

UNIVERSIDADE FEDERAL DO PARANÁ

RODRIGO FANTIN XAVIER DA SILVEIRA

COMPARAÇÃO ENTRE MODELOS PREDITIVOS PARA OTIMIZAÇÃO DA
ESTIMATIVA DE TEMPO EM SERVIÇOS DE GEORREFERENCIAMENTO DE
IMÓVEIS RURAIS

CURITIBA

2025

RODRIGO FANTIN XAVIER DA SILVEIRA

COMPARAÇÃO ENTRE MODELOS PREDITIVOS PARA OTIMIZAÇÃO DA
ESTIMATIVA DE TEMPO EM SERVIÇOS DE GEORREFERENCIAMENTO DE
IMÓVEIS RURAIS

Dissertação apresentada ao curso de Pós-Graduação em Ciências Geodésicas, Setor de Ciências da Terra, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestre em Ciências Geodésicas

Orientador: Prof. Dr. Caio Anjos Paiva

CURITIBA

2025

FOLHA DE FICHA CATALOGRÁFICA

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
UNIVERSIDADE FEDERAL DO PARANÁ
SISTEMA DE BIBLIOTECAS – BIBLIOTECA DE CIÊNCIA E TECNOLOGIA

Silveira, Rodrigo Fantin Xavier da

Comparação entre modelos preditivos para otimização da estimativa de tempo em serviços de georreferenciamento de imóveis rurais / Rodrigo Fantin Xavier da Silveira. – Curitiba, 2025.

1 recurso on-line : PDF.

Dissertação (Mestrado) - Universidade Federal do Paraná, Setor de Ciências da Terra, Programa de Pós-Graduação em Ciências Geodésicas.

Orientador: Caio Anjos Paiva

1. Propriedade rural - Georreferenciamento. 2. Aprendizado do computador. 3. Serviços - Precificação. I. Universidade Federal do Paraná. II. Programa de Pós-Graduação em Ciências Geodésicas. III. Paiva, Caio Anjos. IV. Título.

FOLHA DE APROVAÇÃO



MINISTÉRIO DA EDUCAÇÃO
SETOR DE CIÊNCIAS DA TERRA
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO CIÊNCIAS
GEODÉSICAS - 40001016002P6

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação CIÊNCIAS GEODÉSICAS da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **RODRIGO FANTIN XAVIER DA SILVEIRA**, intitulada: **Comparação entre modelos preditivos para otimização da estimativa de tempo em serviços de Georreferenciamento de Imóveis Rurais**., sob orientação do Prof. Dr. CAIO DOS ANJOS PAIVA, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 01 de Agosto de 2025.

Assinatura Eletrônica

06/08/2025 15:13:17.0

CAIO DOS ANJOS PAIVA

Presidente da Banca Examinadora

Assinatura Eletrônica

06/08/2025 15:24:56.0

ALEXANDRE BERNARDINO LOPES

Avaliador Externo (55001149)

Assinatura Eletrônica

06/08/2025 10:57:59.0

PAULO SERGIO DE OLIVEIRA JUNIOR

Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Centro Politécnico - Caixa Postal 19001 - CURITIBA - Paraná - Brasil

CEP 81531-980 - Tel: (41) 3361-3153 - E-mail: cpge@ufpr.br

Documento assinado eletronicamente de acordo com o disposto na legislação federal Decreto 8539 de 08 de outubro de 2015.

Gerado e autenticado pelo SIGA-UFPR, com a seguinte identificação única: 470507

Para autenticar este documento/assinatura, acesse <https://siga.ufpr.br/siga/visitante/autenticacaoassinaturas.jsp> e insira o código 470507

DEDICATÓRIA

Aos meus pais, José Volpi e Clarice Fantin, por todo suporte, amor e apoio.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por me guiar, me conceder as oportunidades e me dar forças e sorte, necessárias para perseverar em cada etapa desta jornada, desde a inscrição no Mestrado até a conclusão desta dissertação.

Ao meu orientador, Prof. Dr. Caio dos Anjos Paiva, expresso minha mais profunda gratidão. Sua paciência, seus conselhos valiosos, sua liderança e seu rigor acadêmico foram importantes para a condução desta pesquisa, para a minha formação e para a conclusão bem-sucedida do mestrado.

Agradeço aos membros da banca examinadora da defesa, Prof. Dr. Alexandre Bernardino Lopes, Prof. Dr. Paulo Sergio de Oliveira Jr., e da qualificação, Prof. Dr. Eduardo Vedor de Paula e Profa. Silvana Philippi Camboim pelas contribuições, questionamentos e sugestões que tanto enriqueceram a versão final deste trabalho.

Estendo meus agradecimentos à Universidade Federal do Paraná (UFPR) e ao Programa de Pós-Graduação em Ciências Geodésicas (PPCG), pela estrutura oferecida e pelo ambiente de excelência acadêmica. Agradeço também ao Laboratório de Geoprocessamentos e Estudos Ambientais (Lageamb) pelo apoio financeiro, que foi indispensável para a realização desta pesquisa.

Aos meus colegas de trabalho e amigos, agradeço o companheirismo, pelas discussões produtivas, viagens de campo, suporte e apoio mútuo ao longo dos desafios do mestrado.

Por fim, dedico esta conquista à minha família, meu alicerce fundamental. Gostaria de expressar minha imensa gratidão e amor aos meus pais, José Volpi e Clarice Fantin, sem os quais eu não teria o suporte necessário, muito menos a índole e o caráter para encarar os desafios do cotidiano. Ao meu pai, pelo exemplo de homem, trabalhador, honesto e companheiro. À minha mãe, por sua presença constante, seu carinho, sua compreensão nos momentos difíceis e por sempre desejar o melhor para mim.

EPÍGRAFE

“O mundo é um espelho, pois se sorris para ele,
ele sorrirá para ti”.

Gustave Le Bon

RESUMO

A determinação de preços para Serviços de Georreferenciamento de Imóveis Rurais (GIR) no Brasil é marcada por desafios, principalmente devido à falta de critérios objetivos e à alta incerteza na estimativa do tempo de execução em campo, que é a principal variável deste estudo, subjetividade esta, que gera dificuldades às partes envolvidas no processo. Em resposta a essa problemática, esta dissertação propõe, desenvolve e valida uma solução quantitativa baseada em modelagem preditiva. Utilizando um conjunto de dados de observações, obtidas durante os levantamentos topográficos georreferenciados nos imóveis rurais visitados, no interior do Paraná. Foram conduzidas análises comparativas entre diferentes classes de algoritmos, incluindo regressões lineares, modelos espaciais (globais e locais) e algoritmos de *Machine Learning*. A avaliação desses modelos foi realizada através de um protocolo de otimização de hiperparâmetros com validação cruzada (K-Fold), garantindo a seleção da melhor versão de cada algoritmo. Conclui-se que, embora a análise espacial tenha confirmado a presença de autocorrelação nos dados, a capacidade dos algoritmos de conjunto de capturar as interações complexas e não-lineares entre as variáveis de campo se mostrou mais decisiva para a precisão preditiva. Portanto, a adoção de modelos de *Machine Learning* representa uma ferramenta metodologicamente defensável e superior para mitigar a subjetividade nos orçamentos, permitindo uma otimização eficaz do planejamento e da precificação dos serviços de georreferenciamento rural (GIR). A hipótese a de que se for possível medir o tempo de deslocamento entre vértices, identificar e associar a as variáveis que o impactam, bem como determinar o melhor modelo ou algoritmo, então a capacidade dessas variáveis em influenciar o tempo, poderá ser testada e quantificada.

Palavras-chave: Georreferenciamento de Imóveis Rurais; Estimativa Temporal; Aprendizado de Máquina; Precificação de Serviços.

ABSTRACT

The pricing of Rural Property Georeferencing (GIR) services in Brazil is marked by challenges, primarily due to the lack of objective criteria and the high uncertainty in estimating fieldwork execution time, the main variable of this study. This subjectivity creates difficulties for the parties involved in the process. In response to this issue, this dissertation proposes, develops, and validates a quantitative solution based on predictive modeling, using a dataset of observations obtained during georeferenced topographic surveys on visited rural properties in the interior of Paraná. Comparative analyses were conducted across different classes of algorithms, including linear regressions, spatial models (global and local), and Machine Learning algorithms. The evaluation of these models was performed through a protocol of hyperparameter optimization with K-Fold cross-validation, ensuring the selection of the best version of each algorithm. It is concluded that, although spatial analysis confirmed the presence of spatial autocorrelation in the data, the ability of ensemble algorithms to capture the complex and non-linear interactions between the field variables proved to be more decisive for predictive accuracy. Therefore, the adoption of Machine Learning algorithms represents a methodologically defensible and superior tool for mitigating subjectivity in budgeting, allowing for an effective optimization of the planning and pricing of rural georeferencing (GIR) services. The hypothesis is that if it is possible to measure the travel time between vertices, identify and associate the variables that impact it, as well as determine the best model or algorithm, then the capacity of these variables to influence the time can be tested and quantified.

Key-words: Rural Property Georeferencing; Time Estimation; Machine Learning; Service Determination.

LISTA DE FIGURAS

FIGURA 1 – DIAGRAMA DA METODOLOGIA.....	45
FIGURA 2- DISPOSIÇÃO DOS ASSENTAMENTOS VISITADOS ENTRE 2023 E 2025.	46
FIGURA 3 - SEGMENTAÇÃO A PARTIR DA UMA IMAGEM DO SATÉLITE CBERS04A	48
FIGURA 4 - MDT PARA ANÁLISE DA TOPOGRAFIA A SER VISITADA, EXTRAÍDO DO SRTM.....	49
FIGURA 5 - PROJETO DE NAVEGAÇÃO NO APLICATIVO QFIELD.	50
FIGURA 6 - “BASE”, VÉRTICE UTILIZADO COMO REFERÊNCIA DE LEVANTAMENTOS.....	50
FIGURA 7 – VÉRTICE DO TIPO “M” (MARCO), PARA DEFINIÇÃO DOS LIMITES DOS IMÓVEIS.	51
FIGURA 8 - CAMINHAMENTO UTILIZANDO O APLICATIVO “LOCUS MAP”.....	52
FIGURA 9 - PROCESSAMENTO DE VÉRTICES RASTREADOS, NO SOFTWARE MAGNETTOOLS.	54
FIGURA 10 - DIREÇÃO DO DESLOCAMENTO (ID), COM OS VÉRTICES OCUPADOS.....	54
FIGURA 11 – EXEMPLO DA VARIÁVEL USO DO SOLO, DE ACORDO COM A DIFICULDADE IDENTIFICADA EM CAMPO.	55
FIGURA 12 - DISTINÇÃO DAS VARIÁVEIS.....	63
FIGURA 13 - MAPA DE CALOR DAS CORRELAÇÕES ENTRE AS VARIÁVEIS, ATRAVÉS DO ÍNDICE I DE MORAN. .	64
FIGURA 14 – OUTLIER IDENTIFICADO NO DESLOCAMENTO.....	71

LISTA DE GRÁFICOS

GRÁFICO 1 – MÉTRICAS ESPERADAS, RESULTADAS DA BIBLIOTECA <i>PyCaret</i>	66
GRÁFICO 2 - COEFICIENTE DE DETERMINAÇÃO (R^2) PARA ALGORITMOS ML.....	68
GRÁFICO 3 - RMSE E MAE PARA ALGORITMOS DE ML.....	69
GRÁFICO 4 – MAIORES IMPORTÂNCIAS DAS VARIÁVEIS NOS ALGORITMOS	70
GRÁFICO 5 – COMPARAÇÃO DA INFLUÊNCIA DE <i>OUTLIERS</i> , PARA O R^2	72
GRÁFICO 6 – COMPARAÇÃO DA INFLUÊNCIA DE <i>OUTLIERS</i> , ENTRE RMSE E MAE	73

LISTA DE TABELAS

TABELA 1 – EXEMPLO DA TABELA/PLANILHA FINAL DAS OBSERVAÇÕES, CONTENDO APENAS AS VARIÁVEIS DESCRITAS.	53
TABELA 2 – EXEMPLO DA VARIÁVEL QUE IDENTIFICA SE FOI OU NÃO UTILIZADO VEÍCULO.	55
TABELA 3 – EXEMPLO DE VARIÁVEL QUE IDENTIFICA A DISTÂNCIA PERCORRIDA.	55
TABELA 4 – EXEMPLO DA VARIÁVEL QUE IDENTIFICA O TEMPO DE DESLOCAMENTO ENTRE OS VÉRTICES.	56
TABELA 5 – EXEMPLO DA VARIÁVEL QUE IDENTIFICA A DECLIVIDADE DO DESLOCAMENTO.	56
TABELA 6 – EXEMPLO DA TABELA/PLANILHA DAS OBSERVAÇÕES, COMPLETA, COM VARIÁVEIS DEFINIDAS, COMENTÁRIOS FEITOS SOBRE CADA TRECHO, COORDENADAS DOS VÉRTICES E EQUIPE QUE REALIZOU O RASTREIO.	63
TABELA 7 – MÉTRICAS TREINAMENTO/TESTE 67/33%.....	67
TABELA 8 – MÉTRICAS OBTIDAS DO TREINAMENTO/TESTE 80/20%	67
TABELA 9 – MÉTRICAS OBTIDAS DA VALIDAÇÃO CRUZADA	68
TABELA 10 – MÉTRICAS PARA OS MODELOS SAR, GWR E OLS.....	68
TABELA 11 - IMPORTÂNCIA DAS VARIÁVEIS.....	70
TABELA 12 - MÉTRICAS DE DESEMPENHO NAS NOVAS OBSERVAÇÕES.	71
TABELA 13 - MÉTRICAS DE DESEMPENHO NAS NOVAS OBSERVAÇÕES SEM <i>OUTLIERS</i>	72

LISTA DE ABREVIATURAS OU SIGLAS

AHP - *Analytic Hierarchy Process*

ARP ou RPA - Aeronave Remotamente Pilotada

ART - Anotação de Responsabilidade Técnica

EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária

CNIR - Cadastro Nacional de Imóveis Rurais

GIR - Georreferenciamento de Imóveis Rurais

GNSS - *Global Navigation Satellite System*

GPS - *Global Positioning System*

GRP ou RGP - Regressão Geograficamente Ponderada

INCRA - Instituto Nacional de Colonização e Reforma Agrária

MTGIR – Manual Técnico de Georreferenciamento de Imóveis Rurais

CRM ou CRI - Cartório de Registro de Imóveis

LAGEAMB - Laboratório de Geoprocessamento e Estudos Ambientais

MRLC - Modelo de Regressão Linear Convencional

COD – Coeficiente de Dispersão

NTRIP - *Networked Transport of RTCM via Internet Protocol*

NIMA - *National Imagery and Mapping Agency*

NASA - *National Aeronautics and Space Administration*

PPP - Posicionamento por Ponto Preciso

PVG - Planta de Valores Genéricos

RTCM - *Radio Technical Commission for Maritime Services*

RAMT - Relatórios de Análise dos Mercados de Terras

RGC - Revisões Gerais de Cadastro

RTK - *Real-Time Kinematic*

SGB - Sistema Geodésico Brasileiro

SIGEF - Sistema de Gestão Fundiária

SRTM - *Shuttle Radar Topography Mission*

SUMÁRIO

<u>1 INTRODUÇÃO GERAL</u>	<u>15</u>
<u>2 OBJETIVOS.....</u>	<u>18</u>
2.1 OBJETIVO GERAL.....	18
2.2 OBJETIVOS ESPECÍFICOS	18
<u>3 ESTRUTURA DA DISSERTAÇÃO</u>	<u>19</u>
<u>4 A DIFICULDADE NA PRECIFICAÇÃO DE TRABALHOS DE GEORREFERENCIAMENTO DE IMÓVEIS RURAIS: UMA REVISÃO.....</u>	<u>21</u>
4.1 A IMPORTÂNCIA E A REGULAMENTAÇÃO DO GEORREFERENCIAMENTO, NO BRASIL	23
4.2 EVOLUÇÃO E HISTÓRICO DO GEORREFERENCIAMENTO	24
4.3 MÉTODOS DE PRECIFICAÇÃO ATUAIS.....	27
4.4 DESAFIOS E VARIÁVEIS QUE IMPACTAM A PRECIFICAÇÃO.....	28
4.4.1 CARACTERÍSTICAS DA PAISAGEM	28
4.4.2 INFRAESTRUTURA E LOGÍSTICA	29
4.4.3 TECNOLOGIA E EQUIPAMENTOS	29
4.4.4 BUROCRACIAS E PRAZOS	30
4.4.5 VARIABILIDADE REGIONAL	31
4.4.6 NÍVEL DE EXPERIÊNCIA PROFISSIONAL	31
4.5 REVISÃO DE TRABALHOS EXISTENTES DE SERVIÇOS DE TOPOGRAFIA	32
4.5.1 METODOLOGIAS DE ESTUDOS SOBRE PRECIFICAÇÃO	32
4.5.2 ANÁLISE COMPARATIVA DE METODOLOGIAS UTILIZANDO RPA	34
4.5.3 PRINCIPAIS LACUNAS E OPORTUNIDADES	35
4.6 TEMPO COMO PRINCIPAL FATOR DE PRECIFICAÇÃO	35
4.6.1 RELAÇÃO ENTRE TEMPO E CUSTO	35
4.6.2 TEMPO E COMPLEXIDADE DOS SERVIÇOS	37
4.6.3 IMPORTÂNCIA DE MODELOS PREDITIVOS	38

4.7 CONSIDERAÇÕES SOBRE O CAPÍTULO 4	40
4.8 CONCLUSÃO	40

5 COMPARAÇÃO ENTRE MODELOS PREDITIVOS PARA OTIMIZAÇÃO DA ESTIMATIVA DE TEMPO EM SERVIÇOS DE GEORREFERENCIAMENTO DE IMÓVEIS RURAIS..... 42

5.1 METODOLOGIA	44
5.1.1 PREPARAÇÃO DO CONJUNTO DE DADOS.....	46
5.1.2 DESCRIÇÃO DOS PROCESSOS ENVOLVIDOS NO LEVANTAMENTO DE CAMPO	50
5.1.3 DEFINIÇÃO DAS VARIÁVEIS	54
5.1.4 ANÁLISE DE CORRELAÇÃO ENTRE AS VARIÁVEIS	56
5.1.5 DEFINIÇÃO DA MODELAGEM	57
5.1.6 MÉTRICAS DE AVALIAÇÃO, VALIDAÇÃO DOS MODELOS E VISUALIZAÇÃO DOS RESULTADOS ..	60
5.1.7 TECNOLOGIA/BIBLIOTECA E LINGUAGEM UTILIZADA	61
5.2 RESULTADOS E DISCUSSÕES	62
5.2.1 DADOS E OBSERVAÇÕES	62
5.2.2 PREPARAÇÃO DAS VARIÁVEIS	63
5.2.3 ANÁLISE EXPLORATÓRIA	64
5.2.4 NORMALIZAÇÃO DAS OBSERVAÇÕES.....	65
5.2.5 CONJUNTO DE TREINO E TESTE	65
5.2.6 ANÁLISE RESUMIDA DAS MÉTRICAS.....	68
5.2.7 INFLUÊNCIA DAS VARIÁVEIS	69
5.2.8 VALIDAÇÃO EXTERNA DE UM CONJUNTO INDEPENDENTE	71
5.3 CONSIDERAÇÕES SOBRE O CAPÍTULO 5	73
5.4 CONCLUSÃO	74

6 CONCLUSÃO GERAL..... 77

7 REFERÊNCIAS..... 80

8 APÊNDICE – CÓDIGOS IMPLEMENTADOS..... 84

1 INTRODUÇÃO GERAL

A questão da precificação de serviços de Georreferenciamento de Imóveis Rurais (GIR), tem ganhado relevância significativa nas últimas décadas, em um cenário que a regularização fundiária e a gestão eficaz das áreas rurais são fundamentais para a promoção do desenvolvimento agrário e a garantia de direitos de propriedade, no Brasil (DOS SANTOS, 2022).

Segundo Lorenzi *et al.* (2007), empresas prestadoras de serviços têm problemas gerenciais semelhantes enfrentados pelas indústrias, em que procuram associar os custos dos recursos que fornecem receitas geradas pelos produtos e clientes.

Para precificação de um serviço deve-se estimar os custos fixos (independentes da realização do serviço), variáveis (são dependentes da análise, do decorrer e da realização completa do serviço) e a margem de lucro (remuneração pela execução) (HENRIQUE DUARTE *et al.*, 2022).

No site <https://www.amiranet.com.br/simulador> ([Simulador Referencial de Preços de Serviços de Agrimensura - Revista A Mira](#)) é possível ter uma referência de preços aproximados de vários tipos de serviços envolvendo topografia e levantamentos geodésicos georreferenciados, que são fundamentais para o GIR.

A ausência de metodologias padronizadas e a dependência de métodos heurísticos resultam na incerteza de orçamentos, gerando riscos financeiros para os profissionais e falta de transparência para os contratantes.

A principal fonte desta incerteza reside na dificuldade de estimar com precisão a variável que impacta no custo final: o tempo de execução em campo. Fatores como a distância a ser percorrida entre os vértices, as condições do terreno, o tipo de solo, a vegetação e o relevo interagem de forma complexa, tornando a estimativa baseada apenas na experiência profissional um processo altamente subjetivo e suscetível a erros significativos. Esta lacuna metodológica cria um cenário de vulnerabilidade para o setor e justifica a busca por soluções capazes de explicar essa relação do tempo com precificação além de mitigar erros orçamentários.

Segundo Belotto (2022), estimar custos na fase de viabilidade ou inicial de um projeto é reconhecidamente uma das etapas mais críticas e desafiadoras na engenharia, além disso elas devem apresentar um grau de precisão que que indiquem custos reais.

Maués (2020) revela a importância dos modelos de estimativa paramétricos, que, através de parâmetros importantes, buscam estabelecer relações estatísticas entre as características de um projeto e seus custos, com dados de projetos já realizados (MAUÉS *et al.* 2022).

Em resposta a essa problemática, a presente dissertação propõe, desenvolve e valida uma solução quantitativa baseada em modelagem preditiva e Aprendizado de Máquina. O objetivo central é substituir a subjetividade por um sistema com base em observações reais, capaz de estimar o tempo de deslocamento a partir das características mensuráveis de um projeto.

