

CLUSTERIZAÇÃO K-MEANS PARA DECISÃO DE LOCALIZAÇÃO INDUSTRIAL

Lucas Antonio Risso (FCA UNICAMP)

lucasrisso@gmail.com

Luciano Kiniti Issoe (FCA UNICAMP)

miagao@gmail.com

Ailson Renan Santos Picanco (FCA UNICAMP)

ailson.picanco@outlook.com

Alessandro Lucas da Silva (FCA UNICAMP)

alessandro.silva@fca.unicamp.br

Paulo Sergio de Arruda Ignacio (FCA UNICAMP)

paulo.ignacio@fca.unicamp.br



A definição do local de uma nova fábrica é uma relevante questão estratégica, devido a influência direta no tempo de resposta ao cliente e nos custos logísticos. Tendo conhecimento dos pedidos de clientes de uma empresa do segmento alimentício, este artigo tem como objetivo aplicar uma abordagem de aprendizado estatístico não supervisionado, por meio do método K-Means, para estudar a possível localização de uma nova planta industrial com base em carga e distância, bem como em custos de transporte parametrizados. A clusterização faz a atribuição das cidades que possuem clientes entre a matriz e a nova instalação considerando, ao invés de distâncias euclidianas, as distâncias entre coordenadas geográficas como numa esfera, ponderando-as pela somatória das cargas correspondentes aos clientes de cada cidade e pela parametrização dos custos de transporte de matéria-prima e produto acabado. Após serem formados os clusters, a região ideal para a localização da segunda planta é obtida usando-se o método do Centro de Gravidade. Os resultados revelam o impacto da variação dos custos logísticos para a formação dos clusters, e conseqüente, para a localização da nova fábrica.

Palavras-chave: Localização, clusterização k-means, centro da gravidade, logística e distribuição, aprendizado estatístico.

1. Introdução

Em um cenário econômico cada vez mais competitivo, as cadeias de suprimentos do segmento alimentício passaram a adotar a busca constante pela melhoria do fluxo que tem início na produção rural, procurando inovar as políticas de gerenciamento da produção, as técnicas de processamento e as alternativas de canais de distribuição, com o intuito de aumentar a assertividade quanto a disponibilidade de produtos e maximizar os resultados obtidos, (AHUMADA; VILLALOBOS, 2010).

De acordo com Yu e Nagurney (2013), a cadeia alimentícia é complexa, e lida com diversas particularidades e desafios, incluindo custos de descarte de produtos perecíveis e operações com baixas margens de lucro. No caso de produtos de cultivo agrícola, a diferenciação é dada por atributos de frescor, e há casos de produtos sazonais que são demandados durante todas as estações, fato acentuado pelo advento da gastronomia.

Segundo Akkerman, Farahani e Grunow (2010) a qualidade do alimento sofre decréscimo contínuo enquanto percorre os deslocamentos entre os elos da cadeia, principalmente caso não haja o condicionamento adequado, de tal modo que são sugeridos a integralização vertical e o agilizar das atividades sempre que possível.

Parte inerente à configuração e ao êxito da Cadeia de Valor, os problemas de localização abordam diferentes tipos de instalações, tais como centros de distribuição e fábricas, constituindo uma área de pesquisa recorrente no âmbito da Pesquisa Operacional (MELO; NICKEL; SALDANHA-DA-GAMA, 2009).

Sendo assim, o objetivo central deste trabalho é estudar a localização de uma planta industrial do segmento alimentício por meio do método de clusterização K-Means com base em cargas, distâncias e custos parametrizados. Devido a questões climáticas, a matéria-prima continuará sendo produzida e fornecida pela matriz, situada na região norte do território brasileiro, fato que atribui maior complexidade ao fluxo. Após realizar a clusterização das cidades-cliente, a localização da nova fábrica é indicada pelo método do Centro da Gravidade.

2. Revisão da Literatura

O dinamismo do mercado tem ocasionado a redução do ciclo de vida de produtos e a adoção de lotes de transporte menores, com isso a concepção da cadeia de suprimentos tem assumido singular importância, (MELO; NICKEL; SALDANHA-DA-GAMA, 2009), pois o panorama atual caminha para a competitividade entre cadeias de valor, em termos de responsividade, confiabilidade, flexibilidade e comunicação, (CHRISTOPHER, 2009).

