

Recebido em 05/10/2018. Aprovado em 13/01/2019. Avaliado pelo sistema *double blind peer review*. Publicado conforme normas da ABNT.
<http://dx.doi.org/10.22279/navus.2019.v9n2.p173-189.849>

Análise do arranjo empresarial de uma cooperativa de apicultores da região de Sorocaba-SP pelo método de *clustering*

Sandro Eduardo de Souza Mestre em Processos Tecnológicos e Ambientais. Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (IFSP) - Brasil. sandro.eduardo@ifsp.edu.br
Waldemar Bonventi Jr. Doutor em Engenharia Elétrica. Universidade de Sorocaba (UNISO) - Brasil. waldemar.bonventi@prof.uniso.br
Daniel Bertoli Gonçalves Doutor em Engenharia de Produção. Processos Tecnológicos e Ambientais da Universidade de Sorocaba (UNISO) - Brasil. daniel.goncalves@prof.uniso.br

RESUMO

As escolhas de localização são extremamente importantes para as organizações produtivas, por trazerem impactos significativos sobre sua cadeia de valor e sobre os custos operacionais das empresas. A análise de agrupamentos ou *clustering* é uma ferramenta matemática com capacidade de oferecer informações estratégicas para este tipo de problema. Neste sentido, o presente estudo buscou aplicar a análise de agrupamentos para definir oportunidades de redução de custos na rede de distribuição de mel de uma Cooperativa de Apicultores de Sorocaba-SP, da qual participam 357 cooperados distribuídos em 75 municípios, em um raio de atuação de mais de 200 km. A área abrangida pelos municípios e suas localizações foi dividida em grupos, onde foram analisadas distâncias e custos de transportes, através de algoritmo de agrupamentos. Os resultados apontaram que a redução de custos poderá ser obtida mediante instalação de centros de distribuição em pontos estratégicos da área de abrangência da cooperativa.

Palavras chave: Arranjo produtivo. Localização. Agrupamentos. *Clustering*. Mel.

Analysis of the enterprise arrangement of a beekeepers cooperative of the region of Sorocaba-SP by the clustering method

ABSTRACT

Location choices are extremely important for productive organizations because they have significant impacts on their value chain and on the operational costs of companies. Clustering or clustering analysis is a mathematical tool capable of providing strategic information for this type of problem. In this sense, the present study sought to apply cluster analysis to define opportunities for cost reduction in the honey distribution network of a Beekeepers Cooperative of Sorocaba-SP, with the participation of 357 members distributed in 75 municipalities within a radius of operation of more than 200 km. The area covered by the municipalities and their locations was divided into groups, where distances and transport costs were analyzed through clustering algorithms. The results showed that cost reduction can be achieved through the installation of distribution centers at strategic points in the area covered by the cooperative.

Key words: Productive arrangement. Location. Groupings. Clustering. Honey.

1 INTRODUÇÃO

A região de Sorocaba, no interior do estado de São Paulo, abriga uma das mais importantes cooperativas de pequenos produtores rurais do estado, a Cooperativa de Apicultores de Sorocaba e Região, COAPIS, que conta com a participação de 357 cooperados distribuídos em 75 municípios, em um raio de atuação de mais de 200 km, uma característica peculiar que traz à tona desafios logísticos e econômicos, comuns para estas organizações, com grandes oportunidades para melhorias (SOUZA, 2017).

Na ótica do cooperativismo agrícola, Brito (2011), FAO (2012), e Beserra (2013), reforçam a importância das cooperativas no desenvolvimento regional nacional, bem como podem ser instrumento de redução do êxodo rural e da pobreza nos municípios estagnados e de baixa renda no Brasil. A FAO (2012) determina as políticas das Nações Unidas para erradicação da fome e miséria alicerçada no trabalho cooperativo e no fortalecimento da associação de produtores de zonas rurais.

Este trabalho parte da hipótese de que a otimização da rede logística (posicionamento e simulação de entrepostos de produção) permitirá a redução de custos, estoques e de prazos, bem como incremento na escala de produção de mel e hortifrútiis. A problemática apresentada pela direção da cooperativa, direcionou o estudo para a questão dos custos de frete na distribuição de mel, mas também, em como buscar formas de alavancar a comercialização dos produtos agrícolas de seus cooperados através da COAPIS, almejando a geração de renda e manutenção de empregos nas propriedades rurais.

O arranjo empresarial da cooperativa COAPIS, discutido por Souza (2017), sugere uma análise mais estruturada de sua configuração devido a sua dispersão geográfica em 75 municípios. Nesse contexto podem ser adotados diversos tipos de arranjos para a produção solidária de mel, como Arranjo Produtivo Local e *Cluster*, sendo que ambos diferem estruturalmente da cooperativa (MARTINS; SILVA, 2015). Parte-se assim de uma avaliação considerando a COAPIS como uma cadeia produtiva mais complexa, integrando produtores de mel ao mercado paulista, nacional e internacional.

Dados da COAPIS informam que do total de 75 municípios que compreendem os 357 cooperados, Sorocaba tem 97 apicultores, seguido de Votorantim com 16, Araçoiaba da Serra com 6, Itatinga com 6, Botucatu com 5, Capão Bonito com 5, Itu com 5, Tapiraí com 4 e Piedade com 3 apicultores. Nas cidades listadas, encontram-se 41,2% do total de cooperados permitindo nova análise pela sua representatividade para a proposta de instalação de centros distribuidores.

Localização de instalação é o processo que determina os locais geográficos para as operações de uma empresa. As escolhas de localização podem ser extremamente importantes e tem um impacto profundo sobre a cadeia de valor da empresa. [...] a localização tem um impacto significativo sobre os custos operacionais da companhia, os preços que ela cobra por serviços e bens e sua capacidade de competir no mercado e de penetrar em novos segmentos de clientes (RITZMAN; KRAJEWSKI; MALHOTRA, 2009, p. 350).

Para a realização de um processo sistemático de localização de uma unidade de distribuição, inicialmente deve-se evidenciar a melhoria no desempenho global da organização com uma nova planta. A comparação de novos locais para a instalação de uma unidade também deve conter dados relativos as alternativas de uma localização que justifiquem qualitativa e quantitativamente os fatores de localização para a melhoria do canal de distribuição para a organização (RITZMAN; KRAJEWSKI; MALHOTRA, 2009, p. 360). Conforme Ballou (2006), tratar das decisões de localização das instalações e suas estratégias envolve:

Localizar instalações fixas ao longo da rede da cadeia de suprimentos é um importante problema de decisão que dá forma, estrutura e contornos ao conjunto completo dessa cadeia. Essa formulação define alternativas, juntamente com os custos e níveis de investimentos a elas associadas, usadas para operar o sistema. Decisões sobre localização envolvem a determinação do número, local e proporções das instalações a serem usadas (BALLOU, 2006, p. 433).

Nesse contexto, foi objetivo do presente trabalho definir oportunidades de redução de custos na rede de distribuição de mel da Cooperativa de Apicultores de Sorocaba e Região – COAPIS por meio de análise de

agrupamentos (*clustering*) (JAIN, MURTY e FLYNN, 1999) a fim de propor uma configuração de centros de distribuição que satisfaça dois critérios: primeiro, dada a localização dos produtores no Estado de São Paulo, como agrupá-los de modo a ter um número de centros de distribuição ótimo e, segundo, definido o número de grupos de produtores geograficamente próximos entre si, qual a melhor localização destes centros de distribuição.

