

Arranjos urbanos-rurais regionais em São Paulo: um exercício de análise espacial*

Natália Belmonte Demétrio¹

Rosana Baeninger²

Álvaro de Oliveira D'Antona³

Ricardo de Sampaio Dagnino⁴

Resumo

Esse trabalho apresenta uma discussão metodológica de como apreender a heterogeneidade do rural paulista a partir da integração de dados dos Censos Agropecuário e Demográfico. O objetivo é explorar técnicas de análise multivariada e espacial que auxiliem na apreensão do rural como construção socioespacial, estruturado por fluxos que extrapolam a produção estritamente agropecuária e remetem a diferentes dinâmicas de redistribuição de população. O artigo está dividido em cinco partes principais. Na introdução, é explicitado a perspectiva teórica do trabalho. Em seguida, apresentam-se as variáveis utilizadas, o tratamento dado aos casos perdidos, avaliação das distribuições e correlações. Na terceira parte, recorre-se à análise fatorial como de reduzir o conjunto de indicadores iniciais em constructos fracamente correlacionados. No item quatro, os fatores criados são submetidos a uma análise classificatória de cluster que, na quinta e última parte, é interpretada espacialmente. A título de considerações finais, é reiterada a multiplicidade de formas com que o rural regional se conecta ao urbano local, de modo a reforçar a centralidade dos processos históricos na articulação de diferentes espaços rurais.

Palavras-chave: População rural, Economia Rural, Análise Multivariada

*Trabalho apresentado no VII Congresso da Associação Latino-americana de População e XX Encontro Nacional de Estudos Populacionais, ABEP, realizado em Foz do Iguaçu/SP – Brasil, de 17 a 22 de outubro de 2016

¹Aluna do Programa de Pós-Graduação em Demografia – Nível Doutorado – do Instituto de Filosofia e Ciências Humanas (IFCH/Unicamp), com bolsa do CNPq. E-mail: natalia@nepo.unicamp.br

²Professora do Programa de Pós-Graduação em Demografia e do Programa de Pós-Graduação em Sociologia do Instituto de Filosofia e Ciências Humanas - UNICAMP. E-mail: baeninger@nepo.unicamp.br

³Professor da Faculdade de Ciências Aplicadas (FCA-Unicamp) e do Programa de Pós-graduação de Demografia (IFCH-Unicamp). Pesquisador do Núcleo de Estudos de População “Elza Berquó”. E-mail: alvaro.dantona@fca.unicamp.br

⁴Pós-doutorando em Análises Demográficas Espaciais na Faculdade de Ciências Aplicadas da Universidade Estadual de Campinas (FCA/Unicamp), com bolsa da FAPESP (2015/16270-2) vinculada ao Observatório das Migrações em São Paulo. E-mail: ricardo.dagnino@fca.unicamp.br

1. Introdução

O tripé distribuição da população, economia regional e dinâmica espacial constitui o enfoque analítico desse trabalho (BAENINGER, 1999). Ancorado no processo de urbanização, a análise das relações rurais/urbanas subsidia um debate a partir de dois conceitos principais: espaço e rede urbana. Por um lado, pensar o espaço como acumulação desigual de tempos (SANTOS, 1993) leva ao entendimento de suas rugosidades e heranças históricas. Por outro, os estudos sobre rede urbana são fundamentais na apreensão dos espaços rurais em suas relações de interdependência com urbano, em múltiplas escalas (CORRÊA, 1994). O adensamento da rede de cidades, longe de promover a homogeneização espacial e a diluição das diferenças entre rural e urbano, apenas redefine “os padrões de diferenciação espacial, alterando-os em termos da divisão territorial do trabalho”. (CÔRREA, 1994, p.66-67).

À luz do trabalho de Moura (2009) – cuja hipótese dos arranjos urbanos-regionais centra-se na relação globalização/metropolização –, esse artigo avança no debate sobre a reestruturação urbano-regional em áreas não integradas pelos tradicionais eixos de desenvolvimento fordista. Seja pelo crescimento dos investimentos estrangeiros diretos (SASSEN, 1998), seja pelo progresso tecnológico na área de telecomunicações, o chamado curto-circuito da rede urbana (SANTOS, 2013) leva à criação de novas hierarquias urbanas, entrelaçadas por solidariedades organizacionais que atravessam as dinâmicas de divisão territorial do trabalho construídas a nível regional e nacional. Nesse sentido, as regionalizações tradicionais, tecidas por uma solidariedade orgânica entre um núcleo urbano e seu entorno imediato, representam cada vez menos os nexos espaciais no raio de ação do capital globalizado (SANTOS, 2013).

Essa proposição teórica impõe um desafio metodológico fundamental quanto à delimitação de regionalizações através das quais seja possível apreender as metamorfoses do espaço trazidos pela globalização da produção, a partir da consolidação de distintos espaços transnacionais (SASSEN, 1998; SANTOS, 2013). Se Moura (2009) avança na reflexão sobre as regiões de caráter metropolitano, nesse trabalho, a atenção volta-se para a relação entre globalização, mobilidade espacial da população e produção do espaço em regiões onde o desenvolvimento econômico passa principalmente pela produção de commodities.

2. Avaliação das variáveis

Tendo como unidade de análise dos 645 municípios do Estado de São Paulo, foram selecionados 32 indicadores que servem de *proxy* à captação dos seguintes processos: redistribuição da população, estrutura fundiária e dinâmica econômica (Quadro 1). Diante da importância da produção de cana-de-açúcar, laranja e da criação de bovinos para corte no Estado, essas três atividades compuseram blocos específicos de estimativas. Uma dificuldade fundamental da pesquisa diz respeito ao uso de dados de diferentes períodos: Censo Demográfico 2010 e Censo Agropecuário 2006. A rigor, tais estimativas não podem ser comparadas, haja vista a grande oscilação na produção ao longo desse intervalo, resultado da alta influência do mercado externo na determinação da demanda por tais commodities. Contudo, diante do pressuposto de que a dinâmica da população não apresente alterações bruscas, optou-se pela junção dessas bases. Em concordância com uma análise de período, o trabalho reuniu apenas estimativas de estoque.

Também é elementar esclarecer o tratamento dedicado aos casos perdidos. Nos dados extraídos do Censo Demográfico, os *missings* representam municípios sem população classificada como rural. No tocante às estimativas do Censo Agropecuário, a não informação resulta tanto da não existência de estabelecimentos agropecuários⁵ no município, quanto de uma questão de sigilo: indicadores referentes a unidades territoriais com menos de três informantes não são divulgados. Na construção do banco de dados, optou-se pela permanência de tais observações, as quais receberam o valor zero. Essa opção compromete a capacidade explicativa de medidas resumo dos dados, como a média, com impactos no processo de sintetização das informações.

A escolha por manter os *missings* remete, principalmente, ao viés político das definições oficiais de rural e urbano no Brasil (CUNHA, 2005). Como prerrogativa do poder público municipal, essa classificação pode “não apenas variar de uma região para outra”, como também depender de “injunções e interesses que [...] nem sempre obedecem a uma lógica racional ou funcional” (CUNHA, 2005, p.13). Desse modo, ainda que a variável ‘situação de domicílio’ seja importante, a construção dos arranjos urbanos-regionais rurais não está subordinada aos parâmetros oficiais.

⁵ Vale ressaltar que a existência de estabelecimentos agropecuários independe da situação urbana ou rural de domicílio (IBGE, 2006).

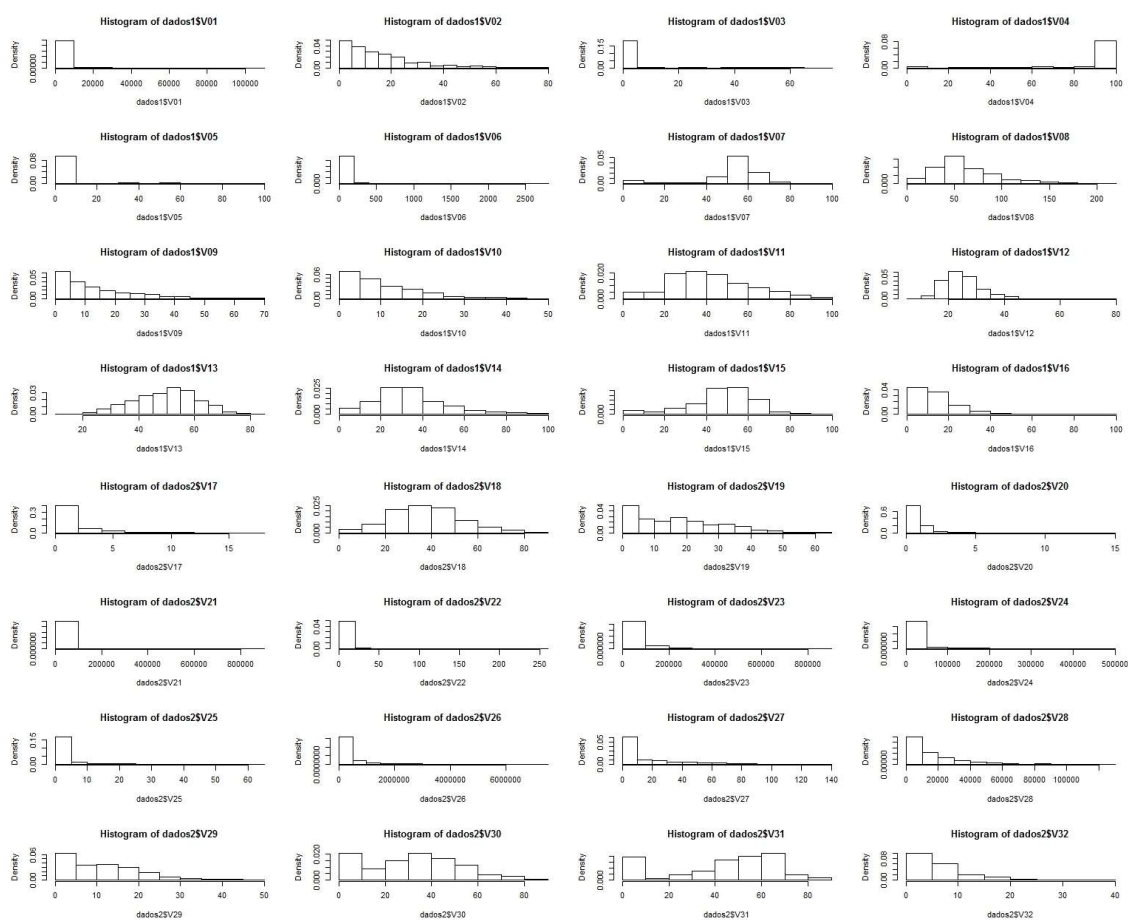
Quadro 1
 Descrição das variáveis utilizadas
 Estado de São Paulo, por município
 2006 e 2010