De acordo com Ballou (2001), existe uma relação interativa entre os problemas de localização, a configuração de redes, o arranjo de cadeia de suprimentos e as estratégias de sistemas logísticos. Em linhas gerais, as decisões associadas lidam com objetivos que envolvem custos conflitantes, no que diz respeito à produção, ao armazenamento, aos níveis de estoque e à distribuição, com metas de lucratividade e de nível de serviço.

Os problemas de localização são divididos em duas categorias principais: problemas de localização de uma única instalação (SFLP, do inglês *single facility location problem*) e problemas de localização de múltiplas instalações (MFLP, do inglês *multi-facility location problem*). O caso do MFLP envolve a alocação de clientes a serem atendidos por cada uma das instalações, sendo que independente de sua categoria, os problemas podem ser explorados em condições de recursos capacitados, (ESNAF; KÜÇÜKDENİZ, 2009). Existem ainda os problemas de localização competitiva (CFL, do inglês *competitive facility location*), sendo que estes consideram questões de atratividade frente aos posicionamentos geográfico e de mercado dos concorrentes, (KUÇUKAYDIN; ARAS; ALTINEL, 2012).

Nos problemas de localização, as demandas são informadas como um conjunto de pontos discretos, e devem ser consideradas variações, devido ao grau de incerteza atrelado. Albareda-Sambola, Fernández, Saldanha-da-Gama (2011), estudam o problema de localização associando a demanda à distribuições de Bernoulli com o intuito de gerar cenários a partir de variações estocásticas. Neste sentido, Wagner, Bhadury e Peng (2009) mencionam que a robustez da alternativa de localização indicada é beneficiada ao se levar em conta cenários extremos, isto é, explorar desde a condição de maior adversidade até a mais provável, principalmente ao envolver variáveis de custos logísticos. Neste sentido, Kosior e Strong (2006) mencionam a relevância das tarifas e taxas associadas, e a influência do uso de economias de escala a fim de obter maior eficiência de distribuição.

Em uma perspectiva mais ampla, Yao *et. al.* (2010) afirmam que a gestão de estoques é uma questão que também deve ser incluída nas decisões de localização, uma vez que as políticas de reposição e os custos de manutenção de inventário detém alto impacto sobre os resultados e sobre o fluxo de caixa. Já Marín (2011) adota uma abordagem particular, a qual busca o balanceamento do número de pontos de demanda a serem atendidos por cada instalação, buscando equiparar o nível de utilização dos recursos.

Diante da complexidade do problema, a clusterização assume papel significativo para o direcionamento das tratativas visando a resolução de um problema de localização de múltiplas instalações. A clusterização de dados, em outras palavras o agrupamento em categorias similares, traz como benefícios a redução do tempo de resolução e o acréscimo da confiabilidade dos resultados para a alocação de recursos industriais, assim como ocorre em problemas de diversas áreas, (NG; LAM, 2011).

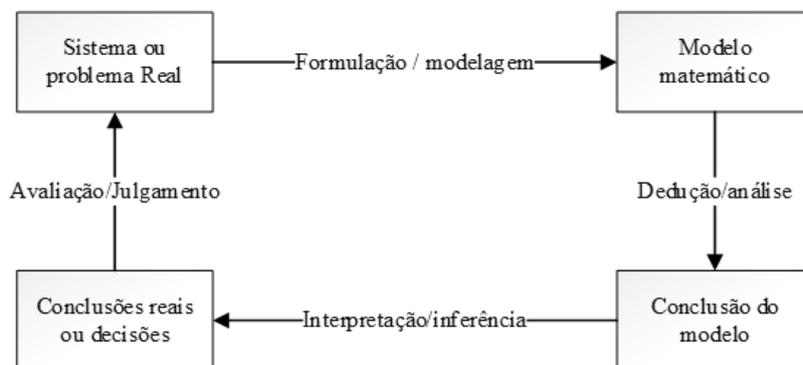
Quanto aos métodos quantitativos para localização de uma instalação, disponíveis em literatura, destacam-se o método do momento, o método de carga-distância e o método do centro da gravidade. Em contraste aos dois primeiros, o método do centro da gravidade tem como vantagem o fato de estar apto a indicar localidades que não são pontos de demanda, em contrapartida, também pode indicar locais inviáveis, como por exemplo localidades em alto mar, (BALLOU, 2006).

3. Métodos e procedimentos

A pesquisa empírica normativa quantitativa visa o desenvolvimento de políticas, estratégias e ações que visam melhorar a situação corrente. Esse tipo de pesquisa baseia-se em modelos que prescrevem uma decisão para o problema, (MIGUEL, 2011).

