O Potencial de Uso de Variáveis Geográficas Derivadas de Imagens Orbitais em Modelos de Avaliação em Massa de Imóveis

The Potential Use of Geographic Variables Derived from Orbital Imagery in Mass Real Estate Valuation Models

Carlos Antonio Oliveira Vieira[1]

Caroline Bernardo Silva[2]

Hatan Pinheiro Silva [1]

Everton da Silva [1]

RESUMO: A utilização de variáveis geográficas derivadas de imagens orbitais tem se mostrado uma alternativa eficiente para enriquecer modelos de avaliação em massa de imóveis. Enquanto informações sobre o contexto urbano e ambiental frequentemente dependem de levantamentos de campo ou bases cadastrais incompletas e de difícil atualização, o sensoriamento remoto (SR) fornece dados padronizados, atualizados e com ampla cobertura espacial. A partir dessas imagens, é possível extrair indicadores relacionados à cobertura e uso da terra, densidade da ocupação, vegetação, impermeabilização do solo e outros atributos que caracterizam a vizinhança dos imóveis. Essas variáveis permitem representar, de forma objetiva, aspectos do contexto espacial que influenciam a formação dos preços imobiliários, contribuindo para aumentar a precisão, a robustez e a capacidade de generalização dos modelos de avaliação em massa. O objetivo deste artigo é analisar o potencial das imagens orbitais na extração de variáveis para valoração imobiliária. Como aplicação prática, foi realizado um estudo de caso no município de Curitiba, comparando modelos de Random Forest com e sem a inclusão de variáveis derivadas de imagens orbitais e processamentos realizados no *Google Earth Engine*. Os resultados demonstram que a inclusão dessas variáveis elevou o R^2 de 0,541 para 0,841 para apartamentos, de 0,460 para 0,676 para casas e de 0,066 para 0,650 para lotes, demonstrando sua contribuição para a estimação dos valores unitários.

Palavras-chave: Sensoriamento Remoto; Valoração Imobiliária; Variáveis Geográficas.

ABSTRACT: The use of geographic variables derived from satellite imagery has emerged as an effective approach to enhance mass real estate appraisal models. While information describing the urban and environmental context often relies on field surveys or incomplete and outdated cadastral databases, remote sensing provides standardized, up-to-date, and spatially comprehensive data. From satellite imagery, it is possible to derive indicators related to land cover and land use, building density, vegetation, impervious surfaces, and other neighborhood characteristics that influence property values. This study investigates the potential of orbital imagery for extracting geographic variables to support real estate valuation. A case study was conducted in the municipality of Curitiba, Brazil, comparing Random Forest models with and without the inclusion of variables derived from satellite imagery processed in Google Earth Engine. The results demonstrate that incorporating these variables substantially improved model performance, increasing the coefficient of determination (R^2) from 0.541 to 0.841 for apartments, from 0.460 to 0.676 for houses, and from 0.066 to 0.650 for vacant lots. These findings highlight the significant contribution of remotely sensed geographic variables to improving the accuracy, robustness, and generalization of mass appraisal models.

Keywords: Remote Sensing; Real Estate Valuation; Geographic Variables.

[1] Universidade Federal de Santa Catarina - Pós-Graduação em Engenharia de Transportes e Gestão Territorial – PPGTG - ECV - CTC / E-mail: carlos.vieira@ufsc.br, hatan.pinheiro@grad.ufsc.br, everton.silva@ufsc.br

[2] Grupo de Observação e Transformação do Território – GOTT - UFSC / E-mail: carolinebernardosilva@gmail.com

1 INTRODUÇÃO

A incorporação de variáveis geográficas derivadas de imagens orbitais em modelos de avaliação em massa de imóveis representa uma alternativa promissora para ampliar a capacidade de caracterização dos fatores que influenciam o valor dos bens imobiliários. Tradicionalmente, a obtenção de variáveis relacionadas ao contexto urbano, às condições ambientais e à qualidade da vizinhança depende de levantamentos de campo, cadastros detalhados ou bases de dados cadastrais frequentemente incompletas, desatualizadas ou indisponíveis em muitas localidades, tornando sua utilização onerosa e de difícil padronização. Em contrapartida, o sensoriamento remoto (SR) oferece uma fonte contínua, padronizada e de ampla cobertura espacial, permitindo extrair, de forma objetiva e sistemática, informações sobre cobertura e uso da terra, densidade da ocupação, presença de vegetação, impermeabilização do solo, proximidade de áreas verdes, corpos d'água e outros elementos que refletem o contexto espacial dos imóveis. Essas variáveis geográficas podem ser empregadas como indicadores indiretos da qualidade ambiental, da infraestrutura e da dinâmica urbana, contribuindo para representar de maneira mais consistente os atributos extrínsecos que afetam a formação dos preços imobiliários e, consequentemente, aumentar a precisão e a capacidade de generalização dos modelos de avaliação em massa.

Assim, o sensoriamento remoto oferece um potencial disruptivo para a valoração territorial. A integração do SR com Sistemas de Informação Geográfica (SIG) permite a extração de variáveis geográficas — como índices de vegetação (NDVI), densidade de ocupação e características topográficas — que exercem impacto direto no valor de mercado dos imóveis. Essas tecnologias possibilitam captar a heterogeneidade espacial e padrões de localização de forma mais ágil e econômica do que os métodos de avaliação tradicionais.

Como o objetivo de verificar a potencialidade das variáveis geográficas, apresenta-se um estudo de caso no município de Curitiba, no qual são comparados modelos de estimação do valor unitário com e sem a inclusão de variáveis derivadas de imagens orbitais e de processamentos digital de imagens em nuvem (*Google Earth Engine*). A análise contempla os segmentos de apartamentos, casas e lotes, permitindo avaliar o ganho de desempenho associado à incorporação de informações sobre a estrutura urbana e as características do entorno.

Embora o sensoriamento remoto seja amplamente empregado para a caracterização da paisagem urbana, sua aplicação na extração de variáveis geográficas para modelos de avaliação em massa de imóveis ainda permanece pouco explorada. A maioria dos modelos de valoração imobiliária utiliza predominantemente atributos cadastrais, estruturais e de localização, enquanto o potencial de índices espectrais e outras variáveis derivadas de imagens orbitais para representar, de forma objetiva e padronizada, o contexto urbano e ambiental dos imóveis ainda carece de investigações sistemáticas. Além disso, são escassos os estudos que quantificam o impacto dessas variáveis no desempenho preditivo de modelos baseados em técnicas de aprendizado de máquina, constituindo uma lacuna que este trabalho busca preencher.

