

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO



Carolina Cerqueira Le Brun de Vielmond

**Localização ótima de unidades de extração de óleos
vegetais para o suprimento a usinas produtoras de
biodiesel**

Monografia apresentada ao Departamento de Engenharia Elétrica da
PUC/Rio como parte dos requisitos para a obtenção do título de
Especialização em *Business Intelligence*.

Orientadores:

Marley Maria B. Rebuzzi Vellasco

Marco Aurélio C. Pacheco

Rio de Janeiro

22/03/2010



A minha família

AGRADECIMENTOS

A minha família, pelo incentivo ao meu contínuo aperfeiçoamento profissional e pessoal. Ao João que tem um papel muito especial em minha vida e que tem me apoiado em todos os momentos.

Aos professores do curso pelos conhecimentos transmitidos.

Aos meus colegas de trabalho pela ajuda na formulação do problema e nas discussões sobre o trabalho.

RESUMO

Esta monografia apresenta uma solução para a localização ótima de unidades de extração de óleos vegetais para suprimento a usinas produtoras de biodiesel da Petrobras Biocombustível. O problema representado consiste em determinar, a partir de um conjunto de municípios candidatos, quais os melhores locais para a instalação de unidades de extração de óleos vegetais, e além disso, os tipos de oleaginosas a serem esmagadas, as tecnologias de esmagamento adotadas, o porte das instalações e o ano de realização do investimento.

A solução do problema deve considerar premissas, restrições e parâmetros reais estabelecidos pela Petrobras Biocombustível, tais como oferta de matéria-prima, atendimento de demandas, opções de porte de esmagadoras e tecnologias de extração e respectivos custos de instalação, custos de frete e de transformação industriais, entre outros.

Este problema foi resolvido através de um modelo de programação matemática pela equipe de Pesquisa Operacional da Petrobras. O modelo atual, com os dados informados, gera cerca de 48.000 restrições e 760.000 variáveis, sendo 9.780 inteiras. Com este número de restrições e variáveis, o modelo fica muito grande e bastante complexo, devido ao elevado número de variáveis inteiras. O tempo computacional para a solução deste problema é muito elevado. Sendo assim, para que o modelo apresentasse bons resultados em um tempo viável foram necessários tratamentos nos dados de entrada, que consistiam em heurísticas para eliminar previamente alguns candidatos antes de fornecê-los ao modelo. Este tratamento nos dados, se não feito corretamente, pode eliminar candidatos que fazem parte da solução ótima. O objetivo deste trabalho é apresentar um modelo alternativo ao atualmente utilizado baseado em algoritmo genético. Resultados mostraram que o modelo proposto apresenta resultados muito próximos ao modelo atualmente utilizado, encontrando soluções viáveis em pouco tempo de execução e sem necessidade de realizar pré-processamento nos dados de entrada.

ABSTRACT

This paper presents a solution to the problem of determining an optimum location for units of vegetable oil extraction, in order to supply Petrobras Biocombustível biodiesel plants. The problem consists in finding the best places to install the vegetable oil extraction units from a set of candidate cities; the types of raw materials to be processed; the extraction technologies to be adopted; the size of the facilities; and the year when the investment will be realized.

The solution should consider assumptions, constraints, and actual parameters established by Petrobras Biocombustível, such as: supply of raw materials; demand fulfillment; size, extraction technologies and installation costs for the facilities; freight and industrial processing costs, among others.

The problem was solved using a mathematical programming model made by the Operational Research team at Petrobras. The current model generates about 48,000 constraints and 760,000 variables, from which 9780 are integers with the reported data. The model gets very large and complex with this great amount of constraints and variables, due to the large number of integer variables. The computational time needed for solving this problem is very high. Thus the input data needed some heuristics to eliminate some candidates in advance before providing them to the model to present good results in a timely manner. This treatment could eliminate candidates who are part of the optimal solution if not done properly. The goal of this paper is to present an alternative model to the one being currently used, based on genetic algorithms. Results showed that the proposed model produces results that are very close to the ones provided by the current model. The new model finds viable solutions in a short amount of time and with no need for input data preprocessing.

SUMÁRIO

| | | |
|------|--|----|
| 1. | INTRODUÇÃO | 7 |
| 1.1. | MOTIVAÇÃO..... | 8 |
| 1.2. | OBJETIVOS DO TRABALHO..... | 10 |
| 1.3. | DESCRIÇÃO DO TRABALHO | 11 |
| 1.4. | ORGANIZAÇÃO DA MONOGRAFIA..... | 11 |
| 2. | DESCRIÇÃO DO PROBLEMA | 12 |
| 3. | ESTADO DA ARTE | 18 |
| 3.1. | FUNCIONAMENTO DO ALGORITMO GENÉTICO | 18 |
| 3.2. | MECANISMOS DOS ALGORITMOS GENÉTICOS..... | 20 |
| 3.3. | REPRESENTAÇÃO DAS SOLUÇÕES | 22 |
| 3.4. | OPERADORES GENÉTICOS | 23 |
| 3.5. | CONVERGÊNCIA DO ALGORTIMO GENÉTICO | 26 |
| 4. | ARQUITETURA DO SISTEMA PROPOSTO..... | 28 |
| 4.1. | REPRESENTAÇÃO DO ALGORITMO GENÉTICO | 29 |
| 4.2. | DECODIFICAÇÃO DA SOLUÇÃO | 30 |
| 4.3. | FUNÇÃO DE AVALIAÇÃO | 32 |
| 4.4. | TECNOLOGIAS ADOTADAS NA IMPLEMENTAÇÃO DO MODELO | 39 |
| 5. | RESULTADOS | 40 |
| 5.1. | MODELO DE PROGRAMAÇÃO INTEIRA MISTA..... | 40 |
| 5.2. | MODELO DE PROGRAMAÇÃO INTEIRA MISTA COM PRÉ- PROCESSAMENTO PARA REDUÇÃO DO NÚMERO DE LOCAIS CANDIDATOS E PONTOS DE OFERTA..... | 40 |
| 5.3. | MODELO UTILIZANDO ALGORTIMO GENÉTICO | 42 |
| 6. | CONCLUSÕES E TRABALHOS FUTUROS | 48 |
| | REFERÊNCIAS BIBLIOGRÁFICAS | 50 |

1. INTRODUÇÃO

Este trabalho tem como objetivo determinar uma solução otimizada para localização de unidades de extração de óleos vegetais, também denominadas neste trabalho como esmagadoras, para suprimento a usinas produtoras de biodiesel da Petrobras Biocombustível. O problema consiste em determinar, a partir de um conjunto de municípios candidatos, quais os melhores locais para a instalação de unidades de extração de óleos vegetais. Além disso, deve-se também determinar os tipos de oleaginosas a serem esmagadas, as tecnologias de esmagamento adotadas, o porte das instalações e o ano de realização do investimento. A Figura 1 mostra o esquema geral do problema, onde a matéria-prima (oleaginosas) deve ser transportada para as unidades de extração de óleos vegetais e os óleos produzidos devem ser entregues nos pontos de demanda. Vale ressaltar que é importante que a solução do problema contenha os fluxos de movimentação da matéria-prima e dos óleos produzidos.

A solução do problema deve considerar premissas, restrições e parâmetros reais estabelecidos pela Petrobras Biocombustível, tais como oferta de matéria-prima, atendimento de demandas, opções de porte de esmagadoras e tecnologias de extração e respectivos custos de instalação, custos de frete e de transformação industriais, entre outros.

No problema apresentado pela Petrobras Biocombustível, foram considerados cerca de 1.400 municípios produtores de oleaginosas, localizados em estados da região Nordeste e em Minas Gerais. Foram avaliados 96 municípios potenciais para a instalação das esmagadoras, selecionados de acordo com critérios de localização, densidade populacional e IDH (Índice de Desenvolvimento Humano). Em termos de tecnologia de esmagamento, foram consideradas duas: prensagem mecânica e extração por solvente, cada uma delas com rendimentos e custos diferenciados. A seleção da capacidade das esmagadoras foi feita tendo como base cinco portes possíveis variando entre 5.000 e 80.000 toneladas de óleo produzido por ano, cada uma das capacidades com custos operacionais e de investimento distintos. São considerados neste problema dois tipos de matéria-prima e quatro tipos de produtos finais (produzidos pelas unidades de extração de óleos vegetais).

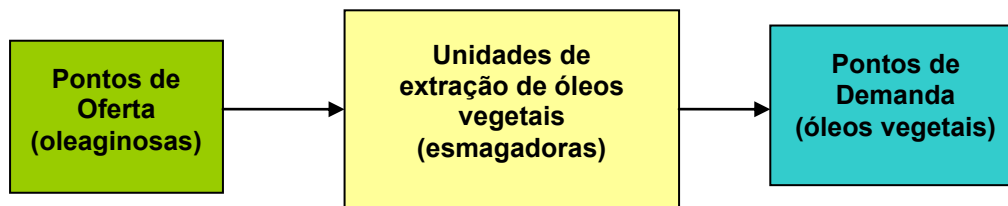


Figura 1: Esquema geral do problema

Este problema foi resolvido através de um modelo de programação matemática (Taha, 2007) pela equipe de Pesquisa Operacional da Petrobras. Este modelo, com os dados informados, gera cerca de 48.000 restrições e 760.000 variáveis, sendo 9.780 inteiras. Com este número de restrições e variáveis, o modelo fica muito grande e bastante complexo, devido ao elevado número de variáveis inteiras. O tempo computacional para a solução deste problema é muito elevado. Sendo assim, tratamentos nos dados de entrada foram necessários para que o modelo apresentasse bons resultados em um tempo viável. Os tratamentos nos dados de entrada consistiam em heurísticas para: (1) eliminar candidatos; (2) reduzir o número de variáveis de fluxo de matéria-prima.

Este tipo de problema é classificado como problema de localização e é muito estudado na literatura. O presente trabalho tem como objetivo apresentar um modelo alternativo, utilizando algoritmo genético, que resolva o problema sem que haja a necessidade de tratar os dados de entrada e que apresente resultados satisfatórios em tempo viável.

As subseções a seguir descrevem a motivação deste trabalho, os objetivos, a descrição e a organização deste trabalho.

