



CAIO PEIXOTO CHAIN

**ABORDAGEM GEOESTATÍSTICA PARA IDENTIFICAÇÃO
DE POTENCIAIS *CLUSTERS* INDUSTRIAIS**

**LAVRAS – MG
2018**

CAIO PEIXOTO CHAIN

**ABORDAGEM GEOESTATÍSTICA PARA IDENTIFICAÇÃO DE POTENCIAIS
CLUSTERS INDUSTRIAIS**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Administração, área de concentração em Gestão de Negócios, Economia e Mercados, para a obtenção do título de Doutor.

Prof. Dr. Antônio Carlos dos Santos
Orientador

Prof. Dr. Luiz Gonzaga de Castro Junior
Coorientador

**LAVRAS – MG
2018**

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Chain, Caio Peixoto.

Abordagem geoestatística para identificação de potenciais
clusters industriais / Caio Peixoto Chain. - 2018.

98 p. : il.

Orientador(a): Antônio Carlos dos Santos.

Coorientador(a): Luiz Gonzaga de Castro Junior.

Tese (doutorado) - Universidade Federal de Lavras, 2018.

Bibliografia.

1. Concentração de firmas. 2. Métodos de análise regional. 3.
Viés direcional. I. Santos, Antônio Carlos dos. II. Castro Junior,
Luiz Gonzaga de. III. Título.

CAIO PEIXOTO CHAIN

**ABORDAGEM GEOESTATÍSTICA PARA IDENTIFICAÇÃO DE POTENCIAIS
CLUSTERS INDUSTRIAIS**

**GEOSTATISTICAL APPROACH FOR IDENTIFICATION OF POTENTIAL
INDUSTRY CLUSTERS**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Administração, área de concentração em Gestão de Negócios, Economia e Mercados, para a obtenção do título de Doutor.

APROVADA em 02 de Fevereiro de 2018.

Prof. Dr. Cléber Carvalho de Castro	UFLA
Prof. Dr. Francisval de Melo Carvalho	UFLA
Prof. Dr. Marcelo Silva de Oliveira	UFLA
Prof. Dr. Cezar Augusto Miranda Guedes	UFRRJ

Prof. Dr. Antônio Carlos dos Santos
Orientador

Prof. Dr. Luiz Gonzaga de Castro Junior
Coorientador

**LAVRAS – MG
2018**

AGRADECIMENTOS

Aos orientadores Antônio Carlos e Gonzaga que despertaram meu interesse por estudos voltados para a economia e gestão agroindustrial. A convivência e os conhecimentos transmitidos foram fundamentais para minha formação.

Aos professores membros da banca, Cleber, Cezar, Marcelo e Francisval, pelas relevantes revisões e sugestões, tanto na etapa de qualificação, quanto na defesa da tese. Certamente contribuirão para a publicação dos artigos em bons periódicos.

Ao professor Scalon e, novamente, ao professor Marcelo do Departamento de Estatística (DES) da UFLA, pelos ensinamentos sobre estatística espacial, especialmente em Geoestatística e Processos Pontuais. Igualmente agradeço ao professor Pedro, da UFMG, que possibilitou complementar meus estudos sobre o assunto por meio da disciplina de Econometria Espacial cursada no Centro de Desenvolvimento e Planejamento Regional (CEDEPLAR). Esse conjunto de disciplinas foi essencial para a elaboração da tese.

Aos professores e funcionários do Departamento de Administração e Economia (DAE), pelos ensinamentos transmitidos e pela assistência prestada nos últimos anos. Agradeço também pela oportunidade de atuação simultânea como doutorando e professor substituto de Teoria Econômica e Economia Industrial que confirmou o meu interesse em trabalhar com ensino, pesquisa e extensão.

Aos colegas do Centro de Inteligência em Mercados (CIM) e, novamente, ao prof. Gonzaga, pela oportunidade de desenvolver estudos e assessorias que materializaram o conhecimento acadêmico na resolução de problemas práticos.

A CAPES pelo apoio financeiro disponibilizado para cursar o doutorado.

Por último, e mais importante, aos meus familiares e amigos.

RESUMO

O objetivo do presente estudo foi propor uma abordagem geoestatística para a identificação de potenciais *clusters* industriais. Parte-se da premissa que novos índices de concentração espacial são fundamentais para integrar a economia e a geografia na elaboração de políticas de desenvolvimento regional e aumento da competitividade empresarial. A teoria dos *clusters* foi dividida em três vertentes em que cada uma possui sua forma de estimação empírica. No primeiro caso as Aglomerações Puras são identificadas por índices locacionais, na visão seguinte, dos Complexos Industriais, o agrupamento setorial é definido por meio de relações insumo-produto e, por fim, os *Clusters* de Porter integram as duas abordagens anteriores. Por meio de uma análise bibliométrica, foi identificado que os métodos baseados em índices locacionais e estatística espacial, principalmente em processos pontuais, representam o *mainstream* no campo de conhecimento sobre a mensuração de *clusters* de firmas. Essa geração de estudos se tornou proeminente por contornar o viés de agregação, porém foi confirmado que a problematização do viés direcional (anisotropia) continua negligenciada, visto que as pesquisas assumem isotropia. A proximidade e a concentração de firmas foram examinadas por meio da geoestatística. Essa abordagem foi capaz de atender aos princípios já consolidados pela literatura *mainstream*, além de agregar a análise do viés direcional, o zoneamento dos potenciais *clusters* industriais em mapa e a estimativa da concentração industrial no nível da firma. A análise direcional representou melhor a aglomeração de firmas do ponto de vista estatístico, com menor nível de erro, e econômico, com o agrupamento de firmas em regiões com perfil homogêneo que tende a facilitar a coordenação estratégica do *cluster*. A abordagem geoestatística foi aplicada na indústria de torrefação do café em Minas Gerais e os potenciais *clusters* identificados foram nas regiões conhecidas como Matas de Minas, Capelinha e Sul de Minas.

Palavras-chave: Concentração de firmas. Métodos de análise regional. Viés direcional.

ABSTRACT

The aim of the present study was to develop a geostatistical approach for the identification of potential industry clusters. It is based on the premise that new spatial concentration indices are fundamental to integrate economics and geography in the formulation of regional development and business competitiveness policies. In the first part of the paper, cluster theory was divided into three strands in which each one has its own form of estimation. In the first case, the Pure Agglomerations are identified by locational indices, in the following view, the Industrial Complexes, the sectorial grouping is defined by means of input-output relations and, finally, clusters of Porter integrate the two previous approaches. Through a bibliometric analysis, it was identified that methods based on indices and spatial statistics, especially in point processes, represent the mainstream in the field of knowledge on the measurement of firm clusters. This generation of studies became prominent because it circumvented the aggregation bias, but it was confirmed that the problem of directional bias (anisotropy) remains neglected, since the researches assume isotropy. Questions of proximity and concentration of firms were examined through geostatistics. This approach was able to meet the principles already consolidated by the mainstream literature, as well as aggregated the directional bias analysis, the zoning of potential industry clusters on maps, and the estimation of firm-level industry concentration. Directional analysis represented better the clustering of firms from a statistical point of view, with a lower level of error, and economic, grouping the firms in regions with a homogeneous profile that tends to facilitate cluster strategic coordination. The geostatistical approach was applied in the roasted coffee industry in Minas Gerais and the potential clusters identified were in the regions known as Matas de Minas, Capelinha and Sul de Minas.

Keywords: Concentration of firms. Methods of regional analysis. Directional bias.

SUMÁRIO

1	AGLOMERAÇÕES GEOGRÁFICAS DE FIRMAS.....	8
1.1	INTRODUÇÃO.....	8
1.2	OBJETIVOS.....	14
1.3	REFERENCIAL TEÓRICO SOBRE AGLOMERAÇÕES DE FIRMAS.....	15
1.3.1	Distritos, <i>Milieux</i> e <i>Clusters</i>	15
1.3.2	Aglomerações, Complexos e Redes Sociais.....	24
1.4	EXPERIÊNCIAS DE POLÍTICAS DE <i>CLUSTERS</i> NO BRASIL.....	27
1.5	CONSIDERAÇÕES FINAIS.....	31
	REFERÊNCIAS.....	32
2	MÉTODOS QUANTITATIVOS APLICADOS À MENSURAÇÃO DE <i>CLUSTERS</i> INDUSTRIAIS.....	36
2.1	INTRODUÇÃO.....	38
2.2	REVISÕES DE LITERATURA.....	39
2.3	PROCEDIMENTOS METODOLÓGICOS.....	42
2.4	RESULTADOS E DISCUSSÃO.....	45
2.4.1	Resultados descritivos.....	46
2.4.2	Resultados das redes.....	52
2.5	CONCLUSÕES.....	59
	REFERÊNCIAS.....	60
3	MAPEAMENTO DE POTENCIAIS <i>CLUSTERS</i> INDUSTRIAIS PELA ABORDAGEM GEOESTATÍSTICA.....	67
3.1	INTRODUÇÃO.....	69
3.2	REFERENCIAL TEÓRICO.....	71
3.3	METODOLOGIA.....	74
3.3.1	Objeto de estudo - a indústria do café.....	74
3.3.2	Banco de dados.....	75
3.3.3	Abordagem geoestatística.....	76
3.3.3.1	(Semi)variograma.....	77
3.3.3.2	Anisotropia.....	78
3.3.3.3	Krigagem de variáveis categóricas.....	79
3.4	RESULTADOS E DISCUSSÃO.....	81
3.4.1	Análise simulada.....	81
3.4.1.1	Cenário 1.....	81
3.4.1.2	Cenário 2.....	82
3.4.2	Análise empírica da indústria do café em Minas Gerais.....	87
3.5	CONCLUSÕES.....	93
	REFERÊNCIAS.....	95

1 AGLOMERAÇÕES GEOGRÁFICAS DE FIRMAS

1.1 INTRODUÇÃO

A questão da proximidade entre firmas ocupa papel de destaque nas investigações sobre desenvolvimento econômico e competitividade empresarial desde que Marshall (1890) formalizou suas observações sobre esse fenômeno na Inglaterra. A localização geográfica das firmas de um mesmo setor e dos demais agentes econômicos correlatos foi considerada o determinante dos ganhos de produtividade das firmas e do aumento da capacidade de inovação setorial em uma região. Essa percepção formou a base da Teoria das Aglomerações.

Os desdobramentos da globalização e das Tecnologias de Informação e Comunicação (TICs) nas últimas décadas alteraram o papel da proximidade geográfica na economia e nos negócios, sem torná-la irrelevante. Alguns autores advogam que a maior integração entre economias internacionais e o avanço tecnológico reduziu significativamente a importância da localização de atividades econômicas (AHN et al., 2009; CAIRNCROSS, 1997). Alguns autores preferem relativizar esta relação de acordo com a indústria em que as firmas estão inseridas (BUCIUNI; PISANO, 2015). Existem ainda pesquisas que partem do pressuposto que a proximidade entre firmas em uma mesma indústria e em setores correlatos é um dos determinantes do desenvolvimento regional e da competitividade em nível regional e internacional (DELGADO; PORTER; STERN, 2014; PORTER, 1998).

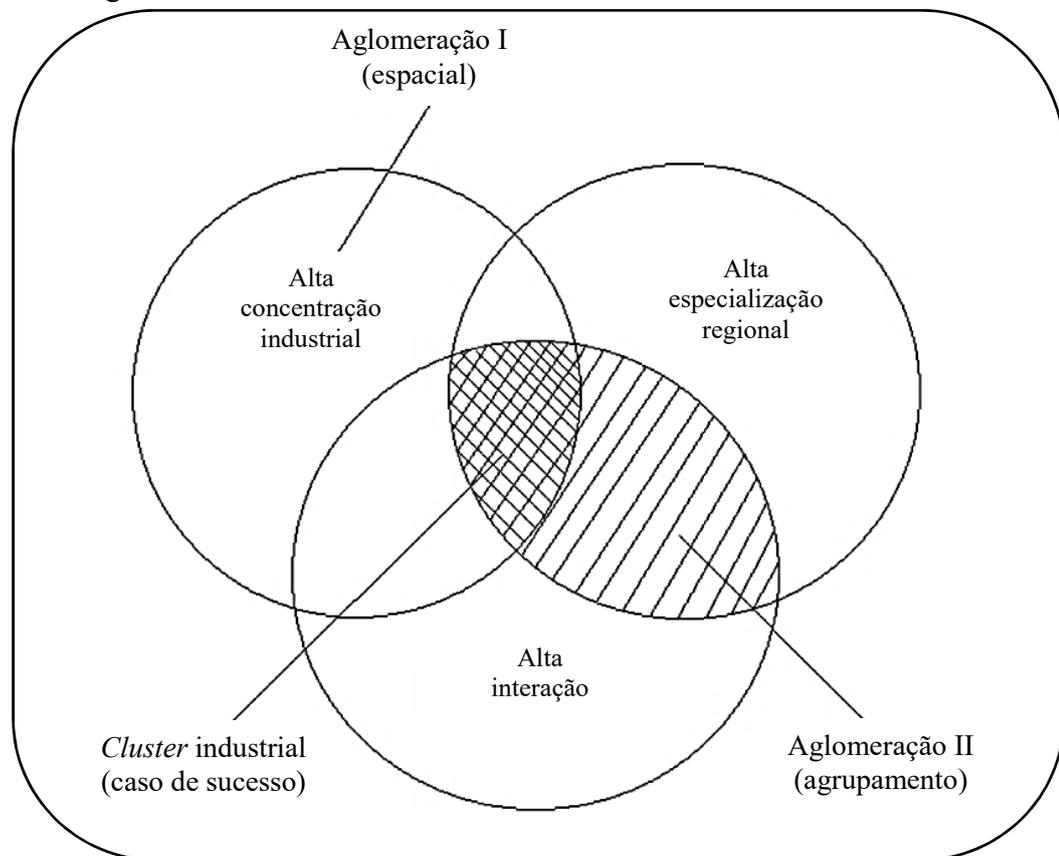
O conceito de *clusters* industriais de Porter (1998) é considerado o mais difundido entre as visões que consideram relevante a localização geográfica de empresas e agentes econômicos. Para esse autor, um *cluster* pode ser definido como uma concentração geográfica onde ocorrem interligações entre as empresas, fornecedores de bens e serviços especializados, empresas em setores correlatos e organizações associadas (universidades, associações de classe, agências de suporte, entre outras) em campos particulares que competem, mas também cooperam (PORTER, 1998). Essa definição teve como base uma versão anterior que considerava que a concentração espacial era apenas uma possibilidade de intensificação dos *clusters* (PORTER, 1990).

Com a visibilidade alcançada pelas pesquisas baseadas na teoria das aglomerações e dos *clusters* entre os anos 80 e 90, surgiram várias vertentes para definir e analisar esse fenômeno, ainda que as perspectivas tendam a convergir para a relação entre a localização, ganhos de produtividade e desempenho das firmas (CRUZ; TEIXEIRA, 2010; HERVAS-

OLIVER et al., 2015; MARTÍNEZ-FERNÁNDEZ; CAPO-VICEDO; VALLET-BELLMUNT, 2012). Dentre essas vertentes, podem ser citadas: “Milieux Innovateurs” (AYDALOT, 1986) da escola de pensamento francesa; “Sistemas Locais de Produção” (STORPER; HARRISON, 1991) e “Novos Espaços Industriais” (STORPER; SCOTT, 1989) da escola californiana; “Sistemas Nacionais de Inovação” (LUNDVALL, 1992) dos estudos em inovação; da escola italiana vieram a “Especialização flexível” (PIORE; SABEL, 1984) e o “Distrito Industrial Neo-marshalliano” (BECATTINI; 1979); por fim, os “Arranjos Produtivos Locais (APLs) no Brasil” (CASSIOLATO; LASTRES, 2003).

O conceito de *clusters* foi originado da teoria das aglomerações (BROSNAN; DOYLE; O’CONNOR, 2016; VOM HOFE; CHEN, 2006), assim *clusters* e aglomerações geralmente são considerados sinônimos (CRUZ; TEIXEIRA, 2010; HERVAS-OLIVER et al., 2015; MALMBERG; MASKELL, 2002). Outros termos referentes ao fenômeno da proximidade espacial entre firmas também podem ser encontrados sem diferenciação na literatura, porém Long e Zhang (2012) propuseram as categorias apresentadas na Figura 1.1.

Figura 1.1 – Relacionamento entre os termos sobre *clusters* industriais.



Fonte: Adaptado de Long e Zhang (2012)

A concentração geográfica de firmas (concentração industrial) é frequentemente tratada como sinônimo de especialização regional. No entanto, o primeiro termo ocorre quando a produção total de uma indústria está concentrada em um número pequeno de regiões, enquanto o segundo descreve o quanto uma região decide focar sua estrutura produtiva em um número limitado de indústrias. Quando a concentração industrial local de uma indústria em particular está acima do nível global desta atividade econômica ela pode ser denominada como indústria localizada, ou seja, após os efeitos da concentração setorial e da concentração total serem controlados (DURANTON; OVERMAN, 2005; FRATESI, 2008).

Por sua vez, o termo “aglomerações” (agrupamentos) enfatiza a especialização conjugada com a proximidade, e conseqüentemente a interação, entre firmas e organizações correlatas em uma região. No entanto, o nível das aglomerações é geralmente mensurado por indicadores de concentração industrial, logo esses termos também costumam ser permutados (aglomerações espaciais). Os ganhos de competitividade e de produtividade das empresas em agrupamentos (Aglomeração II) tende a aumentar a concentração industrial (Aglomeração I) e nesses casos se confirmam a existência de *clusters* industriais de sucesso. Elevados índices de concentração podem então ser entendidos como potenciais *clusters* de sucesso.

Os *clusters* industriais foram utilizados como a modelo organizacional que deu base para o recente desenvolvimento regional alcançado por Estados Unidos, Europa e alguns países emergentes (LAZZERETTI; SEDITA; CALOFFI, 2014). Alguns exemplos de *clusters* de sucesso são os de alta tecnologia no Vale do Silício na Califórnia, aeroespacial em Seattle, automotivo em Detroit-Winsor e móveis na Carolina do Norte (EUA); biotecnologia na França; calçados no norte da Itália; e semicondutores na Malásia (BEHRENS; BOUGNA, 2015; BILLINGS; JOHNSON, 2012; BRAUNERHJLEM; BORGMAN, 2004). Assim, o conceito de *cluster* foi promovido como um quadro analítico e também como uma ferramenta para políticas econômicas.

No entanto, a ausência de uma definição única e a dificuldade de mensuração tornou os conceitos de *clusters* e aglomerações de firmas suscetíveis a críticas (FRATESI, 2008; MALMBERG; MASKELL, 2002; MARTIN; SUNLEY, 2003). Como as metodologias empíricas e estratégias para políticas de mapeamento podem ser variáveis, Martin e Sunley (2003) e Swann (2002) propuseram uma sistematização dos conceitos e das respectivas abordagens (Quadro 1.1).

Quadro 1.1 – Variáveis conceituais dos *clusters* e problemas de mensuração.

Conceito	Metodologia empírica	Profundidade	Tipo de mensuração	Evidência empírica
Co-localização	<i>Top-down</i>	Baixa	Quantitativo	Indireta
Co-localização e proximidade tecnológica	↓	↓	↓	↓
Complementariedade insumo-produto				
Co-localização e alto desempenho				
Externalidades marshallianas				
Redes de empresas				
Colaboração explícita				
Transbordamentos de conhecimento	<i>Bottom-up</i>	Alta	Qualitativo	Direta

Fonte: Adaptado de Martin e Sunley (2003) e Swann (2002)

No extremo superior estão as abordagens em nível nacional (*top-down*) que utilizam dados quantitativos para identificar, indústria por indústria, localizações relevantes de atividades econômicas especializadas ou interligadas. As abordagens *bottom-up* estão preocupadas com identificar *clusters* em uma região em particular, em geral, de forma qualitativa. Mensurações estritamente quantitativas tendem a gerar evidências indiretas sobre a formação dos *clusters* e a profundidade de relacionamentos, enquanto as qualitativas se aproximam das evidências diretas. Dentro do intervalo entre metodologias *top-down* e *bottom-up* existem diversos tipos de combinações possíveis.

Muitos esforços vêm sendo despendidos por pesquisadores em relação aos conceitos iniciais de mapeamento do Quadro 1.1 por representarem a primeira etapa de uma política de *clusters* industriais, uma falha nessa fase tende a fornecer direcionamentos equivocados para todas as demais. A análise quantitativa de aglomerações espaciais de firmas pode fornecer subsídios para que as pesquisas qualitativas e de campo sejam geograficamente menos abrangentes e custosas (AUSTRIAN, 2000; HILL; BRENNAN, 2000; RIVERA; SHEFFI; WELSH, 2014).

Isard, Schooler e Vietorisz (1959) contribuíram de forma pioneira para a análise empírica do tema, por meio de matrizes insumo-produto (I-O) e de métodos locacionais. Esta abordagem fundamentada nas ciências regionais ganhou volume principalmente nos anos 70

com a aplicação complementar de métodos de estatística multivariada proposta por Czamanski (1974). A nomenclatura adotada foi a de *clusters* industriais, definidos como setores interligados por fluxos formais de bens e serviços, independente da proximidade espacial, por entender que as firmas se relacionam também a grandes distâncias. Quando estes *clusters* eram geograficamente concentrados recebiam a denominação de complexos industriais (CZAMANSKI; ABLAS, 1979).

Outra abordagem entre as mais consolidadas é baseada em índices de concentração geográfica e de localização espacial (BICKENBACK; BODE; KRIEGER-BODEN, 2013; COMBES; OVERMAN, 2004). A primeira geração de estudos traçou índices comparáveis entre indústrias e que controlavam a tendência natural que as indústrias possuem para se aglomerarem, como o Índice de Gini Locacional – GL – (KRUGMAN, 1991). A primeira característica permitiu verificar se uma atividade econômica é mais/menos concentrada que suas subatividades, que outro setor, entre regiões/países ou ao longo do tempo. A segunda característica determinou se uma indústria é mais/menos aglomerada do que a tendência média de todas as indústrias.

A geração seguinte de índices obteve sucesso em controlar para a concentração industrial, ou seja, neutralizar o efeito da concentração geográfica induzida pelo tamanho de uma (ou poucas) firma(s) de grande porte. Essa geração ficou conhecida como abordagem *dartboard* e possibilitou comparar a concentração espacial em um setor com uma suposta concentração que existiria se todas as plantas deste setor fossem localizadas aleatoriamente (ELLISON; GLAESER, 1997).

A terceira geração agregou as relações de dependência espacial entre firmas, a significância estatística e contornou o Problema da Unidade de Área Modificável – MAUP, ou seja, o espaço passou a ser tratado sem delimitações geográficas, por meio da função Kd de Duranton e Overman (2005). Esses métodos que consideram o espaço contínuo fornecem as medidas de concentração industrial mais sofisticadas (BICKENBACH; BODE, 2008) e ainda estão em sua infância na ciência econômica, visto que possuem aproximadamente uma década de desenvolvimento e desde então se tornaram os mais explorados pelas pesquisas sobre aglomerações de firmas (MARCON; PUECH, 2017).

A limitação colocada por Duranton e Overman (2005) sobre a função Kd é que ela não é apropriada para identificar padrões que não envolvem aglomeração ou dispersão e pode interpretar como aleatórios alguns padrões não aleatoriamente localizados, como no caso de firmas aglomeradas ao longo de uma ferrovia ou região costeira. Nesse caso, pode-se dizer

que existe um viés direcional na localização das firmas, ou ainda, foi imposto um comportamento anisotrópico pela direção da ferrovia ou da costa. A anisotropia ocorre quando a estrutura de dependência espacial de um fenômeno varia conforme a direção.

Além do problema estatístico, a implicação econômica da anisotropia nas distâncias entre firmas parceiras reside nos possíveis impactos na efetividade de comunicação e de coordenação do *cluster*, assim como no incremento de custos operacionais e redução da eficiência das firmas (AHN et al., 2009). A geoestatística é um método de análise espacial que pode gerar informações adicionais para a teoria das aglomerações e para os agentes que atuam na política e na governança de *clusters* industriais, pois considera o espaço contínuo (baseada em distâncias), possui procedimentos consolidados para modelagem da anisotropia e seus resultados podem ser visualizados em mapas.

1.2 OBJETIVOS

Ao se considerar o problema de pesquisa exposto de que se faz necessário obter evidências empíricas, ainda que indiretas, que confirmem a existência de aglomerações de firmas no espaço contínuo tanto em condições isotrópicas quanto anisotrópicas, o objetivo geral do presente estudo é propor a abordagem geoestatística para o mapeamento de potenciais *clusters* industriais.

Especificamente, buscou-se:

- ✓ Apresentar o referencial teórico que embasa a pesquisa empírica sobre a aglomeração geográfica de firmas;
- ✓ Sistematizar o estado da arte em que se encontra a literatura acerca dos métodos para identificar e mensurar os *clusters* industriais, por meio da bibliometria; e
- ✓ Examinar a proximidade e a concentração de firmas em direções específicas, por meio da abordagem geoestatística aplicada em dados simulados e na indústria do café em Minas Gerais.

As análises e políticas de *clusters* industriais podem ser vistas como a aplicação de um conjunto de teorias consolidadas, em versões renovadas, para explicar como a geografia contribui para direcionar o desenvolvimento econômico (BERGMAN; FESER, 1999). Assim, pode-se considerar como justificativa de pesquisa que o desenvolvimento de índices de concentração espacial é um requisito fundamental para contemplar de forma efetiva o papel do espaço geográfico na economia (MARCON; PUECH, 2003; 2017). Como observado por Alfaro e Chen (2014), Head e Mayer (2004) e Malmberg e Maskell (2002) mensurar aglomerações espaciais de atividades econômicas é o desafio central da geografia econômica.

1.3 REFERENCIAL TEÓRICO SOBRE AGLOMERAÇÕES DE FIRMAS

A ligação entre a teoria e a aplicação é o motivo que direciona e suporta as investigações em estudos regionais. Uma forte base epistemológica é necessária na elaboração de políticas de *clusters* industriais para justificar a condução da pesquisa empírica, porém é comum que interesses políticos tenham papel determinante na natureza dessas ações (BERGMAN; FESER, 1999). Por esse motivo, é relevante abordar o tema de forma científica e teoricamente fundamentada.

Nesta etapa foram revisadas as teorias que fundamentam os diferentes conceitos de *clusters* industriais e duas classificações similares foram similares. A primeira subdivide as visões teóricas em Distritos Marshallianos e Neo-marshallianos, com foco em externalidades locais; *Milieux* Inovadores que enfatizam o desenvolvimento regional; e os *Clusters* Industriais de Porter, interessados na agenda da competitividade empresarial. A segunda proposta sistematizou os tipos ideais de *cluster* em Aglomerações Puras, Complexos Industriais e Redes Sociais.

1.3.1 Distritos, *Milieux* e *Clusters*

Pode-se considerar que uma teoria completa deve responder as questões “o que”, “como” e “o porquê” (WHETTEN, 1989). Inicialmente, “o que” identifica os fatores relevantes para a explicação, como variáveis, conceitos ou construtos. Por sua vez, “como” provê as relações de causalidades entre esses fatores de forma ordenada e explícita. Em conjunto, a resposta para essas duas perguntas constituem o campo/assunto da teoria. Por fim, “o porquê” indica o centro da teoria, ou seja, fornece uma justificativa lógica e convincente para os fatores e relacionamentos (“o que” e “como”). Esta terceira questão também é relevante por gerar proposições que demandam novas possibilidades, desafiam ideias consolidadas e aprofundam a compreensão do fenômeno em estudo.

Em relação especificamente à teoria dos *clusters*, Maskell e Kebir (2006) aprofundaram a questão “o porquê” para captar o ciclo de vida dos *clusters* e a subdividiram em três argumentos: “existência”, “extensão” e “exaustão”. O primeiro argumento busca contemplar os benefícios econômicos e sociais que levam as firmas à localização conjunta. A “extensão” dos *clusters* justifica as deseconomias decorrentes de um excessivo nível de extensão geográfica e/ou abrangência setorial. Por último, a “exaustão” explica as condições

internas e externas levaram os benefícios da co-localização a se tornarem negativos durante o ciclo de vida do *cluster*. Quando o argumento “existência” é prejudicado, a vitalidade do *cluster* pode se tornar comprometida.

Maskell e Kebir (2006) buscaram responder o que se qualifica como uma teoria no campo de investigação de *clusters* industriais. Para tal, esses autores investigaram o arcabouço teórico histórico e propuseram um *framework* que delineou as teorias de *clusters*. As três vertentes teóricas definidas foram: conceito das externalidades locais, a perspectiva territorial e a questão da competitividade.

As economias locais são baseadas nas reflexões de Marshall (1890) sobre Distritos Industriais e tem interesse na distribuição desigual de atividades econômicas no espaço, especificamente na tendência de firmas relacionadas a se localizarem em lugares específicos ao longo do tempo. As principais fontes de economias de aglomeração observadas foram: transbordamento de conhecimento entre firmas, especialização do mercado de trabalho e redução de custos com insumos e o termo “Distrito Industrial” se refere ao local em que ocorre esse tipo este fenômeno. Deve-se considerar que as firmas estão diretamente ligadas pelas relações comerciais de compra/venda e indiretamente por meio do mercado de trabalho e dos serviços públicos e privados providos no ambiente econômico. A síntese da teoria de *clusters* baseada em economias locais encontra-se no Quadro 1.3.