Parte-se da hipótese de que variáveis geográficas derivadas de imagens orbitais, especialmente índices espectrais relacionados à cobertura da terra, vegetação, impermeabilização e densidade da ocupação urbana, capturam características do contexto

espacial que não são representadas pelos atributos cadastrais convencionais. Dessa forma, espera-se que sua incorporação em modelos de avaliação em massa de imóveis, baseados em algoritmos de aprendizado de máquina, resulte em ganhos significativos de desempenho, aumentando a capacidade de explicar a variabilidade dos valores imobiliários e produzindo estimativas mais precisas e robustas.

Nesse contexto, o presente artigo tem como objetivo geral analisar o potencial das imagens orbitais na extração de variáveis para valoração imobiliária. Especificamente, busca-se:

- Explorar metodologias para a derivação de variáveis geográficas, aplicados ao contexto urbano;
- Analisar, a partir do estudo de caso de Curitiba, a contribuição das variáveis derivadas do sensoriamento remoto para o desempenho de modelos de avaliação imobiliária em massa.

2 REVISÃO BIBLIOGRÁFICA

Esta revisão bibliográfica analisa as variáveis derivadas de sensoriamento remoto que exercem maior impacto no valor de um terreno, classificadas por suas dimensões físicas, ambientais e locacionais.

2.1 Imagens de Satélite e as Variáveis Geográficas

A utilização do sensoriamento remoto (SR) e de Sistemas de Informação Geográfica (SIG) revolucionou a avaliação em massa de imóveis, permitindo a extração de variáveis geográficas, que influenciam diretamente o valor da terra (COSTA, 1996; FAGUNDES, 2021).

Essas tecnologias possibilitam captar tendências de expansão urbana, identificar padrões de localização e monitorar alterações ambientais com rapidez e baixo custo em comparação a métodos tradicionais (COSTA, 1996).

A seguir, apresentam-se as variáveis derivadas de sensoriamento remoto que exercem maior impacto no valor de um terreno, classificadas por suas dimensões físicas, ambientais e locacionais.

2.1.1 Uso e Cobertura do Solo (LULC)

A classificação do uso e cobertura do solo é uma das aplicações mais fundamentais do SR na avaliação imobiliária. Imagens de satélites como Landsat, SPOT e CBERS permitem distinguir áreas urbanizadas, solo exposto, vegetação e corpos d'água (COSTA, 1996; BLEDOW, 2023; CHAO, 1998).

O monitoramento da mancha urbana permite identificar áreas em consolidação, no contexto da densidade de ocupação, onde a oferta de terrenos é escassa, elevando os preços (HONDA, 2021; CHAO, 1998).

Alternativamente através do *Modelo Linear de Mistura Espectral* (MLME), é possível extrair a fração sombra, que correlaciona-se negativamente com a fração substrato em áreas adensadas, servindo como indicador de verticalização e valorização do solo (PEDRASOLI,

2016).

2.1.2. Variáveis Ambientais e Índices de Vegetação

A presença de atributos naturais é um dos fatores que mais agrega valor "hedônico" ao imóvel (HONDA, 2021; SILVA, 2026).

Utilizando, por exemplo, *Índices de Vegetação* (NDVI e IVS), o sensoriamento remoto permite calcular a densidade de biomassa e arborização urbana (CHAO, 1998; SILVA, 2026). A proximidade de áreas verdes e parques é consistentemente apontada como um fator de valorização dos imóveis (HONDA, 2021; TRIVELLONI, 2005).

Em relação a *Qualidade Ambiental*, as imagens orbitais auxiliam na extração de variáveis como exposição solar, microclima e presença de corpos d'água (rios e lagos), que influenciam positivamente no bem-estar e no preço dos imóveis (EUGUINO e ERBA, 2020; HONDA, 2021; PEDRASOLI, 2016).

Em cidades litorâneas, a proximidade com a orla (Distância ao Mar), captada por bases cartográficas georreferenciadas, é a variável locacional de maior impacto, apresentando uma relação inversamente proporcional ao preço, ou seja, quanto menor a distância, maior o valor do imóvel (SILVESTRE, 2025; ZILLI, 2020).

2.1.3 Topografia e Riscos Geológicos

As variáveis morfométricas, derivadas de *Modelos Digitais de Terreno* (MDT) como os da missão SRTM, são cruciais para definir o potencial construtivo, naquela região, em análise (BLEDOW, 2023).

Terrenos com alta declividade possuem custos de construção elevados e maior suscetibilidade a escorregamentos, o que deprecia o valor da terra (CHAO, 1998; LIOTTE, 2000).

O sensoriamento remoto pode também ser utilizado para localizar e caracterizar áreas de risco geológico ou de inundação (CHAO, 1998; LIOTTE, 2000). Imóveis em regiões com potencial de desastres naturais recebem, frequentemente, tratamento tributário diferenciado e sofrem desvalorização no mercado (HONDA, 2021).

2.1.4. Acessibilidade e Infraestrutura Urbana

A integração do sensoriamento remoto com o SIG permite calcular variáveis de distância euclidiana ou por rede a polos de interesse (SILVA, 2026; ZILLI, 2020).

A distância a rodovias, estações de metrô e paradas de ônibus, no contexto de proximidade de vias e transporte, é um fator de calibração essencial em modelos de valor da terra (BLEDOW, 2023; SILVA, 2026; ZILLI, 2020).

O valor de um terreno é fortemente impactado pela proximidade a *Shopping Centers* (apontada como a variável de maior impacto em certas metrópoles), escolas, hospitais e centros comerciais, no contexto de serviços e amenidades (HONDA, 2021; SILVA, 2026).

A disponibilidade de redes de saneamento, iluminação pública e pavimentação (Infraestrutura Física), identificáveis através de bases cartográficas integradas a imagens de alta resolução, qualifica o ambiente urbano e eleva o valor venal dos imóveis (EUGUINO e ERBA, 2020).

2.2 Variáveis Geográficas Derivadas Do Sensoriamento Remoto

Imagens de satélite permitem a geração de diversos índices. O objetivo desta seção é explorar possibilidades de derivar variáveis, que poderiam ser usadas no processo de estimação dos valores dos imóveis, a partir de produtos do sensoriamento remoto.

Coleções das imagens do Sentinel-2 estão disponibilizadas gratuitamente na plataforma *Google Earth Engine* (GEE), gerando mosaicos com resolução de 10 metros.

Os produtos mensais são obtidos utilizando a técnica de Mosaico de Qualidade, criando mosaicos pela seleção dos pixels com o maior valor de NDVI (Índice de Vegetação por Diferença Normalizada) para cada posição.

As variáveis são então calculadas a partir desses mosaicos e podem ser classificadas de acordo com sua metodologia: - Variáveis calculadas a partir de índices obtidos de imagens de satélite.