1.1. MOTIVAÇÃO

O problema a ser tratado é conhecido como Problema de Localização (“*Location Problem*”) e tem sido muito estudado na literatura. Foi formalmente introduzido em 1909 por Alfred Weber (Weber, 1909), que considerou o problema de localização de uma única planta para minimizar a distância total de viagem entre a planta e um conjunto de clientes distribuídos espacialmente. Posteriormente a este trabalho, outros autores trabalharam em cima deste tema e foram identificadas diferentes classes deste problema, que foram resolvidos e aplicados em uma variedade de aplicações práticas. Um resumo com mais de 50 tipos de problemas de localização pode ser visto em (Brundeau, 1989).

Segundo (Brundeau, 1989), o problema de localização é um problema de alocação de recursos no espaço. Em geral, um ou mais provedores servem a um conjunto de pontos de demanda (clientes) distribuídos espacialmente. O objetivo é determinar a localização dos provedores (e talvez, alocar os clientes aos provedores) para otimizar um objetivo explicitamente ou implicitamente dependente da localização espacial. Um critério típico de decisão inclui: minimizar o tempo médio de viagem ou a distância entre as demandas e provedores; minimizar o tempo de resposta; minimizar uma função de custo de viagem ou de tempo de resposta; dentre outros.

A classe de problemas de localização no qual o problema a ser tratado neste trabalho se encaixa é no problema simples de alocação de planta (*"Simple plant location problem"*). Uma boa referência a este problema e suas extensões pode ser visto em (Krarup, 1981). Em resumo, este problema é uma versão particular do problema de localização, que procura selecionar uma quantidade indeterminada de provedores a partir de um número finito de potenciais localizações com o objetivo de minimizar o custo fixo das plantas mais o custo variável de transporte. Extensões deste problema incluem restrições de capacidade nas plantas, tratamento de mais de um tipo de produto nas plantas, determinação da localização das plantas ao longo do tempo (visando maximizar o retorno durante um período de tempo), determinar os fluxos dos produtos dos provedores aos clientes; dentre outros. O problema a ser tratado neste trabalho inclui diversas destas extensões: as esmagadoras são capacitadas; a instalação das esmagadoras deve ser feita ao longo do tempo, dentro do período de análise; são tratados dois tipos de matéria-prima e quatro tipos de produtos finais; a esmagadora pode utilizar dois tipos de tecnologia; necessidade de determinação dos fluxos de matéria-prima e produto final; dentre outros.

Este tipo de problema sofre de explosão combinatória e em geral é do tipo NP-Difícil, onde na maioria dos casos, o tempo e recursos computacionais necessários para resolver estes problemas em aplicações práticas são proibitivos.

A literatura sobre este tipo de problema é muito rica, pesquisadores têm trabalhado tanto em soluções utilizando heurísticas quanto em algoritmos exatos. Segundo (Sridharan, 1995), enquanto algoritmos exatos podem resolver problemas de porte médio, com até 50 plantas e 50 clientes, com um razoável esforço computacional, heurísticas são necessárias para resolver problemas com centenas de plantas e clientes. O problema a ser tratado neste trabalho

pode ser considerado como de grande porte, dado que possui 1.400 pontos de oferta de oleaginosas e 96 municípios candidatos.

Algumas das heurísticas mais conhecidas são a Busca Tabu, *Simulated Annealing* e Algoritmos Genéticos. O trabalho (Arostegui et al., 2005) compara estas três técnicas para este tipo de problema e concluiu que a Busca Tabu apresenta bons resultados na maioria dos casos, porém Algoritmos Genéticos obteve melhores resultados nos problemas que possuem mais variáveis de decisão (multi-período e *multi-commodities*) além de apresentar “boas” soluções mais rapidamente do que as demais técnicas.

São poucos os trabalhos que mostram a aplicação de algoritmos genéticos neste tipo de problema. (Jaramillo, 2002) aplicou algoritmos genéticos para resolver um conjunto de problemas de localização e reportou que o comportamento do algoritmo genético foi diferente para cada tipo de problema. Para os três tipos de problema estudados, o algoritmo genético tende a demorar mais tempo para solucionar do que heurísticas especializadas. Porém o algoritmo genético não apresentou soluções piores e, em algumas vezes, apresentou soluções melhores do que com outros métodos da literatura. E mesmo em testes com limite de tempo, o algoritmo genético apresentou “boas” soluções, indicando que ele evolui rapidamente em direção à solução ótima, mas que gasta muito tempo tentando melhorar a solução. (Kratka, 2001) utilizou um algoritmo genético para o problema simples de localização de planta e conseguiu resolver um problema com mais de 1000 plantas e clientes, comparando o resultado com a técnica de programação matemática. (Syarif, 2002) usou um algoritmo genético para resolver problemas envolvendo localização de plantas e alocação de clientes às plantas e reportou resultados satisfatórios. (Chaudry, 2003) aplicou um algoritmo genético para um problema relativamente grande e concluiu que foi possível obter resultados ótimos para a maioria dos problemas, contudo em alguns casos o tempo necessário para resolver o problema foi maior do que 1 hora.

1.2. OBJETIVOS DO TRABALHO

O principal objetivo do presente trabalho foi desenvolver um modelo baseado em algoritmo genético que apresente resultados satisfatórios em tempo viável, como alternativa ao modelo atualmente utilizado.

1.3. DESCRIÇÃO DO TRABALHO

Neste trabalho foram realizadas as seguintes etapas para o desenvolvimento do modelo proposto: estudo do problema e do modelo atualmente utilizado para resolvê-lo; modelagem do problema utilizando algoritmo genético; implementação do modelo e realização de testes comparativos entre o modelo proposto e o atualmente utilizado.

1.4. ORGANIZAÇÃO DA MONOGRAFIA

Este trabalho está dividido em cinco seções adicionais, descritas a seguir:

A seção 2 apresenta a descrição do problema e o modelo atualmente utilizado para resolver o problema.

A seção 3 apresenta o estado da arte da técnica empregada no modelo proposto neste trabalho.

A seção 4 descreve a arquitetura do sistema proposto.

A seção 5 detalha os resultados obtidos.

Finalmente, a seção 6 descreve as conclusões do trabalho e identifica possíveis trabalhos futuros.

2. DESCRIÇÃO DO PROBLEMA

O problema a ser tratado no presente trabalho consiste em indicar propostas de arranjo para futuras instalações de unidades de esmagamento de oleaginosas para a produção de biodiesel levando em conta aspectos de eficiência agrícola, logística e industrial, contemplando-se as seguintes análises: tipos de oleaginosas a serem esmagadas, tecnologias de esmagamento adotadas, porte das instalações, número de unidades e a localização das mesmas, além do ano de realização do investimento.

Para atingir tal objetivo, é necessário um modelo computacional para determinação de soluções otimizadas de localização das unidades. O modelo deve considerar premissas, restrições e parâmetros reais, tais como oferta de matéria-prima, atendimento de demandas, opções de porte de esmagadoras e tecnologias de extração, custos de frete e de transformação industriais, entre outros.

No problema apresentado pela Petrobras Biocombustível, foram considerados cerca de 1.400 municípios produtores de oleaginosas, localizados em estados da região Nordeste e em Minas Gerais. São considerados dois tipos de matéria-prima: mamona e girassol.

Para a instalação das esmagadoras foram avaliados 96 municípios potenciais. Estes municípios foram selecionados de acordo com critérios de localização, densidade populacional e IDH (Índice de Desenvolvimento Humano). A Figura 2 apresenta a distribuição espacial dos locais candidatos. Pode-se verificar que são muitos candidatos e que alguns estão muito próximos.

Em termos de tecnologia de esmagamento, foram consideradas duas: prensagem mecânica e extração por solvente, cada uma delas com rendimentos e custos diferenciados. O rendimento de cada tecnologia de esmagamento também varia de acordo com o tipo de oleaginosa utilizada.

Diferentes capacidades para as esmagadoras foram elencadas para se conhecerem então os investimentos necessários para implantação de unidades de pequeno, médio e grande porte. Foram selecionadas as capacidades de 5.000 e 10.000 t óleo/ano, para as unidades de pequeno porte; 25.000 t óleo/ano, para as de médio porte; e 50.000 e 80.000 t óleo/ano, para as extratoras de grande porte. Cada uma das capacidades possui custos

operacionais fixos e variáveis (depende da quantidade de óleo produzida) e de investimento distintos. Estes custos também variam de acordo com a tecnologia adotada.

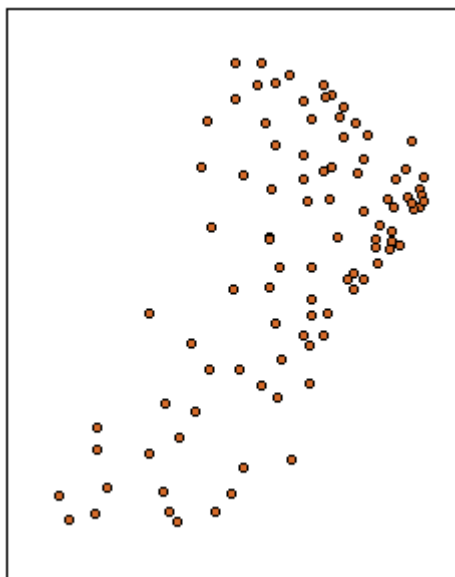


Figura 2: Distribuição espacial dos locais candidatos à instalação das esmagadoras

É premissa do modelo que toda a oferta de oleaginosas deve ser processada, e que a matéria-prima não pode ser transportada para uma esmagadora que esteja localizada a mais de 800 km de distância do ponto de oferta. Uma vez levados às esmagadoras, os grãos produzem tipicamente óleo e torta/farelo. O óleo pode ser vendido às unidades produtoras de biodiesel da própria Petrobras, ou ainda no mercado interno e externo. A torta de mamona é bastante utilizada como adubo orgânico e, a torta de girassol pode ser usada como matéria-prima para rações de ruminantes em geral. O segmento de torta é muito regionalizado e possui perfil de negociação preferencialmente FOB (comprador retira na fábrica). Sendo assim, deve-se considerar que toda a torta produzida é absorvida pelo mercado, gerando receita, mas sem implicar em custos de transporte. Com relação aos óleos produzidos, devem ser considerados os custos de frete de transporte. Alguns pontos de demanda possuem um mínimo para ser atendido, como por exemplo, as unidades produtoras de biodiesel. O óleo excedente pode ser vendido no mercado interno ou levado para os portos, para exportação. Sendo assim, além das unidades produtoras de biodiesel, foram incluídos pontos de demanda, para representar os portos e os pontos de venda de óleo para a indústria nacional. Para o escoamento dos produtos finais não foi considerada distância máxima limite.