Quadro 1.3 - Teoria de *clusters* baseada em economias locais.

O que:	Requisitos para localização das empresas e externalidades.
Como:	Interação entre oferta e demanda por requisitos de localização.
O porquê:	Existência: Dada pelo efeito simultâneo das forças centrípetas e centrífugas. Extensão: Limitada pelas forças centrífugas. Exaustão: Fatores de localização totalmente explorados.

Fonte: Adaptado de Maskell e Kebir (2006).

As variáveis principais (“o que”) são os requisitos de localização das firmas, como fonte de matéria prima, condições climáticas, infraestrutura de transportes, preço da terra, acesso aos serviços, entre outras. Adicionalmente, as externalidades são as economias externas de custos para as firmas decorrentes do crescimento da produção do distrito como um todo, enquanto as economias internas são as fontes de retornos crescentes devido ao crescimento da firma individual (BERGMAN; FESER, 1999). Por sua vez, as economias locais são resultantes de uma combinação bem sucedida entre os requisitos de localização e a oferta dos fatores de localização (“como”).

O argumento da “existência” é dado pela interação entre as variáveis que atraem as firmas (forças centrípetas) e as que repelem as mesmas (forças centrífugas) que conjuntamente determinam a localização geográfica das firmas. As forças centrípetas (ou *spillovers* positivos) podem ser vantagens de custos em transportes, infraestrutura compartilhada, mão de obra e fornecedores especializados, etc. As forças centrífugas, ou *spillovers* negativos, representam o argumento de “extensão” e incluem os custos de congestionamento, elevação do preço da terra, da mão de obra ou bens e serviços locais. Por último, a “exaustão” do *cluster* se encontra quando um ou mais fatores vitais de localização são completamente explorados.

Uma nova geração de origem italiana interessada na base teórica Marshalliana para o estudo das aglomerações de firmas surgiu nos anos 70/80 (BECATTINI, 1979), também conhecidos como Neo-marshallianos. No entanto, conforme Maskell e Kebir (2006), nessa nova abordagem o ciclo de vida apresentado no Quadro 1.3 foi substituído por um modelo unilateral que atualizou o argumento “existência”, mas desconsiderou os demais. Para os Neo-marshallianos, o argumento “existência” é baseado na combinação das vantagens da coerência social, flexibilidade relacional, interação intra-*cluster* e no aprofundamento do conhecimento local. Em outras palavras, o foco deixa de ser nos efeitos econômicos das aglomerações e passa para as fundações sociais, culturais e institucionais do desenvolvimento regional em que a competitividade e a eficiência da produção estão relacionadas com condições socioculturais (TRIPPL; BERGMAN, 2013).

A visão teórica com foco no desenvolvimento regional, ou abordagem do *Milieu* Inovador, está preocupada com tecnologia, organização e, principalmente, o território onde se dá a interação entre inovação e fatores de localização. Estes três fatores (“o que”) compõe um ambiente localizado, sem fronteiras físicas e com certo grau de homogeneidade comportamental. Deve-se considerar que existem trajetórias de desenvolvimento desiguais entre regiões geográficas e na difusão de novas tecnologias (AYDALOT, 1986). A síntese da teoria de *clusters* com foco no desenvolvimento territorial encontra-se no Quadro 1.4.

Quadro 1.4 - Síntese da teoria de *clusters* com foco no desenvolvimento regional.

O que:	Tecnologia, organização e território.
Como:	Agentes independentes, dinâmica de aprendizado e lógica organizacional.
O porquê:	Existência: relacionamentos espontâneos que geram aprendizado coletivo. Extensão: capital relacional, valores locais. Exaustão: comportamento oportunista.

Fonte: Adaptado de Maskell e Kebir (2006).

O “como” na visão do desenvolvimento regional é composto por um conjunto de agentes que tomam decisões estratégicas na gestão de recursos materiais (infraestrutura, maquinário, financiamentos, etc.) e imateriais (conhecimento, instituições, etc.) e que possuem uma capacidade de se adaptar as mudanças no ambiente econômico (dinâmica de aprendizado). Estes atores também estão sujeitos a uma lógica organizacional de cooperação para inovação e desenvolvimento de relacionamentos comerciais e não comerciais.

O argumento da “existência” advém dos relacionamentos em uma área geográfica que geram uma dinâmica de aprendizado coletivo capaz de reduzir as incertezas do processo inovativo, facilitar o conhecimento mútuo, colaboração, troca de informações e relações de confiança. A “extensão” se dá pela cooperação que leva a construção de capital relacional e mobilização de recursos não necessariamente monetários, os valores locais (empreendedorismo, familiares, profissionais, etc.) levam os agentes a inovarem e estabelecerem relações de confiança. Estes relacionamentos delimitam o *Milieu* e seus participantes. Por fim, a “exaustão” do *cluster* se dá quando o interesse individual se sobrepõe ao interesse comum e se observa comportamento oportunista, ou ainda quando o nível de abertura externa se torna insuficiente para garantir a expansão ou a renovação de relações cooperativas ou a atualização das tecnologias.

A vertente promovida por Porter (1990; 1998) buscou fornecer uma teoria de competitividade nacional, estadual e local em um contexto de economia global e não necessariamente uma teoria de *clusters* industriais. Segundo essa teoria, as nações avançam não em indústrias isoladas, mas em *clusters* de indústrias com conexões verticais e horizontais. A síntese da teoria de *clusters* com foco na competitividade encontra-se no Quadro 1.5.

Quadro 1.5 - Síntese da teoria de *clusters* com foco na competitividade.

O que:	Condições dos recursos, condições da demanda, indústrias relacionadas, rivalidade/competição.
Como:	Conexão entre os determinantes supracitados.
O porquê:	Existência: proximidade. Extensão: dualidade Marshalliana da dispersão x localização. Exaustão: forças <i>intracluster</i> e também influências externas.

Fonte: Adaptado de Maskell e Kebir (2006).

O modelo conhecido como “Diamante de Porter” (PORTER, 2000) tem em seus fatores de competitividade o próprio “o que” da teoria e no relacionamento entre eles está o “como”. Os recursos (ou fatores) estão relacionados com as condições de disponibilidade,

custos, qualidade e especialização dos insumos para produção, como recursos naturais, mão de obra, infraestrutura, tecnologia e conhecimento local, entre outros. As condições de demanda abrangem consumidores finais com necessidades de bens e serviços especializados (ou sofisticados) e também as indústrias relacionadas que consomem bens intermediários. A demanda pode ser suprida por bens/serviços domésticos ou internacionais. A presença de firmas, setores correlatos também competitivos e fornecedores especializados contribuem para determinar a competitividade global. Por fim, um contexto de rivalidade e competição no ambiente local leva a investimentos apropriados e inovações sustentáveis.

O argumento de “existência” do *cluster* está relacionado com os benefícios da proximidade geográfica, cultural e transacional, como fluxos comerciais, trocas em Pesquisa & Desenvolvimento, relacionamento com fornecedores, entre outros. Como consequência, os *clusters* atraem profissionais qualificados e empreendedores que reforçam o ambiente local em busca de oportunidades e desenvolvimento (PORTER, 1990; 1998). A “extensão” do *cluster*, no entanto, segue a visão tradicional da teoria Marshalliana baseada forças de dispersão e localização. O último argumento, “exaustão”, parte de fatores internos aos *clusters* (queda na rivalidade local, aumento na rigidez e regulação, entre outros) e influências externas (descontinuidade tecnológica, deterioração da condição dos fatores ou da demanda).

As três teorias de *clusters* definidas por Maskell e Kebir (2006) foram corroboradas e discutidas por Trippel e Bergman (2013) que cobriram, além dos fatores econômicos, as relações sociais e institucionais inerentes à coordenação dos agentes econômicos. O foco da discussão, no entanto, foi nos elementos similares e distintos entre os conceitos de Distritos Industriais Marshallianos e Neo-Marshallianos, os *Milieus* Inovadores e os *Clusters* Industriais de Porter. As dimensões relevantes consideradas são: geografia e espaço; agentes e interações; indústrias e inovações; ambiente, cooperação e competição.

Todos os conceitos são baseados no entendimento do papel do espaço e dos fatores geográficos que os delimitam de uma forma específica (geografia e espaço). O Distrito Marshalliano é delimitado por uma densidade suficientemente alta e distâncias físicas baixas entre os fatores de produção e as firmas, ao ponto de criar uma vizinhança. Na vertente Neo-Marshalliana, o aglomerado é definido por pequenas unidades governamentais, denominadas distritos políticos, que são parceiras e facilitadoras no funcionamento do aglomerado. O *Milieu* é delimitado por um “espaço relacional” interligado a uma rede de contatos entre organizações interessadas na geração de inovações. Por fim, os *Clusters* Industriais apresentam as definições geográficas mais flexíveis, determinadas principalmente pelas

externalidades positivas geradas por transbordamento de conhecimento (Marshall-Arrow-Romer, ou M-A-R) que resultam da co-localização de firmas de uma mesma indústria e geram inovações e crescimento. Nas versões recentes, com o avanço dos transportes e comunicações, podem se expandir para áreas metropolitanas e extensas regiões.

Os atores no Distrito Marshalliano são as pequenas e médias empresas independentes que buscam maximizar o lucro e que encontram mais vantagens de mercado por estarem dentro do distrito do que se estivessem fora dele. Já o Distrito Neo-Marshalliano conta com diversos agentes, como pequenas firmas familiares que expandem seus arranjos contratuais com outras firmas, empregados com mobilidade laboral entre firmas e os distritos políticos locais que estabelecem as diretrizes e práticas internas. A maximização de lucros é um entre vários objetivos, como domínio das técnicas de manufatura, orgulho da região e reconhecimento por parte dos consumidores de alta renda. Os agentes do *Milieu* constituem redes que envolvem diversas organizações sem papéis específicos, embora motivadas pelo orgulho da região e o esforço para inovar. Os agentes secundários são compostos por agentes políticos e do mercado local que também fazem parte do sistema de produção. Por fim, os *Clusters* Industriais são compostos por firmas de todos os tamanhos e formas organizacionais, é possível ainda que o próprio *cluster* se torne um agente e assuma uma forma organizacional de governança própria. As principais motivações são o Retorno sobre o Investimento (ROI), ganhos de produtividade e de parcelas de mercado.

Em relação às interações entre agentes, os distritos tradicionais são interligados por vínculos com fornecedores e fluxos inconscientes de conhecimento, ideias e de trabalhadores (*spillovers*) entre firmas próximas geograficamente. Por sua vez, a nova versão dos distritos apresenta ênfase nas redes colaborativas entre pequenas firmas e organizações de suporte que facilitam a troca de conhecimento tácito, compra conjunta de insumos, ou estratégias de acesso conjunto a tecnologias e serviços financeiros. O foco dos *Milieus* está nas relações sociais informais em nível local que promove o aprendizado coletivo, conseqüentemente, as inovações por parte das empresas, além do papel dos vínculos externos. Os *Clusters* Industriais de Porter consideram as ligações das empresas com consumidores e fornecedores, assim como as trocas verticais entre empresas. As interações ocorrem em fluxos insumo-produto e de conhecimento.

A visão de indústria Marshalliana não se aplica a um setor específico, ao contrário da versão Neo-Marshalliana, focada em indústrias maduras, com produção predominante de bens de consumo que se aproveitam de inovações incrementais e tecnologias básicas, são

intensivas em *design* e estilo, demandam trabalhadores altamente qualificados em artesanato e é direcionada para consumidores exigentes em mercados internacionais. Os *Milieux* tem interesse na relação entre alta tecnologia e inovações industriais para reviver antigos centros produtivos, assim como estimular os novos, sem o aparente foco em setores específicos. Os *clusters* são relevantes para qualquer indústria em que a competitividade é determinada pelas externalidades MAR ou pelo Diamante de Porter. Em geral, essas firmas competem no mercado internacional, embora as concorrentes locais também se beneficiem dos *clusters*.

As inovações (em produtos, processos e organizacionais) observadas por Marshall são apoiadas pela difusão de novas ideias e melhores práticas, transbordamentos de segredos comerciais e pelo estoque de conhecimento acumulado que é constantemente melhorado pela interação entre firmas localizadas e pela transferência de conhecimento entre gerações. A atividade tecnológica Neo-Marshalliana é baixa, visto que é determinada por pequenas empresas, em setores tradicionais e com baixo investimento em pesquisa e desenvolvimento. Assim, o foco está na qualidade, refinamento artesanal e inovações colaborativas e incrementais geradas por organizações flexíveis. Os pesquisadores relacionados aos *Milieux* ressaltam como determinantes das inovações o aprendizado coletivo local e a adoção/exploração de um fluxo interno de competências, tecnologia e conhecimentos complementares gerados externamente. O processo de inovação no *Cluster* Industrial de Porter é resultado das externalidades MAR e das decisões estratégicas tomadas pelas firmas, como *joint-ventures* pré-comerciais e aquisição de conhecimento interno e externo. Os elementos do Diamante de Porter também têm efeitos positivos sobre a inovação (especialmente as radicais) como a troca de conhecimento e desenvolvimento conjunto com fornecedores. A alta competição local e consumidores exigentes também aumentam a pressão por inovações, sejam de produto, processo ou organizacional.

O conceito de ambiente, entendido como contexto cultural e institucional, está relacionado com seu papel no desenvolvimento regional, competitividade e geração de inovações. Em Marshall a noção de “atmosfera industrial” engloba o papel dos fatores sociais e a identidade cultural no suporte aos fluxos de conhecimento locais compartilhados e ao desenvolvimento do Distrito Marshalliano. Os Neo-Marshallianos corroboram a percepção sobre os fatores não econômicos que são fundamentais para o sucesso econômico, como valores comuns, confiança e identidade coletiva com a região e seus produtos. Assim, as relações econômicas e sociais são entrelaçadas, geram colaboração entre os agentes e permitem uma esfera de cooperação e competição. A abordagem dos *Milieux* também segue

esta linha referente aos valores comuns e visões compartilhadas, como desenvolvimento de longo prazo frente a uma maximização de lucros no curto prazo, em que a dimensão sociocultural colaborativa leva a geração de inovações e desenvolvimento. O conceito de *Clusters* Industriais provê pouca ênfase aos fatores não econômicos em comparação aos outros, mas considera que confiança e capital social são fundamentais para interação entre os agentes e para criação de valor.

Um dos temas mais controversos entre as teorias apresentadas é a relação entre competição e cooperação. No Distrito Marshalliano a competição é o principal determinante do aglomerado, enquanto a geração de conhecimento e inovação é resultado da colaboração. Na versão Neo-Marshalliana a disposição para cooperação é fundamental para gerar inovações e eficiência, tão importante quanto os princípios da competição. A vertente do *Milieu* coloca um equilíbrio entre competição e cooperação não muito claro, com uma tendência de considerar a cooperação como dimensão mais relevante. Por último, o *Cluster* Industrial de Porter não enfatiza o papel da cooperação, destacando o processo de competição como determinante da competitividade. No entanto, algumas vantagens da colaboração são reconhecidas principalmente nas relações verticais com indústrias correlatas e com as instituições locais, enquanto a cooperação com os concorrentes é vista como perigosa.

As características das três principais vertentes da teoria de *clusters* foram sistematizadas no Quadro 1.6.

Quadro 1.6 - Principais características das teorias de *clusters*

	Distritos Marshallianos	Distritos Neo-Marshallianos	Milieux Inovadores	<i>Clusters Industriais</i>
Geografia/espço:	Distância física/vizinhança.	Distrito político comum.	Espaço coextensivo em rede.	M-A-R + fatores de globalização.
Principais atores:	Pequenas firmas.	Firmas familiares, funcionários locais.	Instituições políticas e sociais, produtores de conhecimento.	Firmas de todos os tamanhos + grupos industriais.
Motivações:	Maximização do lucro.	Domínio das técnicas de produção, orgulho (lugar/produto), alto valor, mercados de luxo.	Crescimento, orgulho (lugar/produto), referência em desenvolvimento tecnológico.	M-A-R + ROI, oportunidades de mercado e produção oferecidas pela globalização.
Interações (relacionamento entre atores):	M-A-R.	M-A-R, relacionamentos de confiança e de longo prazo.	Relacionamentos sociais informais, aprendizado coletivo, relacionamentos formais externos.	M-A-R, cadeia de valor, trocas verticais, alianças empresariais.
Indústrias, tecnologia e mercados:	Indústrias comuns no século 19.	Consumo discricionário / baixo e média manufatura.	Alta tecnologia.	Indústrias exportadoras, alta e baixa tecnologia.
Inovação:	Difusão de ideias, transbordamento de conhecimento local.	Refinamento manufatureiro, organizações flexíveis, inovações incrementais.	Aproveitamento interno de inovações externas.	M-A-R, <i>joint ventures</i> pré-comerciais, fontes de conhecimento.
Ambiente (cultural e institucional, extra econômico):	Atmosfera industrial.	Valores familiares e comunitários, integração entre comunidade e negócios.	Valores e visões compartilhadas, coesão regional.	Confiança e capital social ligam os agentes do <i>cluster</i> .
Competição e cooperação:	Moderada, competição é a força motriz do distrito e cooperação é a pré-condição para inovação coletiva.	Complexa, alta importância da cooperação como mecanismo de coordenação.	Fraco balanço entre competição e cooperação.	Elevada competição, cooperação menos importante.

Fonte: Adaptado de Trippel e Bergman (2013).

1.3.2 Aglomerações, Complexos e Redes Sociais

McCann (2008, 2013) trabalhou com uma estrutura analítica subdividida em três vertentes teóricas que explicam o fenômeno dos *clusters* industriais (Quadro 1.7): Aglomerações Puras, Complexos Industriais e Redes Sociais. Um *cluster* pode apresentar características de mais de um dos tipos ideais. As pressuposições implícitas nessa classificação são pautadas em uma adaptação da Teoria dos Custos de Transação, porém o interesse está na distinção entre a natureza das transações entre as firmas dentro dos *clusters* e na distinção entre as transações que ocorrem dentro e fora da aglomeração espacial (transações específicas em localização).

Quadro 1.7 - Tipos ideais de *clusters* industriais

Características	Aglomeração Pura	Complexo Industrial	Rede Social
Tamanho da firma:	Atomizada.	Podem haver grandes.	Variável.
Relacionamentos:	Não identificável; fragmentado; instável.	Identificável; estável; trocas.	Confiança; lealdade; <i>joint venture</i> ; não oportunista.
Membros:	Aberto.	Fechado.	Parcialmente aberto.
Acesso:	Aluguel; locação necessária.	Investimento interno; locação necessária.	Experiência; história; locação necessária, mas não suficiente.
Resultados:	Apreciação do aluguel.	Sem efeito no aluguel.	Capitalização parcial do aluguel.
Noção de espaço:	Urbano.	Regional, não urbano.	Regional, não urbano.
Exemplo:	Economia urbana competitiva.	Complexo produtivo de aço.	Novas áreas industriais.
Abordagem analítica:	Modelos de aglomeração pura.	Modelos insumo-produto.	Redes sociais (Granovetter).

Fonte: Adaptado de McCann (2013).

No modelo de Aglomeração Pura as relações entre firmas são momentâneas. Não existem firmas com poder de mercado (mercado atomizado) e o relacionamento com outras firmas e com os clientes responde a oportunidades de arbitragem de mercado, gerando intensa competição local, ausência de lealdade e de relações de longo prazo. Todas as firmas se beneficiam das externalidades positivas, simplesmente devido a sua presença na aglomeração e cujo custo desta presença é o aluguel no mercado imobiliário local. Não existem barreiras a entrada e, como consequência, o indicador de desempenho do *cluster* é a valorização dos aluguéis locais. A estrutura de mercado implícita é a competição monopolística e espaço é

urbano, ou seja, este tipo de aglomeração existe somente em cidades individuais. Este tipo ideal foi mais bem representado por Marshall (1860) e Krugman (1991).

O Complexo Industrial é caracterizado por relações de longo prazo, previsíveis e que envolvem transações frequentes entre as firmas pertencentes ao *cluster*. Este relacionamento é dado principalmente por relações de troca (compra de insumos e venda de produtos) que determinam o padrão de localização. Cada firma realiza investimentos de longo prazo, principalmente em capital físico e imobiliário para se tornarem parte do agrupamento. O acesso é restrito, devido aos elevados custos de entrada e saída e a lógica é que a proximidade espacial minimize os custos de transação entre firmas, em especial com transporte.

Apreciação imobiliária não é uma característica observável, pois o terreno já foi adquirido pela firma e não está a venda. A estrutura de mercado nos Complexos Industriais é oligopolista e a noção de espaço é local, mas não necessariamente urbana, e pode se estender por outras regiões. O transbordamento de conhecimento informal também não é verificado, como no caso das Aglomerações Puras, o Complexo Industrial reflete os retornos internos a escala e a características de polos de crescimento local. Este tipo de *cluster* industrial foi originalmente discutido em Isard, Schooler e Victorisz (1959).

A teoria das Redes Sociais argumenta que relações de confiança entre os tomadores de decisão nas organizações são tão importantes quanto às decisões sobre hierarquia nos modos de coordenação das organizações individuais. Esta confiança se manifesta de várias formas, mas a principal é a ausência do oportunismo, em que as firmas não temem represálias após alguma reorganização das relações entre firmas. O modelo de Rede Social é “aespacial”, porém, pode-se argumentar que a proximidade tende a reforçar as relações de confiança no ambiente de negócios local. Assim, a participação na rede é parcialmente aberta e o pagamento de aluguéis não garante acesso, mas eleva estas chances. O tipo ideal de Redes Sociais está associado às Novas Áreas Industriais de Scott (1988) e também diversos de seus elementos são encontrados no conceito de *Clusters* Industriais de Porter (1990, 1998).

Autores como Palacios (2005), no entanto, consideram que um Distrito Industrial é um *Cluster* Industrial e vice versa, e os dois constituem um Complexo Industrial que, por sua vez, pode ser parte de um *cluster*. Esse autor argumenta que polos de crescimento, distritos e complexos podem ser chamados de *clusters* em uma conotação genérica do termo.

Os fundamentos teóricos que explicam a localização de firmas também foram explorados por Brosnan et al. (2016) com objetivo de identificar as influências comuns e as raízes intelectuais e os contrastes entre os *Clusters* Industriais de Porter e seus antecessores,

em particular os Distritos Industriais e os Complexos Industriais. A teoria de *clusters* pode ainda ser entendida como a forma que os *clusters* industriais são definidos, descritos e explicados. A divisão analítica proposta por Vom Hofe e Chen (2006) sobre as vertentes teóricas comparáveis dos *clusters* está de acordo com a de Brosnan et al. (2016) em relação aos conceitos fundamentados nas Aglomerações Puras, nos Complexos Industriais e em uma base teórica mais abrangente que é a fusão entre a visão locacional (relações horizontais), insumo-produto (relações verticais) e as características da teoria do Diamante de Porter.

O primeiro conceito, Aglomerações Puras, segue os princípios teóricos das economias de localização de Marshall (1890) e pode ser definido de forma generalista como um grupo de estabelecimentos que pertencem ao mesmo setor industrial em uma mesma região geográfica. Embora observada qualitativamente em Marshall (1890), uma forma recorrente de se traduzir a ideia das economias de localização em um conceito de *cluster* é por meio de métodos e medidas de especialização regional e concentração industrial, como o Quociente Locacional e outros índices. Deve-se considerar que transbordamentos de conhecimento, especialização do mercado de trabalho e compartilhamento de insumos potencializa o agrupamento espacial de firmas e a especialização regional.

O segundo conceito, Complexos Industriais, é derivado das relações insumo-produto entre indústrias concebido por Isard, Schooler e Vietorisz (1959). Esses autores expandiram o conceito tradicional de *clusters* em que todas as firmas pertencem ao mesmo setor ao considerar os vínculos entre indústrias distintas. Assim, esse conceito é baseado em encadeamentos para frente e para trás que permitem a formação de um complexo de firmas relacionadas na cadeia de valor, mas não necessariamente pertencem ao mesmo setor industrial. Nessa abordagem, a utilização de matrizes insumo-produto é o ponto de partida para a identificação de setores homogêneos que compõem os Complexos Industriais.

O terceiro conceito, *Clusters* de Porter (1990), abrange um conjunto variado de argumentos que inclui economias de localização e urbanização, retorno de escala interno a firma, relacionamentos na cadeia de valor, inovação tecnológica, entre outros. A contribuição única da teoria de Porter está nas estratégias de competição empresarial como promotoras do desenvolvimento regional. Entretanto, a abrangência do conceito de Porter é mais subjetiva do que lógica, sem uma rigorosa aplicabilidade prática, logo não oferece uma metodologia específica que viabilize a tradução do conceito em ações efetivas. Após críticas, Delgado, Porter e Stern (2014) e Porter (2003) incorporaram índices de concentração, matrizes insumo produto, algoritmos de agrupamento e opiniões de especialistas no mapeamento de *clusters*.

1.4 EXPERIÊNCIAS DE POLÍTICAS DE *CLUSTERS* NO BRASIL

No Brasil, a denominação mais popular derivada do conceito de *clusters* industriais é a de Arranjos Produtivos Locais – APLs. Segundo Cassiolato e Lastres (2003) um APL pode ser descrito como um conjunto de agentes econômicos, políticos e sociais localizados em uma mesma região geográfica, que desenvolvem atividades produtivas correlacionadas e que apresentam vínculos relevantes de produção, interação, cooperação e aprendizagem. Em geral, incluem empresas – produtoras de serviços finais, fornecedoras de equipamentos e outros insumos, prestadoras de serviços, comercializadoras, clientes, cooperativas, associações e representações, etc. – e organizações voltadas para a formação e treinamento de recursos humanos, informação, pesquisa, desenvolvimento e engenharia, promoção e financiamento.

A abordagem de APLs consolidada no final dos anos 90 pela Rede de Pesquisa em Arranjos e Sistemas Produtivos e Inovativos Locais (REDESIST) - grupo alocado na Universidade Federal do Rio de Janeiro - teve uma rápida difusão pelo território nacional e substituiu termos similares na maioria das diretrizes de políticas públicas. Desde então, um esforço significativo para compreender e promover esse conceito vem ocorrendo, seguido de um processo de experimentação e aprendizado. Os APLs foram conduzidos para as demais esferas da gestão pública, por meio da criação de núcleos gestores em nível estadual e/ou municipal, além do envolvimento da iniciativa privada (LASTRES et al. 2014).

A popularização das políticas de estímulo aos APLS se deu em um contexto de transformações relevantes ocorridas na economia mundial e nacional nos 1990/2000. Em ambiente internacional, Matos et al. (2015) destacaram a quebra da bolsa Nasdaq em 2000, a crise financeira em 2008, a emergência de economias asiáticas como a China e aumento do protecionismo tecnológico por parte de países desenvolvidos. Com o processo de abertura econômica da economia nacional nos anos 90 aliado aos eventos em âmbito internacional, foi verificado uma queda da participação industrial na economia, em especial da indústria de transformação, além do esvaziamento dos sistemas produtivos/inovativos e da redução da contribuição de manufaturas para a balança comercial. Esse novo cenário demandou a retomada de políticas públicas como estratégia de desenvolvimento socioeconômico interno, destacando-se o foco nos APLs (MATOS et al., 2015).

As políticas de apoio aos APLs no Brasil perpassaram por vários órgãos, principalmente a partir do ano 2000, e em âmbito federal pode-se destacar: Ministério de

Ciência e Tecnologia (MCT), Ministério da Integração (MI), Serviço de Apoio às Micros e Pequenas Empresas (SEBRAE), Desenvolvimento, Indústria e Comércio Exterior (MDIC) e o Banco de Desenvolvimento econômico e Social (BNDES).

A iniciativa melhor estruturada de cunho governamental é o Grupo de Trabalho Permanente para APLs (GTP-APL) criada no âmbito federal sob a liderança do MDIC, em 2004, para sistematizar a atuação institucional com objetivo de desenvolver APLs, em conjunto com aproximadamente 33 organizações públicas e privadas. O governo, por meio do GTP-APL, deve atuar como direcionador de ações para a criação, consolidação e desenvolvimento de APLs, de forma articulada com as demais políticas públicas voltadas para o setor produtivo e o desenvolvimento econômico regional. Como conquistas do grupo destaca-se que política de APLs se difundiu, a necessidade de coordenação de ações foi reconhecida pelos agentes econômicos nacionais e locais e a cooperação passou a ser entendida pelos empresários como fator chave para o desenvolvimento empresarial, setorial e da economia como um todo (BELLUCCI et al., 2014).

Para articular e coordenar as ações e medidas dos agentes envolvidos no GTP-APL foi conduzido o estudo “Identificação, mapeamento e caracterização estrutural de arranjos produtivos locais no Brasil” pelo Instituto de Pesquisa Econômica e Aplicada - IPEA (SUZIGAN, 2006) que teve como objetivo direcionar políticas públicas e ações institucionais de fomento a atividades produtivas e inovativas em APLs. Esse estudo aplicou métodos quantitativos baseados em indicadores de concentração setorial e especialização regional, como o Gini Locacional e o Quociente Locacional, e utilizou dados da Relação Anual de Informações Sociais do Ministério do Trabalho e Emprego (RAIS/MTE).

