An aerial photograph of a dry, hilly landscape. The terrain is brown and yellow, indicating a lack of vegetation. A winding, dry riverbed or gully runs through the center of the image, following the contours of the hills. The hills are rounded and have a similar brownish color. The overall scene is arid and desolate.

**Modelos de regressão espacial
aplicados à investigação de variáveis
relacionadas aos processos de
mudança de uso e cobertura da terra
no Alto Uruguai (RS e SC)**

Marcos Wellausen Dias de Freitas

Trabalho Final
Disciplina: Análise Espacial
Dezembro/2009

Apresentação

- Problemática contemporânea: identificação de processos de mudanças de uso e cobertura da terra e seus impactos
- Objetivo do trabalho: análise de variáveis sociais e físicas associadas aos processos de mudança com vistas ao uso em modelos de dinâmica da paisagem em área do Alto Uruguai (RS/SC)
- Área de estudo: 18 municípios (206 setores censitários rurais), referência é Lages (SC), área total de aprox. 15.000km²

Questão e hipóteses

- Que variáveis e tipos de variáveis apresentam relação com os processos de MUCT na área em questão?
- Hipóteses
 - Importância do uso integrado de variáveis físicas e sociais em estudos de MUCT e outros que envolvem sistemas complexos
 - Uso de modelos de regressão espacial permitem a incorporação da problemática da autocorrelação espacial existente neste tipo de estudo (Overmars et al., 2003, Soares-Filho et al., 2008)

Modelos de Regressão

- Modelo de regressão linear:

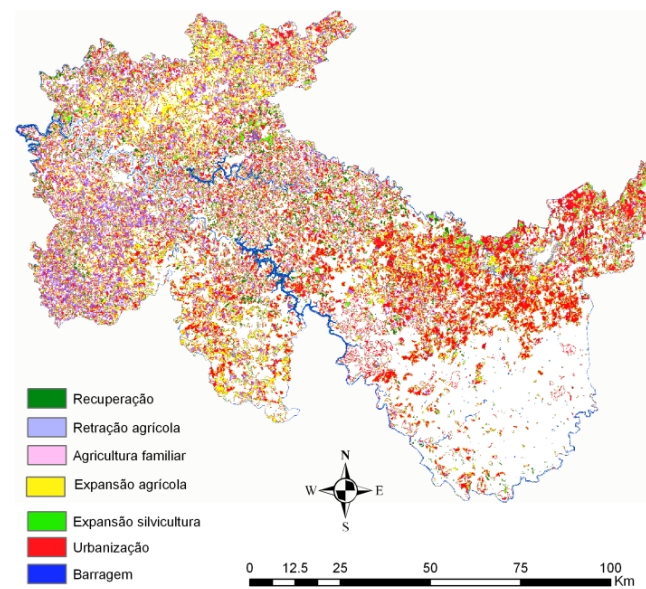
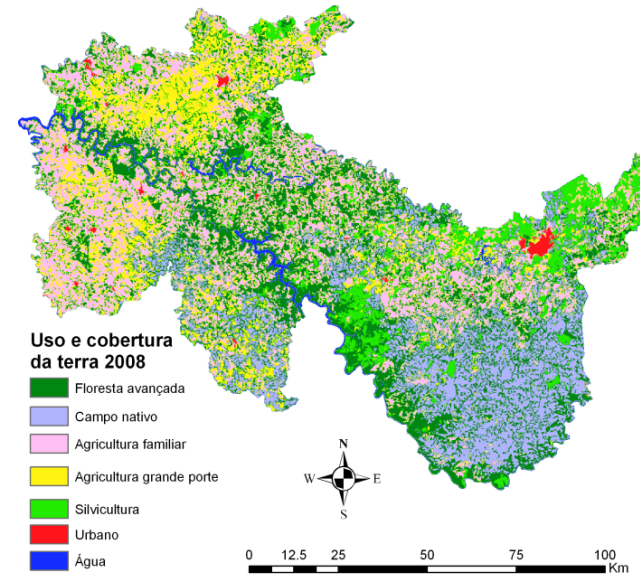
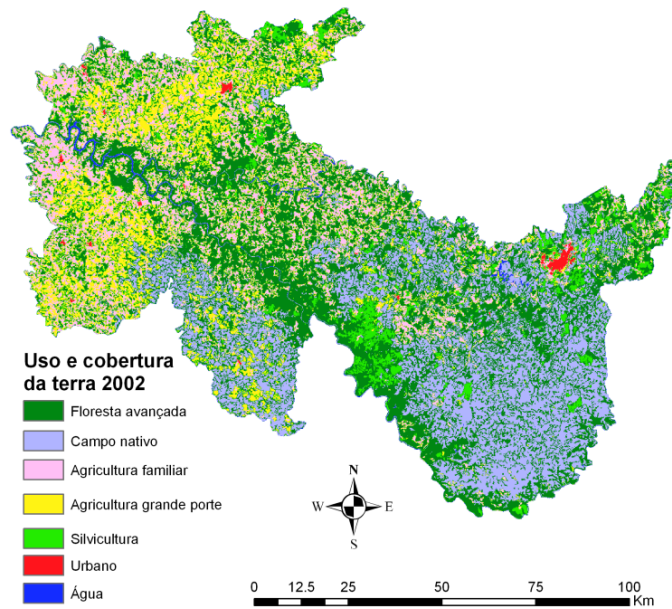
$$Y = \beta_0 + X\beta_1 + \varepsilon$$

- Modelo Spatial Lag (SAR)