Nesse sentido, essa pesquisa é aderente a esta classificação por indicar uma localização ótima, mediante aos parâmetros considerados, utilizando a base de dados real referente aos pedidos processados por uma empresa alimentícia durante um período de 12 meses. Foi desenvolvido um modelo para o problema real adotado, de acordo com a sequência representada na Figura 1.

Figura 1 - Processo de modelagem matemática



Fonte: Miguel (2011)

Foram obtidas informações de demanda e localização de cada um dos clientes que solicitaram ao menos um pedido durante o intervalo analisado, os quais estão situados nas cinco regiões do território brasileiro. Primeiramente, os dados foram agrupados no nível de município, e então foram levantadas as coordenadas geográficas das cidades-cliente. Com esses dados, fez-se então uma clusterização por meio do método K-means.

O método de clusterização K-Means é um algoritmo iterativo, que gera uma solução factível ao alocar um ponto ao cluster k cujo “critério de agrupamento” é menor em relação ao centroide em questão, (BALCAN; EHRLICH; LIANG, 2013). O parâmetro $k = 2$, é definido pelo problema, uma vez que é prospectada a abertura de apenas uma nova fábrica.

No entanto, o presente problema implica em algumas adaptações ao algoritmo padrão:

- os centros dos clusters devem ser atualizados a cada iteração, contudo, no modelo proposto, o algoritmo foi utilizado de maneira que a primeira fábrica se mantenha no mesmo local, devido às restrições reais, isto é, a localização da fábrica atual, que não será transferida;
- os centros dos clusters são tipicamente calculados usando-se a média aritmética, enquanto o modelo proposto usa a média ponderada pela somatória das cargas associadas aos clientes agrupados, envolvendo ainda uma parametrização relacionada aos custos;
- as distâncias entre os pontos foram calculadas como sendo as distâncias entre pontos de uma esfera, ao invés de se usar distâncias euclidianas. A modificação se faz

necessária devido às distorções causadas pelas representativas distâncias entre as cidades-cliente consideradas no problema.

Sendo a Terra uma esfera de raio R e dois pontos de latitude φ_1 e φ_2 e longitude λ_1 e λ_2 , respectivamente, a distância d entre coordenadas pode ser calculada pela fórmula do “haverseno”, cujo procedimento é mostrado em (1), (2) e (3).

$$a = \sin^2\left(\frac{\Delta\varphi}{2}\right) + \cos(\varphi_1) \cdot \cos(\varphi_2) \cdot \sin^2\left(\frac{\Delta\lambda}{2}\right) \quad (1)$$

$$c = 2 \cdot \tan^{-1}\left(\frac{\sqrt{a}}{\sqrt{1-a}}\right) \quad (2)$$

$$d = R \cdot c \quad (3)$$

A clusterização divide os clientes em dois grupos, atribuindo aqueles que são atendidos pela Fábrica 1, bem como pela nova fábrica com base em uma função de custo relativo de transporte de matéria-prima e produtos acabados, usando a distância calculada a partir de coordenadas geográficas, conforme indica a equação 4:

$$f_{custo_{Ci}} = d_{(F_k-F_1)} \cdot Cg_i \cdot C_{TMP} + d_{(C_i-F_k)} \cdot Cg_i \cdot C_{TPF} \quad (4)$$

Onde:

$d_{(F_k-F_1)}$ = distância em km da fábrica k para a Fábrica 1, fornecedora de matéria-prima;

Cg_i = carga em kg a ser entregue para os clientes localizados na cidade i ;

C_{TMP} = custo de transporte de matéria-prima em $\frac{R\$}{km.kg}$;

$d_{(C_i-F_k)}$ = distância em km da cidade i para a Fábrica k que atende o seu pedido;

C_{TPF} = custo de transporte de produto final em $\frac{R\$}{km.kg}$;

$k = (1; 2)$, indica a qual cluster uma cidade i pertence;

$i = (1; 91)$, atribui a cidade i do cliente para o qual a carga de produto final será entregue;

Nota-se que caso o cliente seja atendido pela Fábrica 1 o primeiro termo da soma apresentada

na equação 4 torna-se nulo, uma vez que não há necessidade de transporte de matéria-prima e o termo correspondente à distância entre as plantas, $d_{(F_k-F_1)}$ assumirá valor zero. Por sua vez, a localização da Fábrica 2, ou seja, na nova instalação, é atualizada a cada iteração, até que haja a convergência a partir de um valor inicial escolhido aleatoriamente. Adota-se $k = 2$, pois atualmente a empresa possui uma fábrica e busca a abertura de uma planta, ao passo que i assume o respectivo intervalo devido ao fato de os clientes estarem localizados em 91 cidades do território brasileiro.

