



# **ANÁLISE DA INFLUÊNCIA DE FATORES GEOGRÁFICOS NA IDENTIFICAÇÃO DE ÁREAS DE CULTIVO DE CACAU NO PA ITATÁ: UMA ABORDAGEM COM MODELOS DE REGRESSÃO**

**ESPACIAL**

ORIENTADORES:

ISABEL ESCADA

SIDNEI SANT'ANNA



# **INTRODUÇÃO**

# INTRODUÇÃO

## *Cacau & Economia Nacional*

### **Estado Pará**

- *Maior produtor de cacau do Brasil, contribuindo com mais de 50% da produção nacional*
- **Histórico**
  - *Aumento de mais de 100% nas áreas plantadas (1990-2016)*

### **Investimentos**

- *Projeto Inova Cacau 2030: foco na sustentabilidade com investimentos que visam atingir 400 mil toneladas de produção anual de cacau até 2030*



# INTRODUÇÃO

## *Cacau & Sensoriamento*

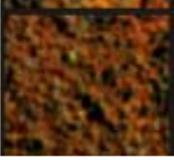
### *Remoto*

#### **Desafios**

- *Semelhança espectral do cacau cultivado em sistemas agroflorestais (com sombreamento) com floresta primária e secundária*
- *Predominância de cobertura de nuvens em regiões tropicais*

#### **Alternativas à imagens de satélite**

- *Uso de variáveis contextuais para identificação das áreas potenciais de cultivo de cacau*

Class	Planet Scope (R4G2B3)	
<b>Cocoa</b>		
		
<b>Forest</b>		
		
<b>Secondary Vegetation</b>		
		



# INTRODUÇÃO

## Cacau & Sensoriamento

### Remoto

#### Desafios

- *Semelhança espectral do cacau cultivado em sistemas agroflorestais (com sombreamento) com floresta primária e secundária*
- *Predominância de cobertura de nuvens em regiões tropicais*

#### Alternativas à imagens de satélite

- *Uso de variáveis contextuais para identificação das áreas potenciais de*

*cultivo de cacau  
distância à estradas...*

*concentração de imóveis...*

*uso e cobertura do solo...*

Class	Planet Scope (R4G2B3)	
Cocoa		
		
Forest		
		
Secondary Vegetation		
		



# INTRODUÇÃO

## Cacau & Sensoriamento

### Remoto

#### Desafios

- *Semelhança espectral do cacau cultivado em sistemas agroflorestais (com sombreamento) com floresta primária e secundária*
- *Predominância de cobertura de nuvens em regiões tropicais*

#### Alternativas à imagens de satélite

- *Uso de variáveis contextuais para identificação das áreas potenciais de cultivo de cacau*  
*distância à estradas...*

*concentração de imóveis...*

*uso e cobertura do solo...*



Class	Planet Scope (R4G2B3)	
Cocoa		
Forest		
Secondary Vegetation		



# INTRODUÇÃO

## *Avaliação da importância das variáveis contextuais*

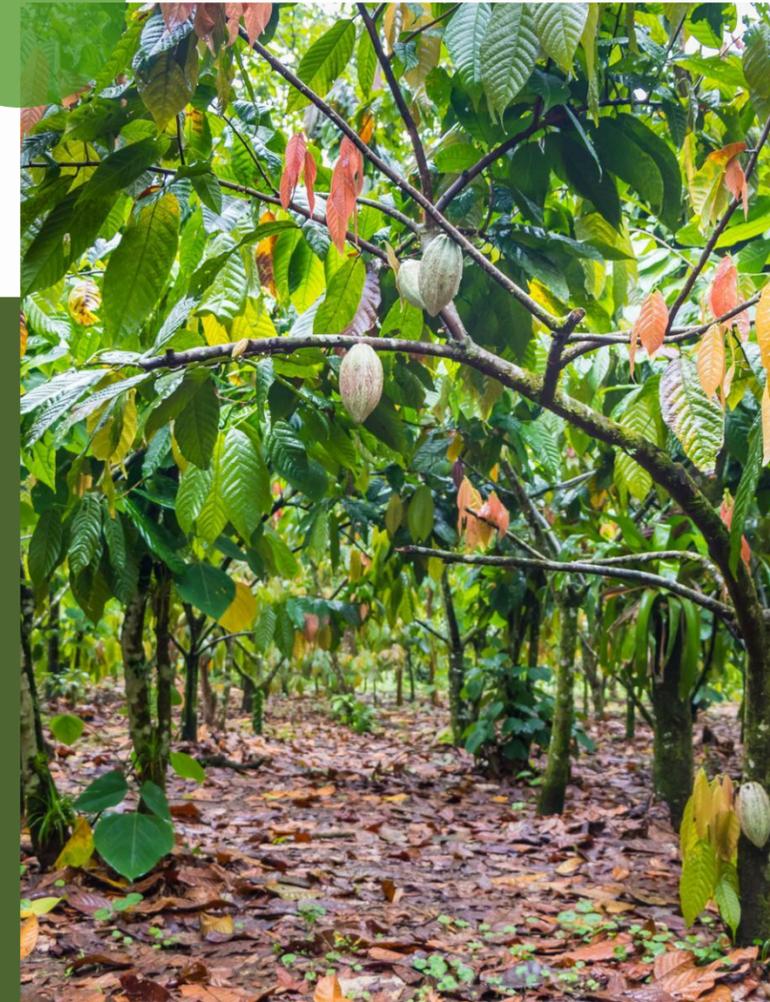
### ***Dependência Espacial & Primeira Lei da Geografia***

- *Coisas mais próximas no espaço tendem a ser mais similares (ou influenciadas) do que coisas distantes.*

*A relação entre as variáveis explicativas e o cultivo de cacau pode não ser uniforme em toda a área de estudo*

***Modelos Globais:*** *consideram a dependência espacial, porém assumem que a relação entre as variáveis é constante em toda a área de estudo.*

***Modelos Locais:*** *permitem que os coeficientes variem espacialmente, capturando diferenças locais na importância das variáveis*

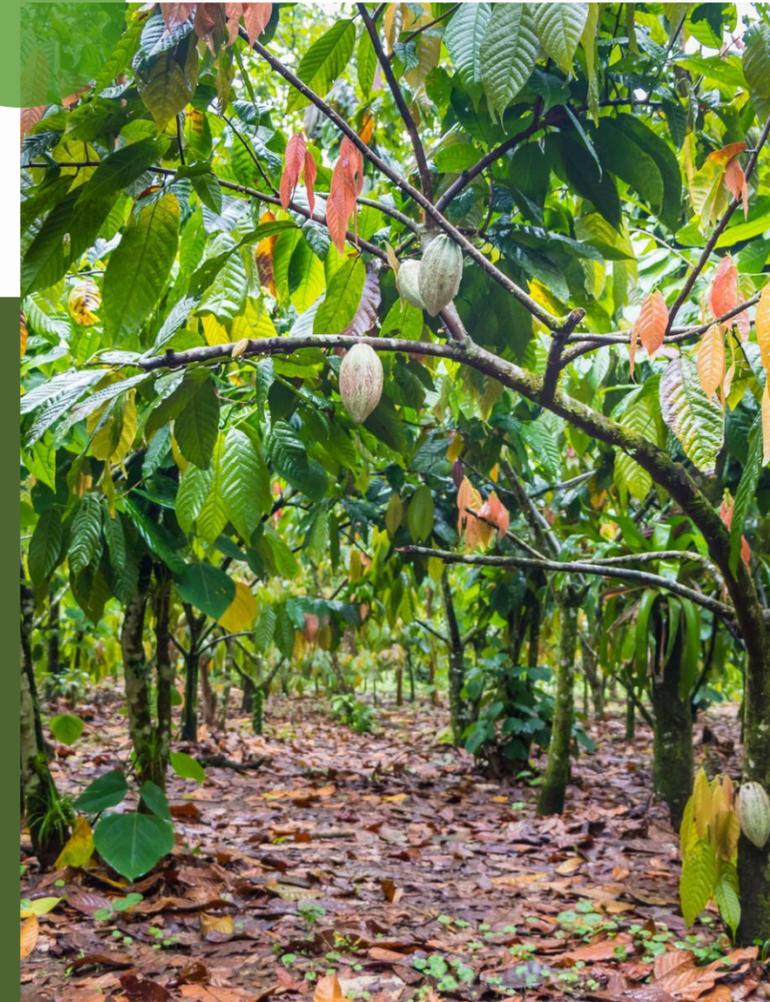


# INTRODUÇÃO

## Objetivo

*Avaliar a variação espacial da importância das variáveis geográficas relacionadas ao cultivo de cacau ao longo do Projeto de Assentamento Itatá (PA), utilizando modelos de regressão espacial.*

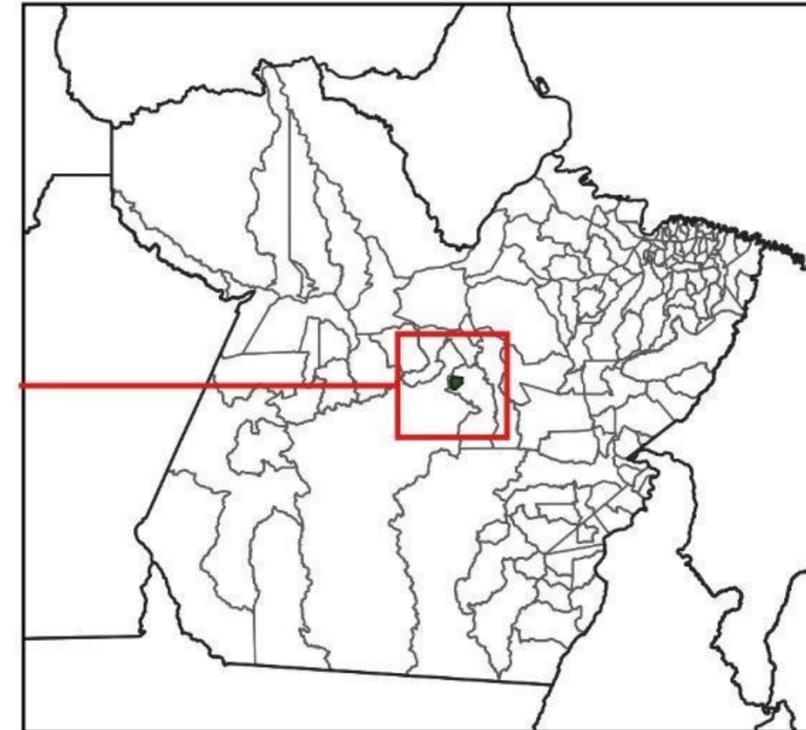
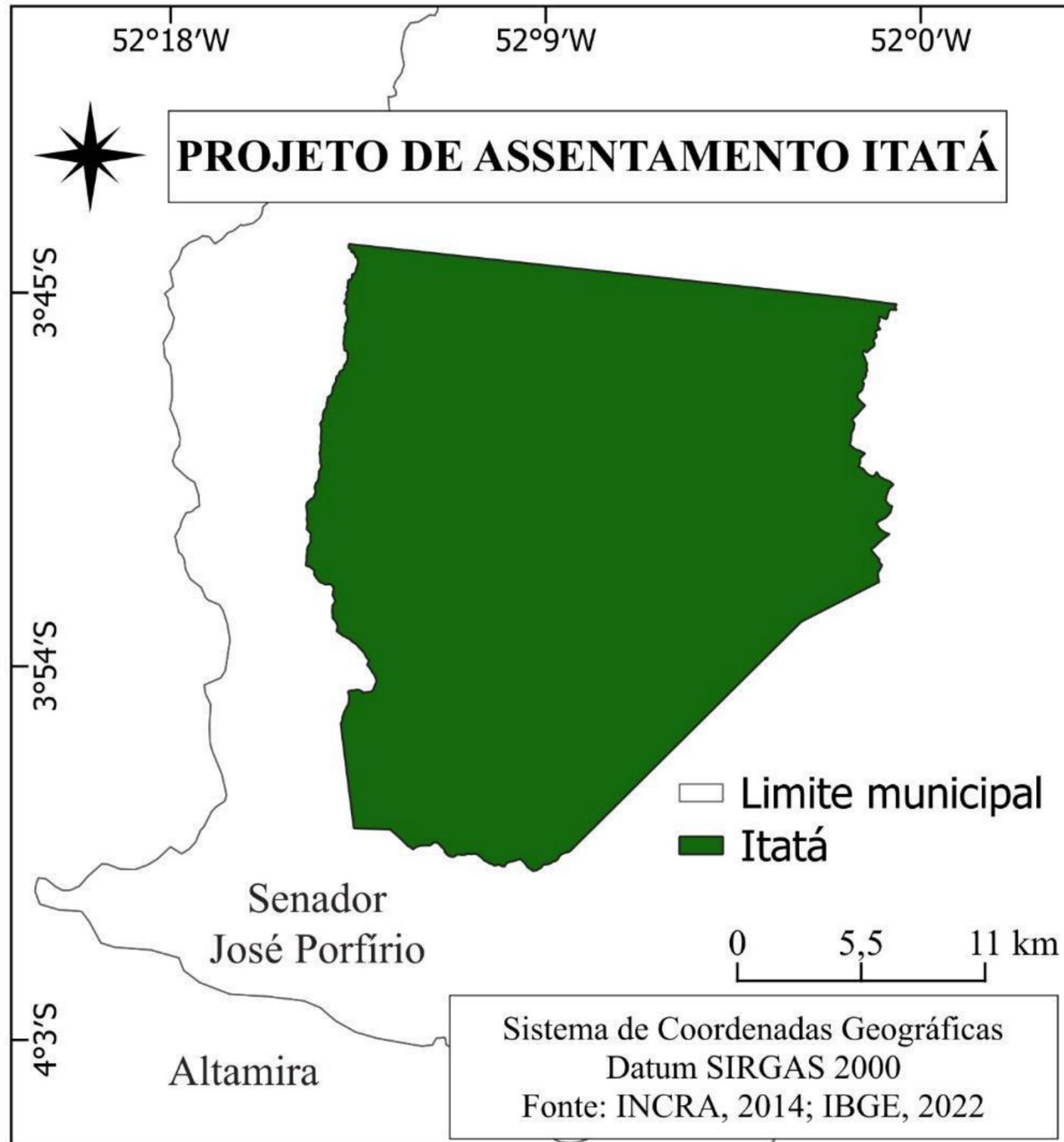
- *Comparar o desempenho do modelo global (CAR) e local (GWR) na representação da variação espacial das relações entre variáveis geográficas e o cultivo de cacau, identificando padrões globais e nuances locais ao longo da área de estudo.*