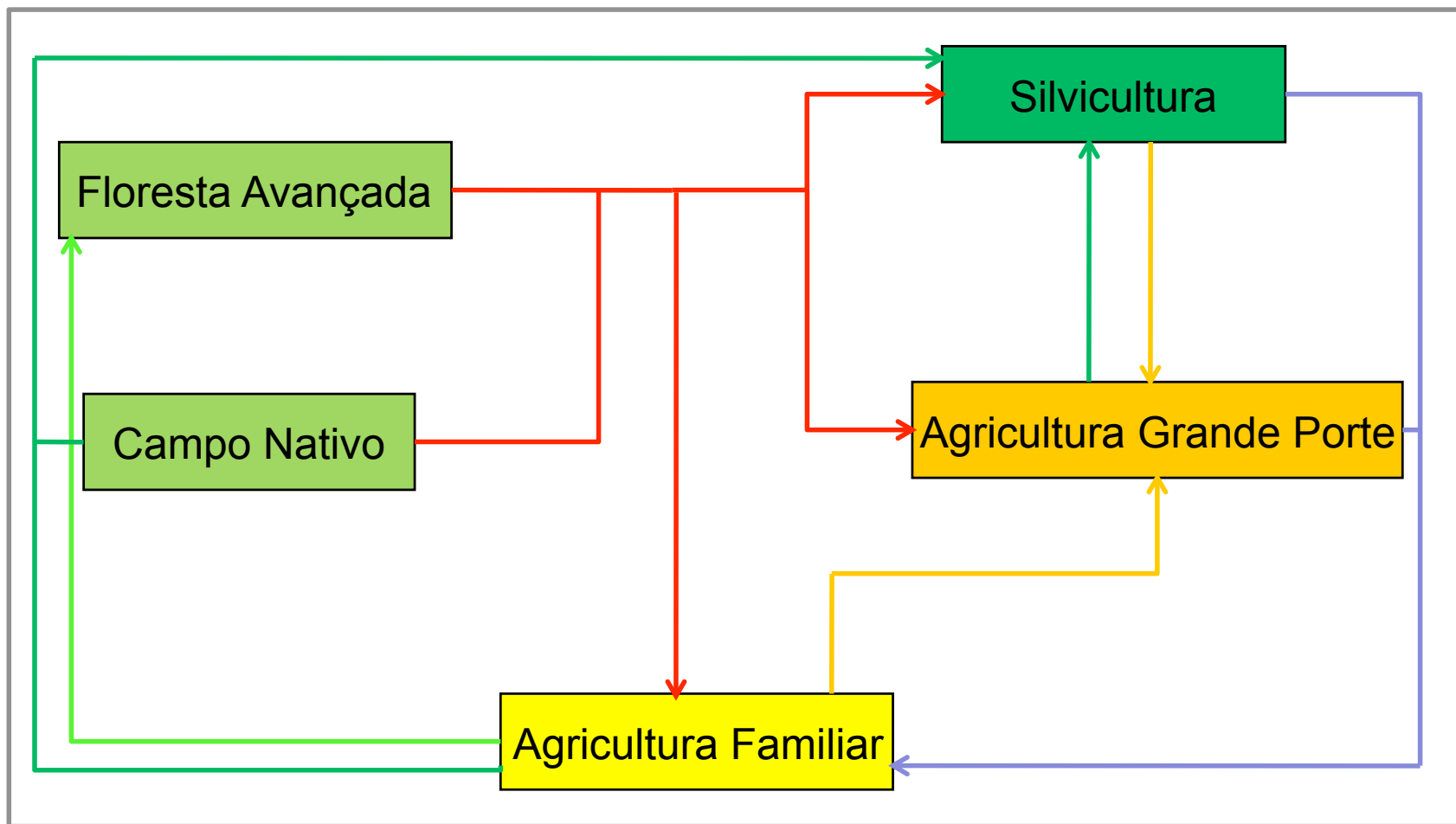
$$Y = \rho WY + X\beta + \varepsilon$$

- Y é a variável dependente,
- ρ é o coeficiente espacial autoregressivo (medida de correlação espacial)
- W = matriz de vizinhança espacial
- WY expressa a dependência espacial em Y
- X é a matriz de observações das variáveis independentes
- β_0 é o intercepto populacional
- $\beta_1 \dots$ são coeficientes da regressão
- ε é o erro aleatório, com distribuição normal e variância constante

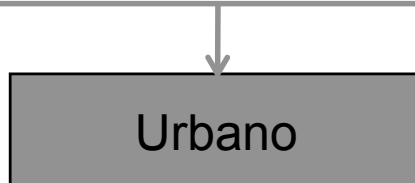
Uso e cobertura da terra 2002-2008



Processos de mudanças de uso e cobertura da terra



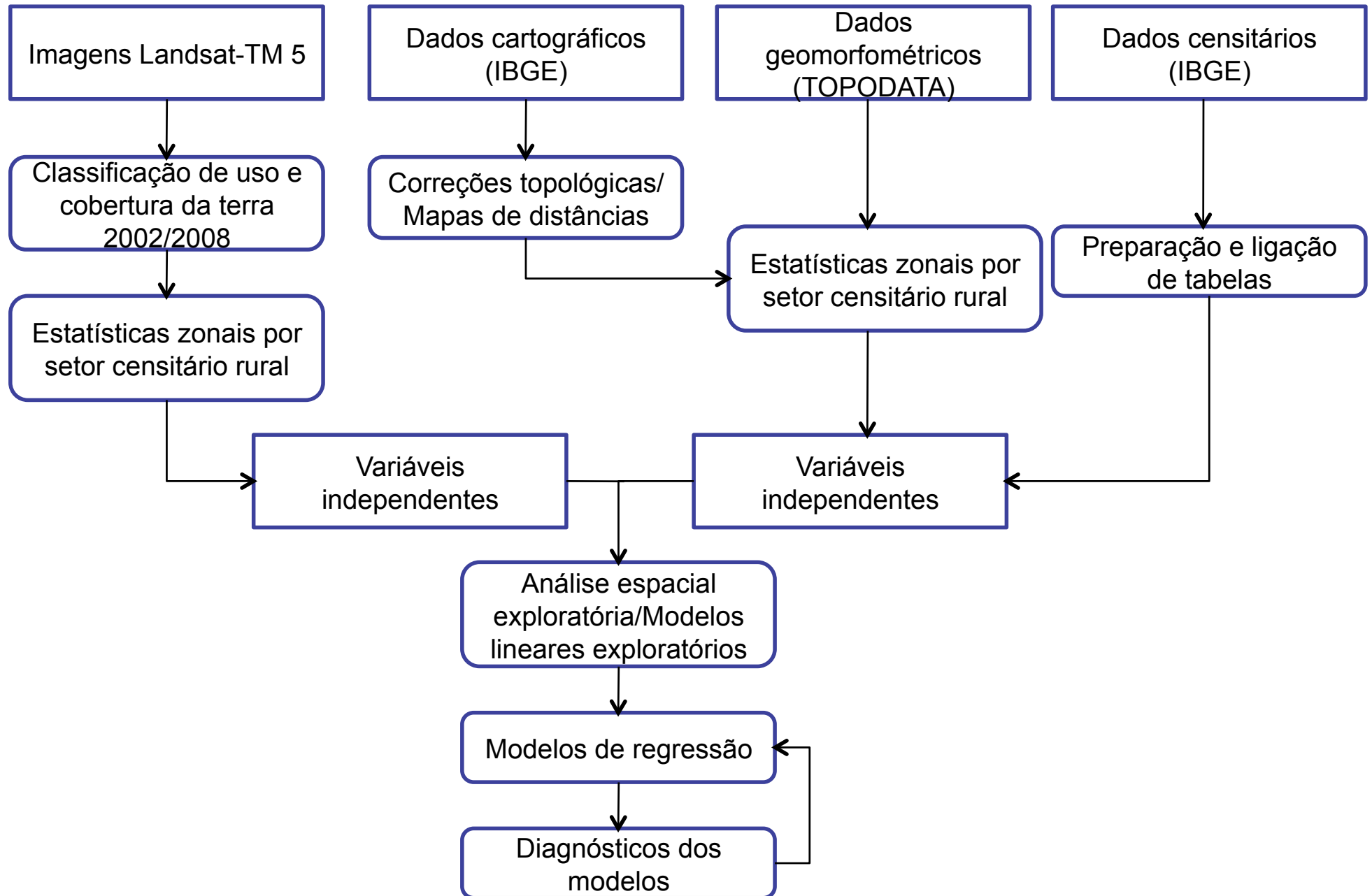
- Recuperação
- Retração agrícola
- Degradação
- Expansão agrícola
- Expansão silvicultura
- Urbanização