Bloco	Código	Descrição	Fonte	Ano
Demográficas	V01	População rural	FIBGE, Censo Demográfico 2010 (Sidra, Tabela 1378)	2010
	V02	Grau de ruralização ¹	FIBGE, Censo Demográfico 2010 (Sidra, Tabela 1378)	2010
	V03	Proporção de população residente em situação de setor 4 (rural extensiva-urbana) no total de população rural	FIBGE, Censo Demográfico 2010 (tabulação própria)	2010
	V04	Proporção de população residente em situação de setor 8 (rural exclusiva aglomerados) no total de população rural	FIBGE, Censo Demográfico 2010 (tabulação própria)	2010
	V05	Proporção de população carcerária (população residente em setor especial tipo 6) no total de população rural	FIBGE, Censo Demográfico 2010 (tabulação própria)	2010
	V06	Razão de sexo da população rural ²	FIBGE, Censo Demográfico 2010 (Sidra, Tabela 1378)	2010
	V07	Razão de dependência da população rural ³	FIBGE, Censo Demográfico 2010 (Sidra, Tabela 1378)	2010
	V08	Índice de Envelhecimento da população rural ⁴	FIBGE, Censo Demográfico 2010 (Sidra, Tabela 1378)	2010
	V09	Pendularidade rural ⁵	FIBGE, Censo Demográfico 2010 (tabulação própria)	2010
	V10	Proporção de trabalhadores residentes no urbano e que estavam ocupados no setor primário	FIBGE, Censo Demográfico 2010 (tabulação própria)	2010
	V11	Proporção de trabalhadores residentes no rural e que estavam ocupados nos setores secundário e terciário	FIBGE, Censo Demográfico 2010 (tabulação própria)	2010
	V12	Proporção de migrantes que chegaram nos últimos três anos no total da população não natural	FIBGE, Censo Demográfico 2010 (tabulação própria)	2010
	V13	Proporção de nascidos no município que nunca migrou no total de população residente	FIBGE, Censo Demográfico 2010 (tabulação própria)	2010
Estrutura fundiária	V14	Proporção de estabelecimento de até 10 ha no total de estabelecimentos agropecuário do município	FIBGE, Censo Agropecuário (Sidra, Tabela 837)	2006
	V15	Proporção de estabelecimento de 10 a 100 ha no total de estabelecimentos agropecuário do município	FIBGE, Censo Agropecuário (Sidra, Tabela 837)	2006
	V16	Proporção de estabelecimento de 100 a 1000 ha no total de estabelecimentos agropecuário do município	FIBGE, Censo Agropecuário (Sidra, Tabela 837)	2006
	V17	Proporção de estabelecimento com mais de 1000 ha no total de estabelecimentos agropecuário do município	FIBGE, Censo Agropecuário (Sidra, Tabela 837)	2006
Econômicas	V18	Proporção de estabelecimento de agricultura não familiar no total de estabelecimentos agropecuários do município	FIBGE, Censo Agropecuário (Sidra, Tabelas 777 e 1109)	2006
	V19	Participação do valor adicionado bruto a preços correntes da agropecuária no valor adicionado bruto a preços correntes total (Percentual)	FIBGE em parceria com os Órgãos Estaduais de Estatística, Secretarias Estaduais de Governo e Superintendência da Zona Franca de Manaus - SUFRAMA	2006
	V20	Número médio de tratores por estabelecimento agropecuário	FIBGE, Censo Agropecuário (Sidra, Tabelas 777 e 860)	2006
	V21	Valor dos financiamentos obtidos pelos estabelecimentos agropecuários (em mil reais)	FIBGE, Censo Agropecuário (Sidra, Tabela 1893)	2006
	V22	Média de pessoal ocupado por estabelecimento agropecuário (em 31/12/2006)	FIBGE, Censo Agropecuário (Sidra, Tabelas 802 e 777)	2006
	V23	Renda monetária bruta dos estabelecimentos agropecuários (Mil Reais)	FIBGE, Censo Agropecuário (Sidra, Tabela 1258)	2006
Laranja	V24	Quantidade produzida (em toneladas) nos estabelecimentos com mais de 50 pés	FIBGE, Censo Agropecuário (Sidra, Tabela 1177)	2006
	V25	Proporção de área plantada de laranja no total de área ocupada pela agropecuária do município (em ha) ⁶	FIBGE, Censo Agropecuário (Sidra, Tabelas 837 e 1271)	2006
Cana-de-açúcar	V26	Quantidade produzida de cana-de-açúcar (em toneladas)	FIBGE, Censo Agropecuário (Sidra, Tabela 822)	2006
	V27	Proporção de área plantada de cana-de-açúcar no total de área ocupada pela agropecuária do município (em ha)	FIBGE, Censo Agropecuário (Sidra, Tabelas 837 e 1272)	2006
Pecuária	V28	Rebanho de bovinos	FIBGE, Censo Agropecuário (Sidra, Tabela 922)	2006
	V29	Proporção de estabelecimentos com pecuária de corte no total de estabelecimentos do município	FIBGE, Censo Agropecuário (Sidra, Tabelas 1262 e 777)	2006
	V30	Proporção de estabelecimentos agropecuários com até 10 ha de pastagens no total de estabelecimentos com pastagem do município	FIBGE, Censo Agropecuário (Sidra, Tabelas 777 e 783)	2006
	V31	Proporção de estabelecimentos agropecuários de 10 a 100 ha de pastagens no total de estabelecimentos com pastagem do município	FIBGE, Censo Agropecuário (Sidra, Tabelas 777 e 783)	2006
	V32	Proporção de estabelecimentos agropecuários de 100 a 200 ha de pastagens no total de estabelecimentos com pastagem do município	FIBGE, Censo Agropecuário (Sidra, Tabelas 777 e 783)	2006

Fonte: FIBGE, Censos Demográfico 2010 e Censo Agropecuário 2006.
¹(População rural/população total)*100
²(População de homens/população de mulheres)*100
³(População entre 0 e 14 anos/População de 60 anos ou mais)*100
⁴(População de 60 anos ou mais/população entre 0 e 14 anos) * 100
⁵(População que trabalha em município diferente do de residência que tem como atividade principal o trabalho na agricultura, pecuária, produção florestal, pesca e aquicultura/população que trabalha em município diferente do de residência)*100
⁶Total da área ocupada pela agropecuária: somatória da área ocupada pela lavoura temporária, horticultura e floricultura, lavoura permanente, produção de sementes, mudas e outras formas de propagação vegetal e pecuária e criação de outros animais

Em análise multivariada, o ponto de partida para o entendimento de qualquer variável é caracterizar a forma de sua distribuição (HAIR et al, 2005). Assim, a Figura 1 apresenta o histograma dos 32 indicadores considerados. Como é possível observar, é nítida a falta de normalidade em quase todas as distribuições. Apenas as variáveis V13 (proporção de nascidos no município que nunca migrou no total de população residente) e V18 (proporção de estabelecimento de agricultura não familiar no total de estabelecimentos agropecuários do município) apresentam distribuições que se aproximam da normal.

Figura 1
Histograma das variáveis selecionadas

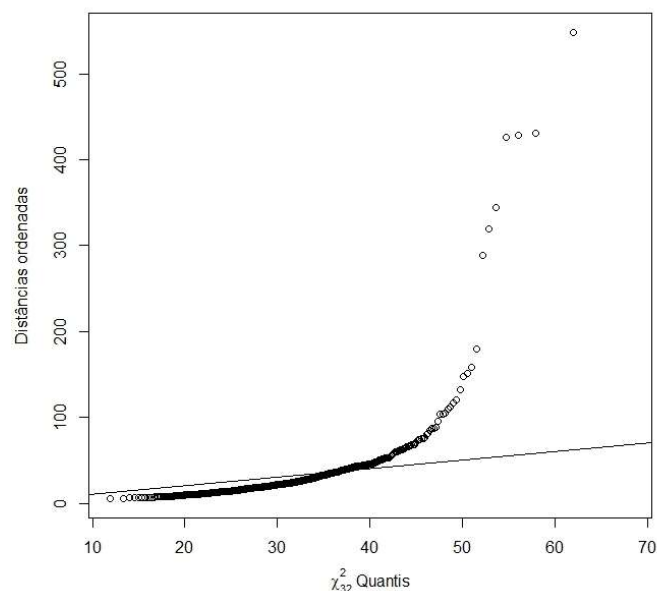


Fonte: FIBGE, Censo Demográfico 2010 e Censo Agropecuário 2006

O Gráfico 1 corrobora a falta de normalidade multivariada das informações. Quando a normalidade é coerente aos dados, esse gráfico deve resultar em algo próximo a uma reta (MINGOTI, 2005). Essa forma de representação gráfica também é útil na identificação de valores discrepantes (*outliers*). Ainda que a distribuição normal não seja

um pressuposto das análises multivariadas, sua verificação em muito colabora na aplicação de suas técnicas (MINGOTI, 2005).

Gráfico 1
Q-Q plot da distribuição



A Tabela 1 auxilia no entendimento da distribuição dos dados ao apresentar os valores mínimo, máximo, média e desvio padrão para cada informação. Vale lembrar que municípios sem população rural e sem estabelecimentos agropecuários (ou com menos de três estabelecimentos) receberam o valor zero para as estimativas construídas a partir de tais indicadores. A análise dessa tabela revela, ainda, uma inconsistência na variável proporção de área plantada de cana-de-açúcar no total da área ocupada pela agropecuária no município (V27), que ultrapassa 100%. Esse indicador – assim como a estimativa de área ocupada pela laranja (V25) – foi construído tendo como denominador a somatória das variáveis do Censo Agropecuário sobre tamanho de área ocupada com lavoura temporária, horticultura e floricultura, lavoura permanente, produção de sementes, mudas e outras formas de propagação vegetal, pecuária e criação de outros animais⁶. Esses dados são obtidos em campo, a partir das informações prestadas pelo entrevistado. Podem, portanto, não ser precisos (principalmente tendo em vista que uma mesma área pode ser ocupada por mais de um cultivo). Levando-se em consideração essa situação, acredita-se que a inconsistência relatada decorra desse fato.

⁶ Estimativas extraídas da Tabela 837 do Sidra. Foi desconsiderada a produção florestal proveniente tanto de florestas nativas como de florestas plantadas.