# **MATERIAL E MÉTODOS**

# ÁREA DE ESTUDO



# MATERIAL E MÉTODOS

## SELEÇÃO E PREPARAÇÃO DAS DADOS

Mapa de probabilidade da classe cacau

Variável Dependente (VD)

- Distância de estradas
- Uso e Cobertura de Imóveis
- Concentração de Imóveis
- Tamanho Imóveis
- Textura do Solo
- Declividade

Variáveis Independentes (VIs)

## VERIFICAÇÃO DEPENDÊNCIA ESPACIAL

Matriz de Pesos Espaciais

Regressão dos Mínimos Quadrados Ordinários (OLS)

Índice de Moran aplicado aos resíduos

Dependência Espacial

## DEFINIÇÃO DA MATRIZ DE VIZINHANÇA

Grids Hexagonais (de 0 a 500m, passo 50m)

Estatísticas Zonais (EZ)

VERIFICAÇÃO DEPENDÊNCIA ESPACIAL

Grid Hexagonal de 400m com EZ moda da VD e mediana das VIs

Fontes

Datasets

Informação Complementar

Etapa de Processamento

Produtos & Estatísticas

## COMPARATIVO MODELOS

Mapas dos Valores Preditos

Akaike Index Criteria (AIC)

Coefficiente de Determinação  $R^2$

SAÍDAS COMUNS

## AVALIAÇÃO ATRIBUTOS

Testes de significância T de Student

Obtenção dos Coeficientes Locais

## REGRESSÃO ESPACIAL GLOBAL

Conditional Autoregressive Regression (CAR)

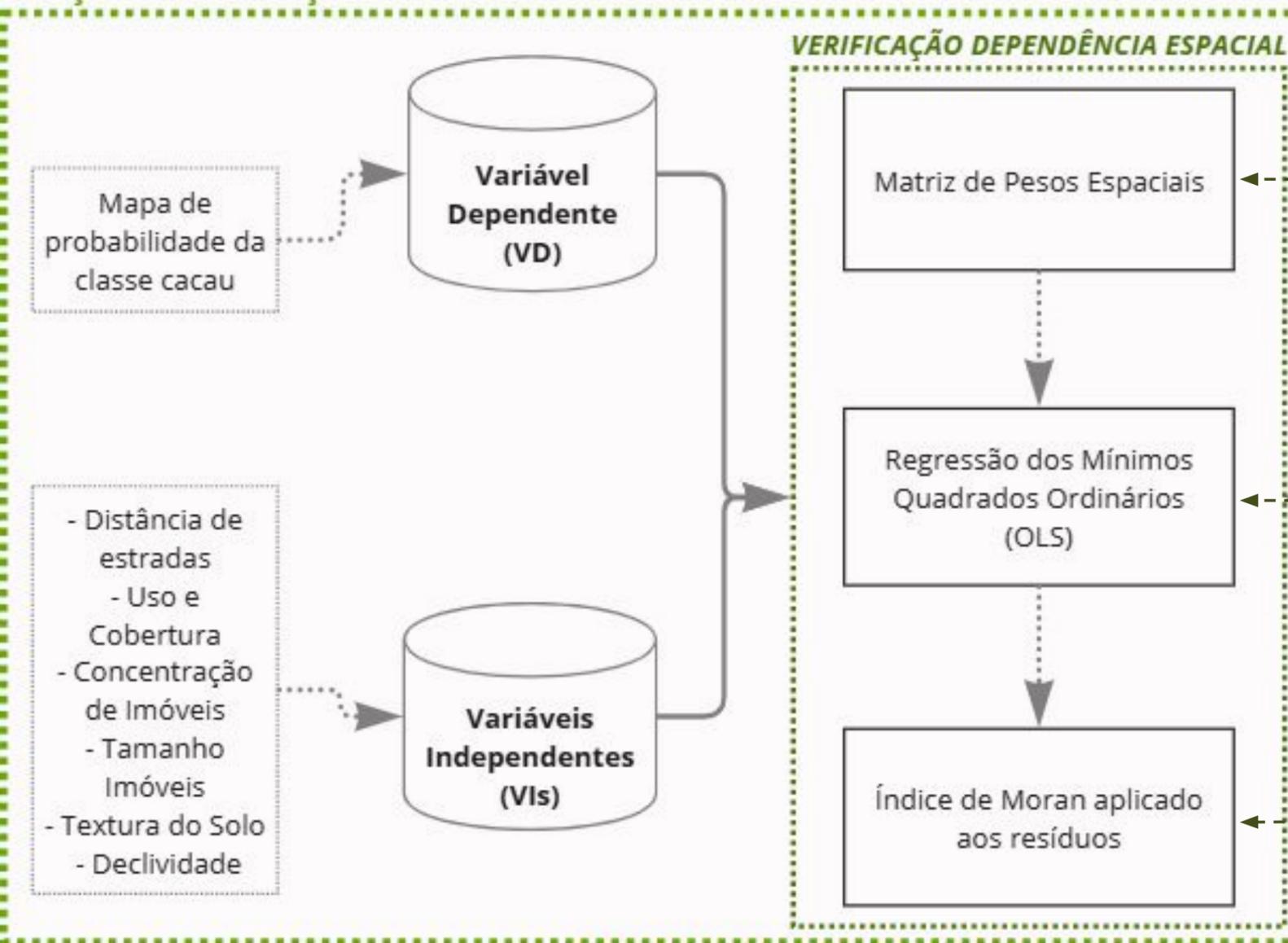
Teste Lagrange Multiplier (LM)

## REGRESSÃO ESPACIAL LOCAL

Geographically Weighted Regression (GWR)

# MATERIAL E MÉTODOS

## SELEÇÃO E PREPARAÇÃO DAS DADOS



### 1º Matriz de vizinhança:

- Critério "queen": baseado no compartilhamento de vértices ou arestas entre os polígonos.

### 2º Matriz de ponderação de contiguidade:

- Tipo: binária
- Normalização dos pesos por linha: a soma de cada linha é 1

$$Y_i = \beta_0 + \beta_1 X_i + \xi_i$$

$Y_i$  = Variável dependente (resposta).

$X_i$  = Variável independente (explicativa).

$\beta_1$  = Coeficiente angular da reta (indica a taxa de variação de Y em relação a X).

$\beta_0$  = Intercepto ou constante (valor de Y quando X = 0).

$\xi_i$  = erro ou termo residual;

$$I = \frac{e' W e}{e' e}$$

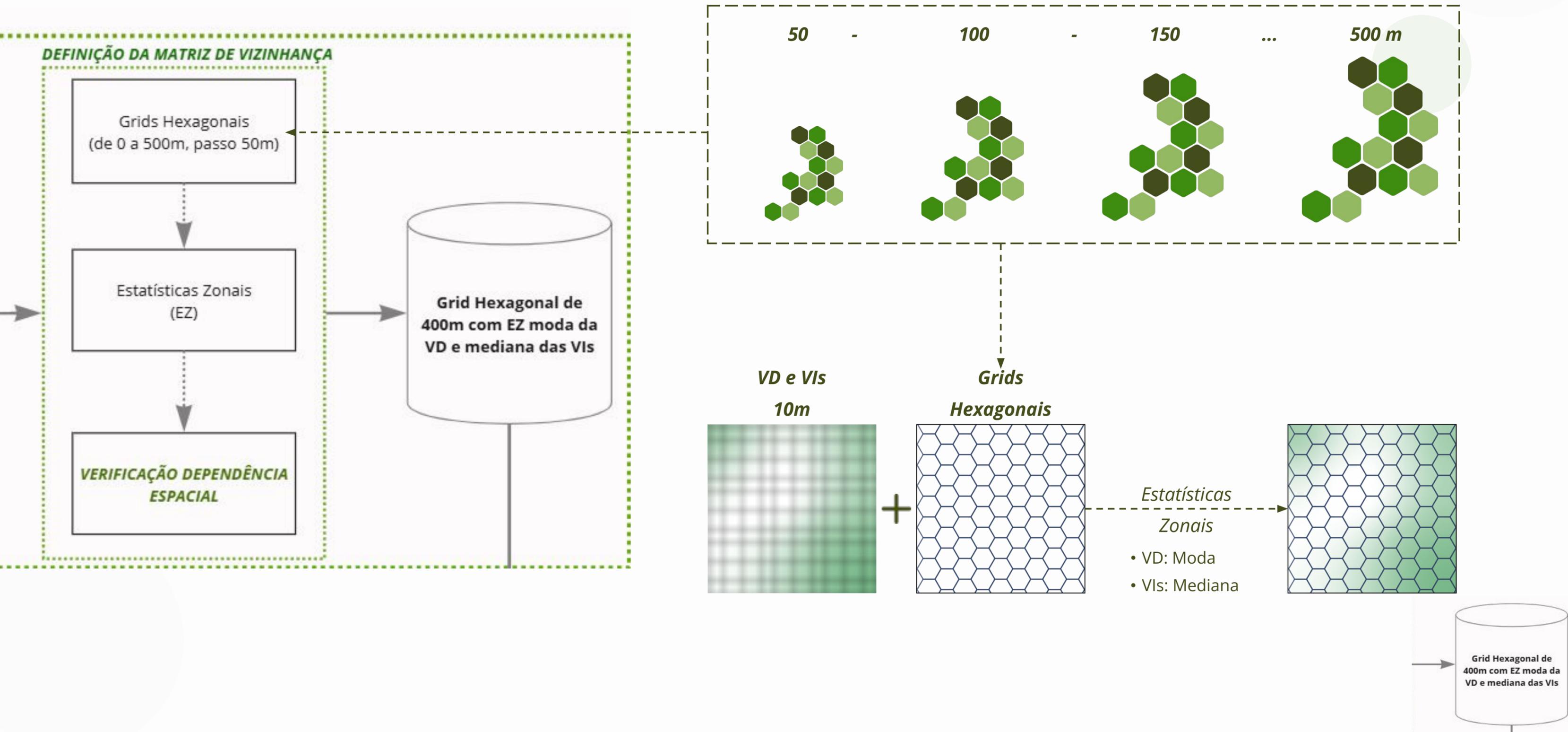
$e$  = vetor de resíduos.

$W$  = matriz de pesos espaciais.

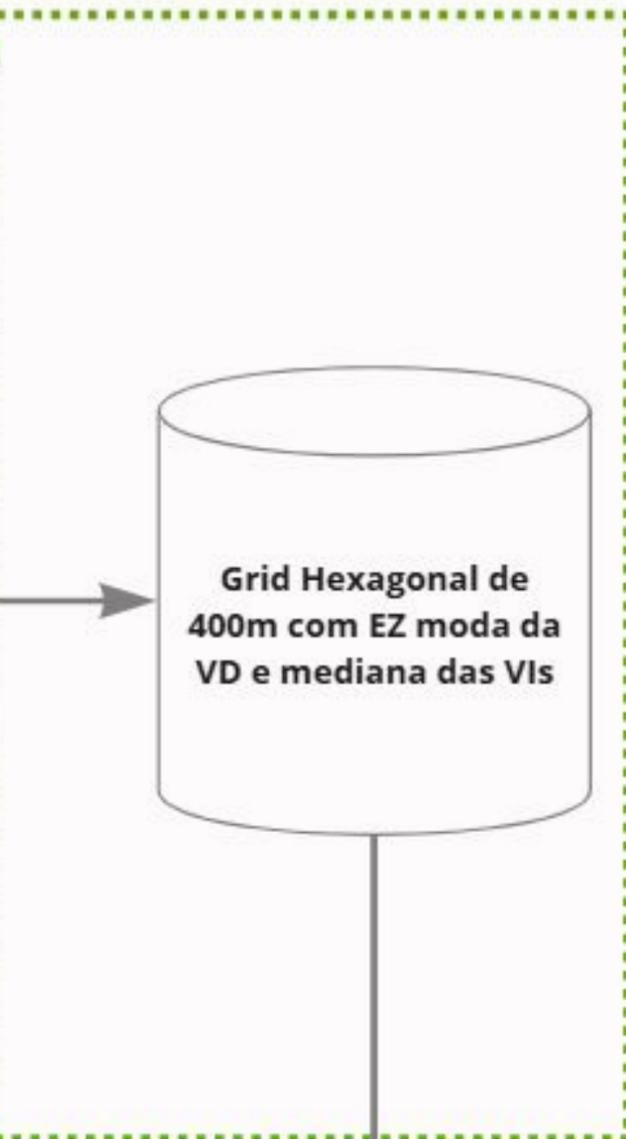
$e'$  = transposta do vetor de resíduos.



# MATERIAL E MÉTODOS



# MATERIAL E MÉTODOS



**Variável Resposta**.....  
Probabilidade de cultivo  
de cacau

**Fatores Condicionantes**



Uso e Cobertura do Solo.....

Texturas de Solo.....