Um dos pontos cruciais para qualidade na elaboração e gestão de projetos é a qualificação dos gestores e dos formuladores das políticas de desenvolvimento regional. Assim, o Plano Nacional de Capacitação em Gestão de APLs (PNC), também no âmbito do GTP-APL, buscou mitigar esse gargalo ao seguir experiências em âmbito internacional, como os indicadores de melhores práticas de gestão de *clusters* industriais que orientam os programas de treinamento da *European Cluster Excellence Initiative*. Outras tendências do PNC foram o estímulo às possibilidades de interação entre os agentes, consolidação das redes de relacionamentos e a integração de recursos complementares no âmbito dos arranjos (BELLUCCI et al., 2014).

Outro agente importante para os APLs no Brasil é o BNDES. Um dos eixos de orientação política do BNDES é a redução de desequilíbrios e a promoção de

desenvolvimento socioeconômico regional, de forma integrada, dinâmica e sustentável e o foco em APLs é um meio de potencializar estas ações (LASTRES et al., 2014). Em 2007, foi criado o Comitê de arranjos produtivos, inovação, desenvolvimento local, regional e socioambiental (CAR-IMA) com objetivo de buscar e propor mecanismos de apoio aos APLs. Contemplar demandas de diferentes regiões, arranjos e empresas representa um desafio para o BNDES por estar fora de suas atuações tradicionais, assim, o banco estabeleceu duas vertentes de atuação para o desenvolvimento: uma no entorno de grandes projetos e outra em regiões menos atendidas pelo banco (LASTRES et al., 2014).

Outra contribuição importante do BNDES é o envolvimento em estudos, pesquisas, seminários e oficinas, seja como realizador ou patrocinador, com objetivo de desenvolver o planejamento e a execução de políticas e instrumentos de suporte aos APLs. O principal projeto financiado foi o “Mapeamento e análise das políticas para Arranjos Produtivos Locais no Brasil” desenvolvido entre 2009 e 2010 que além de identificar APLs apoiados em 22 estados, também apontou casos “invisíveis”, ou seja, não contemplados por políticas públicas (LASTRES et al., 2014).

A análise do potencial dos APLs ainda é limitada, segundo Bellucci et al. (2014), pela falta de dados necessários para pesquisas sobre territórios. Dados sobre os arranjos e sobre as empresas participantes são escassos, assim como informações sobre organizações informais que não estão vinculadas às redes de relacionamentos. A confiabilidade dos dados é necessária para gerar indicadores precisos que apontem as demandas territoriais e a falhas das políticas públicas de APLs (BELLUCCI et al., 2014). O Observatório Brasileiro de APL é uma tentativa de gerar este tipo de informações.

Outras participações relevantes foram destacadas por Tatsch et al. (2015), em especial a do MCT, MI e SEBRAE. O MCT inaugurou o início da política de APLs no final dos anos 90, com o financiamento de estudos temáticos. O tema também passou a fazer parte do Plano Plurianual 2000/2003. No ano 2000, a indústria de base mineral foi um dos primeiros setores com objetivo de identificar e caracterizar empiricamente os APLs com base em dados quantitativos. O MI começou sua atuação no estímulo aos APLs em 2003 por meio do Plano Nacional de Desenvolvimento Regional (PNDR), o foco foi em regiões prioritárias de baixa renda, estagnadas ou dinâmicas. Em 2012, o PNDR II propôs a atualização da política baseando-se em redes de APLs setorialmente interligados. O SEBRAE buscou promover a competitividade e a sustentabilidade de micro e pequenas empresas ao apresentar como uma de suas prioridades a atuação em APLs. A partir de 2007, passou a atuar por meio de projetos

em diversas formas de organizações produtivas, como APLs, polos e redes. As etapas seguidas pelo SEBRAE são: identificação, seleção e atendimento/apoio (via projetos).

A política voltada para APLs no âmbito do GTP-APL também está integrada a outras iniciativas governamentais como o PNDR e o Plano “Brasil Maior” 2013/2015 (política industrial) que apresentam objetivos comuns de reversão de desigualdades regionais e desenvolvimento econômico. Porém, segundo Bellucci et al. (2015), uma limitação é que essa articulação ainda não é suficiente para evitar esforços duplicados, promover ações coordenadas e chegar em resultados mais efetivos.

A síntese da experiência brasileira com as políticas voltadas para APLs foi apresentada no Quadro 1.2 (TATSCH et al., 2015).

Quadro 1.2 – Análise das políticas para aglomerados produtivos no Brasil.

Itens	Características
Objetivo e metas	Desenvolvimento econômico; Redução das desigualdades sociais e regionais; Inovação tecnológica; Expansão e modernização da base produtiva; Crescimento do nível de emprego e renda; Redução da taxa de mortalidade de micro e pequenas empresas; Aumento da escolaridade e da capacitação; Aumento da produtividade e da competitividade; e Aumento das exportações.
Focos de atuação	Crédito e financiamento; Governança e cooperação; Tecnologia e inovação; Formação e capacitação; Acesso aos mercados nacional e internacional; Fomento à atuação sistêmica; Fortalecimento de capacitações produtivas e inovativas; Coesão com o desenvolvimento local; e Sustentabilidade econômica, política / institucional, social e ambiental.
Aparato legal e instrumentos	Portarias interministeriais; Plano de desenvolvimento do APL; e Editais.
Estágio do arranjo a ser apoiado	Arranjos já existentes.

Fonte: Adaptado de Tatsch et al. (2015).

1.5 CONSIDERAÇÕES FINAIS

Exposto o problema de pesquisa sobre a limitação dos métodos vigentes para o mapeamento de aglomerações de firmas em direções específicas, o objetivo geral do presente trabalho foi propor uma abordagem geoestatística para o mapeamento de potenciais *clusters* industriais. Essa nova abordagem pode contribuir com o avanço científico do campo de conhecimento sobre aglomerações de firmas e também fornecer ferramentas para subsidiar a elaboração políticas econômicas de estímulo à formação de *clusters* industriais.

Em termos teóricos, foi identificado que a vertente dos Distritos Industriais considera que a localização de firmas é um indicador da ocorrência das economias de aglomerações, então geralmente é analisada por índices de concentração. A corrente teórica dos Complexos Industriais é fundamentada nos encadeamentos entre setores similares, por isso baseia suas análises em matrizes insumo-produto. Por fim, os *Clusters* Industriais de Porter buscam agregar os quadros analíticos de concentração industrial e de encadeamentos interindustriais.

Como limitações das teorias de aglomerações e dos *clusters*, pode-se dizer que ambas estão preocupadas com o fenômeno da concentração geográfica de firmas. Porém, a visão da economia das aglomerações está preocupada em explicar como as questões geográficas influenciam a vantagem competitiva das regiões, com pouco detalhamento no nível das firmas. Por sua vez, a visão estratégica dos *clusters* se concentra nos determinantes de competitividade das firmas, mas com pouca atenção em questões geográficas, como as relações de vizinhança.

Tais limitações podem ser contempladas com uso da abordagem geoestatística. Para justificar tal afirmativa, a seguir, é sistematizada a literatura internacional acerca dos métodos para identificar e mensurar os *clusters* industriais, por meio da bibliometria. Bem como é aplicada a abordagem geoestatística na análise das relações de proximidade entre firmas e de concentração setorial. Essa abordagem foi aplicada em dados simulados e na indústria do café em Minas Gerais.

REFERÊNCIAS

- AHN, M. J. et al. Death of distance?—biotechnology agglomeration patterns, alliance proximity, and firm performance. **International Journal of Innovation and Technology Management**, v. 6, n. 03, p. 247-264, 2009.
- ALFARO, L.; CHEN, M. X. The global agglomeration of multinational firms. **Journal of International Economics**, v. 94, n. 2, p. 263-276, 2014.
- AUSTRIAN, Z. Cluster case studies: The marriage of quantitative and qualitative information for action. **Economic Development Quarterly**, v. 14, n. 1, p. 97-110, 2000.
- AYDALOT, P. **Milieux Innovateurs en Europe**. Paris: Groupe de Recherche Europe' en sur les Milieux Innovateurs (GREMI), 1986.
- BECATTINI, G. Dal settore industriale al distretto industriale: Alcune considerazioni sull'unita` di indagine dell'economia industriale. **Rivista di Economia e Politica Industriale**, v.1, p.7–21, 1979.
- BEHRENS, K.; BOUGNA, T. An anatomy of the geographical concentration of Canadian manufacturing industries. **Regional Science and Urban Economics**, v. 51, p. 47-69, 2015.
- BELLUCCI, A. C. S. et al. GTP APL: dez anos de avanços e perspectivas de futuro. **Boletim Regional, Urbano e Ambiental - IPEA**, v. 10, n.2, p. 13-17, 2014.
- BERGMAN, E. M.; FESER, E. J. **Industrial and regional clusters: concepts and comparative applications**. Morgantown: WVU Regional Research Institute, 1999.
- BICKENBACH, F.; BODE, E. Disproportionality measures of concentration, specialization, and localization. **International Regional Science Review**, v. 31, n. 4, p. 359-388, 2008.
- BICKENBACH, F.; BODE, E.; KRIEGER-BODEN, C. Closing the gap between absolute and relative measures of localization, concentration or specialization. **Papers in Regional Science**, v. 92, n. 3, p. 465-479, 2013.
- BILLINGS, S. B.; JOHNSON, E. B. A non-parametric test for industrial specialization. **Journal of Urban Economics**, v. 71, n. 3, p. 312-331, 2012.
- BRAUNERHJELM, P.; BORGMAN, B. Geographical concentration, entrepreneurship and regional growth: Evidence from regional data in Sweden, 1975-99. **Regional Studies**, v. 38, n. 8, p. 929-947, 2004.
- BROSNAN, S.; DOYLE, E.; O'CONNOR, S. From Marshall's Triad to Porter's Diamond: added value?. **Competitiveness Review**, v. 26, n. 5, p. 500-516, 2016.
- BUCIUNI, G.; PISANO, G. P. **Can Marshall's Clusters Survive Globalization?**. Harvard Business School Working Paper, n. 15-088, 2015.
- CAIRNCROSS, F. **The Death of Distance**. Cambridge, MA: Harvard Business School Press, 1997.

CASSIOLATO, J. E.; LASTRES, H. M. M. O foco em Arranjos Produtivos e Inovativos Locais de micro e pequenas empresas. In: LASTRES, H. M. M.; CASSIOLATO, J. E.; MACIEL, M. L. (Orgs.). **Pequena empresa: cooperação e desenvolvimento local**. Rio de Janeiro: Relume Dumará, 2003. p. 21-34.

COMBES P.; OVERMAN H. G. The spatial distribution of economic activities in the European Union. IN: HENDERSON J. V.; THISSE J. F. (Orgs.). **Handbook of Regional and Urban Economics**. Amsterdam: Elsevier, 2004. p. 2845–2909.

CRUZ, S. C. S.; TEIXEIRA, A. A. C. The evolution of the cluster literature: shedding light on the regional studies–regional science debate. **Regional Studies**, v. 44, n. 9, p. 1263-1288, 2010.

CZAMANSKI, S. **Study of clustering of industries**. Halifax, NS: Institute of Public Affairs, Dalhousie University, 1974.

CZAMANSKI, S.; ABLAS, L. A. Q. Identification of industrial clusters and complexes: a comparison of methods and findings. **Urban Studies**, v. 16, n. 1, p. 61-80, 1979.

DELGADO, M.; PORTER, M. E.; STERN, S. Clusters, convergence, and economic performance. **Research Policy**, v. 43, n. 10, p. 1785-1799, 2014.

DURANTON, G.; OVERMAN, H. G. Testing for localization using micro-geographic data. **The Review of Economic Studies**, v. 72, n. 4, p. 1077-1106, 2005.

ELLISON, G.; GLAESER, E. L. Geographic concentration in US manufacturing industries: a dartboard approach. **Journal of Political Economy**, v. 105, n. 5, p. 889-927, 1997.

FRATESI, U. Issues in the measurement of localization. **Environment and Planning A**, v. 40, n. 3, p. 733-758, 2008.

HEAD, K.; MAYER, T. The empirics of agglomeration and trade. In: HENDERSON, J. V.; THISSE, J. F. (Orgs.). **Handbook of Regional and Urban Economics**. Amsterdam: Elsevier, 2004. p. 2609–2669.

HERVAS-OLIVER, J. et al. Clusters and industrial districts: Where is the literature going? Identifying emerging sub-fields of research. **European Planning Studies**, v. 23, n. 9, p. 1827-1872, 2015.

HILL, E. W.; BRENNAN, J. F. A methodology for identifying the drivers of industrial clusters: the foundation of regional competitive advantage. **Economic Development Quarterly**, v. 14, n. 1, p. 65-96, 2000.

ISARD, W.; SCHOOLER, E. W.; VIETORISZ, T. **Industrial Complex analysis and Regional Development: A Case Study of Refinery-Petrochemical-Synthetic Fiber Complexes and Puerto Rico**. Cambridge, MA: MIT Press, 1959.

KRUGMAN, P. **Geography and Trade**. Cambridge, MA: MIT Press, 1991.

LASTRES, H. M. M. et al. Políticas para APLs: a experiência do BNDES. **Boletim Regional, Urbano e Ambiental - IPEA**, v.10, n. 2, p. 19-32, 2014.

LAZZERETTI, L.; SEDITA, S. R.; CALOFFI, A. Founders and disseminators of cluster research. **Journal of Economic Geography**, v. 14, n. 1, p. 21-43, 2013.

LONG, C.; ZHANG, X. Patterns of China's industrialization: Concentration, specialization, and clustering. **China Economic Review**, v. 23, n. 3, p. 593-612, 2012.

LUNDEVALL, B. A. **National Systems of Innovation**: Towards a theory of innovation and interactive learning. London: Pinter, 1992.

MALMBERG, A.; MASKELL, P. The elusive concept of localization economies: towards a knowledge-based theory of spatial clustering. **Environment and Planning A**, v. 34, n. 3, p. 429-449, 2002.

MARCON, E.; PUECH, F. Evaluating the geographic concentration of industries using distance-based methods. **Journal of Economic Geography**, v. 3, n. 4, p. 409-428, 2003.

MARCON, E.; PUECH, F. A typology of distance-based measures of spatial concentration. **Regional Science and Urban Economics**, v. 62, p. 56-67, 2017.

MARSHALL, A. **Principles of Economics**. London: Macmillan, 1890.

MARTIN, R.; SUNLEY, P. Deconstructing clusters: chaotic concept or policy panacea?. **Journal of Economic Geography**, v. 3, n. 1, p. 5-35, 2003.

MARTÍNEZ-FERNÁNDEZ, M. T.; CAPÓ-VICEDO, J.; VALLET-BELLMUNT, T. The present state of research into industrial clusters and districts. Content analysis of material published in 1997–2006. **European Planning Studies**, v. 20, n. 2, p. 281-304, 2012.

MASKELL, P.; KEBIR, L. What qualifies as a Cluster Theory?. In: ASHEIM, B.; COOKE, P.; MARTIN, R. (Orgs.). **Clusters and Regional Development**: critical reflections and explorations. London and New York: Routledge, 2006. p. 30-49.

MATOS, M. P. et al. Evolução de Arranjos Produtivos Locais em uma década. In: MATOS, M. P.; BORIN, E.; CASSIOLATO, J. E. (Orgs.). **Uma Década de Evolução dos Arranjos Produtivos Locais**. Rio de Janeiro: E-Papers, 2015, p. 23-64.

MCCANN, P. Agglomeration economics. In: KARLSSON, C. (Org.). **Handbook of Research on Cluster Theory**. Cheltenham, UK and Northampton, USA: Edward Elgar, 2008. p. 23-38.

MCCANN, P. Agglomeration and Clustering. In: MCCANN, P. (Org.). **Modern Urban and Regional Economics**. Oxford, UK: Oxford University Press, 2013. p. 49-72.

PALACIOS, J. Economic agglomeration and industrial clustering in developing countries: the case of the Mexican Silicon Valley. In: KUCHIKI, S. J.; JUAN, S.; PALACIOS, J. (Eds.). **Joint Research Program Series**, v. 137, pp. 161-271, 2005.

PIORE, M. J.; SABEL, C. F. **The Second Industrial Divide: Possibilities for Prosperity**. New York: Basic Books, 1984.

PORTER, M. E. **The Competitive Advantage of Nations**. New York: The Free Press, 1990.

PORTER, M. E. Clusters and the new economics of competition. **Harvard Business Review**, v. 76, n. 6, p. 77–90, 1998.

PORTER, M. E. The economic performance of regions. **Regional Studies**, v. 37, n. 6-7, p. 549-578, 2003.

PORTER, M. E. Location, competition, and economic development: Local clusters in a global economy. **Economic Development Quarterly**, v. 14, n. 1, p. 15-34, 2000.

RIVERA, L.; SHEFFI, Y.; WELSCH, R. Logistics agglomeration in the US. **Transportation Research Part A: Policy and Practice**, v. 59, p. 222-238, 2014.

SCOTT, A. J. Flexible production systems and regional development: the rise of new industrial spaces in North America and Western Europe. **International Journal of Urban and Regional Research**, v. 12, n. 2, p. 171-186, 1988.

STORPER, M.; HARRISON, B. Flexibility, hierarchy and regional development: the changing structure of industrial production systems and their forms of governance in the 1990s. **Research Policy**, v. 20, n. 5, p. 407-422, 1991.

STORPER, M.; SCOTT, A. The geographical foundations and social regulation of flexible production complexes. In: WOLSH, J.; DEAR, M. (Orgs.). **The Power of Geography: How Territory Shapes Social Life**. Boston, MA: Allen & Unwin, 1989. p. 21–40.

SUZIGAN, W. **Identificação, mapeamento e caracterização estrutural de arranjos produtivos locais no Brasil**. Brasília: IPEA, 2006.

SWANN, G. M. P. **The implications of clusters: some reflections**. Manchester: Manchester Business School, 2002.

TATSCH, A. L. et al. Análise de políticas para aglomerações no Brasil e em países europeus selecionados. **Planejamento e Políticas Públicas**, n. 44, 2015.

TRIPPL, M.; BERGMAN, E. M. Clusters, Local Districts, and Innovative Milieux. In: FISCHER, M.; NIJKAMP, P. (Orgs.). **Handbook of Regional Science**, Berlin Heidelberg: Springer, 2013. p. 439-456.

VOM HOFE, R.; CHEN, K. Whither or not industrial cluster: conclusions or confusions?. **Industrial Geographer**, v. 4, n. 1, 2006.

WHETTEN, D. A. What constitutes a theoretical contribution?. **Academy of Management Review**, v. 14, n. 4, p. 490-495, 1989.

**2 MÉTODOS QUANTITATIVOS APLICADOS À MENSURAÇÃO DE *CLUSTERS*
INDUSTRIAIS**

Normas do periódico *Journal of Economic Surveys* (versão preliminar)

Resumo

A literatura sobre metodologias para análise de interdependência setorial e concentração geográfica de firmas começou a se multiplicar a partir dos anos 90. O objetivo deste capítulo foi sistematizar a literatura sobre métodos e medidas aplicadas à análise de *clusters* industriais, assim como identificar as tendências nesse campo de conhecimento. O método utilizado foi a bibliometria que consistiu em uma análise de frequência das publicações e da rede de relacionamento entre elas. Foi verificado um aumento exponencial no número de artigos publicados, baseados principalmente na teoria da Administração Estratégica e na Nova Geografia Econômica. Ao longo dos anos a literatura passou a se concentrar na questão da concentração geográfica em relação às análises de encadeamentos setoriais e a fronteira do conhecimento passou de métodos tradicionais de análise regional para áreas como estatística espacial, econofísica e inteligência artificial. Existem questões relevantes em evidência como o Problema da Unidade de Área Modificável (viés de agregação), no entanto, questões como a anisotropia (viés de direção) ainda não foram exploradas.

Palavras-chave

Teoria das aglomerações; métodos insumo-produto; estatística espacial; bibliometria.

2.1 Introdução

A partir dos anos 90, o nome *cluster* se tornou o mais difundido para descrever o fenômeno originado da aglomeração de empresas, seja no sentido de especialização setorial ou concentração regional. Ao se apresentar como uma eficiente estratégia de desenvolvimento econômico e regional, as políticas de *clusters* industriais atraíram o interesse por parte dos acadêmicos e formuladores de políticas, principalmente após os trabalhos desenvolvidos por Porter (1990, 1998), Saxenian (1990, 1994), Krugman (1991a), Storper (1995, 1997), entre outros.

Nathan e Overman (2013, p. 389) tratam a Teoria de *Clusters* como as pesquisas que “specifically seeks to explain the formation and evolution of geographically concentrated sets of linked firms in the same or closely related sectors”. Para implementar uma pesquisa e uma política de *clusters*, Delgado et al. (2016) estabeleceram uma definição operacional desse fenômeno, subdividida em *clusters* comparáveis e *clusters* regionais. A primeira definição é referente aos *clusters* que tem base em ligações interindustriais inferidas a partir de análises multi-regionais, enquanto a segunda contempla os *clusters* baseados em ligações observadas entre indústrias ou empresas em uma mesma região.

O conceito de *clusters* se originou da apropriação de diferentes campos como a economia, administração, sociologia, geografia, planejamento urbano e regional e gestão da inovação e conforme exposto por Lazzarotti et al. (2013) essa abrangência gerou ganhos no desenvolvimento do embasamento teórico e nas aplicações em políticas públicas. Na visão de Ellison e Glaeser (1997, p. 890) “researchers who are primarily interested in international trade, growth, industrial organization, and business strategy have joined geographers and urban economists in investigating why agglomerations exist”.

O crescimento da produção científica em conjunto com a pluralidade entre as áreas de conhecimento geraram uma fragmentação na literatura de *clusters* de atividades econômicas. A sistematização da literatura, por meio de uma pesquisa bibliométrica, realizada por Cruz e Teixeira (2010) propôs a categorização dos principais temas pesquisados em: abordagem evolucionária; economias de aglomerações; teorias do conhecimento e aprendizado; sistemas de inovação; política industrial e desenvolvimento regional; internacionalização e redes globais; abordagem institucional; e métodos e mensuração.

Segundo Cruz e Teixeira (2010, p. 1272) o campo de métodos e mensuração engloba “all the statistical methods and technical tools that have been developed to provide more

objective ways to identify, classify, and explain clustering processes”. Assim, o objetivo principal deste estudo foi sistematizar, por meio de uma análise bibliométrica, a evolução da literatura sobre métodos e medidas aplicadas na mensuração de *clusters* industriais. De particular interesse foram as tendências de pesquisa, a identificação e análise dos artigos mais citados, autores mais produtivos, principais periódicos, a relação entre autores/citações, campos de conhecimento das publicações e, por fim, as palavras-chave que caracterizam o campo de conhecimento.

A análise bibliométrica pode contribuir para a tomada de decisão em políticas industriais de *clusters*, facilitar a ligação entre os métodos e teorias, assim como indicar o desenvolvimento de novos métodos de mensuração de *clusters* de empresas. Segundo Do Prado et al. (2014) essa análise pode ser entendida como a primeira condição para a realização de pesquisas inovadoras uma vez que possibilita a identificação de lacunas e oportunidades dentro do campo, tanto de novas construções teóricas como de pesquisas empíricas. A leitura das limitações e das sugestões de pesquisas futuras dos artigos mais relevantes também contribui para indicar novos caminhos a serem pesquisados.

Além desta seção introdutória, o artigo foi estruturado por: (1) *Insights* gerais acerca de revisões sistemáticas prévias sobre de *clusters* industriais; (2) critérios metodológicos usados na presente pesquisa bibliométrica; (3) resultados e discussões sobre os padrões deste campo de pesquisa e (4) principais conclusões.

2.2 Revisões de literatura

A maioria das revisões sistemáticas existentes se concentrou em diferentes conceitos e teorias sobre *clusters* industriais (Hervas-Oliver et al., 2015; Lazzeretti et al., 2014; Martínez-Fernández et al., 2012; Howells e Bessant, 2012; Cruz e Teixeira, 2010). No entanto, foram apresentadas nesta seção apenas as revisões de literatura que abordaram especificamente os métodos de identificação de *clusters* e aglomerações, ainda que sem o caráter sistêmico proveniente da bibliometria.

A revisão conduzida por Czamanski e Ablas (1979) examinou os métodos de análise empregados com base nas ferramentas matemáticas utilizadas e foco no espaço. Os autores efetuaram a seguinte classificação: Análise multivariada; triangulação e teoria dos grafos e análise de redes; estudos descritivos (que não recorreram a um método rigoroso); e estruturação de complexos em setores específicos. A maioria dos estudos revisados analisou

os fluxos inter-industriais por meio de matrizes insumo-produto. A combinação de abordagens, como análise multivariada e teoria dos grafos, produziu resultados considerados mais completos.

A demanda por uma nova revisão de literatura sobre os métodos de identificação de *clusters* industriais foi identificada por Bergman e Feser (1999), pois “only since the early 1990s have industry *cluster* applications become numerous enough to begin to discern trends in methods and approaches”. Os métodos foram divididos pelos autores com base em análises *micro-level* e *meso-level*, sendo esta segunda subdividida em abordagens *top-down* e *bottom-up*. Estas subdivisões buscaram facilitar a elaboração de estratégias por parte dos analistas de políticas de desenvolvimento regional de acordo com o tempo, recursos financeiros e dados disponíveis.

A análise *micro* possui um perfil predominantemente qualitativo - baseado em entrevistas de campo, método Delphi e grupos focais - enquanto a *meso* é geralmente quantitativa. O nível *meso* de análise de *clusters*, baseado em ramificações setoriais, foi o considerado o mais importante para as pesquisas sobre *clusters* realizadas a partir dos anos 90, segundo a revisão de Stejskal (2011).

Em relação aos métodos quantitativos, Bergman e Feser (1999) enfatizam a utilização do Quociente Locacional (QL) como medida de especialização setorial principalmente para abordagens *bottom-up*. Devido às críticas sobre a fragilidade do QL, os autores recomendaram que o mesmo fosse utilizado em conjunto com outras técnicas e dados como número de estabelecimentos, participação relativa do setor na economia, taxas de crescimento do nível de especialização, entre outras. Outro método aplicado na abordagem *bottom-up* foi a análise de redes baseada em dados insumo-produto ou em dados qualitativos para identificar conexões entre empresas e setores.

No entanto, a predominância da literatura identificada por Bergman e Feser (1999) ainda foi de investigações *top-down* que utilizaram análises multivariadas com base em matrizes I-O. A caracterização deste perfil das pesquisas científicas foi corroborada por Feser e Sweeney (2000):

While input-output based analysis of economic association has been explored extensively, methods for investigating the spatial manifestation of economically linked sectors has received less attention. This is in some sense unsurprising since the study of economic association benefits from a broader literature on input-output that is not direct concerned with location (Feser e Sweeney, 2000, p. 350).

Na década seguinte, um novo cenário foi encontrado na revisão de Nakamura e Paul (2009) que destacaram o crescimento dos estudos interessados em mensurar aglomerações e coaglomerações produtivas por meio de medidas de concentração espacial de atividades econômicas, gerado pelo primeiro índice robusto desenvolvido por Ellison e Glaeser (1997), posteriormente denominado índice EG ou abordagem *dartboard*. Essa mudança pode ser explicada pelo resgate do conceito de aglomeração espacial e o interesse em sua mensuração, em detrimento do conceito de *cluster* isardiano fundamentado em relacionamentos entre indústrias baseados em dados insumo-produto.

As medidas de mensuração de aglomerações industriais analisadas por Nakamura e Paul (2009) foram: Localização industrial em termos de empregos; Especialização regional em termos de empregos; índices de localização em termos de tamanho da planta; índices de localização em termos de nível de produto; Índice de Ellison e Gleaser; Co-aglomeração industrial; escala espacial das aglomerações; e diversidade regional de aglomerações.

Kominers (2008) também indicou um crescimento na mensuração de aglomerações produtivas por meio de índices e enfatizou os avanços obtidos por meio de estatística espacial. A classificação utilizada pelo autor foi a de índices discretos e índices contínuos, a primeira considerou a presença de bordas e teve como expoente o trabalho de Ellison e Glaeser (1997), enquanto no segundo caso houve um refinamento metodológico em que espaço é tratado como sem fronteiras baseado em distâncias entre processos pontuais e teve em a contribuição mais relevante a pesquisa de Duranton e Overman (2005), posteriormente denominado índice DO ou função Kd.

Os estudos interessados na análise da distribuição espacial de atividades econômicas foram divididos em três gerações por Combes e Overman (2004) e De Dominicis et al. (2012). A primeira é fundamentada em índices baseados em proporções como Gini, Herfindhal, Theil, QL, entre outros. A geração seguinte foi baseada na abordagem *dartboard*, como o índice EG. A terceira geração foi composta por estudos baseados na distância entre firmas, como o índice DO.

Embora tenha sido verificada uma proeminência recente dos métodos baseados em índices de concentração e estatística espacial, Vom Hofe e Chen (2006) destacaram que nos anos 2000 as análises de *clusters* baseadas em tabelas insumo-produto ainda eram populares. Deve-se considerar que a presença de aglomerações e coaglomerações implicam em ligações com outras indústrias, então Nakamura e Paul (2009, p.323) concluíram que “to capture these patterns and their implications for outsourcing more directly it will be desirable to estimate

linkage effects between manufacturing and service industries using regional input–output tables”.