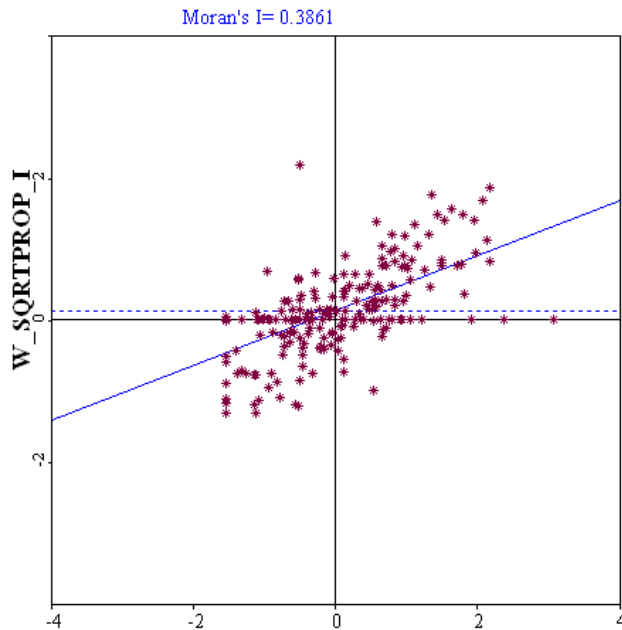
Variáveis independentes

- **Renda:**
 - Rendimento nominal mensal por pessoa responsável por domicílio particular permanente
 - Média do rendimento nominal mensal das pessoas responsáveis por domicílios particulares permanentes
- **Educação:**
 - Média do número de anos de estudo das pessoas responsáveis por domicílios particulares permanentes
 - Taxa de alfabetização (Pessoas alfabetizadas com 5 ou mais anos de idade*100)/Pessoas com 5 anos ou mais
- **Domicílio:**
 - Média do número de moradores em domicílios particulares permanentes
- **Demográficos:**
 - Densidade populacional (população residente no setor/área do setor censitário em Km²)
 - Razão de dependência ((população com menos de 15 anos + população com mais de 64 anos)/população entre 15 e 64 anos)*100
 - Porcentagem de população com menos de 15 anos
 - Porcentagem de população entre 15 e 64 anos
 - Porcentagem de população com mais de 64 anos
- **Infraestrutura**
 - Distância euclidiana média às estradas principais e vicinais
 - Desvio-padrão da distância euclidiana às estradas principais e vicinais
- **Geomorfométricas**
 - Amplitude altimétrica
 - Altitude média
 - Declividade média
 - Declividade mediana
 - Desvio-padrão da declividade

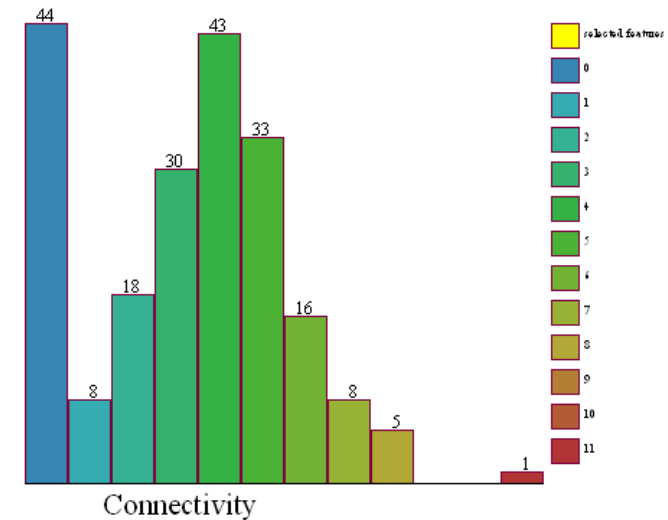
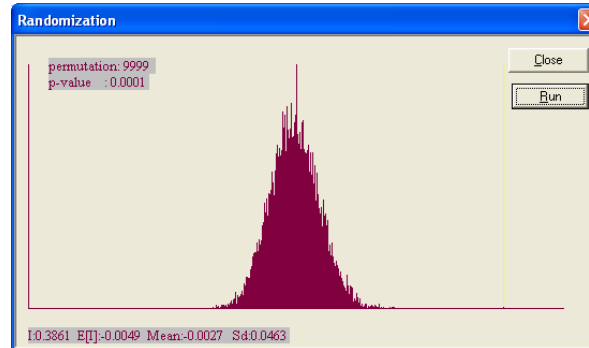
Metodologia geral



Análise espacial exploratória



Expansão agrícola



Matriz de vizinhança - Rook

Variáveis	Moran's I	p-valor
Degradação	0,2999	0,0001
Recuperação	0,3856	0,0001
Retração agrícola	0,8159	0,0001
Expansão agrícola	0,3861	0,0001
Expansão silvicultura	0,3935	0,0001
Urbanização	0,0557	0,0951

Obs.: Todas variáveis são proporções da área total e foram transformadas por raiz quadrada