Tabela 1
Valores mínimos, valores máximos, média e desvio padrão das variáveis selecionadas
Estado de São Paulo, segundo municípios
2006 e 2010

Variáveis	N	Mínimo	Máximo	Média	Desvio Padrão
V01 População rural	645	0,00	101.159,00	2.599,92	5.302,31
V02 Grau de ruralização	645	0,00	75,09	15,68	14,28
V03 Proporção de população residente em situação de setor 4 (rural extensão-urbana) no total de população rural	645	0,00	74,89	2,20	10,01
V04 Proporção de população residente em situação de setor 8 (rural exclusive aglomerados) no total de população rural	645	0,00	100,00	90,36	24,07
V05 Proporção de população carcerária (população residente em setor especial tipo 6) no total de população rural	645	0,00	92,87	2,70	12,01
V06 Razão de sexo da população rural	645	0,00	2.721,35	127,34	145,89
V07 Razão de dependência da população rural	645	0,00	93,10	52,56	15,31
V08 Índice de envelhecimento da população rural	645	0,00	218,18	59,16	32,02
V09 Pendularidade rural	645	0,00	67,12	14,01	13,72
V10 Proporção de trabalhadores residentes no urbano e que estavam ocupados no setor primário	645	0,14	46,25	10,63	8,96
V11 Proporção de trabalhadores residentes no rural e que estavam ocupados nos setores secundário e terciário	645	0,00	99,49	41,65	20,56
V12 Proporção de migrantes que chegaram nos últimos três anos no total da população não natural	645	9,08	77,36	25,12	7,49
V13 Proporção de nascidos no município que nunca migrou no total de população residente	645	13,17	80,28	49,56	11,38
V14 Proporação de estabelecimento de até 10 ha no total de estabelecimentos agropecuário do município	645	0,00	100,00	34,66	18,38
V15 Proporação de estabelecimento de 10 a 100 ha no total de estabelecimentos agropecuário do município	645	0,00	90,27	47,36	16,00
V16 Proporação de estabelecimento de 100 a 1000 ha no total de estabelecimentos agropecuário do município	645	0,00	100,00	13,32	9,86
V17 Proporação de estabelecimento com mais de 1000 ha no total de estabelecimentos agropecuário do município	645	0,00	16,67	1,30	1,93
V18 Proporção de estabelecimento de agricultura não familiar no total de estabelecimentos agropecuários do municípios	645	0,00	87,50	37,70	15,62
V19 Participação do valor adicionado bruto a preços correntes da agropecuária no valor adicionado bruto a preços correntes total (Percentual)	645	0,00	61,15	16,70	13,17
V20 Número médio de tratores por estabelecimento agropecuário	645	0,00	14,69	0,77	0,85
V21 Valor dos financiamentos obtidos pelos estabelecimentos agropecuários (em mil reais)	645	0,00	899.013,00	5.787,56	38.592,72
V22 Média de pessoal ocupado por estabelecimento agropecuário (em 31/12/2006)	645	0,00	243,94	5,55	13,88
V23 Renda monetária bruta dos estabelecimentos agropecuários (Mil Reais)	645	0,00	883.397,00	41.415,65	59.582,47
V24 Quantidade produzida de laranja (em toneladas) nos estabelecimentos com mais de 50 pés	645	0,00	463.711,00	15.554,30	45.483,78
V25 Proporção de área plantada de laranja no total de área ocupada pela agropecuária do município (em ha)	645	0,00	62,77	3,23	7,69
V26 Quantidade produzida de cana-de-acucar (em toneladas)	645	0,00	7.248.703,00	370.879,63	780.143,35
V27 Proporção de área plantada de cana-de-açúcar no total de área ocupada pela agropecuária do município (em ha)	645	0,00	131,52	15,73	22,69
V28 Rebanho de bovinos	645	0,00	126.676,00	16.288,26	18.633,73
V29 Proporção de estabelecimentos com pecuária de corte no total de estabelecimentos do município	645	0,00	47,17	11,27	9,70
V30 Proporção de estabelecimentos agropecuários com até 10 ha de pastagens no total de estabelecimentos com pastagem do município	645	0,00	83,72	31,26	21,11
V31 Proporção de estabelecimentos agropecuários de 10 a 100 ha de pastagens no total de estabelecimentos com pastagem do município	645	0,00	88,96	43,07	24,65
V32 Proporção de estabelecimentos agropecuários de 100 a 200 ha de pastagens no total de estabelecimentos com pastagem do município	645	0,00	37,50	6,14	5,94

Fonte: FIBGE, Censo Demográfico 2010; FIBGE, Censo Agropecuário 2006.

A despeito dessa inconsistência, optou-se por manter tais indicadores (V25 e V27), na medida em que se tratam de uma importante *proxy* do grau de especialização produtiva, nos diferentes municípios, quanto a essas atividades. Desde essa perspectiva, pode-se supor que as localidades onde a proporção de área plantada de cana-de-açúcar foi maior que toda a área ocupada pela agropecuária no município destacam-se pela alta especialização produtiva e quase monocultura desse produto⁷.

Além de uma análise da distribuição, é fundamental ressaltar a forma com que as diferentes variáveis se relacionam. A Tabela 2 apresenta, portanto, os coeficientes de correlação de Pearson. O simples exame das correlações já é capaz de apontar para diferentes arranjos da população rural. A correlação positiva entre população carcerária (V05), razão de sexo (V06) e migrantes recém-chegados (V12), por exemplo, é indicativo de uma importante dimensão da dinâmica da população rural do Estado de São Paulo: o rural dos presídios. As especificidades dessa mobilidade são discutidas por Cescon (2012), em pesquisa que destaca os impactos locais da construção de unidades prisionais em pequenos municípios do Oeste Paulista.

Outra associação interessante é identificada a partir das correlações positivas entre população rural (V01) e proporção de população residente em setores rurais-expansão urbana (V03). Esse último indicador também está correlacionado positivamente com proporção de trabalhadores residentes no rural, mas ocupados nos serviços e na indústria (V11) (Tabela 2). Esse é o rural das novas atividades agrícolas (SILVA, 1999), típico de regiões densamente habitadas, com problemas flagrantes de classificação de situação de domicílio, que vão desde a não atualização do perímetro urbano até questões fiscais de arrecadação de impostos (RODRIGUES, 2001; CUNHA, 2005).

Por outro lado, grau de ruralização (V02) apresenta correlação positiva com participação do PIB agropecuário no PIB total do município (V19). Esse indicador, por sua vez, também registra correlação com pendularidade rural (V09) e estabelecimentos de médio porte (entre 10 e 100 ha) (V15) (Tabela 2). Tendo em vista a importância das atividades agropecuárias na dinâmica da população e na geração de divisas locais, esse rural pode ser pensado como estruturado por municípios cujas sedes, embora definidas oficialmente como urbanas, remetam às cidades imaginárias de Veiga (2003).

⁷ Esses municípios são: Dobrada, Dumont e Macatuba, todos na porção centro leste do Estado, importante região canavieira de São Paulo (PERES; BAENINGER, 2013).