O período de análise considerado foi de 2011 a 2020, com a possibilidade de investimento em qualquer ano do cenário. A distância entre os pontos de oferta, locais candidatos a instalação de esmagadoras e pontos de demanda foi calculada considerando a menor distância rodoviária, segundo um algoritmo de caminho mínimo.

A solução proposta pela área de Pesquisa Operacional da Petrobras foi o desenvolvimento de um modelo matemático que utiliza a técnica de programação matemática (Taha, 2007), uma das mais utilizadas em Pesquisa Operacional. Para resolver o problema foi utilizado um sistema gerador de modelos desenvolvido internamente pela equipe e o otimizador CPLEX (IBM, 2009).

A técnica de programação matemática surgiu na década de 40, motivada pela necessidade de se resolver problemas de planejamento de operações militares e houve uma crescente aplicação em diversos setores da indústria no período do pós-guerra. Atualmente, a programação matemática é utilizada para resolver diversos problemas, como: alocação ótima de recursos; análise de investimentos; planejamento e operação da produção; e problemas de rede (transporte, telecomunicação, etc). Este problema se encaixa numa classe importante de problemas estudados na programação matemática que é a dos problemas de Programação Inteira / Mista. É dito problema de programação inteira aquele que possui uma ou mais variáveis de decisão que só assumem valores inteiros.

Neste problema, a variável de decisão do investimento são do tipo 0 ou 1, que quando iguais a zero indicam que não será feito o investimento e, quando iguais a 1, determinam o investimento em determinado local, faixa de capacidade, tecnologia e período. Para resolver esta classe de problemas o método mais utilizado são os Métodos *Branch-and-Bound* (B&B) (Taha, 2007), onde são resolvidos diversos problemas contínuos, até que se obtenha alguma solução que atenda a integralidade e que seja a solução ótima. A vantagem do B&B é que, apesar de não resolver exaustivamente todos os problemas contínuos possíveis, a árvore de busca é inteiramente pesquisada. Sendo assim, este método garante que a solução encontrada seja a ótima, porém para problemas complexos, com muitas variáveis inteiras, esta busca pela solução ótima pode levar um tempo muito elevado. Segundo (Taha, 2007), embora seja comprovado teoricamente esse algoritmo converge em um número finito de iterações, sua implementação em computador (com seu inerente erro de

arredondamento) pode não encontrar a solução ótima independentemente do tempo em que for executada no computador. Durante as iterações deste algoritmo é possível obter uma indicação da proximidade da solução atual em relação à ótima (estimada) através da distância relativa (*gap*), em termos da função objetivo, entre a melhor solução inteira conhecida e a melhor solução contínua do problema inteiro relaxado. O *gap* é usualmente utilizado com critério de parada da otimização e o seu valor padrão no otimizador CPLEX (IBM, 2009) é 0,01%.

Nesta técnica, o problema é representado através de um conjunto de variáveis e equações, denominadas restrições (vide Tabela 1). Softwares específicos resolvem o sistema de equações, visando maximizar uma equação representativa do Resultado Econômico (a função objetivo). No caso em questão, a função objetivo é a maximização do lucro resultante dos custos logísticos e industriais e de receitas representativas de venda de óleo e torta. Embora o valor desta função não possa ser avaliado de maneira absoluta, é um indicador importante, especialmente para análises comparativas entre diversas soluções possíveis.

O modelo possui todas as suas variáveis positivas e suas equações são todas estritamente lineares e, portanto, apresenta simplificações em relação à realidade, mas que em geral não comprometem a qualidade dos resultados para fins de análise.

O modelo, com os dados informados, gera cerca de 48.000 restrições e 760.000 variáveis, sendo 9.780 inteiras. Com este número de restrições e variáveis, o modelo fica muito grande e bastante complexo, devido ao elevado número de variáveis inteiras. O tempo computacional para a solução deste problema é muito elevado. Este modelo mesmo após uma rodada com tempo limite de 48 horas, não apresentou solução satisfatória, retornando um resultado econômico negativo. Estudos preliminares da Petrobras Biocombustível indicam que existem arranjos de esmagadoras onde o resultado econômico é positivo.

| Variáveis | Restrições |
|---|--|
| Fluxos – de matéria-prima dos pontos de origem para as esmagadoras e de produto final das esmagadoras para os pontos de destino. | Balancos que garantem a compatibilização de ofertas, fluxos, cargas, produção e demanda. |
| Carga de Matéria-prima nas Esmagadoras. | Atendimento aos limites de oferta (100% de atendimento) e de demanda. |
| Produção de Produtos Finais. | Restrições de carga nas plantas, conforme capacidade selecionada para investimento. |
| Variáveis de Decisão de Investimento – são variáveis 0 ou 1 que, quando zero indicam que não será feito o investimento e, quando iguais a 1, determinam o investimento em determinado local, faixa de capacidade, tecnologia e período. | Em cada local, somente é possível colocar planta de uma capacidade e tecnologia por período. |
| | Uma vez selecionada uma tecnologia para a planta em um dado local, quaisquer ampliações realizadas neste local devem ter a mesma tecnologia. |

Tabela 1: Principais variáveis e restrições do modelo atual

Sendo assim, tratamentos nos dados de entrada foram necessários para que o modelo apresentasse bons resultados em um tempo viável. O primeiro tratamento realizado nos dados de entrada foi o agrupamento dos pontos de oferta para a redução do número de variáveis de fluxo. Este agrupamento foi realizado através de outro modelo de programação matemática que minimiza o momento de transporte. De um total de 1.400 municípios ofertantes do problema original, foram selecionados 150 para representarem todo o conjunto. A oferta de cada município será agregada a aquele, dentre os 150 selecionados, no qual o custo de frete é o menor. Uma desvantagem deste procedimento é a perda de precisão no cálculo do frete, dado que o modelo irá enxergar apenas 150

municípios e não os 1.400 originais. Outro tratamento nos dados de entrada foi a eliminação de alguns candidatos. Foram realizadas rodadas preliminares utilizando a maior demanda existente no horizonte de estudo, limitando a instalação de uma faixa de capacidade por vez. Um ranking é realizado com o resultado destas rodadas e os candidatos selecionados mais vezes nas rodadas preliminares são aqueles que serão passados para a rodada final. Com isso, a rodada final ficou limitada a possuir um número menor de locais candidatos, por volta de 15 ao invés de 96. Mesmo assim, esta última rodada necessita de aproximadamente 24 horas para obter um bom resultado. Uma desvantagem desta heurística é a possibilidade de eliminação de candidatos que fazem parte da solução ótima. Maiores detalhes sobre os resultados encontrados com este modelo serão apresentados na seção 5.

3. ESTADO DA ARTE

O algoritmo genético é um método computacional de busca e otimização baseado em mecanismos de seleção natural e genética. Segundo (Davis, 1991), os algoritmos genéticos empregam um processo adaptativo e paralelo de busca de soluções em problemas complexos, com difícil formulação matemática, ou com grande espaço de busca (grande número de soluções).

O algoritmo genético foi proposto inicialmente por John H. Holland em 1975 no trabalho intitulado *Adaptation in Natural and Artificial Systems*. Baseia-se no princípio da sobrevivência dos mais aptos e na reprodução das espécies, inspirado no mecanismo de evolução das espécies apresentado nos trabalhos de Charles Darwin (Darwin, 1859) e na herança genética proposta por Mendel (Mendel, 1865).

De acordo com (Michalewicz, 1995), os algoritmos genéticos são particularmente adequados para problemas complexos de otimização, de difícil modelagem matemática como aqueles que apresentam diversos parâmetros ou características que precisam ser combinadas em busca da melhor solução; problemas com muitas restrições ou condições que não podem ser representadas matematicamente; e problemas com grandes espaços de busca. Este método tem sido aplicado em diversas áreas tais como: Otimização de Funções Matemáticas, Otimização Combinatorial, Otimização de Planejamento, Problema do Caixeiro Viajante, Problema de Otimização de Rota de Veículos, Otimização de Layout de Circuitos, etc.

3.1. FUNCIONAMENTO DO ALGORITMO GENÉTICO

Fundamentalmente, a inspiração biológica resulta, no contexto computacional, em um método de resolução de problemas. Primeiramente é criado um conjunto de soluções possíveis para o problema. Cada solução se torna um indivíduo de uma população cuja evolução será simulada através de mecanismos estocásticos. Os indivíduos mais aptos e “melhores” têm maior probabilidade de sobreviver, de reproduzir e de transmitir suas características (carga genética) às próximas gerações (Goldberg, 1989). A adaptação de cada indivíduo ao meio ambiente é dada através do critério de otimização utilizado no problema que se deseja solucionar. Os membros da nova população sofrem transformação através dos operadores genéticos, para formar novas soluções. Existem, basicamente, dois tipos de transformação dos indivíduos

operacionalizados pelos operadores genéticos: mutação e cruzamento, que serão detalhados adiante. No entanto, a atuação de ambos cria uma nova população formada, na média, por indivíduos mais adaptados do que na geração anterior. Depois de algumas gerações o programa converge esperando-se que a melhor solução individual represente a solução próxima do ótimo global.

O processo de resolução de problemas envolve os seguintes elementos:

- Especificação das informações relevantes do problema a ser resolvido.
- Caracterização das informações que definem uma solução para o problema.
- Codificação computacional da solução do problema.
- Identificação dos critérios que permitem a avaliação da qualidade de uma solução.
- Construção de uma Função de Avaliação, que permite expressar quão boa é uma solução.