As observações/dados utilizadas na pesquisa, foram obtidas nos levantamentos topográficos georreferenciados de áreas no interior do Paraná, visitadas pelo autor, com o registro de algumas variáveis simultaneamente ao rastreo dos vértices. A manipulação e avaliação desses dados foram feitas em escritório.

O processo envolve identificar as variáveis independentes que influenciam no tempo do levantamento de cada vértice relevante, em particular, coletar dados relacionados a essas variáveis. Busca também realizar transformações necessárias, construir o modelo de predição, avaliar criticamente as variáveis e a adequação do modelo e verificar sua aplicabilidade para prever e auxiliar na estimativa do tempo necessário para os levantamentos, e por consequência, auxiliar na definição dos preços de serviços de GIR em novos casos.

Por fim, comparar a estimação do modelo proposto com levantamentos já realizados, para que possa ter uma maior confiabilidade nos resultados dessa modelagem, possibilitando a manutenção de preço e a conclusão dos serviços de GIR.

A hipótese da pesquisa é de encontrar as variáveis que influenciam o tempo de deslocamento entre vértices, nos serviços de GIR, e modelá-las, de forma que se possa estimar o tempo de deslocamento de trabalhos futuros.

Se for possível desenvolver um modelo preditivo para a estimativa de tempo em serviços de GIR, através do tempo mensurado em campo, identificar as variáveis e associá-las a esse tempo, além de determinar o melhor modelo de inferência, que se adeque a elas.

Então o tempo do levantamento, associado às variáveis identificadas como influenciadoras desse tempo poderão alimentar diferentes modelos de inferência e a

capacidade individual dessas variáveis em influenciar esse tempo do levantamento, poderá ser testada, conseqüentemente quantificada.

Além disso, a capacidade de explicação do modelo poderá ser mensurada e, com isso, uma estrutura de inferência para determinação desse tempo necessário para realização dos serviços de georreferenciamento poderá ser estabelecido

Além de auxiliar na precificação, procura também, aos que realizam esse tipo de serviço, subsídio para estipular os custos dos trabalhos de campo, bem como programação e planejamento com melhores precisões acerca de itinerários e tempo de materialização/levantamento dos vértices.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Este trabalho busca mitigar a subjetividade e o risco relacionados à precificação de serviços de GIR, com o objetivo central de desenvolver um modelo preditivo robusto para a estimativa do tempo de execução dos trabalhos de campo. Por fim, validar a abordagem de modelagem mais acurada como uma ferramenta de apoio à decisão, servindo de subsídio para a elaboração de orçamentos e planejamento logístico dos serviços.

2.2 OBJETIVOS ESPECÍFICOS

- Mensurar o tempo necessário para os levantamentos topográficos georreferenciados dos diferentes tipos de limites de imóveis rurais presentes na NTGIR, para que possa estimar o tempo de deslocamento entre vértices em levantamentos futuros.
- Estudar, definir e mensurar um conjunto de variáveis espaciais para representar as características dos terrenos, e como estas características impactam o tempo de levantamento dos trabalhos de georreferenciamento. Além de testar a capacidade e aplicabilidade dessas variáveis, em diferentes modelos e algoritmos de inferência.
- Apresentar um sistema para verificação do tempo necessário para execução dos levantamentos topográficos georreferenciados, visando subsidiar a tomada de decisão no que se refere a estimativa de recursos financeiros.

3 ESTRUTURA DA DISSERTAÇÃO

Este capítulo detalha a organização estrutural da presente dissertação, com o objetivo de orientar o leitor através da linha de raciocínio e da articulação entre os seus componentes. A estrutura adotada segue um formato híbrido, que combina os capítulos convencionais de uma dissertação com a apresentação de dois artigos científicos que aprofundam, respectivamente, a fundamentação do problema e o desenvolvimento da solução metodológica. A seguir, cada capítulo é descrito, destacando seu papel no desenvolvimento da pesquisa.

A dissertação inicia-se com uma estrutura convencional. O **Capítulo 1 (INTRODUÇÃO GERAL)**, apresenta o tema, contextualizando a problemática central da precificação de serviços, mais especificamente no âmbito do GIR e aborda a maneira que será tratado a modelagem dos dados e as hipóteses levantadas para resolução desta problemática. Neste capítulo também, é possível verificar a existência de um modelo de precificação para trabalhos envolvendo levantamentos topográficos.

O **Capítulo 2 (OBJETIVOS)**, estabelece o objetivo central da pesquisa: desenvolver e validar um modelo preditivo para a estimativa do tempo de execução em serviços de GIR, com finalidade de subsidiar uma precificação mais clara e objetiva. Os objetivos específicos definem a rota metodológica, que parte da definição de variáveis preditoras com base em observações de campo, passa pela comparação sistemática entre diferentes classes de modelos (como aprendizado de máquina e regressão espacial) e culmina na apresentação do modelo de melhor desempenho como uma ferramenta de apoio à decisão para a otimização de orçamentos.

O **Capítulo 4** é apresentado no formato de artigo, intitulado "**A Dificuldade Na Precificação De Trabalhos De Georreferenciamento De Imóveis Rurais: Uma Revisão**" (Silveira, Paiva e Goulart, 2025). Este capítulo constitui uma revisão crítica aprofundada da literatura e da prática de mercado, investigando a evolução do GIR no Brasil, analisando as metodologias de precificação convencionais e suas limitações. A análise identifica a subjetividade na estimativa do tempo de execução em campo como a principal fonte de incerteza e risco. O capítulo conclui estabelecendo a lacuna de conhecimento e justificando a necessidade de uma solução quantitativa, que é o foco do capítulo/artigo seguinte.

Respondendo diretamente ao desafio estabelecido, o **Capítulo 5** apresenta-se o segundo artigo, "**Comparação entre Modelos Preditivos para Otimização da**

Estimativa de Tempo em Serviços de Georreferenciamento de Imóveis Rurais".

Este capítulo detalha o desenvolvimento e a validação da solução quantitativa proposta. Nele, são descritos os materiais e métodos, desde a organização pre-campo, passando pela coleta e o tratamento dos dados de campo, a análise exploratória e, principalmente, a metodologia de análise comparativa entre as diferentes classes de modelos: regressão global (OLS), modelos espaciais (SAR, SEM e GWR) e algoritmos de Aprendizado de Máquina (XGBoost, Random Forest, etc.). É demonstrado o protocolo de otimização de hiper parâmetros e validação cruzada empregado para garantir a validação dos resultados.

O **Capítulo 5.2 (RESULTADOS E DISCUSSÕES)** consolida e discorre sobre os achados apresentados no segundo artigo. Esta seção apresenta tabelas comparativas, discute a capacidade preditiva dos modelos utilizados e interpreta os *insights* fornecidos pelos modelos espaciais, como a presença de autocorrelação e não-estacionariedade. A discussão conecta os resultados quantitativos com as implicações práticas para a otimização de orçamentos e planejamento no setor de GIR.

Finalmente, o **Capítulo 6 (CONCLUSÃO GERAL)** sumariza as principais contribuições da dissertação. Ele reafirma a validação dos modelos de aprendizado de máquina como a abordagem mais eficaz, discute as limitações do estudo e, crucialmente, aponta direções para trabalhos futuros, sugerindo a incorporação de novas variáveis e a exploração de modelos híbridos para avançar ainda mais na precisão e confiabilidade das estimativas.

4 A DIFICULDADE NA PRECIFICAÇÃO DE TRABALHOS DE GEORREFERENCIAMENTO DE IMÓVEIS RURAIS: UMA REVISÃO

O Georreferenciamento de Imóveis Rurais (GIR), criado com o objetivo de organização territorial, é fundamental para a regularização fundiária no Brasil, conforme apresentado pela Lei 10.267/2001, conhecida como a "Lei do Georreferenciamento de Imóveis Rurais" (Brasil 2002). Esta legislação estabeleceu a necessidade do georreferenciamento como requisito para o registro de imóveis rurais no Cadastro Nacional de Imóveis Rurais (CNIR). O principal objetivo desta regulamentação é garantir a precisão na definição dos limites das propriedades rurais, prevenindo sobreposições e proporcionando maior segurança jurídica tanto para os proprietários quanto para o Estado.

De acordo com Toledo e Bertotti (2014), o Instituto Nacional de Colonização e Reforma Agrária (INCRA) estabeleceu novas regras para a regularização fundiária visando unificar e garantir a precisão do CNIR. A base para essas diretrizes é a 3ª Norma Técnica para Georreferenciamento (3ª NTGIR), que define os padrões técnicos e o novo conceito de 'imóvel rural' a serem seguidos. Para detalhar os procedimentos, a norma é complementada por dois manuais técnicos: um focado no posicionamento geográfico e o outro na definição dos limites e confrontações da propriedade.

A ferramenta que opera todo este sistema é o Sistema de Gestão Fundiária (SIGEF), uma plataforma online que automatiza a validação de dados e a certificação de imóveis de forma ágil e eficiente (Toledo & Bertotti 2014).

Segundo Santos (2022), a colaboração entre o INCRA e o Cartório de Registro de Imóveis (CRI) é determinante, pois envolve a supervisão de um profissional qualificado e treinado para realizar a medição dos imóveis rurais.

De acordo com Duarte *et al.* (2022), o GIR é vital para complementar a segurança jurídica e a gestão eficaz das áreas rurais, contribuindo diretamente para o progresso agrário e territorial do país. Contudo, a avaliação dos custos desses serviços ainda apresenta lacunas devido à falta de critérios uniformes e à complexidade das variáveis envolvidas no processo, que, por sua vez, influenciam os preços ao longo de todo o processo (Duarte *et al.* 2022).

Além disso, Felipetto *et al.* (2023) destacam que, na perspectiva da Agrimensura e Cartografia, o Georreferenciamento é o processo de determinar as

posições de pontos na superfície terrestre com base em suas coordenadas tridimensionais (X, Y, Z) ou através de coordenadas geodésicas, como latitude, longitude e altitude, dentro de um Sistema de Referência específico, neste caso, o SIRGAS 2000, utilizado no Brasil. Com base neste princípio, é viável estabelecer a forma, o tamanho e a localização de áreas utilizando diversos métodos de levantamentos geodésicos e topográficos.

Segundo Borges Júnior *et al.* (2017), o Sistema Geodésico Local (SGL) é um sistema de referência único para cada imóvel, definido por um plano cartesiano local com origem no centroide do polígono da propriedade. Este plano é tangente à superfície da Terra em sua origem, uma característica que minimiza as distorções de área comumente encontradas em projeções cartográficas mais extensas. O processo de cálculo de área é automatizado pelo SIGEF; o sistema primeiro converte as coordenadas geodésicas dos vértices do imóvel para o SGL e, subsequentemente, aplica a Fórmula de Gauss para determinar a área final.

Os obstáculos encontrados na avaliação da precificação incluem as características do terreno, a logística necessária para alcançar os pontos de medição e os gastos com equipamentos e tecnologias apropriadas (Ok 2025), bem como os prazos e exigências, definidos pelo INCRA. A falta de previsibilidade nos custos afeta tanto os profissionais responsáveis pela execução dos serviços quanto os clientes, podendo também impactar os processos licitatórios.

Assim, este artigo tem como objetivo conduzir uma revisão de literatura sobre a precificação dos serviços de GIR, destacando a importância da previsão do tempo de execução como critério fundamental na definição de custos. Subsequentemente, serão discutidos os aspectos legais que sustentam a técnica do GIR, juntamente com um breve histórico do GIR no país e os métodos de precificação existentes e seus desafios.

Algumas dessas variáveis constituem verdadeiros obstáculos para quem executa o GIR, consequentemente levando a variações na precificação do serviço. Alguns obstáculos comuns encontrados no GIR incluem: características da paisagem (encostas íngremes, áreas úmidas, vegetação densa, para citar alguns), a logística necessária para alcançar os pontos de medição, os gastos com equipamentos e tecnologias apropriadas, bem como os prazos e requisitos estabelecidos pelo INCRA. A falta de previsibilidade dos custos afeta tanto os profissionais responsáveis pela

execução dos serviços quanto os clientes, o que também pode impactar a precificação final dos processos licitatórios.

4.1 A IMPORTÂNCIA E A REGULAMENTAÇÃO DO GEORREFERENCIAMENTO, NO BRASIL

Segundo Figueiredo Filho, Disperati & Watzlawick (2003), a configuração da propriedade rural no Brasil passou por mudanças substanciais com o crescimento do agronegócio, levando a uma maior demanda por uma gestão eficaz das áreas rurais. A Lei nº 10.267/2001 (Brasil, 2002) introduziu o georreferenciamento e a certificação de imóveis rurais, criando o CNIR, que, de acordo com Fontes (2018), é uma base de dados unificada do INCRA e da Secretaria da Receita Federal, permitindo a identificação e o compartilhamento de informações sobre propriedades rurais. O INCRA é a entidade responsável por validar essas certificações, garantindo que os limites das propriedades cumpram a legislação vigente e evitando a sobreposição desses imóveis. O georreferenciamento não apenas proporciona uma visualização clara dos limites da propriedade, mas também permite a prevenção de conflitos de terras, sobreposições e disputas territoriais, especialmente em áreas com alta concentração de propriedades rurais (Fontes 2018; Felipetto *et al.* 2023).

De acordo com Fontes (2018), o processo de transcrição de dados georreferenciados para as matrículas dos imóveis serve como subsídio para a atualização de bases fundiárias e para a eficácia de políticas públicas voltadas à distribuição e ao acesso à terra.

Assim, segundo Fontes (2018), quando realizado de forma adequada, o GIR pode garantir a legalidade dos imóveis rurais e ter relevância econômica e social, servindo como subsídio para a regularização fundiária e a gestão territorial. A regularização fundiária, por sua vez, facilita o acesso a financiamentos rurais, promove o crescimento econômico local e ajuda a minimizar conflitos relacionados à posse da terra. No entanto, para que o processo seja eficaz, é vital que esses serviços sejam executados com a precisão técnica adequada, envolvendo profissionais de boa reputação, evitando sobreposições de áreas e garantindo que os custos sejam consistentes com a realidade de cada serviço com base na região em que são realizados.

Além da "Lei do Georreferenciamento de Imóveis Rurais", vale ressaltar que outras leis, decretos e regulamentos complementam e destacam o GIR, como a Lei

nº 6.015 (Brasil 1973), que dispõe sobre os registros públicos, e a Norma Técnica NBR 13.133 (ABNT 1994), para citar algumas (Santos 2022).

O georreferenciamento vai além de um procedimento técnico normativo; ele é a base fundamental para o Cadastro Rural. Em diversos países, como Alemanha, Holanda e Suíça, essa base cadastral é utilizada para manter os mapeamentos de referência atualizados, fornecendo suporte para projetos de infraestrutura pública. No Brasil, essa prática se torna ainda mais relevante, considerando que apenas 5,5% do território nacional possui mapeamento na escala 1:25.000 (Sluter *et al.* 2018). Além disso, há uma divergência nesse processo de mapeamento, pois o mapeamento urbano é realizado pelo IBGE e pela DSG, enquanto o mapeamento rural é de responsabilidade do INCRA e do CNIR.

O georreferenciamento desempenha um papel crucial na publicidade e garantia do direito de propriedade, facilitando a identificação de sobreposições de áreas e, assim, prevenindo conflitos de terras, ocupações irregulares e protegendo áreas de conservação.

4.2 EVOLUÇÃO E HISTÓRICO DO GEORREFERENCIAMENTO

Os levantamentos topográficos têm diversas finalidades, como a verificação de áreas para compra e venda, a abertura de matrículas, as divisões amigáveis, a reabertura de linhas de divisa e os levantamentos de uso da terra, entre outros (Figueiredo Filho, Disperati & Watzlawick 2003).

Com o avanço das geotecnologias, o Sistema de Posicionamento Global (GPS) tornou-se uma ferramenta essencial nos levantamentos topográficos, complementando ou até substituindo métodos tradicionais que utilizam teodolitos ou estações totais. Essa tecnologia facilita a aquisição de coordenadas geográficas e a criação de mapas digitais e georreferenciados, proporcionando maior eficiência no trabalho de campo. No entanto, seu desempenho pode ser afetado por fatores ambientais, como a presença de cobertura vegetal densa, que interfere na precisão do sinal. A integração de receptores GPS Diferencial (DGPS) com estações totais melhorou significativamente a confiabilidade dos levantamentos, garantindo maior precisão na gestão de terras e contribuindo para a segurança e a exatidão das informações geoespaciais (Figueiredo Filho, Disperati & Watzlawick 2003).

Segundo Santos (2022), até 2001, como as matrículas e/ou transcrições não estabeleciam precisão e eram insuficientes para determinar a localização dos vértices

dos imóveis, as transações imobiliárias eram registradas apenas em Cartórios de Notas, com a possibilidade de retificação judicial de inconsistências até 2003. De acordo com o referido autor, os proprietários eram responsáveis por atualizar o cadastro no INCRA, enquanto o agrimensor era responsável pela qualidade posicional e pela incerteza dos vértices medidos.

Dessa forma, o INCRA não conseguia produzir um mosaico do conjunto de propriedades devido à falta de padronização dos levantamentos e dados existentes (Santos 2022).

De acordo com Toledo e Bertotti (2014), a promulgação da Lei 10.267/2001 (Brasil, 2002) representou um marco, pois estabeleceu normas para o georreferenciamento e instituiu a necessidade de certificação pelo INCRA, instituindo o georreferenciamento de imóveis rurais e a integração de dados fundiários com outros órgãos federais. Este processo passou por transformações significativas ao longo dos anos, utilizando novas tecnologias da informação que tornaram o Georreferenciamento de Imóveis Rurais (GIR) mais automatizado e eficiente. A implementação da 3ª NTGIR (INCRA 2013) e do SIGEF representa um marco importante neste cenário, facilitando o mapeamento do território rural e aumentando a responsabilidade técnica dos profissionais envolvidos, que devem atualizar constantemente seus conhecimentos para garantir a precisão do levantamento e a qualidade do processo de certificação.

Segundo Ballar (2021), até 2013, este procedimento era feito manualmente e consumia um tempo considerável; após essa data, com a criação do SIGEF, o processo foi digitalizado, melhorando a eficiência e reduzindo o prazo de certificação. Com a combinação do INCRA e do CNIR através do SIGEF, mais de 40 milhões de hectares foram certificados nos primeiros dez meses de operação da ferramenta.

O SIGEF é uma ferramenta eletrônica desenvolvida pelo INCRA para apoiar a governança fundiária no território brasileiro. Através dele, são realizadas a recepção, validação, organização, regularização e disponibilização de informações georreferenciadas dos limites de imóveis rurais, tanto públicos quanto privados. Por meio do SIGEF, são realizadas a certificação de dados referentes aos limites dos imóveis rurais e a gestão de contratos de serviços de georreferenciamento com a administração pública (INCRA 2022).

Devido à gestão pelo SIGEF e à regulamentação do GIR pelo INCRA, é viável associar a relevância do tempo na definição dos vértices para a certificação de

imóveis e sua influência nas avaliações de serviços. A digitalização proporcionada pelo SIGEF não apenas aumentou a eficácia da certificação, reduzindo significativamente o tempo necessário e melhorando a precisão, mas também estabeleceu padrões técnicos através dos manuais do INCRA (Ballar 2021).

Toledo e Bertotti (2014) comentaram que a 3ª NTGIR representou um avanço não apenas na automação, mas também na especificação dos procedimentos de campo, observando que a norma foi complementada por manuais técnicos específicos, como o Manual Técnico de Posicionamento e o Manual Técnico de Limites e Confrontações, que detalham como os limites das propriedades devem ser caracterizados. Além disso, foram estabelecidos critérios rigorosos de precisão que os profissionais devem atender.

Conforme detalhado por Borges Júnior *et al.* (2017), a precisão posicional exigida para os vértices varia de acordo com seu tipo e localização: 50 cm para limites artificiais, 3,0 m para limites naturais e 7,5 m para pontos em locais de difícil acesso ou inacessíveis (INCRA 2013).

Essa padronização técnica é fundamental para a qualidade dos dados inseridos no SIGEF e para a segurança jurídica e a confiabilidade do cadastro fundiário nacional, garantindo que não ocorram sobreposições de propriedades (Toledo e Bertotti 2014; Borges Júnior *et al.* 2017).

A atualização do Manual Técnico para o Georreferenciamento de Imóveis Rurais (MTGIR), segundo Felipetto *et al.* (2023), visou principalmente integrar novas tecnologias e procedimentos, aprimorando as diretrizes da primeira edição do manual. Nesse contexto, foram incorporadas diversas técnicas que podem ser aplicadas em situações específicas, como o posicionamento por sensoriamento remoto via fotogrametria aérea, utilizando aeronaves remotamente pilotadas, como as Aeronaves Remotamente Pilotadas (RPAs), bem como radar aerotransportado, varredura a laser aerotransportada e sensores orbitais em satélites.

Para cada método de posicionamento adotado, a acurácia posicional absoluta das coordenadas dos vértices que delimitam o imóvel rural deve ser avaliada. Para essa avaliação, podem ser utilizados ortomosaicos e ortofotocartas, além de combinações com Modelos Digitais de Terreno (MDT), Modelos Digitais de Superfície (MDS) ou curvas de nível (Felipetto *et al.* 2023).

4.3 MÉTODOS DE PRECIFICAÇÃO ATUAIS

De acordo com Bomfim e Passareli (2011), é necessário compreender a estrutura de custos, mesmo que esta já seja definida pelo mercado, pois só é possível expandir ou melhorar as margens de lucro se os custos forem reduzidos. A falta de conhecimento sobre métodos de custeio e a ausência de um planejamento estruturado dificultam o alcance de margens de lucro sustentáveis. Isso torna importante a educação gerencial e o desenvolvimento de estratégias orientadas a dados para uma precificação eficiente (Purificação *et al.* 2022).

A definição de preços é um fator essencial para qualquer empresa ou profissional autônomo, garantindo o lucro na prestação de serviços. Ela também demonstra, através de indicadores financeiros, a saúde da empresa (Pissinati & Dias, 2022). Embora uma estrutura de custos seja de fato muito importante para a lucratividade, o processo deve ser bem-estruturado com o pessoal envolvido. Isso significa que os funcionários de uma empresa devem aplicar e reforçar, por mais tedioso e exaustivo que seja, o registro de tudo o que entra e sai, proporcionando assim o conhecimento do lucro, como mencionado por Bomfim e Passareli (2011), e a formação de um preço justo, como observado por Pissinati e Dias (2022).