Tabela 2
Matriz de correlação de Pearson

Variáveis	Demográficas													Estrutura Fundiária				Econômicas					Laranja		Cana-de-açúcar		Pecuária						
	V01	V02	V03	V04	V05	V06	V07	V08	V09	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	V21	V22	V23	V24	V25	V26	V27	V28	V29	V30	V31	V32	
Demográficas	V01	1,00	0,13	0,36	-0,08	0,05	0,02	0,06	-0,09	-0,15	-0,22	0,29	-0,10	0,18	0,29	-0,16	-0,16	-0,10	-0,07	-0,19	-0,04	-0,01	-0,04	0,11	0,02	-0,05	-0,04	-0,13	0,06	-0,13	0,13	-0,04	-0,06
	V02	0,13	1,00	-0,07	0,21	0,19	0,26	0,21	0,19	0,13	0,09	0,06	0,44	0,04	-0,03	0,21	-0,04	-0,11	-0,19	0,36	-0,22	-0,07	-0,12	-0,08	-0,11	-0,10	-0,23	-0,30	0,11	0,13	0,15	0,32	0,12
	V03	0,36	-0,07	1,00	-0,37	-0,01	-0,02	0,01	-0,17	-0,13	-0,17	0,35	-0,12	-0,05	0,15	-0,15	0,03	-0,07	-0,01	-0,19	-0,02	0,00	-0,01	0,02	0,00	-0,02	-0,01	-0,05	-0,06	-0,07	-0,01	-0,11	0,04
	V04	-0,08	0,21	-0,37	1,00	0,05	0,18	0,64	0,45	0,25	0,29	0,16	0,15	0,25	-0,10	0,47	0,19	0,13	0,22	0,34	0,13	0,02	0,05	0,12	0,07	0,03	0,09	0,16	0,20	0,25	0,25	0,36	0,19
	V05	0,05	0,19	-0,01	0,05	1,00	0,76	-0,46	0,15	0,04	0,01	0,13	0,39	-0,15	-0,04	0,02	0,09	0,08	0,07	0,04	-0,03	-0,01	-0,03	-0,01	0,01	-0,05	0,00	-0,04	0,14	0,19	-0,01	0,09	0,09
	V06	0,02	0,26	-0,02	0,18	0,76	1,00	-0,22	0,22	0,08	0,07	0,17	0,45	-0,12	-0,08	0,11	0,15	0,08	0,12	0,07	0,00	-0,01	-0,01	-0,01	0,00	-0,03	0,00	-0,01	0,07	0,25	0,00	0,14	0,17
	V07	0,06	0,21	0,01	0,64	-0,46	-0,22	1,00	0,29	0,12	0,14	0,27	-0,07	0,31	0,02	0,36	0,07	0,03	0,10	0,18	0,06	0,03	0,04	0,09	0,00	0,00	0,02	0,07	0,05	0,02	0,20	0,22	0,10
	V08	-0,09	0,19	-0,17	0,45	0,15	0,22	0,29	1,00	0,14	0,17	0,06	0,10	-0,05	-0,06	0,43	-0,13	-0,18	-0,16	0,28	-0,11	0,01	-0,08	-0,04	0,01	0,05	-0,03	0,00	0,21	0,22	0,18	0,36	0,00
	V09	-0,15	0,13	-0,13	0,25	0,04	0,08	0,12	0,14	1,00	0,80	-0,29	0,15	0,12	-0,24	0,29	0,17	0,12	0,11	0,48	0,08	-0,01	-0,01	0,08	0,14	0,17	0,04	0,07	0,06	0,19	0,10	0,17	0,05
	V10	-0,22	0,09	-0,17	0,29	0,01	0,07	0,14	0,17	0,80	1,00	-0,38	0,19	0,09	-0,25	0,32	0,19	0,12	0,18	0,63	0,10	-0,01	0,01	0,03	0,11	0,24	0,00	0,11	-0,02	0,20	0,04	0,14	0,10
	V11	0,29	0,06	0,35	0,16	0,13	0,17	0,27	0,06	-0,29	-0,38	1,00	0,01	-0,08	0,33	-0,08	-0,08	0,01	0,08	-0,31	0,07	-0,02	0,12	-0,04	-0,10	-0,10	0,00	0,06	-0,11	-0,12	0,02	-0,09	-0,01
	V12	-0,10	0,44	-0,12	0,15	0,39	0,45	-0,07	0,10	0,15	0,19	0,01	1,00	-0,16	-0,14	0,16	0,13	0,05	0,01	0,30	-0,12	-0,06	-0,08	-0,10	-0,12	-0,10	-0,08	-0,07	0,05	0,22	0,02	0,21	0,14
	V13	0,18	0,04	-0,05	0,25	-0,15	-0,12	0,31	-0,05	0,12	0,09	-0,08	-0,16	1,00	-0,04	0,09	0,11	0,02	0,07	0,03	0,06	0,06	-0,02	0,22	0,13	-0,01	0,21	0,13	0,12	-0,05	0,24	0,11	0,06
Estrutura Fundiária	V14	0,29	-0,03	0,15	-0,10	-0,04	-0,08	0,02	-0,06	-0,24	-0,25	0,33	-0,14	-0,04	1,00	-0,54	-0,53	-0,28	-0,22	-0,35	-0,09	0,01	0,00	-0,07	-0,14	-0,07	-0,15	-0,14	-0,24	-0,41	0,18	-0,38	-0,31
	V15	-0,16	0,21	-0,15	0,47	0,02	0,11	0,36	0,43	0,29	0,32	-0,08	0,16	0,09	-0,54	1,00	0,13	0,01	0,15	0,41	0,01	0,00	-0,06	0,09	0,14	0,13	0,10	0,14	0,33	0,32	0,09	0,53	0,15
	V16	-0,16	-0,04	0,03	0,19	0,09	0,15	0,07	-0,13	0,17	0,19	-0,08	0,13	0,11	-0,53	0,13	1,00	0,48	0,59	0,26	0,28	-0,01	0,14	0,13	0,09	0,00	0,18	0,15	0,15	0,49	-0,12	0,26	0,54
	V17	-0,10	-0,11	-0,07	0,13	0,08	0,08	0,03	-0,18	0,12	0,12	0,01	0,05	0,02	-0,28	0,01	0,48	1,00	0,45	0,15	0,38	0,07	0,25	0,19	0,15	0,05	0,30	0,27	0,08	0,25	-0,14	0,00	0,20
Econômicas	V18	-0,07	-0,19	-0,01	0,22	0,07	0,12	0,10	-0,16	0,11	0,18	0,08	0,01	0,07	-0,22	0,15	0,59	0,45	1,00	0,13	0,44	0,01	0,22	0,18	0,19	0,20	0,21	0,25	-0,08	0,22	-0,10	0,01	0,31
	V19	-0,19	0,36	-0,19	0,34	0,04	0,07	0,18	0,28	0,48	0,63	-0,31	0,30	0,03	-0,35	0,41	0,26	0,15	0,13	1,00	0,04	-0,03	-0,05	0,10	0,17	0,22	0,04	0,10	0,09	0,29	0,06	0,31	0,20
	V20	-0,04	-0,22	-0,02	0,13	-0,03	0,00	0,06	-0,11	0,08	0,10	0,07	-0,12	0,06	-0,09	0,01	0,28	0,38	0,44	0,04	1,00	0,07	0,85	0,21	0,18	0,19	0,30	0,46	-0,15	-0,06	-0,07	-0,21	-0,05
	V21	-0,01	-0,07	0,00	0,02	-0,01	-0,01	0,03	0,01	-0,01	-0,01	-0,02	-0,06	0,06	0,01	0,00	-0,01	0,07	0,01	-0,03	0,07	1,00	0,05	0,15	0,04	0,00	0,23	0,17	0,01	-0,03	0,07	-0,06	-0,04
	V22	-0,04	-0,12	-0,01	0,05	-0,03	-0,01	0,04	-0,08	-0,01	0,01	0,12	-0,08	-0,02	0,00	-0,06	0,14	0,25	0,22	-0,05	0,85	0,05	1,00	0,06	0,02	0,04	0,13	0,28	-0,10	-0,08	-0,12	-0,15	-0,08
V23	0,11	-0,08	0,02	0,12	-0,01	-0,01	0,09	-0,04	0,08	0,03	-0,04	-0,10	0,22	-0,07	0,09	0,13	0,19	0,18	0,10	0,21	0,15	0,06	1,00	0,39	0,18	0,62	0,34	0,15	0,00	0,22	0,06	-0,01	
Laranja	V24	0,02	-0,11	0,00	0,07	0,01	0,00	0,00	0,01	0,14	0,11	-0,10	-0,12	0,13	-0,14	0,14	0,09	0,15	0,19	0,17	0,18	0,04	0,02	0,39	1,00	0,63	0,25	0,13	0,03	-0,03	0,17	0,00	-0,06
	V25	-0,05	-0,10	-0,02	0,03	-0,05	-0,03	0,00	0,05	0,17	0,24	-0,10	-0,10	-0,01	-0,07	0,13	0,00	0,05	0,20	0,22	0,19	0,00	0,04	0,18	0,63	1,00	0,06	0,09	-0,13	-0,14	0,04	-0,16	-0,16
Cana-de-açúcar	V26	-0,04	-0,23	-0,01	0,09	0,00	0,00	0,02	-0,03	0,04	0,00	0,00	-0,08	0,21	-0,15	0,10	0,18	0,30	0,21	0,04	0,30	0,23	0,13	0,62	0,25	0,06	1,00	0,67	0,02	-0,04	0,12	-0,06	-0,04
	V27	-0,13	-0,30	-0,05	0,16	-0,04	-0,01	0,07	0,00	0,07	0,11	0,06	-0,07	0,13	-0,14	0,14	0,15	0,27	0,25	0,10	0,46	0,17	0,28	0,34	0,13	0,09	0,67	1,00	-0,20	-0,15	0,04	-0,23	-0,10
Pecuária	V28	0,06	0,11	-0,06	0,20	0,14	0,07	0,05	0,21	0,06	-0,02	-0,11	0,05	0,12	-0,24	0,33	0,15	0,08	-0,08	0,09	-0,15	0,01	-0,10	0,15	0,03	-0,13	0,02	-0,20	1,00	0,51	0,02	0,55	0,28
	V29	-0,13	0,13	-0,07	0,25	0,19	0,25	0,02	0,22	0,19	0,20	-0,12	0,22	-0,05	-0,41	0,32	0,49	0,25	0,22	0,29	-0,06	-0,03	-0,08	0,00	-0,03	-0,14	-0,04	-0,15	0,51	1,00	-0,04	0,62	0,57
	V30	0,13	0,15	-0,01	0,25	-0,01	0,00	0,20	0,18	0,10	0,04	0,02	0,02	0,24	0,18	0,09	-0,12	-0,14	-0,10	0,06	-0,07	0,07	-0,12	0,22	0,17	0,04	0,12	0,04	0,02	-0,04	1,00	0,30	0,06
	V31	-0,04	0,32	-0,11	0,36	0,09	0,14	0,22	0,36	0,17	0,14	-0,09	0,21	0,11	-0,38	0,53	0,26	0,00	0,01	0,31	-0,21	-0,06	-0,15	0,06	0,00	-0,16	-0,06	-0,23	0,55	0,62	0,30	1,00	0,53
	V32	-0,06	0,12	0,04	0,19	0,09	0,17	0,10	0,00	0,05	0,10	-0,01	0,14	0,06	-0,31	0,15	0,54	0,20	0,31	0,20	-0,05	-0,04	-0,08	-0,01	-0,06	-0,16	-0,04	-0,10	0,28	0,57	0,06	0,53	1,00

Correlação moderada para franca (entre ±0,3 e ±0,5) (HAIR et al, 2005)

Correlação moderada (entre ±0,5 e ±0,7) (HAIR et al, 2005)

Correlação alta (acima de ±0,7) (HAIR et al, 2005)

A proporção de população residente em setores rural-exclusive aglomerados (V04) registra correlação positiva com índice de envelhecimento (V08), proporção de estabelecimentos entre 10 e 100 ha (V15), PIB agropecuário (V19) e estabelecimentos com 10 a 100 ha de pastagens (V31). Em relação à população residente em setores rural-expansão urbana, esse indicador tem correlação negativa (Tabela 2). Essas interações reforçam a tendência apontada pela bibliografia de masculinização das populações rurais que sobrevivem primordialmente da atividade agrícola (CAMARANO; ABRAMOVAY, 1997).

A agricultura não familiar (V18) destaca-se pelas correlações positivas com proporção de estabelecimentos entre 100 e 1000 ha (V16), com mais de 1000 ha (V15), média de tratores (V20) e estabelecimentos com 100 a 200 ha de pastagens (V32) (Tabela 2). No tocante à presença de mega latifúndios (V17), vale destacar a correlação com produção de cana-de-açúcar (V26). A proporção de área plantada dessa cultura (V27) está correlacionada, por sua vez, com renda monetária dos estabelecimentos (V23) e média de tratores (V20). Tais evidências reforçam o processo de concentração de riqueza que vem no reboque da expansão dos canaviais (CASTILLO, 2015).

Assim como a cana, a produção de laranja (V24) também tem correlação positiva com renda monetária dos estabelecimentos (V23) (Tabela 2). Contudo, a existência de não correlação desse indicador com estabelecimentos de grandes proporções (mais de 100 ha) pode ser indicativo da não necessária relação entre esse cultivo e o processo de concentração fundiária.

Por fim, rebanho de bovinos (V28) tem correlação positiva com estabelecimentos entre 10 e 100 ha de pastagens (V31). A proporção de estabelecimentos com pecuária de corte (V29) está correlacionada positivamente com estabelecimentos de 10 a 100 ha (V15) e de 100 a 1000 ha (V16), e negativamente com estabelecimentos de até 10 ha (Tabela 2).

3. A Análise Fatorial

Todos os procedimentos executados nessa seção foram feitos no SPSS, versão 22. Na análise fatorial, o primeiro passo é o exame das correlações (HAIR et al, 2005). Se o propósito da técnica é condensar variáveis em um conjunto menor de dimensões (fatores), no caso dos dados não serem correlacionados, a variação de uma informação não implica na variação da outra e, portanto, não é possível identificar um fator comum que descreva

a variabilidade de ambas as estimativas. Por essa razão, Hair et al (2005) sugerem a exclusão de variáveis que não tenham correlações substanciais com as demais informações do banco (abaixo de $\pm 0,3$). Diante desse pressuposto, a variável V21 (valor dos financiamentos obtidos pelos estabelecimentos agropecuários) foi removida (Tabela 2).

Uma forma de quantificar o grau de intercorrelações entre as variáveis e a adequação da análise fatorial é através do teste de Kaiser-Meyer-Olkin (KMO) (CORRAR, PAULO, DIAS FILHO, 2011; HAIR et al, 2005). Esse índice varia entre 0 e 1, “alcançando 1 quando cada variável é perfeitamente prevista sem erro pelas outras variáveis” (HAIR et al, 2005, p.98). Vale constar que as correlações médias aumentam quando o tamanho da amostra aumenta. Por essa razão, o teste KMO também se trata de uma medida de adequação da amostra (HAIR et al, 2005). Para os dados selecionados (sem V21), esse índice foi de 0,700, o que representa adequação mediana dos dados à técnica⁸.

Depois de avaliadas as correlações, decisões devem ser tomadas quanto ao método de extração dos fatores e ao número de dimensões selecionadas para representar as informações originais (HAIR et al, 2005). A seleção do método de extração depende do objetivo da pesquisa. Se o propósito é apenas resumir as informações originais a um mínimo de fatores, o autor indica a análise de componentes principais (esse é o caso dessa pesquisa). Se, por outro lado, o objetivo é construir novas variáveis que, além de sintetizar as informações originais, também expressam o que elas têm em comum, é sugerida a análise dos fatores comuns (HAIR et al, 2005).