A Figura 3 apresenta um fluxograma com o funcionamento de um algoritmo genético:

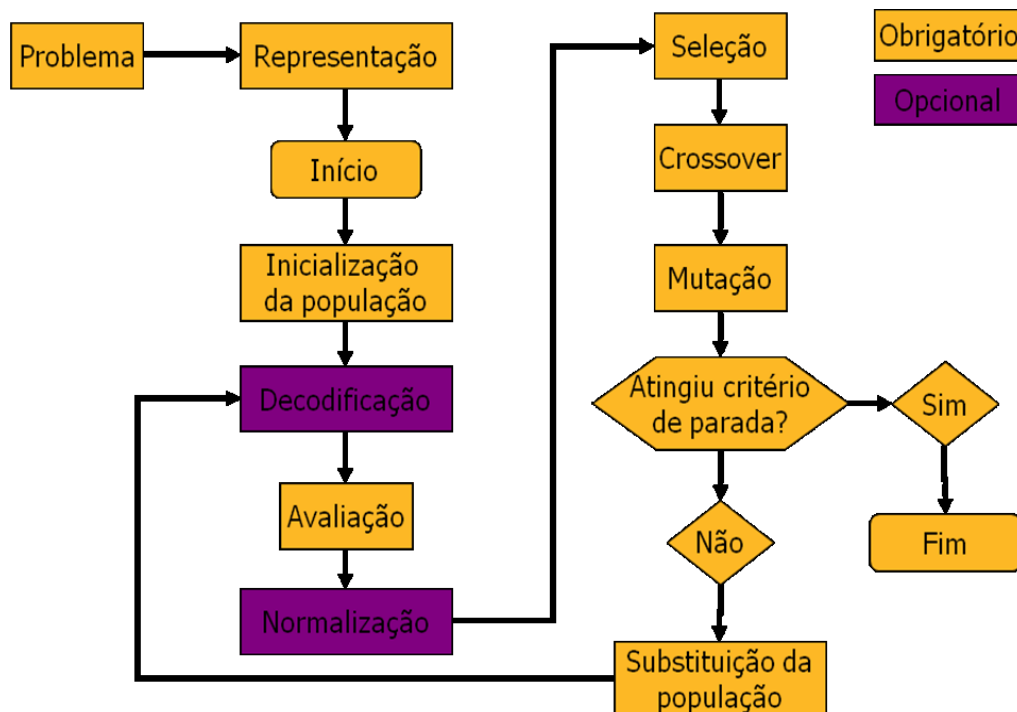


Figura 3: Fluxograma do funcionamento de um algoritmo genético

Os processos descritos no fluxograma da Figura 3 podem ser resumidos por:

- **Representação:** Definição dos parâmetros do algoritmo tais como o tamanho da população, a regra de parada e a representação dos cromossomos.
- **Inicialização da população:** Criação de uma população de N soluções para o problema. Cada solução é gerada aleatoriamente.
- **Decodificação:** Decodificar o cromossomo.
- **Avaliação:** Avaliar cada candidato através da função de avaliação. Quanto maior a função de avaliação melhor será a solução.
- **Normalização:** Realizar a normalização se necessário.
- **Seleção:** Selecionar probabilisticamente, dois indivíduos da população atual. A probabilidade de um indivíduo ser escolhido deve ser diretamente proporcional a sua função de avaliação. Ou seja, as melhores soluções têm maiores chances de serem selecionadas.
- **Crossover:** Recombinar pares de pais (indivíduos selecionados no passo anterior) - cruzamento (*crossover*) - (utiliza mecanismos de probabilidade).
- **Mutação:** Modificar os filhos - mutação (*mutation*) - (utiliza mecanismos de probabilidade);
- **Critério de parada:** Avaliar novos candidatos;
 - Se atingiu o critério de parada, finaliza;
 - Senão, seleciona os indivíduos para a próxima geração (substituição da população).

3.2. MECANISMOS DOS ALGORITMOS GENÉTICOS

Os algoritmos genéticos utilizam mecanismos que garantem a busca de uma boa solução em todo o espaço de soluções possíveis, a saber:

- **Aptidão:** É a quantificação da adaptação do indivíduo (cromossomo) ao ambiente. Ela mede o grau de adaptação de cada indivíduo através da função a ser otimizada. Portanto, se o problema a ser resolvido é de maximização, o indivíduo mais adaptado é aquele que apresenta o maior

valor da função quando substituído na mesma. Para a minimização o conceito se aplica para os indivíduos que reduzem o valor da função.

Definido o conceito de aptidão, os operadores podem ser aplicados, caracterizando o funcionamento e evolução da técnica:

- Os indivíduos da população, candidatos à solução são avaliados através da substituição na função avaliação.
 - Os resultados obtidos indicam o grau de adaptação de todos os indivíduos.
 - Os indivíduos são hierarquizados de acordo com o critério de otimização desejado (minimização, maximização).
 - Escolha dos cromossomos (indivíduos) sobre os quais serão aplicados os operadores genéticos para a determinação da próxima geração. Este mecanismo é denominado seleção e é usado o grau de aptidão de cada indivíduo em relação à aptidão média da população (mecanismo probabilístico).
 - Definição da próxima geração de cromossomos (indivíduos) através da aplicação de operadores genéticos nos indivíduos escolhidos na etapa de seleção como “pais”. É importante ressaltar que dada a natureza probabilística da seleção, alguns indivíduos menos aptos poderão ser selecionados.
- **Aplicação dos Operadores Genéticos:** Os operadores genéticos utilizados para garantir a geração dos novos cromossomos (soluções do problema) são:
 - **Reprodução** – Consiste na duplicação dos cromossomos mais adaptados na composição de uma nova geração.
 - **Cruzamento (Crossover)** – Consiste na composição do cromossomo de solução a partir do cruzamento dos genes existentes nos cromossomos dos pais. Sua tarefa é explorar a bagagem genética existente nos pais, ou seja, uma Busca Local. É dado por um valor percentual, que representa a probabilidade de ocorrência de cruzamento, denominada taxa de cruzamento.
 - **Mutação (Mutation)** – Tem a função de repor o material genético perdido em gerações anteriores e também a introdução de material inexistente, promovendo uma busca global. É dado por

um valor percentual, que representa a probabilidade de ocorrência de mutação denominada taxa de mutação.

Logo, o cruzamento busca a solução a partir do conhecimento dos indivíduos já existente (*exploitation*); a mutação promove uma avaliação em áreas do espaço ainda não avaliadas (*exploration*).

Os mecanismos de cruzamento e mutação apresentam sub-mecanismos próprios de operação que devem ser escolhidos de acordo com o problema a ser solucionado. Porém, todos eles são determinados por mecanismos estocásticos (Beasley, 1993). A taxa de cruzamento reflete a probabilidade que os futuros cenários ou “filhos” possuam informações da geração imediatamente precedente (“pais”).

A taxa de mutação reflete a probabilidade de que futuros cenários contenham alguns valores definidos randomicamente. Uma alta taxa de mutação significa que mais mutações ou valores aleatórios de “genes” serão introduzidos na população.

3.3. REPRESENTAÇÃO DAS SOLUÇÕES

A representação na qual as possíveis soluções aparecem dentro dos cromossomos é uma decisão crítica na utilização dos algoritmos genéticos. A dificuldade na identificação da codificação mais adequada para um dado problema é um dos principais fatores que determinam a impropriedade na utilização da técnica de algoritmos genéticos. Segundo (Pacheco, 1999), os métodos de codificação podem ser classificados da seguinte forma:

- **Codificação Binária:** Os genes só podem assumir os valores 0 ou 1. Foi o primeiro tipo de representação em algoritmos genéticos. É o tipo de representação mais simples e que permite uma fácil manipulação dos cromossomos através dos operadores genéticos. Muito utilizada em problemas numéricos por ser fácil de ser transformada em inteiro ou real. Um exemplo de aplicação deste tipo de codificação é em problemas de maximização de funções matemáticas.
- **Codificação com Números Reais:** Os genes podem assumir qualquer valor real (ponto flutuante). É utilizada em problemas numéricos e tem a vantagem de incorporar o conhecimento no domínio do problema. É uma

representação mais adequada em problemas de otimização com variáveis sobre domínio contínuo.

- **Codificação com Permutação de Símbolos:** Os genes podem assumir valores dentro de um alfabeto especificado e não podem apresentar valores repetidos. É uma representação utilizada em problemas onde a busca da solução depende de diversas combinações (ordem) dos elementos considerados. Alguns exemplos de aplicações são: Problema do Caixeiro Viajante, Problemas de Alocação de Recursos e Problema de Coloração de grafos.
- **Codificação com Símbolos Repetidos:** Os genes podem assumir valores dentro de um alfabeto especificado e podem apresentar valores repetidos. É uma representação utilizada em problemas onde a busca da solução consiste em associar valores (de uma lista especificada) a determinados elementos. Estes problemas são classificados como problemas de agrupamento.

A escolha da melhor representação é dada pela natureza do problema estudado. Escolhida a melhor codificação, os mecanismos de reprodução e evolução podem ser aplicados. Porém, quando se alcança a etapa de avaliação, as soluções devem ser substituídas na função de avaliação para o cálculo da sua adequação. Este processo é denominado Decodificação.

3.4. OPERADORES GENÉTICOS

Nesta subseção serão descritos os operadores genéticos mais utilizados para cada tipo de codificação da solução.