Em tempo, os custos relacionados à estoques não são contemplados pelo modelo, visto que dada a agregação dos dados, não são consideradas variações da demanda, o giro e os estoques de segurança necessários, de modo a não ser possível computar o custo do estoque em trânsito.

Por sua vez, os termos C_{TMP} e C_{TPF} são aproximações que pressupõem que os custos de transporte variam em razão da distância entre as fábricas e os clientes. No entanto, sabe-se que na prática os custos logísticos são dependentes de distâncias rodoviárias, do tipo de fracionamento de carga, além de fatores locais, tais como preços de combustíveis e pedágios, condições da malha rodoviária, entre outros. Ainda assim, a aproximação funciona como um balizador satisfatório para a estimativa, corroborando com o processo de decisão dada a relação estabelecida pela razão r , determinada por (5).

$$r = \frac{C_{TMP}}{C_{TPF}} \rightarrow C_{TMP} = r \cdot C_{TPF} \quad (5)$$

Substituindo-se C_{TMP} em (4), obtém-se:

$$f_{custo} = d_{(F_k-F_1)} \cdot Cg_i \cdot r \cdot C_{TPF} + d_{(C_i-F_k)} \cdot Cg_i \cdot C_{TPF} \quad (6)$$

$$f_{custo} = C_{TPF} \cdot d_{(F_k-F_1)} \cdot Cg_i \cdot r + C_{TPF} \cdot d_{(C_i-F_k)} \cdot Cg_i \quad (7)$$

$$f_{custo} = C_{TPF} (d_{(F_k-F_1)} \cdot Cg_i \cdot r + d_{(C_i-F_k)} \cdot Cg_i) \quad (8)$$

Logo, a clusterização determina a fábrica responsável por atender as cidades-cliente consideradas de acordo com a seguinte condição:

$$C_i \in k = \begin{cases} 1 & d_{(C_i-F_1)} \cdot Cg_i < d_{(F_2-F_1)} \cdot Cg_i \cdot r + d_{(C_i-F_2)} \cdot Cg_i \\ 2, & \text{caso contrário} \end{cases}$$

Isto é, a cada iteração, a cidade-cliente i é alocada ao cluster k se o múltiplo da carga atribuída pela distância percorrida é menor caso a cidade i seja atendida pela fábrica k , considerando a movimentação adicional de matéria-prima para o caso da Fábrica 2. A cada iteração, a localização da nova fábrica é atualizada por meio do método do Centro da Gravidade, considerando as coordenadas de latitude e longitude, bem como a demanda dos clientes então alocados ao Cluster 2, de acordo com (9) e (10):

$$LAT_{F_2} = \frac{\sum_{i \in k=2}^{91} Cg_i \cdot Cord_{LATi}}{\sum_{i \in k=2}^{91} Cg_i} \quad (9)$$

$$LONG_{F_2} = \frac{\sum_{i \in k=2}^{91} Cg_i \cdot Cord_{LONi}}{\sum_{i \in k=2}^{91} Cg_i} \quad (10)$$

Onde:

LAT_{F_2} = latitude da localização indicada para a Fábrica 2;

$LONG_{F_2}$ = longitude da localização indicada para a Fábrica 2;

$Cord_{LATi}$ = coordenada de latitude da cidade cliente i ;

$Cord_{LONi}$ = coordenada de longitude da cidade cliente i ;

Afim de orientar o processo de tomada de decisão, é possível saber o custo correspondente a um cenário ao realizar a somatória dos custos de cada cluster:

$$f_{custo-total} = \sum_{i \in k=2}^{91} d_{(F_k-F_1)} \cdot Cg_i \cdot C_{TMP} + \sum_{k=0}^2 \sum_{i \in k}^{91} d_{(C_k-F_k)} \cdot Cg_i \cdot C_{TPF} \quad (11)$$

Fazendo a substituição mostrada nas passagens (5), (6) e (8), tem -se

$$f_{custo-total} = C_{TPF} \left(\sum_{i \in k=2}^{91} d_{(F_k - F_1)} \cdot C g_i \cdot r + \sum_{k=0}^2 \sum_{i \in k}^{91} d_{(C_k - F_k)} \cdot C g_i \right) \quad (12)$$

E assim é possível obter um comparativo percentual em relação ao cenário atual a partir do fator de esforço logístico parametrizado em função de r , termo entre parênteses em (12), o qual pode ser assumido como um indicativo da carga de trabalho requerida para movimentação de produtos e de matéria-prima.