Esses índices são obtidos utilizando técnicas algébricas com as bandas de sensoriamento remoto. Os índices utilizados seguiram a metodologia proposta pelo IDECOR (2025), e podem ser assim listados: o *Índice de Vegetação por Diferenças Normalizadas* (NDVI), o *Índice de Composição Biofísica* (BCI), o *Índice de Solo por Diferença Normalizada* (RNDSI), o *Índice Urbano* (UI), e o *Índice de Edificações por Diferença Normalizada* (NDBI).

Em todos os casos, após o desenvolvimento destes índices, espera-se que possam ser derivadas variáveis geográficas, considerando cada posição de amostras de mercado, com um raio de 500 metros, de modo que cada posição indique uma média e/ou mediana para seu entorno, de acordo com a metodologia proposta por IDECOR (2025).

2.2.1 Índice de Vegetação por Diferenças Normalizadas (NDVI)

O *Índice de Vegetação por Diferenças Normalizadas (NDVI)* é usado para medir e monitorar o crescimento das plantas, a cobertura vegetal e a produção de biomassa a partir de dados multiespectrais detectados pelo sensoriamento remoto. É calculado como uma divisão entre a diferença e a soma das bandas da região do Infravermelho Próximo (IVP) e da região do Vermelho (V), atmosfericamente corrigidas, do conjunto de dados do sensoriamento remoto:

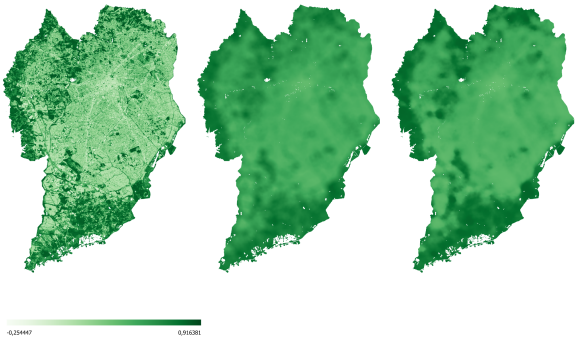
$$NDVI = (IVP - V) / (IVP + V) \quad (1)$$

O índice NDVI é calculado desta maneira, porque a vegetação saudável reflete muito bem na região do infravermelho próximo do espectro eletromagnético, mantendo baixa refletância na região do vermelho, e assim, produzindo altos valores para o NDVI.

Os valores obtidos para o NDVI variam entre [-1 a +1], mas a terra vegetada tem um NDVI variando de 0,1 a 0,7-0,8. Valores acima de 0,5 indicam vegetação densa e vigorosa.

A Figura 1 apresenta a imagem NDVI gerada para a área urbana do Município de Curitiba. Da esquerda para a direita, a primeira imagem é o NDVI em cada pixel. A segunda e a terceira imagens são, respectivamente, o valor do NDVI da média e mediana, para um raio de 500m a partir do pixel analisado, seguindo a metodologia proposta pelo IDECOR (2025).

Figura 1 – Imagem de NDVI para o Município de Curitiba.



Fonte: Autores (2026).

2.2.2 Índice de Composição Biofísica (BCI)

O *Biophysical Composition Index* (BCI) é um índice derivado de sensoriamento remoto usado para caracterizar a composição biofísica da superfície urbana, destacando a proporção relativa de vegetação, solo exposto e áreas construídas. Ele é bastante aplicado com imagens do satélite Sentinel-2.

A lógica do BCI não é uma simples razão entre bandas (como NDVI), mas sim uma combinação de componentes espectrais obtidos por transformação estatística.

2.2.2.1 Pré-processamento da imagem

Inicialmente utiliza-se imagens do Sentinel-2 nível Level-2A (reflectância de superfície), sendo as Bandas mais usadas: B2 (azul), B3 (verde), B4 (Vermelho), B8 (NIR), B11 (SWIR1), B12 (SWIR2). A coleção contendo 49 imagens obtidas, para o ano 2023, foram inicialmente corrigidas utilizando a máscara de nuvens SCL (do sistema Sentinel) e posteriormente reduzidas pelo valor da mediana de cada pixel na imagem. Como essas bandas possuem resolução espacial (escalas) diferentes, após essa seleção, tornou-se necessário reamostrar as bandas para mesma resolução espacial de 20m (IDECOR, 2025).

2.2.2.2. Transformação PCA (Análise de Componentes Principais)

O BCI é baseado na Análise de Componentes Principais. Para gerar o BCI, torna-se necessário empilhar as bandas espectrais e aplicar a transformação PCA. Obtém-se os três primeiros componentes: PC1 → brilho (*built-up*/solo); PC2 → vegetação; e, PC3 → umidade/sombra.

A fórmula do índice BCI normalmente utilizada, é:

$$BCI = (PC1 - PC2) / (PC1 + PC2) \quad (2)$$

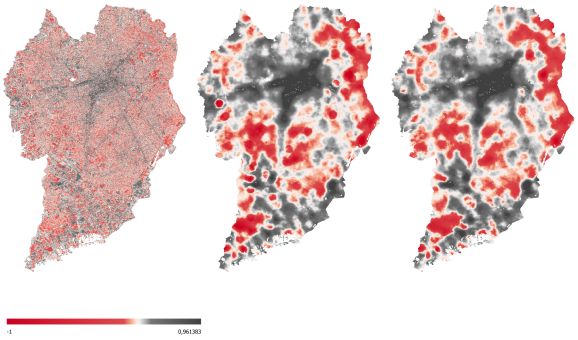
2.2.2.3. Interpretação do BCI

Os valores altos do BCI (próximos a +1), representam áreas construídas e/ou superfícies impermeáveis. Os valores baixos do BCI (próximos a -1), representam vegetação densa na imagem. E os valores intermediários do BCI, representam áreas com solo exposto e/ou transições.

O índice funciona porque o PC1 (brilho) aumenta com superfícies urbanas e solo, já o PC2 (vegetação) aumenta com cobertura vegetal. Logo, o BCI mede o contraste entre acidentes artificiais e vegetação.

A Figura 2 apresenta a imagem BCI gerada para a área de estudo do município de Curitiba, apresentando o valor de BCI, para cada pixel, média e mediana para um raio de 500m em torno do pixel analisado.

Figura 2 – Imagem de BCI para o Município de Curitiba.



Fonte: Autores (2026).

2.2.3 Índice de Solo por Diferença Normalizada (RNDSI)

O *Ratio Normalized Difference Soil Index* (RNDSI) é um índice espectral voltado para realçar solo exposto, reduzindo a influência da vegetação e da umidade. Este índice pode ser calculado diretamente com bandas do Sentinel-2, sem necessidade de aplicar o PCA.