- **Crossover:**
 - **Crossover de um ponto:** O *crossover* de um ponto corta os dois genitores em uma posição aleatoriamente escolhida, criando dois novos possíveis descendentes. A Figura 4 ilustra um exemplo deste operador, onde P1 e P2 são os genitores e F1 e F2, os descendentes. Utilizado na codificação binária e com número real.

| | | | | | | |
|-------|---|---|---|---|---|---|
| P_1 | 1 | 0 | 1 | 1 | 0 | 1 |
| P_2 | 0 | 0 | 1 | 1 | 0 | 0 |
| F_1 | 1 | 0 | 1 | 1 | 0 | 0 |
| F_2 | 0 | 0 | 1 | 1 | 0 | 1 |

ponto de corte aleatório ↗

Figura 4: Crossover de um ponto. Fonte: (Pacheco, 2008)

- **Crossover de dois pontos:** Este tipo de *crossover* é semelhante ao *crossover* de um ponto, onde é realizada a recombinação de dois indivíduos a partir de dois pontos escolhidos aleatoriamente. Utilizado na codificação binária e com número real.
- **Crossover de uniforme:** O *crossover* uniforme é capaz de recombinar quaisquer posições entre dois genitores. Este operador utiliza um padrão (palavra binária) escolhido aleatoriamente para designar os bits selecionados em cada gerador na criação dos descendentes. A Figura 5 ilustra um exemplo deste operador, onde P1 e P2 são os genitores e F1 e F2, os descendentes. Utilizado na codificação binária e com número real.

| | | | | | | | |
|--------|---|---|---|---|---|---|---|
| P_1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 |
| P_2 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| Padrão | 1 | 1 | 0 | 1 | 0 | 0 | 1 |
| F_1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 |
| F_2 | 0 | 1 | 0 | 1 | 0 | 1 | 1 |

Figura 5: Crossover uniforme. Fonte: (Pacheco, 2008)

- **Crossover aritmético:** O *crossover* aritmético é uma combinação linear de dois vetores (genitores). Utilizado na codificação com número real.
- **Crossover uniforme baseado em ordem:** Este operador recombina a ordem e a posição dos genes. Utilizado na codificação com permutação de símbolos. Dados os genitores P1 e P2, cria o descendente F1 da seguinte forma:
 - Gera um padrão de bits do comprimento dos genitores;

- Preenche F1, copiando o genitor P1 nas posições em que o padrão é igual a “1”;
- Faz uma lista dos elementos de P1 associados com os bits “0” do padrão;
- Permuta estes elementos de modo que eles apareçam na mesma ordem em que aparecem em P2;
- Preenche as lacunas de F1 com os elementos ordenados no passo anterior;

A Figura 6 apresenta um exemplo deste operador.

| | | | | | | | | |
|----------------|---|---|---|---|---|---|---|---|
| P ₁ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| P ₂ | 8 | 6 | 4 | 2 | 7 | 5 | 3 | 1 |
| Padrão | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 |
| F ₁ | - | 2 | 3 | - | 5 | 6 | - | - |
| F ₂ | 8 | - | - | 2 | - | - | 3 | 1 |

Elementos de P₁ associados a “0”: 1, 4, 7, 8. Ordenados segundo P₂: 8, 4, 7, 1
 Elementos de P₂ associados a “1”: 6, 4, 7, 5. Ordenados segundo P₁: 4, 5, 6, 7

| | | | | | | | | |
|----------------|---|---|---|---|---|---|---|---|
| F ₁ | 8 | 2 | 3 | 4 | 5 | 6 | 7 | 1 |
| F ₂ | 8 | 4 | 5 | 2 | 6 | 7 | 3 | 1 |

Figura 6: Crossover uniforme baseado em ordem. Fonte: (Pacheco, 2008)

• **Mutação:**

- **Mutação de binário:** Realiza a troca cada gene de um cromossomo se o teste de probabilidade for verdadeiro. A Figura 7 ilustra a aplicação do operador de mutação em um cromossomo. Utilizado na codificação binária.

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 1 | 0 | 1 | 1 |
| 1 | 0 | 0 | 1 | 0 | 1 | 0 |



Figura 7: Mutação. Fonte: (Pacheco, 2008)

- **Mutação de real:** Substitui cada número real em um cromossomo por um número real aleatório se o teste de probabilidade for verdadeiro. Utilizado na codificação com número real.

- **Mutação CREEP:** Consiste em modificar o valor do gene ao somar/subtrair um valor aleatório, que varia de acordo com a probabilidade de ocorrer a mutação. Busca uma solução próxima através de ajustes aleatórios em ambas as direções. Utilizado na codificação com número real.
- **Mutação por embaralhamento:** Seleciona aleatoriamente um trecho do cromossomo e o embaralha. na codificação com permutação de símbolos. A Figura 8 mostra um exemplo deste tipo de mutação:

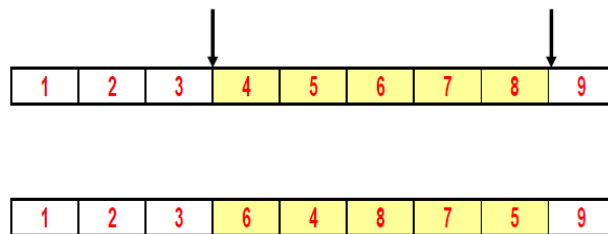


Figura 8: Crossover uniforme baseado em ordem. Fonte: (Pacheco, 2008)

3.5. CONVERGÊNCIA DO ALGORITMO GENÉTICO

O algoritmo genético é um algoritmo de busca que utiliza conceitos de probabilidade para direcionar a busca pela melhor solução. Se for corretamente implementado, a população irá evoluir ao longo de sucessivas gerações de modo que a aptidão do melhor e do indivíduo médio, em cada geração aumenta em direção ao ótimo global. Segundo (Beasley, 1993), a convergência é a progressão para a uniformidade crescente e um gene é dito ter convergido quando 95% da população compartilha o mesmo valor. Diz-se que a população convergiu quando todos os genes convergiram. A Figura 9 mostra evolução típica da aptidão de um algoritmo genético. Conforme a população converge, a aptidão do indivíduo médio evolui para a do melhor indivíduo.

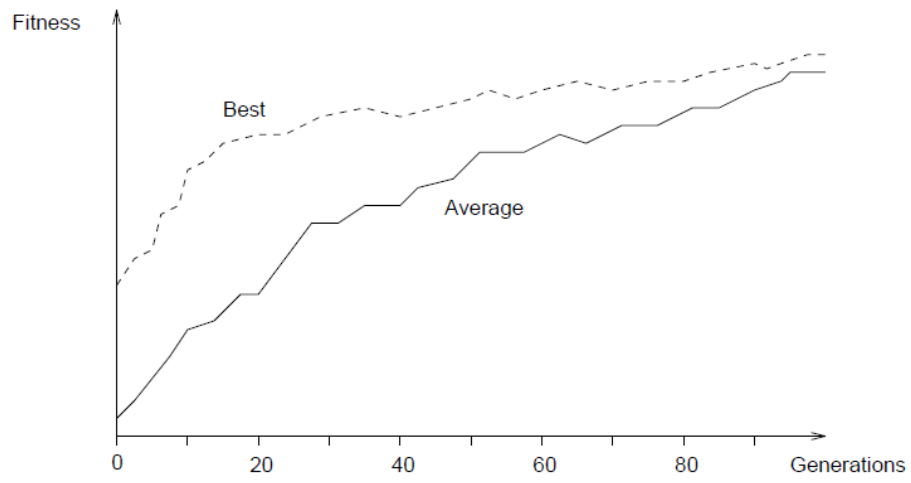


Figura 9: Evolução típica de um algoritmo genético. Fonte: (Beasley, 1993)

4. ARQUITETURA DO SISTEMA PROPOSTO

O sistema proposto consiste basicamente de um programa que implementa a técnica de algoritmo genético. O algoritmo genético é o responsável por explorar o espaço de busca de soluções de possíveis instalações de esmagadoras, definindo além da localização, suas capacidades e tecnologias e o ano de implantação. O módulo responsável pela execução do algoritmo genético comunica-se com um banco de dados que possui os dados de entrada do problema tais como: locais candidatos, tecnologias, capacidades, período de análise, ofertas e demandas, custos de frete e de investimento, dentre outros. Devido às características do problema de localização, não utilizamos o algoritmo genético para determinar os fluxos de matéria-prima e produtos finais. Uma vez que o algoritmo genético encontra uma determinada configuração de instalação das esmagadoras, a determinação dos fluxos de matéria-prima e produtos finais é realizada, de maneira ótima, por um modelo de programação linear (Taha, 2007). Este modelo auxiliar é a função de avaliação das soluções encontradas pelo algoritmo genético. A Figura 10 apresenta um diagrama com a arquitetura proposta para o sistema:

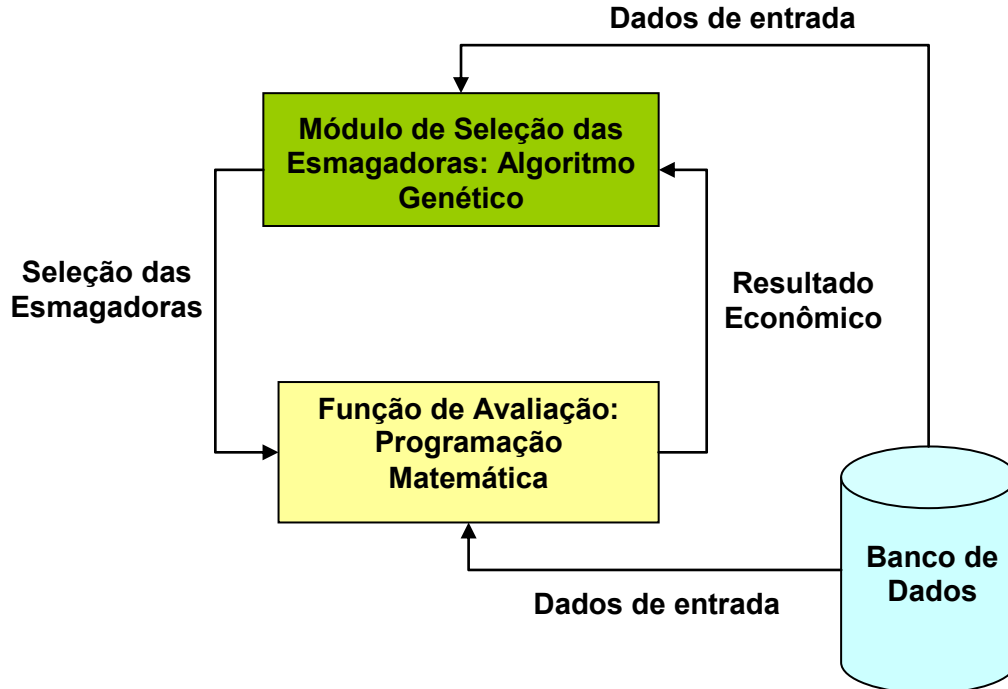


Figura 10: Arquitetura do Sistema

Nas subseções a seguir serão apresentados em detalhes a modelagem do algoritmo genético proposto, da função de avaliação utilizada e as tecnologias adotadas na implementação do modelo.