O algoritmo foi aplicado aos dados estudados com programação em GNU Octave, de acordo com o seguinte pseudocódigo:

```

Escolher aleatoriamente k-1 locais
Enquanto (Verdadeiro )
    Crie k agrupamentos atribuindo cada cliente à fábrica mais próxima
    Calcule o local da nova fábrica aplicando o método do Centro de Gravidade aos elementos do grupo  $k > 1$ 
Se o local não mudar
    Retorne
    
```

Na próxima sessão, são mostrados os resultados de uma aplicação com o uso dos dados reais, sendo que os cenários são obtidos com a variação de r de 0 a 1 com incrementos centesimais, os infográficos foram elaborados a partir da ferramenta CartoDB.

4. Resultados e discussão

Inicialmente, a fábrica localizada no estado do Pará, atende clientes distribuídos em 91 cidades de 17 estados brasileiros, conforme informa o Quadro 1, que traz a distribuição percentual de cargas e a de cidades por estados e regiões, mostrando também potenciais locais para a expansão da participação de mercado.

Quadro 1 – Distribuição da demanda por estados e regiões

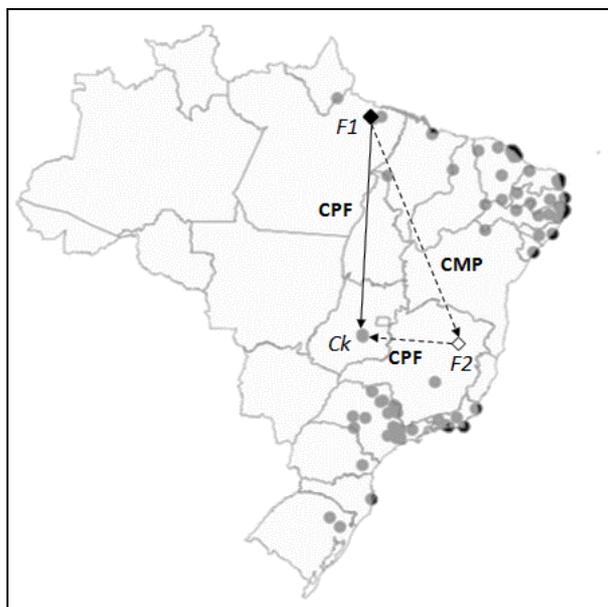
Região	Estado	UF	Por estado		Por região	
			Carga de produtos	Cidades cliente	Carga de produtos	Cidades cliente
Norte	Amazonas	AM	-	-	5,8%	3
	Roraima	RR	-	-		

	Amapá	AP	0,8%	1		
	Pará	PA	5%	2		
	Tocantins	TO	-	-		
	Rondônia	RO	-	-		
	Acre	AC	-	-		
Nordeste	Maranhão	MA	2%	3	38%	39
	Piauí	PI	2%	3		
	Ceará	CE	11%	10		
	Rio Grande do Norte	RN	2%	5		
	Pernambuco	PE	18%	12		
	Paraíba	PB	2%	3		
	Sergipe	SE	0,1%	1		
	Alagoas	AL	0,9%	2		
Bahia	BA	-	-			
Centro Oeste	Mato Grosso	MT	-	-	0,2%	2
	Mato Grosso do Sul	MS	-	-		
	Goiás	GO	0,2%	2		
Sudeste	São Paulo	SP	26%	30	50%	43
	Rio de Janeiro	RJ	9,4%	12		
	Espírito Santo	ES	-	-		
	Minas Gerais	MG	14,6%	1		
Sul	Paraná	PR	0,1%	1	6%	4
	Rio Grande do Sul	RS	5,6%	2		
	Santa Catarina	SC	0,3%	1		

Fonte: elaborado pelos autores

Com o intuito de expandir a sua participação de mercado e reduzir o tempo de resposta aos clientes, bem como os custos de transporte de produto acabado, a empresa deseja abrir uma nova planta. Devido a questões estratégicas, a nova fábrica deve ser localizada em uma região de alta concentração de clientes, caso das regiões nordeste e sudeste. Entretanto, o fornecimento de matéria-prima será realizado pela Fábrica 1, de acordo com o fluxo ilustrado na Figura 2.