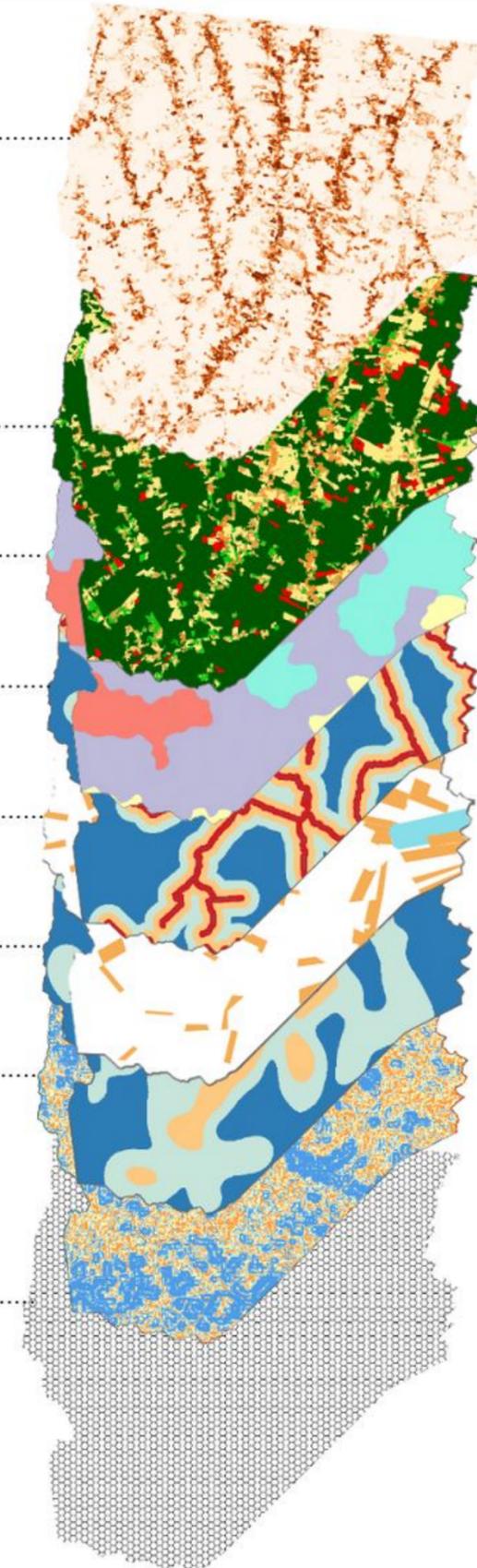
Distância de Estradas (m).....

Tamanho Imóveis.....

Concentração Imóveis.....

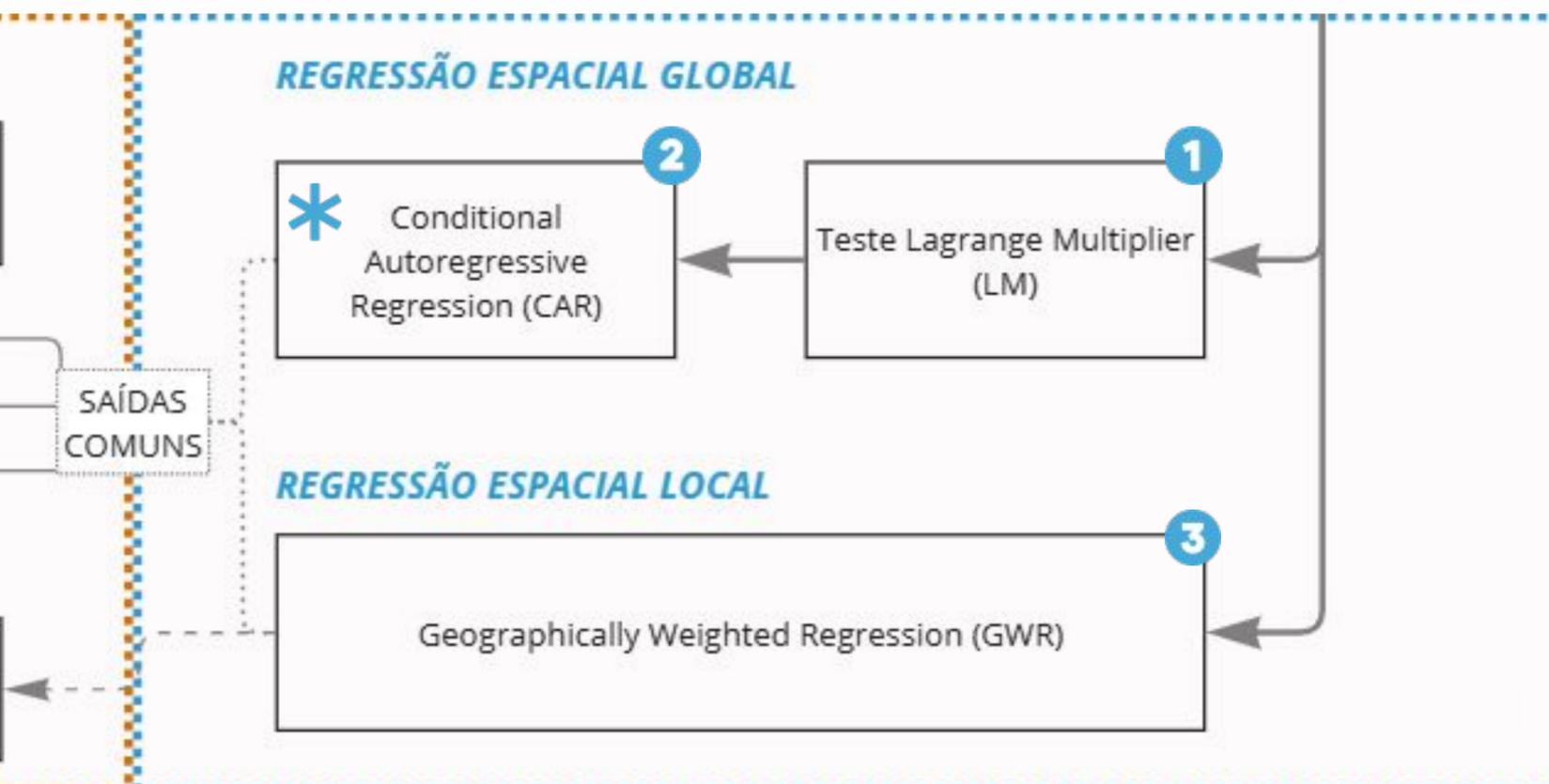
Declividade (%).....

**Grade (400m)**.....



# MATERIAL E MÉTODOS

\* Remoção de variáveis com p-valor não significativo



## 1 Testes de Lagrange Multiplier (LM)

### Simple

- **LMerr**: para dependência do erro
- **LMlag**: para uma variável dependente espacialmente defasada

### Robusto

- **RLMerr**: para dependência do erro
- **RLMlag**: para uma variável dependente espacialmente defasada

## 2 Modelo de Erro Espacial (CAR)

$$Y = X\beta + \varepsilon$$

$Y$  = Variável dependente (resposta).  
 $X$  = Matriz de variáveis explicativas.  
 $\beta$  = Vetor de Coeficientes  
 $\varepsilon$  = erro ou termo residual;

$$\varepsilon = \rho W\varepsilon + \xi$$

- $\rho$  = Parâmetro de correlação espacial.
- $W$  = Matriz de pesos espaciais (relaciona as observações vizinhas).
- $\xi$  = Componente aleatório do erro, com variância constante e não correlacionada.

## 3 Geographically Weighted Regression (GWR)

$$Y_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)X_{1i} + \beta_2(u_i, v_i)X_{2i} + \dots + \beta_k(u_i, v_i)X_{ki} + \varepsilon_i$$

$Y_i$  = Variável dependente (resposta).  
 $X_i$  = Valores das variáveis explicativas.  
 $\beta$  = Vetor de Coeficientes  
 $\varepsilon$  = erro ou termo residual;

$\beta_0(u_i, v_i), \beta_1(u_i, v_i), \dots, \beta_k(u_i, v_i)$  = coeficientes espaciais (ou locais) que variam de acordo com a localização geográfica da  $i$ -ésima observação

A notação  $(u_i, v_i)$  representa as coordenadas espaciais (por exemplo, latitude e longitude) da  $i$ -ésima observação

# MATERIAL E MÉTODOS

## \* Mapas dos valores preditos

Permite comparar visualmente a acurácia e a capacidade de cada modelo em capturar os padrões espaciais da variável analisada

### 1 Akaike Information Criterion (AIC)

- Qualidade de diferentes modelos ajustados.
  - Penaliza a complexidade para evitar overfitting.
- Ajuste/Simplicidade

### 2 Coeficiente de Determinação $R^2$

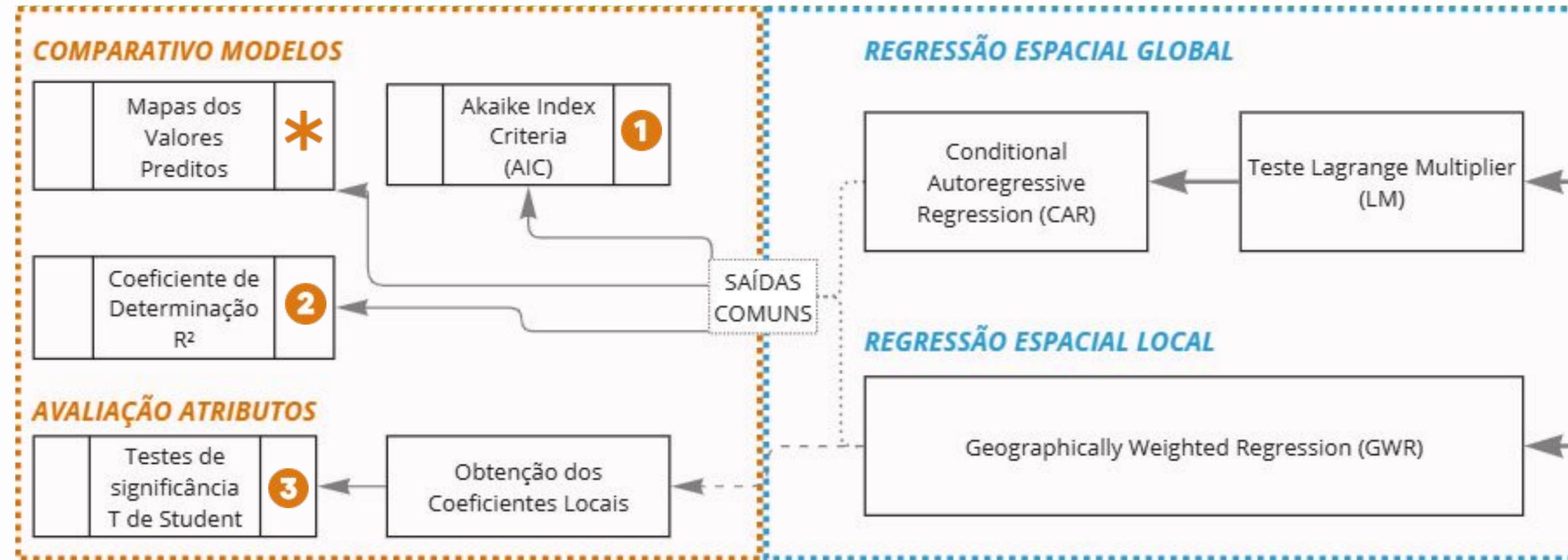
Quanto da variabilidade de Y é explicada pelo modelo.

### 3 Testes de significância T de Student

Verifica se os coeficientes  $\beta$  são significativamente diferentes de zero.

$$t^* = \frac{\hat{\beta}_1}{S(\hat{\beta}_1)}$$

$S(\beta_1)$  = erro-padrão da estimativa de  $\beta_1$ .