Os métodos quantitativos para a identificação de *clusters* sistematizados por Stejskal e Hajek (2012) e Stejskal (2011) corroboraram a literatura supracitada e foram divididos em: quociente de especialização e localização; Análise I-O; análise shift-share; Coeficiente locacional de Gini; e os índices de Ellison-Gleaser e Maurel-Sédillot (MS); Índice relativo de diversidade; Índice de concentração geográfica e distribuição espacial. A revisão de Webster (2013) simplificou a síntese dos métodos em duas categorias, identificação das indústrias líderes por meio de matrizes insumo-produto e medidas para identificar *clusters* industriais baseadas em estatística multivariada de agrupamentos.

Com o aumento das pesquisas baseadas no índice DO (Duranton e Overman, 2005) diversas medidas estatísticas baseadas em distância foram desenvolvidas e aplicadas na análise da localização de firmas. Marcon e Puech (2017) contabilizam 11 funções diferentes e propuseram uma tipologia para os métodos baseados em distância que foram categorizados de acordo com a distância da contagem de vizinhos (função densidade ou cumulativa) e o número médio de vizinhos (medida topográfica, relativa ou absoluta).

2.3 Procedimentos metodológicos

A bibliometria pode ser definida como a análise de um conjunto de documentos publicados que tem interesse em informações quantitativas e na rede de relacionamentos entre essas publicações. Os principais indicadores utilizados para mensurar os fluxos de informações são a frequência - número de ocorrências das citações ou publicações - e o índice de centralidade – uma propriedade gráfica-teórica que quantifica a importância (número de ligações) de um elemento em uma rede (Chen, 2006; Wei et al., 2015).

A análise bibliométrica de um campo de conhecimento é útil para a compreensão da sua dinâmica e visualização de tendências na produção científica de forma robusta. Esse tipo de método organiza a literatura existente, demonstra a trajetória das publicações, os campos de pesquisa tradicionais e emergentes e o seu desenvolvimento no tempo.

O conjunto final de documentos analisado pode ser denominado *research front* enquanto o conjunto de registros bibliográficos que o suportam é chamado de *intellectual base*. O *research front* indica as tendências emergentes da temática e os novos tópicos enquanto a *intellectual base* é representada pela rede de cocitações (Chen, 2006). Segundo

Wei et al. (2015), a análise de cocitações ainda é pouco explorada em outras áreas além da Ciência da Informação e permite o estudo da estrutura científica por meio de similaridades semânticas que extraem relacionamentos entre documentos.

O *framework* utilizado pela presente pesquisa foi adaptado do trabalho de Do Prado et al. (2014) que estabeleceu as etapas de busca, seleção, organização e análise dos dados. Estes procedimentos tiveram por objetivo garantir o caráter científico e a transparência da presente pesquisa de forma a permitir que qualquer pesquisador tenha acesso ao mesmo conjunto de dados (Quadro 1).

Quadro 1. Framework da análise bibliométrica.

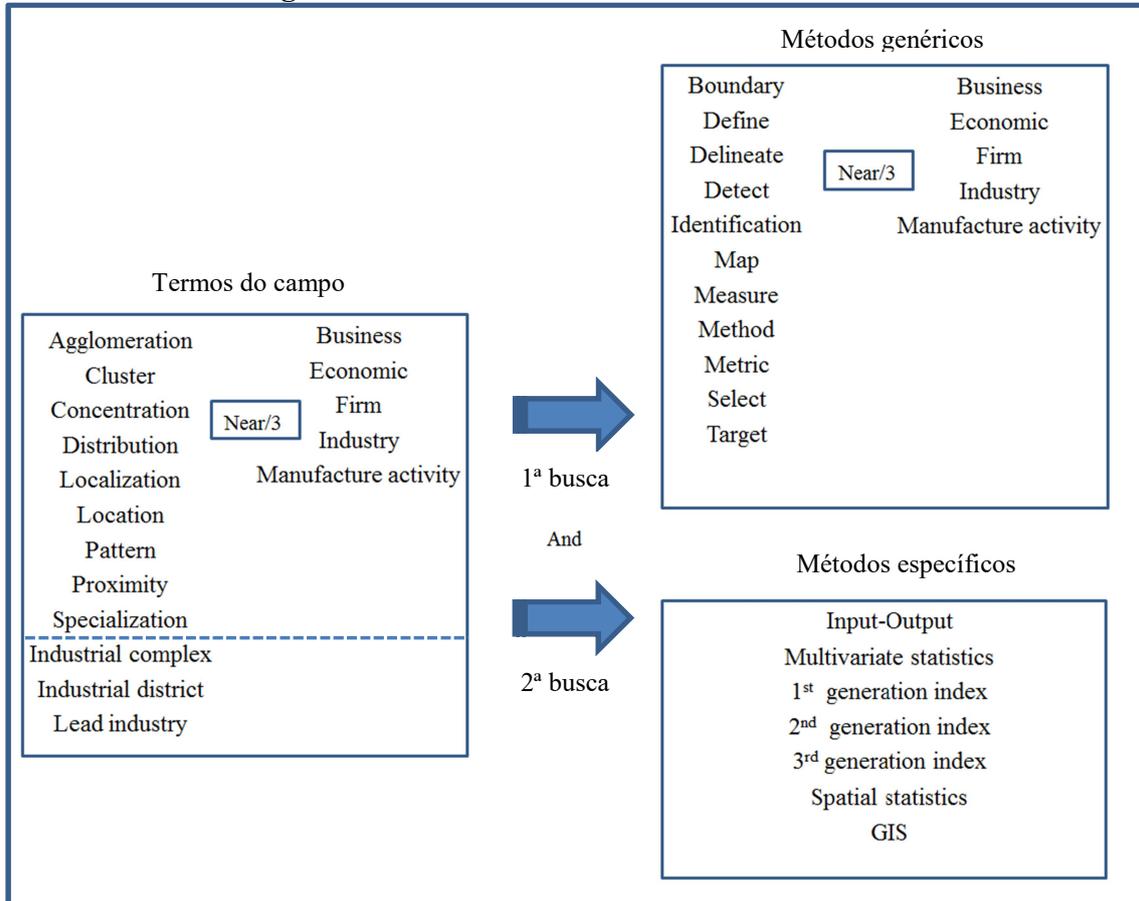
Etapa	Procedimento	Descrição
1	Operacionalização da pesquisa	1.1 Escolha da(s) base(s) científica(s) ou periódicos
		1.2 Delimitação dos termos que representam o campo
		1.3 Delimitação de outros termos para apurar os resultados
2	Procedimentos de busca (filtros)	2.1 Tópico (título, resumo e palavras-chave)
		2.2 Uso de caracteres (“*”, “\$”, “_”) e conectores (“and”, “near”, “or”)
		2.3 Filtro 1: somente artigos
		2.4 Filtro 2: a partir de 1990
		2.5 Filtro 3: Todas as áreas
		2.6 Filtro 4: Todos os idiomas
3	Procedimentos de seleção (Banco de dados)	3.1 Download das referências - <i>software Mendeley</i>
		3.2 Download das referências em formato planilha eletrônica
		3.3 Download das referências para utilização no <i>CiteSpace</i>
		3.4 Organização das referências no <i>Mendeley</i>
		3.5 Organização de matriz de análise em planilha eletrônica
		3.6 Importação dos dados para <i>softwares</i> de análise
4	Adequação e organização dos dados	4.1 Eliminação dos artigos duplicados no banco de dados
		4.2 Eliminação por meio da análise da polissemia dos termos
		4.3 Eliminação de artigos por meio de leitura flutuante
		4.4 Busca dos artigos completos em .pdf
5	Análise da produção Científica	5.1 Análise do volume das publicações e tendência temporal
		5.2 Análise dos artigos mais citados
		5.3 Análise dos autores que mais publicaram
		5.4 Análise da rede de cocitação de autores
		5.5 Análise dos periódicos com maior frequência de publicação
		5.6 Análise das categorias (áreas) das publicações
		5.7 Análise das palavras-chave
		5.8 Descrição, estudo das relações e tendências

Fonte: Adaptado de Do Prado et al. (2014).

A principal coleção da *Web of Science* (WoS) da *Thomson Reuters* foi a base científica escolhida, assim como em Hervas-Oliver et al. (2015), Wei et al. (2015) e Lazzeretti et al. (2014). A WoS é uma das fontes bibliográficas mais extensas e confiáveis do mundo devido a sua padronização nas publicações acadêmicas e pelos critérios de avaliação rigorosos necessários para indexação. A WoS possui aproximadamente 12.000 *journals* indexados e 150.000 anais de congressos, em mais de 250 disciplinas (Wei, 2015).

Os termos de busca foram escolhidos com base nas recomendações dos trabalhos prévios apresentados na revisão de literatura (Figura 1).

Figura 1. Procedimento de busca utilizado na base WoS



Fonte: Elaborado pelo autor.

Em todas as combinações possíveis as buscas se basearam em um conceito do campo de *clusters* industriais em conjunto com um método (genérico ou específico). Os termos conceituais como *cluster*, *agglomeration*, *concentration* e similares foram atrelados a outros de ordem econômica (*industry*, *firm* e similares) a uma distância de até três palavras para formar os termos do campo, assim como os métodos genéricos (*map*, *identification*, *measure* e similares). Esta estratégia precisou ser utilizada, pois a abrangência das palavras utilizadas como conceitos e métodos genéricos elevaram exponencialmente os resultados da busca tornando-a inviável. As exceções foram os termos *industrial complex*, *industrial district*, *lead industry* e os métodos específicos que não demandaram o conector de proximidade.

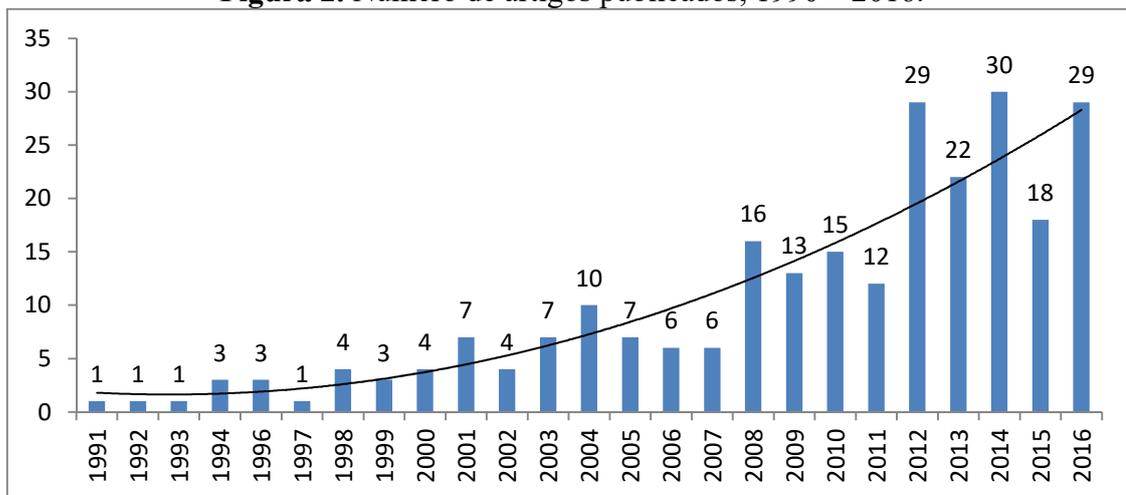
Ao final do procedimento de busca e adequação foram selecionados 252 artigos que compuseram o *research front*. Esta amostra, publicada em um total de 106 periódicos internacionais, utilizou 30.953 referências bibliográficas válidas que formaram a *intelectual*

base, além de 396 palavras-chave distintas utilizadas 1.520 vezes. Para organização de referências foi utilizado o software *Mendeley* e para a elaboração das redes o software *CiteSpace* (Chen, 2006).

2.4 Resultados e discussão

A tendência de crescimento no número de publicações sobre a mensuração de *clusters* industriais entre 1990 e 2016 cresceu a uma taxa de aproximadamente 14% ao ano (Figura 2). A partir de 2008 a série subiu de patamar com uma maior frequência de artigos publicados, possivelmente beneficiados pelo aumento recente na disponibilidade de dados e maior poder de processamento computacional. A maior quantidade de publicações se deu em 2014 (30 artigos), seguida dos anos de 2012 e 2016 (29 artigos) e em média foram 10 artigos por ano. Exceto em 1990 e 1995, todos os anos apresentaram pelo menos uma contribuição científica para a formação deste campo de pesquisa.

Figura 2. Número de artigos publicados, 1990 – 2016.



Fonte: Elaborado pelo autor.

A tendência do número de publicações sobre mensuração de *clusters* pode ser explicada pela crescente importância que a especialização regional e os *clusters* industriais alcançaram em termos acadêmicos e de políticas públicas, assim como explicado por Cruz e Teixeira (2010) ao fazerem referência à literatura do campo como um todo. Hervas-Oliver et al. (2015, p. 1845) constataram, também em uma análise bibliométrica, que o crescimento do campo foi “impressively prominent since the beginning of the 1990s. Only in 2011, more than 400 documents addressing industrial districts and *clusters* are observed”.

2.4.1 Resultados descritivos

O Quadro 2 apresentou as pesquisas mais relevantes para o campo de conhecimento em estudo. A classificação foi baseada nos artigos mais citados pelos demais artigos indexados na base WoS, seja dentro e fora da amostra analisada. Além do número de citações, foi também indicada a média de citações por ano como informação complementar.

Em geral, os artigos foram publicados entre o final dos anos 90 e o início dos anos 2000 e representaram a proposição ou aprimoramento de alguma das categorias metodológicas definidas na revisão de literatura. As classificações seguidas para a análise de resultados foram as propostas por De Dominicis et al. (2013) e Combes e Overman (2004) para os índices de 1ª, 2ª, 3ª geração e por Czamanski e Ablas (1979) para o método tradicional baseado dados insumo-produto (I-O).

Quadro 2. Artigos mais citados.

Artigo	Autor	Periódico	Citações	Citações (média)
Geographic concentration in US manufacturing industries: A dartboard approach	Ellison e Glaeser (1997)	Journal of Political Economy	624	31,2
The economic performance of regions	Porter (2003)	Regional Studies	372	26,6
The determinants of agglomeration	Rosenthal e Strange (2001)	Journal of Urban Economics	279	17,44
Testing for localization using micro-geographic data	Duranton e Overman (2005)	Review of Economic Studies	220	18,33
What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns	Ellison, Glaeser e Kerr (2010)	American Economic Review	167	24
National industry <i>cluster</i> templates: A framework for applied regional <i>cluster</i> analysis	Feser e Bergman (2000)	Regional Studies	134	7,94
The geography of firm births in Germany	Audrescht (1994)	Regional Studies	125	5,43
A measure of the geographic concentration in French manufacturing industries	Maurel e Sedillot (1999)	Regional Science and Urban Economics	107	5,94
Localization of industry and vertical disintegration	Holmes (1999)	Review of Economics and Statistics	93	5,17
A methodology for identifying the drivers of industrial <i>clusters</i> : The foundation of regional competitive advantage	Hill e Brennan (2000)	Economic Development Quarterly	78	4,59
Evaluating the geographic concentration of industries using distance-based methods	Marcon e Puech (2003)	Journal of Economic Geography	73	5,21
Geographic concentration and establishment scale	Holmes e Stevens (2002)	Review of Economics and Statistics	72	4,8

Fonte: Elaborado pelo autor.

O artigo de Ellison e Glaeser (1997), aplicado à indústria americana, e a versão modificada de Maurel e Sedillot (1999), aplicada ao setor manufatureiro da França, foram citados 624 e 107, respectivamente e representaram o início da segunda geração de índices para análise da distribuição espacial de atividades econômicas. A contribuição principal dos autores foi o desenvolvimento de um teste para verificar se os níveis de concentração geográfica observados são maiores do que era de se esperar ao acaso. Até então, os índices utilizados eram baseados na proporção de uma indústria em uma região em relação à economia total, assim outra contribuição foi controlar o efeito de poucas firmas que detêm a maior parte de determinados setores.

Deve-se considerar que a primeira geração de índices, baseada no Gini Locacional, foi proveniente do trabalho pioneiro de Krugman (1991a) e complementada por Audretsch e Feldman (1996). Porém, a primeira referência bibliográfica é um livro não indexado e a segunda um artigo que não apresenta os campos “resumo” e “palavras-chave” em sua formatação, logo não foram captadas pelo sistema de busca apresentado nos procedimentos metodológicos. Por sua vez, Audretsch e Fritsch (1994), com 125 citações, analisaram a economia da Alemanha com interesse na relação entre a os retornos crescentes à escala e a tendência de aumento na concentração de atividades econômicas, medidas pela taxa de nascimento de firmas.

Após críticas a subjetividade e falta de aplicabilidade prática de sua teoria inicial, Porter (2003), com 372 citações, propôs uma metodologia quantitativa que examinou o padrão da especialização regional de empregos em diferentes setores industriais dos Estados Unidos. As indústrias consideradas geograficamente concentradas foram correlacionadas com objetivo de mensurar os padrões de co-localização e revelar potenciais externalidades de demanda, conhecimento, trabalho, entre outras. Essa metodologia é a base do *US Mapping Cluster Project* da *Harvard Business School*, ainda em andamento e atualizada pela equipe do professor Porter nos artigos Delgado et al. (2014, 2016).

Citados por 279 artigos, Rosenthal e Strange (2001) estimaram os determinantes das aglomerações da indústria manufatureira do EUA. A concentração espacial mensurada pelo índice EG foi utilizada como variável dependente em um modelo econométrico que teve como variáveis explicativas *proxies* para as economias de aglomeração (transbordamento de conhecimento, mercado de trabalho, compartilhamento de insumos) além das variáveis de controle como custos de transporte e vantagens naturais. Todas essas fontes se mostraram significativas, em especial o mercado de trabalho.

Para estudar a tendência de aglomeração de indústrias no Reino Unido, Duranton e Overman (2005) - 220 citações - desenvolveram um teste de localização baseado na distância entre firmas que pode ser considerada a principal contribuição, em conjunto com o artigo sobre a indústria francesa de Marcon e Puech (2003) - 73 citações -, para o início da terceira geração de índices. Por meio da estatística espacial de padrões pontuais, os autores utilizaram a função Kd e as funções L e D, respectivamente, para tratar o espaço de forma contínua, ou seja, não viesadas em relação a escala espacial e de agregação (ex. bairros, municípios, microrregiões, estados, etc.). Essa foi uma solução para o Problema da Unidade de Área Modificável (MAUP) que torna os resultados das análises de concentração regional dependentes da escala geográfica escolhida.

Em uma atualização de sua proposta metodológica, Ellison et al. (2010) enfatizaram o papel da aglomeração de indústrias correlatas por meio do índice de coaglomeração propostos por Ellison e Glaeser (1997) para a indústria manufatureira dos EUA. Assim como em Rosenthal e Strange (2001), esses índices foram utilizados como variáveis dependentes em modelos econométricos em função das fontes marshallianas de economias de aglomeração (trabalho, insumos e tecnologia) que indicaram um efeito significativo.

O modelo de Feser e Bergman (2000), citado 134 vezes, propôs a utilização de informações sobre relações inter-industriais para identificar potenciais *clusters* regionais, seguindo a abordagem tradicional dos anos 70 relatada por Czamanski e Ablas (1979). As medidas de ligações diretas e indiretas entre cada setor extraídas das contas desagregadas insumo-produto foram utilizadas como variáveis em uma Análise Fatorial de Componentes Principais. Os setores membros de cada *cluster* foram definidos de acordo com as maiores cargas em cada fator que indicaram a força da ligação entre as indústrias, em uma aplicação na economia da Carolina do Norte/EUA.

O artigo de Holmes (1999) - 93 citações - utilizou o índice EG para identificar as indústrias geograficamente concentradas nos EUA, o objetivo do autor foi confirmar que estabelecimentos localizados em regiões concentradas tendem a ser verticalmente desintegrados. Posteriormente, Holmes e Stevens (2002) - 72 citações - mostraram que plantas industriais são, em média, maiores em regiões em que o setor é concentrado, novamente com o auxílio do índice EG.

Hill e Brennan (78 citações) apresentaram um método baseado em uma combinação entre as análises de *cluster* hierárquico e discriminante para identificação de *clusters* de indústrias na qual uma determinada região apresenta vantagem competitiva quando não se

tem noção dos setores mais importantes. As variáveis referentes à região de Cleveland-Akron/EUA foram a competitividade regional, exportações, centralidade (ligações insumo-produto) e a especialização dos empregos (Quociente Locacional), todas derivadas de duas teorias de desenvolvimento econômico: vantagem competitiva e base de exportação.

Ao analisar as técnicas utilizadas pelos artigos mais citados foi confirmada a tendência de comparação e/ou utilização de métodos complementares. No caso do Índice EG, o efeito da concentração regional da indústria foi mitigado por meio do índice Herfindhal e quando esse tende a zero, o índice EG nada mais é que o próprio índice de Gini. Maurel e Sedillot (1999) compararam os resultados dos índices MS, EG e Gini para o setor industrial francês e americano, visto que, segundo estes autores, Ellison e Glaeser (1997) não realizaram essa verificação. Foi encontrada pouca diferença entre os dois métodos de segunda geração e uma grande diferença entre estes e o Gini Locacional.

Para tornar seu método mais robusto, Porter (2003) utilizou o Quociente Locacional em conjunto com Coeficiente de Gini, ambos com dados referentes ao número de empregos setoriais como filtros para categorizar as indústrias. Para evitar correlações espúrias, este autor utilizou a matriz I-O dos EUA para identificar as interligações relevantes entre indústrias. Por sua vez, Hill e Brennan (2000) foram os únicos entre os artigos mais relevantes a complementar os resultados com análise qualitativa, por meio de grupos focais e entrevistas com especialistas.

Duranton e Overman (2005) compararam os índices de segunda e terceira geração (EG e DO, respectivamente) e verificaram que no primeiro, 94% das indústrias do Reino Unido foram consideradas localizadas enquanto pelo segundo índice, apenas 52%. Na mesma linha, Marcon Puech (2003) fizeram uma comparação da função K de Ripley com a Função Kd de Duranton e Overman (2005) apresentada anteriormente em versão *working paper* em 2002. Os autores concluíram que a Função Kd poderia contribuir para o aperfeiçoamento da primeira em relação ao controle para o tamanho da firma e para a concentração industrial, além de tornar desnecessária a correção do efeito de borda. Estes dois estudos foram atualizados em Marcon e Puech (2010) e Duranton e Overman (2008). Além de trabalhar com o índice EG (discreto), Ellison et al. (2010) propuseram uma aproximação com o índice DO (contínuo) para fins de comparação de resultados.

Para complementar a visualização do *research front* os autores que mais contribuíram para o crescimento do campo foram identificados, seja como primeiro autor ou coautor. O

Quadro 3 apresentou o autor, número de artigos publicados, título do artigo e do periódico mais recente. Apenas oito autores publicaram quatro ou mais obras neste campo de pesquisa.

Quadro 3. Autores com maior número de publicações.

Autor	Nº	Artigo mais recente	Periódico	Ano
ARBIA G.	7	Spatio-temporal <i>clustering</i> in the pharmaceutical and medical device manufacturing industry: A geographical micro-level analysis	Regional Science And Urban Economics	2014
FESER E.	7	The Rural Role in National Value Chains	Regional Studies	2009
ESPA G.	6	Spatio-temporal <i>clustering</i> in the pharmaceutical and medical device manufacturing industry: A geographical micro-level analysis	Regional Science And Urban Economics	2014
FIGUEIREDO O.	5	Accounting for neighboring effects in measures of spatial concentration	Journal Of Regional Science	2011
GIULIANI D.	5	Weighting Ripley's K-Function to Account for the Firm Dimension in the Analysis of Spatial Concentration	International Regional Science Review	2014
GUIMARAES P.	5	Accounting for neighboring effects in measures of spatial concentration	Journal Of Regional Science	2011
WOODWARD D.	5	Accounting for neighboring effects in measures of spatial concentration	Journal Of Regional Science	2011
AKGUNGOR S.	4	Public Policies and Development of the Tourism Industry in the Aegean Region	European Planning Studies	2009

Fonte: Elaborado pelo autor.

Os professores Giuseppe Arbia e Edward Feser foram os autores com maior volume de publicações sobre métodos para identificação de *clusters* industriais. Em geral, a produção do professor Espa et al. (2013) foi em coautoria com Arbia e com Giuliani et al. (2014) que formaram na Itália um grupo de estudos com contribuições relevantes para o tema pela abordagem da estatística espacial de padrões pontuais (Arbia et al., 2014; Arbia et al., 2012; 2010; Arbia et al., 2008; De Dominicis et al., 2013).

Feser e Isserman (2009) utilizaram a análise multivariada em conjunto com dados IO para mostrar que áreas rurais podem contribuir para a competitividade de sistemas nacionais de produção. Embora especialista nos métodos tradicionais para identificação de *clusters* (Feser e Bergman, 2000; Feser, 2003; Renski et al., 2007) este autor também trabalhou com os métodos estatísticos de dados de área e de pontos (Feser et al., 2008; Feser et al., 2005; Sweeney e Feser, 1998).

Os pesquisadores Figueiredo, Guimaraes e Woodward, de Portugal e Estados Unidos, respectivamente, possuem cinco publicações e todas foram em conjunto entre os autores. No artigo mais recente, os autores propuseram um método para considerar a autocorrelação espacial em índices não espaciais para dados de área, como o EG, Gini, Herfindhal, entropia, etc (Guimarães et al. 2011). Em geral, este grupo de autores se debruçou em propor refinamentos metodológicos para os de primeira e segunda geração de índices (Guimarães et

al., 2009; 2007) e na relação do tamanho dos estabelecimentos e desintegração vertical com a teoria de Marshall (Figueiredo et al., 2010;2009).

Akgüngör aplicou o QL para localizar as regiões na Turquia em que o setor de turismo era mais especializado (Gülcan et al. 2009). Em suas publicações anteriores, o autor trabalhou também com o Gini Locacional e com métodos multivariados baseados em dados I-O (Akgüngör r, 2006; Akgüngör et al., 2003; Falcioğlu e Akgüngör, 2008).

A próxima fase dos resultados da bibliometria está relacionada com a quantidade e a qualidade dos periódicos que abarcaram as publicações sobre o tema (Tabela 1). Foram identificados 106 periódicos em que foram veiculados os 252 artigos que compuseram a amostra deste estudo. A concentração das publicações se deu em 14 revistas que publicaram 131 artigos, ou seja, 52% do total de *papers*. O fator de impacto médio foi de 1,41.

Tabela 1. Periódicos com maior número de publicações.

Periódico	Frequência	Proporção	JCR
<i>Regional Studies</i>	32	12,7%	1,99
<i>Regional Science and Urban Economics</i>	13	5,2%	1,02
<i>Annals of Regional Science</i>	11	4,4%	0,57
<i>European Planning Studies</i>	11	4,4%	1,06
<i>Journal of Economic Geography</i>	9	3,6%	3,43
<i>Journal of Regional Science</i>	8	3,2%	1,63
<i>Papers in Regional Science</i>	8	3,2%	1,14
<i>Growth and Change</i>	7	2,8%	0,88
<i>Journal of Urban Economics</i>	7	2,8%	2,12
<i>Chinese Geographical Science</i>	6	2,4%	1,15
<i>Economic Development Quarterly</i>	6	2,4%	1,00
<i>Urban Studies</i>	5	2,0%	1,93
<i>Journal of Geographical Systems</i>	4	1,6%	1,18
<i>Journal of Economic And Social Geography</i>	4	1,6%	0,68
Parcial	131	52%	-
Outros periódicos	121	48%	-
Total	252	100%	1,41

Fonte: Elaborado pelo autor.

Quatro periódicos se destacaram por suas contribuições para o campo por meio de uma frequência de publicações superior a 10 artigos, foram eles *Regional Studies* (32 artigos), *Regional Science and Urban Economics* (13 artigos), *Annals of Regional Science* (11 artigos) e a *European Planning Studies* (11 artigos). O *Regional Studies* também se destacou por publicar três entre os artigos mais relevantes da área, assim como o *Regional Science and Urban Economics*, *Journal of Economic Geography*, *Journal of Urban Economics* e o *Economic Development Quarterly* que publicaram um artigo entre os mais citados (Quadro 01).

O *Journal of Economic Geography* foi o periódico que apresentou maior fator de impacto JCR, na ordem de 3,43, e publicou 3,6% dos artigos da amostra em estudo, seguido do *Journal of Urban Economics* com fator de impacto 2,12 e 2,8% das publicações. A revista *Regional Studies* que mais contribuiu para o desenvolvimento do campo (12,7%) obteve um fator de impacto de 1,98, seguida da *Regional Science and Urban Economics* (5,2%) com fator de impacto de 1,02.

Em relação aos métodos presentes nos artigos publicados pelos periódicos foi identificado que o *Regional Studies*, *Annals of Regional Science*, *Journal of Economic Geography*, *Journal of Regional Science* e *Papers in Regional Science* foram os mais abrangentes e publicaram estudos com todos os tipos de técnicas. Por sua vez, o *Journal of Economic and Social Geography*, *Journal of Urban Economics* e *Regional Science and Urban Economics* publicaram somente trabalhos com base nas três gerações de índice. O *Economic Development Quarterly*, *Urban Studies* e *European Planning Studies* se concentraram em métodos tradicionais e nos índices de primeira geração, enquanto o *Journal of Geographical Systems* publicou métodos tradicionais e índices de terceira geração. Por fim, os *journals Growth and Change* e *Chinese Geographical Science* publicaram todos os métodos, exceto os baseados na abordagem *dartboard*.