Figura 2 - Fluxo de matéria-prima para a nova fábrica



Fonte: elaborado pelos autores

Sob a perspectiva geométrica, a princípio o caminho mais viável seria a Fábrica 1 atender o cliente selecionado para o exemplo, pois a indicação do trajeto de transporte de matéria-prima sugere um esforço logístico adicional. Entretanto, sabe-se que o transporte de um mesmo volume de matéria-prima tem preço inferior ao transporte de produto acabado, devido ao fracionamento de cargas, ao valor agregado, aos riscos atrelados e as taxas e tarifas inerentes, além da menor economia de escala do modal quando da utilização de carga paletizada, ou do custo de transporte refrigerado de produtos acabados.

Por isso, utiliza-se o parâmetro r , descrito em (5), a fim de considerar diferentes cenários, por meio da variação da razão entre os custos de transporte nos diferentes estágios da cadeia. A Tabela 1 traz os resultados:

Tabela 1 - Resultados obtidos para diferentes valores de r

Nível	Razão C_{MP}/C_{PF}	Localização da nova Fábrica	Número de cidades de cada cluster	Divisão do volume F1-F2
-------	-----------------------	-----------------------------	-----------------------------------	-------------------------

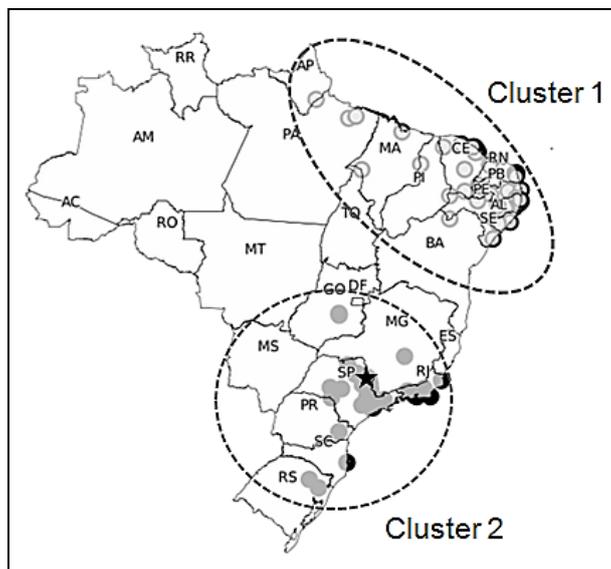
N1	0,01 – 0,43	Itapira-SP	43-48	43,6 – 56,4
N2	0,44	Itapira-SP	44-47	43,6 – 56,4
N3	0,45-0,73	Itapira-SP	45-46	43,8 – 56,2
N4	0,74	Boituva-SP	47-44	58,3 – 41,7
N5	0,75	Boituva-SP	48-43	58,3 – 41,7
N6	0,76	Piedade-SP	51-40	65,5 – 34,5
N7	0,77	Juquiá-SP	55-36	68,7- 31,3
N8	0,78	Juquiá-SP	56-35	68,7- 31,3
N9	0,79	Tapiraí-SP	58-33	69,3 – 30,7
N10	0,80 - 0,86	Ipê-RS	88-3	93,7 – 6,3
N11	0,87-0,98	Vila Flores-RS	89-2	94,4 – 5,6
N12	0,99	Nova Prata-RS	90-1	94,9 – 5,1
N13	1	-	91-0	100 - 0

Fonte: elaborado pelos autores

Dada a distribuição geográfica dos atuais clientes com quem a empresa tem contrato de fornecimento de médio e longo prazo, bem como os volumes entregues no período em análise, a localidade indicada para 73% dos casos é a cidade de Itapira-SP, situada na mesorregião de Campinas. A variação centesimal de r deu margem a 13 diferentes níveis de formação de clusters. Um caso particular é a formação N13, que não apresentou convergência, o que significa que a solução indicada do ponto de vista logístico é expandir a capacidade da planta atual, ao invés de abrir uma nova instalação.

A Figura 3 mostra o resultado da clusterização nos cenários em que o Cluster 1 engloba 43 cidades, e o Cluster 2, 48, resultados observados quando $r \geq 0,01$ e $r \leq 0,43$, a localização da cidade de Itapira-SP é assinalada com uma estrela.