Uma fórmula comum do RNDSI é:

$$RNDSI = (SWIR - NIR) / (SWIR + NIR) \quad (3)$$

onde, para o Sentinel-2 utilizam-se as bandas: NIR → Banda B8 (842nm) e a SWIR → Banda B11 (1610nm) (mais comum) (ou B12 em algumas variações)

Pode-se, neste caso, interpretar que os valores altos de RNDSI (próximos a +1) representam solo exposto seco e/ou áreas degradadas. Já os valores próximos a 0 representam mistura (solo + vegetação). E, os valores negativos para o RNDSI (próximos a -1) representam vegetação ou água.

Isso ocorre porque: Solo reflete mais no SWIR e Vegetação reflete mais no NIR. A Figura 3 apresenta a imagem RNDSI gerada para a área urbana do município de Curitiba, apresentando o valor de RNDSI, para cada pixel, média e mediana, respectivamente - para um raio de 500m em torno do pixel analisado.

2.2.4 Índice Urbano (UI)

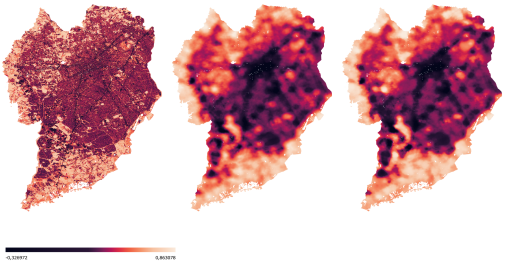
O *Urban Index (UI)* é um índice espectral simples e bastante usado para realçar áreas urbanizadas (superfícies impermeáveis) a partir de imagens do Sentinel-2.

A fórmula do *Urban Index* (UI) é:

$$UI = (SWIR - NIR) / (SWIR + NIR) \quad (4)$$

Observa-se que essa expressão é idêntica ao NDBI — na prática, muitos autores usam os termos como equivalentes ou muito próximos (IDECOR, 2025).

Figura 3 – Imagem de RNDSI para o Município de Curitiba.



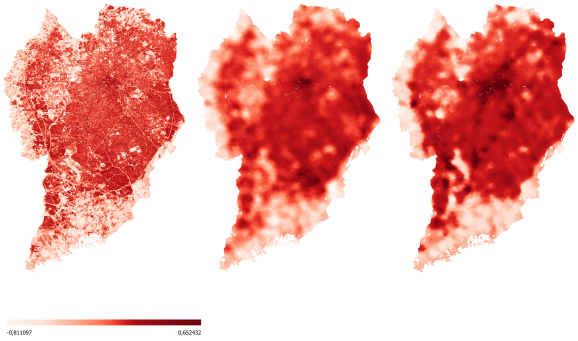
Fonte: Autores (2026).

Para o Sentinel-2 as bandas utilizadas são: NIR → B8 (10m); e, SWIR → B11 (20m) (mais comum) (ou B12 em algumas variações). Recomenda-se trabalhar na resolução espacial de 20m, para compatibilizar as bandas.

Pode-se neste caso interpretar que os valores altos de UI (próximos a +1) representam áreas urbanizadas (telhados, asfalto, concreto); Já os valores de UI próximos de 0 representam áreas mistas; e, os valores negativos de UI (próximos a -1) representam vegetação ou água.

Isso ocorre porque as superfícies urbanas refletem mais no SWIR e a cobertura de vegetação, reflete mais no NIR. A Figura 4 apresenta a imagem com o índice UI, gerada para a área urbana do município de Curitiba, apresentando o valor de UI, para cada pixel, média e mediana - respectivamente, para um raio de 500m em torno do pixel analisado.

Figura 4 – Imagem de UI para o Município de Curitiba.



Fonte: Autores (2026).

2.2.5 Índice de Edificações por Diferença Normalizada (NDBI)

O *Normalized Difference Built-up Index* (NDBI) é um dos índices mais usados para detectar áreas urbanizadas (superfícies construídas) a partir de imagens do Sentinel-2.

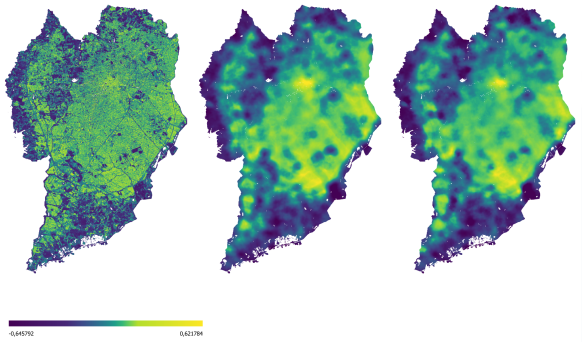
A fórmula do NDBI normalmente utilizada é:

$$NDBI = (SWIR - NIR) / (SWIR + NIR) \quad (5)$$

As bandas utilizadas do Sentinel-2 para gera o NDBI são: a banda NIR (Near Infrared) → B8 (10 m); e, a banda SWIR (*Shortwave Infrared*) → B11 (20m) (mais comum) (ou B12 como alternativa). Trabalha-se preferencialmente com resolução de 20m para compatibilizar as escalas.

Os valores positivos NDBI (próximos a +1) representam áreas construídas (urbano); já os valores de NDBI próximos de 0 representam áreas mistas; e, os valores de NDBI negativos (→ -1) representam vegetação e água. Isto porque as superfícies urbanas refletem mais no SWIR e, a vegetação, reflete mais no NIR. A Figura 5 apresenta a imagem NDBI gerada para a área urbana do município de Curitiba, apresentando o valor de BCI, para cada pixel, média e mediana - para um raio de 500m em torno do pixel analisado.

Figura 5 – Imagem de NDBI para o Município de Curitiba.



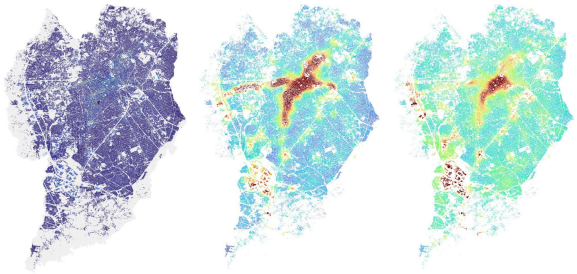
Fonte: Autores (2026).

2.2.6 Variáveis Relacionadas à Morfologia Urbana (*Open Building 2.5D*)

A caracterização da morfologia urbana é importante pois refletem a intensidade do uso do solo e a centralidade econômica de uma região (PEDRASSOLI, 2016). Para capturar essas características, utilizam-se dados da plataforma *Open Buildings 2.5D Temporal*, que disponibiliza estimativas de altura e presença de edificações a partir de imagens de satélite e modelos de aprendizado (*deep learning*) (GOOGLE, 2024). A partir desse conjunto de dados, foram derivadas seis variáveis morfológicas para o processo de estimação de valores: Altura estimada das edificações, Presença de pixel edificado, Densidade Percentual Construtiva, Fragmentação Urbana e Média e Mediana em um raio de 500m.