Regressões backward stepwise

Processo	Nº de variáveis	Variáveis	R ²	P (estatística F)	CN	P (Jarque-Bera)	P(Koenker-Bassett)
Degradação	7	pop+64/ renda/ dens_pop/ amp_alt/ dec_mediana / dec_med / dec_desv	0.20	8.50E-08*	118.4	0.00	0.23****
Recuperação	4	renda/ tx_alf/ alt_med / dec_med	0.33	4.60E-17*	52.7	0.00	0.66****
Retração agrícola	5	dens_pop/ amp_alt/ alt_med/ dec mediana / dist med	0.35	2.16E-17*	26.8**	0.00	0.00
Expansão de silvicultura	3	pop+64/ dens_pop/ alt_med	0.18	9.58E-09*	16.9**	0.00	0.05****
Urbanização	4	med_estudo/ tx_alf/ dens_pop / dec desv	0.64	1.40E-44*	48.9	0.00	0.00
Expansão agrícola	5	raz_dep/ renda_media/ dens_pop/ dec_mediana/ dist_desv	0.27	7.56E-13*	17.2**	0.10***	0.01

* modelos significantes ($\alpha=0,05$)

** sem problemas de multicolinearidade (CN<30)

*** resíduos com distribuição normal ($\alpha=0,05$)

**** resíduos homocedásticos ($\alpha=0,05$)

Regressão linear simples – Modelo Linear 1

REGRESSION

SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION

```

Data set      : setores_topo_dados_trans
Dependent Variable : SQRTPROP_I  Number of Observations: 206
Mean dependent var : 0.200706   Number of Variables : 6
S.D. dependent var : 0.129785   Degrees of Freedom : 200

R-squared      : 0.278734   F-statistic      : 15.458
Adjusted R-squared : 0.260702   Prob(F-statistic) : 7.55762e-013
Sum squared residual: 2.50273   Log likelihood    : 161.98
Sigma-square    : 0.0125136   Akaike info criterion : -311.959
S.E. of regression : 0.111864   Schwarz criterion  : -291.992
Sigma-square ML  : 0.0121492
S.E of regression ML: 0.110223
    
```

Significância do modelo

Medidas de ajuste

Variable	Coefficient	Std. Error	t-Statistic	Probability
CONSTANT	0.4234906	0.05348975	7.917229	0.0000000
RAZ_DEP	-0.001408288	0.0006528802	-2.157039	0.0321955
DENS_POP	-9.895512e-005	2.779341e-005	-3.56038	0.0004627
VAR06	8.623844e-005	4.823016e-005	1.78806	0.0752804
DEC_MEDIAN	-0.01198294	0.001659423	-7.221153	0.0000000
DIST_DESV	3.574658e-005	1.940477e-005	1.842154	0.0669330

REGRESSION DIAGNOSTICS

MULTICOLLINEARITY CONDITION NUMBER 17.27017

TEST ON NORMALITY OF ERRORS

TEST	DF	VALUE	PROB
Jarque-Bera	2	4.553387	0.1026230

Resíduos com dist. normal

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	5	16.58229	0.0053641
Koenker-Bassett test	5	14.46462	0.0129125

Resíduos heterocedásticos

SPECIFICATION ROBUST TEST

TEST	DF	VALUE	PROB
White	20	42.7709	0.0021917

DIAGNOSTICS FOR SPATIAL DEPENDENCE

FOR WEIGHT MATRIX : setores_topo_dados_trans.GAL (row-standardized weights)

TEST	MI/DF	VALUE	PROB
Moran's I (error)	0.234735	5.3353759	0.0000001
Lagrange Multiplier (lag)	1	40.9717577	0.0000000
Robust LM (lag)	1	15.7615154	0.0000718
Lagrange Multiplier (error)	1	25.4005163	0.0000005
Robust LM (error)	1	0.1902739	0.6626887
Lagrange Multiplier (SARMA)	2	41.1620317	0.0000000

AE

Escolha do modelo SAR

Regressão espacial – Modelo SAR 1

REGRESSION

SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION

```

Data set          : setores_topo_dados_trans
Spatial Weight    : setores_topo_dados_trans.GAL
Dependent Variable : SQRTPROP_I  Number of Observations: 206
Mean dependent var : 0.200706  Number of Variables : 7
S.D. dependent var : 0.129785  Degrees of Freedom : 199
Lag coeff. (Rho)  : 0.639799

R-squared         : 0.435204  Log likelihood : 187.067
Sq. Correlation   : -         Akaike info criterion : -360.134
Sigma-square      : 0.00951355  Schwarz criterion : -336.839
S.E of regression : 0.0975374
    
```

Modelo SAR
melhor
ajustado

Variable	Coefficient	Std.Error	z-value	Probability
W_SQRTPROP_I	0.6397994	0.04812328	13.29501	0.0000000
CONSTANT	0.2256669	0.04925777	4.581347	0.0000046
RAZ_DEP	-0.0007749845	0.0005718423	-1.355242	0.1753407
DENS_POP	1.671196e-005	2.599772e-005	0.6428239	0.5203382
VAR06	7.533054e-005	4.207828e-005	1.790248	0.0734140
DEC_MEDIAN	-0.007526182	0.001485853	-5.065226	0.0000004
DIST_DESV	-8.2576e-006	1.709769e-005	-0.4829657	0.6291201

Perda de
significância

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

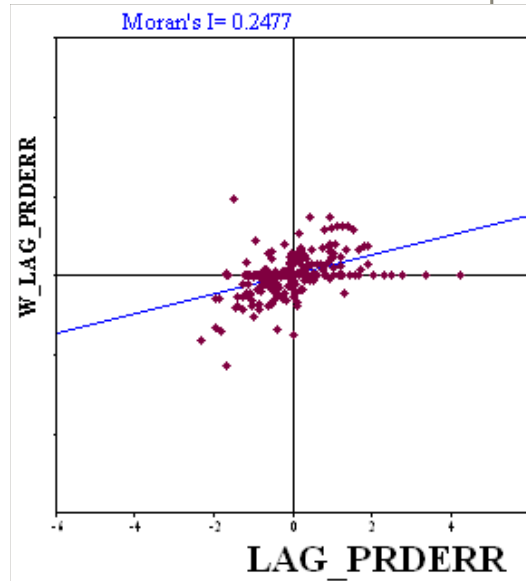
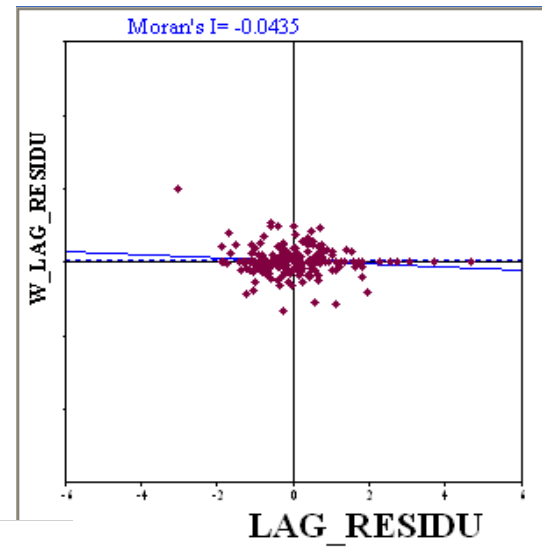
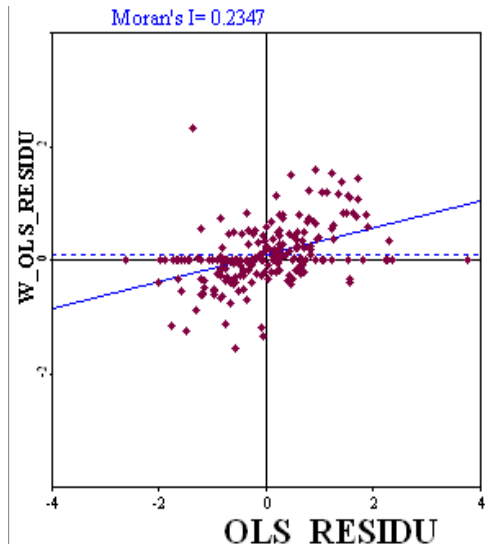
TEST	DF	VALUE	PROB
Breusch-Pagan test	5	24.47674	0.0001758

