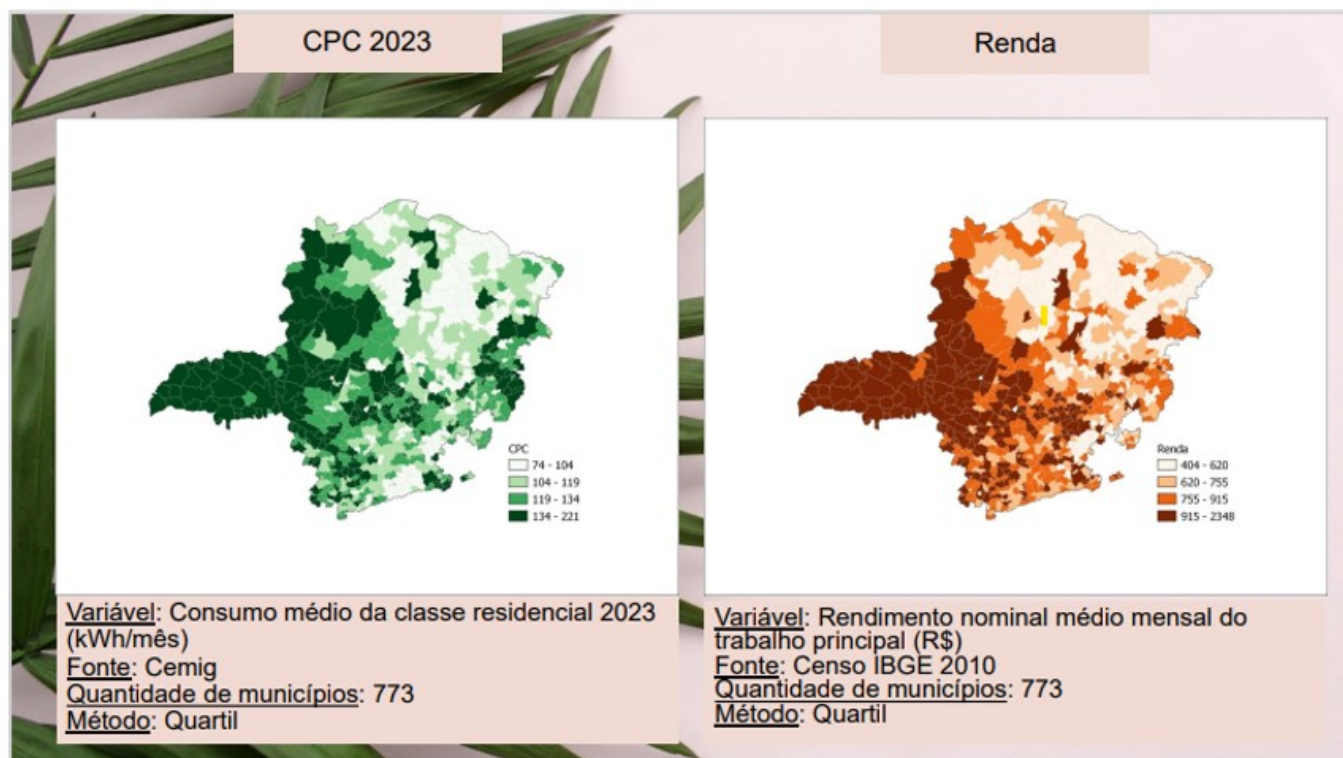


Figura 2: Comparando a distribuição espacial do Consumo médio residencial e da Renda

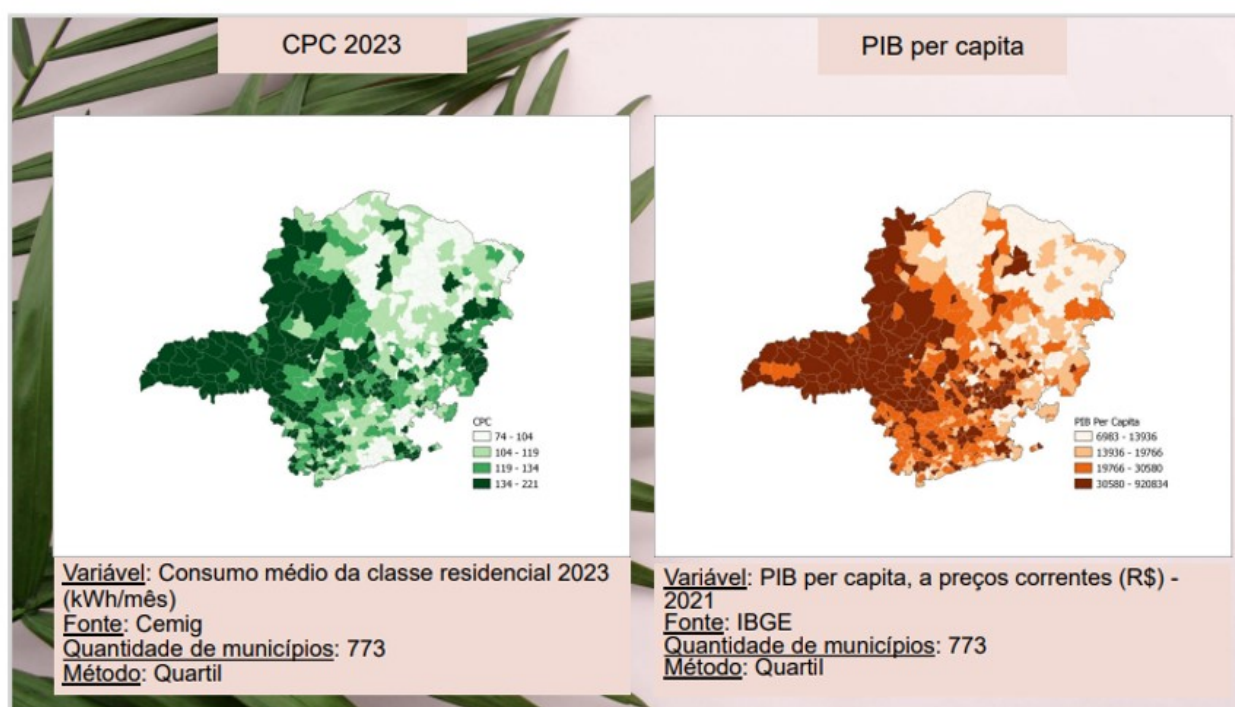


Figura 3: Comparando a distribuição espacial do Consumo médio residencial e do PIB per capita