Quanto à escolha do número ideal de fatores necessário para explicar o maior montante possível da variância original dos dados, Hair et al (2005) e Rogerson (2012) discutem diferentes possibilidades. “(...) escolher o número de fatores a serem interpretados é como focalizar um microscópio. Um ajuste muito alto ou muito baixo irá obscurecer uma estrutura que é óbvia quando o ajuste está correto” (HAIR et al, 2005, p.101). Para os autores, a regra mais convencional é dos autovalores maiores que 1. O raciocínio desse critério fundamenta-se no pressuposto de que os componentes criados devem explicar a variância de pelo menos uma variável original. “Logo, apenas os fatores que têm **raízes latentes** ou **autovalores** maiores que 1 são considerados significantes”

⁸ Segundo Hair et al (2005, p.98), o teste KMO pode ser interpretado a partir das seguintes orientações: “0,8 ou acima, admirável”; entre 0,7 e 0,9, “mediano”; entre 0,6 e 0,7, “mediocre”; entre 0,4 e 0,5 ruim; “e abaixo de 0,5, inaceitável”

(HAIR et al, 2005, p.101). Nesse caso, são construídos 10 componentes que condensam 74% da variância acumulada das informações iniciais (Quadro 2).

Quadro 2
Variância total explicada

Componente	Valores próprios iniciais			Somadas de extração de carregamentos ao quadrado			Somadas rotativas de carregamentos ao quadrado		
	Total	% de variância	% cumulativa	Total	% de variância	% cumulativa	Total	% de variância	% cumulativa
1	5,151	16,616	16,616	5,151	16,616	16,616	2,885	9,305	9,305
2	3,767	12,152	28,768	3,767	12,152	28,768	2,878	9,284	18,589
3	2,702	8,717	37,485	2,702	8,717	37,485	2,758	8,896	27,485
4	2,360	7,613	45,098	2,360	7,613	45,098	2,486	8,018	35,503
5	2,114	6,818	51,916	2,114	6,818	51,916	2,430	7,837	43,341
6	1,827	5,895	57,811	1,827	5,895	57,811	2,342	7,556	50,897
7	1,517	4,894	62,706	1,517	4,894	62,706	1,975	6,372	57,269
8	1,320	4,257	66,963	1,320	4,257	66,963	1,837	5,926	63,195
9	1,112	3,588	70,551	1,112	3,588	70,551	1,690	5,452	68,647
10	1,052	3,392	73,943	1,052	3,392	73,943	1,642	5,297	73,943
11	,950	3,064	77,007						
12	,875	2,822	79,829						
13	,788	2,544	82,373						
14	,575	1,855	84,227						
15	,538	1,735	85,962						
16	,529	1,706	87,668						
17	,461	1,487	89,155						
18	,414	1,335	90,490						
19	,390	1,259	91,749						
20	,336	1,084	92,833						
21	,319	1,030	93,863						
22	,304	,981	94,844						
23	,294	,950	95,794						
24	,270	,872	96,666						
25	,217	,701	97,367						
26	,184	,594	97,961						
27	,174	,562	98,523						
28	,139	,447	98,970						
29	,134	,433	99,402						
30	,098	,318	99,720						
31	,087	,280	100,000						

Método de Extração: Análise de Componente Principal.

O próximo passo é examinar as correlações entre os fatores criados e as variáveis originais. “Esse é um passo chave na análise, pois é onde o ‘significado’ e a interpretação de cada fator ocorre” (ROGERSON, 2012, p.300). Cada ‘loading’ ou carga fatorial representa a correlação linear entre uma componente e a variável original. Por meio dessa relação define-se a importância de cada variável na construção dos fatores (HAIR et al, 2005). “O primeiro fator pode ser visto como o melhor resumo de relações lineares registrada pelos dados originais. O segundo fator é definido como a segunda melhor combinação linear das variáveis”, e assim por diante (HAIR et al, 2005, p.103).

Para facilitar a interpretação dessas cargas, a matriz fatorial é rotacionada de modo que cada fator tende a ser aumentado (próximos de ± 1) ou reduzido (próximos a zero) (ROGERSON, 2012). O critério rotacional recomendado pelos autores (VARIMAX) maximiza a soma das variâncias das cargas. A lógica é que a interpretação se torna mais fácil quando as correlações variável-fator são próximas de ± 1 , indicando assim uma clara

associação positiva ou negativa entre a variável e o fator, ou próximas a 0, apontando para uma nítida falta de associação entre elas (HAIR et al, 2005). Uma carga fatorial deve exceder o valor de 0,7 para que o fator explique 50% da variância da informação original (HAIR et al, 2005). Contudo, cargas maiores que $\pm 0,3$ já apresentam significância mínima, enquanto valores na casa de $\pm 0,4$ são considerados mais importantes e superiores a $\pm 0,5$ apresentam significância prática (HAIR et al, 2005). O Quadro 3 apresenta a matriz fatorial rotacionada.

Quadro 3
Matriz de componente rotativa^a

	Componente									
	1	2	3	4	5	6	7	8	9	10
V01	-0,16	0,00	-0,13	0,06	-0,06	-0,08	0,07	0,02	0,40	0,65
V02	0,32	0,13	-0,13	0,38	0,25	-0,30	-0,11	-0,19	0,23	0,16
V03	-0,09	-0,04	0,04	-0,07	-0,10	0,04	-0,06	0,01	-0,15	0,88
V04	0,15	0,14	0,18	0,12	0,76	0,04	0,10	0,02	0,30	-0,27
V05	-0,05	0,12	0,02	0,89	-0,14	0,03	0,03	0,05	-0,04	0,00
V06	0,00	0,08	0,11	0,88	0,07	0,01	0,01	0,03	-0,04	0,01
V07	0,12	0,00	0,11	-0,36	0,79	-0,02	0,02	-0,08	0,26	0,12
V08	0,10	0,32	-0,35	0,22	0,60	0,01	-0,02	0,10	-0,09	-0,15
V09	0,83	0,05	0,04	0,04	0,02	0,03	0,06	0,07	0,11	-0,04
V10	0,88	-0,02	0,11	0,02	0,08	-0,01	0,05	0,09	0,02	-0,12
V11	-0,47	-0,23	0,08	0,21	0,48	-0,01	0,09	-0,08	-0,03	0,47
V12	0,30	0,00	0,09	0,63	0,10	-0,08	-0,14	-0,22	-0,05	-0,05
V13	0,10	0,06	0,10	-0,23	0,06	0,18	0,02	-0,04	0,66	0,00
V14	-0,34	-0,50	-0,35	0,05	0,00	-0,23	0,05	-0,05	0,35	0,15
V15	0,34	0,49	-0,02	0,01	0,54	0,19	-0,06	0,13	-0,18	-0,05
V16	0,16	0,24	0,83	0,04	-0,02	0,10	0,07	-0,02	-0,06	0,00
V17	0,06	0,03	0,61	0,07	-0,06	0,24	0,27	0,04	-0,02	-0,07
V18	0,03	-0,10	0,77	0,05	0,12	0,12	0,19	0,23	-0,03	-0,01
V19	0,75	0,16	0,12	0,11	0,21	0,03	-0,06	0,11	-0,02	-0,09
V20	0,05	-0,12	0,25	-0,04	0,04	0,24	0,87	0,13	-0,02	-0,01
V22	-0,02	-0,05	0,08	-0,03	0,02	0,06	0,94	-0,03	-0,06	0,01
V23	0,02	0,14	0,06	-0,01	-0,05	0,66	0,04	0,27	0,36	0,11
V24	0,07	0,06	0,08	-0,02	-0,03	0,18	0,02	0,85	0,16	0,01
V25	0,20	-0,16	0,02	-0,05	0,05	0,00	0,05	0,87	-0,07	-0,01
V26	-0,02	0,01	0,12	0,00	-0,01	0,90	0,06	0,05	0,13	-0,01
V27	0,08	-0,23	0,12	-0,03	0,15	0,80	0,23	-0,03	-0,08	-0,08
V28	-0,05	0,83	-0,03	0,04	-0,01	0,00	0,02	-0,02	0,13	0,00
V29	0,13	0,69	0,42	0,18	0,06	-0,11	-0,06	-0,08	-0,08	-0,05
V30	0,03	0,06	-0,14	0,07	0,20	0,08	-0,14	0,12	0,65	0,00
V31	0,14	0,77	0,13	0,10	0,28	-0,11	-0,16	-0,08	0,18	-0,03
V32	0,03	0,43	0,63	0,08	0,07	-0,15	-0,17	-0,15	0,09	0,05

Método de Extração: Análise de Componente Principal.

a. Rotação convergida em 8 iterações.

Valores das cargas

Superior a $\pm 0,7$	Entre $\pm 0,4$ e $\pm 0,5$
Entre $\pm 0,5$ e $\pm 0,7$	Entre $\pm 0,3$ e $\pm 0,4$

Uma forma de avaliar qual a quantia de variância explicada pela solução fatorial para cada variável original é por meio da avaliação das comunalidades (HAIR et al, 2005). Essa estimativa serve para “identificar variáveis que não foram sublinhadas e que, portanto, não carregam sobre qualquer fator” (HAIR et al, 2005, p.108). Para que pelo menos a metade da variância de cada informação seja levada em conta pelos fatores criados, a comunalidade de cada variável não pode ser inferior a 0,5. Para nenhum dos indicadores consideradas, a comunalidade foi menor que isso (Quadro 4).

Quadro 4															
Comunalidades															
V01	0,65	V05	0,84	V09	0,71	V13	0,56	V17	0,52	V22	0,91	V26	0,85	V30	0,53
V02	0,56	V06	0,79	V10	0,82	V14	0,68	V18	0,73	V23	0,68	V27	0,80	V31	0,80
V03	0,83	V07	0,87	V11	0,79	V15	0,75	V19	0,68	V24	0,80	V28	0,72	V32	0,67
V04	0,84	V08	0,69	V12	0,59	V16	0,79	V20	0,92	V25	0,83	V29	0,74		

Método de Extração: Análise de Componente Principal.

A partir dos valores das cargas fatoriais expressos no Quadro 3, o Quadro 5 apresenta a importância de cada variável original na composição dos 10 fatores criados. Essa análise subsidiou a qualificação de cada fator, de modo a refinar as associações entre as variáveis originais discutidas no item dois. Em linhas gerais, a análise fatorial facilita a compreensão da heterogeneidade do rural paulista, que vai desde o rural da produção de commodities (no caso laranja, cana e carne), passa pelos presídios, expansão urbana e uma série de outros arranjos, dentre os quais se destaca áreas de pastagens subutilizadas, característica recorrente da modernização agrícola conservadora brasileira, subjugada à especulação imobiliária (WANDERLEY, 2011).

Vale destacar a correlação negativa entre o fator 6 – rural da cana-de-açúcar – e a variável grau de ruralização, reiterando os fatores de expulsão associados à expansão dessa atividade (SINGER, 1980; CASTILLO, 2015). Além disso, reforça o alto nível de urbanização demandado pelo agronegócio globalizado (ELIAS, 2003) e a não necessidade de população no rural desse arranjo para a dinamização dessa atividade. Na interpretação dos diferentes fatores, também fica evidente a heterogeneidade do rural agrícola baseado na pequena e média propriedade (fatores 1, 5 e 9), marcado por diferentes dinâmicas de redistribuição de população.