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : setores_topo_dados_trans.GAL

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	50.17453	0.0000000

Dependência espacial dos resíduos

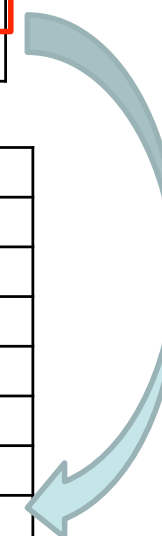


Modelos de regressão 2 e 3

Modelo	R ²	P (estatística F)	CN	P (Jarque-Bera)	P(Koenker-Bassett)	P(Breusch-Pagan)	P(White)	Log	AIC
Linear 2	0.18	6.24E-010*	11.6**	0.76***	0.0001	0.0003	0.0001	149.832	-293.665
SAR 2	0.29	-	11.6**	-	-	0.006	-	164.043	-320.086
Linear 3	0.15	2.45E-008*	9.1**	0.78***	0.0000617	0.0001469	0.0000344	146.107	-286.214
SAR 3	0.28	-	9.1**	-	-	0.0069067	-	162.899	-317.797

Modelo	Variável	Coefficiente	Erro	Estatística t	Valor-p
Linear 2	Constante	0.4649756	0.04445539	10.45937	0.0000000
	raz_dep	-0.00197805	0.0006639551	-2.979192	0.0032425
	dec_mediana	-0.009847137	0.001637443	-6.01373	0.0000000
Linear 3	Constante	0.3315638	0.03664631	9.047671	0.0000000
	renda_med	5.710116e-005	4.916125e-005	1.161507	0.2467997
	dec_mediana	-0.009685661	0.001705135	-5.68029	0.0000000

Modelo	Variável	Coefficiente	Erro	Estatística z	Valor-p
SAR 2	Lag coef. (rho)	0.4585812	0.05398901	8.493974	0.0000000
	Constante	0.3916105	0.04298717	9.109939	0.0000000
	raz_dep	-0.001588582	0.0006198047	-2.563037	0.0103761
	dec_mediana	-0.01148263	0.0015178	-7.565313	0.0000000
SAR 3	Lag coef. (rho)	0.5106849	0.05276368	9.678719	0.0000000
	Constante	0.2482819	0.03516266	7.060953	0.0000000
	renda_med	0.000107097	4.532846e-005	2.362688	0.0181429
	dec_mediana	-0.01107374	0.001559601	-7.100368	0.0000000



Ajuste dos Modelos espaciais

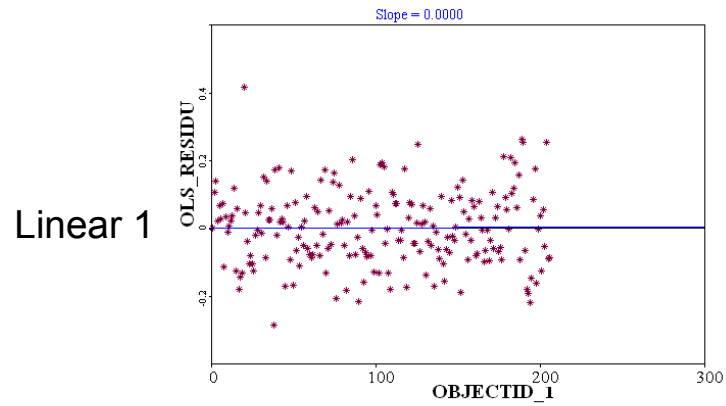
- Anselin (2005): Ordem decrescente dos teste Wald (W), Razão de verossimilhança (LR) e LM-Lag (LM)

$$W > LR > LM$$

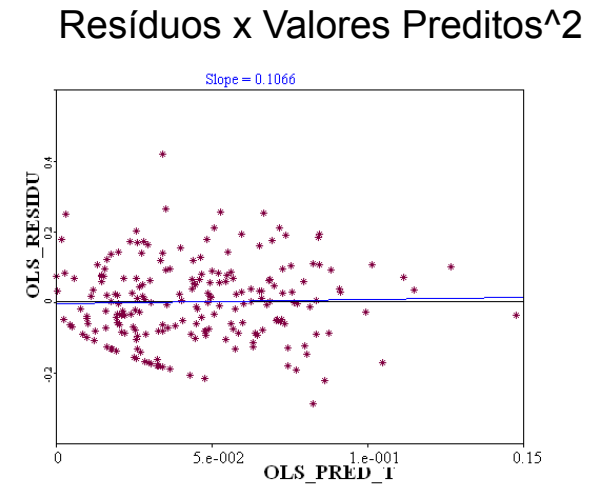
Modelo	Wald	LR	LM
SAR 1	176.76	50.17	40.97
SAR 2	93.67	33.58	27.95
SAR 3	72.14	28.42	23.54