# RESULTADOS

# RESULTADOS

## TESTES DE LAGRANGE MULTIPLIER

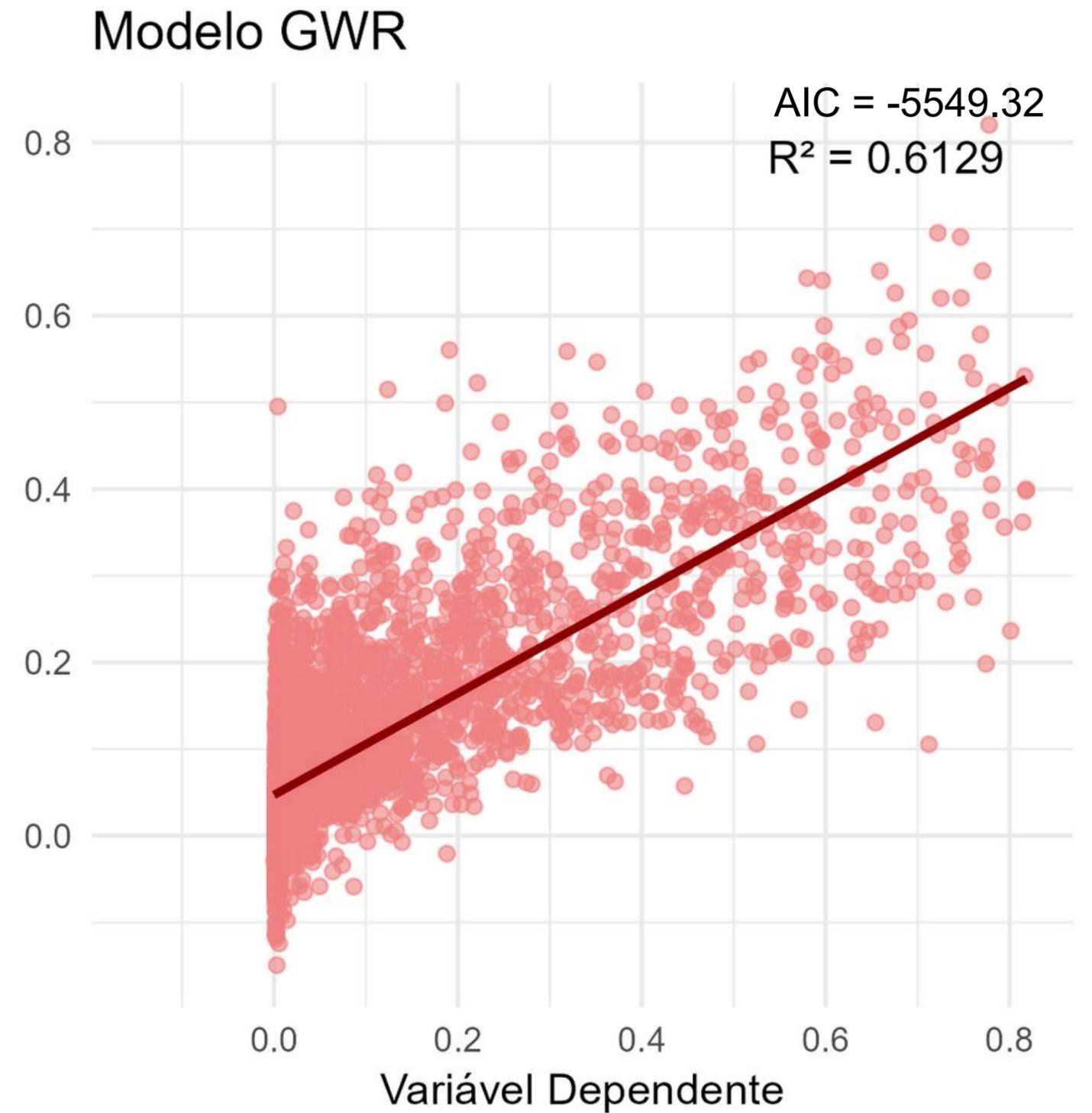
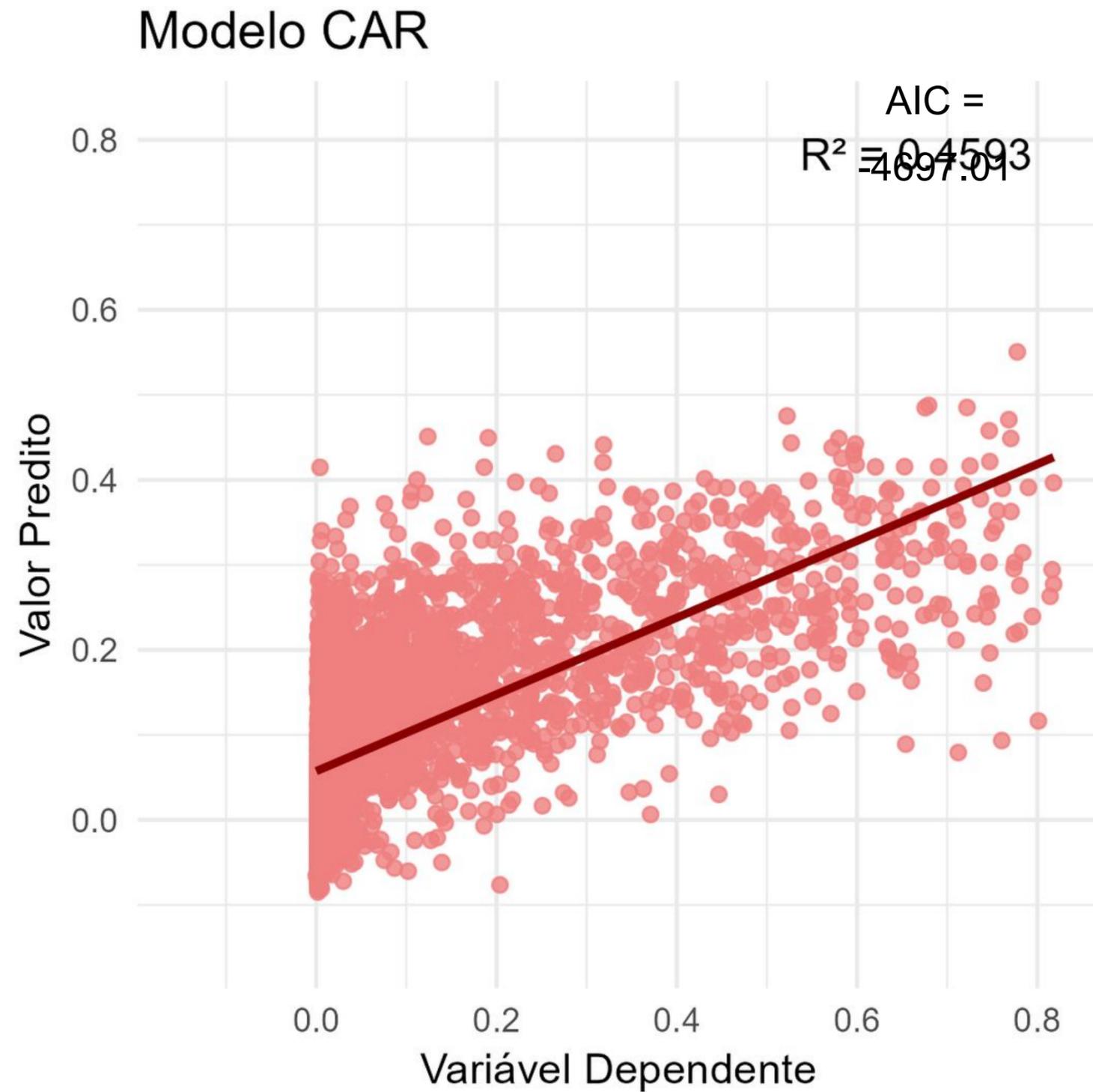
Teste	Estatística	p-valor
RSerr (LMerr)	484.64	< 2.2e-16
RSlag (LMlag)	373.42	< 2.2e-16
<b>adjRSerr (RLMerr)</b>	<b>119.53</b>	<b>&lt; 2.2e-16</b>
adjRSlag (RLMlag)	8.3038	0.003956

←----- SIGNIFICATIVO!