4.1. REPRESENTAÇÃO DO ALGORITMO GENÉTICO

O modelo de algoritmo genético é o responsável por determinar a localização das esmagadoras dentro de um conjunto de locais candidatos, e, além disso, a tecnologia adotada, capacidade instalada e o ano de instalação de cada uma delas. Para isso a codificação da solução (cromossomo) foi dividida em três partes, cada uma responsável por determinar uma característica da solução. A primeira parte representa a tecnologia adotada em cada esmagadora. A segunda representa o incremento de capacidade das esmagadoras a cada período. E, finalmente, a terceira parte determina a ordem de implantação das esmagadoras em cada período. A Figura 11 apresenta um diagrama representando o cromossomo proposto, onde P_i indica que o segmento corresponde ao período i .

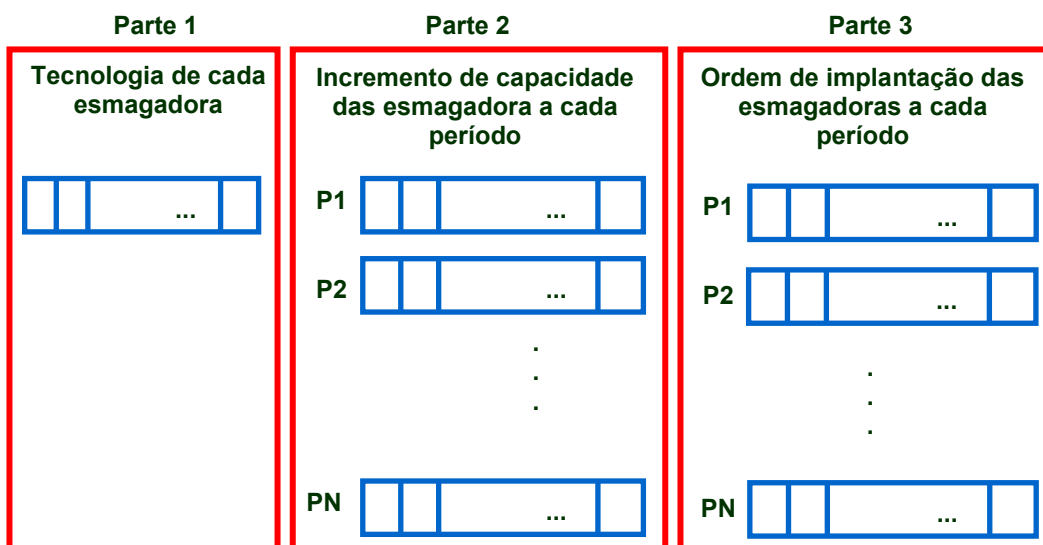


Figura 11: Cromossomo proposto

Como a tecnologia não pode ser alterada após a instalação da esmagadora, cada gene representa a tecnologia de uma esmagadora. No

problema estudado só existem duas opções de tecnologia, então o valor do gene pode assumir os valores 0 ou 1.

É permitido ao modelo ampliar uma instalação já existente, sendo assim, ao invés de representar a capacidade total a cada período das esmagadoras instaladas, indica-se o incremento de capacidade das mesmas. Para isso, é necessário um segmento deste tipo para cada período. Os genes podem assumir as capacidades de instalação disponíveis, que no caso do problema estudado são 5.

A ordem de implantação das esmagadoras será representada por um segmento para cada período. A cada período deve-se considerar a capacidade já instalada no período anterior e verificar se é necessário instalar mais alguma esmagadora para atender a demanda de produto final.

A seguir apresentamos um resumo da representação do GA:

- 1 segmento para cada período com 96 genes (1 para cada local candidato) representando o incremento de capacidade:
 - Valores: 1 a 5 (inteiros)
- 1 segmento para cada período com 96 genes (1 para cada local candidato) com a ordem de implantação das esmagadoras:
 - Valores: alfabeto {1, 2, ..., 96}
- 1 segmento com a tecnologia de cada local candidato:
 - Valores: 0 ou 1 (binário)

Com esta representação temos como resultado um cromossomo com um total de 2.016 genes ($96 \times (10 + 10 + 1) = 96 \times 21 = 2016$).

4.2. DECODIFICAÇÃO DA SOLUÇÃO

O algoritmo de decodificação transforma o cromossomo em uma solução viável de configuração de instalação de esmagadoras. A base central para a decodificação do cromossomo proposto são os segmentos que determinam a ordem de implantação de esmagadoras (Figura 11). Porém, a solução mais econômica não é ampliar todas as esmagadoras em todos os períodos. O que devemos determinar é uma configuração que absorva toda a oferta disponível, atendendo pelo menos a demanda mínima, e que possua menor custo. Sendo assim, a heurística de decodificação da solução utilizada seleciona, a cada período, novas esmagadoras para instalar ou amplia a capacidade das existentes, para atender a demanda daquele período.

A capacidade das esmagadoras é representada pela quantidade de óleo produzido no ano, então se a demanda de um determinado período é maior do que a capacidade já instalada, deve-se realizar uma ampliação ou nova instalação segundo o segmento de ordem de implantação do período estudado. Uma particularidade deste problema impede a determinação exata da demanda a cada período: a quantidade de óleo produzido depende da tecnologia adotada e da oleaginosa que será processada em cada esmagadora. Tanto a tecnologia quanto os fluxos de matéria-prima para cada esmagadora fazem parte da solução do problema, sendo assim, não pode ser determinado a priori. Realizamos então uma estimativa da quantidade máxima de óleo que pode ser produzida a cada período. Como toda oferta tem que ser absorvida, consideramos que cada tipo de oleaginosa será encaminhado para uma esmagadora que possua a tecnologia que proporcione o maior rendimento de óleo. Usamos então esta quantidade máxima que consegue ser produzida como limitante da capacidade instalada de esmagadoras. Testes realizados limitando a capacidade instalada a essa quantidade máxima calculada, mostraram que este valor super-estima a capacidade necessária para processar toda a oferta. Sendo assim, utilizamos uma fração deste valor como limitante da capacidade. Verificou-se que a fração da quantidade máxima que trouxe melhores resultados foi 75%.

Sendo assim, a idéia do algoritmo de decodificação é instalar ou ampliar esmagadoras segundo o segmento de ordem de implantação de esmagadoras, onde a tecnologia é dada pela primeira parte do cromossomo e o incremento de capacidade, pela segunda parte. Apresentamos a seguir o pseudo-código do algoritmo de decodificação:

Algoritmo de decodificação:

```
QuantidadeMaxOleo = soma(Oferta total de cada tipo de
    oleaginosa
                        * Melhor rendimento possível)
CapacidadeAcumulada = 0
Fator = Fração da quantidade máxima estimada através de
    testes

Para cada período P
    Para cada esmagadora E segmento de ordem das esmagadoras
        CapacidadeEscolhida = valor no segmento de incremento
        de capacidade correspondente ao período P e esmagadora
        E

        Se CapacidadeAcumulada >= (QuantidadeMaxOleo *
        Fator) então
            Vá para o próximo período
        Fim Se

        CapacidadeAcumulada += CapacidadeEscolhida
        Capacidade(E,P) = CapacidadeEscolhida
        Investimento(E,P) = 1
    Fim Para
Fim Para
```

Para cada cromossomo, o algoritmo de decodificação determina os locais onde houve investimento a cada período, a tecnologia adotada e o incremento de capacidade. Por fim, esta solução será passada à função de avaliação, que será detalhada na próxima seção.

4.3. FUNÇÃO DE AVALIAÇÃO

Dada uma configuração de instalação de esmagadoras, não é tarefa imediata determinar o resultado econômico desta solução. Isto ocorre, pois mesmo se a configuração for a ótima, se não determinarmos a melhor maneira de transportar os grãos e óleos (para minimizar custos de frete), a melhor distribuição dos grãos nas esmagadoras de acordo com o rendimento (maximizando a receita dos produtos finais), podemos estar valorando erradamente uma determinada solução. A solução adotada para este problema foi utilizar a técnica de programação matemática para determinar a melhor

distribuição dos grãos e óleos dada uma configuração de instalação de esmagadoras. Como o problema é linear, e a técnica de programação matemática nos retorna a solução ótima (Taha, 2007), não estaríamos então avaliando erradamente nenhuma solução encontrada pelo algoritmo genético. Diferentemente do modelo atualmente utilizado, que determina também as configurações das esmagadoras através de um modelo de programação matemática, para a função de avaliação não são necessárias variáveis inteiras. Esta característica permite encontrar a solução ótima com baixo esforço computacional.

O modelo consiste em determinar os fluxos de matéria-prima e produto final entre os pontos de oferta, esmagadoras e pontos de demanda, de maneira a maximizar a receita e minimizar os custos. Existem restrições para o atendimento da demanda mínima, absorção de toda a oferta e de balanço, para garantir que todo grão que entra numa esmagadora é transformado em óleo e que todo óleo produzido é transportado para algum ponto de demanda. Algumas restrições extras foram incorporadas durante os testes para melhorar a solução:

- Não permitir que a produção de uma esmagadora diminua no horizonte: não é desejável que se faça um investimento em um local e que este seja “abandonado” com o tempo, causando problemas sócio-econômicos na região. Esta restrição tem o objetivo de piorar a avaliação de uma solução que instala uma esmagadora com capacidade maior do que a necessária.
- A utilização de todas as esmagadoras no último período deve ser maior do que um determinado valor, i. e. maior do 50 %. Esta restrição tem o mesmo objetivo que a anterior: piorar a avaliação de uma solução que instala uma esmagadora com capacidade maior do que a necessária.
- A restrição de absorção de toda a oferta não é rígida pois o modelo de programação linear poderia retornar como resposta que o problema é inviável, caso fosse passado a ele uma configuração de esmagadoras que não consiga absorver toda a oferta. Uma solução seria assumir um valor muito negativo para o resultado econômico quando ocorresse inviabilidade, porém ao associar um mesmo valor a qualquer solução inviável não cria-se uma distinção entre por exemplo, uma solução que não consiga absorver 1 Kg de matéria-prima de outra que não absorve 10 toneladas. Sendo assim foi introduzida uma variável de folga, que indica

a quantidade de matéria-prima não absorvida pelas esmagadoras e que está sujeita a uma penalidade.