Antes de passar à análise de agrupamentos, é necessário destacar que os fatores criados representam “uma função forte dos dados de entrada” (ROGERSON, 2012, p.302). Ou seja, os componentes “que emergem da análise fatorial não são necessariamente os ‘mais importantes’, mas sim aqueles que captam a natureza do conjunto de dados” originais (ROGERSON, 2012, p.12). Essa observação ressalta a centralidade do processo de escolha das variáveis, etapa profundamente articulada aos interesses da pesquisa.

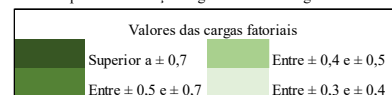
Em vista do objetivo do trabalho – captar a heterogeneidade do rural paulista a partir da dinâmica de suas principais commodities (cana, laranja e carne bovina) –, o entendimento do rural não estruturado por tais atividades fica comprometido pela não adequação das variáveis selecionadas originalmente. Essa limitação torna-se evidente na discussão dos agrupamentos construídos no item a seguir.

Quadro 5
Interpretação dos fatores

Fator 1 variância explicada: 16,6%	Rural agrícola: média propriedade		
	V09	Pendularidade rural	
	V10	Proporção de trabalhadores residentes no urbano e que estavam ocupados no setor primário	
	V19	Participação do PIB agropecuária no PIB total do município	
	V11*	Proporção de trabalhadores residentes no rural e ocupados nos setores secundário e terciário	
	V02	Grau de ruralização	
	V12	Proporção de migrantes recentes	
	V14*	Proporção de estabelecimento de até 10 ha	
	V15	Proporção de estabelecimento de 10 a 100 ha	
	Fator 2 variância explicada: 12,1%	Rural da pecuária	
		V28	Rebanho de bovinos
		V31	Proporção de estabelecimentos agropecuários de 10 a 100 ha de pastagens
		V29	Proporção de estabelecimentos com pecuária de corte
		V14*	Proporção de estabelecimento de até 10 ha
		V15	Proporção de estabelecimento de 10 a 100 ha
V32		Proporção de estabelecimentos agropecuários de 100 a 200 ha de pastagens	
Fator 3 Variância explicada: 8,7%	Rural das pastagens/especulação		
	V16	Proporção de estabelecimento de 100 a 1000 ha	
	V18	Proporção de estabelecimento de agricultura não familiar	
	V17	Proporção de estabelecimento com mais de 1000 ha	
	V32	Proporção de estabelecimentos agropecuários de 100 a 200 ha de pastagens	
Fator 4 variância explicada: 7,6%	Rural dos presídios		
	V05	Proporção de população carcerária	
	V06	Razão de sexo da população rural	
	V12	Proporção de migrantes recentes	
	V02	Grau de ruralização	
V07*	Razão de dependência da população rural		

Fator 5 variância explicada: 6,1%	Rural em avançado processo de envelhecimento	
	V04	Proporção de população residente em situação de setor 8 (rural exclusive-aglomerados)
	V07	Razão de dependência da população rural
	V08	Índice de envelhecimento da população rural
	V15	Proporção de estabelecimento de 10 a 100 ha
	V11	Proporção de trabalhadores residentes no rural e ocupados nos setores secundário e terciário
Fator 6 variância explicada: 5,8%	Rural da cana-de-açúcar	
	V26	Quantidade produzida de cana-de-açúcar (em toneladas)
	V27	Proporção de área plantada de cana-de-açúcar no total de área ocupada pela agropecuária do município (em ha)
	V23	Renda monetária bruta dos estabelecimentos agropecuários (Mil Reais)
Fator 7 variância explicada: 4,8%	Rural altamente capitalizado	
	V20	Número médio de tratores por estabelecimento agropecuário
	V22	Média de pessoal ocupado por estabelecimento agropecuário (em 31/12/2006)
Fator 8 variância explicada: 4,25%	Rural da laranja	
	V24	Quantidade produzida de laranja (em toneladas) nos estabelecimentos com mais de 50 pés
	V25	Proporção de área plantada de laranja no total de área ocupada pela agropecuária do município (em ha)
Fator 9 variância explicada: 3,5%	Rural agrícola: pequena propriedade	
	V13	Proporção de nascidos no município que nunca migrou no total de população residente
	V30	Proporção de estabelecimentos agropecuários com até 10 ha de pastagens
	V01	População rural
	V04	Proporção de população residente em situação de setor 8
	V14	Proporção de estabelecimento de até 10 ha
Fator 10 variância explicada: 3,3%	Rural expansão urbana	
	V03	Proporção de população residente em situação de setor 4 (rural extensão-urbana)
	V01	População rural
	V11	Proporção de trabalhadores residentes no rural e que estavam ocupados nos setores secundário e terciário

Nota: Variáveis marcadas com * apresentam relação negativa com as cargas fatoriais



4. A análise de *cluster*

Enquanto “a análise fatorial reduz as colunas do banco de dados”, o *cluster* “reduz as linhas de dados”, juntando-as em grupos semelhantes internamente, mas diferentes entre si (ROGERSON, 2012, p.297). O processo de classificação das observações se torna tão mais complexo quanto maior for o número de variáveis utilizadas para a montagem dos grupos. Por essa razão, via de regra, a análise de cluster é precedida de uma análise fatorial, com a finalidade de reduzir o número de atributos originais em constructos fracamente correlacionados, evitando redundâncias (HAIR et al, 2005). Além disso, a existência de não correlação entre os dados é fundamental para que os agrupamentos construídos se configurem “por si mesmos”, sem interferência de “qualquer tipo de dependência entre as variáveis” (CORRAR, PAULO, DIAS FILHO, 2011, p.327).

Everitt e Hothorn (2011) apresentam dois métodos de análise de agrupamentos: técnicas de aglomerados hierárquicos e clusters baseados em modelos. O primeiro produz uma classificação na qual os dados são divididos em um número de agrupamentos que varia de acordo com a escala de análise utilizada. Assim, esse método conta com um primeiro grupo, dentro do qual estão todas as observações do banco. À medida que diminuimos as distâncias, mais agrupamentos vão surgindo até que, no limite, cada observação refere-se a um único agrupamento. A questão central dessa técnica é decidir qual o número ideal de agrupamentos aos dados utilizados. Os autores apresentam alguns algoritmos de partição que ajudam nessa escolha. Ainda assim, o método tem limitações relacionadas ao efeito escala.

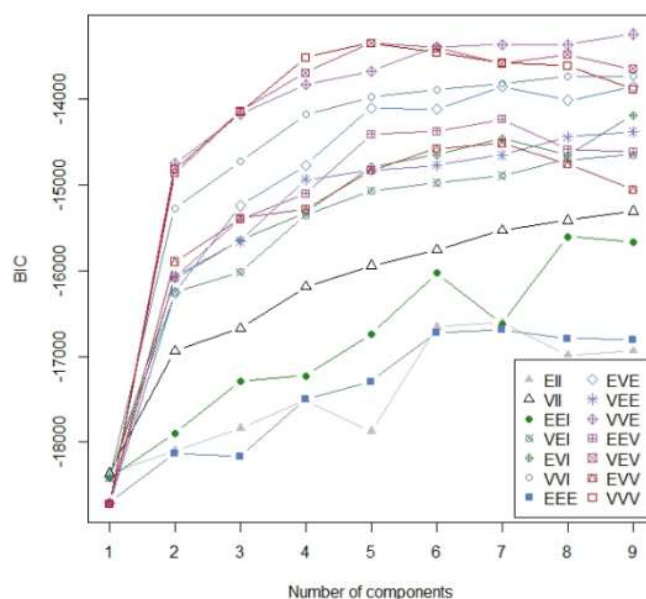
De outro modo, os *clusters* baseados em modelo não são construídos a partir das distâncias, e sim com base na criação de um ponto médio no plano cartesiano (centroide). Essa técnica parte do pressuposto que a população de interesse é composta por subpopulações latentes (que são os próprios *clusters*), cada uma das quais descritas por variáveis com diferentes curvas de probabilidade. No conjunto, essas probabilidades resultam na chamada curva de probabilidades mistura finita (referente à população como um todo) que provê o modelo a partir do qual os *clusters* são estimados (Everitt e Hothorn (2011).

Na aplicação dessa técnica, a principal questão é escolher os parâmetros utilizados para calcular a probabilidade de cada observação compor um determinado agrupamento. Se comparados aos *clusters* hierárquicos, essa técnica tem a vantagem de produzir grupos mais consistentes do ponto de vista estatístico. Além disso, o modelo mostra também a

melhor forma de os analisar. No entanto, a presença de *outliers* dificulta o processo de agrupamento (MINGOTI, 2005). Ainda assim, o trabalho optou por essa técnica.

À luz dessa limitação, os escores resultantes da análise fatorial foram padronizados. A aplicação da técnica foi feita a partir do algoritmo ‘mclust’ no R. De acordo com o critério bayesiano (BIC), o parâmetro mais adequado à análise dos agrupamentos trata-se do modelo elipsoidal, igual orientação (VVE), com 9 componentes (Gráfico 2).

Gráfico 2
Delimitação do número de *cluster* de acordo com o BIC



A Tabela 3 apresenta os valores médios de cada fator padronizado na composição dos *clusters*. À luz das médias mais altas e mais baixas, os agrupamentos criados foram interpretados, nomeados e representados no Quadro 6. É importante ressaltar que alguns dos *clusters* (como o 6 e 9) foram interpretados a partir das dimensões que não os compõem. Essa situação dificulta a análise dos resultados e explicita a não adequação das variáveis selecionadas na compreensão da dinâmica desse rural. No caso do *cluster* 6, a interpretação foi facilitada pelo fato dos municípios agrupados não apresentarem população classificada como rural⁹. Por outro lado, a compreensão do agrupamento 9 explicita a importância de outras dimensões, não captadas pelos indicadores originais.

⁹Dos 30 municípios paulistas sem população rural, segundo o Censo Demográfico 2010, 29 foram agrupados nesse cluster.