Nesta pesquisa foi adotado o método do estudo de caso (YIN, 2015), para observação do cenário tão particular da COAPIS. Não apenas sua abrangência e dispersão geográfica demandaram uma análise mais detalhada, mas também as problemáticas envolvidas do ponto de vista da apicultura, logística e organização.

Para a aplicação do estudo, foi definido o delineamento observacional através de coleta de dados sobre produção, demandas e características da cadeia produtiva da cooperativa diretamente com sua gerência administrativa, por meio de uma entrevista estruturada.

De acordo com Lakatos e Marconi (2005), esta técnica permite estabelecer uma técnica de coleta de dados para conseguir informações relevantes através do exame de fatos e fenômenos estudados.

A entrevista é um encontro entre duas pessoas, a fim de que uma delas obtenha informações a respeito de determinado assunto, mediante uma conversação de natureza profissional. É um procedimento utilizado na investigação social, para a coleta de dados ou para ajudar no diagnóstico ou no tratamento de um problema social. (LAKATOS; MARCONI, 2005, p. 203).

Para a estruturação do levantamento de dados, posteriormente realizou-se a observação direta extensiva com a aplicação de questionário enviado por correio eletrônico para a gerência da empresa para identificação de oportunidades na pesquisa.

Paralelamente às observações supracitadas, foi fundamental a realização de uma pesquisa bibliográfica não apenas para embasamento teórico, mas para nortear e justificar as propostas futuras à COAPIS. Tal modalidade de pesquisa, de acordo com os mesmos autores, tem por finalidade colocar o pesquisador em contato direto com tudo o que foi escrito, dito ou filmado sobre determinado assunto. De acordo com Manzo (1971, p. 32), a bibliografia “oferece meios para definir, resolver, não somente problemas já conhecidos, como também explorar novas áreas onde os problemas não se cristalizaram suficientemente”.

O presente estudo está estruturado partindo da identificação dos tipos de arranjos empresariais, suas características e dos arranjos aplicados à produção de mel, seguido dos pilares que norteiam o conceito de cooperativas, sua origem na agricultura rural familiar brasileira e do cooperativismo como modelo de apoio aos pequenos produtores, abordando, em seguida a metodologia de simulação matemática empregada, seguida dos resultados e considerações finais do estudo.

2 FUNDAMENTAÇÃO TEÓRICA

A fundamentação teórica faz uma abordagem de três tópicos principais: os arranjos empresariais, o Arranjo Produtivo Local (APL) e as cooperativas.

2.1 Arranjos Empresariais

Existem vários termos utilizados na literatura para identificar os diversos tipos de Arranjos Empresariais, tais como *Clusters*, Distritos Industriais, Arranjos Produtivos Locais (APL), Complexos Industriais, Redes de Empresas e Sistemas Produtivos e Inovativos Locais (SPIL), mas comumente estes são interpretados como sinônimos. No Brasil, o termo amplamente utilizado é APL, inclusive pelo próprio Governo Federal (MARTINS; SILVA, 2015).

Há várias lógicas que exploram as definições, características e aplicações conceituais dos tipos de Arranjos e que dificultam o entendimento de aspectos que possam distinguir de maneira clara um Arranjo de outro. Logo, distinguir esses relacionamentos entre organizações não é tarefa fácil, uma vez que determinadas características podem ser

comuns a alguns tipos de Arranjos, entretanto, a visualização de diferenças estruturais parece ser distintiva em alguns formatos. (GONÇALVES; LEITE; SILVA, 2012, p. 829).

Em cada caso há características distintas em termos de história, evolução, organização institucional, contextos sociais e culturais, da forma de organização da produção e da forma de governança local. Nesse sentido, é possível encontrar diferenças significativas entre muitas das diversas experiências no que se refere a graus de desenvolvimento, de articulação e interação entre agentes e instituições locais e de capacidades sistêmicas para a inovação (SUZIGAN, 2006).

Partindo destas considerações, a seguir apresentam-se os principais tipos de arranjos empresariais e, posteriormente, uma breve comparação entre suas estruturas de gestão como base para as análises de dados iniciais propostas pela presente pesquisa.

O termo Arranjo Produtivo Local (APL) é comumente utilizado para aglomerações ou concentrações territoriais e setoriais de agentes econômicos, políticos e sociais em torno de uma atividade econômica específica, nas quais se estruturam vínculos e relações de interação, interdependência, cooperação e aprendizagem, voltadas para a capacitação contínua, essencial para geração de competitividade e sustentabilidade dos seus membros, como também para a promoção do dinamismo econômico local da região em que o APL está inserido, diminuindo as diferenças regionais (MARTINS; SILVA, 2015).

Cluster ou Agrupamento denomina uma aglomeração geograficamente concentrada de organizações com características similares, que trabalham direta ou indiretamente para o mesmo mercado final, com relações verticais e horizontais, mostrando uma clara tendência de cooperação e de compartilhamento de competências, valores e conhecimentos entre os seus membros, com ênfase na concorrência como fator de dinamismo (MARTINS; SILVA, 2015).

Rede de Empresas são formatos organizacionais, definidos a partir de um conjunto de articulações entre entidades empresariais independentes e geograficamente dispersas, que atuam através de ações de coordenação, interação e cooperação, baseadas ou não em contratos formais, compartilhando recursos, pessoas, tecnologias, informações, conhecimentos, sob uma única forma de atuação e uma mesma estratégia, com o objetivo de obter maior capacidade competitiva para lidar com a complexidade do atual ambiente de negócios (GONÇALVES; LEITE; SILVA, 2012).

Já o termo Cooperativa denomina uma organização de pessoas que se reúnem voluntariamente, em igualdade de direitos, com o objetivo comum de desenvolver uma atividade econômica ou prestar serviços, a partir da criação de uma empresa de propriedade coletiva gerida democraticamente, como parte integrante das empresas individuais de cada membro. Apresentam particularidades em sua constituição legal, com a adoção de princípios e doutrinas com reflexos na sua estruturação, delegação e exercício de poder e, conseqüentemente, na forma como ocorre o processo decisório (BRASIL, 2006).

Cadeia de suprimentos ou *Supply Chain*, consiste em uma rede de organizações que mantêm relações mútuas, agregando valor aos produtos e/ou serviços, desde os fornecedores de matéria-prima, passando pelos distribuidores, até os consumidores finais, incluindo também, eventualmente, os movimentos de retorno de produtos não consumidos ou descartados. Consiste em uma sequência de processos e fluxos que acontecem dentro e entre estágios da cadeia, em um formato contínuo e sistêmico, no qual o princípio básico está voltado à integração de informações entre seus membros (BALLOU, 2006).

Cadeia Produtiva consiste em um conjunto de operações sucessivas responsáveis por transformações dissociáveis desde a obtenção de matéria-prima até a comercialização dos respectivos produtos finais. Cada membro da cadeia é responsável pela realização de diferentes etapas do processo e podem estar localizados em regiões ou localidades distintas (GONÇALVES; LEITE; SILVA, 2012).