FOI ESCOLHIDO O  
MODELO MODELO  
GLOBAL CAR

# RESULTADOS

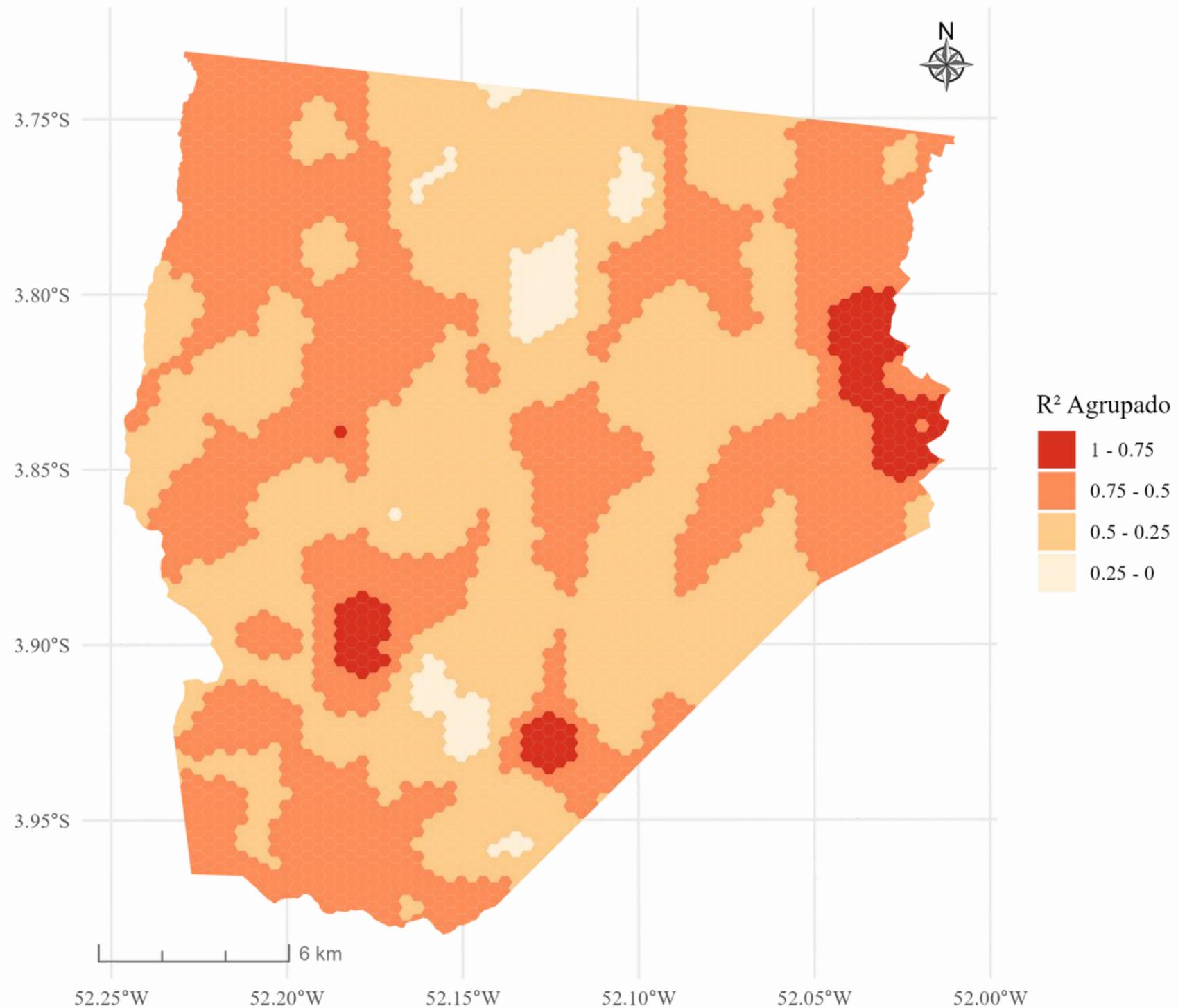
## COEFICIENTE DE DETERMINAÇÃO $R^2$



# RESULTADOS

## COEFICIENTE DE DETERMINAÇÃO $R^2$

$R^2$  =  
**0.6129**

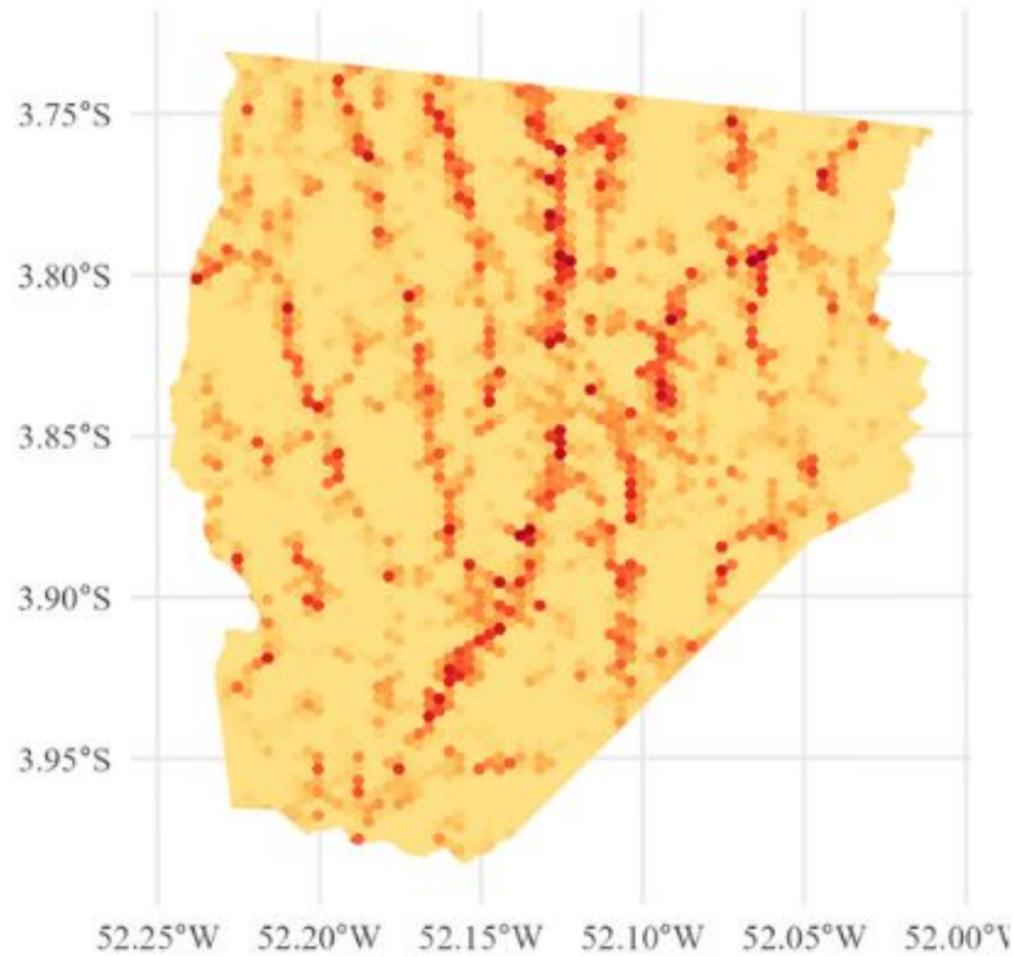


***R<sup>2</sup> DO MODELO  
GWR VARIA AO  
LONGO DA ÁREA***

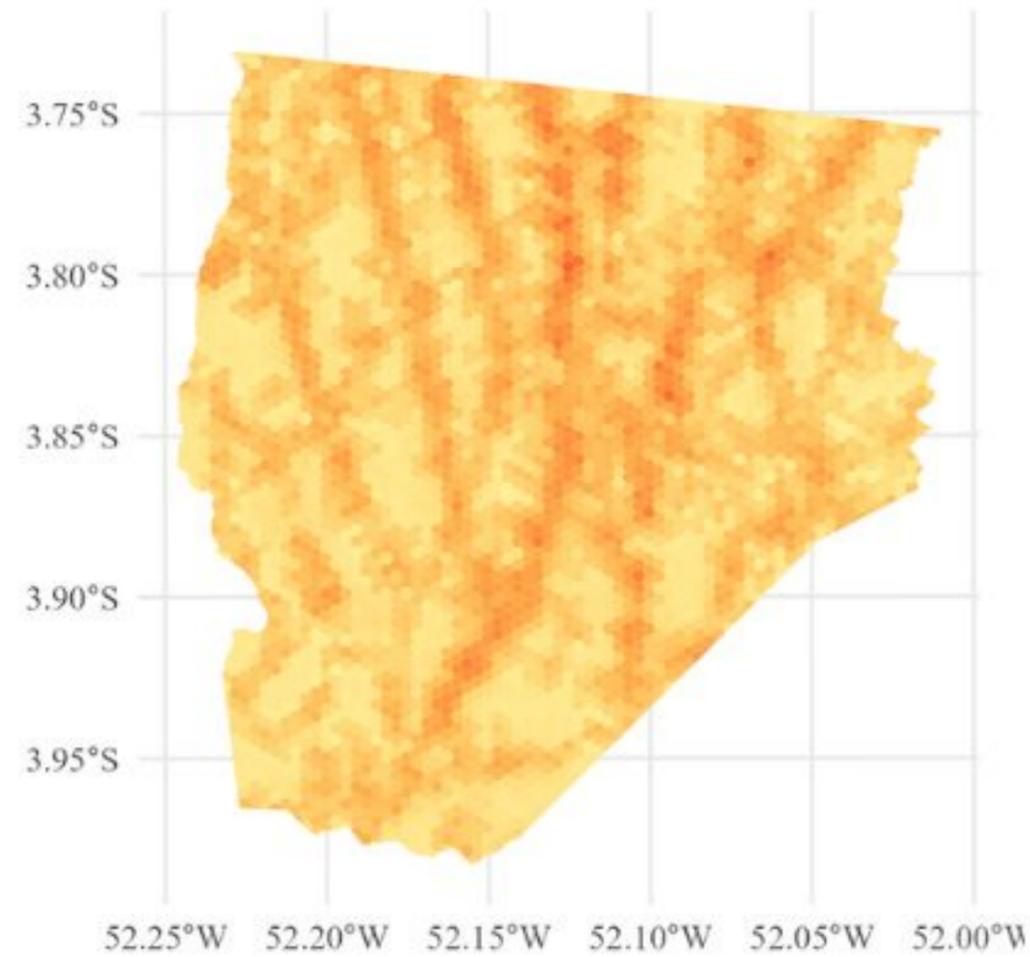
# RESULTADOS

## MAPAS DOS VALORES PREDITOS

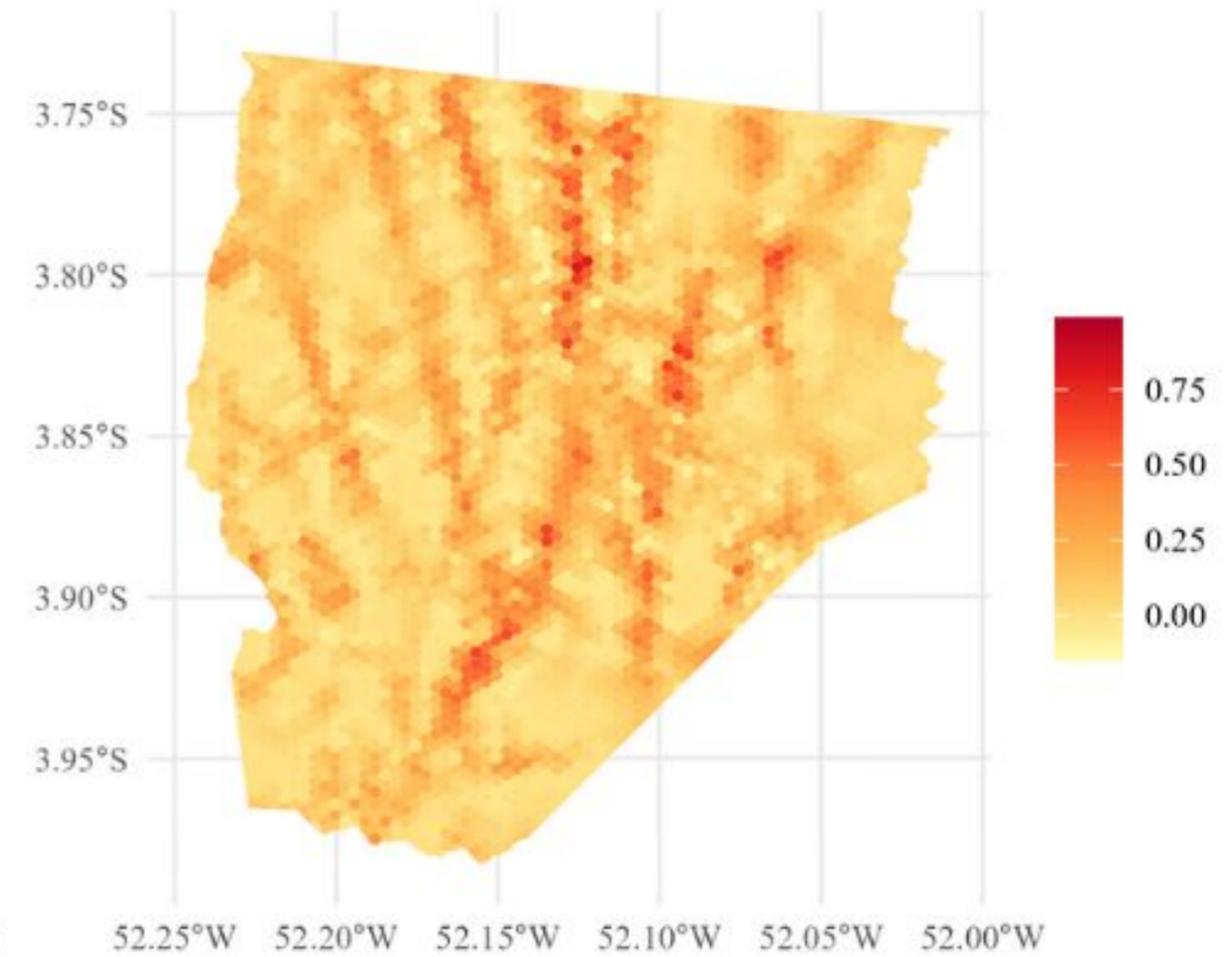
Valor Observado (Variável Resposta)