Buscando alternativas de precificação que envolvam mais de uma variável, Malaman e Amorim (2017) indicaram que a combinação de abordagens tradicionais e inovadoras resulta em melhores estimativas, e os métodos de precificação utilizados para a avaliação de imóveis foram a Lógica Fuzzy e o método tradicional de Regressão Linear presente nas normas da ABNT. Verificou-se que o segundo método, a Regressão Linear, apresentou resultados superiores para a amostra utilizada. Segundo Duarte *et al.* (2022), para buscar um preço justo por um serviço, deve-se estimar os custos fixos (independentes da execução do serviço), os custos variáveis (dependentes da análise, do andamento e da execução completa do serviço) e a margem de lucro (remuneração pela execução do serviço). Dito isso, é benéfico para qualquer empresa, independentemente do porte, manter o conhecimento e o equilíbrio desses fatores.

A precificação de produtos e serviços é um fator extremamente importante para as empresas, impactando diretamente a lucratividade e a competitividade no mercado. Adicionalmente, Araújo *et al.* (2023) demonstram que, embora a questão da precificação seja antiga, a academia produz estudos relevantes que auxiliam na formulação de estratégias para a definição de preços mais eficientes, tanto para

produtos quanto para serviços. Algoritmos e métodos estatísticos são eficazes para apoiar decisões de precificação, tornando o processo mais preciso e adaptável às mudanças do mercado, que são mais complexas devido à variabilidade dos custos e à influência do ambiente digital no caso de serviços (Araújo *et al.* 2023).

Assim, como destacado por Malaman e Amorim (2017), Duarte *et al.* (2022) e Araújo *et al.* (2023), enfatiza-se que, embora não exista um melhor método ou um método único que justifique ou defina a melhor precificação possível, é saudável para uma empresa buscar um método existente ou criar um que a aproxime de um preço mais competitivo no mercado, tanto para produtos quanto para serviços.

4.4 DESAFIOS E VARIÁVEIS QUE IMPACTAM A PRECIFICAÇÃO

Conforme discutido até o momento, a definição de preços para os serviços de GIR é relativa, devido às variáveis que tendem a influenciar o custo e, consequentemente, a duração do serviço. Esses fatores abrangem desde o estudo das características físicas das propriedades até as regiões onde estão registradas, mudando consideravelmente de estado para estado.

4.4.1 Características da Paisagem

A qualidade dos levantamentos georreferenciados está diretamente relacionada a uma série de variáveis, incluindo o tipo de equipamento utilizado, a densidade da cobertura vegetal, as condições atmosféricas e a precisão desejada. Em áreas abertas, o GPS pode substituir eficientemente a topografia convencional, mas em regiões de floresta, por exemplo, os sinais podem ser degradados, comprometendo a precisão das medições (Figueiredo Filho, Disperati & Watzlawick 2003).

Segundo Ok (2025), as características da paisagem são um fator determinante no levantamento topográfico, pois o terreno rege a seleção do método mais apropriado. Por exemplo, enquanto as RPAs (Aeronaves Remotamente Pilotadas) são particularmente eficazes para o levantamento de terrenos acidentados, sua utilidade diminui em áreas com vegetação densa. Nessas condições, os métodos de levantamento tradicionais são aconselhados devido à sua maior confiabilidade e precisão (Ok 2025).

As características do terreno podem ser uma das variáveis mais influentes na formulação de preços. Aspectos como a topografia (comparando áreas planas e

regiões montanhosas), a presença e a densidade da vegetação e o acesso aos pontos de interesse devem ser estudados se influenciarem diretamente o tempo necessário para a realização do serviço.

Essas características, como relevo irregular e/ou vegetação densa, levam a equipe técnica a percorrer longas distâncias em condições desafiadoras, dificultando a funcionalidade dos equipamentos e aumentando o tempo de execução e, consequentemente, os gastos financeiros.

4.4.2 Infraestrutura e Logística

A malha viária da área, o escopo do trabalho e a frequência das medições também influenciam a escolha dos métodos de levantamento (Figueiredo Filho, Disperati & Watzlawick 2003). A infraestrutura da região onde a propriedade está localizada é também um dos fatores que influenciam a precificação. A falta de estradas adequadas ou a necessidade de transportar equipamentos para locais remotos impacta os custos de transporte e acomodação da equipe.

A logística envolvida no deslocamento de equipamentos, como receptores do Sistema Global de Navegação por Satélite (GNSS), drones, estações totais, entre outros (Oliveira *et al.* 2024), consome uma porção considerável do orçamento (tanto em termos de aquisição quanto de aluguel de equipamentos).

Dito isso, é possível avaliar o impacto que uma malha viária de alta qualidade pode ter, não apenas na segurança dos envolvidos, mas também na estimativa do período de serviço, afetando assim a precificação.

4.4.3 Tecnologia e Equipamentos

A evolução da cartografia trouxe equipamentos mais sofisticados, como estações totais, aparelhos de GPS e *softwares* de pós-processamento, que permitem maior controle de erros e precisão nos resultados. Embora a decisão sobre quais ferramentas utilizar deva considerar um equilíbrio entre a acurácia e os custos operacionais, bem como o bom senso e a experiência do profissional (Figueiredo Filho, Disperati & Watzlawick 2003).

Enfatizando a importância do GNSS RTK (*Real Time Kinematic*), Rohr (2023) mostra que esta é a metodologia mais confiável para o georreferenciamento, pois envolve a determinação de dados altimétricos em projetos que exigem precisão e

encontram obstruções constantes, como em áreas rurais onde há vegetação densa, por exemplo.

A seleção de uma metodologia de levantamento topográfico é uma decisão crítica. Conforme comparado por Ok (2025), métodos convencionais, como a Estação Total e o GNSS, oferecem precisão superior e custos iniciais de equipamento reduzidos. Essas técnicas são adequadas para projetos que requerem precisão subcentimétrica em áreas localizadas e acessíveis. Em contrapartida, tecnologias como as RPAs (Aeronaves Remotamente Pilotadas) exigem um investimento inicial maior tanto em equipamentos quanto em *software*. No entanto, elas se mostram mais econômicas para projetos de maior escala e em terrenos de difícil acesso, ao reduzir significativamente o tempo em campo (Ok 2025). Portanto, os custos de aquisição e manutenção são despesas que devem ser consideradas na formação de preços. Além disso, a escolha da tecnologia deve também levar em conta a complexidade e os requisitos específicos do projeto.

O progresso tecnológico trouxe novas opções para o georreferenciamento, se permitidas pelas regulamentações e se alcançarem a precisão necessária, como drones e *softwares* de processamento de imagens, que oferecem maior eficácia na coleta de dados, impactando a precificação.

4.4.4 Burocracias e Prazos

A complexidade do processo se deve às regulamentações técnicas do INCRA, que incluem o uso do SIGEF para a certificação. A apresentação de documentos e a frequente necessidade de ajustes para atender às exigências regulatórias podem estender os prazos e incorrer em despesas inesperadas para todas as partes envolvidas. Além disso, a adesão aos prazos exigidos por contratantes ou instituições reguladoras demanda um planejamento meticuloso, especialmente em projetos maiores ou mais complexos.

Neste caso, embora possa parecer se qualificar como um custo fixo por ter um prazo específico ou um possível retorno para reapresentação de documentos, é considerado um custo variável, uma vez que, em alguns casos, deve-se contabilizar uma extrapolação dessa data/prazo ou a devolução final dos documentos.

4.4.5 Variabilidade Regional

O território brasileiro contribui para influenciar as variações de preço entre as diferentes regiões. Aspectos como clima, densidade populacional, custo de vida e disponibilidade de mão de obra especializada podem tornar certos serviços consideravelmente mais caros em algumas localidades. Segundo Oliveira e Piccinini (2011), na teoria constitucional do mercado de trabalho, a formação e a remuneração não seguem a lógica da competência ou da dinâmica de mercado. Empresas de médio e grande porte tendem a definir regras, como a preparação de funcionários para cargos específicos, a forma de remunerá-los, suas funções e promoções. As profissões e os ofícios são guiados por regras institucionais em vez de pela oferta e demanda.

De acordo com Veloso (1992), o estado do Paraná é composto por dois biomas distintos: a floresta tropical pluvial (Mata Atlântica) e a savana (Cerrado). O primeiro bioma é formado por diferentes contextos físicos, tanto nos aspectos físicos quanto nos econômicos, sociais e culturais. Os climas alternam entre seco e chuvoso em diferentes épocas do ano, assim como a alta variação de altitudes em todo o estado (Santos 2022).

Estes fatores, combinados com a falta de literatura científica que aborde esses desafios, evidenciam a complexidade do processo de definição de preços, onde a diversificação de custos torna a padronização de valores desafiadora e aumenta a demanda por modelos preditivos que levem em conta essas disparidades regionais. Portanto, o preço de um serviço de GIR na região metropolitana de Curitiba/PR será diferente daquele em Sorriso/MT, mesmo que o tamanho da propriedade e o número de marcos sejam idênticos, por exemplo.

Essas dificuldades ilustram quão complicado é o processo de determinação de preços e enfatizam ainda mais a importância de métodos que integrem fatores técnicos, econômicos, logísticos e, principalmente, físicos. A aplicação de modelos que tratam o tempo como um elemento fundamental pode proporcionar uma melhor previsão e precisão na avaliação de despesas, garantindo competitividade e clareza no setor de georreferenciamento.

4.4.6 Nível de Experiência Profissional

De acordo com Silva *et al.* (2017), pelo fato de uma grande parte dos profissionais pesquisados desconhecer o procedimento correto de assepsia, o

resultado para os pacientes poderia ter sido muito mais favorável. Este fator, entre outros, levou os gestores a planejarem medidas de melhoria, como cursos de capacitação e educação em saúde.

A relação entre a experiência profissional e a autopercepção de competência está ligada ao desenvolvimento de habilidades no ambiente de trabalho, pois a experiência prática não apenas enriquece o conhecimento técnico, mas também aumenta a confiança e a adaptabilidade do profissional diante dos desafios. Segundo Batista, Matos e Graça (2011), profissionais mais experientes tendem a refletir criticamente sobre suas vivências, o que lhes permite integrar novos aprendizados à sua prática diária. A competência deve, portanto, ser compreendida de forma multidimensional, abrangendo não apenas o conhecimento formal, mas também os aprendizados informais que surgem das interações e práticas cotidianas.

Assim, é possível perceber que o nível de experiência, tanto profissional, como abordado por Silva *et al.* (2017), quanto as vivências de vida, como destacado por Batista, Matos e Graça (2011), podem influenciar positivamente e prever resultados.

4.5 REVISÃO DE TRABALHOS EXISTENTES DE SERVIÇOS DE TOPOGRAFIA

A pesquisa sobre avaliação de preços para serviços de georreferenciamento no Brasil é bastante limitada, especialmente no que tange a modelos padronizados e preditivos que levem em conta as particularidades regionais e técnicas. Em muitos casos, a precificação é realizada de forma empírica. No entanto, algumas investigações apresentam métodos e abordagens que auxiliam na compreensão das variáveis envolvidas e no desenvolvimento de estratégias mais eficientes.

4.5.1 Metodologias de Estudos sobre Precificação

Gazola (2002) sugere a implementação de modelos de regressão para analisar custos, vinculando as variáveis envolvidas na metodologia, desde a coleta de dados e a análise exploratória até, por fim, o ajuste do modelo e sua aplicabilidade a novos dados. Ao definir uma metodologia, é possível antecipar custos com base em variáveis controláveis, como logística e características físicas do terreno, permitindo uma previsão mais precisa da execução do serviço.

O estudo de Lorenzi e Rosa (2007) aborda o controle de custos em empresas prestadoras de serviços, enfatizando a complexidade da precificação devido às diferentes naturezas dos serviços, sejam eles repetitivos ou específicos. Destaca a importância da alocação adequada de custos diretos e indiretos, considerando fatores como mão de obra, materiais e estrutura organizacional, entre outros.

A proposta de Lorenzi e Rosa (2007) foi a implementação do método de Custeio Baseado em Atividades (ABC) em uma empresa de topografia, visando atribuir com precisão os custos a cada atividade executada. O levantamento detalhado dos custos diretos e indiretos, a identificação dos direcionadores de custo e a análise da relação entre as atividades e a geração de despesas permitiram maior transparência e controle financeiro.

A prestação de serviços de topografia exige um alto nível de precisão e qualidade para garantir a eficiência dos levantamentos realizados. No entanto, empresas que atuam neste setor enfrentam desafios relacionados ao retrabalho, que impactam diretamente a produtividade e a precificação. Cunha, Nascimento e Parreira (2014) destacam que o retrabalho em pequenas empresas de topografia pode ser causado por problemas de comunicação interna, falhas no entendimento do escopo e falta de reuniões organizacionais. Esses fatores geram desperdício de tempo e recursos, afetando a lucratividade do projeto e a competitividade da empresa no mercado.

O retrabalho não apenas compromete a entrega dos serviços no prazo, mas também eleva os custos operacionais, pois exige a repetição de etapas já executadas. De acordo com a análise de Pareto realizada no estudo de Cunha, Nascimento e Parreira (2014), 90,48% das causas de retrabalho estão relacionadas à falta de informação entre os setores da empresa, falhas de comunicação com o cliente e erros na documentação técnica.

Esses problemas impactam diretamente a precificação dos serviços, pois a necessidade de refazer algo compromete as margens de lucro e dificulta a previsão orçamentária. A aplicação de metodologias de gestão, como o ciclo *Plan-Do-Check-Act* (PDCA), é sugerida como estratégia para minimizar o retrabalho e aprimorar a organização dos processos internos. Essa abordagem ajuda a padronizar atividades, mitigar falhas e garantir a conformidade dos serviços de topografia. Além disso, a implementação de ferramentas de qualidade, como *checklists* e formulários de controle, mostra-se eficaz para reduzir inconsistências operacionais sem a

necessidade de grandes investimentos financeiros. Assim, a adoção dessas práticas pode contribuir significativamente para a correta precificação dos serviços de georreferenciamento, garantindo maior previsibilidade de custos e competitividade no setor (Cunha, Nascimento & Parreira 2014).

A implementação de um método de custeio, como feito por Lorenzi e Rosa (2007), pode ser eficaz na determinação dos custos reais dos serviços e auxiliar em tomadas de decisão específicas. A diferenciação entre os tipos de serviços influencia a formulação de preços e a categorização de custos, tornando a escolha de um sistema de custeio eficiente altamente relevante. De forma similar, como conduzido por Cunha, Nascimento e Parreira (2014), uma metodologia foi aplicada para avaliar elementos que estavam causando distorções financeiras em relação ao que foi planejado.

4.5.2 Análise Comparativa De Metodologias Utilizando RPA

No contexto global, o uso de tecnologias como RPA (Aeronaves Remotamente Pilotadas) e imagens de satélite é amplamente adotado para reduzir despesas em levantamentos topográficos. Barros e Silva (2019) examinam a relação custo-benefício do uso de RPAs no georreferenciamento de imóveis rurais no Brasil, mostrando que, embora o investimento inicial seja considerável, a redução no tempo de execução e na necessidade de mão de obra justifica os custos ao longo do tempo, reduzindo o retrabalho e proporcionando serviços com maior qualidade e menores custos.

Outra metodologia seguida por Chicati *et al.* (2019) utiliza imagens obtidas através de RPA (recobrimento fotogramétrico) para o levantamento de limites naturais em áreas de difícil acesso, comparando-as com dados dos mesmos vértices levantados utilizando posicionamento GNSS/RTK. Eles constataram que os valores obtidos com a ortofoto estão dentro da precisão exigida pela NTGIR para limites naturais ou inacessíveis.

Por outro lado, Ferreira (2019) enfatiza o uso de RPA em vez de Estações Totais ao comparar a diferença de tempo na obtenção de pontos em uma superfície desejada. Devido a essa comparação, foi estabelecido que o uso de imagens ou nuvens de pontos é mais vantajoso para este tipo de levantamento.

A comparação entre métodos tradicionais e contemporâneos destaca a importância de modelos que ajustam a precificação às condições locais, o que é especialmente significativo em nações de grande dimensão como o Brasil.

4.5.3 Principais Lacunas e Oportunidades

Embora os estudos disponíveis forneçam contribuições relevantes, existem lacunas notáveis na literatura, particularmente no que tange à integração de variáveis técnicas, econômicas e logísticas em modelos de precificação. Adicionalmente, poucos estudos investigam o impacto de aspectos como sazonalidade, condições climáticas e infraestrutura local na formação de custos.

De acordo com Fontes (2018), a atual certificação de imóveis rurais alterou a forma como profissionais habilitados conduzem essas certificações, gerando diferentes interpretações e aumentando a necessidade de pesquisas sobre a correta aplicação da legislação.

Uma oportunidade promissora reside no desenvolvimento de modelos preditivos baseados em estatística espacial, conforme indicado por Pimenta (2021). Esses modelos podem incorporar fatores como as características do terreno, a distância entre os vértices e o tempo de execução estimado, proporcionando maior precisão na determinação de preços e, conseqüentemente, auxiliando a competitividade de profissionais e empresas no setor de georreferenciamento.

4.6 TEMPO COMO PRINCIPAL FATOR DE PRECIFICAÇÃO

A duração de um projeto de georreferenciamento é um aspecto influente no cálculo de seus custos. Estimar o tempo necessário para cada fase do trabalho ajuda os profissionais a estabelecerem preços mais precisos, minimizando incertezas e melhorando a alocação de recursos humanos e tecnológicos.

4.6.1 Relação Entre Tempo e Custo

De acordo com Alves, Varotto e Gonçalves (2012), uma abordagem quantitativo-descritiva foi utilizada para investigar os preços e métodos de precificação usados por varejistas em São Paulo. A pesquisa foi conduzida através de um levantamento com uma amostra não-probabilística e por conveniência, considerando estabelecimentos na região em 2011, envolvendo 35 indivíduos.

Os métodos de precificação foram classificados em três categorias: métodos baseados nos custos do fornecedor, métodos baseados na expectativa do cliente e métodos baseados na concorrência do mercado. A pesquisa constatou uma predominância no uso de métodos de custo tradicionais em detrimento de abordagens orientadas ao cliente (Alves, Varotto & Gonçalves 2012).

Dito isso, é possível buscar uma metodologia de precificação para serviços e topografia que vise à maximização dos lucros, à competitividade de mercado e à fidelização de clientes. A análise dos métodos de precificação deve incluir abordagens como a análise de custos (considerando equipamentos e mão de obra), a precificação competitiva (avaliando as taxas dos concorrentes) e a precificação baseada em valor (considerando o benefício percebido pelos clientes). Segundo Souza e Peixe (2019), destaca-se a necessidade de um mapeamento sistemático dos custos associados à prestação de serviços, que sirva como base para a construção de uma estrutura de precificação que reflita a realidade econômica das operações.

A implementação de uma metodologia de precificação eficiente no setor de topografia pode se beneficiar significativamente da análise de dados reais de projetos anteriores, como o que foi realizado em pesquisas do setor hoteleiro. A coleta e análise meticulosa de custos diretos, como materiais e mão de obra, e de custos indiretos, que englobam despesas administrativas e fiscais, são fundamentais para garantir uma avaliação abrangente dos gastos (Souza & Peixe 2019). Assim, a adoção de um método descritivo que combine análise teórica e estudos de caso empíricos permite uma avaliação detalhada das variáveis que impactam tanto o custo quanto o tempo de execução dos serviços de topografia, promovendo uma compreensão mais precisa da dinâmica de mercado e da precificação.

Além disso, a incorporação de uma estrutura de margem de lucro adequada através do método markup é essencial para assegurar que os preços cobrados não apenas cubram os custos operacionais, mas também sustentem a viabilidade financeira do negócio. A gestão eficaz de custos e despesas torna-se uma ferramenta estratégica, com a recomendação de implementar sistemas de controle e planejamento que facilitem a comparação entre os custos reais e os projetados (Souza & Peixe 2019). Esta abordagem integrada não apenas otimiza a definição de preços, mas também contribui para a competitividade e sustentabilidade econômica do setor de topografia, de forma similar ao que é observado na hotelaria.

Conforme sugerido por Tinoco e Silva (2019), a metodologia pode ser definida através da análise de Notas Fiscais de Consumidor Eletrônicas (NFC-e) como fontes de dados. Sugere-se também a identificação de categorias e preços médios para cada categoria, um modelo que considere variáveis como a demanda pelo serviço, custos associados (transporte, por exemplo) e a concorrência. Assim, é possível vincular com as despesas de topografia, controlando as notas fiscais de despesas, bem como categorizando cada tipo de serviço por área e local executado e o preço médio a ser pago/recebido por eles. A validação da metodologia e do modelo é importante para verificar sua realidade com a aceitação do mercado e com o que foi efetivamente gasto, consequentemente ajustando os valores para que sejam tanto competitivos quanto justos (Tinoco & Silva 2019). De forma análoga, este modelo de avaliação pode ser replicado para os serviços de topografia e GIR, onde os valores finais do modelo podem ser ajustados à realidade do mercado, tornando-os mais competitivos.

4.6.2 Tempo e Complexidade dos Serviços

De acordo com Fontes (2018), devido à hierarquia entre as regulamentações e à necessidade de conduzir registros e averbações de forma que respeite as leis envolvidas no processo de certificação de imóveis rurais. Além disso, a complexidade dos serviços é afetada por mudanças nas leis existentes.

A integração de documentos técnicos nos registros existentes, devido a particularidades e exigências, somada à necessidade de que as informações sejam compreendidas e aceitas pelos cartórios, cada um com seus próprios requisitos, aumenta a dificuldade e o tempo de entrega dos documentos necessários (Fontes 2018).

A complexidade aumenta ao se considerar a integração desses documentos em sistemas de registro existentes, que podem ter suas próprias particularidades e exigências. Ademais, a necessidade de garantir que essa informação seja compreendida e aceita por diversos cartórios, que frequentemente diferem em suas práticas, adiciona um elemento adicional de dificuldade.

Apenas no estado do Paraná, podem ser encontrados biomas como a floresta tropical pluvial (Mata Atlântica) e a savana (Cerrado), que apresentam características físicas e culturais muito distintas (Veloso 1992). O clima predominante é o subtropical

e temperado, com possibilidade de chuva no verão e períodos secos indefinidos (Maack 2012).