Condomínio Industrial consiste em uma infraestrutura que se configura quando empresas montadoras reúnem, ao redor de sua fábrica, seus principais fornecedores diretos, chamados de *sistemistas*, os quais abastecem a montadora diretamente ao lado da linha de montagem, mas não participam da linha de montagem final do produto, a qual permanece a cargo da montadora (VENANZI, 2009).

Consórcio Modular consiste em um caso de terceirização radical entre uma montadora e um pequeno número de fornecedores diretos, chamados de *modulistas*, que ficam instalados dentro da planta da própria

montadora. Neste tipo de Arranjo, os fornecedores são responsáveis pela montagem prévia do módulo e pela montagem final na linha de produção da montadora (VENANZI, 2009).

Dentre os tipos de Arranjos em análise, Clusters e APL são definidos com conceitos similares; na verdade, o principal aspecto que os diferencia diante das definições abordadas, é que o APL se trata de um tipo específico de Cluster, que abrange Pequenas e Médias Empresas (GONÇALVES; LEITE; SILVA, 2012).

Quanto às Cooperativas, comparando-as aos Arranjos supracitados, conclui-se que elas apresentam peculiaridades em sua constituição legal, com a adoção de princípios e doutrinas que exercem significativos reflexos na sua estruturação e governança, e na forma como ocorre o processo decisório. Consiste em uma organização em que os seus membros concordam entre si para criar uma empresa e operá-la em conjunto como parte integrante de suas empresas individuais (FENG; HENDRIKSE, 2007), o que não ocorre nos casos de Cluster, APL e Redes de Empresas.

A Cadeia de Suprimentos é conceituada como um conjunto ou rede de organizações que trabalham juntas a fim de, necessariamente, agregar valor ao produto e/ou serviço. Batalha (2007) afirma que Cadeia Produtiva seria um conjunto sucessivo de etapas ou operações ou atividades de transformação da matéria-prima em produto acabado, em que cada membro é responsável pela realização de diferentes etapas do processo.

Desconsiderou-se para fins do presente artigo, os arranjos empresariais relativos a Condomínio Industrial e Consórcio Modular por suas particularidades e constante aplicação em indústrias, principalmente do setor automotivo ou montadoras de veículos.

2.2 Arranjo Produtivo Local

As vantagens observadas em sistemas locais de produção foram pioneiramente apontadas por Marshall (1920), a partir de experiências e estudos realizados nos distritos industriais ingleses no século XIX. Marshall apontou três forças principais que explicam a formação de aglomerações: primeiro, a existência concentrada de mão de obra qualificada com habilidades específicas ao setor; segundo, a presença de fornecedores especializados que garantem baixos custos e fácil acesso a produtos e serviços; e, em terceiro lugar, a ocorrência de transbordamentos de conhecimentos, decorrentes da rápida difusão de conhecimentos decorrente da proximidade entre os agentes.

Nos sistemas locais de produção, a forte especialização dos agentes econômicos permite que ocorra uma densa divisão do trabalho entre as firmas que, por sua vez, possibilita a interação frequente entre os atores locais a partir de relações comerciais e produtivas. Esses processos acarretam o aumento da produtividade das empresas e do sistema como um todo, além de ganhos obtidos na esfera organizacional. Com isso, as economias externas são geradas e intensificadas (GARCIA; MADEIRA, 2013).

Conforme destaca Cabete e Dacol (2008, p. 10) o que caracteriza uma aglomeração industrial para **que seja considerada um APL são: “mesma localização geográfica, interdependência e cooperação, e em consequência destes processos há ainda as características competitividade, difusão do conhecimento, inovação e confiança”**.

Para Lastres e Cassiolato (2005), o conceito de aglomeração tornou-se mais estruturado através dos diversos estudos realizados. Para as autoras um marco para este entendimento foi a relação da ideia de aglomeração com a de redes, particularmente observando cadeias produtivas ao redor de uma empresa **“âncora”, ou nos moldes cooperativistas, através da governança de uma cooperativa central**.

O formato de rede, onde podemos inserir a COAPIS, é o caso de aglomerações de Micro e Pequenas Empresas com governança em forma de redes, que pode ser caracterizado a partir dos chamados distritos **industriais italianos (LASTRES; CASSIOLATO, 2005). Estes são estruturados em torno de “setores” como calçados, vestuário, mel etc.** onde a especialização das aglomerações é predominantemente setorial e voltada para resultados comuns de participação em mercados mais competitivos aos membros da cooperativa.

Com o objetivo de melhoria do desempenho das Pequenas e Médias Empresas (PMEs) participantes do APL ou Rede de Empresa, destaca-se **que “a cooperação é um instrumento que permite incorporar as PMEs em uma rede, promover o ciclo de inovação tecnológica e determina os resultados do processo de melhoria de um aglomerado de PMEs” (GALDÁMEZ; CARPINETTI; GEROLAMO, 2009, p. 137).**

Observando aspectos relativos a desempenho do APL, Schmitz e Nadvi (1999, p. 1.504) definem como **eficiência coletiva “a vantagem competitiva derivada das economias externas e da ação conjunta”**. Neste contexto, a eficiência coletiva também é destacada por Keller (2008), podendo ser passiva ou ativa. A **eficiência coletiva passiva seria resultado intrínseco da participação no aglomerado e “a eficiência coletiva ativa uma vantagem buscada de forma deliberada por meio da ação conjunta, aquela que requer esforços conjuntos dos agentes econômicos” (KELLER, 2008, p. 36).**

A resposta adequada dos clusters industriais nos países em desenvolvimento aos desafios colocados pela globalização econômica pressupõe a eficiência coletiva e a criação de formas de governança local da cadeia em um trabalho conjunto, tanto dos agentes econômicos que integram a cadeia do produto e suas entidades representativas quanto das instituições de apoio. Esta resposta adequada dos clusters industriais (caminhando na direção de um cluster virtuoso) emerge somente em um meio social (cultural e institucional) mais forte que envolve maior confiança entre os agentes econômicos e a presença de instituições sociais consolidadas. (KELLER, 2008, p. 44).

Independente da nomenclatura adotada como referência para ações entre empresas participantes de um aglomerado com objetivos comuns e de alguma forma estruturada para obtenção de vantagens competitivas, seu modelo de governança, dispersão geográfica, territorialização e abrangência de mercado irão influenciar em seus resultados (LASTRES; CASSIOLATO, 2005).

Devida à característica intrínseca da Cooperativa de Apicultores de Sorocaba e Região, desde a sua criação, trataremos dos fundamentos que caracterizam as cooperativas e sua estrutura, independente de práticas em regiões produtoras de mel com o Arranjo Produtivo Local, *Cluster* ou Rede de Empresas.