A altura das edificações é um dado nativo obtido através da matriz em formato raster disponibilizada pelo projeto *Open Building*, e mensura a altura aproximada das estruturas em metros para identificar o nível de verticalização do pixel. O Pixel edificado é uma variável binária que indica a existência de uma estrutura construída ou não em cada pixel dos limites urbanos e a densidade percentual é calculada a partir de um raio de 500m para cada pixel edificado. A Figura 6 apresenta o resultado da classificação dos pixels edificados.

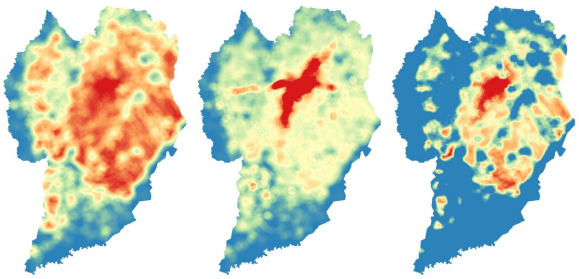
Figura 6 – Imagem da classificação dos pixels edificados para o Município de Curitiba.



Fonte: Autores (2026).

A fragmentação urbana classifica espacialmente pixels por sua densidade, seguindo os critérios de Angel et al. (2010), que estabelece um raio de 586m como de influência do pixel e o classifica sua densidade como Urbano Edificado Compacto (>50%), Urbano Edificado Disperso (entre 50% e 10%) ou Rural Edificado (< 10%) para os pixels edificados. O entorno dos pixels que não são edificados, mas apresentam densidades acima ou abaixo de 10% são classificados como Espaço Aberto Urbanizado e Espaço Aberto Rural, respectivamente, e a zona de transição entre essas áreas é uma borda de 100m chamada de Borda Urbana. A partir dos cálculos de altura das edificações foram extraídas as médias e medianas em um raio de 500m que definiram a altura média e mediana de seu entorno. A Figura 7 apresenta os resultados da densidade de áreas edificadas para o Município de Curitiba.

Figura 7 – Imagem da densidade das edificações e sua média e mediana para o Município de Curitiba.



Fonte: Autores (2026).

3 ESTUDO DE CASO: APLICAÇÃO DAS VARIÁVEIS DERIVADAS DO SENSORIAMENTO REMOTO EM CURITIBA

Com o objetivo de avaliar o potencial das variáveis derivadas de imagens orbitais na

valoração imobiliária, desenvolveu-se um estudo de caso para o município de Curitiba, no Estado do Paraná. A escolha do município justifica-se pela disponibilidade de base amostral imobiliária saneada e pela possibilidade de integração com variáveis geográficas extraídas por meio de processamento em ambiente *Google Earth Engine* (GEE).

A análise teve como finalidade verificar se atributos obtidos a partir de sensoriamento remoto, tais como índices espectrais, densidade urbana e altura estimada de edificações, são capazes de acrescentar informação relevante aos modelos de estimação do valor unitário dos imóveis. Dessa forma, o estudo de caso complementa a discussão teórica apresentada nas seções anteriores, demonstrando uma aplicação prática das variáveis orbitais no contexto da avaliação em massa de imóveis.

3.1 Base de dados e variáveis utilizadas

A base utilizada no estudo corresponde a uma amostra de imóveis anunciados no município de Curitiba, previamente tratada e saneada, integrada a uma camada geoespacial contendo atributos extraídos de imagens orbitais e processamentos derivados.

A camada geoespacial utilizada continha 469.707 pontos e 26 colunas. A integração com a base imobiliária foi relacionada a seleção dos anúncios de Curitiba, onde foram identificados 103.985 registros, correspondentes a 66.625 apartamentos, 27.897 casas e 9.463 lotes.

Para a modelagem, adotou-se a análise de casos completos, sendo excluídas as observações com valores ausentes na variável dependente, nas variáveis do *baseline* ou em qualquer uma das variáveis geográficas adicionadas. Após esse procedimento, foram utilizados 18.476 apartamentos, 6.831 casas e 1.954 lotes. A redução ocorreu principalmente devido à ausência de classificação tipológica e das métricas de altura média e mediana no entorno de 500 metros.

A base imobiliária foi constituída por anúncios de venda extraídos do portal Viva Real por meio de procedimento automatizado de raspagem de dados (*web scraping*) (VIVA REAL, [s.d.]). Os valores utilizados correspondem aos preços de oferta informados pelos anunciantes, não a valores efetivamente transacionados. Os registros presentes na amostra de Curitiba possuem datas de anúncio e de coleta compreendidas entre 6 de junho de 2023 e 1º de março de 2026.

As variáveis derivadas do sensoriamento remoto incluíram índices espectrais associados à vegetação, urbanização, solo exposto e composição biofísica da superfície. Foram considerados o Índice de *Vegetação por Diferença Normalizada* (NDVI), o *Índice de Composição Biofísica* (BCI), o *Ratio Normalized Difference Soil Index* (RNDSI), o *Urban Index* (UI) e o *Normalized Difference Built-up Index* (NDBI). Além dos valores pontuais, também foram utilizadas estatísticas de vizinhança, como médias e medianas dos índices no entorno das amostras (500m), seguindo a metodologia proposta pelo IDECOR (2025).

Também foram incorporadas variáveis relacionadas à estrutura urbana, como: altura estimada de edificações, presença de pixel edificado, densidade percentual, segmentação urbana e métricas de altura média e mediana em raio de 500 metros. Essas variáveis buscam representar características do entorno que podem influenciar o valor imobiliário, como intensidade construtiva, grau de urbanização, presença de cobertura vegetal e composição da

superfície urbana.

3.2 Estratégia metodológica

A análise foi conduzida em duas etapas complementares. Inicialmente, avaliou-se a associação individual entre as variáveis derivadas do sensoriamento remoto e o logaritmo do valor unitário dos imóveis. Essa etapa permitiu identificar sinais preliminares de relação entre os atributos geográficos e os preços observados, separadamente por segmento de mercado: apartamentos, casas e lotes.

Em seguida, foi realizado um teste de ganho incremental por meio de modelos de aprendizado de máquina. Para isso, foram comparados dois cenários: um modelo de referência, denominado *baseline*, contendo apenas variáveis básicas dos imóveis, e um segundo modelo, com a mesma estrutura, acrescido das variáveis derivadas do sensoriamento remoto.