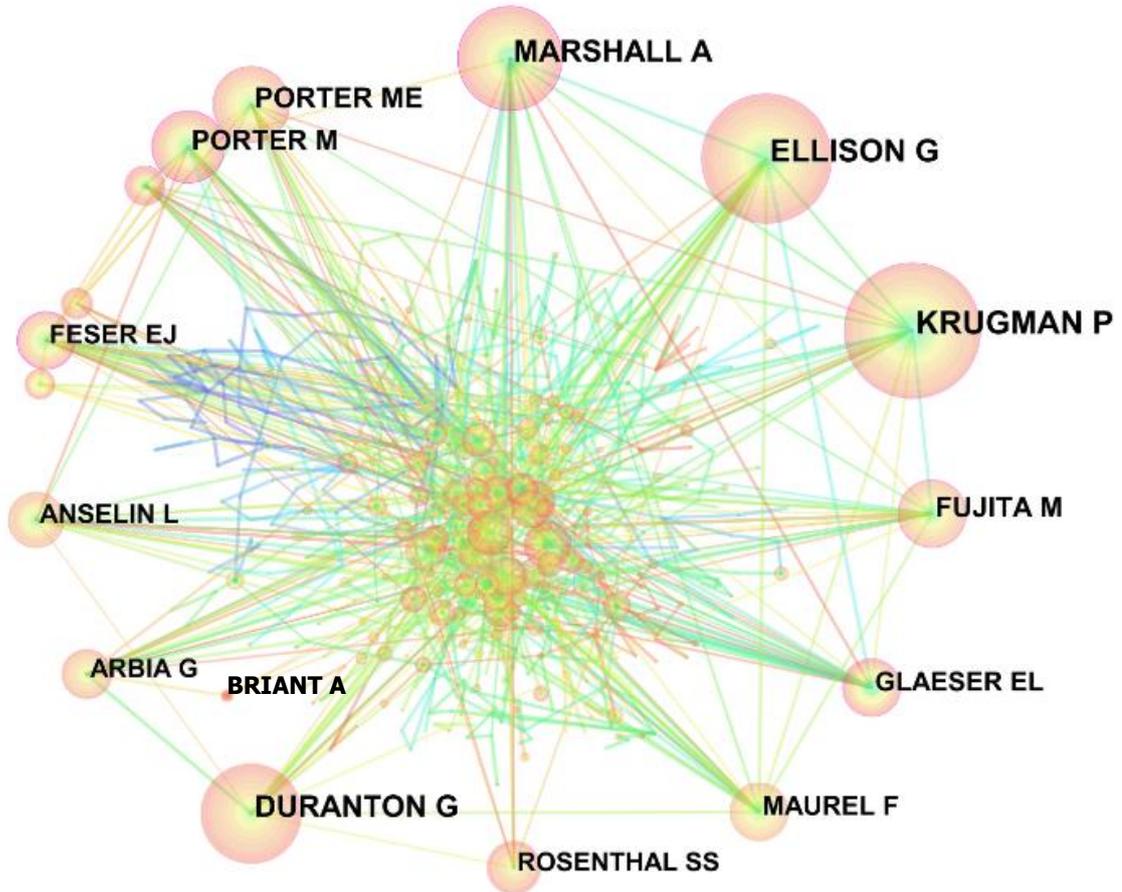
É relevante mencionar que o *Journal of Geographical Systems* publicou 2 artigos relevantes sobre a temática, porém na data de publicação ainda não se encontravam indexados na WoS. No primeiro artigo, Feser e Sweeney (2000) desenvolveram uma metodologia que integrou a análise insumo-produto com estatística espacial de padrões pontuais e pode ser considerada uma notável ligação entre os índices de terceira geração e a metodologia tradicional. O segundo, escrito por Arbia (2001), foi considerado pioneiro em trazer a estatística espacial de dados de área para a análise de concentração regional. O autor demonstrou a importância em se considerar a dependência espacial de atividades econômicas, por meio do Índice de Moran e da estatística de Getis-Ord. Este artigos foram localizados com auxílio da Rede de cocitação de autores (Figura 3).

2.4.2 Resultados das redes

A rede de co-citação de autores elaborada a partir da *intellectual base* relacionou os autores mais citados (acima de 40 citações) pelos artigos da amostra (*research front*) independente da obra referenciada (Figura 3). O interesse desta análise foi revelar também as

principais referências da base teórica que sustentaram o campo de conhecimento em estudo, visto que a base metodológica relevante foi identificada previamente no Quadro 02.

Figura 3. Rede de cocitação de autores.



Fonte: Elaborado pelo autor.

O autor que mais influenciou o campo foi Michael Porter (Porter, M. e Porter, M.E.) com uma frequência de 152 citações, principalmente pelas suas contribuições à teoria da Administração Estratégica pelas obras *The competitive advantage of nations*, de 1990, e *Clusters and the new economics of competitiveness* na *Harvard Business Review* de 1998.

Em segundo lugar, Paul Krugman (110 citações) expôs as aglomerações de atividades econômicas como aspecto central da geografia econômica em seu livro *Geography and Trade*, de 1991, identificada como principal contribuição para o tema em estudo. Além de propor a primeira geração de índices, o Gini Locacional, o livro de Krugman também é um dos principais pilares teóricos da Nova Geografia Econômica (Krugman 1991a; 1991b). Esta abordagem das aglomerações consolidou o papel da dimensão espacial das atividades econômicas. Muitos trabalhos empíricos buscam verificar quanto as aglomerações elevam o

desempenho econômico de regiões em diferentes escalas espaciais, seja país, estado ou municípios (Briant et al., 2010).

Alfred Marshall foi citado 84 vezes pelo livro *Principles of Economics*, de 1890, considerado o primeiro autor a observar as economias da aglomeração com mais ênfase no ponto de vista teórico do que metodológico. Segundo o autor a proximidade entre firmas do mesmo setor e agentes correlatos eleva a produtividade e a capacidade de inovação da indústria na região. Essa contribuição para o campo de conhecimento das aglomerações rendeu a Marshall o segundo maior índice de centralidade (0.20), mesma posição alcançada por Porter. O Autor com maior centralidade na rede foi Feser (0.25).

A contribuição de Masahisa Fujita (1999), com 55 citações, tem sua base em *The Spatial Economy: Cities, Regions, and International Trade*, de 1999, livro também de coautoria de Krugman e base da corrente teórica denominada Nova Geografia Econômica. Por sua vez, Luc Anselin (frequência 46) criou uma extensão do Índice de Moran para dados de área, denominada Índice de Moran Local no artigo *Local Indicators of Spatial Association – LISA*, na *Geographical Analysis* de 1995, que foi amplamente utilizado na segunda geração de índices, embora o próprio autor não tenha apresentado nenhum trabalho diretamente ligado ao mapeamento de aglomerações produtivas. Os demais autores mais relevantes da rede foram Ellison (105 citações), Feser (97), Duranton (82), Glaeser (59), Maurel (48), Rosenthal (44) e Arbia (41) e tiveram suas contribuições apresentadas nos Quadros 01 e 02.

Na rede de co-citação de autores foi detectado um elemento atípico (*citation boom*) representado pelo círculo vermelho, próximo ao nome Arbia, e o responsável por tal evento foi Briant et al. (2010) como consequência de ser citado 11 vezes em apenas 4 anos consecutivos. O problema central investigado foi o MAUP e seu efeito nos índices de Gini e EG para a economia francesa utilizando diversas agregações geográficas. Os autores concluíram que o MAUP causa menos distorções quanto menor a unidade geográfica e que a escolha do índice causa mais distorções nos resultados do que as distintas agregações geográficas de um mesmo índice.

Foi verificada uma tendência de aumento da preocupação do *research front* com problemas e soluções para análises espaciais. No caso do MAUP foram encontrados 43 artigos que fizeram referência a esse fenômeno (Huisman e Wissen, 2004; Bertinelli e Decrop, 2005; Menon, 2012; Nielsen e Hennerdal, 2014; Scholl e Brenner, 2016, entre outros). No entanto, outra característica de dados espaciais que pode distorcer os resultados denominada anisotropia (viés direcional) foi negligenciada. Esse viés ocorre quando a relação

espacial de um fenômeno varia de acordo com a direção da configuração de pontos. Foram encontrados 11 artigos que fizeram menção a esse problema, porém todos assumiram isotropia, ou seja, ausência de anisotropia (Sweeney e Ferser, 1998; Kosfeld et al., 2011; Arbia et al., 2012; Bonneu e Thomas-Agnan, 2015).

A interdisciplinaridade inerente ao campo de *clusters* industriais foi identificada na introdução e abrangeu áreas como a Economia, Administração, Geografia, entre outras (Lazzeretti et al., 2014; Ellison e Glaeser 1997). A linha do tempo das áreas de conhecimento (categorias definidas pela WoS) em que foram publicados dois ou mais dentre os 252 artigos indicou o triênio em que cada área iniciou a discussão sobre métodos e mensuração de *clusters* insutriais (Figura 4). Os círculos à esquerda indicaram publicações próximas do ano de 1990, enquanto os círculos ao lado direito apontaram as áreas dos artigos publicados mais próximos do ano de 2016. O tamanho da fonte das letras é proporcional ao volume de publicações de cada categoria.

Figura 4. Linha do tempo das publicações (categorias WoS).



Fonte: Elaborado pelo autor.

As áreas de *Business & Economics* (150 artigos da amostra publicados até 2016), *Economics* (137), *Environmental Sciences & Ecology* (108) *Geography* (99 artigos) e *Environmental Studies* (98 artigos) foram as primeiras a aparecerem na linha do tempo, pois o periódico *Regional Studies* está indexado nelas e foi o primeiro a contribuir para o tema, em 1991. As áreas *Public Administration* (41) e *Planning & Development* (40) entraram na linha

do tempo em 1992 com uma publicação da revista *Growth and Change* e a área *Urban Studies* (48) entrou em 1994 com um artigo publicado em um *journal* homônimo. Deve-se ressaltar que um periódico geralmente está indexado em mais de uma categoria.

Os artigos que introduziram as categorias supracitadas na linha do tempo são baseados nos métodos de identificação tradicionais. O artigo de Ó'Huallacháin (1991) identificou a associação espacial de empregos setoriais da economia norte americana por meio da análise fatorial, enquanto Anderson e Johnston (1992) utilizaram a abordagem de ligações para frente e para trás para identificar indústrias alvo para iniciativas de desenvolvimento econômico. Braunerhjelm e Carlsson (1999) também propuseram um método baseado em tabelas insumo-produto e no QL de empregos para identificar *clusters* de atividades econômicas. A exceção foi Fingleton (1994) trabalhou com o método *shift-share* para identificar a concentração da indústria de alta tecnologia em nível municipal.

Para acompanhar a tendência da produção acadêmica sobre os métodos e aplicações emergentes foram analisados os artigos publicados no último triênio (a partir de 2014) nas áreas que foram inseridas na linha do tempo recentemente.

Guo et al. (2015) e Oliva et al. (2016) utilizaram as contribuições da Física nos periódicos *Physica A* e *The European Physical Journal*, respectivamente, ao avaliar as relações insumo-produto entre agrupamentos de setores industriais por meio de proposições de gráficos direcionais, redes de fluxo aberto e medidas *clusters* hierárquicos. A área de Ciência e Tecnologia identificou o trabalho de Xing et al. (2016), na revista *PloS One*, que avaliaram os setores industriais que apresentam maior impacto no desenvolvimento regional pela ótica da econofísica e da teoria dos buracos estruturais revisitados, também baseados em dados insumo-produto.

As áreas de matemática, estatística e probabilidade emergiram com os trabalhos de Bersimis et al. (2014) e Bocci e Rocco (2016) que publicaram nos periódicos *Applied Stochastic Models in Business and Industry* e *Journal of Applied Statistics*, respectivamente, análises de estatística espacial por padrões de pontos de dados de área, aliados a Sistemas de Informações Geográficas, para detectar *clusters* de atividades econômicas e a tendência de novas firmas se aglomerarem. Seguindo essa linha de raciocínio. Publicado no *International Journal of Remote Sensing*, o artigo de Rybnikova e Portnov (2014) mapeou a concentração de atividades econômicas utilizando técnicas avançadas de Sensoriamento Remoto como a intensidade das luzes elétricas durante a noite extraídas de imagens de satélite.

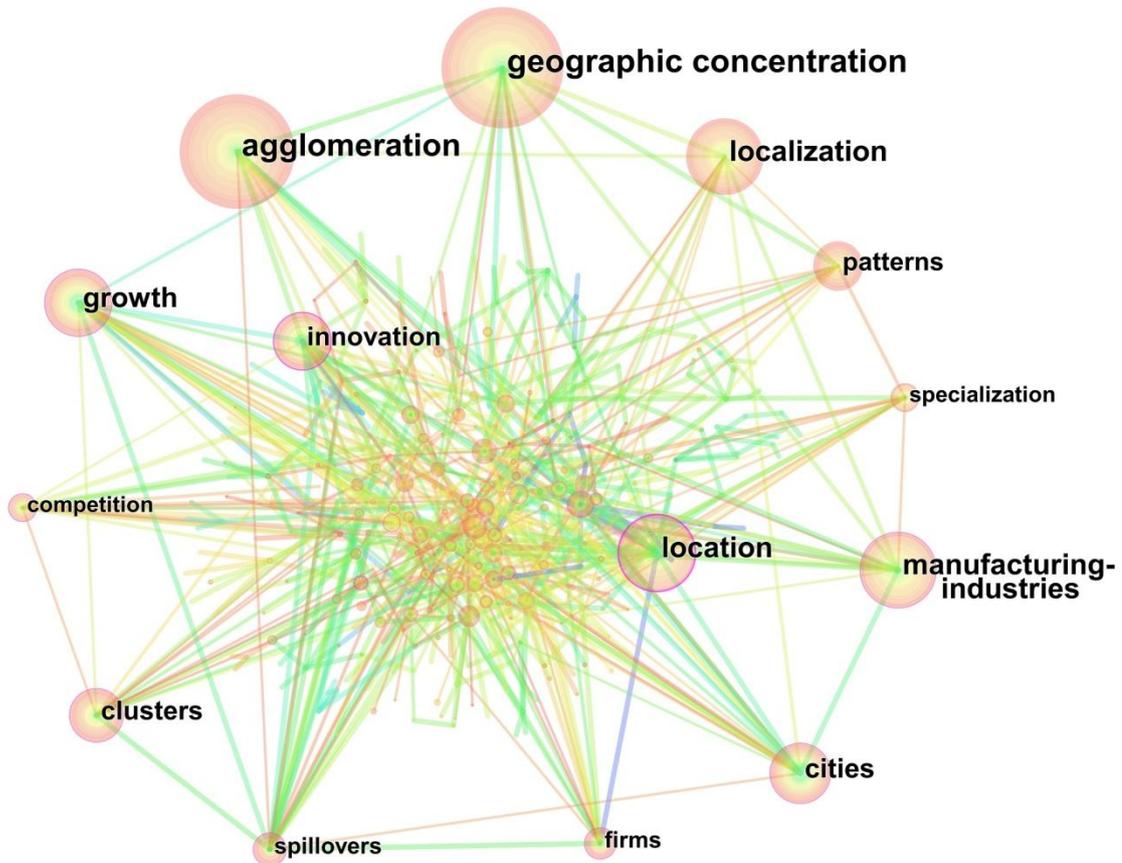
No *Pattern Recognition Letters*, Carlei e Nuccio (2014) ampliaram o campo da Computação ao utilizar a inteligência artificial (redes neurais e mapas auto organizáveis) para propor um índice para mapear aglomerações aplicado aos distritos industriais de vestuário na Itália. Rivera et al. (2014) buscaram identificar *clusters* no setor de logística no EUA por meio do QL modificado em conjunto com a participação de estabelecimentos logísticos em um município proporcionalmente ao total no país e com isso o periódico *Transportation Research Part A* colocou a área de Transportes na linha do tempo.

Com base na linha do tempo das categorias foi possível verificar que a classificação dos artigos em métodos tradicionais (I-O e estatística multivariada) e índices (1ª, 2ª e 3ª geração) adotadas no início de trabalho não foi suficiente para acompanhar a evolução deste campo de conhecimento. Uma nova proposição pode ser subdividida em:

- (1) métodos baseados em matrizes I-O e redução de dados (estatística multivariada, análise de redes ou inteligência artificial);
- (2) Índices de 1ª geração (Gini, QL, HH e suas evoluções);
- (3) índices de 2ª geração (abordagem *dartboard* e suas evoluções);
- (4) métodos espaciais (dependência espacial e Sistemas de Informações Geográficas).

A Figura 5 apresentou a rede das 396 palavras-chave utilizadas nos artigos da amostra (*co-occurring author keywords and WoS KeyWords Plus*) com ênfase nas que foram utilizadas em mais de 15 artigos. A ocorrência destes termos está correlacionada com a delimitação da frente de pesquisa apresentada na metodologia, assim a rede foi útil para quantificar as terminologias mais frequentes e suas relações com as demais palavras-chave.

Figura 5. Rede de co-ocorrência de palavras-chave.



Fonte: Elaborado pelo autor.

Um conjunto de palavras-chave relevante para o entendimento da temática indicado na rede foi formado por *geographic concentration* (71 ocorrências), *agglomeration* (67), *localization* (45)/ *location* (44), *patterns* (29), *clusters* (31) e *specialization* (17) que indicou uma tendência de estudos interessados em mensurar a distribuição espacial de atividades econômicas. A conotação das palavras-chave foi brevemente explanada de acordo com o artigo mais citado a utilizá-la. Porter (2003) utilizou o termo *geographic concentration* em conjunto com *of linked industries* pra definir o termo *clusters*. *Agglomeration* foi utilizado como sinônimo de *concentração* e antônimo de *dispersão*, assim como *localization* (Marcon e Peuch, 2003; Ellison et al., 2010, respectivamente). *Location* em conjunto com *choice* foi outro nome utilizado para o índice de Ellison e Glaeser (1997) e em conjunto com *pattern* para o índice D (Mori, 2005). Por fim, *specialization* foi considerado um efeito local decorrente das economias de aglomeração (Viladecans-Marsal, 2004).

A palavra-chave *location* apresentou o maior número de ligações com outras palavras-chave (índice de centralidade 0,41), possivelmente devido à abrangência alcançada pelo modelo de escolha locacional de Ellison e Glaeser (1997). Enquanto o termo *innovation*, com

a segunda maior centralidade (0,32), indicou que os artigos analisados tendem a associar as palavras-chave (concentração, aglomeração, *clusters*, entre outras) com o processo inovativo, assim como outros termos de elevada centralidade como *growth* (0,20), *cities* (0,17) e *spillovers* (0,15).

A análise das palavras-chave referentes aos métodos indicou uma tendência para a utilização de métodos de estatística espacial (frequência de 73), sobre métodos insumo-produto (37). Dentre as palavras-chave originadas da estatística espacial a ênfase foi na análise de padrões pontuais, sendo *distance-based methods* a mais utilizada (8). Deve-se considerar que a frequência de menções a estas palavras-chave foram agrupadas devido à existência de múltiplos sinônimos (ex.: *point process*, *point patterns*, *k-function*).

2.5 Conclusões

O objetivo deste estudo foi explorar a literatura sobre métodos e medidas aplicadas na mensuração de *clusters* industriais, por meio de uma análise bibliométrica.

Foi identificada uma tendência de crescimento nas pesquisas quantitativas publicadas sobre o campo de conhecimento de *clusters* industriais a partir dos anos 1990. Entre os anos 1970 e 1990 predominou a mensuração de *clusters*/complexos industriais com base em ligações setoriais insumo-produto, como a metodologia proposta Feser e Bergman. A partir de 1997, a literatura passou a se concentrar em medidas locais para identificar aglomerações regionais com o índice de Ellison e Glaeser. Em 2005, uma nova onda de estudos emergiu com base no índice de Duranton e Overman que testou as aglomerações com base na distância entre empresas em um espaço contínuo.

Diversos autores e periódicos contribuíram para o crescimento do *research front* analisado. Foram destacadas as contribuições de Edward Feser, com métodos insumo produto, estatística espacial com dados de área e pontos, e Giuseppe Arbia, com as publicações concentradas em mensurar aglomerações de firmas por meio de processos pontuais. Em relação aos periódicos, a maior quantidade de publicações foi do *Regional Studies* e o maior fator de impacto do *Journal of Economic Geography*. A *intellectual base* da amostra de artigos indicou Alfred Marshall, Michael Porter e Paul Krugman como principais teóricos que fundamentaram as pesquisas quantitativas sobre *clusters* e aglomerações de atividades econômicas.

A análise da linha do tempo das categorias em que foram publicados os trabalhos do *research front* indicou uma tendência de inclusão de novas áreas temáticas. Assim, foi

proposta uma nova classificação dos métodos, não excludentes, de mensuração de *clusters* setoriais composta por (1) métodos baseados em matrizes I-O e redução de dados; (2) índices de 1ª geração; (3) índices de 2ª geração; e (4) métodos espaciais.

Por fim, a rede de palavras-chave indicou uma maior ocorrência dos termos *geographic concentration* e *agglomeration*, em relação aos demais sinônimos *localization*, *location*, *patterns*, *clusters* e *specialization*. A elevada centralidade da palavra *innovation* indicou que os estudos do *research front* tendem a estar ligados a este tema. Em relação aos métodos, foi verificada a predominância dos métodos de estatística espacial para mensurar o fenômeno objeto deste estudo.

O crescimento dos métodos que pressupõem o espaço contínuo e a concentração é medida com base na distância entre pontos predominou no campo de aglomerações industriais após as contribuições de Duranton e Overman (2005) onde foram apresentados como solução para o MAUP. Essa nova geração de estudos problematizou o MAUP e colocou em evidência seus impactos nos resultados, no entanto, esta revisão bibliométrica indicou que outro viés permanece inexplorado na literatura sobre aglomerações de firmas, denominado anisotropia (ou viés direcional).

Referências

- Akgüngör, S. (2006). Geographic concentrations in Turkey's manufacturing industry: Identifying regional highpoint *clusters*. *European Planning Studies*, 14(2), 169-197.
- Akgüngör, S., Kumral, N., & Lenger, A. (2003). National industry *clusters* and regional specializations in Turkey. *European Planning Studies*, 11(6), 647-669.
- Anderson, D., & Johnston, S. A. (1992). A linkage approach to industrial location. *Growth and Change*, 23(3), 321-334.
- Anselin, L. (1995). Local Indicators of Spatial Association—LISA. *Geographical Analysis*, 27(2), 93-115.
- Arbia, G. (2001). The role of spatial effects in the empirical analysis of regional concentration. *Journal of Geographical Systems*, 3(3), 271-281.
- Arbia, G., Espa, G., Giuliani, D., & Dickson, M. M. (2014). Spatio-temporal *clustering* in the pharmaceutical and medical device manufacturing industry: A geographical micro-level analysis. *Regional Science and Urban Economics*, 49, 298-304.
- Arbia, G., Espa, G., Giuliani, D., & Mazzitelli, A. (2010). Detecting the existence of space-time *clustering* of firms. *Regional Science and Urban Economics*, 40(5), 311-323.

- Arbia, G., Espa, G., Giuliani, D., & Mazzitelli, A. (2012). *Clusters of firms in an inhomogeneous space: The high-tech industries in Milan. Economic Modelling, 29(1)*, 3-11.
- Arbia, G., Espa, G., & Quah, D. (2008). A class of spatial econometric methods in the empirical analysis of *clusters* of firms in the space. *Empirical Economics, 34(1)*, 81-103.
- Audretsch, D. B., & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *The American Economic Review, 86(3)*, 630-640.
- Audretsch, D. B., & Fritsch, M. (1994). The geography of firm births in Germany. *Regional Studies, 28(4)*, 359-365.
- Bergman, E. M., & Feser, E. J. (1999). *Industrial and regional clusters: concepts and comparative applications*. Morgantown: WVU Regional Research Institute
- Bersimis, S., Chalkias, C., & Anthopoulou, T. (2014). Detecting and interpreting *clusters* of economic activity in rural areas using scan statistic and LISA under a unified framework. *Applied Stochastic Models in Business and Industry, 30(5)*, 573-587.
- Bertinelli, L., & Decrop, J. (2005). Geographical agglomeration: Ellison and Glaeser's index applied to the case of Belgian manufacturing industry. *Regional Studies, 39(5)*, 567-583.
- Bocci, C., & Rocco, E. (2016). Modelling the location decisions of manufacturing firms with a spatial point process approach. *Journal of Applied Statistics, 43(7)*, 1226-1239.
- Bonneu, F., & Thomas-Agnan, C. (2015). Measuring and testing spatial mass concentration with micro-geographic data. *Spatial Economic Analysis, 10(3)*, 289-316.
- Braunerhjelm, P., & Carlsson, B. (1999). *Industry Clusters in Ohio and Sweden, 1975-1995. Small Business Economics, 12(4)*, 279-293.
- Briant, A., Combes, P. P., & Lafourcade, M. (2010). Dots to boxes: Do the size and shape of spatial units jeopardize economic geography estimations?. *Journal of Urban Economics, 67(3)*, 287-302.
- Carlei, V., & Nuccio, M. (2014). Mapping industrial patterns in spatial agglomeration: A SOM approach to Italian industrial districts. *Pattern Recognition Letters, 40*, 1-10.
- Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology, 57(3)*, 359-377.
- Combes P. & Overman H. G. (2004). The spatial distribution of economic activities in the European Union. In: Henderson J. V. & Thisse J. F. (Eds) *Handbook of Regional and Urban Economics*, pp. 2845-2909. Elsevier, Amsterdam.
- Cruz, S. C., & Teixeira, A. A. (2010). The evolution of the *cluster* literature: shedding light on the regional studies-regional science debate. *Regional Studies, 44(9)*, 1263-1288.

- Czamanski, S., & Ablas, L. A. D. Q. (1979). Identification of industrial *clusters* and complexes: a comparison of methods and findings. *Urban Studies*, 16(1), 61-80.
- De Dominicis, L., Arbia, G., & De Groot, H. L. (2013). Concentration of manufacturing and service sector activities in Italy: accounting for spatial dependence and firm size distribution. *Regional Studies*, 47(3), 405-418.
- Delgado, M., Porter, M. E., & Stern, S. (2014). Clusters, convergence, and economic performance. *Research Policy*, 43(10), 1785-1799.
- Delgado, M., Porter, M. E., & Stern, S. (2016). Defining *clusters* of related industries. *Journal of Economic Geography*, 16(1), 1-38.
- Do Prado, J. W., de Castro Alcântara, V., de Melo Carvalho, F., Vieira, K. C., Machado, L. K. C., & Tonelli, D. F. (2016). Multivariate analysis of credit risk and bankruptcy research data: a bibliometric study involving different knowledge fields (1968–2014). *Scientometrics*, 106(3), 1007-1029.
- Duranton, G., & Overman, H. G. (2005). Testing for localization using micro-geographic data. *The Review of Economic Studies*, 72(4), 1077-1106.
- Duranton, G., & Overman, H. G. (2008). Exploring the detailed location patterns of UK manufacturing industries using microgeographic data. *Journal of Regional Science*, 48(1), 213-243.
- Ellison, G., & Glaeser, E. L. (1997). Geographic Concentration in US Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy*, 105(5), 889-927.
- Ellison, G., Glaeser, E. L., & Kerr, W. R. (2010). What causes industry agglomeration? Evidence from coagglomeration patterns. *The American Economic Review*, 100(3), 1195-1213.
- Espa, G., Arbia, G., & Giuliani, D. (2013). Conditional versus unconditional industrial agglomeration: disentangling spatial dependence and spatial heterogeneity in the analysis of ICT firms' distribution in Milan. *Journal of Geographical Systems*, 15(1), 31-50.
- Falcioglu, P., & Akgüngör, S. (2008). Regional Specialization and Industrial Concentration Patterns in the Turkish Manufacturing Industry: An Assessment for the 1980–2000 Period. *European Planning Studies*, 16(2), 303-323.
- Feser, E. J. (2003). What regions do rather than make: A proposed set of knowledge-based occupation *clusters*. *Urban Studies*, 40(10), 1937-1958.
- Feser, E. J., & Bergman, E. M. (2000). National industry *cluster* templates: a framework for applied regional *cluster* analysis. *Regional Studies*, 34(1), 1-19.
- Feser, E., & Isserman, A. (2009). The rural role in national value chains. *Regional Studies*, 43(1), 89-109.