2.3 Cooperativas

Em se tratando de desenvolvimento agrário, historicamente é impossível não destacar Marx e seus estudos quanto à agricultura familiar, particularmente em **“O capital” (MARX, 1988)**. Em seu diálogo com os populistas russos, Marx coloca uma questão fundamental para compreender o desenvolvimento da **agricultura familiar, qual seja, “as particularidades históricas”, uma vez que, aquilo que Marx observou na Inglaterra e na França, não é, necessariamente, válido para compreender a agricultura familiar em todos os contextos (PLEIN; FILIPPI, 2011).**

As teses marxistas apontavam para o desaparecimento da agricultura familiar com o desenvolvimento das relações capitalistas na agricultura, com o surgimento das grandes corporações. **“Quanto mais a agricultura se identifica com os padrões capitalistas, tanto mais se diferenciam qualitativamente as diferenças técnicas empregadas pelos grandes estabelecimentos das empregadas pelos pequenos” (KAUSTKY, 1986 apud PLEIN; FILIPPI, 2011, p. 99).**

Entretanto, constatou-se o predomínio da agricultura familiar na estrutura agrária nos países onde o modo de produção capitalista mais se desenvolveu, sobretudo no período pós-guerra (1945-1973). De acordo com Veiga (1991, p. 188), **“durante o grande impulso do desenvolvimento capitalista (meados da década de 1930 ao início da década de 1970) foi a agricultura familiar que acabou se firmando em todos os países do chamado Primeiro Mundo”**.

Não é tarefa fácil definir e situar o lugar da discussão sobre agricultura familiar em meio à teoria social contemporânea. Parece que a construção do conhecimento científico nesse campo disciplinar decorre, em larga medida, dos princípios teóricos e metodológicos que guiam as opções e convicções de vários autores (SCHNEIDER; NIEDERLE, 2009).

Antes de explicitar os fundamentos dessa perspectiva analítica, vale a pena comentar sobre o interesse recente e cada vez mais frequente no Brasil em torno da agricultura familiar. Afinal, ainda que não se tenha uma definição rigorosa e consensual sobre o estatuto conceitual da agricultura familiar, há uma certa generalização em torno da ideia de que o agricultor familiar é todo aquele sujeito que vive no meio rural e trabalha na agricultura juntamente com sua família. Assim definido pelo senso comum, o agricultor familiar abarca

uma diversidade de formas de fazer agricultura que se diferencia segundo tipos diferentes de famílias, o contexto social, a interação com os diferentes ecossistemas, sua origem histórica, entre outras. Tomando-se o Brasil de norte a sul, é possível encontrar uma diversidade muito grande de agricultores familiares, muitos deles obedecendo a denominações locais e regionais, como as de colono, sitiante, posseiro, morador, ribeirinho, entre outras. (SCHNEIDER; NIEDERLE, 2009, p. 990).

Nas unidades familiares de produção, o sistema produtivo em geral se assenta no trabalho da terra, realizado por uma família, e na produção primária, destinada prioritariamente à satisfação das necessidades internas da propriedade e do grupo doméstico (composto pelos membros da família). Nessa forma de organização do trabalho e da produção, as atividades agrícolas quase sempre coexistem e se complementam com outras atividades não agrícolas, como o artesanato e o comércio. O processo de mercantilização crescente da vida social e econômica leva a uma crescente interação e integração das famílias aos mercados. (SCHNEIDER; NIEDERLE, 2009).

Historicamente, identificou-se o crescente movimento dos produtores agrícolas durante a revolução industrial na busca de alternativas de participação nos mercados muitas vezes monopolizado pelas grandes corporações capitalistas industriais que surgiam, as quais buscavam a máxima exploração do capital aplicado seguindo as premissas de Adam Smith. Segundo Fleury (1983), as associações se opunham a essa figura capitalista, buscando assim através de sistemas cooperativos de rateio de custos e agregação de volumes produzidos no campo, melhorar a competitividade em um mercado que cada vez mais obrigava o trabalhador a migrar para as grandes cidades para os sistemas produtivos industriais em massa.

As associações e o uso de práticas solidárias e cooperativas têm sido entendidos como mecanismos que podem facilitar o processo produtivo, além de propiciar melhores relacionamentos com o mercado, instituições públicas e com a sociedade de maneira geral, o que contribuiu para a emergência, desde a década de 1980, de inúmeras iniciativas associativistas ou comunitárias que se propõem ao fortalecimento da agricultura familiar, resgatando a experiência democrática e as atividades de cooperação (BESERRA, 2013).

Portanto, foi no início dos anos 1900 que o cooperativismo começou a se delinear no Brasil, influenciado pela religiosidade e pelo pensamento político dos imigrantes. O movimento seguiu **principalmente o chamado “modelo alemão”, que defendia a educação cooperativista para estimular a solidariedade entre as pessoas, a união de todo o sistema na defesa dos interesses comuns e a distinção entre o cooperativismo e a economia de mercado, sendo o primeiro marcado pelo comprometimento com a justiça social (BRASIL, 2006).**

Com potência para gerar renda e para irradiar educação, o cooperativismo vem sendo um aliado dos governos e, por isso, tem recebido incentivos e amparo legal. No Brasil, o setor é incluído na legislação pela primeira vez no século XIX, na Constituição Federal de 1891, que garantia aos trabalhadores o direito de se associarem em cooperativas e em sindicatos.

Dadas as principais definições e conceitos envolvidos nesta pesquisa, apresenta-se no tópico a seguir a metodologia de análise a ser empregada no estudo.

3 METODOLOGIA DE ANÁLISE

A análise de agrupamentos pertence a um ramo da análise multivariada cujo objetivo é sintetizar ou simplificar a estrutura de variabilidade dos dados. O objetivo dessa técnica é organizar elementos, sejam essas observações ou variáveis, em grupos, com base nas informações quanto à sua proximidade. O objetivo da análise de agrupamentos é identificar grupos de objetos similares, cujas características sejam muito próximas entre aqueles pertencentes ao mesmo grupo e sejam bem distintas entre objetos pertencentes a grupos diferentes. Isto define uma tarefa de classificação, ou seja, objetos separados em grupos aos quais podem ser associadas (e rotuladas) classes ou categorias.

Para se definir a similaridade entre objetos, é necessário executar o levantamento das características relevantes para o problema abordado. Por exemplo, a análise da distribuição dos cooperados foi feita por critérios geográficos, baseados na localização descrita em valores de latitude e longitude. Com isto, pode-se identificar grupos de cooperados próximos entre si para sugerir o posicionamento de centros de coleta do mel

bruto e sua redistribuição para os centros beneficiadores ou consumidores. Em linhas gerais, a análise de agrupamentos inicia-se com uma base de dados de N objetos (registros) e M características (campos). Neste trabalho, $M=2$ (latitude e longitude).

O cálculo da distância ou dissimilaridade entre objetos é executado entre cada objeto e cada centroide. As métricas utilizadas para essa distância são euclidiana, manhattan, correlação de Pearson, de Mahalanobis, coeficiente de Jaccard, cossenos e muitas outras (JAIN; MURTY; FLYNN, 1999). Algumas são apropriadas para dados contínuos, outras para nominais e outras para dados binários.

Neste trabalho, foi utilizada a distância euclidiana, que corresponde à distância em linha reta entre dois pontos em um plano, o que é bem adequada devido aos dados provenientes de localização geográfica. Foram utilizados os valores em graus de arco, pois o diâmetro da Terra é considerado constante nas escalas de distância aqui abrangidas. Assim, existe uma proporção direta entre a distância na superfície terrestre (em km) e o deslocamento (em graus) na latitude e longitude.

Quando as variáveis apresentam escalas de medida muito diferentes entre si, é necessário empregar uma etapa de normalização, antes do processo de agrupamento propriamente dito. Porém, devido ao fato de que as únicas duas dimensões aqui utilizadas são proporcionais às distâncias geográficas (graus e km), não há necessidade de normalização.