A seguir apresentamos a descrição matemática do modelo:

- **Índices ou Conjuntos:**
 - **Matérias-Primas – MP:** {Mamona e Girassol}
 - **Produtos Finais – PF:** {Óleo de mamona, óleo de girassol, torta de mamona e torta de girassol}
 - **Origens/Locais de Oferta – O:** {Municípios ofertantes}
 - **Destinos/Locais de Demanda – D:** {Municípios com as usinas de biodiesel da Petrobras, portos para exportação, indústrias, ...}
 - **Esmagadoras – E (selecionadas pelo AG):** { Municípios dentre os possíveis candidatos}
 - **Período(ano) – P:** {2010, 2011, ..., 2020}
 - **Tecnologias de esmagamento – T:** { Prensagem mecânica ou extração por solvente}
 - **Tipos de capacidade da esmagadora – C:** {5.000 ton óleo/ano, 10.000 ton óleo/ano, ..., 80.000 ton óleo/ano}

- **Dados Gerais:**
 - **PenalidadeFic** – Penalidade fictícia, para evitar inviabilidades no modelo e piorar a avaliação de determinadas soluções.
 - **Taxa(P)** – Taxa de desconto a ser aplicada no resultado econômico.

- **Dados das Ofertas:**
 - **Oferta(MP,O,P)** – Volume da matéria-prima MP ofertada no local de oferta O no período P.
 - **OfertaDistancia(O,E)** – Distância entre o local de oferta O e a esmagadora E.
 - **OfertaFrete(MP,O,E)** – Frete da matéria-prima MP do local de oferta O a esmagadora E.

- **DistanciaMax** – Máxima distância que pode ser percorrida de um local de oferta a uma esmagadora.
- **Dados das Demandas:**
 - **DemandaMin(PF,D,P)** – Demanda mínima do produto final PF no local de demanda D no período P.
 - **DemandaMax(PF,D,P)** – Demanda máxima do produto final PF no local de demanda D no período P.
 - **DemandaPreco(PF,P)** – Preço do produto final PF no período P.
 - **DemandaFrete(PF,E,D,P)** – Frete do produto final PF da esmagadora E ao local de demanda D no período P.
 - **Oleoso(PF)** – Indica se o produto final PF é oleoso.
- **Dados das Esmagadoras (selecionados pelo AG):**
 - **E_Capacidade(E,P)** – Incremento de capacidade da esmagadora E no período P.
 - **E_Tecnologia(E)** – Tecnologia adotada pela esmagadora E.
 - **E_Inv(E,P)** – Indica que houve investimento na esmagadora E no período P.
- **Dados das Esmagadoras:**
 - **Rendimento(MP,PF,T)** – Rendimento do produto final PF produzido a partir da matéria-prima MP com a tecnologia T.
 - **CapacidadeEsmag (C)** – Capacidade da esmagadora em ton óleo / ano.
 - **CapacidadeAcumulada(E,P)** – Capacidade total da esmagadora E no período P. Este é um dado calculado a partir dos parâmetros de investimento e capacidade das esmagadoras.

$$CapacidadeAcumulada(E, P) = \sum_{P1|P1 \leq P} CapacidadeEsmag(E, P1)$$

- **CustoFixo(E,P)** – Custo fixo da esmagadora E no período P. Este custo é calculado e depende do custo operacional da tecnologia e capacidade selecionada para a esmagadora E.
- **CustoVariavel(E,P)** – Custo variável da esmagadora E no período P, que é aplicado em função da produção total em óleo. Este custo é calculado e depende do custo operacional da tecnologia e capacidade selecionada para a esmagadora E.
- **CustoInstalacao(C,T,P)** – Custo de instalação de uma nova esmagadora com capacidade C e tecnologia T, no período P.

- **Variáveis:**

- **FluxoMP(MP,O,E,P)** – Fluxo da matéria-prima MP do local de oferta O para a esmagadora E no período P.

Esta variável só será criada se a distância do ponto de oferta O a esmagadora E for menor do que a distância máxima.

- **FluxoPF(PF,E,D,P)** – Fluxo do produto final PF da esmagadora E ao local de demanda D no período P.
- **DeficitOferta(MP,O,P)** – Déficit de absorção da oferta da matéria-prima MP no local de oferta O no período P.
- **UtilizacaoFinalEsmag(E)** – Utilização final da esmagadora E.

- **Restrições:**

- **BalancoOferta(MP,O,P)** – O total que sai de cada ponto de origem para todas as esmagadoras tem que ser igual à oferta do ponto, sujeito a uma penalidade.

$$\sum_E FluxoMP(MP, O, E, P) + DeficitOferta(MP, O, P) = Oferta(MP, O, P)$$

- **AtendimentoDemandaMinima(PF,D,P)** – O total que chega em cada ponto de demanda vindo de todas as esmagadoras tem que ser pelo menos o mínimo.

$$DemandaMin(PF, D, P) \leq \sum_E FluxoPF(PF, E, D, P)$$

- **AtendimentoDemandaMaxima(PF,D,P)** – O total que chega em cada ponto de demanda vindo de todas as esmagadoras tem que ser pelo menor do que a demanda máxima.

$$\sum_E FluxoPF(PF, E, D, P) \leq DemandaMax(PF, D, P)$$

- **BalancoEsmagadoras(PF,E,P)** – A produção total de um produto final PF na esmagadora é igual à carga total de matéria-prima * rendimento na tecnologia selecionada para a esmagadora.

$$\sum_D FluxoPF(PF, E, D, P) = \sum_{O,MP} FluxoMP(MP, O, E, P) \times Rendimento(MP, PF, E_Tecnologia(E))$$

- **CapacidadeEsmagadora(E,P)** – A produção total do produto final PF, onde PF é um produto oleoso, na esmagadora tem que ser menor ou igual à capacidade da esmagadora,

$$\sum_{D,PF|Oleoso(PF)} FluxoPF(PF, E, D, P) \leq CapacidadeAcumulada(E, P)$$

- **ProducaoEsmagadora(E,P)** – A produção total de uma esmagadora não pode diminuir no horizonte de análise.

$$\sum_{D,PF|Oleoso(PF)} FluxoPF(PF, E, D, P) \geq \sum_{D,PF|Oleoso(PF)} FluxoPF(PF, E, D, P-1)$$

- **ProducaoFinal(E)** – A utilização das esmagadoras no último período deve ser maior do que 50%, sujeito a uma penalidade.

$$\frac{\sum_{D,PF|Oleoso(PF)} FluxoPF(PF, E, D, \acute{U}timo_Per\acute{í}odo)}{CapacidadeAcumulada(E, P)} + UtilizacaoFinalEsmag(E) \geq 0.5$$

- **Função Objetivo:** Receita – Custos – Penalidades

$$\begin{aligned} & \sum_P Taxa(P) \times [\\ & \quad \sum_{D,PF,E} FluxoPF(PF, E, D, P) \times Demanda\ Pr\ eco(PF) \\ & \quad - \sum_{O,MP,E} FluxoMP(MP, O, E, P) \times OfertaFrete(MP, O, E) \\ & \quad - \sum_{D,PF,E} FluxoPF(PF, E, D, P) \times DemandaFrete(PF, E, D) \\ & \quad - \sum_E CustoFixo(E, P) \times \sum_{PI|PI \leq P} E_Inv(E, P) \\ & \quad - \sum_E CustoVariavel(E, P) \times \sum_{PF,D|Oleoso(PF)} FluxoPF(PF, E, D, P) \times \\ & \quad \sum_{PI|PI \leq P} E_Inv(E, P) \\ & \quad - \sum_E CustoInstalacao(E_Capacidade(E, P), E_Tecnologia(E), P) \times \\ & \quad E_Inv(E, P) \\ & \quad - \sum_{MP,O} DeficitOferta(MP, O, P) \times PenalidadeFic \\ & \quad] \\ & \quad - \sum_E UtilizacaoFinalEsmag(E) \times PenalidadeFic \end{aligned}$$

4.4. TECNOLOGIAS ADOTADAS NA IMPLEMENTAÇÃO DO MODELO

Para este trabalho foi implementado um protótipo para realização de testes comparativos, com os dados reais do problema. O protótipo não teve como objetivo alcançar alta performance e sim auxiliar na prova de conceito de que um modelo baseado em algoritmo genético seria capaz de obter soluções satisfatórias para um problema de localização. O modelo proposto foi implementado em ambiente .NET, empregando-se a linguagem C#. A escolha das tecnologias adotadas foi voltada para aquelas que permitissem uma rápida implementação e que fornecessem resultados confiáveis.

Para a implementação do algoritmo genético foi utilizado o componente GACOM (ICA, 2009), que tem como finalidade permitir o desenvolvimento de programas utilizando técnicas de computação evolucionária. Esta biblioteca conta com as seguintes funcionalidades: definição de diferentes tipos de representação de genes, segmentação do indivíduo para permitir cromossomos compostos por mais de um tipo de representação, formas de elitismo, seleção por roleta, normalização linear, operadores evolutivos (*crossover*, mutação e *crossover* especial), multiobjetivos com pareto, execução em *pipeline* para permitir inserção de blocos de código entre os passos da otimização, definição de diversos critérios de parada.

O modelo de programação linear foi implementado na ferramenta de modelagem matemática AIMMS (Paragon, 2009) e o otimizador utilizado foi o CPLEX (IBM, 2009). O AIMMS fornece um componente que permite que a otimização seja disparada através do ambiente .NET.

Para armazenar os dados do problema foi utilizado um banco de dados MS-Access.

5. RESULTADOS

Nesta seção apresentamos os resultados encontrados, utilizando os dados do problema real da Petrobras Biocombustível, com o modelo proposto na seção 4 e os modelos de programação inteira mista, utilizados atualmente para resolver o problema, que foram apresentados na seção 2. Serão apresentados apenas resultados comparativos, sem os valores absolutos e a identificação dos locais selecionados, devido à confidencialidade das informações.