Na comparação com as variáveis de Automóveis per capita (figura 4) e Veículos PIB per capita (figura 5), o padrão de dependência espacial observado é parecido com o do consumo de energia e com as regiões Norte e Jequitinhonha ainda sendo as regiões com as menores quantidades. No entanto, há diferença nas distribuições nas demais regiões, com aparente maior autocorrelação na quantidade de Automóveis por município do que na quantidade de Veículos.

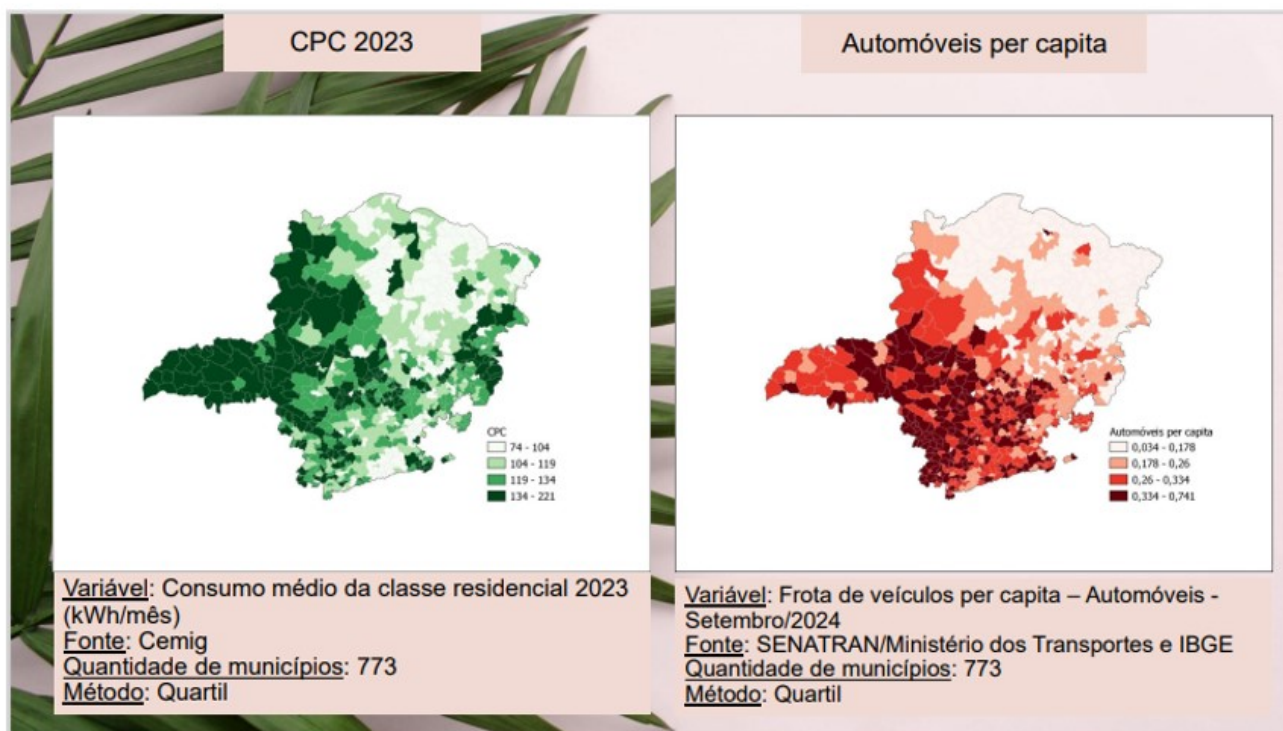


Figura 4: Comparando a distribuição espacial do Consumo médio e de Automóveis per capita