Valor Predito pelo Modelo Global



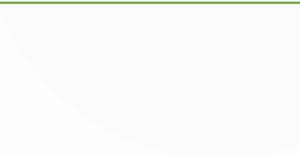
Valor Predito pelo Modelo Local



# RESULTADOS

## COEFICIENTES LOCAIS E TESTES T DE STUDENT

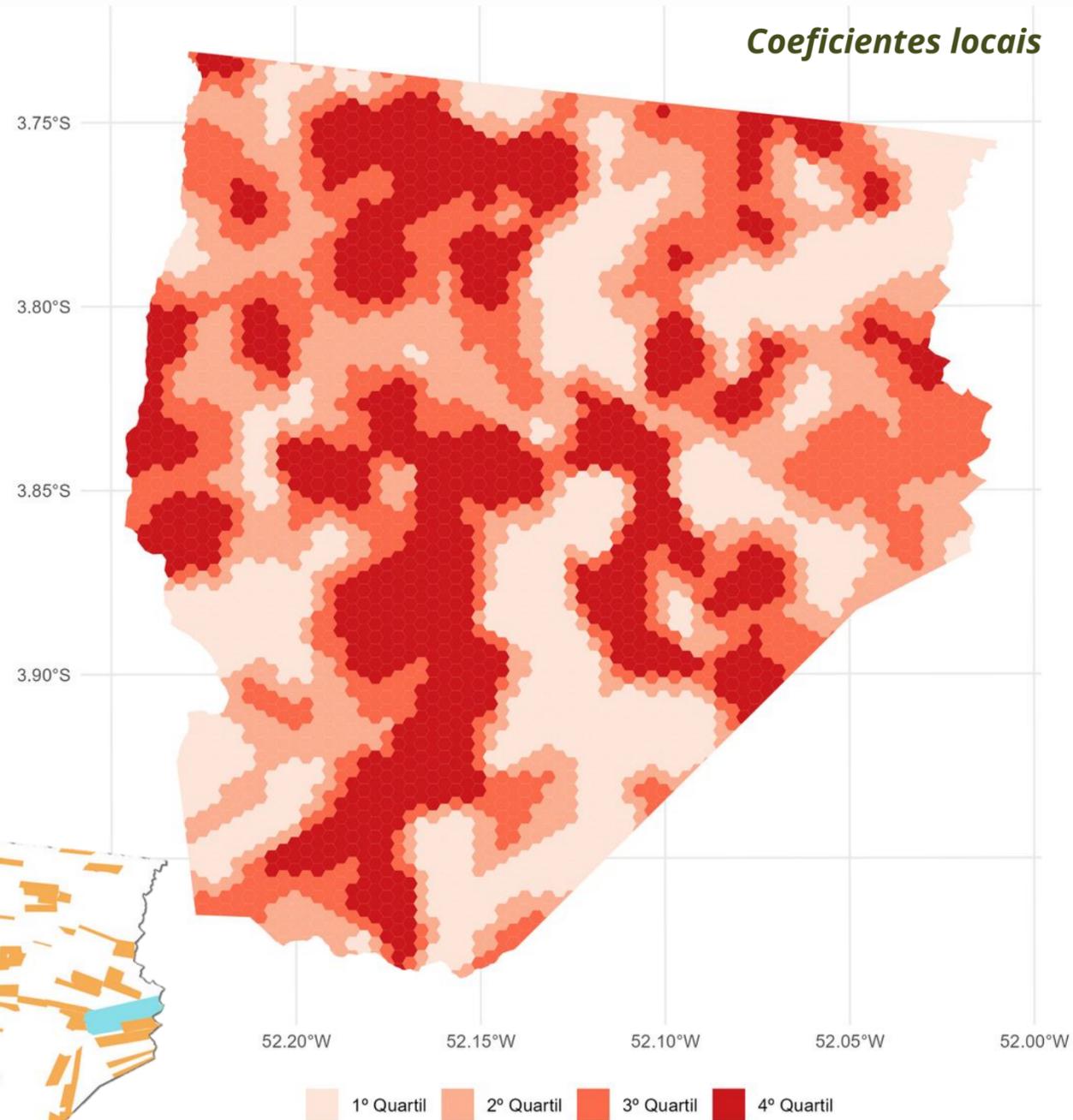
### Quadro Geral



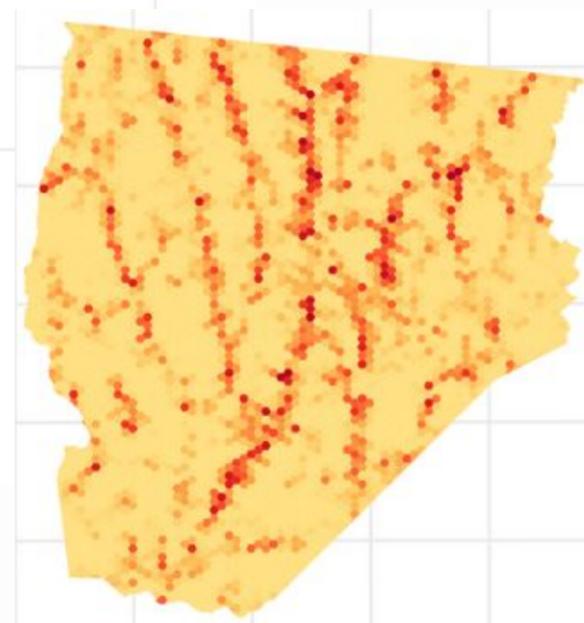
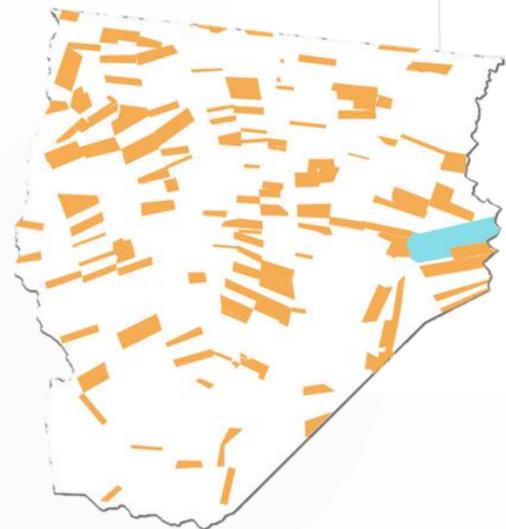
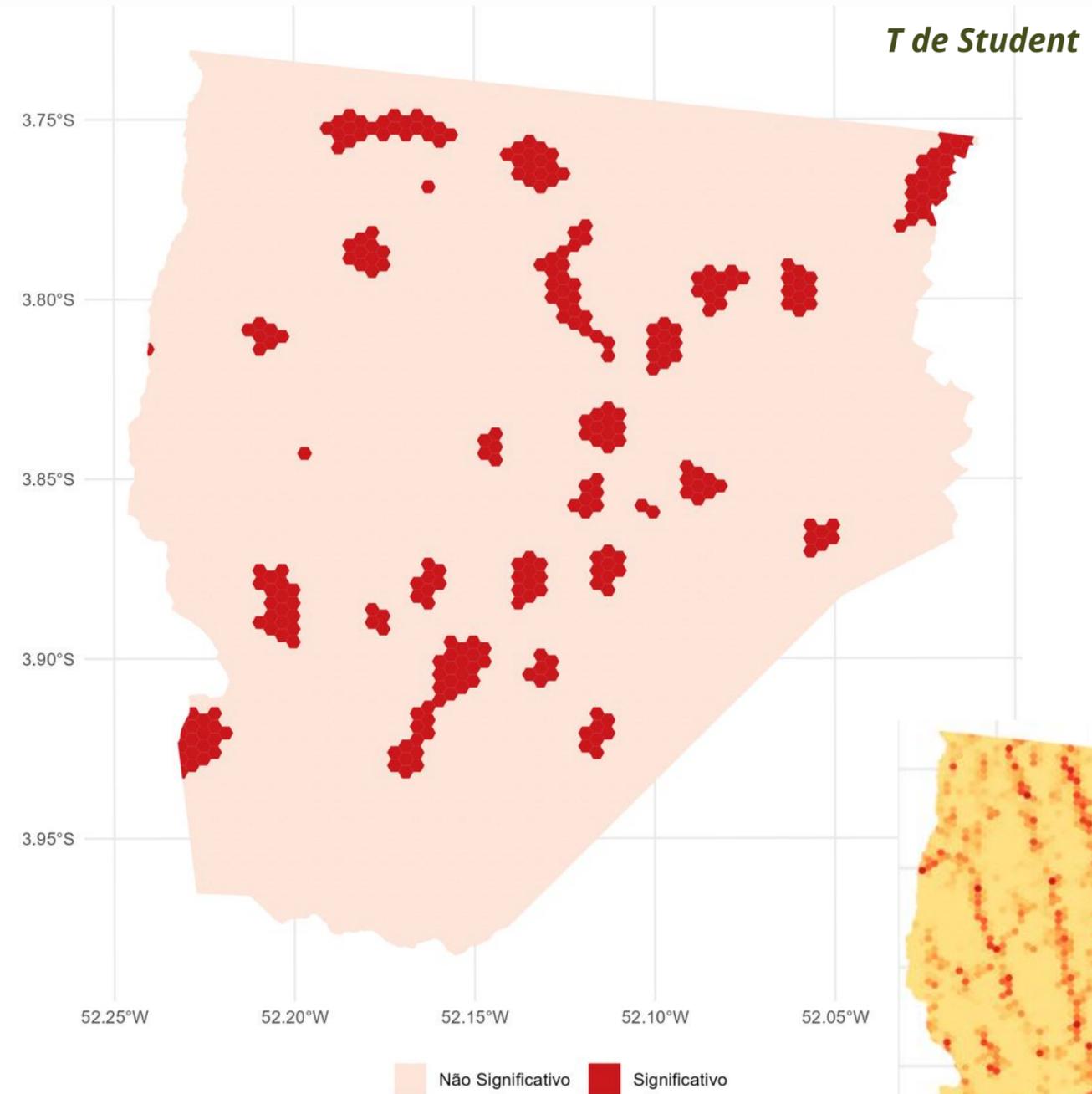
Variável	R2 Agrupado	Estimativa		Erro Padrão		Z-valor		P-valor	
		GWR	CAR	GWR	CAR	GWR	CAR	GWR	CAR
<b>Concentração Imóveis</b>	0-0.25	-0.0030		0.0570		-0.0526		9.5807e-01	
	0.25-0.5	0.0123		0.0498		0.2467		8.0517e-01	
	0.5-0.75	0.0072	0.0260	0.0450	0.0052	0.1602	5.0187	8.7273e-01	<b>5.2000e-07</b>
	0.75-1	-0.0011		0.0401		-0.0267		9.7866e-01	
<b>Uso e Cobertura do Solo</b>	0-0.25	0.0440		0.0323		1.3648		1.7231e-01	
	0.25-0.5	0.1098		0.0356		3.0874		2.0189e-03	
	0.5-0.75	0.1746	0.1408	0.0347	0.0048	5.0345	29.4323	<b>4.7900e-07</b>	<b>0.0000e+00</b>
	0.75-1	0.2501		0.0343		7.2959		<b>2.9700e-13</b>	
<b>Textura do Solo</b>	0-0.25	-0.0122		0.0484		-0.2519		8.0115e-01	
	0.25-0.5	-0.0115		0.0331		-0.3476		7.2816e-01	
	0.5-0.75	-0.0012	0.0088	0.0282	0.0040	-0.0441	2.2151	9.6485e-01	<b>2.6751e-02</b>
	0.75-1	-0.0439		0.0276		-1.5889		1.1208e-01	
<b>Tamanho Imóveis</b>	0-0.25	0.0141		0.0274		0.5170		6.0513e-01	
	0.25-0.5	-0.0010		0.0355		-0.0278		9.7786e-01	
	0.5-0.75	0.0050	0.0060	0.0342	0.0048	0.1467	1.2451	8.8340e-01	2.1308e-01
	0.75-1	-0.0202		0.0356		-0.5665		5.7107e-01	
<b>Distância de Estradas</b>	0-0.25	0.0079		0.0225		0.3502		7.2618e-01	
	0.25-0.5	0.0372		0.0290		1.2833		1.9940e-01	
	0.5-0.75	0.0521	0.0727	0.0312	0.0122	1.6678	5.9593	<b>9.5352e-02</b>	<b>2.5300e-09</b>
	0.75-1	0.0811		0.0357		2.2741		<b>2.2962e-02</b>	

### Tamanho de Imóveis

*Coefficientes locais*

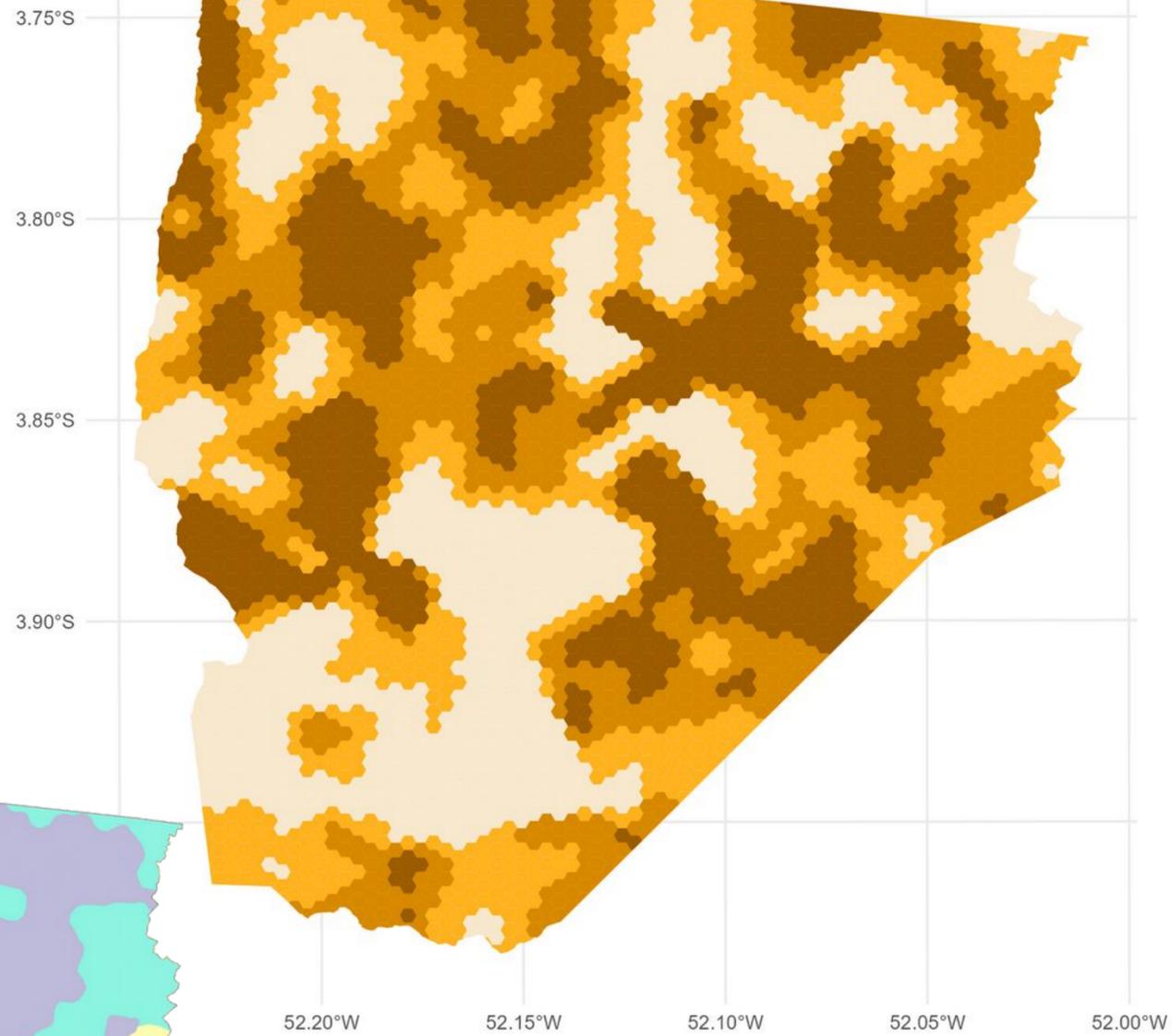


*T de Student*



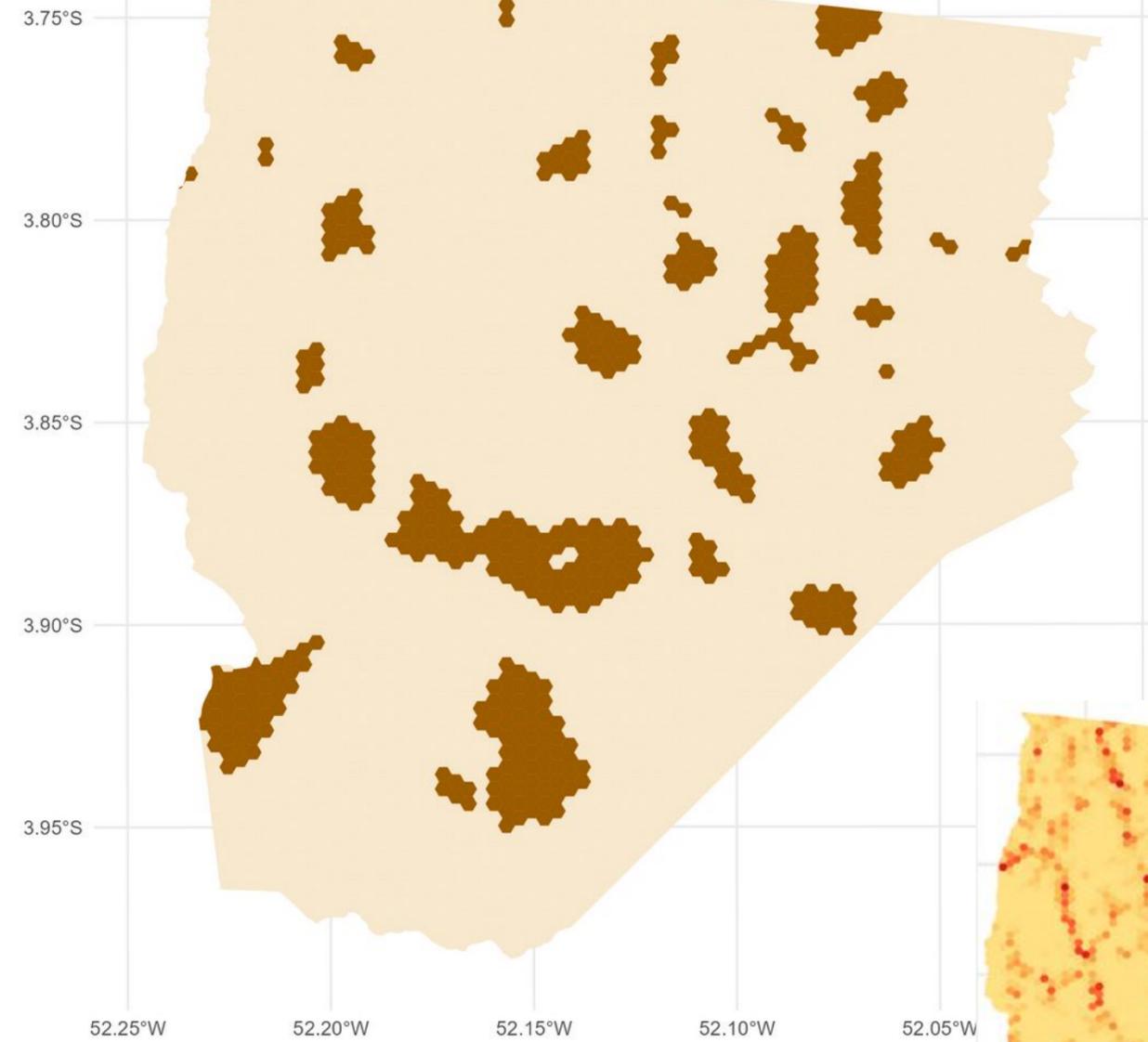
### Textura do Solo

*Coefficientes locais*

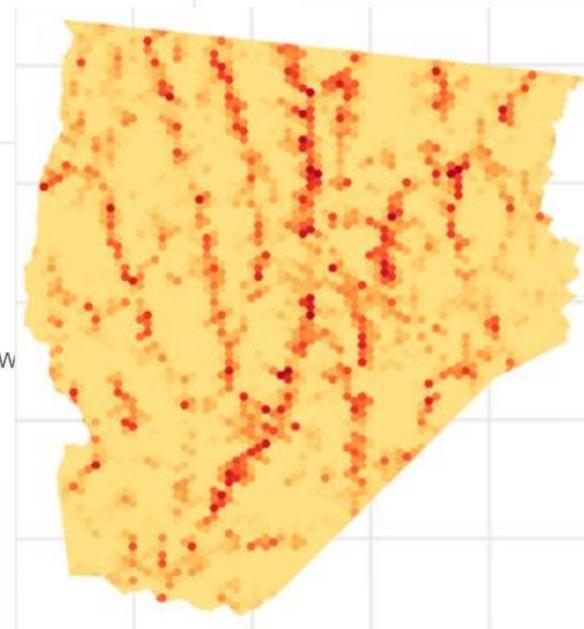
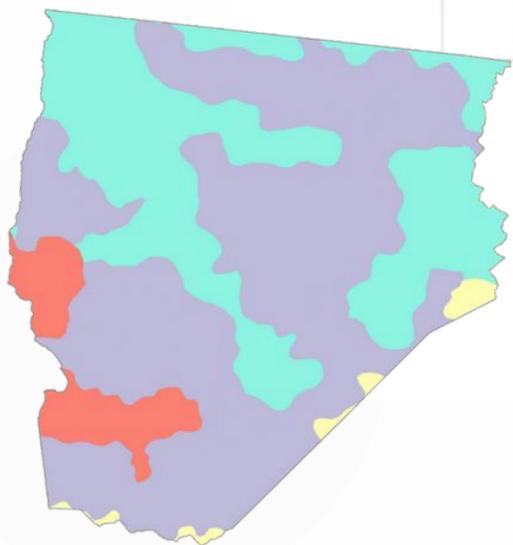