O modelo *baseline* foi definido como uma referência simples. Para apartamentos e casas, foram utilizados o logaritmo da área, o número de vagas de garagem, o número de banheiros e o número de suítes. Para os lotes, foi utilizado o logaritmo da área. Não foram incluídas no *baseline* variáveis locacionais convencionais, como distância ao centro, renda do entorno, acessibilidade ou proximidade de equipamentos urbanos. Consequentemente, os resultados representam o ganho incremental das variáveis geográficas em comparação com essa referência parcimoniosa, não demonstrando que as imagens orbitais superem todas as demais formas de representação da localização.

O algoritmo utilizado foi o *Random Forest Regressor*, método de aprendizado de máquina baseado na combinação de múltiplas árvores de decisão. Cada árvore é ajustada sobre subconjuntos aleatórios dos dados e das variáveis explicativas, e a predição final é obtida pela média das predições individuais. Essa abordagem reduz a variância do modelo, permite capturar relações não lineares e interações entre variáveis, sendo adequada para problemas em que os fatores explicativos apresentam comportamento espacial e urbano complexo.

Para cada segmento, a amostra foi dividida aleatoriamente em 80% para treinamento e 20% para teste, utilizando-se a semente aleatória 42. A mesma divisão foi aplicada ao modelo baseline e ao modelo acrescido das variáveis geográficas. Foram utilizadas 120 árvores, mínimo de cinco observações por folha, seleção de *sqrt(p)* variáveis candidatas em cada divisão e amostragem de 80% das observações para o ajuste de cada árvore. Não foi estabelecido limite máximo para a profundidade das árvores.

Assim, para apartamentos, foram utilizadas 14.780 observações no treinamento e 3.696 no teste. Para casas, foram utilizadas 5.464 observações no treinamento e 1.367 no teste. Para lotes, foram utilizadas 1.563 observações no treinamento e 391 no teste.

Como métricas de avaliação foram utilizados o coeficiente de determinação (R^2), o erro quadrático médio na raiz (RMSE) e o erro absoluto médio (MAE). O R^2 mede a proporção da variabilidade do valor unitário explicada pelo modelo, enquanto o RMSE e o MAE indicam a magnitude dos erros de predição. Assim, um ganho consistente é caracterizado pelo aumento do R^2 e pela redução simultânea dos erros RMSE e MAE.

3.3 Resultados do teste de ganho incremental

Os resultados indicaram que a inclusão das variáveis geográficas aumentou o

desempenho dos modelos nos três segmentos analisados. Para apartamentos, o R^2 passou de 0,541 no modelo *baseline* para 0,841 no modelo ampliado. Para casas, aumentou de 0,460 para 0,676 e, para lotes, de 0,066 para 0,650. As reduções observadas no RMSE e no MAE confirmam que o aumento do R^2 foi acompanhado pela diminuição dos erros de predição, conforme apresentado na Tabela 1.

Tabela 1 – Comparação entre modelo *baseline* e modelo com variáveis geográficas em Curitiba

Segmento	Modelo	R^2	RMSE	MAE
Apartamento	Baseline	0,541	0,305	0,235
Apartamento	Baseline + variáveis geográficas	0,841	0,179	0,131
Casa	Baseline	0,460	0,282	0,216
Casa	Baseline + variáveis geográficas	0,676	0,219	0,164
Lote	Baseline	0,066	0,452	0,356
Lote	Baseline + variáveis geográficas	0,650	0,277	0,206

Fonte: Autores (2026).

Esses resultados demonstram que as variáveis extraídas de imagens orbitais e processamentos geoespaciais acrescentam informação relevante aos modelos de valoração imobiliária. O ganho observado para apartamentos indica que características do entorno, como intensidade construtiva, densidade e composição biofísica da superfície, ajudam a explicar diferenças de valor unitário que não são capturadas apenas por atributos internos do imóvel.

No caso das casas, o aumento de desempenho também foi relevante, embora inferior ao observado para apartamentos e lotes (Figura 8). Esse resultado pode estar associado à maior heterogeneidade do segmento de casas, no qual características internas, padrão construtivo, tamanho do terreno e atributos locais tendem a apresentar maior variabilidade. Ainda assim, a melhora do modelo evidencia que o contexto urbano e ambiental, captado pelas imagens orbitais, possui capacidade explicativa adicional.

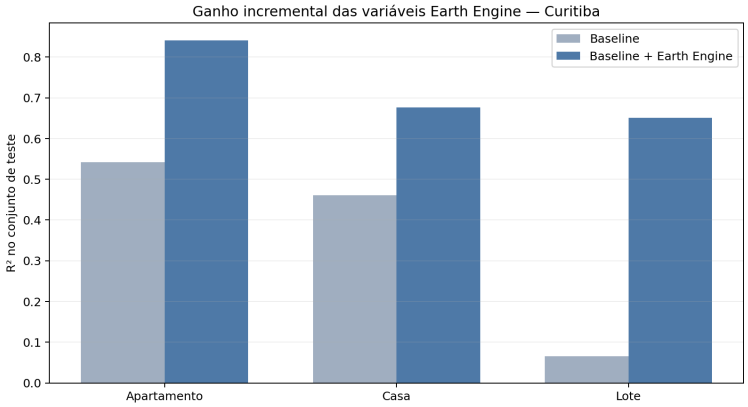
Para lotes, o ganho foi o mais elevado entre os segmentos analisados. Esse resultado é coerente com a natureza do mercado de terrenos, cujo valor depende fortemente das características locacionais e do potencial de uso do entorno. Como lotes não possuem atributos internos de edificação comparáveis aos de apartamentos e casas, variáveis como densidade urbana, altura do entorno, vegetação, solo exposto e grau de urbanização tornam-se especialmente relevantes para a diferenciação dos valores unitários.

3.4 Importância das variáveis geográficas no modelo

A importância por permutação foi calculada sobre o conjunto de teste do modelo acrescido das variáveis geográficas, utilizando o R^2 como função de pontuação. Cada variável foi embaralhada três vezes, com semente aleatória igual a 42.

Uma análise de importância por permutação tem a vantagem de avaliar a contribuição da variável em dados não utilizados no treinamento (capacidade de generalização do modelo), permitindo uma leitura mais robusta da relevância prática dos atributos. Entretanto, é importante destacar que essa técnica não indica causalidade, mas apenas a contribuição da variável para a predição dentro do modelo ajustado.

Figura 8 – Comparação entre modelo *baseline* e modelo com variáveis geográficas orbitais para Curitiba



Fonte: Autores (2026).