Existem grandes variações de altitude, como em Guarapuava/PR, por exemplo (Santos 2022). Este cenário complexo, apresentando uma diversidade de condições, pode influenciar não apenas o tempo necessário para realizar os levantamentos de campo, mas também a mobilidade necessária nessas regiões.

Inicialmente, Duarte *et al.* (2022) enfatizam a necessidade de realizar um levantamento da situação registral do imóvel, o que envolve a obtenção de uma certidão atualizada junto ao Cartório de Registro de Imóveis (CRI) competente. Para a regularização fundiária, o CRI estabelece uma série de requisitos judiciais ou extrajudiciais, conforme a legislação de registros públicos, que o proprietário ou possuidor do imóvel deve seguir. Se o registro estiver regularizado, o CRI emitirá um documento que descreve os passos a serem seguidos e lista toda a documentação necessária para o processo.

No desenvolvimento da execução do georreferenciamento de imóveis rurais, com base nas normas e na legislação estabelecida pelo INCRA, o cumprimento de diversos requisitos é essencial, incluindo a apresentação de arquivos e documentos específicos (Duarte *et al.* 2022).

4.6.3 Importância de Modelos Preditivos

De acordo com Rêgo e Pena (2012), a regressão espacial é fundamental para a compreensão e representação da dependência espacial, utilizando dados observados de áreas próximas para prever ou explicar fenômenos em regiões que ainda não foram analisadas. Modelos de regressão espacial, como o Modelo Regressivo Espacial (SAR), o Modelo de Erro Espacial (SEM) e o Modelo Espacial Geral (SAC), são utilizados para essa representação.

O modelo SAR considera a dependência entre as variáveis, quantificando-a em relação aos seus vizinhos e covariáveis, enquanto o SEM deriva informações da dependência do erro aleatório do modelo. O SAC é uma generalização dos modelos anteriores, eficaz em contextos de alta dependência espacial. O SAR é preferível em situações de dependência espacial, embora o SEM seja relevante quando múltiplos parâmetros são considerados. Em um estudo conduzido pelos autores, os modelos SAR e SEM foram aplicados para avaliar a dependência espacial entre vértices, com

a possibilidade de incorporar o SAC para comparação de qualidade se for identificada forte dependência espacial nas variáveis analisadas (Rego & Pena 2012).

Em uma análise comparativa de três métodos de inferência espacial, Lombardi (2001) investiga o uso do Processo Analítico Hierárquico (AHP), da lógica booleana e do Gama Fuzzy para a geração de mapas de áreas de interesse. A integração de mapas temáticos e numéricos por meio desses métodos resultou na criação de um mapa-síntese, que fornece uma visão abrangente e eficiente da situação, essencial para uma tomada de decisão eficaz.

O método AHP destacou-se por sua objetividade na espacialização de áreas, utilizando a ponderação de variáveis como uso do solo, declividade, malha viária e drenagem. Em contraste, a abordagem booleana mostrou-se inadequada, pois tratou todas as variáveis com igual importância, sem esclarecer as contribuições individuais de cada uma na identificação das áreas de interesse. Assim, Lombardi (2001) enfatiza a importância da escolha do método apropriado, destacando a superioridade do AHP na análise espacial (Lombardi 2001).

Pimenta, Ribeiro e Junior (2021) discutem e avaliam o uso de métodos estatísticos, como a regressão linear, para aumentar a arrecadação municipal, por exemplo. A comparação de diferentes modelos de regressão para estimar variáveis dependentes demonstrou a eficácia de se considerar a autocorrelação espacial por meio de modelos mais avançados. A análise estatística, especialmente a regressão linear múltipla, permitiu a construção de um Modelo de Regressão Linear Convencional (MRLC), reduzindo variáveis com base em testes de significância e utilizando dados cartográficos e de registros imobiliários. A precisão do modelo foi avaliada de acordo com as recomendações da norma NBR 14.653-2, através de testes estatísticos. Aplicando essa mesma metodologia ao campo do georreferenciamento, é possível determinar o tempo necessário para materializar vértices, por exemplo, na delimitação de imóveis rurais.

Cada modelo responde de forma apropriada, onde o MRLC é eficiente na previsão de preços devido à normalidade dos resíduos e à autocorrelação dos erros, e a Regressão Geograficamente Ponderada (GWR) considera a autocorrelação espacial e a heterogeneidade espacial, fazendo previsões onde a dependência da variável espacial influencia o resultado com maior precisão. Portanto, ao considerar o componente geográfico, é inevitável que um modelo de regressão espacial seja

utilizado, obtendo-se assim confiabilidade nos resultados e avaliando a influência das variáveis independentes (Pimenta, Ribeiro & Junior 2021).

4.7 CONSIDERAÇÕES SOBRE O CAPÍTULO 4

O Artigo 01 (O Desafio da Precificação de Serviços de Georreferenciamento de Imóveis Rurais: Uma Revisão Crítica) estabelece a fundamentação para toda a pesquisa. A principal contribuição deste artigo foi realizar um diagnóstico aprofundado do setor de GIR no Brasil, identificando uma lacuna metodológica crítica em sua prática comercial: a ausência de critérios objetivos para a precificação dos serviços. Através de uma revisão fundamental, este artigo demonstrou que a dependência de métodos heurísticos e da experiência subjetiva para estimar o tempo de execução em campo, o principal componente de custo, cria um cenário de incerteza e risco para todas as partes envolvidas. Portanto, a consideração central deste artigo é que ele caracteriza e valida a existência de um problema relevante, tanto do ponto de vista prático quanto acadêmico, e conclui ao apontar a necessidade de desenvolver um modelo quantitativo como uma solução viável, estabelecendo assim a base para a investigação empírica subsequente.

4.8 CONCLUSÃO

A evolução do Georreferenciamento de Imóveis Rurais (GIR), tanto em procedimentos quanto na aquisição de dados, no Brasil tem sido um marco para a regularização fundiária, promovendo segurança jurídica e uma gestão de terras eficiente. Com a crescente demanda por práticas agrárias organizadas, a importância do GIR tornou-se evidente, facilitando a definição precisa dos limites das propriedades e prevenindo conflitos de interesse.

No entanto, apesar de sua relevância, a falta de parâmetros e metodologias claras e bem estabelecidas para a precificação dos serviços de georreferenciamento representa um desafio. A avaliação de custos é frequentemente conduzida de forma empírica, levando a variações significativas nos preços cobrados por serviços similares em diferentes regiões. Essa realidade evidencia a ausência de um modelo de precificação consistente que considere as numerosas variáveis envolvidas, incluindo as características do terreno, os requisitos logísticos e as exigências regulatórias.

O tempo necessário para realizar o serviço de georreferenciamento é um dos fatores mais relevantes na determinação do preço. A previsibilidade do tempo de execução não apenas ajuda os profissionais a estimarem custos mais precisos, mas também serve como base para a alocação de recursos. Infelizmente, a falta de uma abordagem de modelagem que incorpore esta variável, juntamente com outras que influenciam o serviço, prejudica a formulação de preços justos e coerentes.

Diante deste cenário, é essencial harmonizar os critérios de precificação no setor de georreferenciamento de imóveis rurais. A criação de um modelo padronizado que integre as diversas variáveis necessárias para a definição de custos é fundamental para promover a transparência e a competitividade neste mercado. Com uma abordagem sistemática e bem definida, será possível garantir que os serviços de GIR sejam avaliados de forma justa, beneficiando tanto os profissionais da área quanto os proprietários de terras, assim como a sociedade como um todo.

5 COMPARAÇÃO ENTRE MODELOS PREDITIVOS PARA OTIMIZAÇÃO DA ESTIMATIVA DE TEMPO EM SERVIÇOS DE GEORREFERENCIAMENTO DE IMÓVEIS RURAIS

O Georreferenciamento de Imóveis Rurais (GIR), procedimento para a regularização fundiária no Brasil (Brasil, 2002), constitui uma atividade técnica cujo sucesso depende não apenas da precisão topográfica, mas também de um planejamento logístico e financeiro eficaz. Conforme estabelecido no estudo sobre os desafios na precificação destes serviços Silveira, Paiva e Goulart (2025), mencionam que um orçamento justo e competitivo é influenciado por variáveis de campo e de difícil quantificação a priori. Fatores como as distâncias de deslocamento entre vértices, as características da cobertura vegetal, a acessibilidade e a topografia do terreno introduzem um grau de incerteza na estimativa do tempo de deslocamento entre os vértices, que é uma componente relevante sobre o desse tipo de serviço. A dificuldade de encontrar ou determinar uma metodologia padronizada para essa estimativa representa, portanto, um gargalo fundamental na gestão e negociação dos serviços de GIR.

A estimativa de custos na fase de viabilidade ou inicial de um projeto é reconhecidamente uma das etapas mais críticas e desafiadoras na engenharia, além disso elas devem apresentar um grau de precisão que não indiquem custos fora da realidade (BELOTTO, 2022).

Diante deste cenário, o problema deste artigo surge da dependência do setor em regras práticas e na experiência individual dos profissionais para a formulação de propostas orçamentárias. Esta abordagem, embora válida, não possui padronização e reprodutibilidade, resultando em discrepâncias entre orçamentos e, conseqüentemente, em elevado risco comercial para quem executa (subestimação de custos) e em falta de transparência para quem contrata (superestimação de valores).

Segundo a pesquisa realizada por Maués *et al.* (2022), as equações determinadas pela regressão linear simples, após testadas em um novo orçamento, apresentaram resultados similares à precisão do orçamento com existência de anteprojeto. Desafios análogos através do desenvolvimento de modelos de estimativa paramétricos, através de parâmetros importantes, que buscam estabelecer relações

estatísticas entre as características de um projeto e seus custos, com dados de projetos anteriores (MAUÉS *et al* 2022).

A lacuna de conhecimento que este artigo busca preencher, está na ausência de um modelo quantitativo de levantamento de dados que possam apoiar, validar ou até mesmo substituir essa estimativa empírica, traduzindo as características físicas de um imóvel em uma previsão de tempo de execução com margens de erro conhecidas.

Analisando a pesquisa realizada por Oliva e Granja (2015), que utilizaram uma prática de gestão que se baseia no custo alvo o *Target Value Design* (TVD), para substituir a subjetividade por um processo integrado e baseado em dados. Assim, o objetivo geral deste artigo é desenvolver e validar um modelo preditivo para estimativa de tempo em serviços de GIR. Para tal, será realizada uma análise comparativa baseada em técnicas de aprendizado de máquina (*Machine Learning*), modelos de regressão espacial e regressão linear, de acordo que seja possível explicar o comportamento da variável dependente “Intervalo_Deslocamento”, que serve para medir o esforço e o tempo despendido, entre vértices, em campo.

Dentre as técnicas de *Machine Learning*, destacam-se os métodos de conjunto (*ensemble methods*), que, segundo Dietterich (2000), combinam múltiplos modelos para obter uma predição final mais acurada através da votação das decisões dos classificadores individuais no conjunto.

Componentes/algoritmos de re-amostragem como o *Bagging* (ex: *Random Forest* e *Extra Trees*), através de amostragem *bootstrap*, gera conjuntos de treinamento a partir de amostragem uniforme dos conjuntos treinando assim um tipo de classificador/regressor (Coelho, 2006). Outro componente/algoritmo de re-amostragem, *Boosting* (ex: *Gradient Boosting* e *XGBoost*), os conjuntos de treinamento não são gerados via amostragem uniforme e a probabilidade de uma certa amostra ser escolhida depende da contribuição dela para determinar o erro dos componentes treinados (Coelho, 2006).

Segundo Silva *et al.* (2018), o método *bootstrap* consiste na re-amostragem dos dados originais, com repetição. Em cada re-amostragem é criado um conjunto de dados, obtendo a média em cada procedimento, após um número “x” de iterações *bootstrap* é calculado a média das medias obtidas anteriormente (Silva *et al.*, 2018).

Como objetivos secundários, busca-se: (a) realizar uma análise comparativa da performance de diferentes classes de algoritmos de regressão, desde modelos

lineares tradicionais até os avançados como o *Random Forest* e *Gradient Boosting*, para identificar a abordagem mais acurada para este problema específico; (b) quantificar a importância relativa das variáveis de campo (distância, tipo de solo, declividade e a necessidade de veículo) na determinação do tempo, validando empiricamente as hipóteses levantadas na pesquisa realizada, sobre precificação, por Silveira, Paiva e Goulart (2025); e, por fim, (c) propor um tipo de modelagem que possa servir como uma ferramenta de apoio à decisão para o planejamento de projetos e a formulação de orçamentos mais precisos e competitivos, através de um sistema de verificação de tempo.

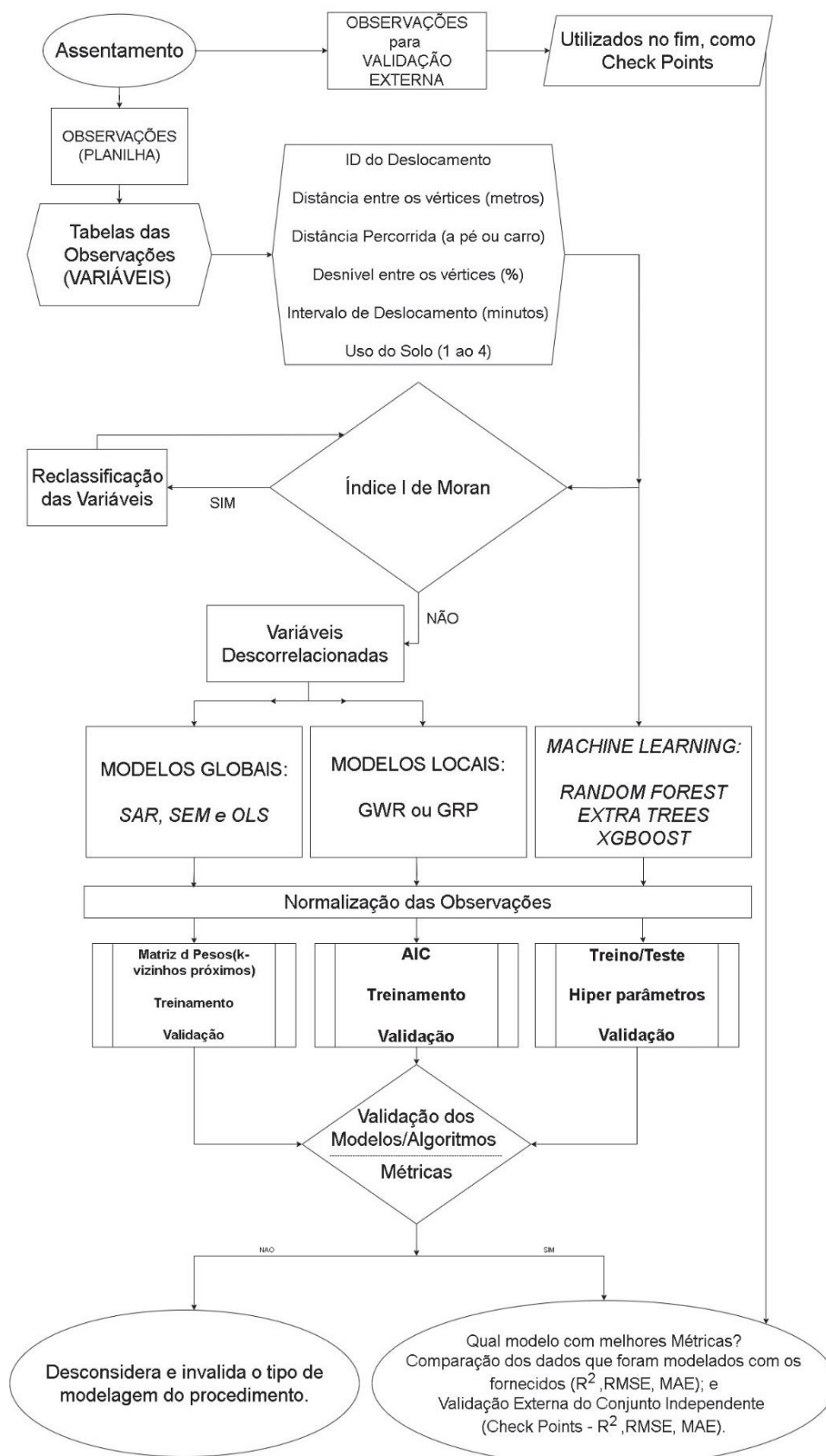
5.1 METODOLOGIA

A metodologia para esta pesquisa e para auxílio na estimação de valores de GIR parte do pressuposto que deverá ser utilizado técnicas de regressão espacial e inferência estatística. Para a elaboração do cálculo aproximado do tempo de deslocamento¹ em campo e conseqüentemente determinação dos vértices de interesse, no planejamento de campo, pode-se estipular um modelo que possa prever o tempo que cada vértice necessita para ser mensurado.

A fim de facilitar o entendimento da metodologia, o diagrama (Figura 1) apresenta o processo desde o registro das observações até a escolha do melhor tipo de modelagem de dados, por meio da validação externa de um conjunto independente.

¹ Distância percorrida entre o ponto de partida e o vértice desejado.

Figura 1 - Diagrama da Metodologia



5.1.1 Preparação do conjunto de dados

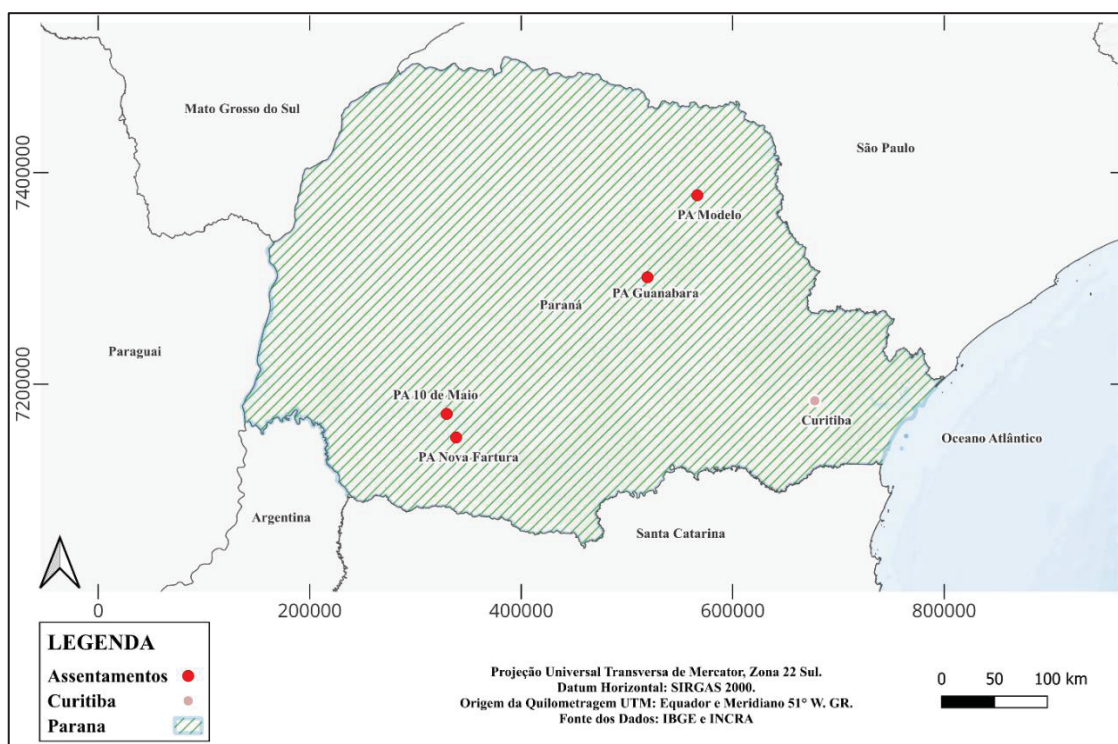
Nesta seção, será apresentado, brevemente, o procedimento pré campo, as áreas escolhidas para pesquisa, avaliando as dificuldades que serão encontradas e por fim a forma que os dados serão visualizados em campo, por meio de uma plataforma digital.

5.1.1.1 DEFINIÇÃO DA ÁREA DE ESTUDO E DESLOCAMENTO

Os estudos de caso utilizados para coleta das observações de campo relevantes para o trabalho, com coordenadas georreferenciadas, são alguns Assentamentos visitados pelo autor da pesquisa, no interior do Estado do Paraná, entre os anos de 2023 e 2025.

A utilização de diferentes áreas como estudo de caso, foi devido a viagens realizadas pelo autor realizando observações de campo, visando o levantamento topográfico georreferenciados de imóveis rurais. Na Figura 2, é apresentada a distribuição das áreas (Assentamentos) utilizadas como caso de estudo no Paraná, onde foram selecionado 4 Regiões.

Figura 2- Disposição dos Assentamentos Visitados entre 2023 e 2025.



Fonte: O Autor, 2025.

5.1.1.2 PLANEJAMENTO DE CAMPO

O INCRA fornece o projeto de parcelamento de cada assentamento a ser Georreferenciado, o qual deve ser respeitado nas próximas etapas, tanto no planejamento quanto no levantamento dos vértices de interesse, em campo.

O planejamento de campo garante que, a demarcação dos limites dos assentamentos e seus lotes, por meio do georreferenciamento, seja realizada de forma precisa e otimizada, levando em consideração elementos como o parcelamento das PA²s em questão, bases cartográficas de hidrografia, informações de propriedades confrontantes, análise do relevo da região e acessos e vias de deslocamento.

O processo envolve uma análise detalhada que abrange desde a compreensão da situação espacial do assentamento até as características de uso do solo da área desejada, passando pela determinação dos vértices a serem materializados em campo, bem como as decisões que devem ser tomadas na execução.

Tópicos são considerados no processo, os de maior relevância para esta pesquisa são:

- Avaliação dos cursos d'água na área;
- Definição das rotas de acesso aos vértices que delimitam as divisas;
- Análise da variação do relevo;
- Identificação preliminar de confrontantes;
- Verificação antecipada das exigências legais para implantação de vias; e
- Localização de potenciais bases de rastreo.

5.1.1.3 ANÁLISE PRÉVIA DAS CARACTERÍSTICAS

O uso de imagens de satélite é essencial no planejamento de campo, proporcionando uma visão detalhada das características físicas, não apenas dos assentamentos, mas como em outras regiões rurais. As imagens são usadas para comparar o projeto de parcelamento do INCRA com a situação real do uso do solo, utilizando técnicas avançadas de segmentação³ para identificar áreas agrícolas, de preservação e cursos d'água.