- Feser, E., Renski, H., & Goldstein, H. (2008). *Clusters and economic development outcomes an analysis of the link between clustering and industry growth. Economic Development Quarterly, 22*(4), 324-344.
- Feser, E. J., & Sweeney, S. H. (2000). A test for the coincident economic and spatial clustering of business enterprises. *Journal of Geographical Systems, 2*(4), 349-373.
- Feser, E., Sweeney, S., & Renski, H. (2005). A descriptive analysis of discrete US industrial complexes. *Journal of Regional Science, 45*(2), 395-419.
- Figueiredo, O., Guimarães, P., & Woodward, D. (2009). Localization economies and establishment size: was Marshall right after all?. *Journal of Economic Geography, 9*(6), 853-868.
- Figueiredo, O., Guimarães, P., & Woodward, D. (2010). Vertical disintegration in Marshallian industrial districts. *Regional Science and Urban Economics, 40*(1), 73-78.
- Fingleton, B. (1994). The location of high-technology manufacturing in Great Britain: changes in the late 1980s. *Urban Studies, 31*(1), 47-57.
- Fujita, M., Krugman, P. R., & Venables, A. (2001). *The spatial economy: Cities, regions, and international trade*. MIT Press.
- Giuliani, D., Arbia, G., & Espa, G. (2014). Weighting Ripley's K-function to account for the firm dimension in the analysis of spatial concentration. *International Regional Science Review, 37*(3), 251-272.
- Guimarães, P., Figueiredo, O., & Woodward, D. (2007). Measuring the localization of economic activity: a parametric approach. *Journal of Regional Science, 47*(4), 753-774.
- Guimarães, P., Figueiredo, O., & Woodward, D. (2009). Dartboard tests for the location quotient. *Regional Science and Urban Economics, 39*(3), 360-364.
- Guimarães, P., Figueiredo, O., & Woodward, D. (2011). Accounting for neighboring effects in measures of spatial concentration. *Journal of Regional Science, 51*(4), 678-693.
- Gülcan, Y., Kuştepe, Y., & Akgüngör, S. (2009). Public policies and development of the tourism industry in the Aegean region. *European Planning Studies, 17*(10), 1509-1523.
- Guo, L., Lou, X., Shi, P., Wang, J., Huang, X., & Zhang, J. (2015). Flow distances on open flow networks. *Physica A: Statistical Mechanics and its Applications, 437*, 235-248.
- Hervas-Oliver, J. L., González, G., MERI, P. C., & Sempere-Ripoll, F. (2015). *Clusters and Industrial Districts: Where is the Literature Going? Identifying Emerging Sub-Fields of Research. European Planning Studies, 23*(9), 1827-1872.
- Hill, E. W., & Brennan, J. F. (2000). A methodology for identifying the drivers of industrial clusters: the foundation of regional competitive advantage. *Economic Development Quarterly, 14*(1), 65-96.

- Holmes, T. J. (1999). Localization of industry and vertical disintegration. *The Review of Economics and Statistics*, 81(2), 314-325.
- Holmes, T. J., & Stevens, J. J. (2002). Geographic concentration and establishment scale. *Review of Economics and Statistics*, 84(4), 682-690.
- Howells, J., & Bessant, J. (2012). Introduction: Innovation and economic geography: a review and analysis. *Journal of Economic Geography*, 12(5), 929-942.
- Huisman, C., & van Wissen, L. (2004). Localization effects of firm startups and closures in the Netherlands. *The Annals of Regional Science*, 38(2), 291-310.
- Kominers, S. D. (2008). Measuring agglomeration. Harvard Urban and Social Economics Seminar.
- Kosfeld, R., Eckey, H. F., & Lauridsen, J. (2011). Spatial point pattern analysis and industry concentration. *The Annals of Regional Science*, 47(2), 311-328.
- Krugman, P. 1991a Geography and Trade. Cambridge, Massachusetts: MIT Press.
- Krugman, P. (1991b) Increasing returns and economic geography. *Journal of Political Economy*, 99(3), pp. 483-499.
- Lazzeretti, L., Sedita, S. R., & Caloffi, A. (2014). Founders and disseminators of *cluster* research. *Journal of Economic Geography*, 14(1), 21-43.
- Marcon, E., & Puech, F. (2003). Evaluating the geographic concentration of industries using distance-based methods. *Journal of Economic Geography*, 3(4), 409-428.
- Marcon, E., & Puech, F. (2010). Measures of the geographic concentration of industries: improving distance-based methods. *Journal of Economic Geography*, 10(5), 745-762.
- Marcon, E., & Puech, F. (2017). A typology of distance-based measures of spatial concentration. *Regional Science and Urban Economics*, 62, 56-67.
- Marshall, A. (1920) Principles of Economics. Revised edn. London: Macmillan; reprinted by Prometheus Books, 1st edn., 1890.
- Martínez-Fernández, M. T., Capó-Vicedo, J., & Vallet-Bellmunt, T. (2012). The present state of research into industrial clusters and districts. Content analysis of material published in 1997-2006. *European Planning Studies*, 20(2), 281-304.
- Maurel, F., & Sédillot, B. (1999). A measure of the geographic concentration in French manufacturing industries. *Regional Science and Urban Economics*, 29(5), 575-604.
- Menon, C. (2012). The bright side of MAUP: Defining new measures of industrial agglomeration. *Papers in Regional Science*, 91(1), 3-28.
- Mori, T., Nishikimi, K., & Smith, T. E. (2005). A divergence statistic for industrial localization. *Review of Economics and Statistics*, 87(4), 635-651.

- Nakamura, R., & Paul, C. J. M. (2009). Measuring agglomeration. In: Capello R. and Nijkamp P. (Eds) *Handbook of Regional Growth and Development Theories*, pp. 305-328, Edward Elgar, Cheltenham, UK.
- Nathan, M., & Overman, H. (2013). Agglomeration, *clusters*, and industrial policy. *Oxford Review of Economic Policy*, 29(2), 383-404.
- Nielsen, M. M., & Hennerdal, P. (2014). MAUPing workplace *clusters*. *Growth and Change*, 45(2), 211-221.
- Ó'Huallacháin, B. (1991). Sectoral *clustering* and growth in American metropolitan areas. *Regional Studies*, 25(5), 411-426.
- Oliva, G., Setola, R., & Panzieri, S. (2016). Critical *clusters* in interdependent economic sectors. *The European Physical Journal*, 225(10), 1929-1944.
- Porter, M. (2003). The economic performance of regions. *Regional Studies*, 37(6-7), 549-578.
- Porter, M. E. (1990). *The Competitive Advantage of Nations*. New York: Free Press.
- Porter, M. E. (1998). *Clusters* and the new economics of competition. *Harvard Business Review*, 76(6), pp. 77-90.
- Renski, H., Koo, J., & Feser, E. (2007). Differences in labor versus value chain industry *clusters*: An empirical investigation. *Growth and Change*, 38(3), 364-395.
- Rivera, L., Sheffi, Y., & Welsch, R. (2014). Logistics agglomeration in the US. *Transportation Research Part A: Policy and Practice*, 59, 222-238.
- Rosenthal, S. S., & Strange, W. C. (2001). The determinants of agglomeration. *Journal of Urban Economics*, 50(2), 191-229.
- Rybnikova, N. A., & Portnov, B. A. (2014). Mapping geographical concentrations of economic activities in Europe using light at night (LAN) satellite data. *International Journal of Remote Sensing*, 35(22), 7706-7725.
- Saxenian, A. (1990) Regional networks and the resurgence of Silicon Valley. *California Management Review*, 33(1), pp. 89-113.
- Saxenian, A. (1994) *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Cambridge, MA: Harvard University Press).
- Scholl, T., & Brenner, T. (2016). Detecting spatial *clustering* using a firm-level *cluster* index. *Regional Studies*, 50(6), 1054-1068.
- Stejskal, J. (2011). Analysis of the applicability of selected methods for industrial *clusters* identifying. *International Journal of Systems Applications, Engineering and Development*, 3(5), 255-262.

- Stejskal, J., & Hajek, P. (2012). Competitive advantage analysis: a novel method for industrial *clusters* identification. *Journal of Business Economics and Management*, 13(2), 344-365.
- Storper, M. (1995). The resurgence of regional economies, ten years later: The region as a nexus of untraded interdependencies. *European Urban & Regional Studies*, 2(3), pp. 191–221.
- Storper, M. (1997). *The Regional World: Territorial Development in a Global Economy*. New York: Guilford Press.
- Sweeney, S. H., & Feser, E. J. (1998). Plant size and *clustering* of manufacturing activity. *Geographical Analysis*, 30(1), 45-64.
- Viladecans-Marsal, E. (2004). Agglomeration economies and industrial location: city-level evidence. *Journal of Economic Geography*, 4(5), 565-582.
- Vom Hofe, R., & Chen, K. (2006). Whither or not industrial *cluster*: conclusions or confusions. *The Industrial Geographer*, 4(1), 2-28.
- Webster, T. J. (2013). Economic development strategies and methods for identifying leading industries and industrial *clusters*. *International Journal of Economics and Business Research*, 5(1), 55-74.
- Wei, F., Grubestic, T. H., & Bishop, B. W. (2015). Exploring the GIS knowledge domain using CiteSpace. *The Professional Geographer*, 67(3), 374-384.
- Xing, L., Ye, Q., & Guan, J. (2016). Spreading Effect in Industrial Complex Network Based on Revised Structural Holes Theory. *PloS One*, 11(5), 1-18.

**3 MAPEAMENTO DE POTENCIAIS *CLUSTERS* INDUSTRIAIS PELA
ABORDAGEM GEOESTATÍSTICA**

Norma NBR 6022 (ABNT 2003)

RESUMO

O objetivo do presente artigo foi examinar a proximidade e a concentração de firmas no espaço contínuo e em direções específicas para o mapeamento de potenciais *clusters* industriais, por meio de uma abordagem geoestatística. Os métodos utilizados foram o semivariograma indicador, a modelagem da anisotropia e a krigagem. Foram aplicados inicialmente em dados simulados e posteriormente validados em dados empíricos, referentes à indústria do café em Minas Gerais. A abordagem desenvolvida foi suficiente para detectar os padrões de aleatoriedade e de dependência espacial, mensurar o alcance da proximidade entre firmas, identificar no mapa as regiões com elevado nível de concentração industrial e estimar um indicador de concentração no nível da firma. Os resultados também indicaram que os modelos anisotrópicos (direcionais) descreveram melhor a realidade dos fenômenos em comparação aos isotrópicos. Concluiu-se que a direção pode ser um determinante da teoria das aglomerações e que a abordagem geoestatística viabilizou o mapeamento de potenciais *clusters* industriais.

Palavras-chave: Teoria das aglomerações. Dependência espacial. Anisotropia.

3.1 INTRODUÇÃO

O conceito de *clusters* industriais pode ser subdividido em um sentido econômico e em um sentido geográfico (FESER; SWEENEY, 2000; FESER; ISSERMAN, 2009). Essa diferenciação objetivou demonstrar que as firmas com maior intensidade nas relações entre fornecedores e compradores estão propensas a se aglomerar no espaço. Segundo esses autores, o *cluster* econômico se refere aos vínculos ou interações entre firmas como resultado de relacionamentos econômicos e tecnológicos que incluem compra e venda de insumos, similaridades nas tecnologias de produção e/ou mercados, compartilhamento da força de trabalho e intercâmbio de informações em ambientes formais e informais. Por sua vez, o *cluster* geográfico refere-se às relações de dependência espacial entre empresas, isto é, clusters econômicos que se manifestam em aglomerações locais.

As economias das aglomerações se manifestam na concentração geográfica de atividades econômicas. Deve-se considerar também que indústrias geograficamente localizadas não garantem a existência de interação ou transbordamentos, porém um padrão de localização consistente entre várias regiões cria uma pressuposição forte de que esses fenômenos estão presentes (DELGADO; PORTER; STERN, 2016; PORTER, 2003). Assim, um conceito de potenciais *clusters* foi proposto por Carrol, Reid e Smith (2008) e definido como uma elevada concentração de firmas geograficamente próximas.

Um dos campos de conhecimento dentro da literatura das ciências regionais é direcionado para pesquisas em métodos e medidas para a identificação, classificação ou explicação de *clusters* de atividades econômicas. A análise do espaço influenciou a literatura deste campo e sustentou uma constante evolução das técnicas de análise regional, como os métodos baseados em Sistemas de Informações Geográficas (SIG) e as recentes contribuições da Econometria Espacial (CRUZ; TEIXEIRA, 2010).

O estado da arte da literatura sobre mensuração de aglomerações locais de firmas tem buscado contornar o Problema da Unidade de Área Modificável (MAUP) pela aplicação de funções de processos pontuais. Esta literatura segue a abordagem de métodos baseados em distâncias proposta no trabalho seminal de Duranton e Overman (2005). Uma revisão de literatura sobre esta abordagem *mainstream*, também conhecida por “Microeconometria Espacial”, e sua evolução pode ser encontrada em Marcon e Puech (2017).

O MAUP pode ser entendido como um fenômeno originado da arbitrariedade das fronteiras geográficas que nem sempre coincidem com as fronteiras econômicas. Como

observado por Krugman (1991, p. 71) “there is no particular reason to think that national boundaries define a relevant region”. Esse problema pode se manifestar em um viés de escala (agregação) ou de zoneamento. No primeiro caso, podem ocorrer variações nos resultados das análises estatísticas dependendo do número de unidades de área utilizado, ou seja, quando as unidades de área são combinadas em unidades maiores ou menores. No segundo, o viés é derivado de diferentes divisões de uma mesma área, mantendo a área e o número de divisões constantes (CRESSIE, 1993; OPENSHAW, 1984; SCHOLL; BRENNER, 2016).

No entanto, dois tópicos relevantes vêm sendo negligenciados pela literatura. O primeiro é o viés direcional, ou anisotropia, que pode se manifestar em dados espacialmente dependentes (ARBIA, 2013; CHAIN et al., 2017), ou seja, quando a distribuição ou a variabilidade de um fenômeno espacial é mais intensa em uma determinada direção, em comparação com as demais direções. A outra crítica é que, embora a abordagem de processos pontuais seja estatisticamente robusta, ela não tem como objetivo zonestar em mapas a localização das aglomerações de firmas (GUILLAIN; LE GALLO, 2010; MARCON; PUECH, 2010).

Diante do exposto, questionou-se a possibilidade de incorporar o viés direcional e o mapeamento na análise da concentração de firmas no espaço sem descartar os fundamentos já estabelecidos pela literatura *mainstream*, especificamente pelo trabalho de Duranton e Overman (2005, p. 1103): “(i) comparability across industries, (ii) control for the uneven distribution of overall manufacturing, (iii) control for industrial concentration, (iv) no aggregation bias, and (v) statistical significance”.

O objetivo do presente artigo foi examinar a proximidade e a concentração de firmas no espaço contínuo e em direções específicas, por meio de uma abordagem geoestatística. Especificamente, buscou-se identificar a estrutura de dependência espacial entre firmas, modelar o viés direcional, estimar a concentração setorial e mapear os potenciais *clusters* de atividades econômicas. Uma abordagem introdutória foi testada em dados simulados e, em seguida, validada com dados reais referentes à indústria do café no estado de Minas Gerais.

Entende-se ainda que o refinamento dos métodos de identificação de *clusters* geográficos se justifica visto que o padrão anisotrópico é mais uma regra do que uma exceção na observação de fenômenos da economia (ARBIA, 2013). Outro ponto observado por Scholl e Brenner (2016) é que a visualização dos mapas é uma ferramenta relevante para os *policy makers* envolvidos em políticas de *clusters* industriais.

3.2 REFERENCIAL TEÓRICO

A estatística espacial é utilizada para analisar dados espacialmente referenciados e pode ser dividida em padrões (processos) pontuais, dados de área (*lattice*) e na geoestatística (superfície). Também conhecida como Teoria das Variáveis Regionalizadas, a geoestatística teve sua origem na prospecção da concentração de ouro em depósitos minerais (MATHERON, 1963) e posteriormente foi incorporada pelas ciências agrárias para análise da concentração de nutrientes no solo (VIEIRA et al., 1983). Atualmente a geoestatística é aplicada em várias áreas do conhecimento. Seu objetivo final é quantificar e explorar o padrão espacial de uma variável regionalizada (LEUANGTHONG; KHAN; DEUTSCH, 2011).

Uma variável regionalizada pode ser definida como uma função espacial numérica, variante no espaço, com uma aparente continuidade e sua variação não pode ser representada por uma função matemática determinística (MATHERON, 1963). A geoestatística pode ser subdividida, principalmente, nos conceitos de variograma e krigagem. O primeiro caso consiste na estimativa da continuidade (ou dependência espacial), enquanto o segundo está interessado na interpolação de valores em qualquer posição dentro de uma determinada região, geralmente mapas, de forma ótima (não tendenciosa e com mínima variância) baseada na estrutura definida pelo variograma.

A presente explanação da teoria geoestatística seguiu principalmente a proposição de Oliver e Webster (2014), no entanto, mais detalhes e as devidas demonstrações encontram-se formalizadas em Cressie (1993), Bierkens e Burrough (1993), Goovaerts (1999), Matheron (1963), Vieira et al. (1983), entre outros.

O valor amostrado “z” em um local “x” (com coordenadas geográficas em duas dimensões x_1 e x_2) é considerado uma entre infinitas realizações de uma variável aleatória $Z(x)$ naquele respectivo local. Então, este valor amostrado (medido) “z” no local “x” pode ser escrito como $z(x)$. A variável $Z(x)$ pertence a um conjunto de variáveis aleatórias denominado processo (campo ou função) aleatório, também representado por $\{Z(x)\}$. Pelo fato de que os valores $\{z(x)\}$ em uma região tendem a apresentar alguma estrutura de dependência espacial, isto é, depender da sua vizinhança, a variável $z(x)$ é chamada de variável regionalizada.

As principais hipóteses sobre $Z(x)$ a serem assumidas estão relacionadas à estacionariedade do processo aleatório. A estacionariedade no espaço assume que o modo de variação é o mesmo de local para local. Um processo aleatório pode ser representado por:

$$Z(x) = \mu(x) + \varepsilon(x) \quad (01)$$

Em que $\mu(x)$ é a média do processo, se $\mu(x)$ é constante, então $\mu(x) = \mu$; $\varepsilon(x)$ é um valor aleatório com média zero e covariância $C(h)$. O termo $\varepsilon(x)$ pode ainda ser subdividido em $\varepsilon'(x)$ e $\varepsilon''(x)$ em que o primeiro é um termo estocástico espacialmente correlacionado e o segundo um ruído aleatório não correlacionado. Na hipótese de estacionariedade de 2ª ordem, a covariância depende somente de h e pode ser representada por:

$$\begin{aligned} C(h) &= E[\{Z(x) - \mu\}\{Z(x+h) - \mu\}] \\ &= E[Z(x)Z(x+h) - \mu^2] \end{aligned} \quad (02)$$

Em que h é a separação entre os pontos em termos de distância e direção, $Z(x)$ e $Z(x+h)$ são os valores de Z nas localidades (x) e $(x+h)$ e E é a esperança. Se a média $\mu(x)$ não é constante então a modelagem se torna mais complexa. Pode-se assumir uma hipótese menos restritiva de estacionariedade, a hipótese intrínseca, na qual a esperança da diferença entre $Z(x)$ e $Z(x+h)$ é zero, e a covariância é substituída pela metade da variância das diferenças (semivariância), como demonstrada na Equação 03.

$$\begin{aligned} \gamma(h) &= \frac{1}{2} \text{var} [Z(x) - Z(x+h)] \\ &= \frac{1}{2} E[\{Z(x) - Z(x+h)\}^2] \end{aligned} \quad (03)$$

A semivariância também depende somente de h , enquanto $\gamma(h)$ pode ser chamada de semivariograma em função de h . O semivariograma passou a ser mais utilizado que a covariância por ser menos exigente estatisticamente. Na hipótese mais restritiva, de um processo estacionário de segunda ordem, as funções de covariância e de semivariância são equivalentes para caracterizar a dependência espacial, como demonstrado na Equação 04:

$$\gamma(h) = C(0) - C(h) \quad (04)$$

Em que $C(0)$ é a variância do processo aleatório. A hipótese de estacionariedade de 2ª ordem, e consequentemente a hipótese intrínseca, são suficientes para a interpolação ótima de uma superfície utilizando a Krigagem Simples e a Krigagem Ordinária.

Uma análise da bibliografia relevante acerca da geoestatística foi elaborada por Hengl, Minasny e Gould (2009) seguindo critérios bibliométricos para sistematização da literatura e indicou que o tema foi pouco explorado pela categoria *business & economics*. Nesse campo, foram identificadas 5 publicações pela palavra-chave *geostatistics* e 151 por *spatial statistics*, em um total de 1.047 publicações referentes a todas as áreas de conhecimento, como ciências da terra, engenharia, matemática, entre outras. Deve-se considerar que o segundo caso abrange basicamente dados de área e processos pontuais.

Algumas contribuições da geoestatística para as ciências sociais podem ser destacadas: Chica Olmo (1995) elaborou uma metodologia espacial para estimar preços e renda no mercado imobiliário, enquanto Gillen, Thibodeau e Wachter (2001) investigaram a autocorrelação anisotrópica em preços residenciais. De forma inovadora, Wang e Zhang (2003) mensuraram o alcance da produtividade de culturas agrícolas na gestão de riscos de seguros agrícolas. Na área de ciência política, Tam Cho e Gimpel (2007) interpolaram doações de campanhas eleitorais para prospectar novos doadores. A contribuição de Fernández-Aviles, Montero e Orlov (2012) foi na análise das distâncias geográficas e financeiras entre países e suas relações com os mercados financeiros, enquanto Arbia e Di Marcantonio (2015) utilizaram o variograma e a krigagem para previsão de taxas de juros.

Na identificação de *clusters* de atividades econômicas foram encontradas duas aplicações da geoestatística, ainda como método secundário. Aguilar e Vlosky (2006) utilizaram o semivariograma isotrópico para analisar a dependência espacial entre firmas de produtos florestais em Louisiana, EUA, entretanto, a presença de *clusters* foi detectada por padrões pontuais. Por sua vez, Rybnikova e Portnov (2015) utilizaram a interpolação do Índice de Moran Local, por meio da krigagem, como tentativa de suavizar o MAUP e aumentar a precisão da delimitação da concentração setorial na Europa.

3.3 METODOLOGIA

Nesta etapa estão os itens que compõem os procedimentos metodológicos da pesquisa. Inicialmente faz-se uma breve contextualização sobre a indústria do café no Brasil e em Minas Gerais, objeto de estudo, em seguida apresenta-se os dados utilizados no cenário simulado e no cenário aplicado à indústria do café em Minas Gerais. Finalmente, apresentam-se os métodos que compõe a abordagem geostatística para o mapeamento de potenciais *clusters* industriais.

3.3.1 Objeto de estudo - a indústria do café

A indústria de torrefação do café possui relevância socioeconômica e um elevado efeito multiplicador insumo-produto. Um novo emprego nesse setor tende a gerar 17 empregos diretos e indiretos e representa o quarto maior efeito multiplicador entre os 42 setores da economia brasileira (BRENE et al. 2014). Segundo Sakon et al. (2012), a indústria do café pode ser considerada setor-chave para o sistema produtivo brasileiro, principalmente por ser uma grande demandante de bens e serviços relativamente aos demais setores.

O processamento da indústria do café no Brasil é concentrado em 10 empresas, com capacidade de produção superior a 480 toneladas/mês cada, que ofertam aproximadamente 57% do café torrado e moído, sendo que 8 delas estão localizadas em São Paulo e Minas Gerais, segundo a Associação Brasileira das Indústrias do Café (ABIC, 2014). No mercado de São Paulo estão localizadas majoritariamente as empresas de médio e grande porte (média de 19 funcionários por empresa), enquanto em Minas Gerais encontram-se principalmente as de pequeno e médio porte (média de 9 funcionários por empresa), segundo dados do Ministério do Trabalho e Emprego (MTE, 2014).

O Brasil apresenta também uma questão regional interessante referente ao agronegócio café. A cafeicultura em Minas Gerais produziu 56% dos grãos, enquanto seu vizinho São Paulo produziu 8%, segundo dados da Companhia Nacional de Abastecimento (CONAB, 2013). Em relação à industrialização, Minas Gerais processou 17% do café torrado e São Paulo 41%, Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE, 2013). Embora grande parte do café verde produzido em Minas Gerais seja direcionada para o mercado externo, nota-se uma desigualdade regional na agregação de valor ao produto final.

Minas Gerais tem se destacado na cafeicultura em decorrência das condições de cultivo favoráveis. A partir dos anos 90, grandes produtores como São Paulo e Paraná perderam participação na área plantada influenciados por problemas climáticos e preço da terra. No mesmo período, a agroindústria do café no Brasil foi desregulamentada e o Instituto Brasileiro do Café, organização reguladora, encerrou suas atividades. Como consequência, houve a saída das torrefadoras pouco competitivas do mercado, abertura para as multinacionais e a concentração da industrialização de café em São Paulo.

Em geral, as torrefadoras fornecem café para o mercado local e se encontram em uma estrutura de mercado próximo da concorrência perfeita, com muitos concorrentes, poucas barreiras à entrada e oferta de produtos praticamente homogêneos. As margens de lucro também se reduzem na medida em que o preço do insumo principal é definido na Bolsa de Mercadorias e Futuros e o preço de venda é altamente influenciado pelas grandes redes de supermercados e varejo.

A política industrial brasileira buscou aperfeiçoar sistemas e mecanismos para promover a sustentabilidade econômica e social na agroindústria do café por meio da estratégia de *clusters*. A necessidade de identificação de potenciais *clusters* (ou Arranjos Produtivos Locais - APLs) foi expressa no “Plano Brasil Maior,” em que uma de suas metas foi “Definir as regiões e validar projetos de APL para a industrialização de café torrado e moído” (MDIC, 2013, p.159).

Como justificativa para escolha do setor de torrefação do café deve-se remeter aos fundamentos de políticas industriais, especificamente os instrumentos verticais (FERRAZ; DE PAULA; KUPFER, 2013). Esse tipo de instrumento privilegia uma indústria específica, ou seja, a partir de medidas estratégicas, o Estado direciona recursos para um conjunto de empresas, indústrias ou cadeias produtivas relevantes para a economia nacional ou regional.

3.3.2 Banco de dados

Os padrões de localização utilizados nos cenários simulados se basearam nos critérios definidos por Scholl e Brenner (2016), para tornar a abordagem geoestatística replicável e comparável.

Todos os pontos foram gerados e referenciados em um quadro de 2x2 km, por meio de coordenadas “x” e “y”. As firmas simuladas como grupo controle (Categoria *benchmark*) foram geradas a partir de um Processo de Poisson Homogêneo (PPH) com intensidade de 500.

As firmas na Categoria 1 também foram geradas com base no PPH de intensidade 75 para testar a capacidade do modelo em detectar o padrão de ausência de dependência espacial. Na Categoria 2 foram geradas firmas seguindo o PPH (intensidade 50) e adicionadas as firmas a partir de um processo Matérn, ou seja, dois *clusters* com raio de aproximadamente 0,1 quilômetro (km). Aos pontos do grupo controle foram atribuídos o valor zero e aos pontos das Categorias 1 e 2 foram atribuídos o valor de 1.

Em relação à abordagem empírica, o grupo controle foi gerado a partir dos endereços de firmas industriais coletadas no Cadastro Industrial de Minas Gerais (CIMG) disponibilizado pelo Centro Industrial Empresarial de Minas Gerais (CIEMG, 2014). O CIMG disponibilizou dados referentes a mais de 30 setores industriais que, embora demandem uma análise computacionalmente intensiva, não são censitários. As firmas de indústria de café tiveram seus endereços disponibilizados pela Relação Anual de Informações Sociais (RAIS) do MTE (2013), mediante solicitação, e abrangeram dados censitários referentes às empresas formais inseridas no grupo 108 (“Torrefação e moagem de café”) da Classificação Nacional de Atividades Econômicas (CNAE) do IBGE. O ano base dos dois bancos de dados foi 2013.

Para coletar as variáveis “latitude” e “longitude”, os endereços das firmas foram transformados em coordenadas geográficas pelo *Google Maps* e projetadas em *Albers Equal Conic* para o georreferenciamento em coordenadas métricas.

Os dados foram processados em ambiente R, com auxílio de diversos pacotes. Os pontos para a análise simulada foram obtidos a partir do *spatstat* (BADDELEY; TURNER, 2005). As análises geoestatísticas foram conduzidas com auxílio do *gstat* (PEBESMA, 2004), com exceção dos envelopes semivariográficos que foram obtidos pelo *geoR* (RIBEIRO Jr; DIGGLE, 2001).

3.3.3 Abordagem geoestatística

A abordagem geoestatística foi composta pela análise do semivariograma, da anisotropia e da Krigagem. A primeira análise indica o alcance da dependência espacial entre firmas, enquanto a modelagem da anisotropia contempla o viés direcional nas relações de proximidade. Por fim, a krigagem demonstra os resultados em mapas e estima o nível de concentração de cada firma individualmente.

3.3.3.1 (Semi)variograma

O semivariograma experimental (empírico ou amostral) é aquele estimado com base no conjunto de dados $z(x_i)$ e na distância h , ou seja, baseado na distância entre pares de pontos. Esta propriedade atende ao pressuposto (iv) – de ausência do viés de agregação (DURANTON; OVERMAN, 2005). O estimador de semivariância ($\hat{\gamma}(h)$) pode ser calculado pelo método dos momentos (MATHERON, 1963), em outras palavras, uma forma de se estimar o semivariograma é pela média do quadrado da diferença de pares de observações que estão separados por uma mesma distância (Equação 05):

$$\hat{\gamma}(h) = \frac{1}{2m(h)} \sum_{\substack{\forall x_i, x_j \\ \text{tal que} \\ |x_i - x_j| = h}} \{z(x_i) - z(x_j)\}^2 \quad (05)$$

Em que $m(h)$ é o número de pontos amostrais a uma distância h . Com os incrementos em h é possível obter um conjunto de pontos que indicam a estrutura de dependência espacial dos dados e que podem ser analisados de forma gráfica (gráfico do semivariograma). Quanto menor a semivariância, mais similares são as informações amostradas e espera-se que os pontos se tornem mais dissimilares na medida em que a distância aumenta.