Para apartamentos, além de variáveis tradicionais como área, suítes e vagas de garagem, destacaram-se atributos associados à intensidade construtiva e à composição biofísica do entorno, como altura média dos imóveis no entorno de 500m de raio, BCI médio, altura média local, e os valores de mediana de RNDSI e BCI (Figura 9). Esse resultado sugere que a verticalização e a estrutura urbana próxima ao imóvel contribuem para diferenciar os valores unitários nesse segmento.

Para casas, as variáveis internas do imóvel, como número de suítes e área, permaneceram relevantes. Contudo, variáveis derivadas do sensoriamento remoto, como NDBI mediano, UI mediano, NDBI médio, alturas agregadas e RNDSI médio, também apresentaram contribuição importante (Figura 10). Esse comportamento indica que o grau de urbanização, a presença de superfícies construídas e a configuração do entorno, ajudam a explicar a variação do valor unitário das casas.

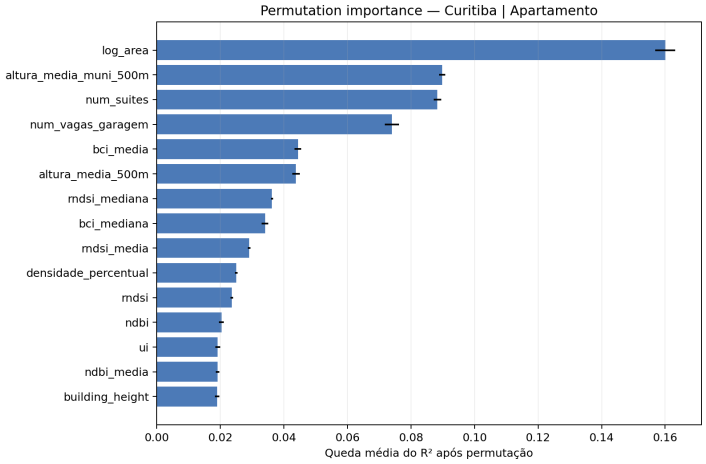
Para lotes, o logaritmo da área manteve elevada importância, mas os atributos de altura do entorno apresentaram forte contribuição adicional. Variáveis como altura média em 500 metros, altura mediana em 500 metros, BCI médio e mediano, RNDSI e NDVI também apareceram entre as mais relevantes (Figura 11). Esse resultado reforça que o valor de terrenos está fortemente associado ao contexto urbano em que estão inseridos, incluindo intensidade de ocupação, padrão de verticalização e características da cobertura do solo.

3.5 Discussão dos resultados

Os resultados obtidos para este estudo de caso realizado para a área urbana do município de Curitiba reforçam a hipótese de que variáveis derivadas de imagens orbitais podem contribuir para a avaliação em massa de imóveis. A melhora de desempenho observada nos

modelos indica que os atributos extraídos por sensoriamento remoto capturam dimensões da localização e da estrutura urbana que não são plenamente representadas por variáveis cadastrais ou características internas dos imóveis.

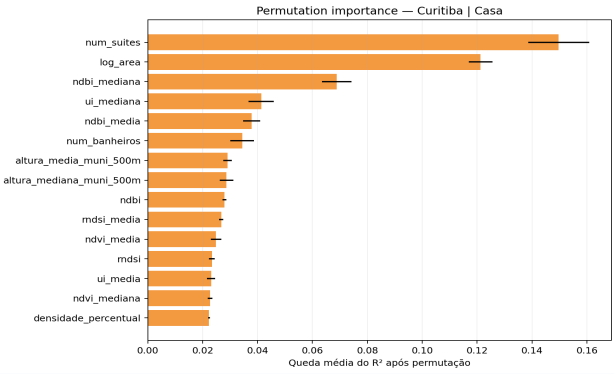
Figura 9 – Importância por permutação das variáveis para apartamentos em Curitiba



Fonte: Autores (2026).

A utilização de índices espectrais permite incorporar ao modelo informações sobre vegetação, solo exposto, áreas construídas e composição biofísica da superfície. Já as variáveis de altura e densidade ajudam a representar a intensidade de ocupação e a verticalização do entorno. Em conjunto, esses atributos ampliam a capacidade do modelo de reconhecer padrões intraurbanos associados à valorização ou desvalorização imobiliária.

Figura 10 – Importância por permutação das variáveis para casas em Curitiba

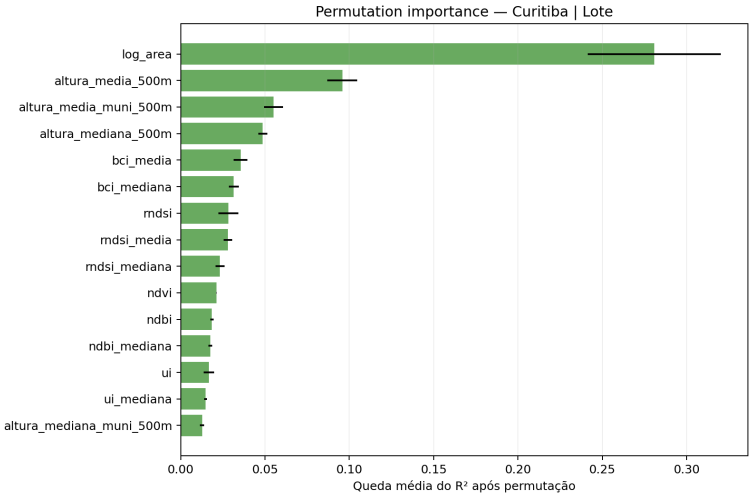


Fonte: Autores (2026).

Entretanto, a divisão entre treinamento e teste foi realizada de forma aleatória e, portanto, não impediu que imóveis espacialmente próximos ou pertencentes ao mesmo setor censitário fossem distribuídos entre os dois conjuntos. A dependência espacial residual pode tornar as métricas mais otimistas, uma vez que padrões locais semelhantes podem estar representados simultaneamente no treinamento e no teste. Dessa forma, os resultados devem ser interpretados como evidências exploratórias para a base analisada e para o contexto urbano de Curitiba, não constituindo demonstração de robustez espacial ou de capacidade de

generalização para outros municípios. Uma validação espacial por blocos e uma validação temporal deverão ser realizadas em trabalhos futuros.

Figura 11 – Importância por permutação das variáveis para lotes em Curitiba



Fonte: Autores (2026).

Ainda assim, este estudo de caso evidencia que o uso de imagens orbitais e plataformas de processamento digital de imagens em nuvem, como o *Google Earth Engine*, pode representar uma contribuição relevante para a construção de Plantas de Valores Genéricos mais aderentes à heterogeneidade espacial das cidades. A incorporação dessas variáveis permite aproximar os modelos de avaliação em massa das características reais do território, especialmente quando associada a métodos de aprendizado de máquina capazes de capturar relações não lineares.