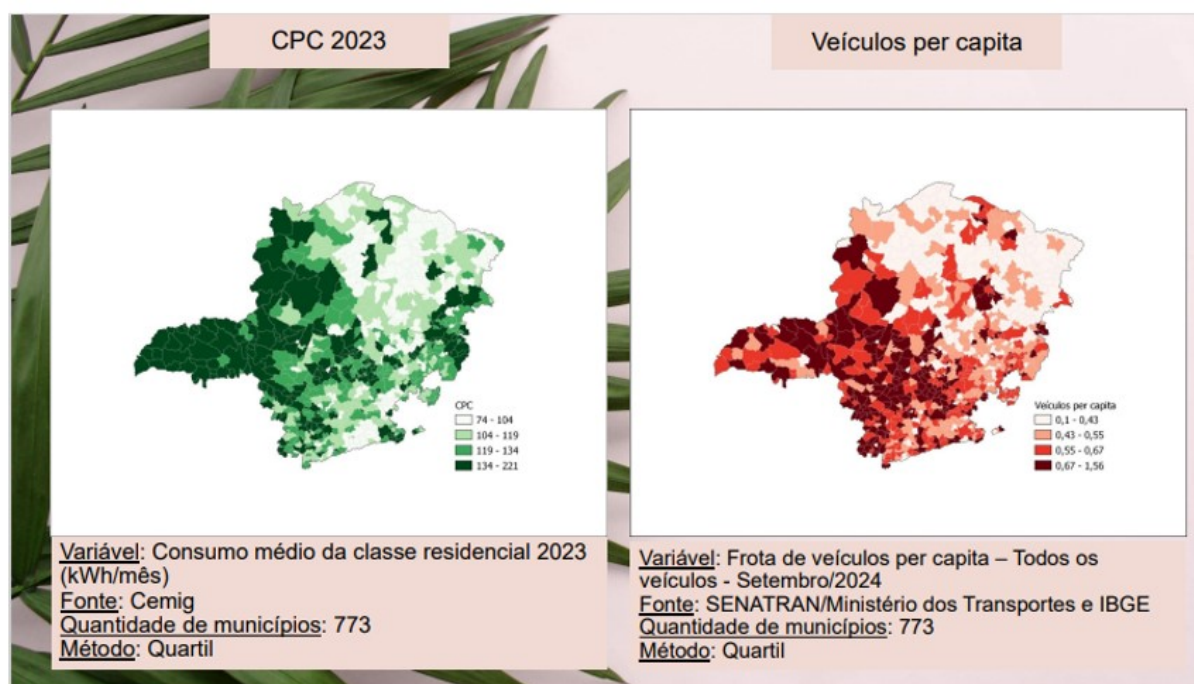


Figura 5: Comparando a distribuição espacial do Consumo médio residencial e de Veículos per capita

Na comparação com as variáveis de Celulares (figura 6), Banda Larga (figura 7) e IBC (figura 8), permanece o padrão de dependência com tendência decrescente do oeste para o leste, mas menos marcante do que o observado no consumo de energia e com um maior destaque para as regiões centrais do estado. As regiões mais a oeste não se destacam tão fortemente como no consumo de energia.

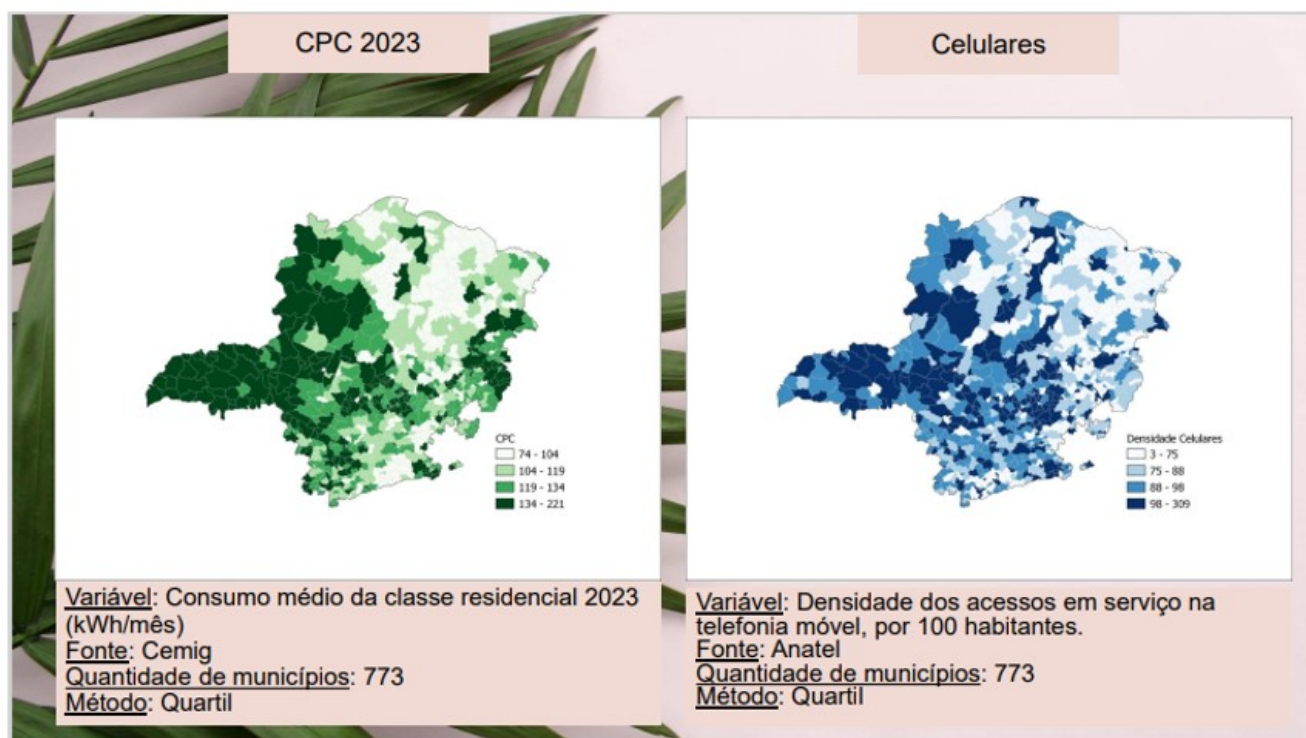


Figura 6: Comparando a distribuição espacial do Consumo médio residencial e de Celulares