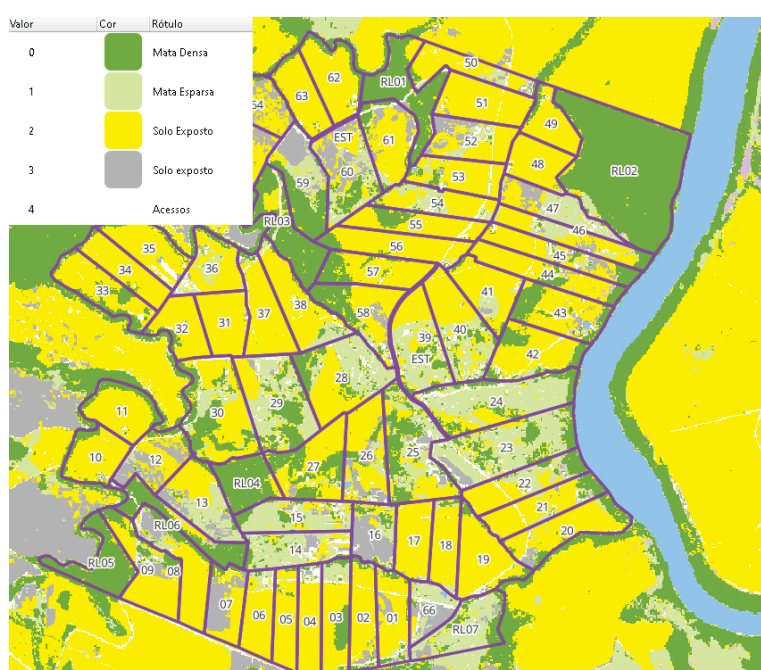
² Projeto de Assentamento, criados por meio de obtenção de terras, pelo INCRA

³ Divide a imagem em pedaços, para tratá-los posteriormente, simplificando o reconhecimento.

Com base na análise das imagens também ajudou a definir rotas eficientes de acesso aos pontos de interesse. A segmentação de uma imagem de satélite CBERS04A⁴, usando *Random Forest*, destacou diferentes classes de vegetação e desafios como a densidade que pode dificultar o acesso aos vértices necessários.

As áreas com vegetação densa, como Reserva Legal⁵ (RL), na coloração verde escuro, enquanto as áreas com vegetação esparsa⁶ estão em tons mais claros do verde. Por fim, áreas com pastagem ou agricultura, possuem cor amarela, consideradas como solo exposto, demonstrado na Figura 3.

Figura 3 - Segmentação a partir da uma imagem do satélite CBERS04A



Fonte: Autor, 2024.

5.1.1.4 ANÁLISE DO RELEVO

A análise detalhada do relevo é crucial no planejamento estratégico para georreferenciamento dos assentamentos, com o objetivo de melhorar a eficiência das operações de campo. Utilizando dados do Modelo Digital de Terreno (MDT) gerado a

⁴ Satélite de sensoriamento remoto de média resolução, com resoluções entre 2m e 60m.

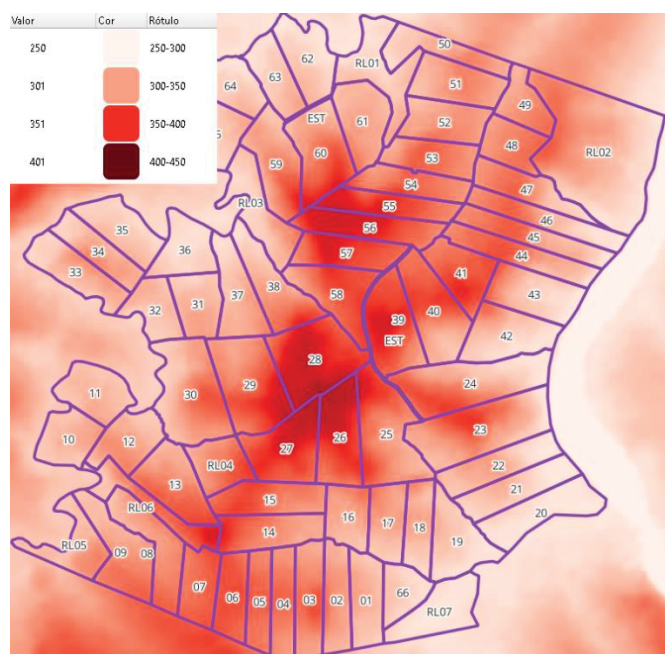
⁵ Área dentro da propriedade rural destinada a garantir o uso sustentável dos recursos naturais, conservar os processos ecológicos, preservar a biodiversidade e oferecer abrigo e proteção a fauna e flora nativa. BRASIL. Lei nº 12.651, de 25 de maio de 2012. Institui o novo código florestal brasileiro.

⁶ Espaços com pouco ou nenhuma vegetação arbustiva, cobrindo no máximo 25% da área.

partir da *Shuttle Radar Topography Mission* (SRTM⁷), disponível no portal da EMBRAPA, é possível entender sobre a topografia da área de estudo, como mostra Figura 4.

O MDT ajuda na seleção estratégica de bases para comunicação eficiente entre a base e as estações móveis (Rovers), melhorando a precisão e a agilidade dos levantamentos georreferenciados. Além disso, a análise do relevo tende a facilitar a elaboração de rotas de acesso otimizadas, minimizando os desafios em terrenos acidentados e garantindo a acessibilidade a todas as partes do assentamento, conforme exigido pelo MTGIR.

Figura 4 - MDT para análise da topografia a ser visitada, extraído do SRTM.



Fonte: Autor, 2024.

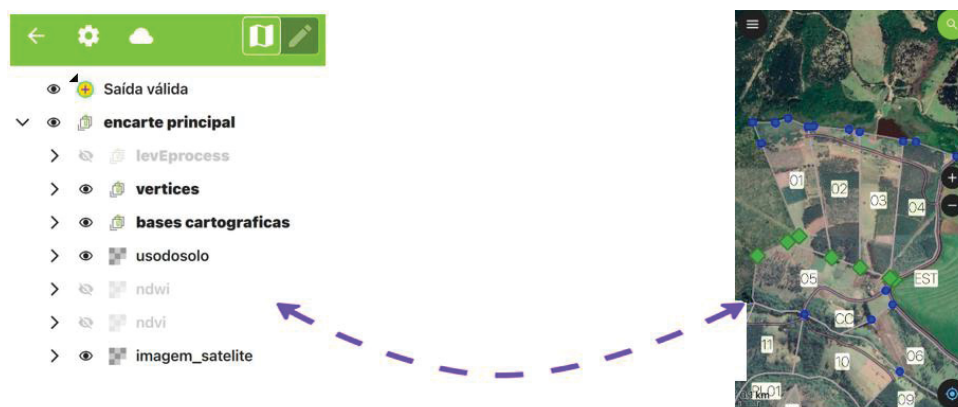
5.1.1.5 PROJETO DE NAVEGAÇÃO

O projeto de navegação visa otimizar as operações de campo e reduzir custos, mantendo a qualidade do georreferenciamento e de toda as etapas citadas no planejamento de campo, citadas anteriormente. Utilizando o aplicativo QField (Figura 5) em dispositivos móveis como smartphones e tablets, a equipe pode navegar, coletar dados de forma eficiente e tomar decisões baseadas no *software*, mesmo em ambientes offline.

⁷ Missão espacial feita em fevereiro de 2000. Auxilia a gerar MDT. NIMA e NASA foram os institutos responsáveis por esta missão.

A aplicação permite visualização detalhada do terreno, atualização de dados em tempo real, integração com informações pré-existentes, e funciona em áreas com conectividade limitada. Essa estratégia potencializa significativamente a logística e a eficácia das atividades de campo, contribuindo para o sucesso dos georreferenciamentos dos assentamentos.

Figura 5 - Projeto de Navegação no aplicativo QField.



Fonte: O Autor, 2025.

5.1.2 Descrição dos processos envolvidos no levantamento de campo

Os métodos de levantamento escolhidos para a pesquisa são as técnicas de posicionamento GPS ou GNSS, permitidos pelo MTGIR.

Diariamente, em campo, é necessário instalar e iniciar o aparelho que permanecerá rastreando um único vértice durante todo o levantamento, registrando as observações desse vértice que serve de referência a todo o levantamento, chamado de “Base” (Figura 6).

Figura 6 - “Base”, vértice utilizado como referência de levantamentos.



Fonte: O Autor, 2025.

Nos primeiros dias de trabalho, não é possível implantar os marcos de divisa dos lotes, realizada pela "locação de marcos"⁸ e/ou materialização "vértices tipo M (marco)", que serão objetos de estudo desse artigo. Isso acontece por alguns motivos importantes, como por exemplo o não conhecimento dos caminhos das estradas, necessário para definir a posição dos marcos, outro exemplo seria a ausência de exatidão do ponto de referência principal (Base) no primeiro dia.

Depois de inicializada a "Base", a próxima etapa é mapear as estradas e os cursos de água que delimitam os lotes. Esses elementos são muito importantes para o GIR, mas não serão tratados nesse artigo, pois o foco está relacionado ao registro e/ou materialização de "vértices tipo M (marco)".

5.1.2.1 LEVANTAMENTO DE CAMPO E MATERIALIZAÇÃO DOS VÉRTICES

No levantamento de campo, são identificados e/ou materializados os pontos importantes no terreno (Figura 7), que são chamados de "vértices". Esses vértices ajudam a definir os limites dos terrenos, áreas, polígonos, lotes e assentamentos.

Figura 7 - Vértice do tipo "M" (marco), para definição dos limites dos imóveis.



Fonte: O Autor, 2025.

⁸ Processo de implantar/materializar vértices por meio de suas coordenadas.

O deslocamento entre vértices é um fator que pode influenciar o tempo e, consequentemente, no custo de um GIR, mais do que a própria implantação física dos vértices. Portanto, a quantidade de tempo e recursos financeiros vai depender principalmente da distância a ser percorrida entre os pontos e da quantidade de pontos a se determinar.

5.1.2.2 REGISTRO DAS VARIÁVEIS

Para registrar o deslocamento a pé e de carro, no momento do trabalho, é possível mapear o caminho com auxílio de um aplicativo de “*tracking*” de celular, denominado “*Locus Map*”, conforme demonstrado na Figura 8.

Dessa maneira ao utilizar carros, que agilizam o deslocamento, pois nem sempre há caminhos ou possíveis vias encontradas no planejamento (que se baseia no uso de solo extraída por imagens orbitais), uma vez que em campo, esse trecho possa não existir mais ou estar coberto de vegetação.

Figura 8 - Caminhamento utilizando o aplicativo “Locus Map”.



Fonte: O Autor, 2025.

O tempo considerado, será a diferença entre o rastreo do vértice em questão e o rastreo do anterior expressado em minutos, ambos os horários são fornecidos pelo processamento que disponibiliza data e horário do rastreo de cada vértice.

Para casos em que foi possível acessar uma determinada região de carro, em seguida feito o caminhamento até o vértice desejado, foi utilizado o horário de saída do carro, registrado no aplicativo “*Locus Map*”.

Para casos em que a coluna “Veículo” das observações teve registro de “carro”, é válido informar que para esta combinação, em particular, só foram considerados os tipos de solo “estrada” e “solo exposto”, não sendo utilizada em nenhum tipo de vegetação.

A Tabela 1 demonstra como a organização de cada variável será registrada na planilha, para que seja possível comparação de tipos diferentes de modelagem. Este tópico apenas aborda a parte realizada em campo, que, com auxílio do aplicativo “Locus Map”, registra o percurso realizado.

Tabela 1 – Exemplo da Tabela/Planilha Final das Observações, contendo apenas as variáveis descritas.

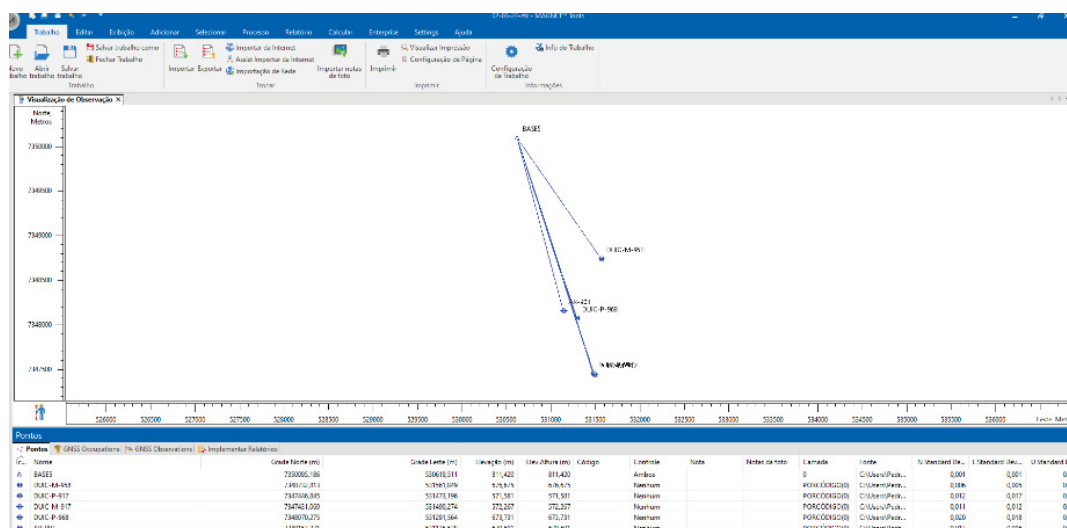
Nome	Vertice Anterior	Solo	Veículo	Distância do anterior (m)	Intervalo Deslocamento (min)	Desnível (%)
SUAV-M-988	BASE	vegetação esparsa	a pé	115	25	-0,16
RIO1	SUAV-M-988	vegetação densa	a pé	12	8	-0,22
RIO2	RIO1	vegetação densa	a pé	239	27	-0,10
SUAV-M-548	RIO2	vegetação densa	a pé	14	24	0,47
SUAV-M-781	SUAV-M-831	estrada	carro	124	14	-0,01
SUAV-M-782	SUAV-M-781	estrada	carro	192	8	0,01
SUAV-M-568	SUAV-M-782	solo exposto	a pé	74	12	-0,07
SUAV-M-551	SUAV-M-568	solo exposto	carro	171	11	-0,00

Fonte: O Autor, 2025.

5.1.2.3 PROCESSAMENTO DOS DADOS LEVANTADOS EM CAMPO

Para processamento de dados, o *software MAGNETTools* é utilizado para essa etapa, possibilitando o vínculo dos dados observados em campo gerando as planilhas do processamento GNSS (RTK e Pós Processado), que subsidiarão a montagem das tabelas, como demonstra a Figura 9.

Figura 9 - Processamento de vértices rastreados, no software MAGNETTools.



Fonte: O Autor, 2025.

5.1.3 Definição das variáveis

As variáveis independentes e dependente foram observadas em campo, registradas e avaliadas pelo autor deste artigo, pontualmente, buscando evitar erros grosseiros no processo ou que não relatassem a realidade de campo.

Visando explicar a tabela/planilha das observações as primeiras 2 colunas constam a Identificação das observações (Nome e Vértice Anterior), como mostra Figura 10.

Figura 10 - Direção do Deslocamento (ID), com os vértices ocupados.

Nome	Vertice Anterior
DUIC-M-948	BASE
DUIC-M-949	DUIC-M-948
RIO1	DUIC-M-949
RIO2	RIO2



Fonte: O Autor, 2025.

A 3ª coluna representa o uso do solo (Solo), referente a característica física do local onde o vértice é determinado, atribuído um peso de 1 a 4, de acordo com a

difficuldade causada no caminharmento, como demonstra a Figura 11, onde: 1 - Estrada; 2- Solo Exposto; 3 - Vegetação Esparsa; e 4 - Vegetação Densa.

Figura 11 - Exemplo da variável Uso do Solo, de acordo com a dificuldade identificada em campo.



Fonte: O Autor, 2025.

A 4ª Esta coluna é referente se foi necessário utilizar carro ou não, como demonstra a Tabela 2.

Tabela 2 – Exemplo da variável que identifica se foi ou não utilizado veículo.

Veículo
a pé
carro

Fonte: O Autor, 2025.

A 5ª coluna é referente a distância percorrida entre vértices (Distância do anterior), em metros, demonstrado pela Tabela 3.

Tabela 3 – Exemplo de variável que identifica a Distância Percorrida.

Distância do anterior (m)
115
12
239
14

Fonte: O Autor, 2025.

A 6ª é referente ao intervalo de deslocamento entre vértices, demonstrado pela Tabela 4.

Tabela 4 – Exemplo da variável que identifica o tempo de deslocamento entre os vértices.

Intervalo Deslocamento (min)
25
8
27
24

Fonte: O Autor, 2025.

A 7ª coluna é o resultado da razão entre a distância horizontal entre dois vértices e a distância vertical entre eles, como demonstra a Tabela 5.

Tabela 5 – Exemplo da variável que identifica a declividade do deslocamento.

Desnível (%)
-0,16
-0,22
-0,10
0,47

Fonte: O Autor, 2025.

5.1.4 Análise de correlação entre as variáveis

A modelagem preditiva é precedida por uma Análise Exploratória de Dados (AED), procedimento fundamental que visa a compreensão da estrutura, distribuição e inter-relações das variáveis coletadas. Conforme preconiza a metodologia para construção de modelos de regressão, esta fase é essencial para a identificação de padrões, tendências, valores atípicos (*outliers*) e para a verificação dos pressupostos dos modelos estatísticos a serem aplicados (GAZOLA, 2002).

De acordo com Zilli *et al.* (2021), a AED é realizada tanto na variável explicada ou dependente, quanto nas variáveis explicativas ou independentes. Nessa etapa, envolve o cálculo de medidas de estatística descritiva, como média, desvio padrão e quartis, e/ou a utilização de ferramentas visuais, como os gráficos *boxplots* e de dispersão, para identificar observações que destoam significativamente do padrão geral dos dados e que podem requerer tratamento específico.

Dado o caráter geográfico envolvido ao problema de pesquisa, a análise vai além da correlação convencional abordando a estatística espacial. A premissa central é a de que variáveis de campo, como o tempo de execução por exemplo, não são

eventos isolados, mas podem exibir dependência espacial, ou seja, seus valores podem ser influenciados pela localização e pelas características de pontos vizinhos.

De acordo com Pimenta *et al* (2021), para quantificar este fenômeno, a metodologia emprega o cálculo do Índice I de Moran, uma das principais ferramentas para a inferência da autocorrelação espacial. Este índice, que varia de -1 (perfeita dispersão) a +1 (perfeita correlação positiva), com um valor próximo de zero indicando aleatoriedade espacial, serve para determinar se a componente geográfica é uma variável significativa e se a aplicação de modelos de regressão espacial, como o *Spatial Autoregressive Model* (SAR) ou o SEM (*Spatial Error Model*), é justificada, consequentemente a comparação entre modelos, assim como feito por Rego e Pena (2012).

Estimando parâmetros de modelos SEM e SAR (Autorregressivo de Defasagem Espacial), Chipenete, Chipenete e Lima (2022) encontraram valores diferentes de zero, tanto para Rho (ρ) quanto para Lambda (λ), isto é, em ambos os modelos encontraram dependência espacial (autocorrelação) nas observações.

A análise exploratória cumpre um duplo objetivo. Primeiramente, ela permite uma análise crítica das variáveis independentes, como a identificação de multicolinearidade entre elas, o que pode levar à decisão de reclassificar ou mesmo desconsiderar uma variável para evitar redundância e instabilidade no modelo (GAZOLA, 2002).

Em segundo lugar, e de forma mais aprofundada, a análise da dependência espacial, utilizando técnicas como o Índice I de Moran e estimadores de densidade como o Kernel, conforme aplicado em estudos de inferência espacial por Rotta (2011), informa se o modelo preditivo final deve incorporar explicitamente as relações de vizinhança para explicar de forma mais acurada a variabilidade da variável dependente. Os resultados desta fase exploratória são, portanto, decisivos para guiar a subsequente etapa de modelagem, garantindo que o algoritmo escolhido seja o mais adequado à estrutura e à natureza espacial dos dados.

5.1.5 Definição da modelagem

A modelagem é uma etapa que define a capacidade de um estudo em capturar a complexidade das relações entre as variáveis. A abordagem metodológica parte de um modelo de base e evolui para técnicas mais sofisticadas, permitindo uma comparação robusta da performance preditiva. A construção do modelo segue um

roteiro estruturado, que abrange desde a identificação e transformação de variáveis até a análise crítica dos resíduos, um processo fundamental para garantir a validade estatística dos resultados (GAZOLA, 2002).

A abordagem inicial para a estimativa de custos e tempo em projetos de engenharia frequentemente se baseia em Modelos de Regressão Linear Convencionais (MRLC), que buscam estabelecer uma relação estatística global entre um conjunto de variáveis preditoras e uma variável resposta (MAUÉS et al 2022).

Assim, Técnicas de Regressão Espacial se dividem em duas categorias principais: global e local. Os modelos espaciais globais, como o SAR e o SEM, assumem que a dependência espacial é um processo contínuo e homogêneo em toda a região. O modelo SAR postula que o valor da variável dependente em um local é influenciado diretamente pelos valores dessa mesma variável em locais vizinhos, enquanto o modelo SEM atribui a dependência espacial aos termos de erro não observados (REGO e PENA, 2012).

Conforme explicado por Sousa *et al.* (2012), o método de regressão por mínimos quadrados, chamado de *Ordinary Least Square* (OLS), é uma forma de estimar alguns parâmetros globais, que possibilitam analisar a relação, por sua vez, constante no espaço, entre 2 ou mais variáveis.

Ao estudar o uso da regressão por mínimos quadrados ordinários (OLS ou MQO), Figueiredo Filho *et al.* (2011), enfatiza que, além das relações entre as variáveis dependente e independente serem representadas por uma função linear, ainda é o modelo estatístico mais empregado na ciência política contemporânea.

Segundo Zilli *et al.* (2021), os modelos globais consideram os dados como estacionários, ou seja, que os processos e as relações são constantes em toda a área de estudo. Em dados georreferenciados, esta premissa é frequentemente violada, tornando necessário o uso de modelos que considerem a componente espacial para se obter resultados confiáveis.