4 CONSIDERAÇÕES FINAIS

A pesquisa demonstrou que as variáveis de localização, em especial aquelas derivadas de sensores remotos, permitem capturar a heterogeneidade espacial que os modelos tradicionais muitas vezes ignoram (SILVA, 2026). Técnicas avançadas como a *Regressão Geograficamente Ponderada* (GWR) e o *Random Forest* demonstram que o valor de um imóvel é fortemente afetado pelo seu entorno geográfico, tornando as variáveis derivadas do sensoriamento remoto, promissoras para a construção de Plantas de Valores Genéricos (PVG) acuradas (ZILLI, 2020; SILVESTRE, 2025; ZILLI et al., 2020).

Essa contribuição foi confirmada no estudo de caso realizado em Curitiba, no qual foram comparados modelos de *Random Forest* com e sem a inclusão de variáveis derivadas de imagens orbitais e de processamentos geoespaciais. Para apartamentos, o R^2 passou de 0,541 para 0,841; para casas, aumentou de 0,460 para 0,676; e, para lotes, passou de 0,066 para 0,650. Os resultados evidenciam que informações relacionadas à vegetação, à densidade urbana, à altura das edificações, à intensidade construtiva e à composição da superfície acrescentam capacidade explicativa aos modelos de estimação do valor unitário.

O ganho mais elevado observado para os lotes também reforça a importância das características locacionais nesse segmento, uma vez que terrenos não possuem atributos internos de edificação comparáveis aos de casas e apartamentos. Entretanto, por se tratar de

uma aplicação específica para Curitiba, os resultados não devem ser generalizados para outros municípios sem validações adicionais, considerando as diferenças de estrutura urbana, ocupação territorial e composição do mercado imobiliário.

Como sugestões para trabalhos futuros, recomenda-se replicar o estudo de ganho incremental em cidades com diferentes estruturas urbanas e características de mercado, a fim de verificar a estabilidade dos resultados e identificar quais variáveis derivadas do sensoriamento remoto apresentam maior capacidade de generalização.

REFERÊNCIAS

ANGEL, S.; PARENT, J.; CIVCO, D. L.; BLEI, A. M. **Atlas of Urban Expansion**. Cambridge: Lincoln Institute of Land Policy, 2010.

BLEDOW, G. D. Modelagem dinâmica espacial: cartografia prospectiva e o valor da terra. 2023. 88 f. **Dissertação** (Mestrado em Sensoriamento Remoto) -- Universidade Federal do Rio Grande do Sul, Porto Alegre, 2023.

CHAO, R. P. P. Análise da dinâmica urbana e de riscos geológicos na província Ciudad de La Habana, Cuba, através de técnicas de Sensoriamento Remoto e SIG. 1998. 138 f. **Dissertação** (Mestrado em Sensoriamento Remoto) -- Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 1998.

COSTA, S. M. F. Metodologia para análise do processo de expansão urbana através da integração de dados de Sensoriamento Remoto e SIG: aplicação à área metropolitana de Belo Horizonte. 1996. 210 f. **Tese** (Doutorado em Sensoriamento Remoto) -- Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 1996.

EGUINO, H.; ERBA, D. **Cadastro, avaliação imobiliária e tributação municipal: experiências para melhorar sua articulação e eficiência**. Washington, D.C.: Banco Interamericano de Desenvolvimento, 2020.

FAGUNDES, C. S. Integração de planta de valores com a base cartográfica municipal e geoprocessamento. 2021. 78 f. **Monografia** (Graduação em Engenharia Cartográfica) -- Universidade Federal do Rio Grande do Sul, Porto Alegre, 2021.

GOOGLE. Open Buildings 2.5D Temporal Dataset. Mountain View: Google Research, 2024. Disponível em: <https://research.google/resources/open-buildings/>. Acesso em: 15 jul. 2026.

HONDA, S. Variáveis relevantes na formação do valor de imóveis urbanos utilizando técnica de inteligência artificial. 2021. 60 f. **Trabalho de Conclusão de Curso** (Especialização em Geoprocessamento) - Universidade Federal de Minas Gerais, Belo Horizonte, 2021.

IDECOR. Estudio del Mercado de la Tierra Urbana de la Provincia de Córdoba 2025. **INFORME IDECOR - Infraestructura de Datos Espaciales**. Gobierno de la Provincia de Córdoba. 97 f., Córdoba, 2025.

LIOTTE, A. Uso de técnicas de geoprocessamento e sensoriamento remoto como subsídio ao planejamento físico-territorial do município de Bananal-SP. 2000. 185 f. **Dissertação** (Mestrado em Sensoriamento Remoto) -- Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2000.

PEDRASSOLI, J. C. Habitação e Sensoriamento Remoto: uma análise da expansão urbana na

RMSP por meio de imagens de satélite aplicando o modelo linear de mistura espectral. 2016. 246 f. **Tese** (Doutorado em Geografia Humana) -Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2016.

SILVA, C. Variáveis de localização na avaliação em massa de imóveis urbanos: sistematização, modelagem conceitual e automação. 2026. 160 f. **Dissertação** (Mestrado em Engenharia de Transportes e Gestão Territorial) - Universidade Federal de Santa Catarina, Florianópolis, 2026.

SILVESTRE, A. G. C. Avaliação em massa de terrenos usando regressão linear múltipla e random forest: um estudo comparativo no bairro de Jurerê, Florianópolis (SC). 2025. 115 f. **Trabalho de Conclusão de Curso** (Graduação em Engenharia Civil) - Universidade Federal de Santa Catarina, Florianópolis, 2025.

TRIVELLONI, C. A. P. Método para determinação do valor da localização com uso de técnicas inferenciais e geoestatísticas na avaliação em massa de imóveis. 2005. 172 f. **Tese** (Doutorado em Engenharia Civil) - Universidade Federal de Santa Catarina, Florianópolis, 2005.

VIVA REAL. Imóveis, casas e apartamentos para compra, venda e aluguel. São Paulo: Grupo OLX, [s.d.]. Disponível em: https://www.vivareal.com.br/

ZILLI, C. A. Regressão geograficamente ponderada aplicada na avaliação em massa de imóveis urbanos. 2020. 192 p. **Dissertação** (Mestrado em Engenharia de Transportes e Gestão Territorial) - Universidade Federal de Santa Catarina, Florianópolis, 2020.

ZILLI, C. A.; DROUBI, L. F. P.; HOCHHEIM, N. Avaliação em massa de apartamentos com uso de regressão geograficamente ponderada. **Valorem - Revista da Sociedade Brasileira de Engenharia de Avaliações**, v. 3, n. 1, p. 10-24, 2022.