Análise de diagramas de dispersão

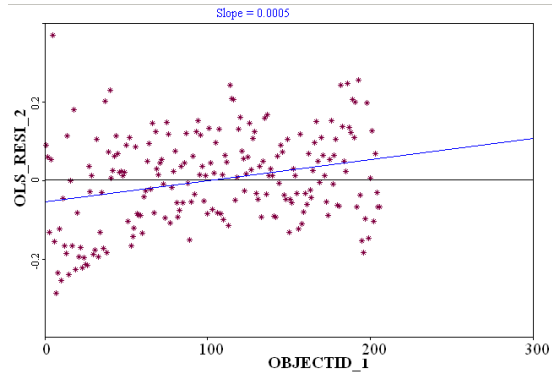
Plot dos resíduos



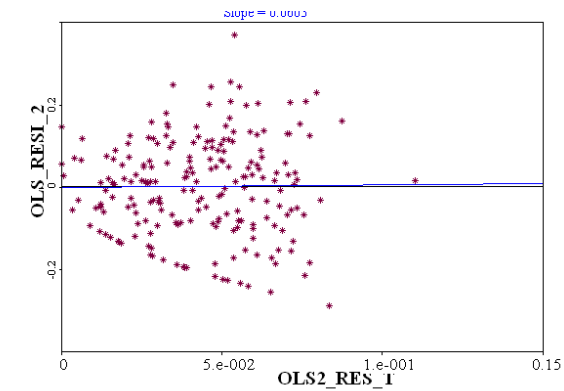
SAR 1



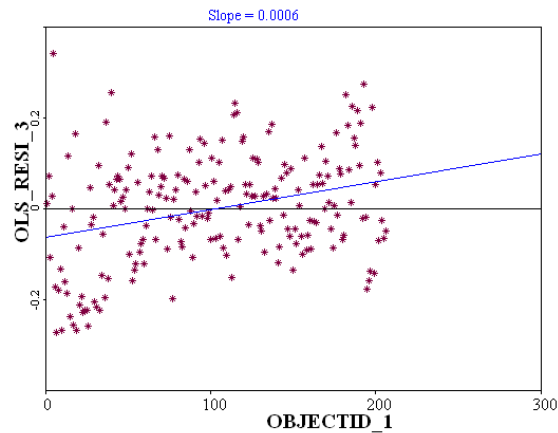
Linear 2



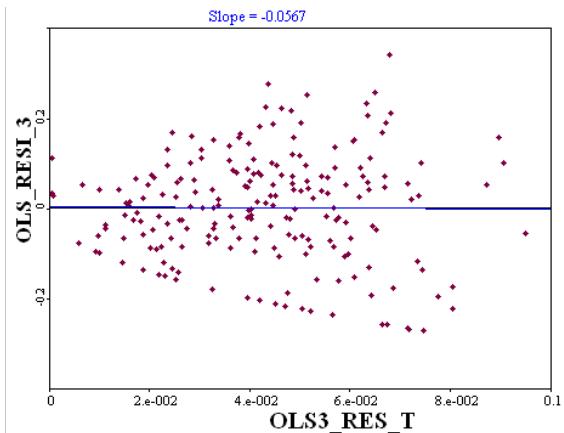
SAR 2



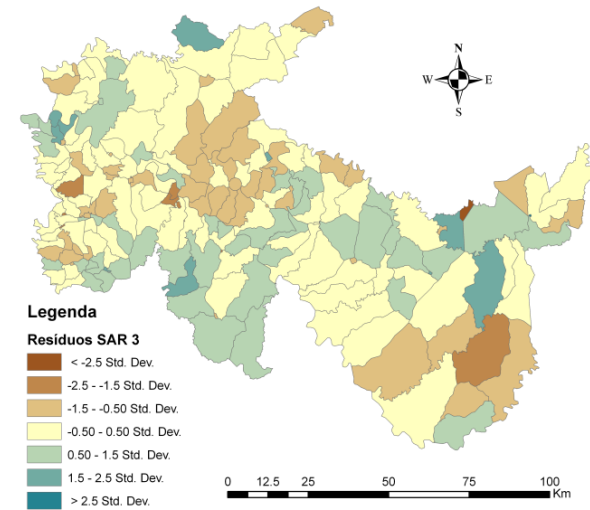
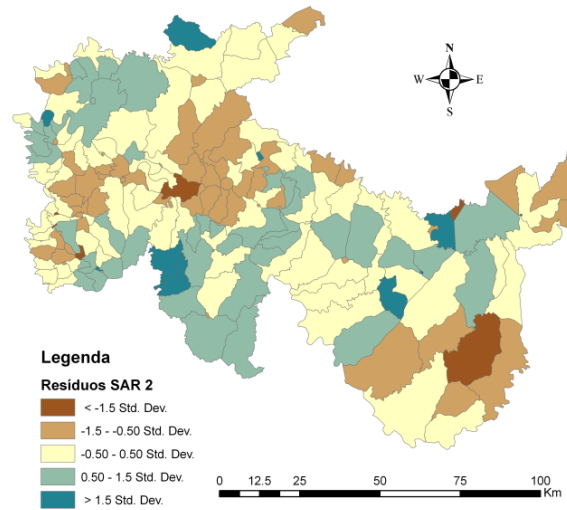
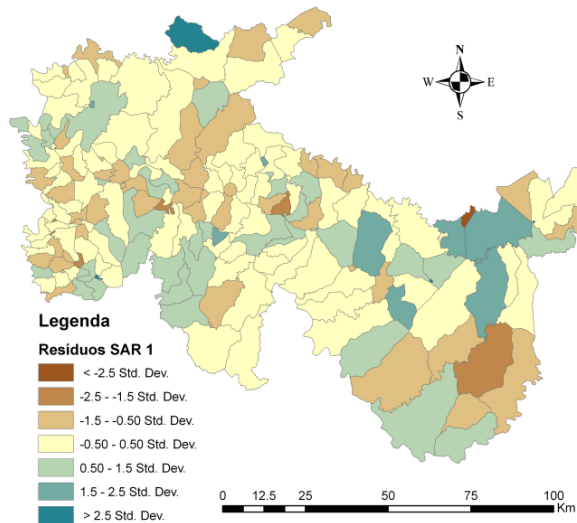
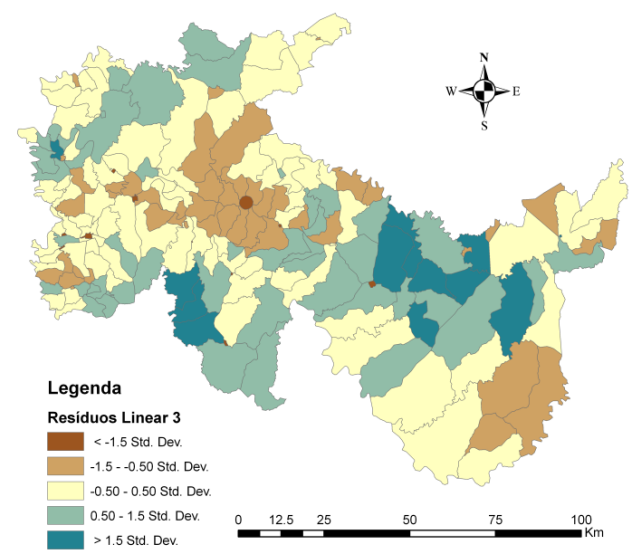
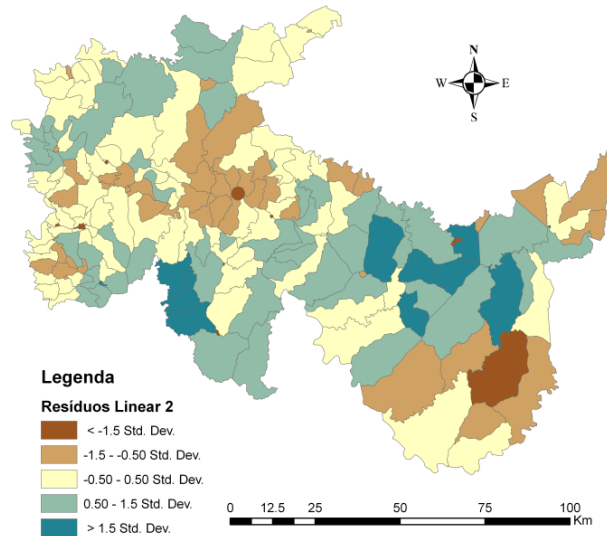
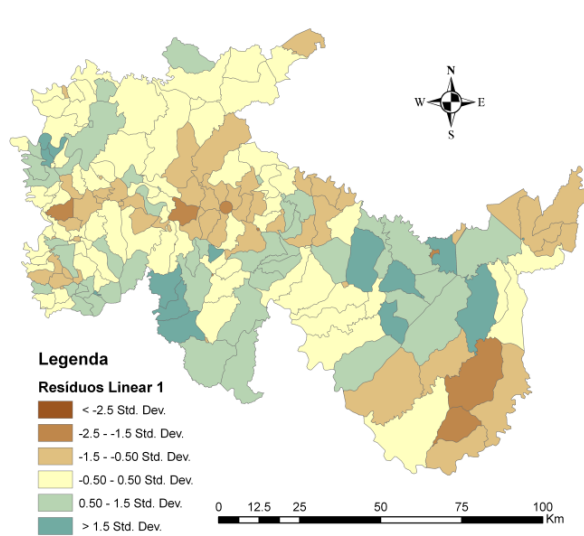
Linear 3