Figura 3 - Formação dos clusters com 43 e 48 cidades



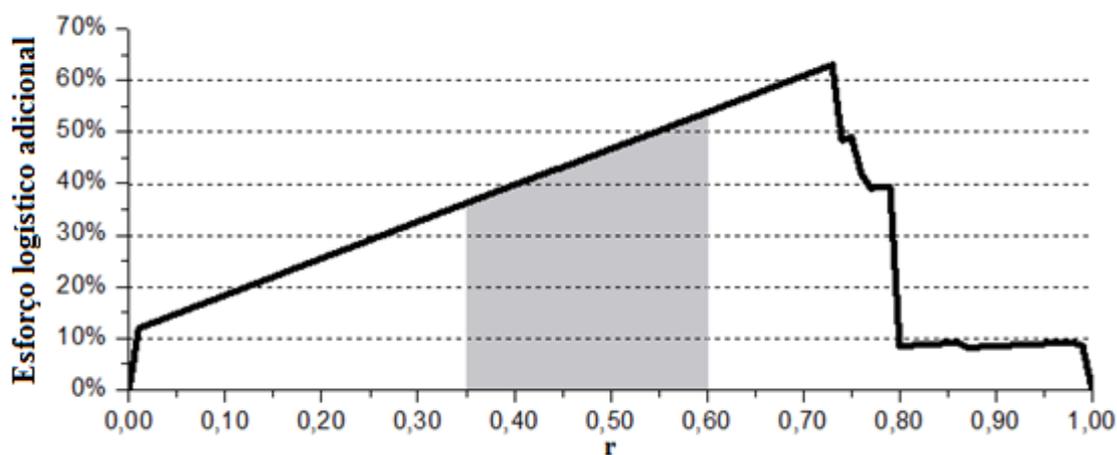
Fonte: elaborado pelos autores

Entretanto, considerando a aplicação do modelo estudado, a tomada de decisão deve ser baseada em cenários em que r assume valores entre 0,35 e 0,60, faixa preenchida em cinza, pois os valores do transporte de matéria-prima e de produto final não possuem acentuada discrepância entre si, fato que faz com que os cenários extremos possam ser desconsiderados.

Além disso, devido às questões operacionais, é desejável que a capacidade produtiva liberada na Fábrica 1 esteja dentro da faixa de 40 e 60%, condição que não é satisfeita pelos cenários com $r \geq 0,74$.

A Figura 4 mostra o acréscimo percentual do esforço logístico que a variação de r incide em comparação com o cenário inicial, dada a formação dos clusters para cada cenário.

Figura 4 – Esforço logístico adicional em relação ao cenário inicial em função de r



Fonte: elaborado pelos autores

Dada a sua maior probabilidade de ocorrência, seleciona-se como referência para análise o caso em que $r = 0,40$. Observa-se um acréscimo de 39,8% nos custos logísticos devido a nova configuração, a qual inclui a transferência de matéria-prima. No entanto, esse valor pode ser assumido como aceitável ao considerar a estratégia de expansão da empresa, com o intuito de alcançar novos clientes, os ganhos de qualidade de serviço e produto, a redução do tempo de resposta ao cliente, e por conseguinte, as reduções da possibilidade de multas por atraso, de perdas por deterioração e de gastos com transporte refrigerado.

5. Considerações finais

A formulação proposta para a abordagem do problema mostrou a aplicabilidade do método de clusterização K-means quanto à designação dos clientes a serem atendidos por cada fábrica com base em seu município de localização, fornecendo consistente base para o problema de cobertura, e em um contexto prático, ainda que indiretamente para o problema de roteamento de veículos.

Por sua vez, o parâmetro r adotado para a representação da análise da segunda fábrica mostrou que sua influência é dependente da distribuição geográfica dos pontos em análise. Já a agregação de dados em termos de produtos e municípios revelou-se uma estratégia satisfatória e conveniente para o propósito do trabalho, reduzindo o tempo de resolução do problema e possibilitando o uso de uma população de 91 cidades clientes sem agregações regionais.

É válido citar que a localização indicada, bem como sua mesorregião, surge como uma boa alternativa, ao passo que a determinação da localização de uma nova instalação deve contemplar outras questões de cunho qualitativo, aqui não contempladas por questões de escopo do trabalho.