1° Quartil 2° Quartil 3° Quartil 4° Quartil

*T de Student*

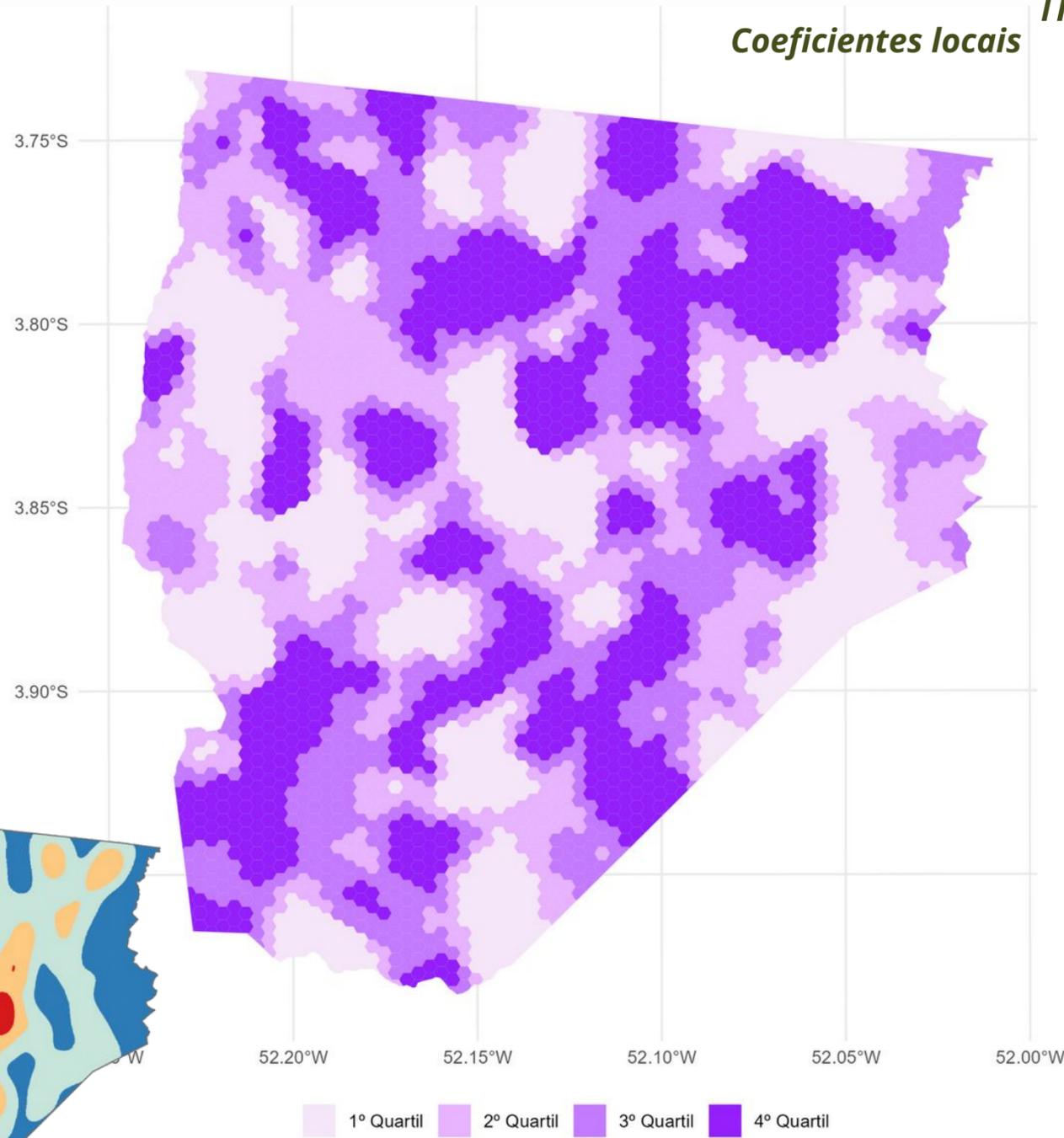


Não Significativo Significativo

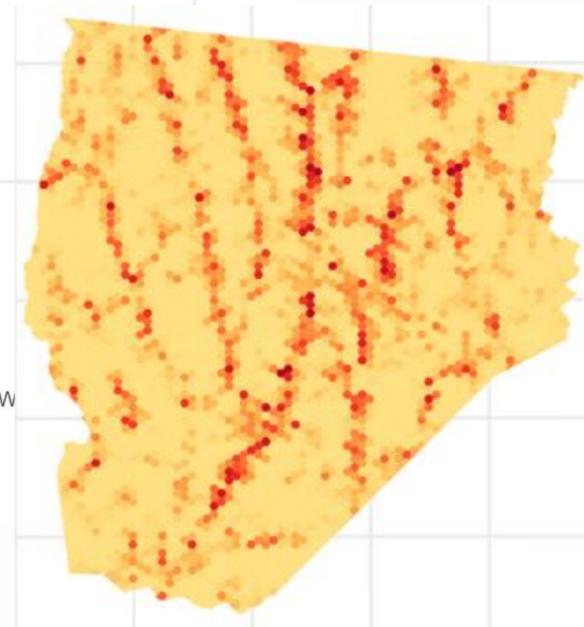
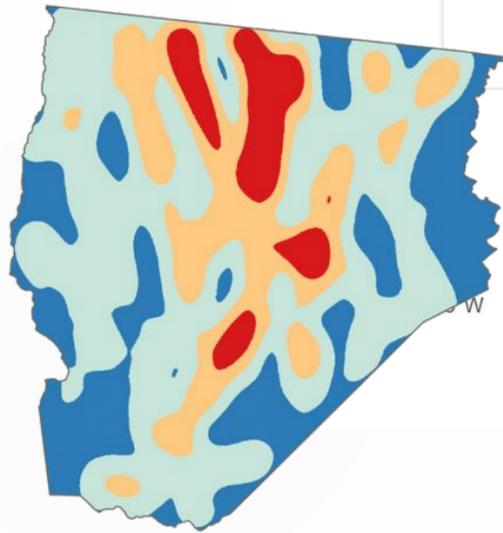
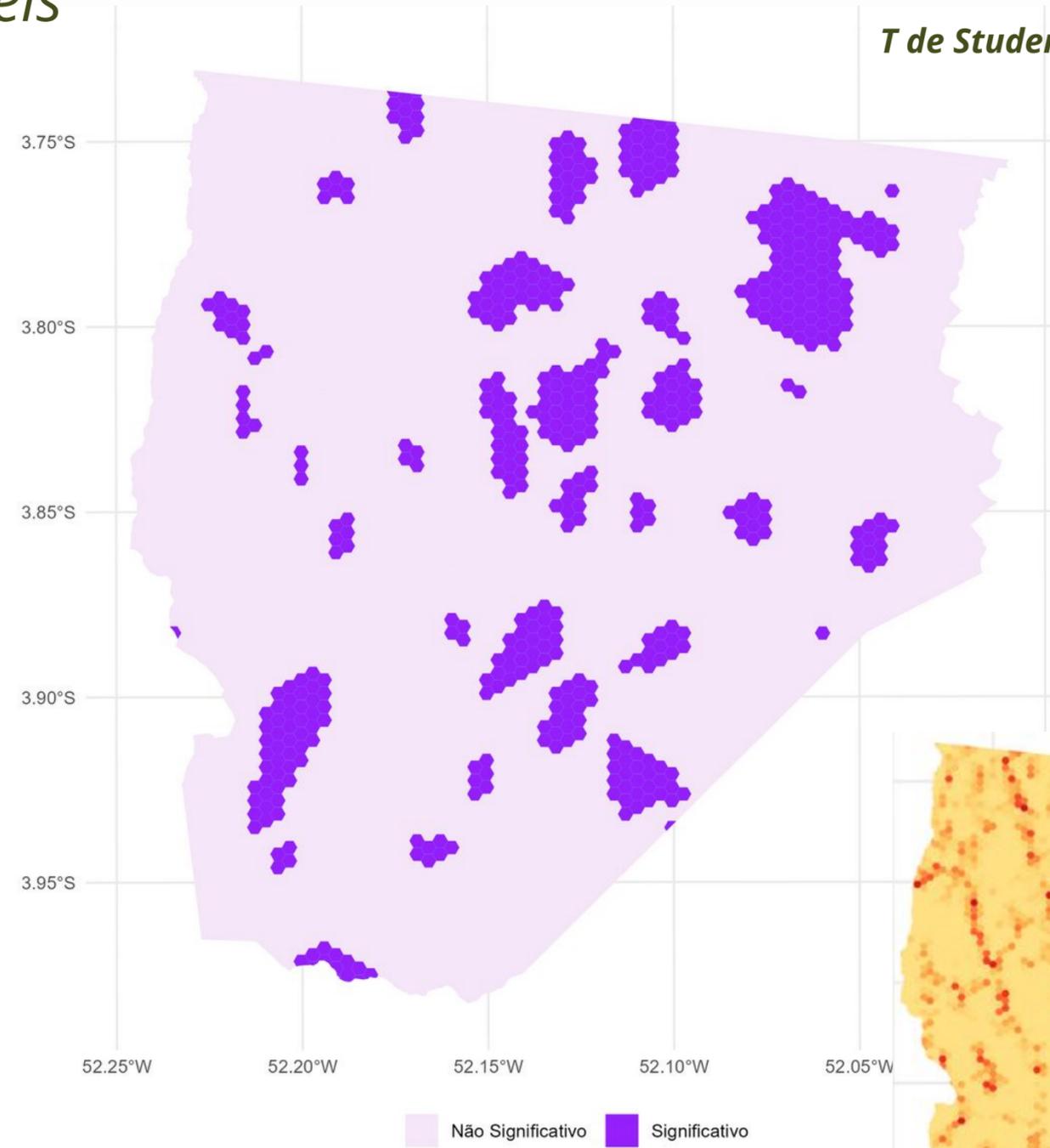