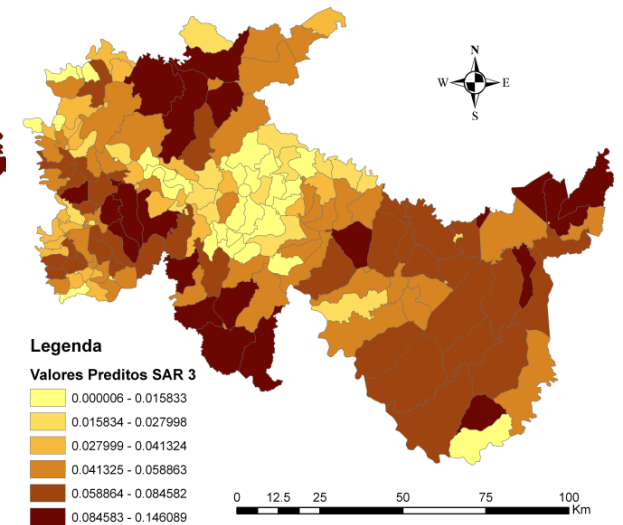
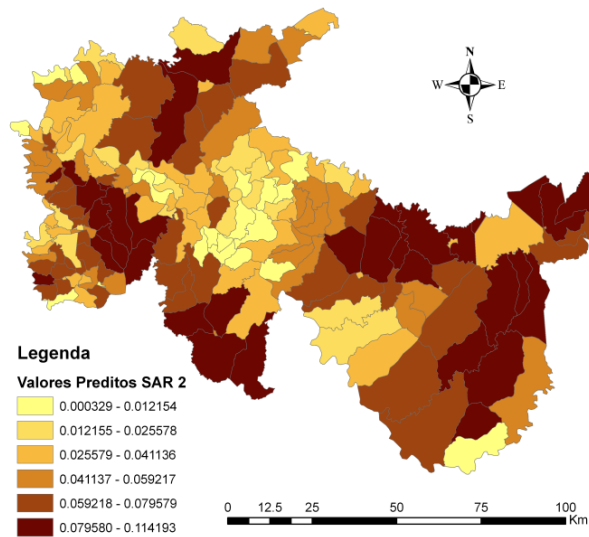
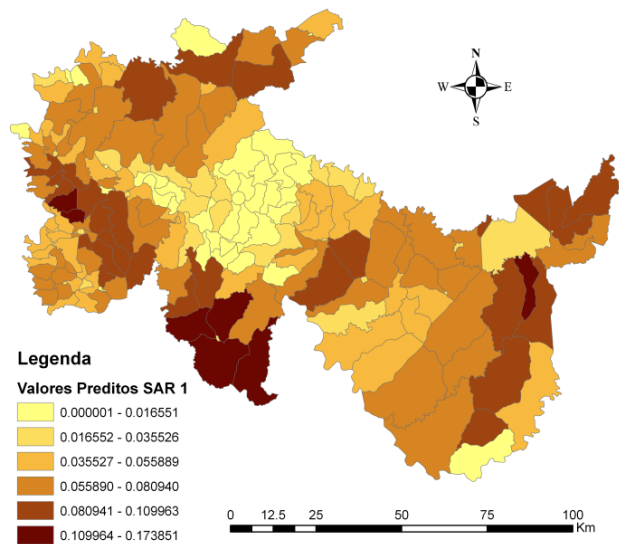
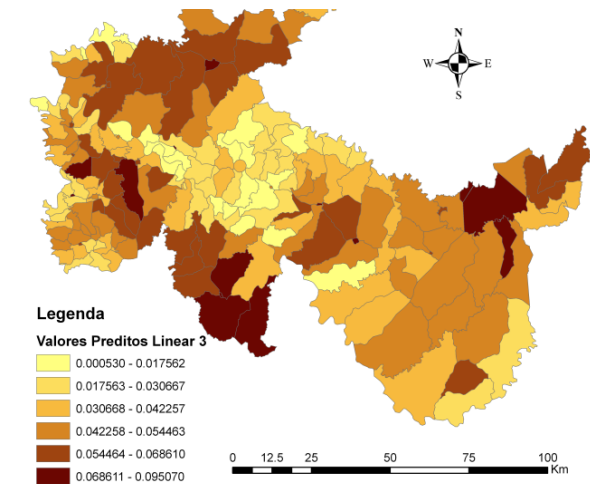
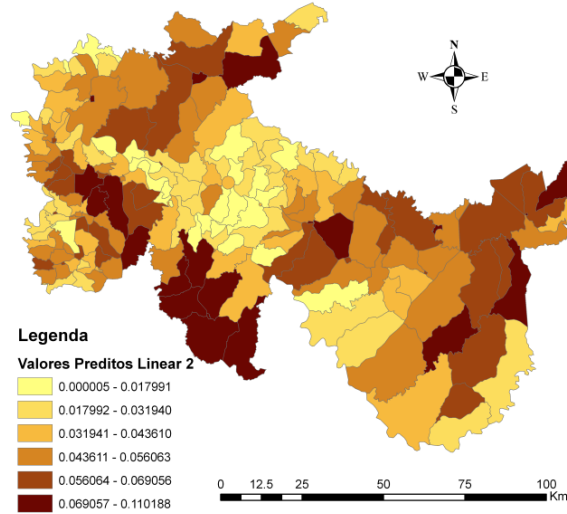
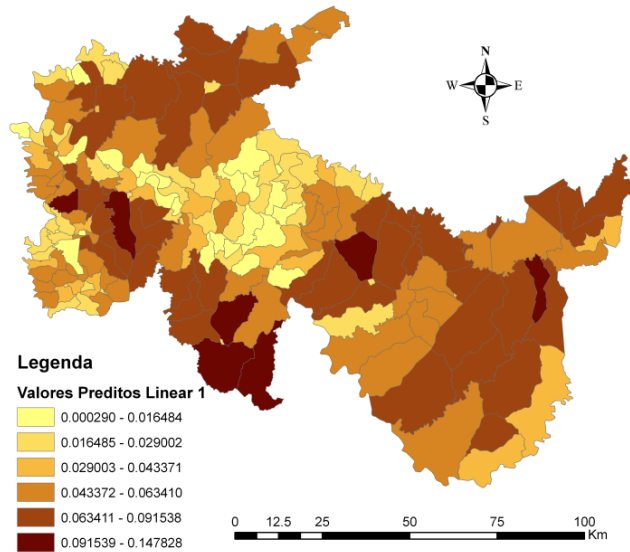
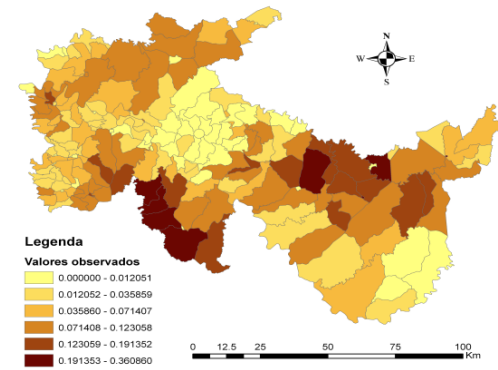
SAR 3