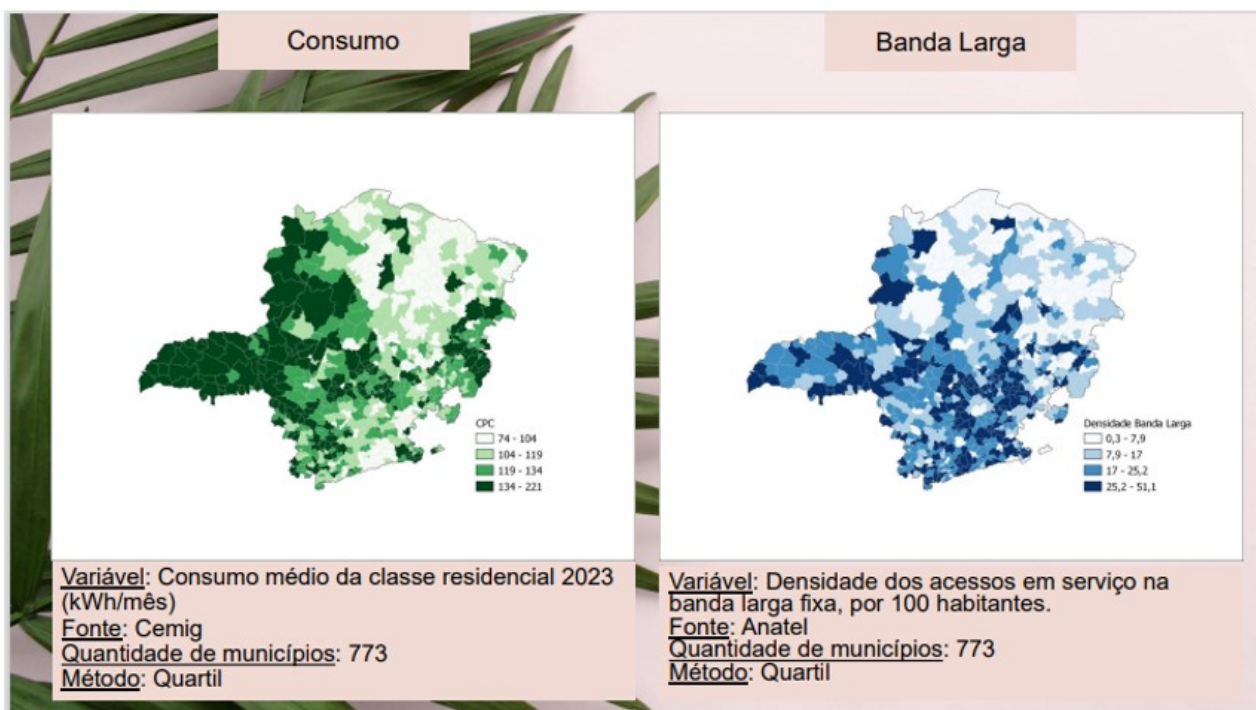


Figura 7: Comparando a distribuição espacial do Consumo médio residencial e de Banda Larga

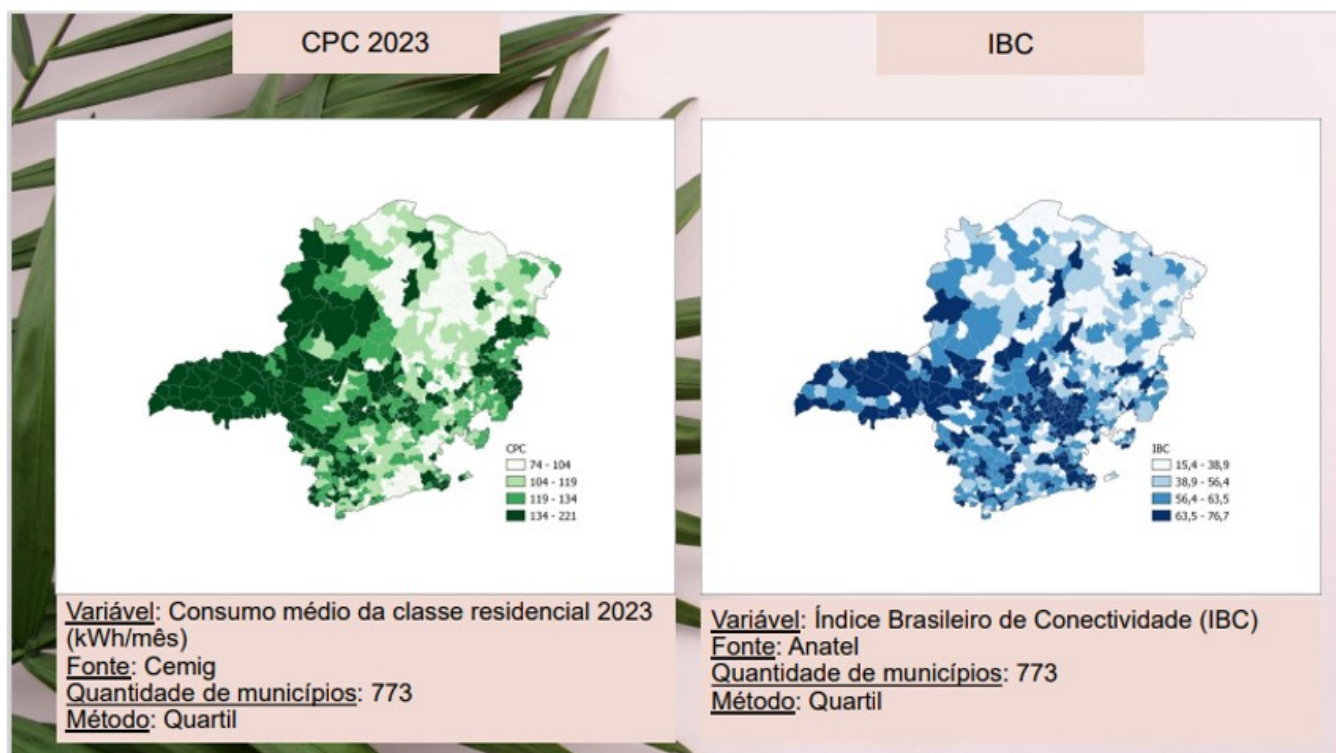


Figura 8: Comparando a distribuição espacial do Consumo médio residencial e do IBC

Para avaliar a relação entre todas as variáveis disponíveis e o consumo residencial médio implementou-se um modelo de regressão no software R com todas as variáveis e depois uma seleção de variáveis pelo modelo Stepwise Backward utilizando como parâmetro de seleção o AIC. O melhor modelo sugerido foi aquele com as variáveis de Renda, IBC, Veículos e Automóveis, sendo que as duas últimas estavam com coeficientes negativos. A variável de Automóveis foi retirada devido a estar com VIF acima de 5 e, na sequência, a variável Veículos foi retirada devido a não ter significância estatística. Assim, o modelo de regressão linear escolhido ficou com as variáveis de Renda e IBC. Este processo está resumido na figura 9. Os resultados deste modelo serão apresentados ao final conjuntamente aos modelos de regressão espacial.