O semivariograma indicador é um caso particular por não trabalhar com variáveis contínuas, mas com atributos categóricos ou pontos de corte, que normalmente são transformados em uma variável indicadora do tipo 1 (pertence a determinada categoria) e 0 (caso contrário). Neste caso, $z(x_i) = i(u_\alpha; s_k)$ e $z(x_i + h) = i(u_\alpha + h; s_k)$, em que $i(u_\alpha; s_k)$ é a variável indicadora i na coordenada geográfica u_α referente ao atributo s da categoria k . Neste semivariograma os pares de pontos – separados por uma distância (h) – pertencentes a categorias opostas são os que mais contribuem para a estimativa da semivariância, quanto menor $\hat{\gamma}(h)$, mais conectada é a categoria tipo 1 (GOOVAERTS, 2009).

A propriedade indicadora atende aos pressupostos (ii) controlar a tendência das indústrias se aglomerarem no espaço e (iii) controlar a concentração industrial (DURANTON; OVERMAN, 2005).

O semivariograma experimental pode ainda ser comparado com um envelope de semivariogramas computados por permutações aleatórias e independentes (normalmente 999) dos dados entre as localizações observadas (DIGGLE; RIBEIRO Jr., 2007). Esse envelope

simulado pelo método de Monte Carlo indica a hipótese de ausência de correlação espacial e um ponto fora dele confirma a significância estatística do semivariograma, ou seja, a presença de correlação espacial. Esta análise buscou atender ao pressuposto – (v) – significância estatística (DURANTON; OVERMAN, 2005).

Após a estimação do semivariograma experimental deve-se ajustar um modelo que descreva as principais características dos dados amostrados e ignore variações erráticas. Esta curva é denominada semivariograma teórico e é expressa por uma equação matemática. Alguns ajustes teóricos são Esférico, Exponencial, *Wave*, entre outros. Para selecionar o modelo mais bem ajustado foi utilizada a menor Soma dos Erros Quadrados Ponderados (SEQP) dos resíduos entre o semivariograma experimental e o teórico (CRESSIE, 1993).

Entre os exemplos mais populares, o modelo esférico isotrópico pode ser representado pela Equação 06 (OLIVER; WEBSTER, 2014):

$$\begin{aligned} \gamma(h) &= 0 \text{ para } |h| = 0 & (06) \\ &= c_0 + c \left\{ \frac{3|h|}{2r} - \frac{1}{2} \left(\frac{|h|}{r} \right)^3 \right\} \text{ para } 0 < |h| \leq r \\ &= c_0 + c \text{ para } |h| > r \end{aligned}$$

Os parâmetros c_0 , c e r são obtidos pela função do semivariograma. O efeito pepita (c_0) é a variação não explicada na distância zero e pode incluir também erros de mensuração, na teoria a semivariância em uma distância zero deveria ser zero; a contribuição – ou patamar parcial – (c) indica a variância espacialmente correlacionada; o patamar ($c_0 + c$) é uma estimativa da variância máxima das realizações do processo aleatório; por fim, o alcance (r) é o limite da correlação espacial, ou seja, uma distância a partir da qual os pares de realizações de um processo aleatório deixam de ser correlacionados (espacialmente dependentes).

3.3.3.2 Anisotropia

O viés direcional, ou anisotropia, pode ser definido como a variação no comportamento do semivariograma devido às mudanças na direção. Quando ignorada ou modelada indevidamente, a anisotropia pode induzir a resultados enganosos em relação ao fenômeno real (OLIVER; WEBSTER, 2014).

A anisotropia pode ser classificada em geométrica, zonal e mista (combinação dos dois casos). No primeiro caso, os semivariogramas direcionais apresentam alcances

diferentes, enquanto no caso seguinte os patamares é que diferem. Quando esse comportamento é similar independente da direção, isto é, depende somente da distância entre pares de pontos, pode-se dizer que o fenômeno em estudo é isotrópico.

A anisotropia geométrica pode ser tratada ao se modelar o alcance como uma elipse definida pelo alcance máximo (maior direção) e mínimo (menor direção), por meio de uma transformação linear das coordenadas que torna a variação isotrópica. O novo parâmetro de alcance anisotrópico (r_ϕ) substitui o alcance no modelo teórico isotrópico (exemplo no modelo esférico da Equação 06), seguindo a Equação 07 (ERIKSSON; SISKKA, 2000):

$$r_\phi = \eta r / \sqrt{\eta^2 \cos^2(\phi - \theta) + \text{sen}^2(\phi - \theta)} \quad (07)$$

Em que r_ϕ é o alcance do semivariograma na direção ϕ que, por sua vez, é o ângulo de separação entre os pares de pontos nas coordenadas originais e transformadas; θ é a direção do semivariograma de maior dependência espacial ($r_{\text{máx}}$); e η é a razão entre o alcance na direção θ e o alcance na direção perpendicular ($r_{\text{mín}}$); se $\eta < 1$, $r_{\text{máx}} = r$ e $r_{\text{mín}} = \eta r$.

As formas de modelar os tipos de anisotropia em uma elipse e os respectivos cálculos foram apresentadas por Eriksson e Siska (2000). Segundo esses autores, a anisotropia zonal pode ser ajustada por meio da estrutura geométrica com uma elevada razão e um alcance tendendo ao infinito na direção do menor patamar. Para anisotropia mista, a modelagem ocorre por meio da adição do modelo geométrico ao zonal (semivariogramas aninhados).

3.3.3.3 Krigagem de variáveis categóricas

A previsão de variáveis categóricas (qualitativas) deve considerar a correlação espacial, a transição de probabilidades entre elas e gerar medidas de confiabilidade nas estimativas (GOOVAERTS, 1999). Para tal, é necessário transformar os dados categóricos em funções indicadoras que classificam dicotomicamente cada observação, sendo 1 (pertence a determinada categoria) e 0 (caso contrário) e assim viabilizar a krigagem indicadora. Este método indica a probabilidade de uma categoria ocorrer em uma localização específica não amostrada com base na função de probabilidade condicional (p) apresentada na Equação 08:

$$p(u; s_k | (n)) = \text{Prob}\{S(u) = s_k | (n)\} \quad k = 1, \dots, K \quad (08)$$

Em que u é a localização não amostrada; s_k é o atributo s da categoria k ; n é o número de observações disponíveis; $|n$ indica que a função está condicionada a informação da vizinhança local de s_k ; e $S(u)$ é um campo aleatório com realizações categóricas na localização u não amostradas. Para estimar os valores da função de probabilidade condicional pode-se utilizar a combinação linear da função indicadora dos vizinhos, por meio do estimador de krigagem ordinária (OK) descrito na Equação 09:

$$[p(u; s_k | n)]_{OK} = \sum_{\alpha=1}^{n(u)} \lambda_{\alpha} i(u_{\alpha}; s_k) \quad (09)$$

Com λ_{α} sendo a solução do sistema de equações lineares que minimiza a variância do erro de previsão, denominado peso de krigagem. Estes pesos levam em consideração a configuração espacial dos dados, a proximidade entre os dados e as localizações não amostradas e o padrão modelado pelo semivariograma experimental (GOOVAERTS, 2009). A variável indicadora observada é dada por $i(u_{\alpha}; s_k)$, como no item 3.2.2. O resultado da krigagem pode ser visualizado no “mapa de krigagem” que é importante para delimitar as regiões (superfícies) com elevada concentração setorial.

A validação da krigagem foi utilizada para verificar a consistência das estimativas e, conseqüentemente, o melhor semivariograma ajustado. Neste processo, os pontos observados são excluídos e comparados com sua própria previsão (*leave one out*). Espera-se que o melhor modelo apresente a menor Soma do Quadrado dos Resíduos (SQR), o Resíduo Médio Padronizado (RMP) próximo de zero e a Raiz do Resíduo Médio Quadrado Padronizado (RRMQP) próxima de um (CRESSIE, 1993).

A probabilidade mapeada e prevista pela validação nada mais é do que um indicador de concentração no nível da firma, aqui denominado Índice Geoestatístico de Concentração (IGC). Em nível setorial o IGC pode ser obtido pela probabilidade média das firmas pertencentes a classe em análise. Esse índice varia entre 0, ausência de concentração, e 1, máxima concentração e buscou atender ao pressuposto (i) – comparabilidade entre indústrias (DURANTON; OVERMAN, 2005).

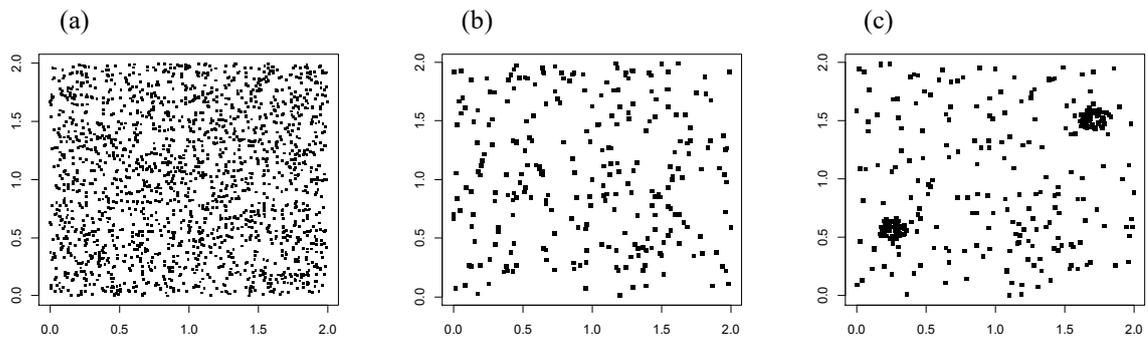
As firmas com elevado IGC indicam o centro dos *clusters* (*cluster core*). O ponto de corte “> 50%” de probabilidade foi selecionado para mapear os potenciais *clusters*, ou seja, maior a probabilidade de ocorrência da categoria em estudo devido ao perfil das firmas na vizinhança, dado o alcance do semivariograma, logo um alto nível de concentração industrial.

3.4 RESULTADOS E DISCUSSÃO

3.4.1 Análise simulada

Na Figura 1(a) estão localizadas as firmas simuladas que formam o grupo controle (Categoria *benchmark*), em um total de 1.904 pontos. A Figura 1(b) apresentou a distribuição espacial das firmas simuladas, na qual todas as 262 firmas seguem o PPH (Categoria 1). Na Figura 1(c) encontra-se a distribuição da Categoria 2, formada por um PPH com 206 firmas acrescidas de um Processo Matérn com dois *clusters* (canto inferior esquerdo e canto superior direito) com 49 firmas cada e distantes entre si em 1,8 km.

Figura 1- Distribuição espacial das firmas simuladas.

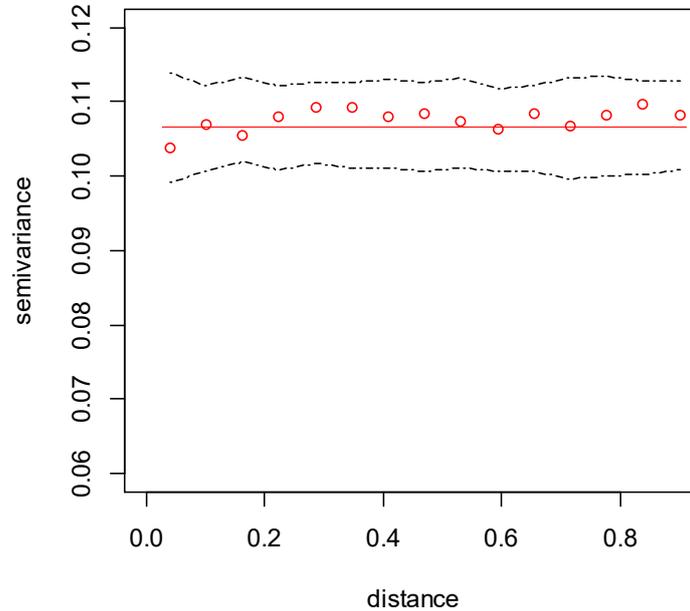


Fonte: Elaborado pelo autor com base em Scholl e Brenner (2016).

3.4.1.1 Cenário 1

O Cenário 1 foi formado pela sobreposição das firmas da Categoria 1 no *benchmark*. O gráfico do semivariograma isotrópico (omnidirecional) da variável indicadora está exibido na Figura 2.

Figura 2 - Semivariograma isotrópico, modelo ajustado e o envelope de confiança estatística, Cenário 01.



Fonte: Elaborado pelo autor.

O comportamento do semivariograma indicou corretamente a ausência de dependência espacial entre as firmas do grupo controle e da Categoria 1, visto que seus pontos se localizaram dentro do envelope de significância estatística. Pode-se afirmar que a Categoria 1 está aleatoriamente distribuída, independentemente da distância entre as firmas simuladas.

A semivariância estimada foi de aproximadamente 0,11 e o modelo teórico ajustável foi o “pepita puro” que não possui o parâmetro de alcance e , conseqüentemente, não possibilitou a identificação de *clusters* espaciais. Segundo Scholl e Brenner (2016) a função Kd também rejeitou o padrão de concentração/dispersão espacial de um conjunto semelhante de dados, visto que a mesma se localizou dentro das bandas do intervalo de confiança.

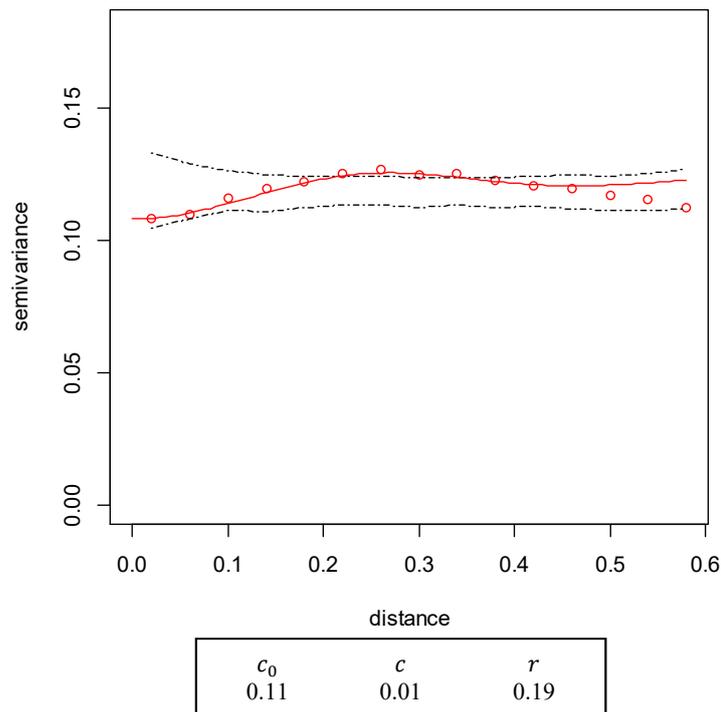
Os semivariogramas direcionais (0° , 45° , 90° e 135°) foram testados e não apresentaram diferenças no padrão de continuidade espacial em relação ao isotrópico. Esse conjunto de resultados, denominado efeito pepita puro, inviabilizou o mapa de krigagem.

3.4.1.2 Cenário 2

O Cenário 2 foi formado pela sobreposição da Categoria 2 ao *benchmark*. Dois modelos teóricos foram ajustados ao semivariograma experimental para fins de comparação:

esférico e *wave*, porém a SEQP indicou um valor menor de erro para o modelo *wave* que foi escolhido para a análise do viés direcional. As medidas de erro foram 67,32 e 41,29, respectivamente. O ponto de corte entre pontos foi de 20% da distância máxima para o variograma omnidirecional e para os direcionais. A Figura 3 apresenta o semivariograma isotrópico da variável indicadora, o ajuste teórico *wave* e o envelope de significância.

Figura 3 - Semivariograma isotrópico, modelo ajustado e o envelope de confiança estatística, Cenário 02.



Fonte: Elaborado pelo autor.

A semivariância observada foi superior ao envelope e confirmou a presença de correlação espacial estatisticamente significativa. Os pontos de 6 a 9 do semivariograma isotrópico foram superiores ao envelope, nas distâncias entre 0,22 km e 0,34 km, respectivamente e após essa distância, o semivariograma retornou ao envelope. Esse resultado foi condizente com a função Kd aplicada à um padrão de pontos semelhantes por Scholl e Brenner (2016) em que a densidade excedeu a banda superior para curtas distâncias, devido ao elevado número de vizinhos dentro dos *clusters*, e retornou para o envelope de confiança estatística em seguida.

Os resultados da função Kd e semelhantes refletem se a contagem de vizinhos de cada ponto, normalizada pelo espaço ou por um *benchmark*, é maior ou menor do que o esperado

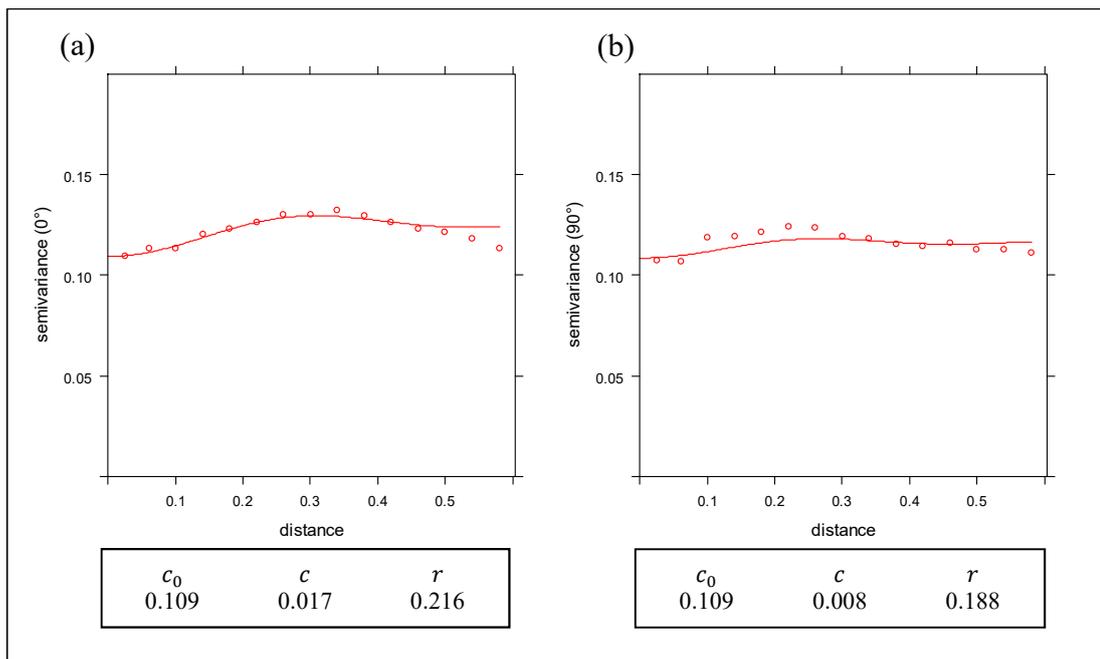
ao acaso, nesse caso a dependência espacial reflete a noção de aglomeração ou dispersão (MARCON; PUECH, 2017). Assim, mesmo com os formatos e resultados semelhantes encontrados entre função Kd e o semivariograma indicador nos cenários simulados (1 e 2), o conceito de dependência espacial é diferente.

O ajuste teórico *wave* do semivariograma isotrópico apresentou um alcance de 0,19 km como era esperado, visto que os *clusters* foram simulados com raio aproximado de 0,1 km. No caso do semivariograma indicador de um fenômeno com a existência de *clusters* espaciais, o alcance corresponde ao tamanho médio destes *clusters* (GOOVAERTS, 2009), ou seja, o diâmetro do *cluster*. No caso da função Kd, o ponto máximo da densidade pode ser interpretado como o próprio raio do *cluster* (MARCON; PUECH, 2010).

O efeito pepita de aproximadamente 0,11 indicou uma considerável heterogeneidade entre a Categoria 2 e o *benchmark* em curtas distâncias e foi aproximadamente o mesmo estimado no Cenário 1. O patamar parcial foi de 0,01.

O modelo *wave* foi ajustado nos semivariogramas da variável indicadora em 2 direções (0° e 90°) para análise da anisotropia. A tolerância utilizada foi de 45° de forma que todos os pontos fossem incluídos na análise. As direções 45° e 135° foram testadas, porém não apresentaram parâmetros consideravelmente diferentes do ajuste isotrópico. Os semivariogramas *wave* ajustados e os respectivos parâmetros se encontram na Figura 4.

Figura 4 - Semivariogramas direcionais (0° e 90°) e modelos ajustados, Cenário 02.

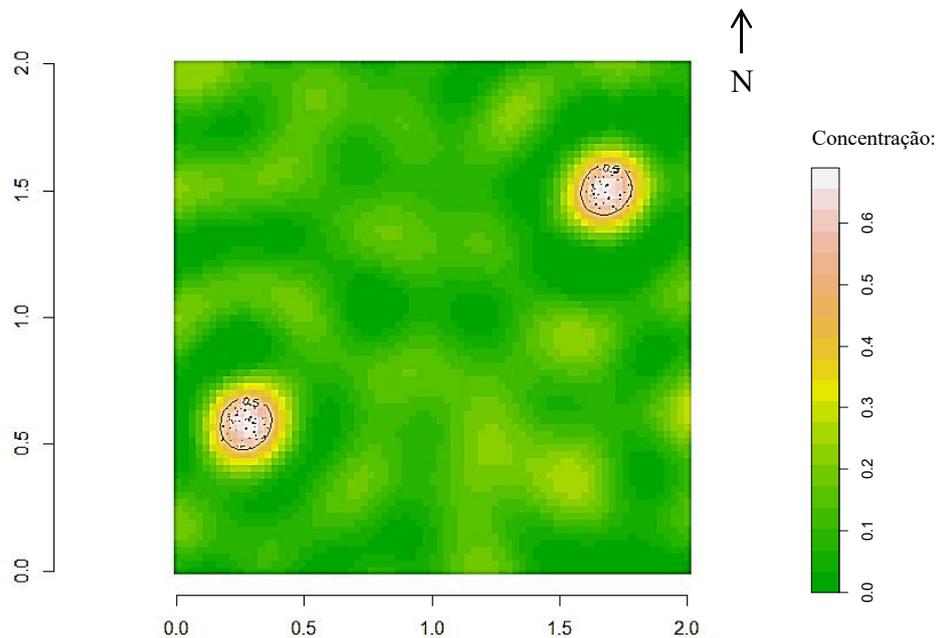


Fonte: Elaborado pelo autor.

Os parâmetros ajustados indicaram alcances distintos nas duas direções, com o maior alcance na direção de 0° (Norte) e o menor a 90° (Leste). Esse padrão espacial de anisotropia geométrica indica que o *cluster* espacial não é circular (GOOVAERTS, 2009). Nesse caso, os *clusters* possuem um formato elíptico com aproximadamente 0,22 km na direção 0° e 0,19 km na direção 90° . O menor patamar na direção de 90° semivariância indicou que os pares de pontos são mais homogêneos a Leste, ou seja, a Categoria 02 é mais conectada.

Para elaboração do mapa de krigagem foi utilizado o semivariograma anisotrópico, com os procedimentos de modelagem de anisotropia mista. A anisotropia zonal foi modelada com um alcance tendendo ao infinito na direção de 90° de forma a nivelar o patamar com a direção de 0° , enquanto a anisotropia geométrica considerou uma razão de 87% (0,188/0,216) entre o menor e o maior alcance. Em seguida, os dois ajustes foram aninhados. O mapa de krigagem (Figura 5) delimitou a superfície da concentração de firmas simuladas.

Figura 5 - Mapa da krigagem indicadora, Cenário 02.



Fonte: Elaborado pelo autor.

Os contornos delimitados na Figura 5 indicam uma probabilidade superior a 50% das firmas pertencerem ao processo Matérn da Categoria 2. Nessa faixa, 92% das firmas simuladas foram identificadas corretamente pelo processo de validação da krigagem e foram classificadas como *clusters cores*. Essas firmas apresentaram alto nível de concentração,

porém as demais firmas no entorno e de mesma categoria também fazem parte do potencial *cluster* geográfico, mesmo com probabilidade inferior a 50%, pois estão dentro do alcance do semivariograma logo, podem ser consideradas geograficamente próximas.

Não foi identificado nenhum caso de “falso positivo” nos contornos acima de 50%, ou seja, firmas com perfil de potencial *cluster*, porém fora da região de elevada concentração.

O Índice Geoestatístico de Concentração (IGC) no Cenário 02 foi de 0,27, valor próximo a 0,29 encontrado pelo índice estimado por Scholl e Brenner (2016) para um padrão simulado equivalente de firmas. Esse valor indica o percentual aproximado de firmas localizadas dentro do *cluster*.

Se utilizado de forma desagregada, o IGC é útil por fornecer uma nova variável de concentração no nível da firma para modelos de regressão que buscam explicar, por exemplo, os determinantes das aglomerações ou do crescimento das firmas (ROSENTHAL; STRANGE, 2001; ALKAY; HEWINGS, 2013). Assim, a abordagem geoestatística pode ser utilizada não somente como uma pesquisa exploratória, mas também como explicativa.

O mapa da krigagem indicadora (Figura 05) apresentou uma interpretação simples em relação a abordagem para dados de área proposta por Guillain e Le Gallo (2010) e facilitou a conciliação entre análise estatística e econômica simultaneamente, visto que uma região é ou não altamente concentrada, dada uma probabilidade como ponto de corte. O mapa de Moran utilizado para a identificação de concentração de atividades econômicas apresenta resultados estatísticos conhecidos como “Alto-Alto”, “Alto-Baixo”, “Baixo-Alto” e “Baixo-Baixo”, no entanto, somente os *clusters* “Alto-Alto” e “Alto-Baixo” possuem interpretação econômica (GUILLAIN; LE GALLO, 2010).

A comparação entre a validação da krigagem utilizando o variograma isotrópico (Figura 3) e anisotrópico (Figura 4) foi apresentada na Tabela 1.

Tabela 1 - Medidas de erro, Cenário 02.

Modelo	SQR	RMP	RRMQP
<i>Wave</i> isotrópico	222	0,00	0,94
<i>Wave</i> anisotrópico	222	0,00	0,94

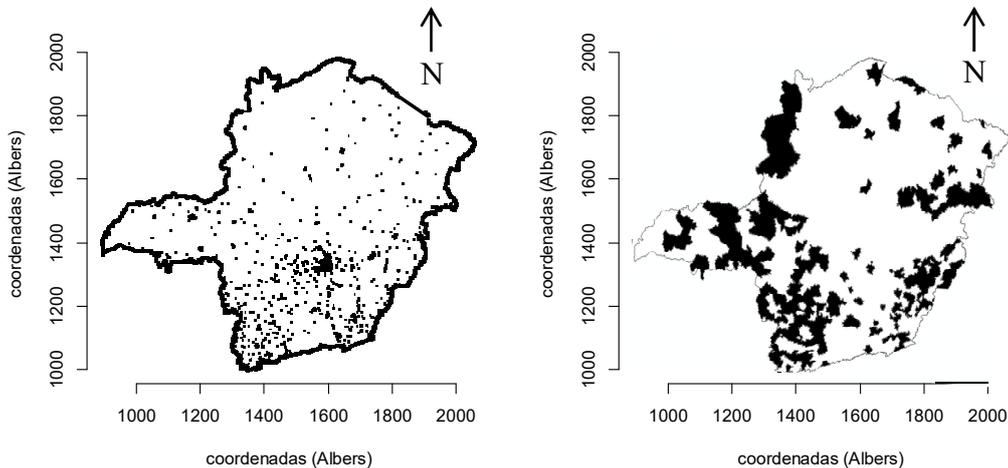
Fonte: Elaborado pelo autor.

Pode-se observar que as medidas de erro foram próximas entre o modelo isotrópico e anisotrópico, apenas diferenças de arredondamento. Os resultados indicaram boas previsões nos dois casos, com o RMP próximo de zero e o RRMQP próximo de 1, no entanto o modelo isotrópico identificou somente 85% das firmas aglomeradas (7% a menos que o anisotrópico).

3.4.2 Análise empírica da indústria do café em Minas Gerais

O mapa à esquerda da Figura 6 indica a localização de 8.533 firmas industriais de Minas Gerais que formaram o grupo controle da análise empírica e as 332 firmas da categoria de torrefadoras de café (mapa à direita da Figura 6) que foram agrupadas em municípios por motivo de confidencialidade dos dados. Visualmente, foi identificado um elevado número de torrefadoras na região Sul de Minas (canto inferior esquerdo), na Zona da Mata (canto inferior direito) e na região metropolitana (centro) que, em geral, aparentam seguir o mesmo padrão de localização do grupo controle.