Tabela 3
Pontos médios de cada fator (padronizado) por cluster

	Fator 1 Rural agrícola: média propriedade	Fator 2 Rural da pecuária	Fator 3 Rural das pastagens	Fator 4 Rural dos presídios	Fator 5 Rural envelhecido	Fator 6 Rural da cana-de- açúcar	Fator 7 Rural altamente capitalizado	Fator 8 Rural da laranja	Fator 9 Rural agrícola: pequena propriedade	Fator 10 Rural expansão urbana
Cluster 1	-0,43	0,12	0,00	1,42	-0,47	0,30	0,80	-0,10	-0,26	1,49
Cluster 2	0,33	-0,32	0,16	-0,15	0,03	0,32	0,08	2,53	0,20	0,03
Cluster 3	0,00	-0,25	-0,10	0,15	0,47	-0,65	-0,34	-0,59	0,70	0,26
Cluster 4	0,07	-0,39	0,03	-0,25	0,19	1,11	0,09	-0,31	-0,11	-0,24
Cluster 5	0,32	0,40	-0,44	-0,10	0,65	-0,14	-0,20	-0,04	-0,39	-0,28
Cluster 6	-0,47	-0,46	-1,00	-0,52	-3,73	-0,27	-0,37	-0,24	-1,12	-0,78
Cluster 7	0,08	0,39	2,16	-0,19	-0,15	-0,24	-0,28	-0,37	-0,07	-0,31
Cluster 8	0,30	0,62	-0,01	-0,30	-0,03	-0,49	0,05	-0,32	0,38	-0,19
Cluster 9	-1,41	-1,05	0,02	-0,12	0,31	-0,71	-0,09	-0,11	0,27	-0,14

Nota: As células destacadas em rosa representam as médias mais altas, enquanto as destacadas em azul apresentam as médias mais baixas

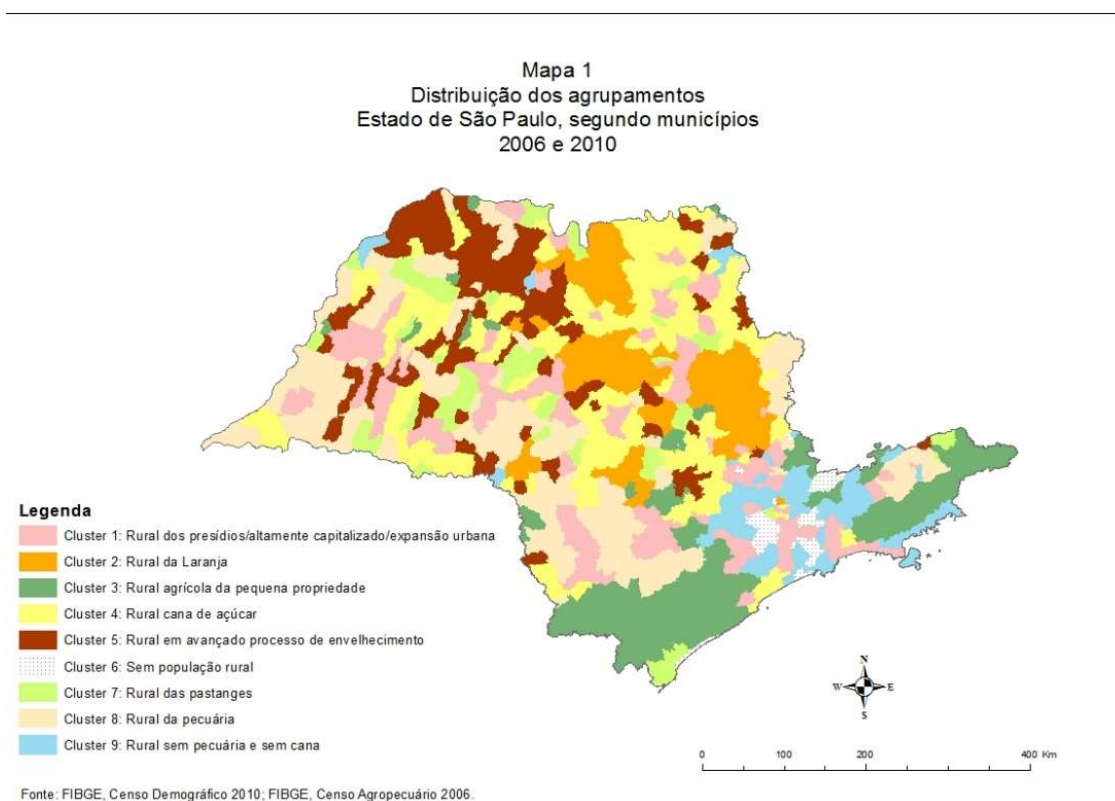
Quadro 6
Fatores que compõe cada cluster e total de municípios nos agrupamentos criados

Cluster	Total de municípios	Fatores que compõe cada cluster
Cluster 1		
Rural dos presídios / altamente capitalizado / expansão urbana	68	<ul style="list-style-type: none"> Fator 4: Rural dos presídios Fator 7: Rural altamente capitalizado (estabelecimentos com alta média de tratores e pessoal empregado) Fator 10: Rural expansão urbana
Cluster 2		
Rural da laranja	56	<ul style="list-style-type: none"> Fator 1: Rural agrícola média propriedade Fator 8: Rural da laranja
Cluster 3		
Rural agrícola da pequena propriedade	72	<ul style="list-style-type: none"> Fator 9: Rural agrícola pequena propriedade *Fator 8: Rural da laranja (ou seja, esse cluster tende a não ter laranja)
Cluster 4		
Rural da cana-de-açúcar	110	<ul style="list-style-type: none"> Fator 6: Rural da de cana-de-açúcar
Cluster 5		
Rural em avançado processo de envelhecimento	127	<ul style="list-style-type: none"> Fator 5: Rural em avançado processo de envelhecimento
Cluster 6		
Municípios sem população rural	29	<ul style="list-style-type: none"> *Fator 4: Rural dos presídios (ou seja, esse cluster tende a não ter população carcerária) *Fator 5: Rural em avançado processo de envelhecimento (ou seja, apresenta uma estrutura etária mais jovem) *Fator 7: Rural altamente capitalizado (ou seja, apresenta uma baixa média de tratores e pessoal ocupado por estabelecimento) *Fator 9: Rural agrícola da pequena propriedade (ou seja, tendência a não ter pequenas propriedades) *Fator 10: Rural expansão urbana (ou seja, tendência a não ter população classificada em setores rural-expansão urbana)
Cluster 7		
Rural das pastagens	36	<ul style="list-style-type: none"> Fator 3: Rural das pastagens
Cluster 8		
Rural da pecuária	103	<ul style="list-style-type: none"> Fator 2: Rural da pecuária
Cluster 9		
Rural com pouca importância das médias propriedades, da pecuária e da cana-de-açúcar	44	<ul style="list-style-type: none"> *Fator 1: Rural agrícola da média propriedade (ou seja, pouca importância do rural formado por estabelecimentos de 10 a 100 ha) *Fator 2: Rural da pecuária (ou seja, tendência a não ter rebanho de bovinos) *Fator 6: Rural da cana-de-açúcar (ou seja, pouca importância da cana)

Nota: Os fatores marcados com (*) apresentam valores médios negativos na composição do cluster

5. A análise espacial

O exame da localização permite refinar o entendimento dos *clusters*, colaborando na compreensão de casos como o agrupamento 9. Esse recurso também ajuda no entendimento do *cluster* 1 que, como consta no Quadro 6, foi formado a partir da junção de fatores que expressam dinâmicas bastante diferentes. O Mapa 1 representa, portanto, a classificação dos 645 municípios paulistas segundo os agrupamentos aos quais pertencem. Como é possível observar, o grupo 9 concentra-se no litoral norte e no entorno da Região Metropolitana de São Paulo. Não captada pelas variáveis selecionadas nesse trabalho, a dinâmica desse rural remete mais aos arranjos urbanos-regionais de Moura (2009).



A espacialização do cluster 1 também facilita a decomposição desse agrupamento em três sub-grupos: a oeste, em população carcerária, conforme contido em Cescon (2012); a nordeste, em rural altamente capitalizado, onde se concentra a produção de laranja e cana-de-açúcar; e mais a sudeste, próximo às regiões metropolitanas, em rural expansão-urbana.

É evidente, ainda, a concentração do *cluster* “rural em avançado processo de envelhecimento demográfico” a oeste do Estado (em especial, no Noroeste paulista). Essa

característica é resultado tanto da baixa taxa de fecundidade da região¹⁰, como do caráter tardio de sua ocupação capitalista. Devido ao fato desse processo se concentrar entre os anos 1940/60, muitas famílias trazidas pela fronteira agrícola ainda sobrevivem nesse rural (DEMÉTRIO, 2013). A persistência dessas famílias, resistentes à agricultura de commodities (MELLO, 2013), cria obstáculos à expansão dessa atividade, representada principalmente pela cana-de-açúcar que, nessa região, só ocorre em convívio com a produção de alimentos para abastecimento local e regional (DEMÉTRIO, 2013).

Também as pastagens aparecem como fenômeno típico do oeste e centro-oeste paulista. Ainda que a pecuária tenha passado por investimentos que a alçou em outro patamar, principalmente com a expansão dos frigoríficos e das exportações de carne brasileira (TARTAGLIA; OLIVEIRA, 1988), a importância dessa atividade, nessa porção do Estado, também remete à dinâmica de sua fronteira agrícola, processo que se fez à custa de baixíssimos coeficientes de capitalização, ou mesmo sem nenhuma capitalização prévia. Juntamente com a produção familiar em pequenas propriedades, a ocupação capitalista do Oeste Paulista resultou na formação de grandes pastagens guardadas à especulação (CHAIA, 1980). A constituição do fator 6, que deu origem ao cluster 7 (rural das pastagens), marcado pela presença de propriedades de mais de 1000 ha, com 100 a 200 ha de pastagens e sem rebanho de bovinos, indica a atualidade dessa dinâmica de valorização de terras no rural paulista.

Em contraste, o rural da cana e da laranja concentra-se na porção central e nordeste, regiões integradas pelos históricos eixos de desenvolvimento do Estado, onde o cultivo de commodities fez-se presente desde o momento histórico de sua ocupação capitalista (TARTAGLIA; OLIVEIRA, 1988). Esse é o rural mais capitalizado do estado, dinamizado pelos complexos agroindustriais (DELGADO, 1985), com a formação de importantes cidades do agronegócio (ELIAS, 2003).

A análise da distribuição espacial dos *clusters* aponta, ainda, um grupo formado de pequenas propriedades cuja produção não passa pelo cultivo da cana-de-açúcar, laranja ou pecuária. Esse agrupamento (*cluster* 4), composto pelo fator 9, concentra-se no extremo sudoeste (Vale do Ribeira) e sudeste (Vale do Paraíba) de São Paulo. Para o seu entendimento, as variáveis selecionadas originalmente também não são suficientes.