Mapeamento dos resíduos



Mapeamentos dos Valores Preditos



Conclusões

- Variáveis selecionadas pelos métodos exploratórios (regressão *stepwise*) apresentaram comportamentos diferentes a partir da aplicação do componente auto-regressivo espacial
- Variáveis sociais e físicas apresentaram associação com os processos
- Variáveis de mudança heterogêneas – necessidade do uso integrado de fatores sociais e ambientais em estudos de mudanças de UCT
- Questões:
 - Será que a premissa de estacionariedade é aplicável neste estudo de caso?
 - Métodos locais não seriam mais adequados ao fenômeno estudado nesta área?
 - Outras variáveis poderão apresentar maior relação com os processos estudados?
- Trabalhos futuros:
 - Análise de autocorrelação espacial por métodos locais (LISA)
 - Aplicação do modelo GWR que assume a heterogeneidade espacial
 - Incorporação de mais variáveis

Bibliografia

- Alves, H. P. F. Fatores demográficos e sócio-econômicos associados às mudanças na cobertura da terra no Vale do Ribeira: análise integrada de dados censitários e de sensoriamento remoto através de um sistema de informação geográfica. In: Hogan, D. J. (Org.). **Dinâmica populacional e mudança ambiental: cenários para o desenvolvimento humano**. Campinas, Nepo/Unicamp, 2007, p.117-143.
- Anselin, L. Exploring Spatial Data with GeoDa™ : A Workbook. Center for Spatially Integrated Social Science, 2005, 226p.
- _____. Under the hood: issues in the specification and interpretation of spatial regression models. **Agricultural Economics**, n. 27, 2002, p.247-267
- Carvalho, M. S; Cruz, O. G.; Nobre, F. F. Perfil de risco: método multivariado de classificação sócio-econômica de microáreas urbanas – os setores censitários da região metropolitana do Rio de Janeiro. **Cadernos de Saúde Pública**, Rio de Janeiro, n.13, v.4, 1997, p.635-645.
- D'Antona, A.; Cak, A. D.; Wey, L. K. V. Efeitos de escala de análise em estudos de mudança da cobertura da terra entre Santarém e Altamira, no Pará, Brasil. In: Hogan, D. J. (Org.). **Dinâmica populacional e mudança ambiental: cenários para o desenvolvimento humano**. Campinas, Nepo/Unicamp, 2007, p.87-113.
- Lesschen, J. P.; Verburg, P. H.; Staal, S. J. **Statistical methods for analysing the spatial dimension of changes in land use and farming systems**. Lucc Report Series No. 7, Nairobi/Wageningen, The International Livestock Research Institute/ Lucc Focus 3 Office, 2005, 80p.
- Overmars, K.P.; Koning G.H.J.; Veldkamp, A. Spatial autocorrelation in multi-scale land use models. **Ecological Modelling**, n.164, 2003, p.257–270.
- Soares-Filho, B. S.; [Garcia, R. A.](#); Rodrigues, H.; Moro, S.; Nepstad, D. Nexos entre as dimensões socioeconômicas e o desmatamento: a caminho de um modelo integrado. In: Batistella, M.; Alves, D. S.; Moran, E.. (Org.). **Amazônia. Natureza e Sociedade em Transformação**. São Paulo: Edusp, 2008.
- Verburg, P. H. Soepboer, W.; Veldkamp, A.; Limpiada, R.; Espaldon, V.; Mastura, S. S. A. Modeling the Spatial Dynamics of Regional Land Use: The CLUE-S Model. *Environmental Management*, v. 30, n. 3, 2002, p. 391– 405