Figura 9: Processo Regressão Múltipla – Seleção de variáveis

Considerando as variáveis selecionadas, realizou-se uma análise de correlação espacial através do cálculo do I de Moran e do LISA Map para avaliar a possibilidade de realizar regressões espaciais. Como indicado pela análise dos mapas temáticos, o Índice de Moran Global foi significativo para as três variáveis indicando um grau de autocorrelação espacial forte para as variáveis de consumo de energia (I de Moran = 0,5662) e de renda (I de Moran = 0,5300) e moderado para a variável IBC (I de Moran = 0,3698). Há, portanto, similaridade entre valores das variáveis em locais geograficamente próximos. Ou seja, a distribuição das variáveis na área de concessão da Cemig Distribuição não é aleatória. Como os índices são positivos, valores semelhantes estão espacialmente agrupados. Para identificar esses padrões de agrupamento e

tendências geográficas foram feitos os mapas no formato LISA MAP para as três variáveis (figuras 10, 11 e 12).

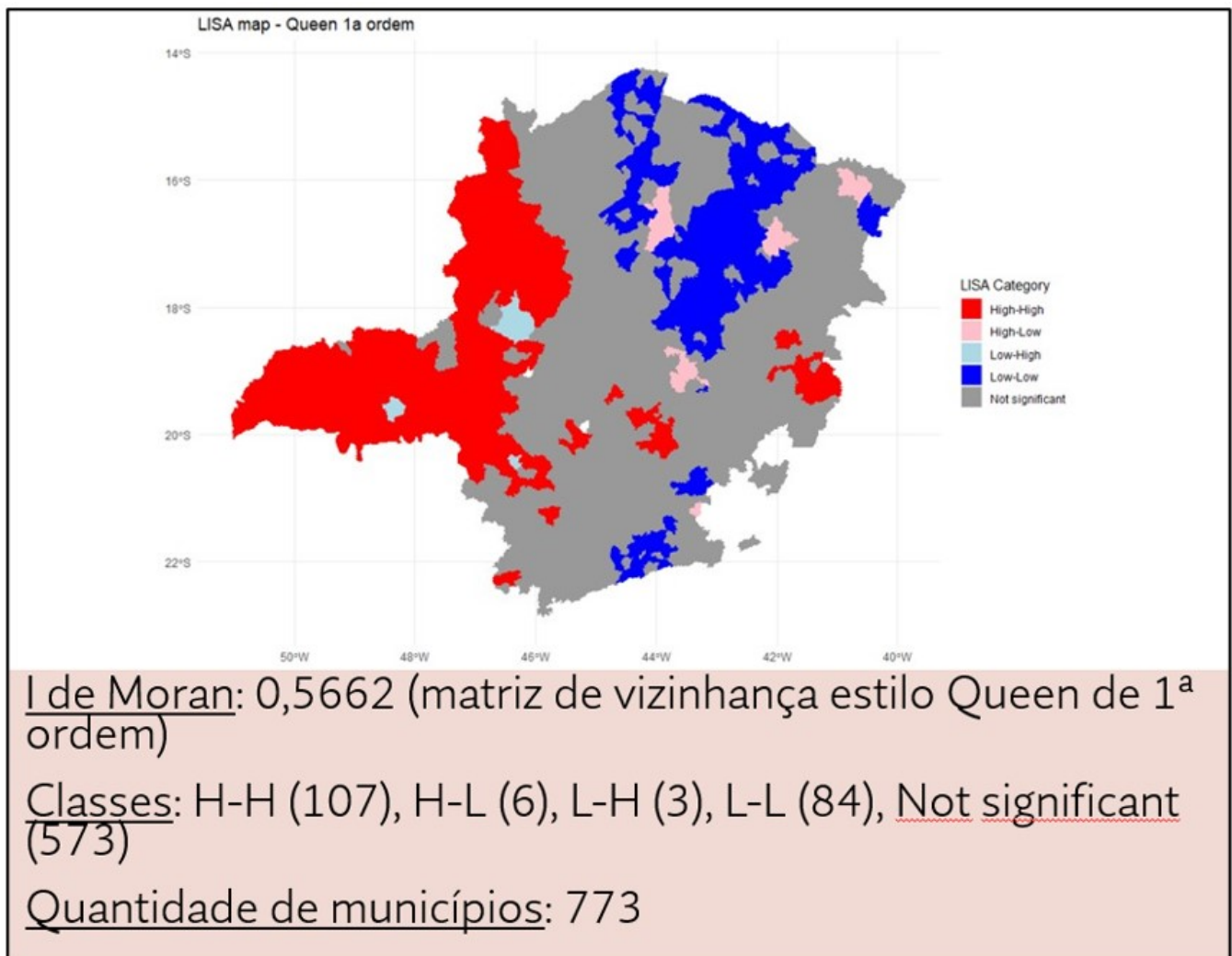


Figura 10: I de Moran Global e Lisa Map do Consumo médio residencial

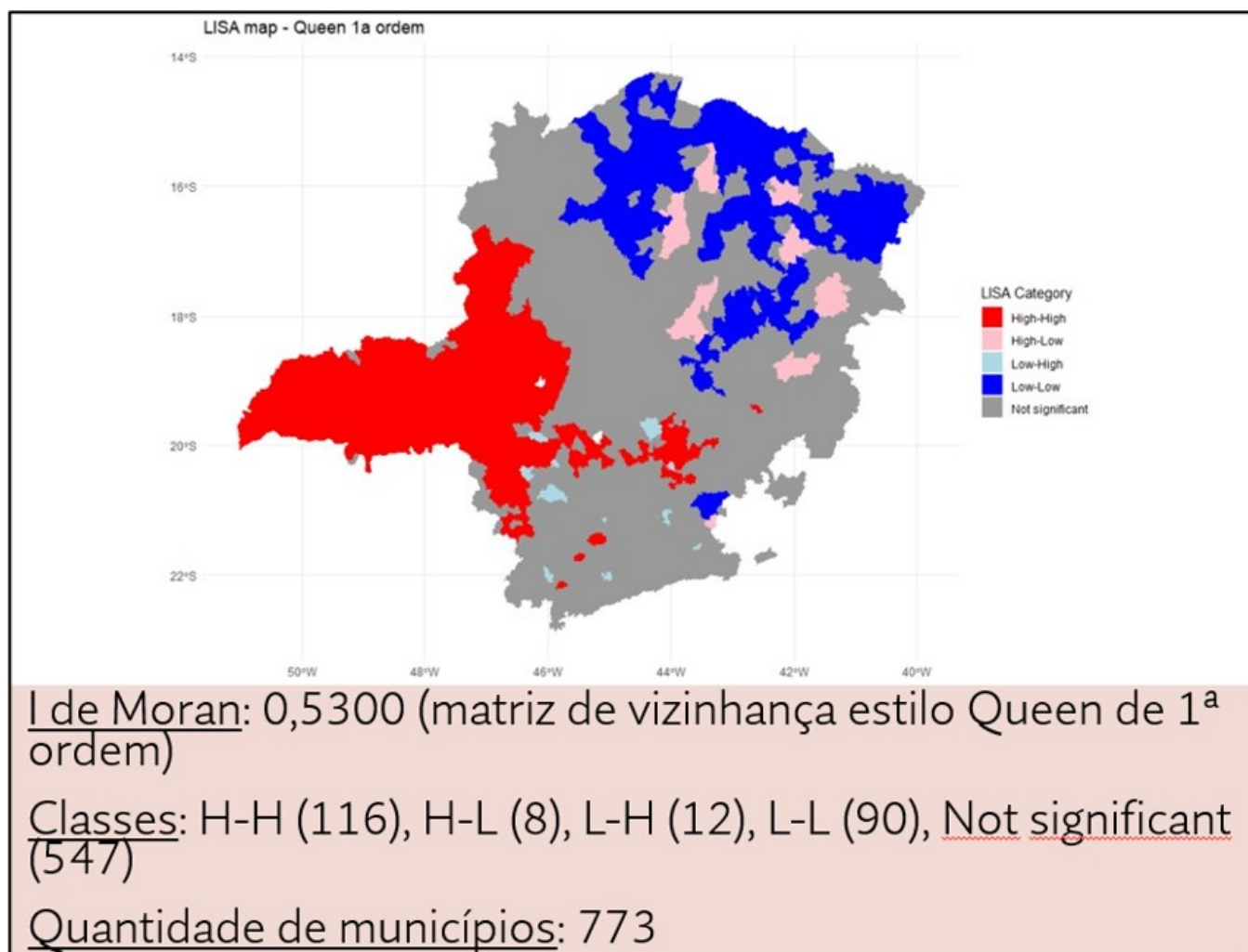


Figura 11: I de Moran Global e Lisa Map da renda média

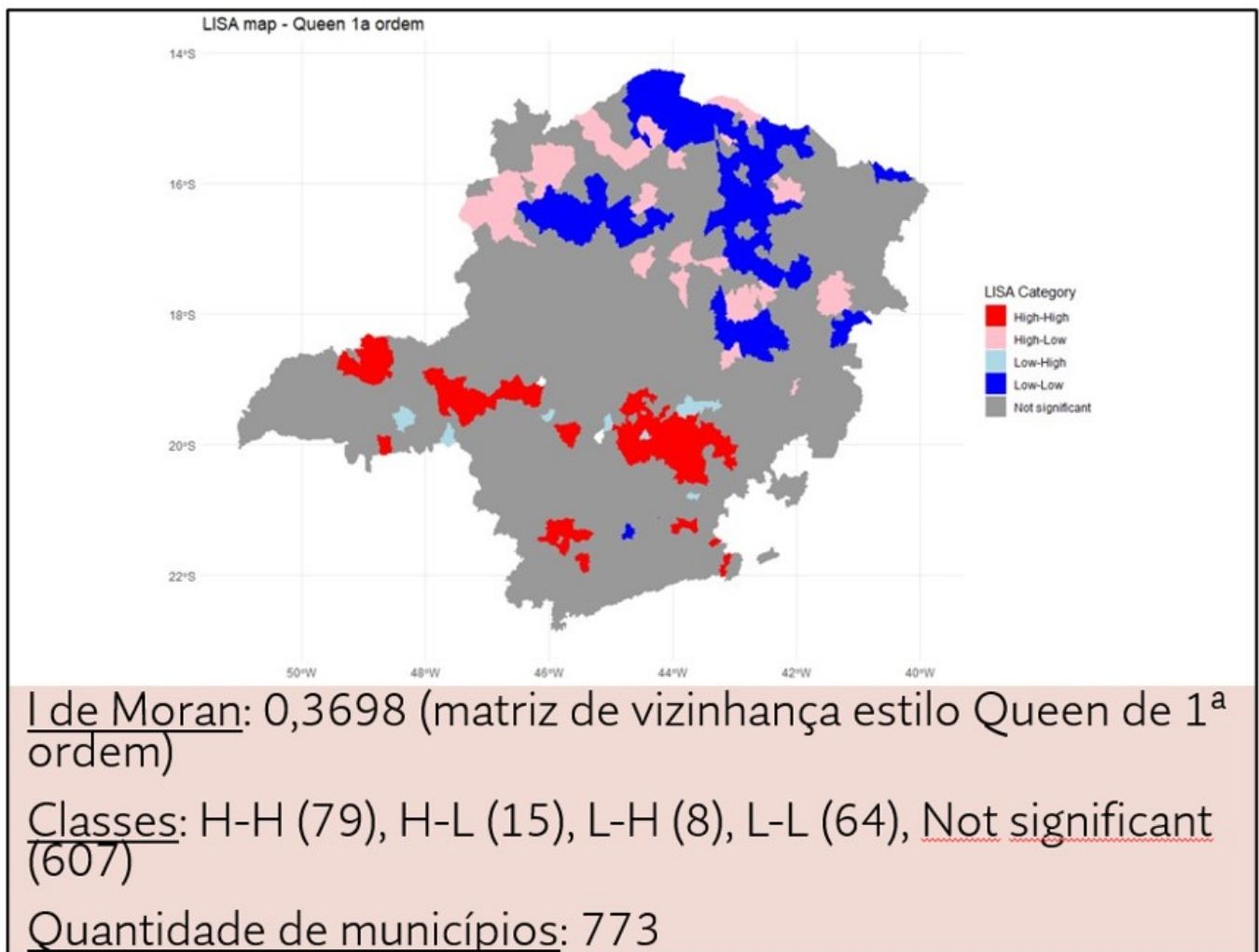


Figura 12: I de Moran Global e Lisa Map do IBC

Corroborando as análises qualitativas dos mapas temáticos das variáveis de consumo de energia e da renda, verifica-se um padrão de dependência espacial entre as regiões do estado de Minas Gerais, com uma tendência decrescente do oeste para o leste. Há um cluster High-High nas mesorregiões do Triângulo Mineiro/Alto Paranaíba e Noroeste de Minas e Low-Low nas regiões Norte e Jequitinhonha.