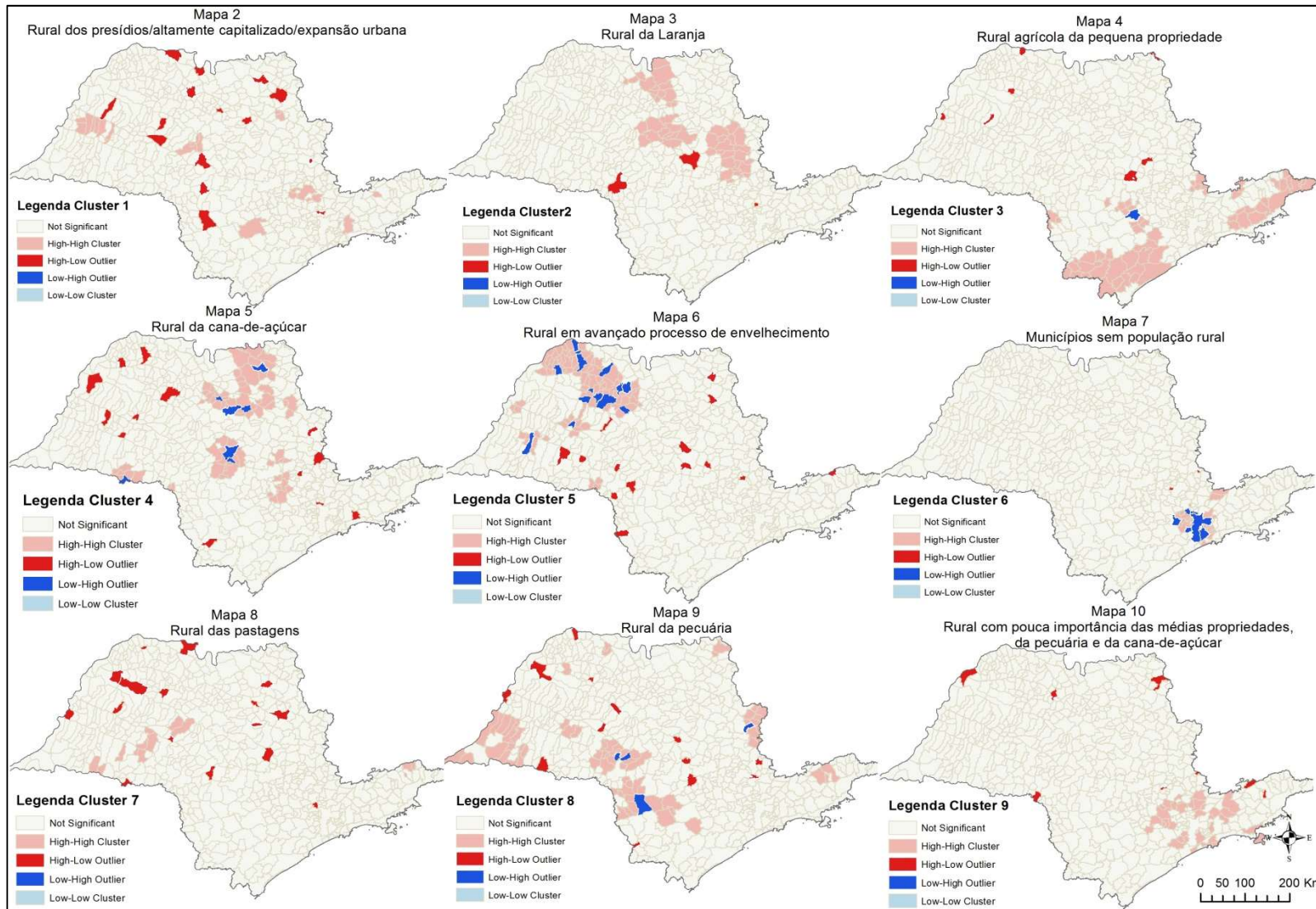
¹⁰ Para o ano de 2009, a Fundação Seade (2011) estimou a TFT da Região Administrativa de São José do Rio Preto (destacada em marrom escuro, no Mapa 1, a noroeste de São Paulo) em 1,5. Essa foi a média mais baixa do Estado.

Para facilitar a identificação de uma correlação espacial na distribuição dos *clusters*, os Mapas 2 a 10 apresentam as aglomerações identificadas a partir da aplicação do Índice Local de Moran para cada agrupamento. Nesse procedimento, não se considerou nenhum raio de distância, bem como foi seguido o padrão do ArcGis (versão 10). Essa técnica mede vizinhança, ou seja, avalia a distribuição de eventos de modo a identificar os locais onde existem aglomerados de áreas com elevados valores cujos polígonos vizinhos apresentam também valores significativamente altos (e ainda áreas com baixos valores circundados por polígonos com valores igualmente baixos).

Segundo Bueno e D'Antona (2012), a vantagem dessa ferramenta é construir mapas que não são apenas descritivos, mas que apresentam uma base estatística por meio do qual se pode assegurar a existência de um cluster espacial. É interessante destacar cana-de-açúcar, laranja e pecuária compõem arranjos urbanos-rurais regionais claramente delimitados. Esse é também o caso do rural em avançado processo de envelhecimento e do rural estruturado por pequenos estabelecimentos da porção sul do Estado (Mapas 2 a 10).

6. Considerações finais

A discussão levantada nesse trabalho permite avançar na análise da relação entre as características de cada região do Estado de São Paulo e a sua dinâmica rural. Nesse debate, é destacado a formação de diferentes arranjos urbanos-rurais regionais, estruturados na interface entre redistribuição de população e produção de commodities (laranja, cana e carne), construções socioespaciais típicas de áreas onde os processos históricos deixaram heranças que possibilitaram a expansão dessas atividades. É reforçado, portanto, as rugosidades do espaço (SANTOS, 1993) e os diferentes vínculos transnacionais que as distintas regiões podem ter (SASSEN, 1998): se nas regiões metropolitanas essas relações passam principalmente pelos serviços (MOURA, 2009), nas regiões não-metropolitanas esses vínculos estão associados principalmente às commodities. A rede urbana paulista redefine-se, assim, em distintos espaços transnacionais.



Fonte: FIBGE, Censo Demográfico 2010; FIBGE, Censo Agropecuário 2006

7. Referências Bibliográficas

- BAENINGER, R. **Região, Metrópole e Interior: Espaços Ganhadores e Espaços Perdedores nas Migrações Recentes – Brasil, 1980-1996**. Tese de doutorado. Campinas: Instituto de Filosofia e Ciências Humanas – IFCH/UNICAMP, 1999
- BUENO, M. C. D.; D'ANTONA, A. O. Utilização de Grades Regulares para Análises Espaciais Intramunicipais de Variáveis Demográficas – Teste para Limeira-SP, 2010. In: **Anais do XVIII Encontro Nacional de Estudos Populacionais**. Águas de Lindóia, 19 a 23 de novembro de 2012.
- CAMARANO, A. A; ABRAMOVAY, R. Êxodo rural, envelhecimento e masculinização no Brasil: panorama dos últimos cinquenta anos. In: **Encontro Anual da Associação Nacional de Pós-Graduação em Ciências Sociais**. Caxambu: ANPOCS, 1997.
- CASTILLO, R. Dinâmicas recentes do setor sucroenergético no Brasil: competitividade regional e expansão para o Bioma Cerrado. *GEOgraphia* – Ano 17 – Nº 35 – Dossiê, 2015.
- CHAIA, V. L. M. **Os conflitos de arrendatários em Santa Fé do Sul – SP (1959-1969)**. São Paulo. Dissertação de Mestrado. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas – FFLCH/USP, 1980.
- CESCON, F. R. P. **Migração e unidades prisionais: o cenário dos pequenos municípios do Oeste paulista**. Dissertação de Mestrado. Campinas: Instituto de Filosofia e Ciências Humanas – IFCH/UNICAMP, 2012.
- CORRAR, L. J.; PAULO, E.; DIAS FILHO, J.M. **Análise multivariada para cursos de administração, ciências contábeis e economia**. São Paulo: Atlas, 2011.
- CORRÊA, R. L. **A rede urbana**. São Paulo: Editora Ática, 1994.
- CUNHA, J. M. P. Migração e urbanização no Brasil: alguns desafios metodológicos para análise. In: **São Paulo em Perspectiva**, v.19, n.4, p.3-20. out/dez, 2005.
- DELGADO, Guilherme. Mudança técnica na agricultura, constituição do complexo agroindustrial e política tecnológica recente. **Cadernos de Ciência & Tecnologia**, Brasília, 2(1): 79-97, jan. /abr., 1985.
- DEMÉTRIO, N. B. **População e dinâmica econômica na Região de Governo de Jales: o outro rural do Oeste Paulista**. Dissertação de Mestrado. Campinas: Instituto de Filosofia e Ciências Humanas – IFCH/UNICAMP, 2013.
- ELIAS, D. Agricultura científica no Brasil: impactos territoriais e sociais. In: SOUZA, M. A. A. de (org.) **Território brasileiro: usos e abusos**. Campinas: Territorial, 2003.
- FUNDAÇÃO IBGE. Censo Agropecuário 2006.
- FUNDAÇÃO IBGE. Censo Demográfico 2010.
- FUNDAÇÃO SEADE. Fecundidade continua em queda em São Paulo. In: **SP Demográfico: resenha de estatísticas vitais do Estado de São Paulo**, ano 11, nº 5, julho de 2011.
- GONÇALVES, M. F. **As engrenagens da locomotiva: ensaio sobre a formação urbana paulista**. Tese de doutorado. Campinas: Instituto de Filosofia e Ciências Humanas – IFCH/UNICAMP, 1998.
- HAIR, J. F. Jr.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C. **Análise multivariada de dados**. Porto Alegre: Bookman, 2005.
- MELO, B. M. de. **História e memória na contramão da expansão canavieira: um estudo das formas de resistência dos sitiantes do extremo noroeste paulista**. Tese de Doutorado. São Carlos: UFSCAR, 2013.

MINGOTI, Sueli Aparecida. **Análise de dados através de métodos de estatística multivariada - uma abordagem aplicada**. UFMG: Belo Horizonte. 2007, 295p.

MOURA, R. **Arranjos urbano-regionais no Brasil: uma análise com foco em Curitiba**. Tese de doutorado. Curitiba: Programa de Pós-Graduação em Geografia. Universidade Federal do Paraná, 2009.

NEGRI, B., GONÇALVES, M. F. e CANO, W. O processo de interiorização do desenvolvimento e da urbanização no Estado de São Paulo (1920-1980). In: CANO, W. (Org.). **A interiorização do desenvolvimento econômico no Estado de São Paulo (1920-1980)**. São Paulo: SEADE, 1988, p.5-93.

PERES, R. G.; BAENINGER, R. Dinâmica populacional nas regiões canavieiras. In: BAENINGER, R. et al (Org.). **Regiões Canavieiras**. Campinas: Núcleo de Estudos de População – Nepo/Centro de Pesquisas Meteorológicas e Climáticas Aplicadas à Agricultura – CEPAGRI/Núcleo de Estudos e Pesquisas em Alimentação – NEPA/Unicamp, 2013.

RODRIGUES, I. A. **A Demografia da Vida Rural Paulista**. Tese de doutorado. Campinas: Instituto de Filosofia e Ciências Humanas – IFCH/UNICAMP, 2001.

ROGERSON, P. A. **Métodos estatísticos para a geografia: um guia para o estudante**. Porto Alegre: Bookman, 2012.

SANTOS, M. **A urbanização brasileira**. São Paulo: HUCITEC, 1993.

SILVA, J. G. da. **O novo rural brasileiro**. Campinas: UNICAMP. IE, 1999.

SINGER, P. Migrações internas: considerações teóricas sobre o seu estudo. In: MOURA, H. (Org.). **Migração interna: textos selecionados**. Fortaleza: Banco do Nordeste do Brasil S.A., 1980.

TARTAGLIA, J.C; OLIVEIRA, O. L. Agricultura Paulista e sua Dinâmica Regional (1920-1980). In: CANO, W. (Org.). **A interiorização do desenvolvimento econômico no Estado de São Paulo (1920-1980)**. São Paulo: SEADE, 1988, p.5-93.

VEIGA, J. E. **Cidades imaginárias: o Brasil é menos urbano do que se calcula**. Campinas: Editora Autores Associados, 2003.

WANDERLEY, M. N. B. **Um saber necessário: os estudos rurais no Brasil**. Campinas: Editora da Unicamp, 2011.