Segundo Zilli *et al.* (2021), os modelos locais, como a Regressão Geograficamente Ponderada (RGP ou GWR), partem da premissa de que o espaço é heterogêneo e que as relações entre as variáveis podem mudar ao longo da região, calibrando uma regressão para cada ponto de observação.

Para o caso em que as relações entre as variáveis de campo podem ser intrinsecamente não-lineares, algoritmos de *Machine Learning* validam alternativas aos modelos paramétricos possibilitando testes.

Segundo Rodrigues (2023), a busca em rede (*Grid Search*), técnica de otimização de parâmetros, almeja encontrar os melhores resultados para o problema específico, testando todas as possíveis variações dos dados, por meio de processo treino/teste.

De acordo com Breiman (2001), o classificador *Random Forest* (Floresta Aleatória), que opera construindo múltiplas árvores de decisão para melhorar a precisão e controlar o sobre ajuste.

O *Gradient Boosting Regressor* (GBR), segundo Friedman (2001), pode ser utilizado tanto na regressão como classificação. Ele constrói árvores de forma sequencial, onde cada nova árvore busca corrigir os erros residuais da anterior, sendo menos influenciável por dados “sujeitos”, isto é, que ainda não passaram por filtragem de especialista.

De acordo com Marinho (2021), o *XGBoost*, algoritmo de *Machine Learning*, funciona bem para problemas de regressão. Este algoritmo se baseia em *gradiente boosting*, técnica de aprendizagem de máquina utilizada para problemas de regressão e classificação, produzindo um modelo de preditivo na forma de árvores de decisão (Marinho, 2021).

Um algoritmo que, segundo de Paula (2022), tem vantagens sobre a técnica *XGBoost* como, maior velocidade de treinamento, maior eficiência e menor uso de memória, melhor precisão e maior capacidade para lidar com dados em larga escala de várias bases de observações simultâneas é o *LightGBM*.

De uma forma resumida, dos Santos (2024), compara e avalia o método *Extra Trees*, com outros métodos como o Random Forest, sendo aplicável em regressão e classificação. A principal característica se dá pela possibilidade de inserir um nível maior de aleatoriedade na criação das árvores de decisão, buscando melhorar a precisão do modelo e reduzir o sobre ajuste, resultando em modelos mais generalizados (dos Santos *et al.* 2024).

A superioridade de tais métodos, reside em sua capacidade de capturar interações complexas e padrões não-lineares que modelos lineares ou espaciais tradicionais podem não identificar. Ao utilizar métodos que não se prendem a regras simples e que tem a capacidade de aprender os padrões complexos e as nuances diretamente dos dados de campo, pode resultar em uma previsão de tempo mais realista e precisa.

5.1.6 Métricas de avaliação, validação dos modelos e visualização dos resultados

A construção de um modelo preditivo, conforme destaca Gazola (2002), exige a adoção de uma estratégia metodológica detalhada e estruturada para garantir sua confiabilidade e capacidade de predição.

A validação de um modelo vai além da simples análise de ajuste aos dados de treinamento, sendo um passo para evitar a criação de modelos com baixo poder de generalização que podem levar a decisões equivocadas (GAZOLA, 2002). Neste contexto, um dos principais desafios é evitar o sobre ajuste (*overfitting*), fenômeno que ocorre quando o modelo se adequa excessivamente aos dados de treinamento, capturando ruídos em vez das tendências gerais (FRIEDMAN, 2001).

A capacidade de predição do modelo em novos dados pode então ser verificada ao se comparar os valores preditos com os valores reais, analisando as discrepâncias geradas (GAZOLA, 2002).

Para realizar essa análise quantitativa, pode estabelecer um conjunto de métricas de desempenho que avaliam diferentes aspectos da performance do modelo (DAM *et al.* 2022).

O poder de explicação geral é medido pelo Coeficiente de Determinação (R^2), que informa a proporção da variabilidade da variável dependente que é explicada pelo modelo (GAZOLA, 2002).

As métricas de performance das técnicas de regressão espacial tanto para o modelo global como local, segundo Zilli *et al.* (2021), podem ser por meio do valor de Critério de Akaike (AIC), Critério de Informação Bayesiano (BIC), *loglikelihood* (LIK), coeficiente de determinação (R^2) e raiz quadrada do erro médio quadrático (RMSE).

Para avaliar a magnitude dos erros de predição, o Erro Médio Quadrático (MSE) e sua raiz, o RMSE, são utilizados, sendo métricas que penalizam fortemente erros grandes e distantes do alvo. O RMSE possui a vantagem de ser expresso na mesma unidade dos dados originais, facilitando sua interpretação (CARMO E SILVA, 2023; DAM *et al.* 2022).

Como alternativa, o Erro Médio Absoluto (MAE) calcula a média das diferenças absolutas e, por não elevar os erros ao quadrado, é menos sensível a valores discrepantes. Conforme observam Carmo e Silva (2023), a proximidade entre os valores de RMSE e MAE sugere uma distribuição de erros homogênea, enquanto uma diferença acentuada indica a presença de *outliers*.

Adicionalmente, o Erro Percentual Médio Absoluto (MAPE) expressa o erro em termos percentuais, o que facilita a comparação entre modelos cujas variáveis dependentes possuam diferentes unidades de medida (CARMO E SILVA, 2023).

A validação cruzada *k-fold*, segundo Silva, Matsuoka e Roffato (2024), é um método utilizado para estimativa de precisão e seleção de técnicas de *Machine Learning* que separa o conjunto de dados em subconjuntos.

A utilização conjunta dessas métricas, frequentemente apresentadas em tabelas e gráficos de desempenho para uma análise quantitativa, permite uma avaliação completa e comparativa da performance de cada modelagem, conforme a abordagem adotada por de Carvalho (2024).

Silva (2020), entende que é importante para validação de modelos realizar a comparação entre os valores experimentais e preditos de um conjunto de dados independentes. Chamado de validação externa ou validação independente, são utilizados para avaliar o desempenho e generalidade de predição dos modelos, por meio de métricas estatísticas (Silva 2020).

5.1.7 Tecnologia/biblioteca e linguagem utilizada

Buscando identificar os melhores desempenhos de modelos e algoritmos, para modelagem de observações, Pereira (2024) utilizou a biblioteca *PyCaret*, que simplifica a criação de algoritmos de *Machine Learning* e auxilia na escolha de quais métodos serão mais relevantes utilizar, bem como uma previa das métricas finais.

Ribeiro, Frazão e Oliveira e Sá (2018), realizaram uma comparação entre algoritmos e tecnologia/biblioteca de *Machine Learning*, considerando as necessidades computacionais de projeto, bem como as características de cada tecnologia/biblioteca, entre elas a popularidade, compatibilidade com GPU, Linguagens de programação, vantagens e desvantagens e algoritmos suportados.

As tecnologias/bibliotecas avaliadas por Ribeiro, Frazão e Oliveira e Sá (2018), foram: *Scikit-Learn*, *TensorFlow*, *Weka*, *Dask*, *Microsoft Cognitive Toolkit*, *OpenCV*, *Chainer*, entre outras.

Tabelando a tecnologia/biblioteca *Scikit-Learn*, Ribeiro, Frazão e Oliveira e Sá (2018) indicam que além de possuir a linguagem em *python*, estão entre suas vantagens a documentação detalhada, *Machine Learning*, bom desempenho e solução para grandes *datasets*. Além disso, Algoritmos como Classificação e Regressão são suportados.

Python é uma linguagem de programação que possui estruturas de dados de alto nível e abordagem simples, afirma Barbosa (2018), para programação orientada a objetos. Assim, se torna uma linguagem ideal para scripts e desenvolvimento de aplicativos, na maioria das plataformas (Barbosa, 2018).

5.2 RESULTADOS E DISCUSSÕES

Este capítulo apresenta os resultados quantitativos das comparações entre os modelos preditivos testados, bem como a análise das variáveis que mais influenciam a estimativa de tempo nos serviços de GIR.

5.2.1 Dados e Observações

As observações registradas em campo, compõe uma planilha organizada com os levantamentos dos vértices.

Inicialmente, mais de 400 observações foram realizadas. Analisando pontualmente essas observações, pode-se perceber que, nem todas elas seriam utilizadas, pois agregaria erros humanos ou particularidades de campo na modelagem. Erros como pausas no deslocamento entre vértices, deslocamentos em direções incorretas, problemas com equipamentos e veículos durante o deslocamento, entre outros que foram mitigados manualmente.

Por fim, foi possível utilizar uma planilha composta por 323 observações, com variáveis estudadas e bem definidas, demonstrada pela Tabela 6.

Além disso, teve-se o cuidado para que os tipos de solo e veículo, respeitassem uma hierarquia e um intervalo definido, como por exemplo evitar que o tipo de solo “estrada” fosse registrado como “estrda”, fazendo com que não modelasse corretamente esses dados, bem como o registro “caro” para o tipo de veículo utilizado.

Vale ressaltar que as observações que demonstrassem ausência de um dos dados seriam excluídas do processo ou revisitadas na análise, reduzindo erro humano.

Tabela 6 – Exemplo da Tabela/Planilha das Observações, completa, com variáveis definidas, comentários feitos sobre cada trecho, coordenadas dos vértices e equipe que realizou o rastreo.

Nome	Vértice Anterior	Solo	Veículo	Distância do anterior (m)	Intervalo Deslocamento (min)	Des-nível (%)	Obs	X	Y	Z	Equipe
SUAV-M-988	BASE	Vegetação esparsa	a pé	115	25	-0,16	Saida da base 9.30 e deixaram o carro as 09.43	3391 69,81	71525 12,24	657 ,55	03
RIO1	SUAV-M-988	Vegetação densa	a pé	12	8	-0,22		3391 71,28	71525 24,52	654 ,95	03
RIO2	RIO1	Vegetação densa	a pé	239	27	-0,10		3392 10,56	71526 21,15	644 ,82	03
SUAV-M-548	RIO2	Vegetação densa	a pé	14	24	0,47		3392 24,71	71526 18,26	651 ,34	03
SUAV-M-559	SUAV-M-548	Vegetação esparsa	a pé	366	25	0,07		3391 43,21	71523 71,59	668 ,69	03
RIO3	SUAV-M-559	Vegetação densa	a pé	5	8	-0,54		3391 39,81	71523 74,99	666 ,00	03
SUAV-M-549	AL-MOÇO	Vegetação esparsa	a pé	401	17	0,07	carro deixado as 13.52	3388 44,38	71526 14,16	693 ,55	03

Fonte: O Autor, 2025.

5.2.2 Preparação das Variáveis

Neste passo, as variáveis são distinguidas em 2 conjuntos de dados, as preditoras ou independentes e as alvo ou independentes, indicadas na Figura 12.

Figura 12 - Distinção das Variáveis.

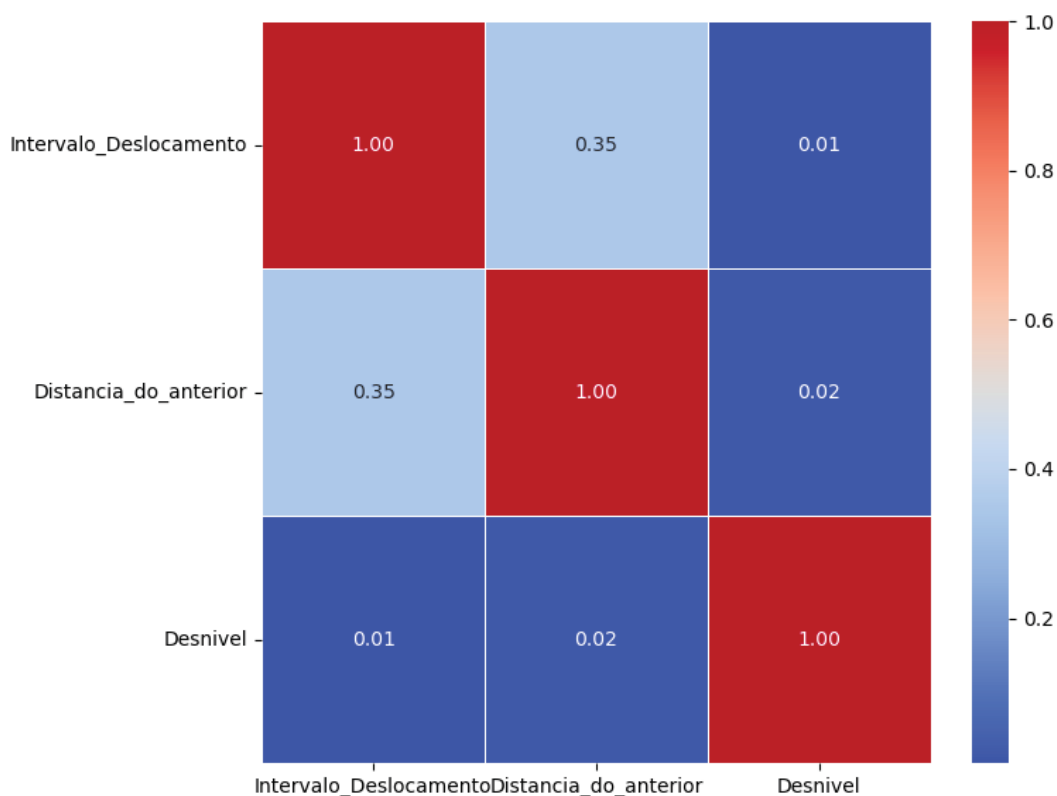


Fonte: O Autor, 2025.

5.2.3 Análise exploratória

Neste passo, avalia-se se há ou não a correlação entre as variáveis utilizadas na modelagem, visto que o valor dessa relação varia entre “1” e “-1”, sendo o valor “0” indica a ausência de correlação (Gazola 2002). O mapa de calor representa o Índice I de Moran, como demonstra a Figura 13, é possível visualizar essa questão.

Figura 13 - Mapa de Calor das Correlações entre as Variáveis, através do Índice I de Moran.



Fonte: O Autor, 2025.

Assim, é possível visualizar que as variáveis independentes não são correlacionadas entre si, muito menos com a variável dependente. Mesmo que o valor de correlação entre as variáveis “Intervalo_Deslocamento” e “Distancia_do_anterior” venha a ser um valor próximo de 0.4, ela é pouco significativa para interferência na modelagem das observações, sendo irrelevante tanto para os modelos globais (SAR e SEM) quanto para o GWR.

5.2.4 Normalização das observações

Segundo Meerholz *et al.* (2023), a normalização é um passo importante, quando se tem um conjunto de dados ou observações onde suas características possuem grandes diferenças de intervalos, sendo necessário a transformação desses dados em valores próximos, evitando distribuição de peso de forma equivocada.

Nesse caso, as variáveis possuem diferentes escalas, como por exemplo a “Distância do anterior” e “Desnível”. Isso quer dizer que a primeira varia entre 1 e 10mil metros, enquanto a segunda varia entre $-0,3\%$ e $+0,3\%$.

Assim SILVEIRA¹ *et al.* (2021), afirmaram que a padronização de um conjunto de dados é comum para estimadores de *Machine Learning*. Estes tipos de estimadores podem ter uma resposta equivocada quando não se encontram em uma distribuição padrão Gaussiana. A padronização faz parte do pré-processamento de dados, possibilitando padronizar as escalas dos conjuntos das observações.

Desta maneira, para a presente pesquisa, utilizou-se a técnica de padronização, implementada pela ferramenta “StandardScaler” na biblioteca scikit-learn, transformando cada variável para que possua média 0 e desvio padrão “1”.

5.2.5 Conjunto de treino e teste

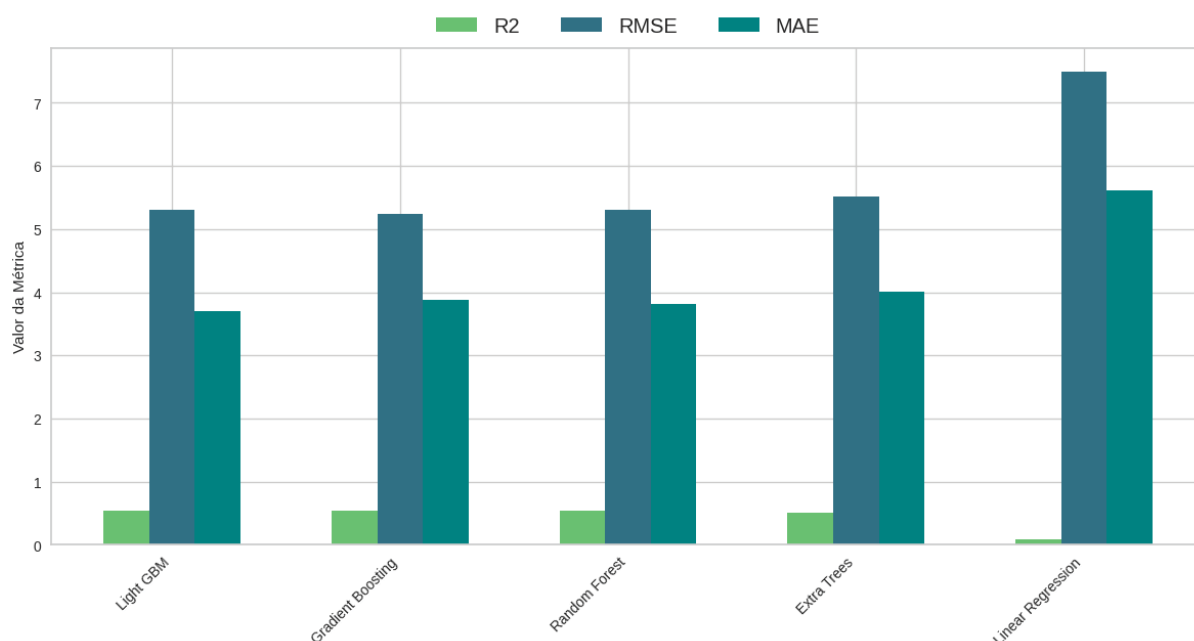
Segundo Friedman (2001), dois terços (67%) das observações são consideradas “amostras de aprendizado”, isto é, elas treinam o modelo. As demais observações, são utilizadas como “amostras teste”, as quais irão testar a modelagem realizada.

Gazola (2002), reitera essa necessidade de deixar algumas observações de fora do treinamento, sendo uma forma de medir valores finais tendenciosos.

Pereira (2024), realiza uma avaliação e revisão bibliográfica de alguns algoritmos de *Machine Learning*, entre eles o *Random Forest*, *Extra Trees Regressor* e *Light GBM*, além de explicar as funcionalidades da biblioteca *PyCaret* e em qual momento ela pode ser utilizada na modelagem.

Seguindo essa sugestão, o Gráfico 1 demonstra o comportamento esperado através da biblioteca *PyCaret* para as observações a serem modeladas.

Gráfico 1 – Métricas esperadas, resultadas da biblioteca PyCaret



Fonte: O Autor, 2025.

Assim, é possível estimar que os melhores algoritmos de *Machine Learning*, segundo a *PyCaret*, podem alcançar, aproximadamente, 55% para R^2 , 5,5 minutos para RMSE e 4 minutos para o MAE.

Já para Regressão Linear, o valor encontrado para R^2 se aproxima de 10% , 7,5 minutos para RMSE e 5,5 minutos para o MAE.

Além disso, a fim de garantir que comparação entre modelos, *Learning Machine* apenas, fosse realizado entre suas melhores versões, foi utilizado um processo de otimização por hiper parâmetros. Por meio da ferramenta *GridSearchCV*, da biblioteca *Scikit-learn*, que segundo Pedregosa *et al.* (2011) foi projetada para explorar sistematicamente um espaço de parâmetros de um modelo. Combinada a validação cruzada 5 *folds* (*k-fold*), são testadas diferentes combinações de parâmetros para cada algoritmo, assegurando que a performance final de cada modelo represente o seu máximo potencial preditivo para o conjunto de dados utilizado.

Vale ressaltar que, os algoritmos de Machine Learning passaram por um processo de hiper parametrização, onde foram testados com diferentes taxas de

aprendizado e números de árvores, resultando o melhor e o mais ajustado modelo, para predição.

Três tipos de utilização das observações foram comparados:

- A Tabela 7 utilizou 67/33, isto é, 67% das observações treinaram os modelos (216), os outros 33% serviram de teste (107). Conforme sugerido por Friedman (2001).

Tabela 7 – Métricas dos Algoritmos de Machine Learning (Divisão 67/33)

Algoritmo de ML	Tipo	R ²	RMSE	MAE
<i>Random Forest</i>	<i>Machine Learning</i>	62,41%	4.67 (min)	3.44 (min)
<i>Extra Trees</i>	<i>Machine Learning</i>	59,92%	4.83(min)	3.37(min)
<i>Gradient Boosting</i>	<i>Machine Learning</i>	58,06%	4.94(min)	3.58(min)
<i>Light GBM</i>	<i>Machine Learning</i>	55,80%	5.07(min)	3.81 (min)
<i>XGBoost</i>	<i>Machine Learning</i>	45,21%	5.64 (min)	4.01 (min)

Fonte: O Autor, 2025.

- A Tabela 8 utilizou 80/20, isto é, 80% das observações treinaram os modelos (258), os outros 20% serviram de teste (65). Conforme sugerido por Degtyarev (2022).

Tabela 8 – Métricas dos Algoritmos de Machine Learning (Divisão 80/20)

Algoritmo de ML	Tipo	R ²	RMSE	MAE
<i>Light GBM</i>	<i>Machine Learning</i>	61,11%	4.22 (min)	3.34 (min)
<i>XGBoost</i>	<i>Machine Learning</i>	58,09%	4.38 (min)	3.41(min)
<i>Extra Trees</i>	<i>Machine Learning</i>	52,57%	4.66 (min)	3.24 (min)
<i>Random Forest</i>	<i>Machine Learning</i>	50,50%	4.76 (min)	3.46(min)
<i>Gradient Boosting</i>	<i>Machine Learning</i>	49,35%	4.82(min)	3.46 (min)

Fonte: O Autor, 2025.

- Outro tipo de método avaliado é a validação cruzada *K-fold*. Conforme realizado por Silva, Matsuoka e Roffato (2024), demonstrado na Tabela 9.