Assim, nestas áreas geográficas há padrões espaciais significativos de autocorrelação, mostrando onde os valores altos ou baixos de consumo de energia ou renda estão espacialmente agrupados. Esses clusters de autocorrelação positiva mostram regiões onde municípios com valores altos de consumo e renda estão próximos de outros municípios com valores altos (High-High) e municípios de baixo consumo e baixa renda estão próximos de outros municípios nas mesmas condições (Low-Low). Há alguns municípios com autocorrelação negativa, indicando áreas onde valores altos estão próximos de valores baixos (ou o contrário), configurando áreas de desigualdade espacial (High-Low ou Low-High).

Além do modelo de regressão linear (OLS), foram desenvolvidos modelos de regressão ponderada geograficamente (GWR) para diferentes tipos de kernel (gaussian, exponential, bisquare, tricube, boxcar) e modelos autorregressivos espaciais (SAR) para diferentes matrizes de vizinhança: k vizinhos mais próximos e adjacência de polígonos de 1ª e 2ª ordens. Lembrando que a variável resposta é o consumo médio residencial, com as variáveis explicativas sendo a renda média e o Índice Brasileiro de Conectividade (IBC). Um resumo dos resultados pode ser visto na tabela 2.

Tabela 2: Resultados dos modelos de regressão por ordem decrescente de R^2

Modelo de Regressão	R^2	Bandwidth (k)	Intercept	Coef. Renda	Coef. IBC	Rho
GWR Bisquare	0.8276	32				
GWR Tri-Cube	0.8197	32				
GWR Boxcar	0.7301	28				
GWR Gaussian	0.6970	21				
GWR Exponential	0.6796	21				
SAR 1st Order	0.6735		8.1196	0.0428	0.1854	0.5692
SAR 2nd Order	0.6452		9.5859	0.0426	0.1797	0.5640