Então, a aplicação da análise de agrupamentos implica em duas escolhas: primeiro, qual a métrica de (des)similaridade a ser utilizada e, segundo, qual o método de particionamento do conjunto de objetos em grupos. Os métodos de particionamento podem ser hierárquicos ou não, os mais conhecidos estão dentre os hierárquicos, que essencialmente se subdividem em aglomerativos e divisivos. Dentre todos estes, neste trabalho foi escolhido o método hierárquico divisivo denominado *k-means*.

O algoritmo *k-means* (MACQUEEN, 1967) consiste em assumir k grupos fixados *a priori*, definindo k *centroides*, um para cada grupo. Cada centroide fica posicionado no valor médio do seu grupo no espaço de **dados, logo, é considerado o “centro de gravidade” do grupo. Nessa definição inicial, a melhor escolha é colocá-los o máximo possível longe um do outro.** O próximo passo é tomar cada ponto pertencente a um determinado conjunto de dados e associá-lo ao centroide mais próximo. Quando nenhum ponto está pendente, o primeiro passo é concluído e um agrupamento inicial é feito. Em seguida, deve-se recalculá-los k novos centroides dos grupos resultantes do passo anterior. Depois, é necessário fazer uma reclassificação dos dados, verificando qual o centroide novo mais próximo de cada dado. Com a reatribuição dos dados aos grupos, os centroides dos novos grupos devem ser recalculados. Esses dois procedimentos geram uma repetição.

Assim, podemos notar que os centroides mudam cada vez menos sua localização até um limiar, logo o algoritmo é convergente. O algoritmo visa minimizar uma função objetiva, neste caso, uma função de erro ao quadrado. O algoritmo também é sensível às escolhas das posições iniciais dos centroides, então, *k-means* pode ser executado várias vezes para reduzir esse efeito. Algumas heurísticas podem ser propostas para uma boa escolha inicial, como por exemplo, distribuir uniformemente os k centroides, evitando posicioná-los onde a densidade de dados é baixa (o que requer um outro algoritmo de inicialização com uma certa complexidade). Muitas vezes, um grupo razoavelmente denso pode ficar associado a dois ou mais centros, com muitos dados na divisa entre os grupos assim formados.

O algoritmo *k-means* converge pela minimização de uma *função objetivo* dada pela soma dos erros quadráticos (ou variâncias intragrupos), como segue:

$$\sum_{c=1}^k \sum_{x_j \in C_c} (\vec{x}_j - \vec{x}_c)^2 \quad (1)$$

em que x_j são as coordenadas de um objeto j (no caso, clientes ou cooperados), x_c são as coordenadas do centroide. As somas são executadas sobre todos os objetos pertencentes ao respectivo centro ($x_j \in C$) e sobre todos os k centros obtidos.

O método *k-means* pode ser pensado como o inverso da análise de variância (Anova) (VIEIRA, 2006) enquanto o teste de significância Anova avalia a variabilidade intergrupos com a variabilidade intergrupos, o *k-means* realoca dados nos grupos iterativamente para otimizar os valores de Anova. É um algoritmo de implementação relativamente simples, geralmente convergindo para uma configuração estável em poucas iterações.

O algoritmo *k-means* pode ser descrito da seguinte maneira:

1. Escolher k distintos valores para centros dos grupos (de forma aleatória ou utilizando alguma heurística baseada na natureza do problema);
2. Associar cada ponto ao centro mais próximo;
3. Recalcular o centro de cada grupo;
4. Repetir os passos 2 e 3 até nenhum elemento mudar de grupo. Se houver a necessidade de saber a distância de cada elemento ao centroide do grupo a que pertence, o critério de parada poderá ser uma variação mínima nas distâncias a cada passo.

Como a divisão em k grupos é um requisito inicial da execução do algoritmo, e não se sabe ao certo quantos grupos resultam na melhor partição da massa de dados, foram desenvolvidos critérios de desempenho. O critério utilizado neste trabalho foi o proposto por Xie e Beni (1991) para a variante difusa do *k-means*, o *fuzzy c-means* (BEZDEK; EHRlich; FULL, 1984). A diferença para e entre estes métodos é que, enquanto o *k-means* atribui simplesmente cada dado ao centroide mais próximo, a versão *fuzzy* atribui um *valor de pertinência* para cada dado em relação a cada um dos grupos. O maior valor de pertinência é o grupo onde o objeto estiver “mais inserido”. Neste caso, existe um **parâmetro de nebulosidade** ($m > 1$) o qual, quanto mais próximo de 1, mais “nítida” é a partição do conjunto de dados em grupos (ou seja, mais próxima dos resultados do *k-means*). Então, para tirar proveito do critério de validação de Xie e Beni, foi aplicado o algoritmo *fuzzy c-means*. No final, o melhor particionamento da massa de dados (melhor número de grupos) é dado pelo menor índice Xie-Beni.

Em resumo, a metodologia aplicada para definir quantos centros de distribuição para a distribuição de cooperados foi a seguinte:

- 1) levantamento geográfico da localização dos cooperados em latitude e longitude;
- 2) aplicação do algoritmo *fuzzy c-means* com os seguintes parâmetros:
 - $N = 75$ cooperados;
 - 2 campos ou variáveis: latitude e longitude;
 - número de grupos a atribuir: variando de 2 até 10 (9 execuções do algoritmo);
 - “fator de nebulosidade” $m = 1,1$;
 - saídas geradas pelo algoritmo: melhor número k^* de grupos a atribuir, posição dos k^* centroides, distância de cada cooperado a cada centroide, grupo atribuído a cada cooperado, valor do indicador de Xie-Beni;
- 3) avaliação e interpretação dos resultados. Análise da coerência obtida.

Os dados primários utilizados para a realização deste estudo foram obtidos a partir de documentos fornecidos pela Cooperativa COAPIS em 2015 e 2016, tais como relatórios e fichas cadastrais. Informações complementares estão detalhadas em Souza (2017).

4 RESULTADOS E DISCUSSÃO

A partir das informações de localização geográfica de cooperados e clientes fornecidos pela cooperativa COAPIS, e da aplicação do método de agrupamento, foram obtidos dois grupos de resultados: por agrupamentos e por clientes, como descritos a seguir.

4.1 Resultados para a rede de cooperados

O resultado da execução do algoritmo está expresso na Tabela 1, na qual vê-se o valor do índice de Xie-Beni em função do número atribuído de grupos. Verifica-se que para $k=5$ foi apresentado o menor índice (em destaque na tabela).

A localização geográfica dos cooperados está no gráfico da Figura 1. Nesta, foram marcadas as localizações dos centroides resultantes do processo de agrupamento (com o símbolo □). Veem-se também linhas demarcatórias para ilustrar os grupos formados. Estas linhas apenas cercam os grupos, não foi efetuado nenhum cálculo a respeito do seu posicionamento. Entretanto, é importante notar que cada cidade está mais próxima do centroide (Figura 1) do grupo à qual pertence, principalmente aquelas limítrofes, como, p. ex., Itatiba, Indaiatuba, Jundiá, Tatuí, Tapiraí e São Miguel Arcanjo.