Este artigo não tem o intuito de indicar uma localização ótima para a nova instalação, uma vez que o modelo desenvolvido leva em consideração variáveis exclusivamente quantitativas e não contempla aspectos mercadológicos e vantagens locais, tais como facilidade de mão de obra e localização de fornecedores.

No entanto, os resultados obtidos servem de subsídio para o planejamento estratégico, sinalizando uma região factível para a instalação de uma nova planta, de modo a favorecer os objetivos de desempenho de qualidade e flexibilidade pretendidos pela empresa a partir da redução das distâncias, e por conseguinte, dos tempos de resposta ao cliente.

Como sugestões para trabalhos futuros tem-se a aplicação da formulação para casos com um maior número de clusters, inclusão de novos fatores atrelados aos custos logísticos, cenários com a prospecção de novos clientes e a análise estendida da cadeia incluindo os principais fornecedores.

6. Referências

- AHUMADA, Omar; VILLALOBOS, J. Rene. Application of planning models in the agri-food supply chain: A review. **European Journal of Operational Research**, v.1, n.195, p. 1-20, 2009.
- AKKERMAN, R.; FARAHANI, P.; GRUNOW, M. Quality, safety and sustainability in food distribution: a review of quantitative operations management approaches and challenges. **OR Spectrum**, v. 32, n., p. 863-904, 2010.
- ALBAREDA-SAMBOLA, Maria; FERNÁNDEZ, Elena; SALDANHA-DA-GAMA, Francisco. A leader-follower game in competitive facility location. **Journal of Computers & Operations Research**, v.1, n.39, p. 437-448, 2012.
- BALCAN, Maria Florina; EHRLICH, Steven; LIANG, Yingyu. Distributed k-Means and k-Median Clustering on General Topologies. **Advances in Neural Information Processing Systems**, n. 26, v.1, 2013
- BALLOU, Ronald H. Unresolved Issues in Supply Chain Network Design. **Information Systems Frontiers**, n. 3, v.4, p.417-426, 2001.
- BALLOU, Ronald H. Gerenciamento da cadeia de suprimentos / logística empresarial. Tradução de Raul Rubenich. 5. ed. Porto Alegre, RS: Bookman, 2006.
- CHRISTOPHER, Martin. (2010). Logistics and supply chain management, creating value-adding networks. 4 ed. United Financial Times/ Prentice Hall, Harlow.
- ESNAF, Sakir.; KÜÇÜKDENİZ, Tarik. A fuzzy clustering-based hybrid method for a multi-facility location problem. **Journal of Intelligent Manufacturing**, v. 20, n. 1, p. 259-265, 2009.

KOSIOR, Jake M.; STRONG, Doug. Supply/demand chain modeling utilizing logistical-based costing. **Journal of Enterprise Information Management**, n. 3,v. 19, p. 346-360, 2006.

KUÇUKAYDIN, Hande; ARAS, Necati; ALTINEL, Kuban. A leader–follower game in competitive facility location. **Journal of Computers & Operations Research**, n. 39, v. 1, p. 437-448, 2012.

MARÍN, Alfredo. The discrete facility location problem with balanced allocation of customers. **European Journal of Operational Research**, n. 210, v. 1, p. 27-38, 2011.

MELO, M. T.; NICKEL, S.; SALDANHA-DA-GAMA, F. Facility location and supply chain management – A review. **European Journal of Operational Research**, n. 196, v. 1, p. 401-412, 2009.

MIGUEL, Paulo Augusto Cauchick. **Metodologia de Pesquisa em Engenharia de Produção e Gestão de Operações**. 2ª Ed Rio de Janeiro: Campus, 2011

NG, Wendy Pei Qin; LAM, Hon Loong. A supply network optimisation with functional clustering of industrial resources. **Journal of Cleaner Production**, n. 71, v. 1, p. 87-97, 2014.

WAGNER, Michael R.; BHADURY, Joy; PENG, Steve. Risk management in uncapacitated facility location models with random demands. **Computers & Operations Research**, v. 36, p. 1002– 1011, 2009.

YAO, Zhishuang; LEE, Loo Hay; JARUPHONGSA, Wikrom; TAN, Vicky; HUI, Chen Fei. Multi-source facility location–allocation and inventory problem. **European Journal of Operational Research**, n. 207, v. 1, p. 750-762, 2010.

YU, Min; NAGURNEY, Anna. Competitive food supply chain networks with application to fresh produce **European Journal of Operational Research**, n. 224, v. 1, p. 273-282, 2013.