SAR k near neigh	0.6359	21	15.0337	0.0431	0.1783	0.5093
LM Model	0.5262		58.9185	0.0641	0.2104	

O modelo linear simples tem um R^2 de 0,5262, indicando que cerca de 52,62% da variabilidade do consumo médio residencial é explicada pela renda média e pelo IBC. Este modelo sugere uma relação positiva entre o consumo de energia e as variáveis explicativas, com coeficientes significativos para a renda média e o IBC.

Os modelos GWR (Geographically Weighted Regression) levam em conta a variação espacial nas relações entre as variáveis explicativas e a variável resposta. Diferentes tipos de kernel foram usados para ajustar esses modelos, e a largura de banda foi selecionada utilizando o critério de informação de Akaike corrigido (AICc):

- GWR Bisquare: Com um R^2 de 0.8276 e uma largura de banda de 32, este modelo mostra a melhor capacidade de explicação entre os modelos GWR, indicando uma alta autocorrelação espacial e uma forte relação espacialmente variável entre o consumo de energia e as variáveis explicativas.
- GWR Tri-Cube: R^2 de 0.8197 e largura de banda de 32, também apresenta uma forte capacidade explicativa, embora ligeiramente inferior ao modelo bisquare.
- GWR Boxcar: R^2 de 0.7301 com largura de banda de 28, indicando uma explicação razoável da variabilidade dos dados, mas inferior aos modelos bisquare e tri-cube.
- GWR Gaussian: R^2 de 0,6970 e largura de banda de 21, demonstrando uma relação significativa, porém menos forte em comparação com os modelos anteriores.
-

GWR Exponential: R^2 de 0,6796 e largura de banda de 21, o que indica uma menor capacidade de explicação da variabilidade do consumo de energia.