Tabela 9 – Métricas dos Algoritmos de Machine Learning (Validação Cruzada)

Algoritmo de ML	Tipo	R ²	RMSE	MAE
<i>Gradient Boosting</i>	<i>Machine Learning</i>	56,41%	5.13 (min)	3.88(min)
<i>Random Forest</i>	<i>Machine Learning</i>	52,54%	5.34 (min)	4.02 (min)
<i>Light GBM</i>	<i>Machine Learning</i>	48,23%	5.62 (min)	4.07 (min)
<i>XGBoost</i>	<i>Machine Learning</i>	45,52%	5.75 (min)	4.37 (min)
<i>Extra Trees</i>	<i>Machine Learning</i>	45,33%	5.75 (min)	4.31 (min)

Fonte: O Autor, 2025.

Os Modelos Espaciais e Regressão Linear, foram treinados com todas as observações, apresentando as métricas de acordo com a Tabela 10.

Tabela 10 – Métricas para os Modelos SAR, GWR e OLS

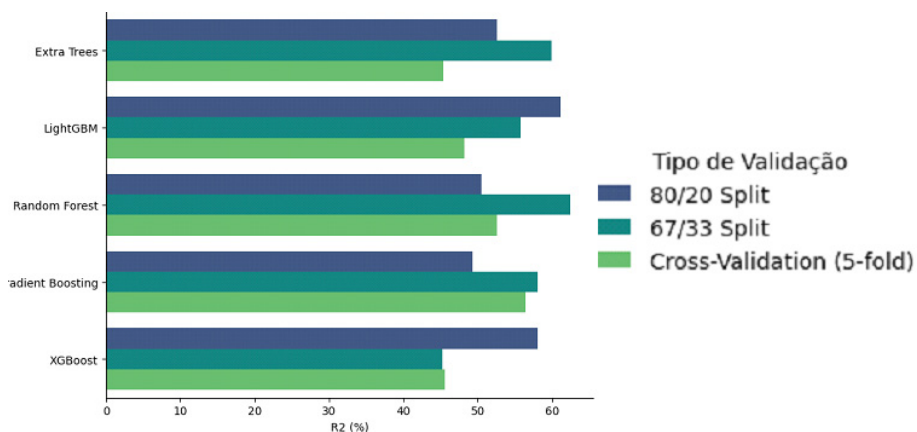
Modelo	Tipo	R ²	R ² Ajustado	AIC	BIC
SAR	Espacial Global	35,46%	33,69%	2153.7465	2183.9677
GWR	Espacial Local	34,83%	32,00%	2166.6672	2221.0956
OLS	Regressão Linear	32,26%	30,97%	2164.3622	2190.8057

Fonte: O Autor, 2025.

5.2.6 Análise Resumida das Métricas

Com base nas métricas apresentadas e no item 5.1.6 (Métricas de avaliação, validação dos modelos e visualização dos resultados), é possível encontrar métricas que são universais para os tipos de técnicas envolvidas, como o R², demonstrado no Gráfico 2.

Gráfico 2 - Coeficiente de Determinação (R²) para algoritmos ML

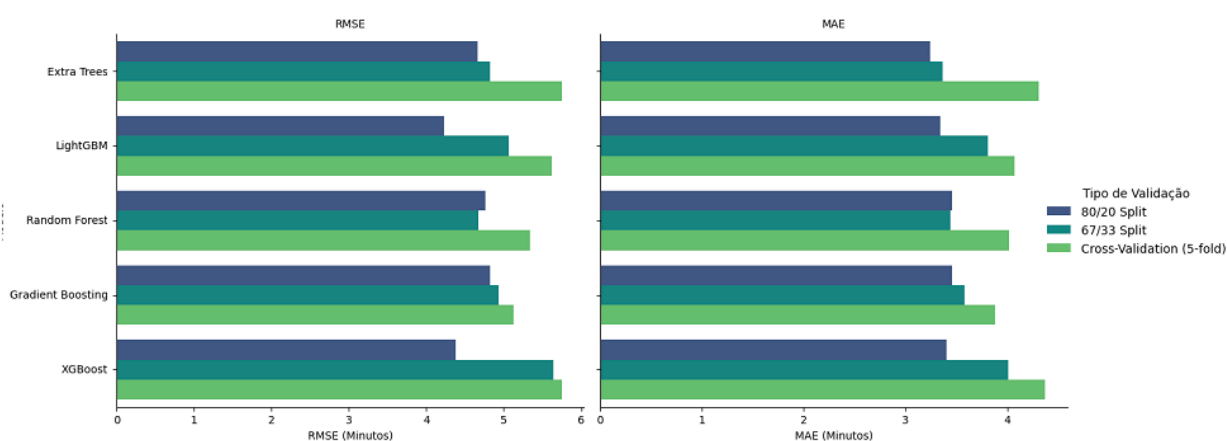


Fonte: O Autor, 2025.

Os Algoritmos de *Machine Learning*, notadamente os baseados em conjuntos de árvores como *XGBoost*, *LightGBM* e *Extra Trees*, otimizados, demonstraram uma superioridade consistente nos cenários de avaliação, como mostra

Com métricas como R^2 alcançando aproximadamente 60% de explicação do tempo de deslocamento entre vértices, RMSE se aproximando de 4 min e o MAE se aproximando de 3 min, como demonstra o Gráfico 3, estes modelos superaram tanto a regressão linear (OLS obtendo $R^2 \approx 32\%$) quanto os modelos espaciais (SAR e GWR obtendo $R^2 \approx 35\%$). A estabilidade deste resultado foi confirmada através de múltiplas divisões de dados (80/20 e 67/33) e por validação cruzada (5-fold), que atestou um desempenho similar, representados no Gráfico 3.

Gráfico 3 - RMSE e MAE para algoritmos de ML



Fonte: O Autor, 2025.

5.2.7 Influência das Variáveis

A análise da influência das variáveis, Tabela 11, foi conduzida através da importância das variáveis (*feature importance*) para algoritmos de *Machine Learning*.

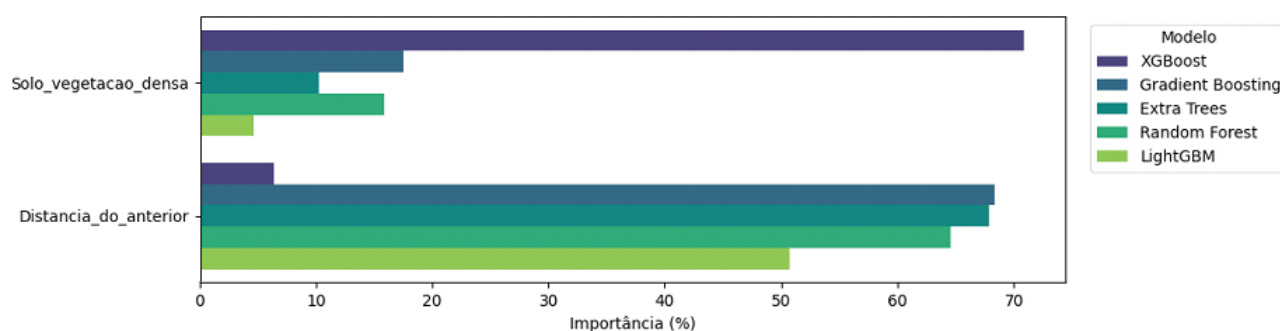
Tabela 11 - Importância das Variáveis.

Variável Modelo	Distancia do anterior	Desnível	Solo exposto	Solo vegetação densa	Solo vegetação esparsa	Veículo carro
<i>Gradient Boosting</i>	68.35%	4.62%	1.79%	17.45%	1.65%	6.14%
<i>Random Forest</i>	64.56%	11.52%	1.46%	15.87%	2.38%	4.21%
<i>LightGBM</i>	50.75%	33.16%	5.04%	4.60%	1.68%	4.77%
<i>XGBoost</i>	6.31%	1.57%	1.95%	70.87%	5.98%	13.32%
<i>Extra Trees</i>	67.82%	12.64%	3.15%	10.17%	2.21%	4.01%

Fonte: O Autor, 2025.

Analisando as tabelas acima, da importância das variáveis, para algoritmos de *Machine Learning*, é possível identificar que a variável independente “Distância do anterior” é o fator mais preditivo para a variável dependente “Intervalo de deslocamento entre vértices”, chega na maioria dos casos entre 60% e 70%, visualizado no Gráfico 4.

Gráfico 4 – Maiores Importâncias das Variáveis nos Algoritmos



Fonte: O Autor, 2025.

Para o caso do *XGBoost*, em particular, foi atribuído a maior parte da predição (aproximadamente 71%) à variável independente “vegetação densa”. Esta técnica pode ter identificado uma interação forte e não-linear entre a vegetação densa e o tempo, diferente de outros algoritmos.

5.2.8 Validação Externa de um Conjunto Independente

Por fim, são avaliados os melhores Algoritmos, dos três tipos treinamento, para observações desconhecidas pela modelagem. Assim é possível verificar e validar se eles estão realmente entregando as métricas esperadas (Silva, 2020).

Utilizando 21 observações na primeira análise, sem remover *outliers*, as métricas obtidas são apresentadas na Tabela 12:

Tabela 12 - Métricas De Desempenho Nas Novas Observações.

Algoritmo de Machine Learning	R²	RMSE	MAE
<i>Random Forest</i> (Treino 67/33)	42,21%	5.58 (min)	4.16 (min)
<i>Extra Trees</i> (Treino 67/33)	40,91%	5.64 (min)	4.27 (min)
<i>Gradient Boosting</i> (Validado por CV)	38,98%	5.73 (min)	4.63 (min)
<i>Random Forest</i> (Validado por CV)	36,81%	5.84 (min)	4.29 (min)
<i>LightGBM</i> (Treino 80/20)	25,73%	6.33 (min)	5.05 (min)
<i>XGBoost</i> (Treino 80/20)	-20,68%	8.07 (min)	6.28 (min)

Fonte: O Autor, 2025.

Embora as métricas tenham sido inferiores as resultadas no 5.2.5, o qual avaliou apenas as observações de treinamento em métodos diferentes de validação, é possível considerar aceitável aos padrões de campo, onde uma leve discrepância entre o RMSE e MAE foi apresentada indicando presença de *outliers*.

Em alguns casos, imprevistos acontecem, como percorrer pequenos trechos em um longo período, demonstrado pela Figura 14, onde foi percorrido 10 metros (pontos em amarelo) em 10 minutos, em solo exposto, consideradas *outliers* para serviços de GIR.

Figura 14 - Outlier identificado no deslocamento.



Fonte: O Autor, 2025.

Assim, retirando alguns *outliers* (6 observações) e refazendo as métricas com 15 observações, é possível verificar a Tabela 13

Tabela 13 - Métricas De Desempenho Nas Novas Observações sem Outliers.

Algoritmo de <i>Machine Learning</i>	R²	RMSE	MAE
<i>Random Forest</i> (Validado por CV)	83,00%	2.74 (min)	2.32 (min)
<i>Random Forest</i> (Treino 67/33)	81,35%	2.87 (min)	2.38 (min)
<i>Gradient Boosting</i> (Validado por CV)	74,94%	3.33 (min)	2.93 (min)
<i>Extra Trees</i> (Treino 67/33)	57,46%	4.34 (min)	3.10 (min)
<i>LightGBM</i> (Treino 80/20)	56,04%	4.41 (min)	3.45 (min)
<i>XGBoost</i> (Treino 80/20)	30,31%	5.55 (min)	4.11 (min)

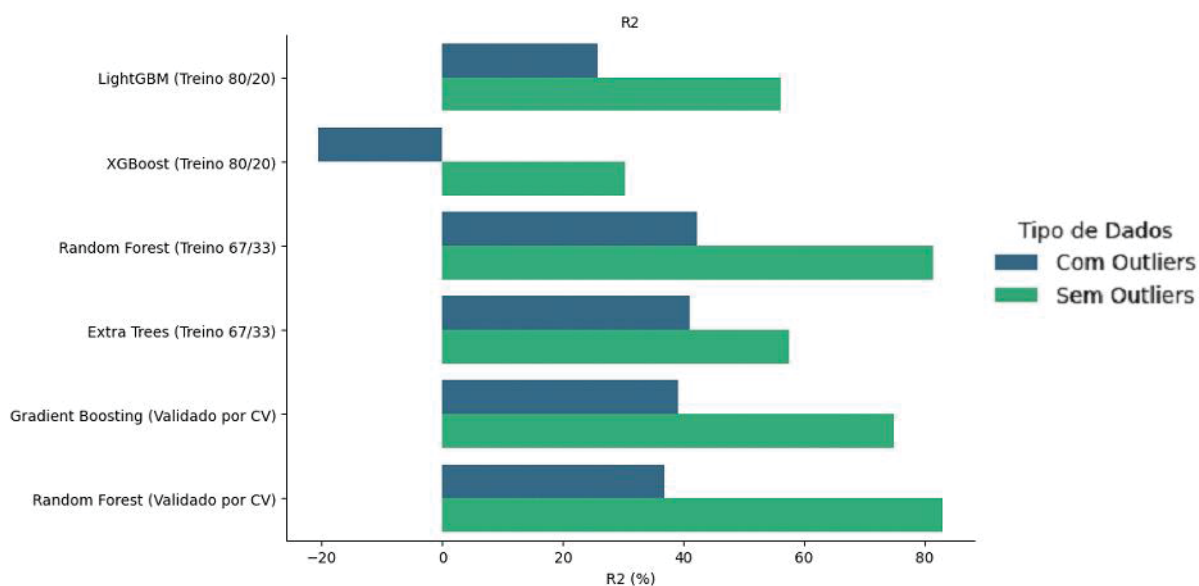
Fonte: O Autor, 2025.

Com isso, os valores aproximados entre o RMSE e MAE, indica que a presença de *outliers* foi tratada e para o caso do *Random Forest*, obteve-se um ótimo R².

Como uma boa parte dos *outliers* identificados e removidos pertenciam a vértices com vegetação densa, foi notável a melhora do resultado do modelo *XGBoost*, validando a relação que foi encontrada do tempo com esse tipo de vegetação.

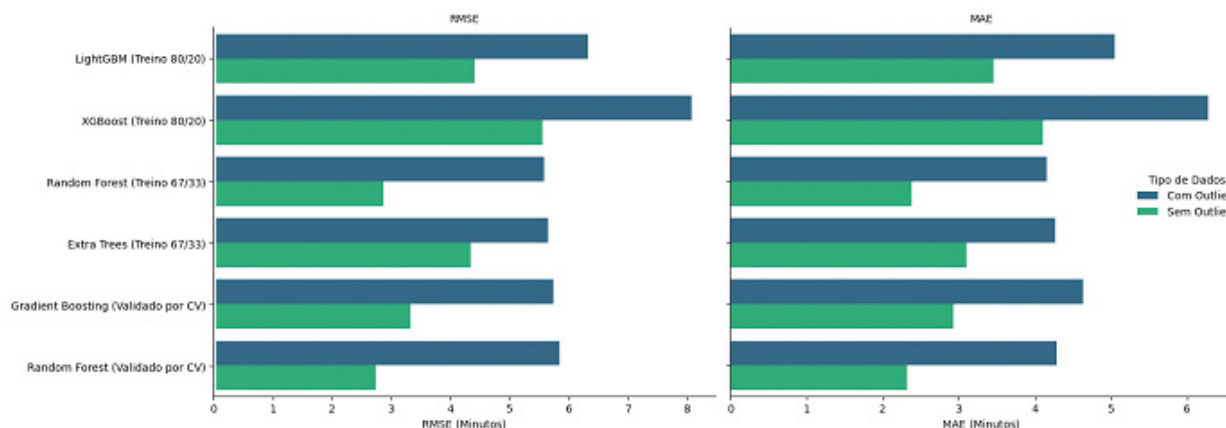
Para ficar mais clara a diferença entre as métricas de desempenho com e sem *outliers*, o Gráfico 5 e Gráfico 6 representam as comparações entre as 2 tabelas anteriores.

Gráfico 5 – Comparação da influência de outliers, para o R²



Fonte: O Autor, 2025.

Gráfico 6 – Comparação da influência de outliers, entre RMSE e MAE



Fonte: O Autor, 2025.

Assim, é possível verificar ainda melhor a qualidade dos algoritmos, testando sua aplicabilidade em novas observações não conhecidas muito menos inseridas na modelagem.

O Gráfico 6 esclarece essa questão, ao registrar uma melhora de aproximadamente 3 minutos e 2 minutos entre as métricas com e sem *outliers*, para RMSE e MAE respectivamente.

5.3 CONSIDERAÇÕES SOBRE O CAPÍTULO 5

O segundo artigo - Comparação entre Modelos Preditivos para Otimização da Estimativa de Tempo em Serviços de Georreferenciamento de Imóveis Rurais - apresentou o desenvolvimento e a validação de uma solução quantitativa baseada em observações, realizadas em campo. A principal contribuição desta pesquisa foi a aplicação e a comparação de múltiplas classes de algoritmos, relatando a superioridade conclusiva dos modelos de Aprendizado de Máquina de conjunto (ensemble methods), como o *XGBoost* e o *Extra Trees*, sobre as abordagens de regressão linear e espacial. O estudo não apenas forneceu um modelo preditivo com desempenho superior aos demais, mas também quantificou a influência das variáveis de campo, validando empiricamente o conhecimento prático, por fim, a validação externa, refinou ainda mais a qualidade da modelagem. As considerações finais desta etapa são que a adoção do aprendizado de máquina, incluindo a otimização de hiper

parâmetros e a validação cruzada, representa uma ferramenta metodologicamente defensável e eficaz para otimizar a precificação do GIR.

5.4 CONCLUSÃO

Este estudo se propôs a estudar e validar uma solução quantitativa para otimizar a estimativa de tempo em serviços de GIR, um dos componentes que impactam diretamente a precificação deste tipo de serviço. A abordagem metodológica se baseou em modelos preditivos, capazes de superar as limitações dos métodos heurísticos tradicionalmente empregados, como regressão linear e previsões empíricas. Para tal, foi realizada uma análise comparativa sistemática de múltiplos algoritmos de regressão, implementados através da biblioteca *scikit-learn*, sobre um conjunto de dados de observações de campo.

Os resultados demonstraram a superioridade dos algoritmos de *Machine Learning* como o *Extra Trees* (Otimizado), *XGBoost* (Otimizado) e *Random Forest* (CV 5-fold) alcançaram um R^2 superior a 0,60, mesmo com diferentes métodos de validação. Indicaram que são capazes de explicar aproximadamente 60% da variabilidade no tempo de deslocamento entre vértices. Em nítido contraste, a Regressão Linear, incorporada pelo modelo de mínimos quadrados ordinários (MQO ou OLS) mostrou-se insuficiente para capturar a complexidade do fenômeno, com um R^2 de apenas 0,32. Este desempenho representa uma capacidade de explicação da variabilidade dos dados, quase duas vezes, superior à do melhor modelo espacial, o SAR, que obteve um R^2 de aproximadamente 0,35.

Esta discrepância valida que a relação entre as variáveis de campo e o tempo de execução é fundamentalmente não-linear, exigindo modelos capazes de aprender interações complexas, conforme teorizado por Breiman (2001) e Friedman (2001) para os algoritmos de *bagging* e *boosting*, respectivamente.

O método de validação externa, assim nomeado por Silva (2020), utilizou os melhores modelos para estimar o tempo de deslocamento, então desconhecidos pelo modelo, apresentou resultados satisfatório após a retirada de alguns *outliers*, validando de forma satisfatória técnicas de *Machine Learning*, mais especificamente o método de conjunto *Random Forest*.

A respeito das interações entre as variáveis de campo (distância, desnível, solo, etc.), é possível verificar que a capacidade dos algoritmos de *Boosting* e *Bagging* de modelar essas relações complexas, provou ser mais decisiva para a precisão

preditiva do que a modelagem da dependência espacial. A análise de importância das variáveis foi validada, apontando a “Distância do anterior” e “Desnível” como os fatores mais influentes na maioria dos algoritmos de *Machine Learning*.

Outra métrica que torna essa modelagem mais aplicável, é o RMSE. Pode-se indicar que a cada vértice visitado pela equipe, um erro estimado de no máximo 5 min, para mais ou menos, já é esperado. Assim é possível afirmar que a cada 12 vértices implantados (estimativa de 1 dia de serviço), o erro estimado pode ser de aproximadamente 1 hora.

A análise de importância das variáveis dos algoritmos vencedores, revelou que na maioria dos casos a distância entre os vértices é a preditora mais influente e a vegetação densa, para um único caso. Este achado oferece uma validação quantitativa para o conhecimento empírico dos profissionais da área e reforça a coerência do modelo, que aprendeu a priorizar os fatores que, intuitivamente, mais impactam o tempo de trabalho em campo.

A Regressão Linear (OLS) e os modelos espaciais como (SAR e GWR) mesmo que exista correlações entre variáveis, obtidas pelos critérios de informação AIC, BIC e Índice I de Moran, obtiveram métricas inferiores aos resultados das técnicas de *Machine Learning*, sendo assim, reprovadas na proposta deste artigo.

As implicações práticas desta pesquisa são diretas e significativas. A aplicação de algoritmos preditivos validados, permitem que as empresas e profissionais da área transitem de um processo de orçamento subjetivo para uma metodologia proativa e orientada a dados. A capacidade de gerar estimativas de tempo com maior precisão, como indicado pelas métricas de erro obtidas, possibilita a elaboração de propostas comerciais mais justas e competitivas, a mitigação de riscos financeiros associados à subestimação de custos e a otimização da alocação de equipes e recursos. Esta abordagem alinha-se à necessidade, já estabelecida em áreas correlatas como a avaliação de imóveis, de se empregar metodologias cientificamente estruturadas para aumentar a confiabilidade e a transparência dos resultados.

Reconhece-se, contudo, as limitações do presente estudo. Embora os algoritmos expliquem uma porção majoritária da variância, cerca de 30% a 40% dela permanece não explicada, sugerindo a influência de fatores não contemplados no escopo inicial. Pesquisas futuras poderiam se beneficiar da inclusão de novas variáveis para refinar ainda mais o poder preditivo da modelagem. Sugere-se a

incorporação de dados meteorológicos (como precipitação e temperatura no dia do serviço), características da equipe (como o nível de experiência dos operadores, desgaste diário da equipe) e uma análise mais detalhada da cobertura vegetal por meio de índices de sensoriamento remoto, por exemplo. A exploração de arquiteturas de modelos ainda mais complexas, como redes neurais, também representa um caminho promissor.