Tabela 1 – Índice de Xie-Beni para várias partições do conjunto de dados

k	2	3	4	5	6	7	8	9	10
Índice XB	0,370	0,367	0,261	0,207	0,461	0,481	0,555	0,420	0,446

* variando-se o número de grupos. A melhor partição está indicada em negrito.

Fonte: elaborado pelos autores (2017)

Os grupos ficaram assim constituídos:

Grupo 1 – 7 cooperados: Angatuba, Botucatu, Carlópolis (PR), Itapetininga, Itatinga, Porangaba, Tupã.

Grupo 2 – 13 cooperados: Apiaí, Capão Bonito, Guapiara, Iguape, Iporanga, Itapeva, Itariri, Juquiá, Miracatu, Registro, Ribeirão Branco, São Miguel Arcanjo, Sete Barras.

Grupo 3 – 32 cooperados: Alambari, Alumínio, Araçoiaba da Serra, Boituva, Cabreúva, Capeia do Alto, Carapicuíba, Cotia, Ibiúna, Indaiatuba, Iperó, Itapeçerica da Serra, Itu, Jundiá, Mairinque, Osasco, Piedade, Pilar do Sul, Pirapora do Bom Jesus, Porto Feliz, Salto, Salto de Pirapora, Santo André, São Bernardo do Campo, São Paulo, São Roque, Sarapuí, Sorocaba, Tapiraí, Tatuí, Vargem Grande Paulista, Votorantim.

Grupo 4 – 18 cooperados: Americana, Amparo, Araras, Artur Nogueira, Campinas, Casa Branca, Espírito Santo do Pinhal, Holambra, Itatiba, Jaguariúna, Mogi Mirim, Munhoz MG, Piracicaba, Rio Claro, São Pedro, Serra Negra, Socorro, Valinhos.

Grupo 5 – 5 cooperados: Jacareí, Salesópolis, São José dos Campos, São Luís do Paraitinga, Ubatuba.

No entanto, nenhum centroide recai exatamente sobre alguma cidade, exceto em Alumínio e Holambra. Pode-se então pensar na instalação de centros de captação e distribuição nas cidades ou às margens de rodovias mais próximas a cada centroide.

As coordenadas dos centros estão na Tabela 2, dispostas em ordem de longitude (oeste para leste), a fim de melhor identificar suas posições na Figura 1.

A Tabela 3 é um complemento da Tabela 2, onde inclui-se uma nova linha indicando se há cliente no grupo e qual o cliente mais próximo, o que pode ser facilmente verificado na Figura 2.

Na Tabela 4 estão as demandas informadas pelos clientes. Em conjunto com a Tabela 3, pode-se analisar a viabilidade de instalação dos centros nos locais indicados.

Tabela 3 – Coordenadas dos centroides obtidos, cidades e clientes mais próximos a cada centroide

	Centro 1	Centro 2	Centro 3	Centro 4	Centro 5
Longitude (°)	-48,90	-48,10	-47,26	-47,05	-45,61
Latitude (°)	-23,08	-24,28	-23,50	-22,58	-23,33
Cidade mais próxima	Itatinga	Sete Barras	Alumínio	Holambra	Jacareí
Cliente mais próximo	Itatinga	Sete Barras	Sorocaba	Rio Claro	–

Fonte: elaborado pelos autores (2017).

Tabela 4 – Demandas informadas pela COAPIS

Município	Demandas (kg)	Grupo
Botucatu	6362	1
Itapetininga	2037	1
Itatinga	1250	1
Capão Bonito	710	2
São Miguel Arcanjo	1025	2
Sete Barras	2500	2
Santo André	1500	3
Sorocaba	1500	3
Tatuí	3000	3
Rio Claro	1000	4
Araçariguama	3000	–
Itaoca	1500	–
Sarutaia	1240	–

Fonte: elaborado pelos autores (2017).

4.3 Aplicação do método de agrupamento

A área abrangida pelos municípios e suas localizações foi dividida em grupos de modo que a soma das distâncias a serem percorridas entre os centros de distribuição e os cooperados seja a menor possível. Em consequência, os custos de transporte tornam-se minimizados contabilizando-se todos os grupos. Isto decorre do fato de que o algoritmo *k-means* converge pela minimização da função objetivo descrita na Equação 1.

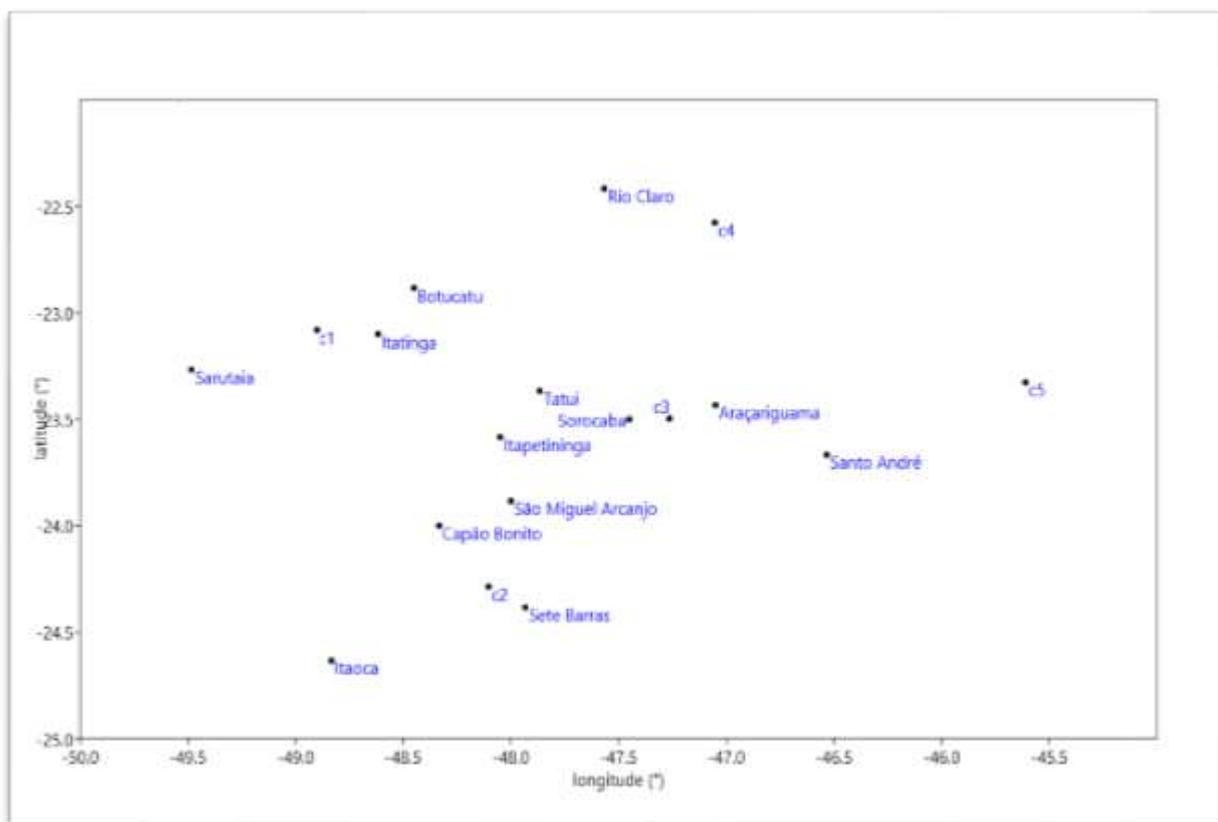
As vantagens de se utilizar o algoritmo de agrupamentos *k-means*, além de obter o centroide de cada grupo (que é o cerne do problema da distribuição), é ser simples e intuitivo; a complexidade computacional para execução é linear para todas as variáveis importantes (número de objetos e número de grupos); é eficaz

para muitos domínios de aplicação, com resultados de interpretação simples. Ainda, é considerado um dos 10 mais influentes algoritmos em *Data Mining* (WU, KUMAR, 2009).