Os modelos SAR levam em conta a dependência espacial entre as observações, ajustando-se com diferentes matrizes de vizinhança.

- SAR 1st Order: R^2 de 0,6735 com intercepto de 8,1196, coeficiente de renda de 0,0428, coeficiente de IBC de 0,1854 e rho (coeficiente de autocorrelação espacial) de 0,5692. A magnitude do coeficiente rho indica uma dependência espacial forte, mostrando que a autocorrelação espacial é significativa e que tanto a renda quanto o IBC têm impactos positivos no consumo de energia.
- SAR 2nd Order: R^2 de 0,6452 com intercepto de 9,5859, coeficiente de renda de 0,0426, coeficiente de IBC de 0,1797 e rho de 0,5640. A magnitude do coeficiente rho é semelhante à do modelo de 1ª ordem, sugerindo uma dependência espacial consistente.
- SAR k near neigh: R^2 de 0,6359 com largura de banda de 21, intercepto de 15,0337, coeficiente de renda de 0,0431, coeficiente de IBC de 0,1783 e rho de 0,5093. Este modelo utiliza uma matriz de vizinhança baseada nos 21 vizinhos mais próximos e mostra uma dependência espacial considerável, mas menor do que os modelos de 1ª e 2ª ordens.

Os modelos GWR, especialmente os que utilizam os kernels bisquare e tri-cube, apresentaram os maiores R^2 , indicando que considerar a variação espacial melhora significativamente a capacidade explicativa do modelo. O modelo OLS, embora simples, ainda fornece uma boa estimativa geral. Os modelos SAR destacam a importância da autocorrelação espacial, com coeficientes rho fortes, mostrando que a inclusão de dependência espacial melhora a explicação dos dados.

3. Conclusão

Tanto este trabalho quanto a tese de Francisco (2010) demonstram que o consumo mensal de energia elétrica residencial coletado pelas distribuidoras pode ser utilizado para estimar a renda dos domicílios devido à alta correlação entre os dois. Essa correlação permite que os dados de consumo de energia funcionem como uma espécie de censo de renda atualizado regularmente.

Os resultados dos modelos analisados indicam que tanto a renda média quanto o Índice Brasileiro de Conectividade (IBC) são preditores significativos do consumo médio residencial de energia. Em modelos de regressão simples, como o modelo linear (OLS), esses preditores já mostram uma influência substancial. Contudo, quando a dependência espacial é incorporada através de modelos espaciais, como a Regressão Ponderada Geograficamente (GWR) e os modelos autorregressivos espaciais (SAR), a complexidade e a precisão das previsões aumentam. A GWR, ao utilizar diferentes tipos de kernel e selecionar a largura de banda com base no critério de informação de Akaike corrigido (AICc), revela como a relação entre as variáveis explicativas e a variável resposta varia espacialmente, destacando regiões específicas onde a correlação é mais forte ou mais fraca.

Por outro lado, os modelos SAR, que consideram a autocorrelação espacial, mostram que a dependência espacial entre as observações é significativa e deve ser levada em conta. Os coeficientes rho nesses mod-

elos indicam uma autocorrelação espacial forte, ressaltando que áreas geograficamente próximas tendem a ter comportamentos similares no consumo de energia. A análise dos diferentes valores dos coeficientes entre os modelos revela que, enquanto a renda média mantém uma influência relativamente constante, o impacto do IBC diminui quando a dependência espacial é considerada. Isso sugere que a influência do IBC pode estar mais dispersa ou ser mais sensível a variações locais do que a renda média.

Os resultados mostraram uma correlação positiva significativa entre as variáveis analisadas, indicando que municípios com maior renda média e maior IBC (Índice Brasileiro de Conectividade) tendem a ter um consumo médio residencial maior. Os modelos de regressão geograficamente ponderada (GWR) e autor-regressivos espaciais (SAR) foram mais robustos na modelagem do consumo médio com base na renda e no IBC. Além disso, foi identificado um padrão de dependência espacial entre as regiões do estado de Minas Gerais, com uma tendência decrescente do oeste para o leste. Estes resultados contribuem para um planejamento mais preciso da compra de energia, da expansão da infraestrutura e da implementação de políticas de eficiência energética, garantindo que a oferta de energia atenda adequadamente às necessidades da população.

4. Referências bibliográficas

AGÊNCIA NACIONAL DE TELECOMUNICAÇÕES. Portal na Internet da Anatel. Disponível em: <<http://www.anatel.gov.br>>. Acesso em: 22 out. 2024.

BRASIL. Ministério dos Transportes. Portal na Internet do Ministério dos Transportes. Disponível em: <<http://www.transportes.gov.br>>. Acesso em: 22 out. 2024.

FRANCISCO, Eduardo de Rezende. Indicadores de Renda baseados em Consumo de Energia Elétrica: Abordagens Domiciliar e Regional na Perspectiva da Estatística Espacial. 2010. Tese (Doutorado em Administração de Empresas) - FGV-EAESP, São Paulo.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Portal na Internet do IBGE. Disponível em: <<http://www.ibge.gov.br>>. Acesso em: 22 out. 2024.

MELO, Rafael Oliveira; FREITAS, Anne Caroline de; FRANCISCO, Eduardo de Rezende; MOTOKANE, Marcelo Tadeu. Impacto das variáveis socioeconômicas no desempenho do Enem: uma análise espacial e sociológica. *Revista de Administração Pública*, 55, 1271–1294, 2022.