### Concentração de Imóveis

*Coefficientes locais*

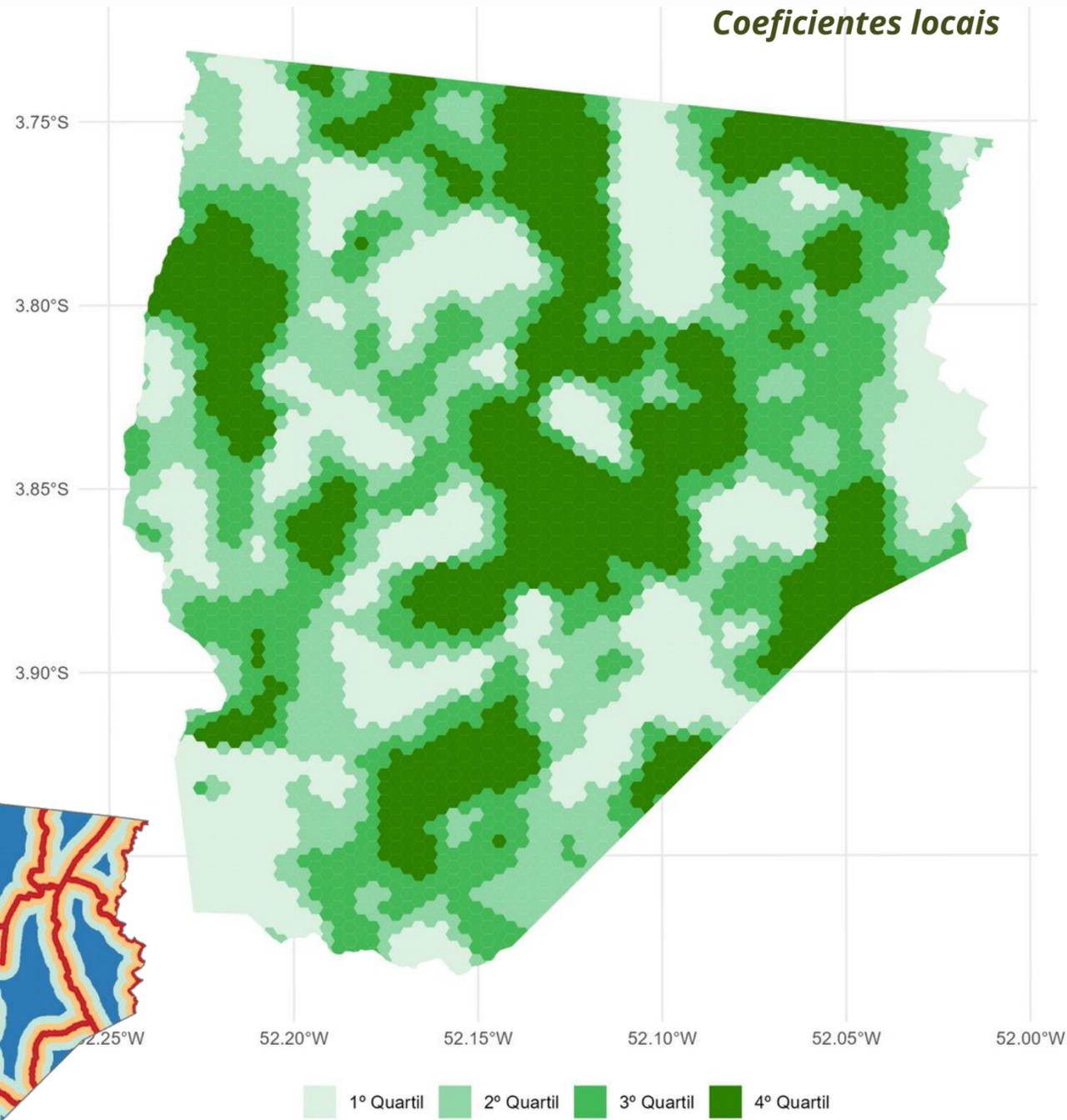


*T de Student*

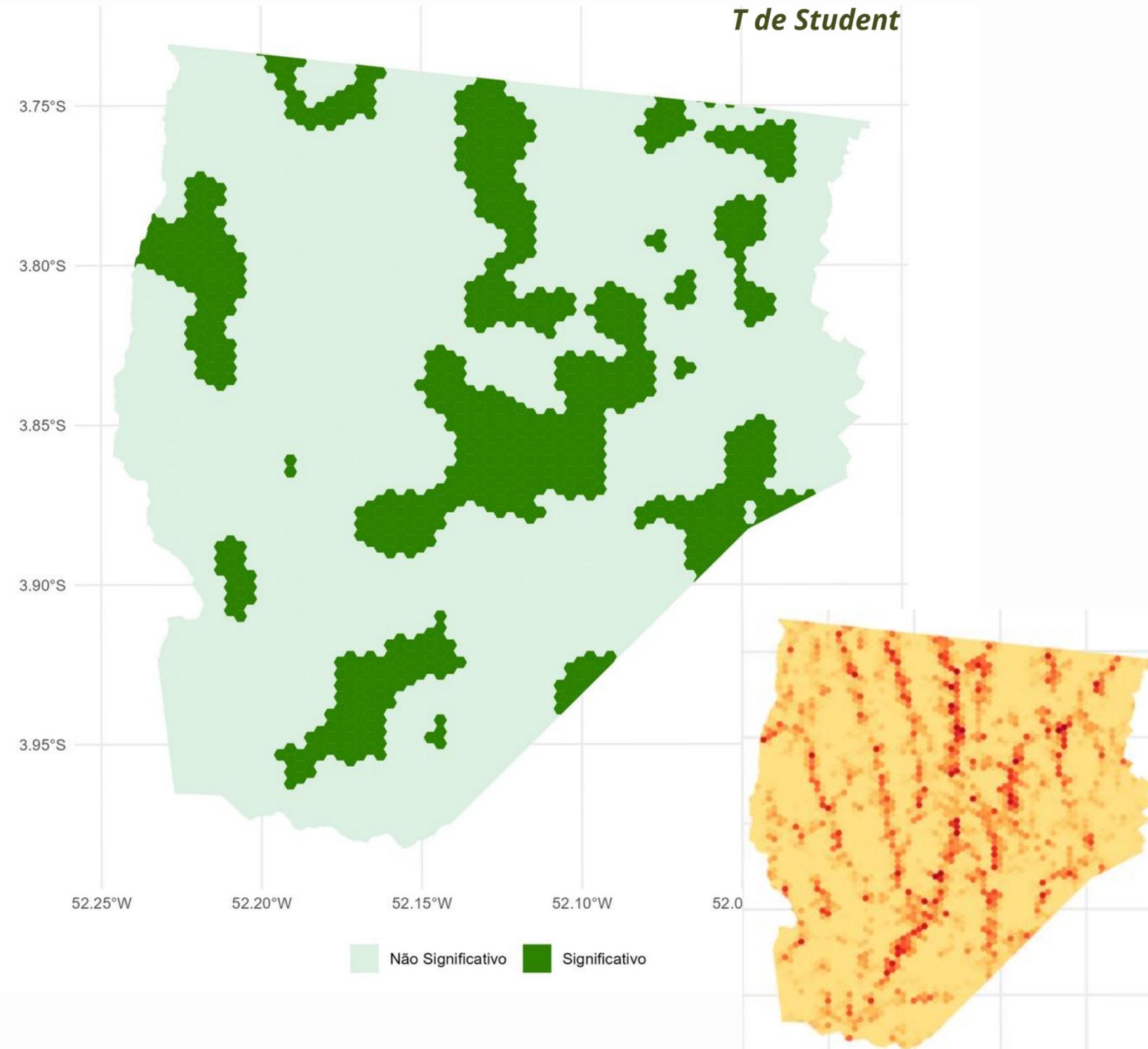


### Distância a Estradas

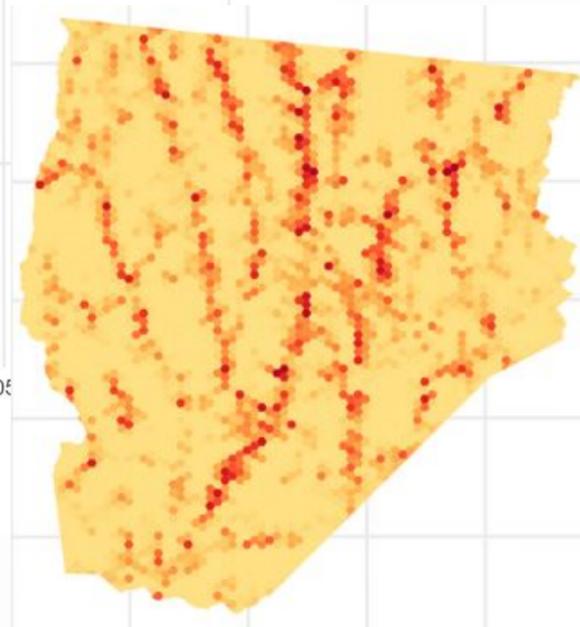
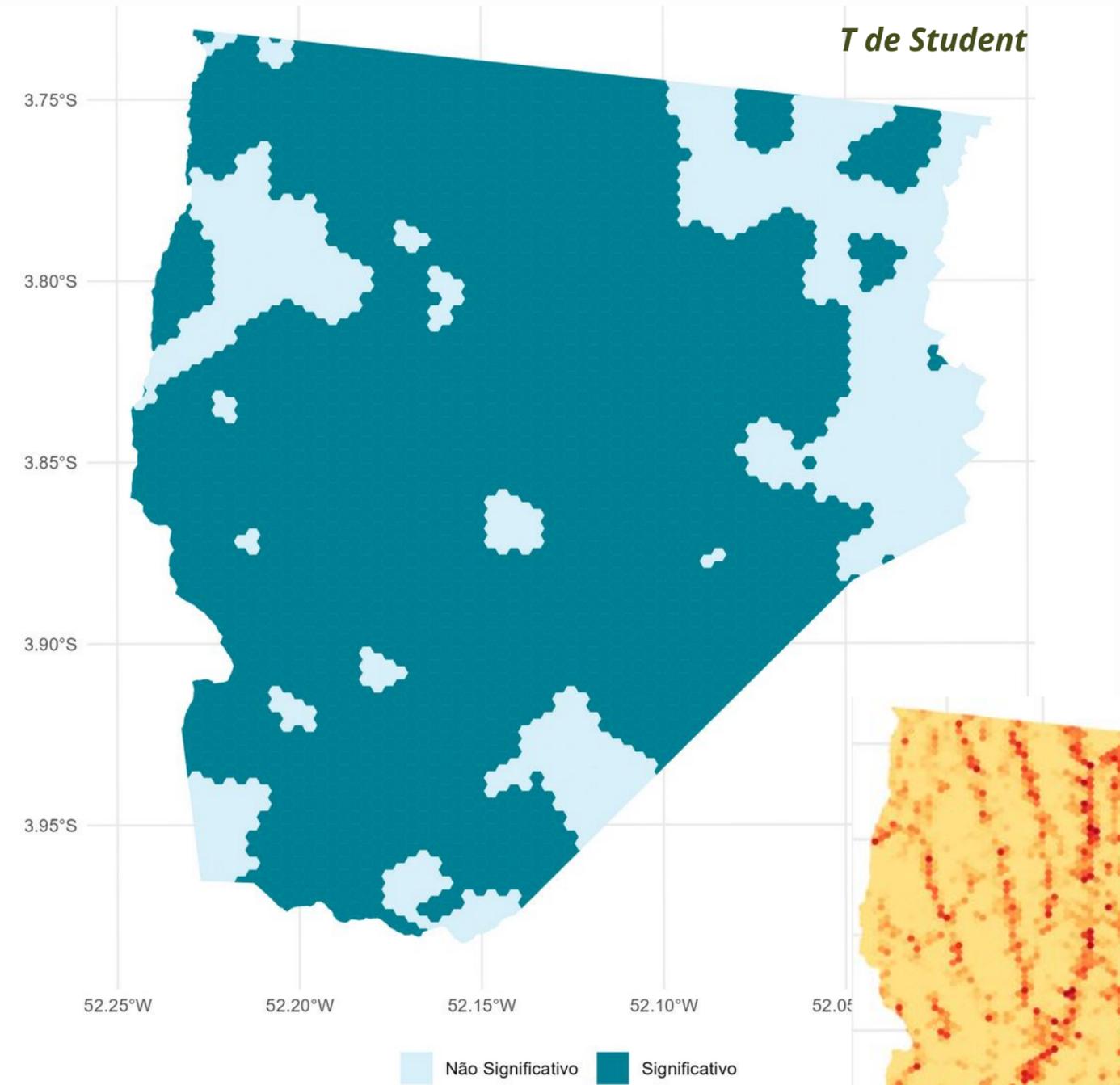
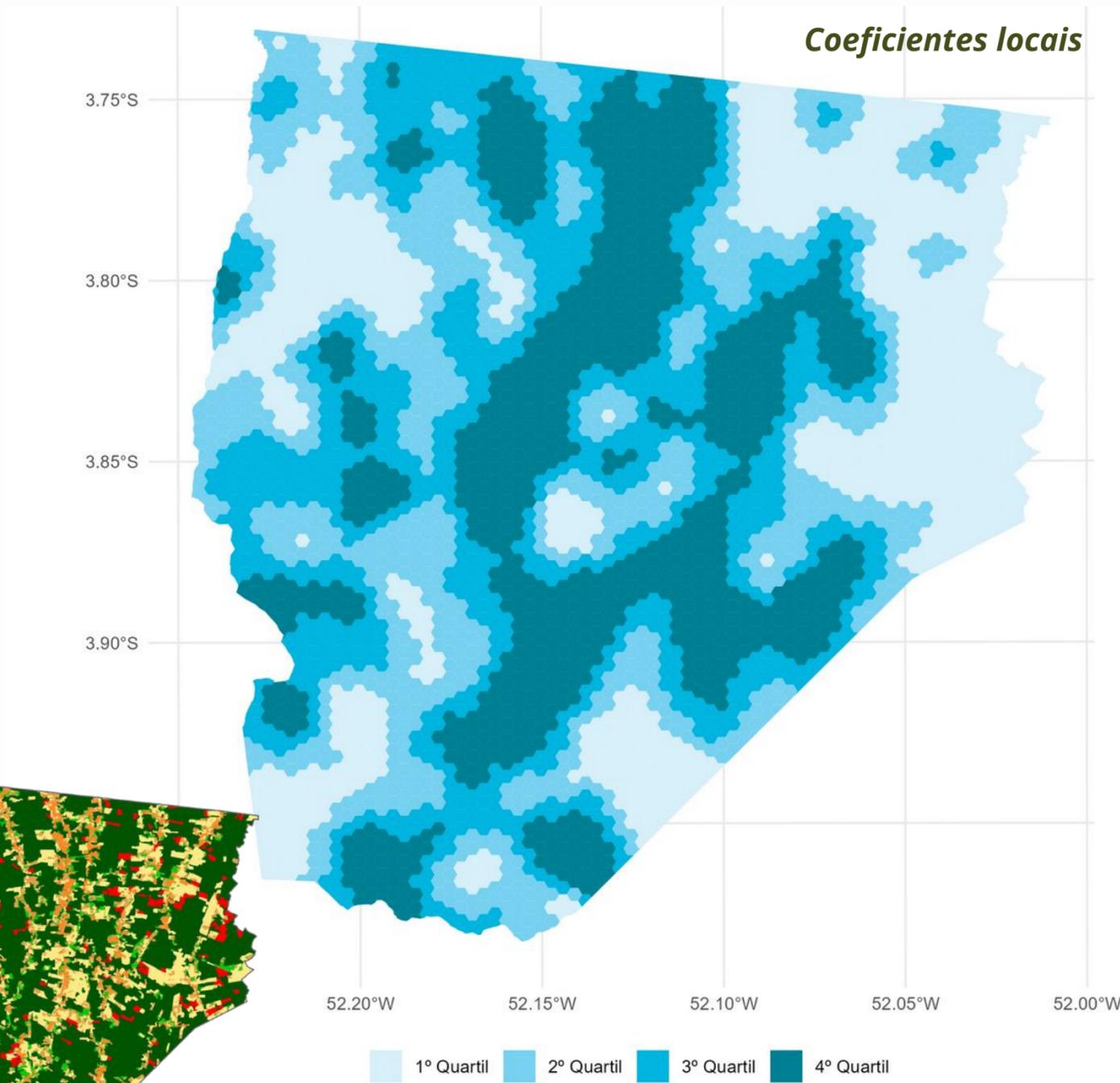
*Coefficientes locais*



*T de Student*



### Uso e Cobertura do Solo



# CONCLUSÃO

# CONCLUSÃO

## MODELOS

- *O modelo local (GWR) apresentou melhor ajuste ( $R^2$  e AIC), evidenciando maior capacidade de captar variações espaciais.*
- *A diferença no ajuste pode ser explicada pela adaptação do GWR às nuances locais, minimizando o impacto de outliers.*
- *O GWR detectou padrões espaciais heterogêneos, enquanto o modelo global simplificou as relações.*

## VARIÁVEIS

- *O modelo global (CAR) identificou mais variáveis significativas, já que assume uma relação uniforme para todas as áreas.*
- *As variáveis 'Uso e Cobertura do Solo' e 'Distância de Estradas' são as mais significativas de acordo com o modelo local*