Em suma, este estudo validou com sucesso a aplicação de técnicas de *Machine Learning* como uma ferramenta superior para a otimização da estimativa de tempo em serviços de GIR. Ao eleger esse tipo de modelagem, com modelos práticos, de alto desempenho e ao identificar um caminho claro para investigações futuras, oferece uma contribuição para a modernização, precisão e eficiência do planejamento e da precificação não apenas para medição de imóveis rurais, mas para levantamentos topográficos em geral.

6 CONCLUSÃO GERAL

A investigação inicial apresentada nesta dissertação estabeleceu com clareza uma lacuna metodológica fundamental no setor de GIR: a ausência de um paradigma quantitativo para a estimativa do tempo de execução dos serviços de campo. Foi demonstrado que a dependência de critérios subjetivos e da experiência profissional, resulta em um processo de orçamentação com incerteza, agregado de alto risco para os envolvidos. Esta constatação serviu como subsídio para a segunda fase da pesquisa, que se dedicou a construir e validar uma metodologia orientada a observações de campo, para solucionar, da melhor forma encontrada, essa problemática.

O mérito central deste trabalho reside na rigorosa análise comparativa entre diferentes filosofias de modelagem. Em vez de se limitar a uma única abordagem, a pesquisa confrontou o desempenho de modelos de regressão linear, modelos espaciais (global e local) e algoritmos de aprendizado de máquina. Através de um protocolo de otimização e validação, foi possível não apenas identificar os melhores algoritmos, mas também entender a natureza que envolve o fenômeno estudado.

O achado central desta pesquisa é que a variabilidade no tempo de execução é dominada por interações complexas e não-lineares entre as variáveis de campo. Isso é evidenciado pela superioridade conclusiva dos algoritmos de aprendizado de máquina de conjunto, que consistentemente superaram os outros modelos em precisão preditiva. Isso sugere que, embora relações lineares e efeitos espaciais estejam presentes, eles são de importância secundária em relação aos efeitos de distância, relevo, tipo de solo e uso de veículo, que apenas os modelos de aprendizado de máquina foram capazes de capturar eficazmente.

Adicionalmente, a pesquisa enriqueceu a compreensão do fenômeno ao confirmar a relevância da dimensão geográfica, mesmo que ela não tenha sido o principal fator preditivo. A análise espacial demonstrou a existência tanto de dependência espacial global (autocorrelação) quanto de não-estacionariedade local, provando que a influência das variáveis de campo não é homogênea em toda a área de estudo. Este achado, por si só, já justifica a inadequação de modelos excessivamente simplistas, como regressão linear.

Portanto, esta dissertação valida a transição de uma metodologia baseado em métricas, modelagens e observações realizadas pelo autor, substituindo o

empirismo na precificação de serviços de GIR. Ao apresentar e validar o aprendizado de máquina como a mais eficaz, este trabalho oferece uma base mais segura e objetiva para a tomada de decisão gerencial e orçamentária.

Além disso, a pesquisa pode servir de subsídio para elaboração de uma ferramenta de precificação ou embasar uma nova metodologia, de 2 etapas, que leve em consideração o procedimento elaborado pelo autor e assim, com a identificação de outras variáveis, em uma segunda etapa, compor uma modelagem que busque explicar o custo envolvido nos serviços de GIR.

Embora este estudo tenha focado em serviços de GIR, em uma região específica, a estrutura metodológica desenvolvida possui potencial de generalização. O fluxo de trabalho (desde a análise exploratória e a engenharia de *features*, até a comparação entre algoritmos de aprendizado de máquina e espaciais com otimização de hiper parâmetros) é diretamente aplicável a outros problemas em diversos ramos da engenharia. Esta abordagem pode ser utilizada em um cenário que busque substituir processos empíricos e subjetivos por modelos preditivos orientados a dados, como na estimativa de custos em obras civis, no planejamento logístico em agronomia ou na previsão de produtividade em recursos naturais, onde variáveis espaciais e não-lineares tem um papel determinante.

Ao reconhecer as limitações sobre este estudo, é válido ressaltar o Coeficiente de Determinação (R^2) de aproximadamente 60% indica que, embora o modelo explique a maior parte da variabilidade no tempo de deslocamento, cerca de 40% dessa variação permanece não explicada por outros fatores não contemplados. Outro fator, a respeito do conjunto de dados, embora representativo, é geograficamente concentrado no estado do Paraná, o que impõe um limite à generalização direta do modelo treinado para outras regiões do Brasil com biomas e características de relevo distintas. Por fim, variáveis como "uso de solo" foram tratadas de forma categórica, uma simplificação que futuras pesquisas poderiam aprimorar com dados mais específicos.

Sugestões para pesquisas futuras, como aprimorar a precisão dos algoritmos, sugere-se a incorporação de variáveis mais específicas, como dados meteorológicos (precipitação e temperatura no dia do serviço), o nível de experiência da equipe de campo (níveis de habilidade, conhecimento e horas trabalhadas no dia) e uma análise mais detalhada da cobertura vegetal através de índices de sensoriamento remoto (como o NDVI), para quantificar a densidade da vegetação de forma mais precisa que

uma simples categorização. A frente de pesquisa mais inovadora seria a exploração de métodos híbridos, que combinam o poder preditivo do *Machine Learning* com os *insights* da análise espacial. Uma abordagem promissora seria utilizar os coeficientes locais gerados pela Regressão Geograficamente Ponderada (GWR) como novas variáveis de entrada para um modelo XGBoost ou Extra Trees, criando assim um modelo que seja, ao mesmo tempo, sensível às complexidades não-lineares e às variações geográficas locais.

7 REFERÊNCIAS

- ABNT. Associação Brasileira de Normas Técnicas 1994. NBR 13133: Execução de levantamentos topográficos. Rio de Janeiro, 35 p.
- Alves, C.A., Varotto, L.F., & Gonçalves, M.N. 2012, Objetivos e métodos de preço no varejo: estudo na zona sul de São Paulo. *Revista de Administração de Empresas*, 52(6). <https://doi.org/10.1590/s0034-75902012000600003>
- Araújo, G.C., do Patrocínio Gomes, V.A., Batista, B.C., & de Freitas, R.R. 2023, Precificação de Produtos e Serviços: Uma Análise Bibliométrica. *Revista Científica Foz*, 6(1), 16-16.
- ASSOCIAÇÃO DOS PROFISSIONAIS DE ENGENHARIA AGRIMENSURA DO ESTADO DE SÃO PAULO. 2015, Simulador Referencial de Preços de Serviços de Agrimensura. Acessado em: 17/08/2025 Disponível em: <https://www.amiranet.com.br/simulador>
- Ballar, V.C.M. 2021, Georreferenciamento de imóvel rural e processo de certificação via SIGEF: estudo de caso do CECA–UFAL.
- Barbosa, F.R.M. 2018, Otimização de hiperparâmetros em algoritmos de árvore de decisão utilizando computação evolutiva.
- Batista, P., Matos, Z., & Graça, A. 2011, Auto-percepção das competências profissionais em profissionais do desporto: efeito da área de intervenção e da experiência profissional. *E-balonmano. com: Revista de Ciencias del Deporte*, 7(2), 117-131.
- Belotto, T. 2022, Modelo paramétrico de previsão de custos de edificações residenciais multifamiliares na fase de viabilidade.
- Bomfim, E.A., & Passarelli, J. 2011, Custos e formação de preços (7th ed.). São Paulo: IOB Thomson.
- Borges Júnior, C.R., Brito, J.L.S., & Schmidt, M.A.R. 2017, Estimativa de Área no Sistema Geodésico Local para Parcelas Rurais com Cálculo de Área Referenciado ao Sistema UTM. *Revista Brasileira de Cartografia*, (69/7), 1411-1425.
- BRASIL. Lei n. 6.015 de 31 de dezembro de 1973. Dispõe sobre Registros Públicos e dá outras providências. Acessado em: 01/07/2025 Disponível em: https://www.planalto.gov.br/ccivil_03/leis/16015compilada.htm.
- BRASIL. Lei nº 10.267, de 28 de agosto de 2001. Altera dispositivos das Leis n: os 4.947, de 6 de abril de 1966, 5.868, de 12 de dezembro de 1972, 6.015, de 31 de dezembro de 1973, 6.739, de 5 de dezembro de 1979, 9.393, de 19 de dezembro de 1996, e dá outras providências. 2001a. Acessado em: 01/07/2025 Disponível em: https://www.planalto.gov.br/ccivil_03/leis/leis_2001/110267.htm.
- Breiman, L. 2001, Random forests. *Machine learning*, 45(1), 5-32.
- Burnham, K.P., & Anderson, D.R. 2004, Multimodel inference: understanding AIC and BIC in model selection. *Sociological methods & research*, 33(2), 261-304.
- Carmo, C.R.S., & de Melo Silva, J.R. 2023, Aprendizado de máquina e prestação de serviços de armazenamento de dados: Métricas para análise e validação de algoritmos previsores. *Revista GeTeC*, 12(38).
- Chicati, M.L., Neto, O.T., Nanni, M.R., Oliveira, R.B.D., Cézar, E., & Chicat, M.S. 2019, Uso de drones no georreferenciamento de imóveis rurais para levantamento de vértices de difícil acesso. *Simpósio Brasileiro de Sensoriamento Remoto*, 19, 283-286.
- Chipenete, C.F., Chipenete, G.H.N., & de Lima, R.R. 2022, Modelos de regressão ajustados a dados espaciais de áreas com sementes melhoradas de milho em Moçambique e Regression models fitted to spatial area data which use d improved maize seeds in Mozambique. *Brazilian Journal of*

- Development*, 8(3), 20017-20034.
- Coelho, G.P. 2006, *Geração, seleção e combinação de componentes para ensembles de redes neurais aplicadas a problemas de classificação* (Doctoral dissertation, Dissertação de Mestrado. Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas. Campinas).
- CUNHA, G., Nascimento, G.B., & Parreira, G.F. 2014, Propostas de melhorias na prestação de serviços de uma empresa de topografia a fim de reduzir os índices de retrabalho. *Ponta Grossa, PR*, 3, 9.
- da Silva, C.M., Vieira, J.W., Neto, N.A., & Gonçalves, G.L. 2018, Controle estatístico de qualidade usando o método bootstrap. *Scientia Plena*, 14(3).
- da Silva, D.J. 2020, Aplicação de espectroscopias vibracionais e métodos quimiométricos avançados no controle de qualidade de blendas poliméricas recicladas de PEBD/PEAD.
- Dalaqua, R.R., Amorim, A., & Flores, E.F. 2010, Utilização de métodos combinados de avaliação imobiliária para a elaboração da planta de valores genéricos. *Boletim de Ciencias Geodesicas*, 16(2).
- de Carvalho, C. D. R. Modelagem e Algoritmos de Controle Robusto Neuro-Adaptativo para Foguetes de Combustível Sólido.
- de Freitas Dam, R.S., Salgado, W.L., Schirru, R., & Salgado, C.M. 2022, Application of radioactive particle tracking and an artificial neural network to calculating the flow rate in a two-phase (oil–water) stratified flow regime. *Applied Radiation and Isotopes*, 180, 110061.
- de Paula, K.G.F. 2022, Sensor virtual de alvura em polpa branqueada de celulose baseado em Inteligência Artificial.
- DE SOUSA, Adélia Maria Oliveira *et al.* Análise da distribuição de fogos ativos detectados pelo sensor ATSR 2 e sua relação com variáveis demográficas, ambientais e infraestruturais: comparabilidade entre métodos de regressão OLS e GWR. **Revista Brasileira de Cartografia**, v. 3, n. 64, p. 704-721, 2012.
- De Souza, M.C., & Peixe, B.C.S. 2019, Custos para definir o preço das diárias da prestação de serviço da hotelaria: estudo empírico de caso aplicado. In *Anais do Congresso Brasileiro de Custos-ABC*.
- Degtyarev, V.V. 2022, Machine Learning Models for Predicting Bond Strength of Deformed Bars in Concrete. *ACI Structural Journal*, 119(5).
- Disperati, A.A., de Oliveira Filho, P.C., Bobrowski, R., & Dlugosz, F.L. 2003, Análise temporal da cobertura florestal do município de Irati-PR, utilizando imagens satelitárias Landsat. *RECEN-Revista Ciências Exatas e Naturais*, 5(1), 115-123.
- dos Santos Felipetto, H., Carvalho Granemann, D., Robson Elias, A., & Fernanda De Faveri Marquesine, M. 2023, Análise do Georreferenciamento de Imóveis Rurais (Lei nº 10.267/2001) na Região Sudoeste do Paraná (2001-2022). *Revista Paranaense de Desenvolvimento*, 44(145).
- dos Santos, J.V., Felipe, V.F., Xavier, E.F.M., de Oliveira, T.A., Jale, J.S., & Junior, S. F.A.X. 2024, A STUDY OF THE STANDARDIZED PRECIPITATION INDEX (SPI) AND MACHINE LEARNING TECHNIQUES FOR DROUGHT PREDICTION IN THE STATE OF PARAÍBA, BRAZIL. *RECIMA21-Revista Científica Multidisciplinar-ISSN 2675-6218*, 5(10), e5105736-e5105736.
- Duarte, A.H., Affonso, M.R., Vicente, R.R., & Ribas, E.B. 2022, Proposta à execução de um georreferenciamento de A a Z em uma propriedade rural. *PARAMÉTRICA*, 14(2).
- Felipetto, H.S., Carvalho Granemann, D., Robson Elias, A., & Fernanda De Faveri Marquesine, M. 2023, Análise do Georreferenciamento de Imóveis Rurais (Lei nº 10.267/2001) na Região Sudoeste do Paraná (2001-2022). *Revista Paranaense de Desenvolvimento*, 44(145).
- Ferreira, R.D.S. 2019, Estudo comparativo entre os métodos de levantamento fotogramétrico (VANT) e o topográfico (estação total) para cálculo de volume em uma mineração.

- Figueiredo Filho, A., Disperati, A.A., & Watzlawick, L.F. 2003, Integração de geotecnologias como topografia, GPS e base cartográfica na empresa florestal. *RECEN-Revista Ciências Exatas e Naturais*, 5(2), 187-199.
- Figueiredo Filho, D., Nunes, F., da Rocha, E.C., Santos, M.L., Batista, M., & Júnior, J.A.S. 2011, O que fazer e o que não fazer com a regressão: pressupostos e aplicações do modelo linear de Mínimos Quadrados Ordinários (MQO). *Revista Política Hoje*, 20(1).
- Fontes, M.L. 2018, Certificação de imóveis rurais gerada pelo INCRA/SIGEF: transcrição da certificação em serviço de registro de imóveis. *Revista Brasileira de Geomática*, 6(3). <https://doi.org/10.3895/rbgeo.v6n3.7114>
- Friedman, J.H. 2001, Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.
- Gazola, S. 2002, Construção de um modelo de regressão para avaliação de imóveis.
- Holanda Nepomuceno Nobre, L., Pereira de Macêdo, Á.F., Takashi Kato, H., & Francisco Maffezzoli, E.C. 2016, ESTRATÉGIAS DE PRECIFICAÇÃO DE NOVOS SERVIÇOS EM PEQUENAS E MÉDIAS EMPRESAS. *Revista Eletrônica de Estratégia & Negócios*, 9(3). <https://doi.org/10.19177/reen.v9e320163-21>
- INCRA - Instituto Nacional de Colonização e Reforma Agrária. Manual Técnico para Georreferenciamento de Imóveis Rurais – 2ª Edição. Brasília-DF: INCRA, 2022. Acessado em: 01/07/2025 Disponível em: https://sigef.incra.gov.br/static/documentos/manual_geo_imoveis.pdf.
- INCRA. Instituto Nacional de Colonização e Reforma Agrária 2013. Manual técnico de limites e confrontações: georreferenciamento de imóveis rurais. Brasília, 24 p.
- Júnior, C.R.B., Brito, J.L.S., & Schmidt, M.A.R. 2017, Estimativa de Área no Sistema Geodésico Local para Parcelas Rurais com Cálculo de Área Referenciado ao Sistema UTM. *Revista Brasileira de Cartografia*, (69/7), 1411-1425.
- Lombardi, R.J.R. 2001, Identificação de áreas vulneráveis à ocorrência de incêndios florestais, provocados por atividades antrópicas, utilizando diferentes métodos de inferência espacial.
- Lorenzi, G.S.G., & Rosa, L.D. 2007, Controle de custos em uma empresa de topografia.
- Maack, R. 2012, *Geografia Física do Estado do Paraná*. 4ª edição. Ponta Grossa: Editora UEPG. ISBN: 978-85-7798-140-3.
- Malaman, C.S., & Amorim, A. 2017, Método para determinação de valores na avaliação imobiliária: comparação entre o Modelo de Regressão Linear e Lógica Fuzzy. *Boletim de Ciências Geodésicas*, 23(1), 87-100.
- MARINHO, Tiago Lima *et al.* 2021, Otimização de hiperparâmetros do XGBoost utilizando meta-aprendizagem.
- Maués, F.C.A., de Melo, K.P., de Oliveira Leão, C.B., & Serra, S.M.B. 2022, Estimativa de custos paramétricos de construção de edifícios usando modelo de regressão linear. *Gestão & Tecnologia de Projetos*, 17(2), 19-37.
- Meerholz, P. V., de Mattos, W. D., de Freitas, D. S., & Inacio, E. C. 2023, Um Estudo Sobre o Uso de Modelos de Machine Learning para Manutenção Preditiva Industrial. *Anais do Computer on the Beach*, 14, 311-318.
- Ok, E. 2025, A Comparative Study of Conventional Survey Methods and UAV Structure-from-Motion for Topographic Data Collection. Acessado em: 01/07/2025 Disponível em: [A-Comparative-Study-of-Conventional-Survey-Methods-and-UAV-Structure-from-Motion-for-Topographic-Data-Collection.pdf](#).

- Oliveira, M., Goulart, A.Á., Borges, U., dos Anjos Paiva, C., & de Paula, E.V. 2024, Levantamento de campo voltados ao georreferenciamento de imóveis rurais: Experiências e ferramentas essenciais.
- Oliveira, S.R.D., & Piccinini, V.C. 2011, Mercado de trabalho: múltiplos (des) entendimentos. *Revista de administração pública*, 45, 1517-1538.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. 2011, Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825-2830.
- Pereira, R.P. 2024, Modelagem preditiva da produtividade da soja utilizando aprendizado de máquina.
- Pimenta, F.de S., Ribeiro, F.V., & Cruz Júnior, D. C. 2021, Modelagem de Regressão Espacial para Estimativa de Valores em Massa a partir de Cartografia Cadastral. *Revista Brasileira de Cartografia*, 73(1). <https://doi.org/10.14393/rbcv73n1-51484>
- Pissinatti, L.G., & Dias, E. 2022, A MARGEM DE CONTRIBUIÇÃO COMO MÉTRICA DE PRECIFICAÇÃO VISANDO LUCRO PARA A EMPRESA. *Revista Ibero-Americana de Humanidades, Ciências e Educação*, 8(10). <https://doi.org/10.51891/rease.v8i10.7491>
- Purificação, É.F.D.S., Neves, S.D.S., Cavalheiro, R.T., & Conceição, G.P.de M. 2022, Precificação no Contexto das Micro e Pequenas Empresas (MPes): Podemos Sofisticar ou Ainda Pecamos no Básico? *Revista de Contabilidade e Controladoria*, 14(3). <https://doi.org/10.5380/rcc.v14i3.84458>
- Rêgo, C.V., & Pena, M.G. 2012, Análise dos modelos de regressão espacial SAR, SEM e SAC. *Mono-grafia. Instituto de Ciências Exatas, Departamento de Estatística, Universidade de Brasília, Brasília*.
- Ribeiro, A.C.F., Frazão, R., & Oliveira e Sá, J. 2018. Quebra-cabeças Machine Learning: Como selecionar Use Cases, Algoritmos e Tecnologias?.
- Rohr, P.B. 2023, Comparação do erro altimétrico da aerofotogrametria com drone com o método de posicionamento RTK GNSS.
- Santana Santos, B., Campos Tinoco Ribeiro, V., & Silva, I. 2019, *UMA METODOLOGIA DE PRECIFICAÇÃO DINÂMICA PARA PRODUTOS E SERVIÇOS EM SMART CITIES*. <https://doi.org/10.17648/sbai-2019-111535>
- Santos, L.A.D. 2022, Análise de dados e modelagem matemática aplicada ao georreferenciamento de imóveis rurais.
- Silva, T.T.D.S.C., Rodrigues, J.L.N., Amaral, G.P., & Peixoto Júnior, A.A. 2017, Conhecimento dos profissionais de enfermagem sobre sepse: estudo em um hospital universitário de Fortaleza/Ceará.
- Silveira¹, I. T., Kakitani, M. T., de Arruda Pereira, M., & Ferreira, A. 2021, Aplicações de técnicas de machine learning nos dados de telemetria do satélite brasileiro SCD-2.
- Sluter, C.R., Camboim, S.P., Iescheck, A.L., Pereira, L.B., Castro, M.C., Yamada, M.M., & Araújo, V.S. 2018, A Proposal for topographic map symbols for large-scale maps of urban areas in Brazil. *The Cartographic Journal*, 55(4), 362-377.
- Tinoco, B.S.S.V.C., & Silva, R.I. 2019, Uma Metodologia de Precificação Dinâmica para Produtos e Serviços em Smart Cities.
- Toledo, B.H.C., & Bertotti, L.G. 2014, Breve histórico da certificação de imóveis rurais no Brasil e apresentação do Sistema de Gestão Fundiária-SIGEF Brief history of certification of rural properties in Brazil and presentation of Landed Property Management System-SIGEF. *Ambiência*, 10(3), 839-847.
- Veloso, H.P. 1992, *Manual técnico da vegetação brasileira*. Rio de Janeiro: IBGE. ISBN: 978-85-240-4272-0.

8 APÊNDICE – CÓDIGOS IMPLEMENTADOS

PyCaret: [PyCaret.ipynb - Colab](#)

https://colab.research.google.com/drive/1Y36PbDp4-ADUGbNsTdMyMqYrv9naZMep#scrollTo=TZwT1_vsszlk

Modelagem: ([Modelagem.ipynb](#) - [Colab](#))

https://colab.research.google.com/drive/1gTxQvBe7xDcG_fZeUYUkvAjkJk7EpNDf?usp=sharing

Validação Externa: ([Validacao Externa.ipynb - Colab](#))

https://colab.research.google.com/drive/1Y36PbDp4-ADUGbNsTdMyMqYrv9naZMep#scrollTo=TZwT1_vsszlk