Como desvantagens, a principal é que não se sabe em quantos grupos deve ser particionado o conjunto de objetos para otimizar a função objetivo. Além disso, é sensível à inicialização dos centroides (pode apresentar vários mínimos locais para a função objetivo). Também, limita-se a encontrar grupos esféricos – para encontrar grupos alongados, substitui-se a distância euclidiana pela de Mahalanobis (BEZDEK; EHRlich; FULL, 1984), no entanto, aumenta o custo computacional. É limitado a atributos numéricos e sensível a *outliers* (pontos extremos).

Neste trabalho, a desvantagem de não se saber o número de grupos *a priori* que melhor separa os dados foi contornada pela múltipla execução do algoritmo com a avaliação do índice Xie-Beni. A desvantagem de melhor separar grupos esféricos (circulares, no caso bidimensional) não ocorre aqui, pois trata-se de dados de localização geográfica, onde a distância euclidiana funciona bem.

Figura 2 – localização dos clientes em relação aos centros distribuidores propostos



Fonte: elaborado pelos autores (2017).

O método do centro de gravidade (BALLOU, 2006), comparado com *clustering*, não serve para particionar objetos em grupos, apenas determina o centroide de um determinado conjunto de dados. Baseado em cálculo de média ponderada, é fácil de implementar no caso de dados com pesos (por exemplo, capacidade prévia de um centro distribuidor, como discutido aqui). No entanto, uma simples modificação no algoritmo *k-means* pode tratar desta particularidade – atribuir pesos aos dados.

4.4 Resultados por agrupamentos

Nesta seção analisam-se os resultados obtidos, discutindo-se as possibilidades de implementação.

Em relação ao grupo 1, nota-se que Itatinga é a cidade mais próxima do centroide do grupo. Mesmo sendo uma cidade pequena para os padrões brasileiros (23 mil habitantes), está localizada muito próxima à Rodovia Castello Branco, que é uma importante via de escoamento comercial. Além disso, a implantação de

um centro de distribuição melífero no município resulta em expansão e diversificação na sua atividade comercial.

Quanto ao grupo 2, o centroide calculado está no meio do Parque Estadual Intervales, tendo como cidade mais próxima Sete Barras. Em termos logísticos, é recomendável propor a instalação do centro na cidade de Registro, que fica às margens da rodovia Régis Bittencourt.

O grupo 3 é o mais denso, compreendendo 32 municípios e o centroide está no perímetro urbano do município de Alumínio, cortado pela rodovia Raposo Tavares. Devido à boa malha viária da região, pode-se propor a instalação do centro em Sorocaba, que é centro regional.

O grupo 4 é o segundo mais denso (18 municípios) e as cidades estão em torno de uma boa malha viária, com alguns pontos de serra. O centro está em Holambra. Poderia se pensar na instalação do centro de distribuição tanto em Campinas (ao sul) ou em Mogi Mirim (ao norte), ambas distantes cerca de 35 km de Holambra. No entanto, Holambra estabeleceu-se como polo floricultor e a associação disto com a produção e a distribuição de mel se torna apropriada, para alavancar a instalação do centro.

No grupo 5, observa-se que o centroide fica próximo a um município não listado (Jambeiro). A análise da malha viária na região indica a instalação do centro distribuidor em Jacareí ou São José dos Campos, ambas às margens da Rodovia Dutra.

Desta forma, são definidas as recomendações para instalação dos centros distribuidores em função de facilidades logísticas.

4.5 Resultados por Clientes

Com os dados da Tabela 4, é interessante uma rediscussão dos centros de distribuição propostos na seção anterior, uma vez que vários clientes estão no mesmo município de cooperados. Os centros propostos, em função da demanda informada, são atualizados da seguinte maneira:

- Grupo 1: Itatinga para Botucatu. A distância de Botucatu ao centroide geográfico 1 é um pouco maior, mas a alteração se justifica pela enorme demanda de Botucatu, mais que o dobro de qualquer outro cliente.
- Grupo 2: mantém-se em Sete Barras, devido a apresentar a maior demanda do grupo.
- Grupo 3: com centroide coincidindo com o município de Alumínio, recomenda-se instalar o centro distribuidor em Sorocaba, devido aos seguintes fatores: grande demanda, forte infraestrutura, localização viária privilegiada (confluência de várias rodovias) e centro regional.
- Grupo 4: apesar de o único cliente estar em Rio Claro, a região apresenta um potencial de consumo futuro bastante elevado. Mudando-se o centro para Rio Claro, há um deslocamento considerável por parte dos cooperados, devido ao afastamento do centroide do grupo. A soma das distâncias percorridas é maior do que em relação ao centroide. Devido à infraestrutura já existente em Holambra para ser um centro de distribuição, no caso flores, recomenda-se manter o centro neste município.
- Grupo 5: não consta nenhum cliente nesta região, no entanto, a instalação de um centro distribuidor em Jacareí ou São José dos Campos possibilita alavancar o desenvolvimento comercial do mel e seus derivados.

Os três últimos municípios da Tabela 4 apresentam demanda considerável e, conforme o mapa mostrado na Figura 2, devem consumir o produto dos respectivos centros recomendados: Sarutaia de Botucatu (C1), Itaoca de Sete Barras ou Registro (C2) e Araçariguama de Sorocaba (C3).

5 CONSIDERAÇÕES FINAIS

O presente estudo buscou aplicar a análise de agrupamentos para definir oportunidades de redução de custos na rede de distribuição de mel de uma Cooperativa de Apicultores, da qual participam 357 cooperados distribuídos em 75 municípios, em um raio de atuação de mais de 200 km.

A análise de agrupamentos por *clustering* mostrou ser uma ferramenta prática para esse tipo de estudos, oferecendo grandes vantagens comparada aos cálculos baseados apenas em centros de gravidade, com diversas possibilidades de ajustes ao longo de sua aplicação.

Os resultados apontaram que poderá se obter redução de custos mediante instalação de centros de distribuição em pontos estratégicos da área de abrangência da cooperativa.

O arranjo empresarial de cooperativa da COAPIS também sugeriu uma análise mais estruturada de sua configuração devido a sua dispersão geográfica. Arranjos para a produção solidária de mel, direcionaram uma avaliação considerando a COAPIS como uma cadeia produtiva mais complexa, integrando produtores de mel ao mercado paulista, nacional e internacional.

Sugerem-se assim novas pesquisas que busquem a elaboração de um plano de negócios e a definição de viabilidade monetária para a expansão da rede de cooperados e a análise da intralogística, que considerem as atividades inerentes à movimentação e armazenagem de materiais nas instalações da COAPIS.