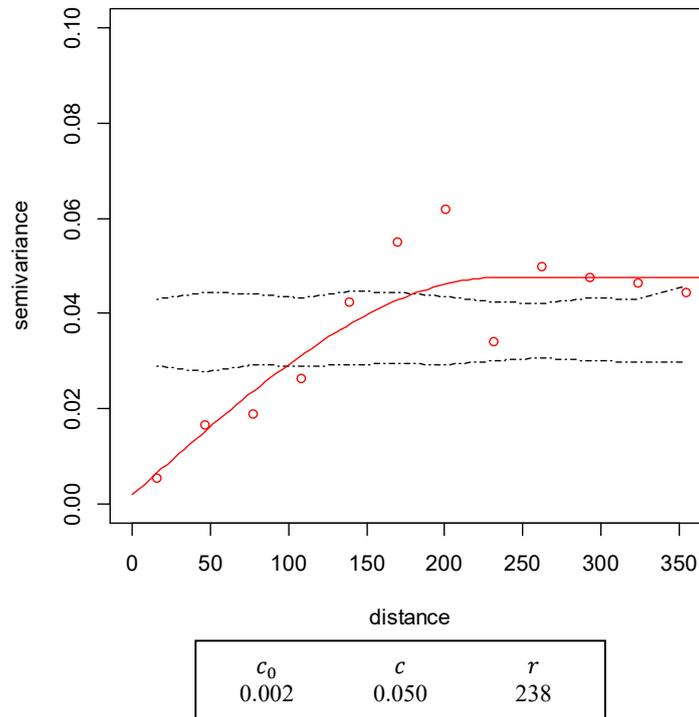
Figura 6 - Distribuição espacial das firmas *benchmark* e das torrefadoras de café.



Fonte: Elaborado pelo autor.

A análise empírica foi conduzida pela sobreposição das firmas de torrefação no *benchmark* formado pelos demais setores industriais da economia de Minas Gerais. Na Figura 7 foi representado o semivariograma isotrópico da variável indicadora, o ajuste esférico e o envelope simulado de significância estatística. O erro (SEQP) dos ajustes *wave* e esférico foram 0,07 e 0,05, respectivamente, logo o ajuste esférico foi utilizado para a análise direcional. O ponto de corte foi de 1/3 da distância máxima entre os pontos, tanto para os variogramas direcionais quanto para o isotrópico.

Figura 7 - Semivariograma isotrópico, modelo ajustado e envelope de confiança estatística, Cenário Empírico.

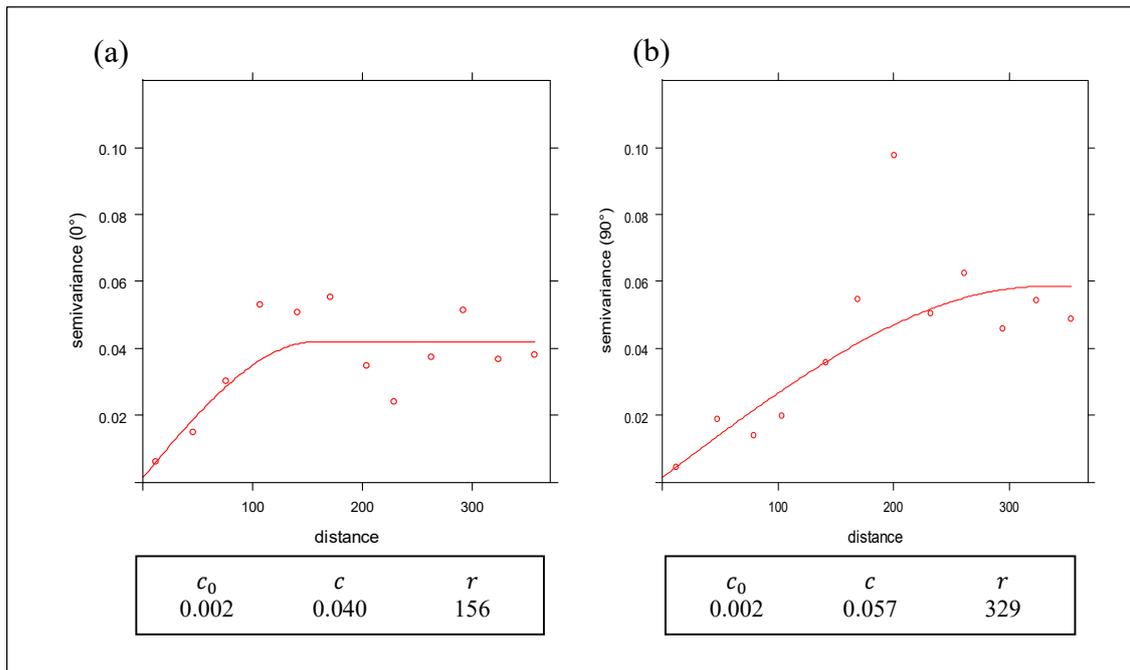


Fonte: Elaborado pelo autor.

Os pontos fora do envelope indicaram uma correlação espacial significativa. O alcance do semivariograma esférico foi de 238 km e confirmou a presença do padrão de *cluster* espacial. O efeito pepita próximo de zero (origem) indicou que o grupo controle também apresentou um padrão aglomerado, ou seja, em distâncias curtas, as firmas *benchmark* estão próximas de outras da mesma categoria, assim como as firmas da categoria torrefação. O patamar parcial foi de 0,05.

O modelo esférico foi ajustado nos semivariogramas de 0° e 90° para análise do viés direcional da variável indicadora (Figuras 8a e 8b, respectivamente) com tolerância de 45° para incluir todas as firmas nas estimativas. As direções 45° e 135° foram testadas, porém não apresentaram estruturas espaciais consideravelmente distintas da isotropia.

Figura 8 - Semivariogramas direcionais (0° e 90°) e modelos ajustados, Cenário Empírico.



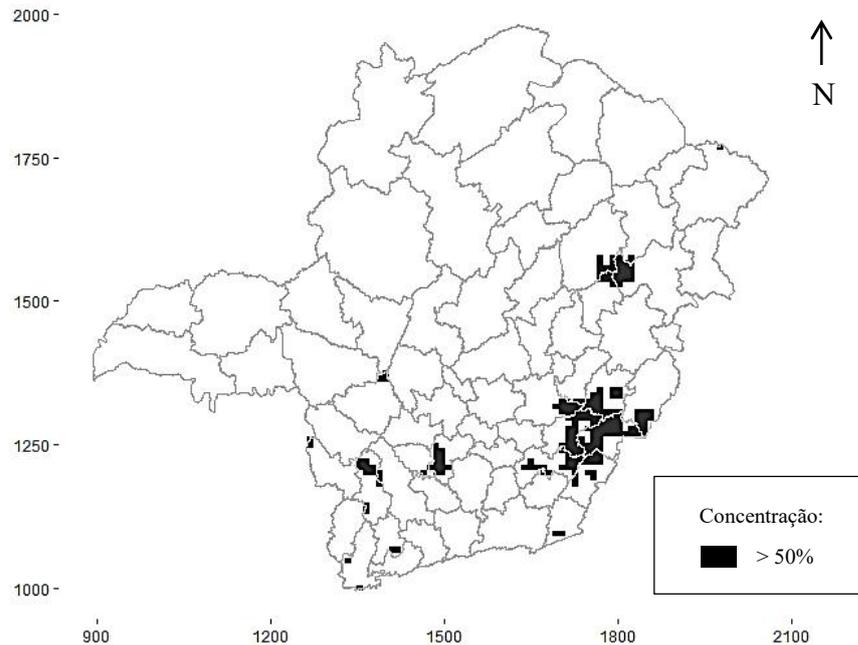
Fonte: Elaborado pelo autor.

A análise dos semivariogramas para indústria do café indicou uma anisotropia geométrica na direção de 90° com um alcance de 329 km contra 156 km na direção de 0° . Esse resultado indicou um *cluster* elíptico de maior alcance na direção Leste. A anisotropia zonal na direção de 0° também foi detectada, com um patamar de 0,04, enquanto na direção perpendicular foi de aproximadamente 0,06. Assim, na direção Norte as firmas são mais homogêneas, logo as torrefadoras são espacialmente mais conectadas. Na direção de 90° foi identificado um *outlier* espacial ao alcance de 200 km que indicou a elevada predominância da categoria de torrefadoras de café sobre o *benchmark*.

Para o mapa de krigagem indicadora foi utilizado o semivariograma anisotrópico. A anisotropia zonal foi modelada na direção de 0° de forma a nivelar o patamar com a direção de 90° , enquanto a razão de 47% ($156/329$) entre o menor e o maior alcance foi utilizada para modelagem da anisotropia geométrica. Os dois semivariogramas ajustados foram aninhados para modelagem da anisotropia mista.

Foi mapeada, na Figura 9, a superfície com probabilidades acima de 50% de firmas pertencerem à categoria de torrefação de café, por meio da krigagem indicadora e do semivariograma anisotrópico. O mapa de krigagem pode ser interpretado como as regiões com perfil de elevada concentração na indústria do café (*cluster core*).

Figura 9 - Mapa de krigagem Indicadora, Cenário Empírico.



Fonte: Elaborado pelo autor.

A maior concentração ocorreu nas microrregiões limítrofes Caratinga, Itabira, Manhuaçu, Ponte Nova e Viçosa. Outra região proeminente foi a da fronteira entre Capelinha e Teófilo Otoni. Em extensões menores, os demais *clusters core* se encontram no Sul de Minas, em especial Alfenas e Oliveira. Esses resultados foram condizentes com o mapeamento de *clusters* da indústria de café em Minas Gerais realizado por Chain et al. (2016), por meio do Índice de Moran Local, em nível municipal, que teve como insumo o Quociente Locacional do valor da produção. Segundo esses autores, as regiões de Caratinga, Manhuaçu, Ponte Nova, Viçosa, Capelinha e, por fim, Varginha (também Sul de Minas e limítrofe a Alfenas) indicaram a presença de concentração estatisticamente significativa

Os resultados semelhantes encontrados entre os mapas na abordagem geoestatística e a análise de dados de área no cenário empírico eram esperados, visto que a formulação da covariância e da semivariância na geoestatística são análogas aos Índices de Moran e de Geary, respectivamente (BIVAND, 1998). Em relação à interpretação dos resultados, essas duas abordagens compartilham o conceito de correlação espacial, ou seja, valores semelhantes estão mais próximos entre si do que era de se esperar ao acaso.

Essas regiões são reconhecidas como grandes produtoras do café *in natura* e ao mesmo tempo, tem no setor de torrefação uma participação do número de empresas acima da média em relação aos demais setores da economia. O elevado número de unidades

geográficas em Minas Gerais (853 municípios) tendeu a aproximar os resultados da análise de área (Índice de Moran) e da abordagem geoestatística. É relevante mencionar que a abordagem geoestatística indicou a concentração do número de firmas, enquanto o Índice de Moran indicou a concentração de produção regional (produção agregada de todas as firmas em cada município), e mesmo assim os resultados foram condizentes entre si.

O IGC para o cenário empírico foi de 0,19 e indicou uma concentração setorial fraca, mesma classificação encontrada por Chain et al. (2016) para a indústria do café em Minas Gerais pelo Índice de Moran Global. O IGC da indústria do café pode ser comparado com o índice da análise simulada que, por sua vez, se apresentou mais concentrada. No nível da firma, o maior IGC foi encontrado em uma torrefadora em Manhuaçu (0,93).

A Tabela 2 apresentou os resultados da validação da krigagem baseada no semivariograma isotrópico (Figura 7) e anisotrópico (Figura 8).

Tabela 2 - Medidas de erro, Cenário Empírico.

Modelo	SQR	RMP	RRMQP
Esférico isotrópico	301	0,00	3,56
Esférico anisotrópico	294	0,00	2,52

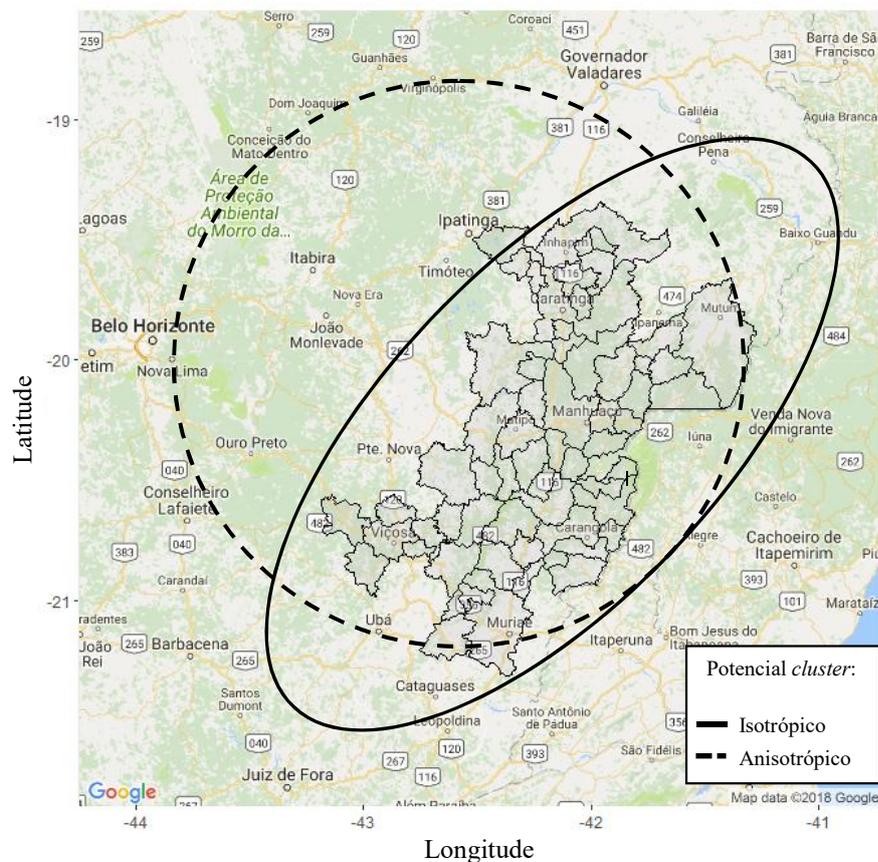
Fonte: Elaborado pelo autor.

Os critérios de erro apresentaram bons ajustes, no entanto a análise anisotrópica gerou estimativas mais robustas, com a SQR menor, a RRMQP mais próxima de 1 e aproximadamente 30% inferior que o RRMQP isotrópico.

A região das Matas de Minas, demarcada na Figura 10, foi selecionada para fins de visualização do viés direcional e da delimitação do potencial *cluster* industrial, pois apresentou o maior número de firmas com elevado IGC em Minas Gerais. A Figura 10 pode então ser interpretada como o mapa de krigagem indicadora no nível da firma, pois mostra a probabilidade de cada firma de pertencer ao *cluster*. Nesse mapa, está visível o direcionamento das probabilidades, interpretadas como o indicador de concentração (IGC).

Segundo Porter (2000), delimitar os limites do *cluster* envolve certo grau de criatividade, assim, para reduzir a subjetividade, os potenciais *clusters* geográficos foram delimitados com base no alcance dos semivariogramas isotrópico (círculo) e direcionais (elipse) para comparação. No caso da elipse, a inclinação a 90° tem uma tolerância de 45° utilizada no cálculo do semivariograma anisotrópico. O alcance do semivariograma definiu a distância que as firmas podem ser consideradas próximas.

Figura 10 – Potencial *cluster* industrial de torrefação de café, Cenário Empírico.



Fonte: Elaborado pelo autor.

A implicação econômica do mapeamento isotrópico em relação ao anisotrópico foi o agrupamento de regiões com perfis distintos. Enquanto o *cluster* anisotrópico alcançou as firmas na região das Matas de Minas, o *cluster* isotrópico alcançou, além da região das Matas de Minas, a região de João Monlevade até Belo Horizonte. O potencial *cluster* anisotrópico foi composto por 78 torrefadoras, sendo 27 o *cluster core* ($IGC > 0,5$) e 51 firmas periféricas. Todas as firmas com IGC acima de 50% na região abrangida pela Figura 10 se localizam nas Matas de Minas. A extensão do *cluster core* foi de aproximadamente 200 km, condizente com o *outlier* espacial no semivariograma anisotrópico da variável indicadora na Figura 8(b).

A região das Matas de Minas é regulamentada pelo Conselho das Entidades das Matas de Minas (CEMM), uma organização sem fins lucrativos de caráter representativo, científico, educacional e cultural, voltada ao desenvolvimento dos produtores de café em 63 municípios (demarcados na Figura 10) que compõem a região. Desde 2010, houve uma intensificação no processo de mobilização dos atores na busca pelo registro da Indicação de Procedência, uma estratégia de reconhecimento da singularidade do café da região, depositado em setembro de 2016 pelo CEMM no Instituto Nacional de Propriedade Intelectual (CARVALHO, 2016).

3.5 CONCLUSÕES

A abordagem geoestatística para identificação de potenciais *clusters* industriais foi condizente com a teoria das aglomerações ao mensurar a proximidade entre firmas, por meio do semivariograma indicador, e estimar a concentração industrial, por meio da krigagem. Além dos pressupostos consolidados pela literatura *mainstream*, foi agregada a análise do viés direcional, a exibição dos resultados em mapas e um indicador de concentração para cada firma.

No caso da indústria do café em Minas Gerais, firmas podem ser consideradas próximas a uma distância de 238 km em termos isotrópicos ou 329 km a leste e 156 km a norte, em termos direcionais. O IGC indicou um nível de concentração setorial fraco, na magnitude de 0,19, com maior potencial de *cluster* localizado na região cafeeira das Matas de Minas. O maior IGC no nível da firma foi encontrado em Manhuaçu (0,93).

Como limitação do estudo coloca-se que não foram investigados os fatores que determinam a concentração da indústria do café em Minas Gerais. Alguns fatores conhecidos na literatura são as economias de urbanização (relacionadas com a infraestrutura da região) e as de localização (associadas à proximidade entre firmas). Assim, novos estudos podem ser conduzidos relacionando os índices de concentração em nível da firma (isotrópicos e anisotrópicos) em função de variáveis que determinam as aglomerações. Estudos futuros podem também analisar outros setores de atividades econômicas com diferentes padrões de localização das firmas.

Somente por meio de dados secundários e métodos quantitativos não é possível afirmar que o potencial *cluster* anisotrópico é viável e o isotrópico é inviável em termos de política industrial. Deve-se considerar que o modelo anisotrópico descreveu melhor a realidade dos dados com menor erro de predição (até 7% no cenário simulado e até 30% no cenário empírico) e agrupou firmas inseridas em uma região de perfil histórico, social e cultural semelhante que tende a facilitar a coordenação do *cluster*. Assim, a variável “direção” pode ser considerada um determinante da teoria das aglomerações nos casos analisados.

A limitação da abordagem geoestatística está relacionada com a identificação do padrão de dispersão (regularidade) entre firmas que é viabilizada pela função Kd e derivadas. Nesse sentido, a estatística de processos pontuais pode complementar o semivariograma se o objetivo da pesquisa é analisar o padrão de distribuição geográfica das firmas, mas caso o

objetivo seja mapear potenciais *clusters* industriais, o semivariograma e a krigagem são suficientes.

Para pesquisas futuras, recomenda-se a inclusão do tamanho das firmas como, por exemplo, o número de empregos no eixo de profundidade (3ª dimensão). Outra possibilidade analítica é a de se trabalhar com indústrias correlatas, por meio da variograma cruzado e da cokrigagem. No caso da indústria de torrefação do café, pode-se trabalhar com o mapeamento de produtores rurais, cafeterias e supermercados. Para maior robustez dos resultados também é interessante trabalhar com geoestatística espaço-temporal, novos testes estatísticos para anisotropia e dependência espacial.

REFERÊNCIAS

- AGUILAR, F. X.; VLOSKY, R. P. Spatial analysis of forests products manufacturer clusters in Louisiana. **Wood and Fiber Science**, v. 38, n. 1, p. 121-131, 2007.
- ALKAY, E.; HEWINGS, G. J. D. The determinants of agglomeration for the manufacturing sector in the Istanbul metropolitan area. **The Annals of Regional Science**, v. 48, n. 1, p. 225-245, 2012.
- ARBIA, G.; BEE, M.; ESPA, G. Testing isotropy in spatial econometric models. **Spatial Economic Analysis**, v. 8, n. 3, p. 228-240, 2013.
- ARBIA, G.; DI MARCANTONIO, M. Forecasting interest rates using geostatistical techniques. **Econometrics**, v. 3, n. 4, p. 733-760, 2015.
- ASSOCIAÇÃO BRASILEIRA DA INDÚSTRIA DO CAFÉ. **Indicadores da indústria de café no Brasil**. Rio de Janeiro: ABIC, 2014. Disponível em: <<http://www.abic.com.br>>. Acesso em: 10 jul. 2016.
- BADDELEY, A.; TURNER, R. Spatstat: an R package for analyzing spatial point patterns. **Journal of Statistical Software**, v. 12, n. 6, p. 1-42, 2005.
- BIERKENS, M. F. P.; BURROUGH, P. A. The indicator approach to categorical soil data. **European Journal of Soil Science**, v. 44, n. 2, p. 361-368, 1993.
- BIVAND, R. S. **A review of spatial statistical techniques for location studies**. Bergen: Norwegian School of Economics and Business Administration. 1998
- BRENE, P. R. A. et al. Mudança estrutural e dualismo no Brasil: Uma análise insumo-produto para os anos 2000 e 2007. **Revista Capital Científico**, v. 12, n. 4, p. 29-46, 2014.
- CARROLL, M. C.; REID, N.; SMITH, B. W. Location quotients versus spatial autocorrelation in identifying potential cluster regions. **The Annals of Regional Science**, v. 42, n. 2, p. 449-463, 2008.
- CARVALHO, J. P. **Campo organizacional e adoção de Indicação Geográfica (IG): um estudo sobre a produção de cafés especiais no Brasil**. Tese de Doutorado. Universidade de São Paulo. 2016.
- CENTRO INDUSTRIAL EMPRESARIAL DE MINAS GERAIS. **Cadastro Industrial de Minas Gerais**. Belo Horizonte: CIEMG, 2014.
- CHAIN, C. P. et al. Concentração espacial na indústria do café em Minas Gerais. **Organizações Rurais & Agroindustriais**, v. 18, n. 2, p. 111-124, 2016.
- CHAIN, C. P. et al. Revisão bibliométrica dos métodos quantitativos aplicados à mensuração de *clusters* industriais. In: **XV Encontro Nacional da Associação Brasileira de Estudos Regionais e Urbanos**, São Paulo, 2017.

- CHICA OLMO, J. Spatial estimation of housing prices and locational rents. **Urban Studies**, v. 32, n. 8, p. 1331-1344, 1995.
- COMPANHIA NACIONAL DE ABASTECIMENTO. **Safras: séries históricas**. Brasília: CONAB, 2013. Disponível em: <<http://conab.gov.br/site/>>. Acesso: 28 jun. 2016.
- CRESSIE, N. A. C. **Statistics for Spatial Data**. New York: Wiley & Sons, 1993.
- CRUZ, S. C. S.; TEIXEIRA, A. A. C. The evolution of the cluster literature: shedding light on the regional studies–regional science debate. **Regional Studies**, v. 44, n. 9, p. 1263-1288, 2010.
- DELGADO, M.; PORTER, M. E.; STERN, S. Defining clusters of related industries. **Journal of Economic Geography**, v. 16, n. 1, p. 1-38, 2016.
- DIGGLE, P. J.; RIBEIRO, P. J. **Model-based Geostatistics**. Springer, 2007.
- DURANTON, G.; OVERMAN, H. G. Testing for localization using micro-geographic data. **The Review of Economic Studies**, v. 72, n. 4, p. 1077-1106, 2005.
- ERIKSSON, M.; SISKKA, P. P. Understanding anisotropy computations. **Mathematical Geology**, v. 32, n. 6, p. 683-700, 2000.
- FERNÁNDEZ-AVILÉS, G.; MONTERO, J. M.; ORLOV, A. G. Spatial modeling of stock market comovements. **Finance Research Letters**, v. 9, n. 4, p. 202-212, 2012.
- FERRAZ, J. C.; DE PAULA, G. M.; KUPFER, D. Política Industrial. In: KUPFER, D; HASENCLEVER, L. (Org.). **Economia Industrial: Fundamentos Teóricos e Prática no Brasil**. 2ª Ed. Rio de Janeiro: Campus, 2013, p. 313-323.
- FESER, E.; ISSERMAN, A. The rural role in national value chains. **Regional Studies**, v. 43, n. 1, p. 89-109, 2009.
- FESER, E. J.; SWEENEY, S. H. A test for the coincident economic and spatial clustering of business enterprises. **Journal of Geographical Systems**, v. 2, n. 4, p. 349-373, 2000.
- GILLEN, K.; THIBODEAU, T.; WACHTER, S. Anisotropic autocorrelation in house prices. **The Journal of Real Estate Finance and Economics**, v. 23, n. 1, p. 5-30, 2001.
- GOOVAERTS, P. Geostatistics in soil science: state-of-the-art and perspectives. **Geoderma**, v. 89, n. 1, p. 1-45, 1999.
- GOOVAERTS, P. Combining area-based and individual-level data in the geostatistical mapping of late-stage cancer incidence. **Spatial and Spatio-temporal Epidemiology**, v. 1, n. 1, p. 61-71, 2009.
- GUILLAIN, R.; LE GALLO, J. Agglomeration and dispersion of economic activities in and around Paris: an exploratory spatial data analysis. **Environment and Planning B: Planning and Design**, v. 37, n. 6, p. 961-981, 2010.

HENGL, T.; MINASNY, B.; GOULD, M. A geostatistical analysis of geostatistics. **Scientometrics**, v. 80, n. 2, p. 491-514, 2009.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Pesquisa Industrial Anual**: Produto. Brasília: IBGE, 2013.

KRUGMAN, P. **Geography and Trade**. Cambridge, MA: MIT Press, 1991.

LEUANGTHONG, O.; KHAN, K. D.; DEUTSCH, C. V. **Solved problems in geostatistics**. John Wiley & Sons, 2011.

MARCON, E.; PUECH, F. Measures of the geographic concentration of industries: improving distance-based methods. **Journal of Economic Geography**, v. 10, n. 5, p. 745-762, 2010.

MARCON, E.; PUECH, F. A typology of distance-based measures of spatial concentration. **Regional Science and Urban Economics**, v. 62, p. 56-67, 2017.

MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA E COMÉRCIO EXTERIOR. **Plano Brasil Maior**: acompanhamento das medidas setoriais. Brasília: MDIC, 2013.

Disponível em:

<<http://www.brasilmaior.mdic.gov.br/images/data/201310/1b9c7442c23cf1e63033ac26f41ae903.pdf>>. Acesso em: 03 out. 16.

MINISTÉRIO DO TRABALHO E EMPREGO. **Relação Anual de Informações Sociais**.

Brasília: MTE, 2014. Disponível em: <<http://bi.mte.gov.br/bgcaged/rais.php>>. Acesso em: 23 nov. 2016.

MATHERON, G. Principles of geostatistics. **Economic Geology**, v. 58, n. 8, p. 1246-1266, 1963.

OLIVER, M. A.; WEBSTER, R. A tutorial guide to geostatistics: Computing and modelling variograms and kriging. **Catena**, v. 113, p. 56-69, 2014.

OPENSHAW S. **The Modifiable Areal Unit Problem (Concepts and techniques in modern geography)**. Norwich: Geo Books, 1984.

PEBESMA, E. J. Multivariable geostatistics in S: the gstat package. **Computers & Geosciences**, v. 30, n. 7, p. 683-691, 2004.

PORTER, M. E. *Clusters and competition*. in Gordon E. Clark, eds., **Oxford Handbook of Economic Geography**. Oxford: Oxford University Press, 2000, pp. 197-2

RIBEIRO JR, P. J.; DIGGLE, P. J. geoR: a package for geostatistical analysis. **R news**, v. 1, n. 2, p. 14-18, 2001.

ROSENTHAL, S. S.; STRANGE, W. C. The determinants of agglomeration. **Journal of Urban Economics**, v. 50, n. 2, p. 191-229, 2001.

RYBNIKOVA, N. A.; PORTNOV, B. A. Using light-at-night (LAN) satellite data for identifying clusters of economic activities in Europe. **Letters in Spatial and Resource Sciences**, v. 8, n. 3, p. 307-334, 2015.

SAKON, F. M. et al. A produção e industrialização do café no Brasil e a economia nacional: simulações sob um modelo inter-regional de insumo-produto. **Revista Econômica do Nordeste**, v. 43, n. 3, p. 443-462, 2012.

SCHOLL, T.; BRENNER, T. Detecting spatial clustering using a firm-level cluster index. **Regional Studies**, v. 50, n. 6, p. 1054-1068, 2016.

TAM CHO, W. K.; GIMPEL, J. G. Prospecting for (campaign) gold. **American Journal of Political Science**, v. 51, n. 2, p. 255-268, 2007.

VIEIRA, S. R. et al. Geostatistical theory and application to variability of some agronomical properties. **California Agriculture**, v. 51, n. 3, p. 1-75, 1983.

WANG, H.; ZHANG, H. On the possibility of a private crop insurance market: A spatial statistics approach. **Journal of Risk and Insurance**, v. 70, n. 1, p. 111-124, 2003.