REFERÊNCIAS

- BALLOU, R. H. Gerenciamento da cadeia de suprimentos. 5. ed. Porto Alegre: Bookman, 2006.
- BATALHA, M. O. (coord.). Gestão agroindustrial, GEPAl: grupo de estudos e pesquisas agroindustriais. 3. ed. São Paulo: Atlas, 2007.
- BESERRA, M.R. ASSOCIATIVISMO RURAL: estratégia de participação dos agricultores familiares para afirmação e defesa dos seus direitos no mercado. // JORNADA INTERNACIONAL DE POLÍTICAS PÚBLICAS, 6., 2013, São Luís do Maranhão. Anais [...]. São Luís do Maranhão: Universidade Federal do Maranhão, 2013.
- BEZDEK, J. C.; EHRLICH, R.; FULL, W. F. C. M: The fuzzy c-means clustering algorithm. Computers & Geosciences, v. 10, n. 2-3, p. 191-203, 1984.
- BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Evolução do cooperativismo no Brasil: DENACOOOP em ação. Brasília: MAPA, 2006.
- BRITO, F. E. M. Agricultores, cooperativas e a organização da produção de mel no território nordeste II. // CONGRESSO LUSO AFRO BRASILEIRO DE CIÊNCIAS SOCIAIS., 11., 2011, Salvador. Anais [...]. Salvador: Universidade Federal da Bahia, Campus de Ondina, 2011.
- CABETE, N. P. F.; DACOL, S. Identificação das características dos Arranjos Produtivos Locais. // ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO (ENEGEP), 28., 2008, Rio de Janeiro. Anais [...]. Rio de Janeiro: ABEPRO, 2008.
- FAO. Organização das nações unidas para a alimentação e agricultura. Cooperativas agrícolas, a chave para reduzir a fome e a pobreza. 2012. Disponível em: <http://humanitasdhecidania.blogspot.com/2012/09/cooperativas-agricolas-chave-para.html>. Acesso em: 18 nov. 2015.
- FENG, Li; HENDRIKSE, George. On the nature of a cooperative: a system of attributes perspective. Netherlands: ERIM Report Series Research in Management, 2007.
- FLEURY, M. T. L. Cooperativas agrícolas e capitalismo no Brasil. São Paulo: Global, 1983.
- GALDÁMEZ, E. V. C.; CARPINETTI, L. C. R.; GEROLAMO, M. C. Proposta de um sistema de avaliação do desempenho para arranjos produtivos locais. Gestão e Produção, São Carlos-SP, v. 16, n. 1, p. 133-151, jan./mar. 2009.
- GARCIA, R.; MADEIRA, P. Uma avaliação da difusão de práticas de gestão da produção entre pequenas empresas em sistemas locais de produção. Produção, São Paulo, SP, v. 23, n. 1, p. 20-30, jan./mar. 2013.

GONÇALVES, A. T. P.; LEITE, M. S. A.; SILVA, R. M. Um estudo preliminar sobre as definições e as diferenças dos principais tipos de arranjos empresariais. *Revista Produção Online*, Florianópolis, SC, v. 12, n. 3, p. 827- 854, jul./set. 2012.

JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data clustering: a review. *ACM computing surveys (CSUR)*, New York, NY, USA, v. 31, n. 3, p. 264-323, 1999.

KELLER, P. F. Clusters, distritos industriais e cooperação interfirmas: uma revisão da literatura. *Revista Economia & Gestão*, Minas Gerais, v. 8, n. 16, p. 30-47, 2008.

LAKATOS, E.; MARCONI, M. A. Fundamentos de metodologia científica. 6. ed. São Paulo: Atlas, 2005.

LASTRES, H. M. M.; CASSIOLATO, J. E. Mobilizando conhecimentos para desenvolver arranjos e sistemas produtivos e inovativos locais de micro e pequenas empresas no Brasil. *RedeSist*, 2005. Disponível em: <http://www.ie.ufrj.br/redesist/P3/NTF2/Cassiolato%20e%20Lastres.pdf>. Acesso em: 15 dez. 2018.

MACQUEEN, J. B. Some Methods for classification and Analysis of Multivariate Observations. [1967]. Disponível em: <https://projecteuclid.org/euclid.bsm/1200512992>. Acesso em: 15 dez. 2018.

MANZO, A. J. Manual para la preparación de monografías: una guía para presentar informes y tesis. Buenos Aires: Humanitas, 1971.

MARSHALL, A. Princípios de economia. São Paulo: Nova Cultura, 1920.

MARTINS, E. R.; SILVA, S. Estudo sobre as definições e as diferenças dos principais tipos de arranjos empresariais. // ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO, 35., 2015, Fortaleza. Anais [...]. Fortaleza, CE: UFC, 2015.

MARX, K. O capital. São Paulo: Nova Cultural, 1988.

PLEIN, C.; FILIPPI, E. E. Capitalismo, agricultura familiar e mercados. *REDES*, Santa Cruz do Sul, v. 16, n. 3, p. 98-121, set/dez 2011.

RITZMAN, L. P.; KRAJEWSKI, L. J.; MALHOTRA, M. Administração de produção e operações. 8. ed. São Paulo: Person Prentice Hall, 2009.

SCHMITZ, H.; NADVI, K. Clustering and industrialization: introduction. *World Development*, Oxford, v. 27, n. 9, p. 1503-1514, 1999.

SCHNEIDER, S.; NIEDERLE, P. A. Agricultura familiar e teoria social: a diversidade das formas familiares de produção na agricultura. // SIMPÓSIO NACIONAL CERRADO, 9., 2009, Planaltina, DF. Anais [...]. Planaltina, DF: Embrapa Cerrados, 2009. p. 988-1014.

SOUZA, S. Otimização do arranjo empresarial da COAPIS - Cooperativa de Apicultores de Sorocaba e região. 2017. 116 f. Dissertação (Mestrado em Processos Tecnológicos e Ambientais) - Universidade de Sorocaba, Sorocaba, SP, 2017.

SUZIGAN, W. Identificação, mapeamento e caracterização estrutural de arranjos produtivos locais no Brasil. Relatório Consolidado. Rio de Janeiro: IPEA, 2006.

VEIGA, J. E. da. O desenvolvimento agrícola: uma visão histórica. São Paulo: Hucitec, 1991.

VENANZI, D. Uma análise dos modelos de consórcio modular e condomínio industrial na indústria automobilística brasileira sob a perspectiva de estratégia de operações. 2009. 207 f. Dissertação (Mestrado) - Universidade Estadual Paulista, Faculdade de Engenharia de Bauru, 2009. Disponível em: <http://hdl.handle.net/11449/93025>. Acesso em: 15 dez. 2018.

VIEIRA, S. Análise de variância (Anova). São Paulo: Atlas, 2006.

WU, X.; KUMAR, V. The top ten algorithms in data mining. Estados Unidos: Chapman & Hall/CRC, 2009.

XIE, X. L.; BENI, G. A validity measure for fuzzy clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 13, p. 841–847, 1991.

YIN, R. K. Estudo de caso: planejamento e métodos. 5. ed. Porto Alegre: Bookman, 2015.