5.1. MODELO DE PROGRAMAÇÃO INTEIRA MISTA

Este modelo consiste de um problema de programação matemática, utilizando variáveis inteiras e contínuas, onde foram passados para o modelo todos os 96 locais candidatos e as possíveis capacidades e tecnologias, para que o mesmo determinasse a melhor solução. Testes realizados com este modelo mostraram que mesmo após 48 horas não obteve solução satisfatória. O resultado econômico encontrado era negativo e muito longe da solução ótima (*gap* de 122,6%).

5.2. MODELO DE PROGRAMAÇÃO INTEIRA MISTA COM PRÉ-PROCESSAMENTO PARA REDUÇÃO DO NÚMERO DE LOCAIS CANDIDATOS E PONTOS DE OFERTA

Como descrito na seção 2, este modelo é o mesmo apresentado na seção anterior, mas com um pré-processamento para redução dos dados de entrada. Uma heurística foi utilizada para eliminar candidatos, resultando em 13 locais. Todo o processo de eliminação de candidatos leva cerca de 1 dia e exige a intervenção de um especialista. Além disso, também foi realizado um agrupamento dos pontos de oferta através outro modelo de programação matemática, que leva aproximadamente 30 minutos. Com os novos dados de entrada, o modelo de programação inteira obteve resultado econômico positivo com uma rodada com tempo limite de 18 horas. A rodada terminou por limite de tempo e com *gap* de 2,5%. O modelo indicou 7 locais para a instalação de esmagadoras. A Figura 12 apresenta a distribuição espacial dos locais escolhidos.

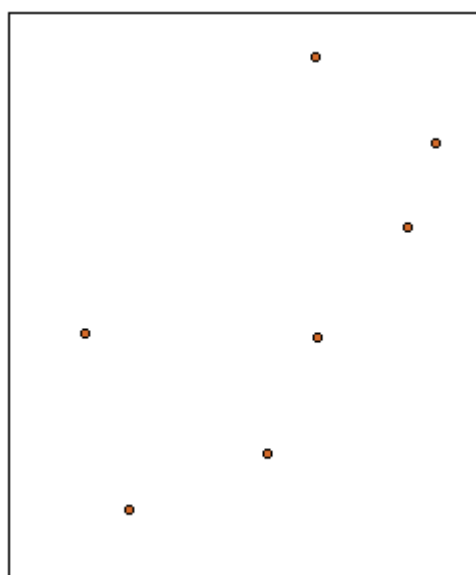


Figura 12: Distribuição espacial dos locais escolhidos para a instalação das esmagadoras

Em dois locais foi indicada a instalação de esmagadoras com tecnologia de prensagem mecânica e os outros 5 com a tecnologia de extração por solvente. A Tabela 2 apresenta a capacidade instalada das esmagadoras a cada período. Pode-se verificar que foram indicadas esmagadoras de vários portes, desde 5.000 ton / óleo ano a 80.000 ton / óleo ano. Também houveram indicações de ampliações de algumas esmagadoras, como por exemplo, a 2 e a 6. A Tabela 3 mostra a utilização das esmagadoras a cada período, que foi calculada como a quantidade de óleo produzido dividida pela capacidade da esmagadora. A maioria das esmagadoras tem utilização maior que 50% nos períodos analisados e no último período, todas tem uma alta utilização.

| Local\ Período | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 |
|----------------|------|------|------|------|------|------|------|------|------|------|
| 1 | 80 | 80 | 80 | 80 | 80 | 80 | 80 | 80 | 80 | 80 |
| 2 | | | | | | | | 25 | 25 | 30 |
| 3 | | | | | 25 | 25 | 25 | 25 | 25 | 25 |
| 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 5 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| 6 | | | | | | 25 | 25 | 75 | 75 | 75 |
| 7 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 |

Tabela 2: Capacidade instalada nas esmagadoras a cada período (em mil ton / óleo ano)

| Local\ Período | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 |
|----------------|------|------|------|------|------|------|------|------|------|------|
| 1 | 44% | 55% | 64% | 75% | 73% | 81% | 95% | 91% | 98% | 100% |
| 2 | | | | | | | | 64% | 69% | 74% |
| 3 | | | | | 95% | 96% | 100% | 77% | 83% | 89% |
| 4 | 83% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% |
| 5 | 62% | 70% | 84% | 100% | 100% | 82% | 100% | 87% | 94% | 100% |
| 6 | | | | | | 100% | 100% | 62% | 73% | 84% |
| 7 | 65% | 78% | 91% | 100% | 100% | 100% | 100% | 100% | 100% | 100% |

Tabela 3: Utilização das esmagadoras a cada período

5.3. MODELO UTILIZANDO ALGORITMO GENÉTICO

Inicialmente foram realizados testes para ajustes dos parâmetros do algoritmo genético. Na melhor solução encontrada, utilizamos os seguintes parâmetros do algoritmo genético:

- **Crossover:** Varia com taxa linear iniciando em 0,95 e decrescendo até 0,90. Foram utilizados os operadores disponíveis no GACOM.
- **Mutação:** Varia com taxa linear iniciando em 0,01 e crescendo até 0,50. Foram utilizados os operadores disponíveis no GACOM.
- **Steady State:** Varia com taxa linear iniciando em 0,95 e decrescendo até 0,90.
- **Número de indivíduos em cada geração:** 100.
- **Número de gerações:** 100.
- **Número de experimentos:** 1.

Quanto aos parâmetros do algoritmo genético, como os segmentos eram binários ou inteiros, utilizamos os operadores disponíveis no GACOM (ICA,2009), com seleção pelo método de roleta. Não foi foco de estudo realizar testes com apenas um subconjunto dos operadores disponíveis.

Cada avaliação de cromossomo através do modelo de programação linear demorava cerca de 7 segundos, então cada experimento com 100 gerações demorava cerca de um dia para ser executado.

O algoritmo genético, mesmo nas primeiras rodadas, sem os parâmetros ajustados já apresentou solução com resultado econômico positivo nas primeiras gerações, o que é uma boa vantagem com relação à técnica de programação matemática. Além disso, as soluções encontradas já eram da mesma ordem de grandeza da melhor solução obtida com o outro modelo. A Figura 13 apresenta a evolução da solução com 100 gerações, comparada com a melhor solução encontrada com o modelo de programação matemática. A ordem de grandeza do resultado econômico é centenas de milhões de reais.

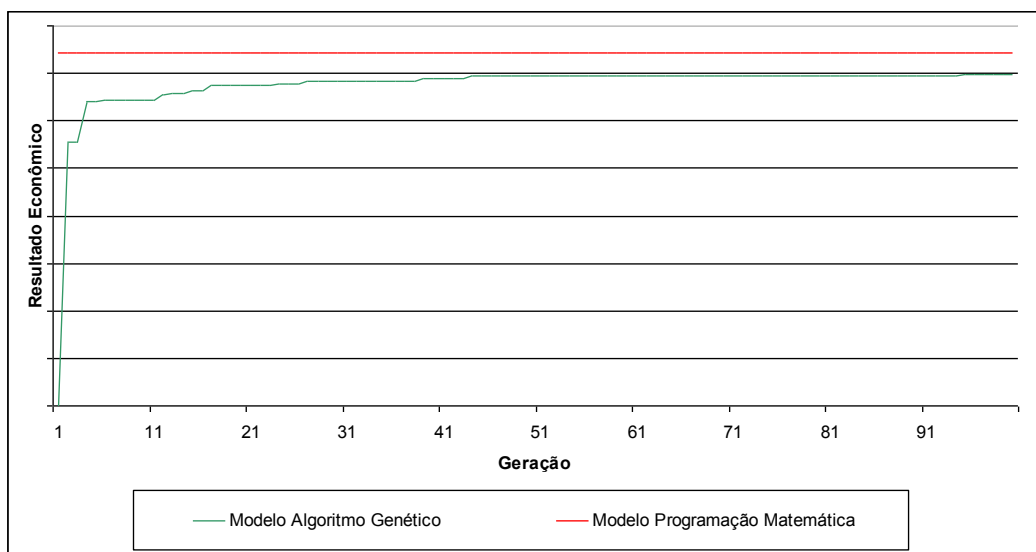


Figura 13: Comparação da evolução da solução no modelo que utiliza algoritmo genético com a melhor solução encontrada com o modelo de programação matemática

Pode-se verificar através da Figura 13, que o modelo que utiliza algoritmo genético obteve uma solução pior do que o modelo que realiza pré-processamento nos dados de entrada. Esta solução encontrava-se 6% abaixo da solução do modelo de programação matemática. Porém, vale ressaltar que o modelo de algoritmo genético com apenas 10 gerações já havia encontrado uma boa solução. Após a 50ª geração, a solução não melhorou, mas já encontrava-se muito próxima a do modelo de programação matemática.

Esta solução determinou a instalação de esmagadoras em 10 locais. A Figura 14 apresenta a distribuição espacial dos locais selecionados. É possível verificar que esta solução aponta locais mais ao centro do espaço de busca do que o modelo de programação matemática.

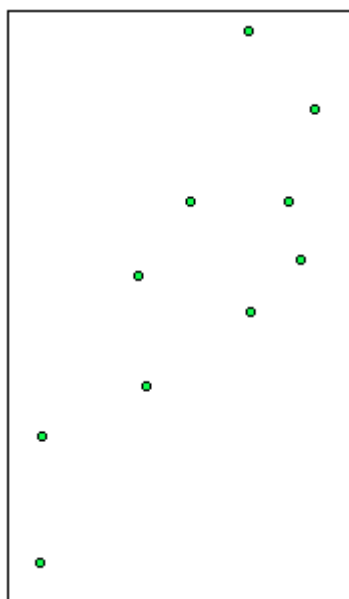


Figura 14: Distribuição espacial dos locais escolhidos para a instalação das esmagadoras

As tecnologias selecionadas foram: 5 locais com prensagem mecânica e nos outros 5, extração por solvente. A Tabela 4 apresenta a capacidade instalada das esmagadoras a cada período. Foram indicadas esmagadoras de vários portes, desde 5.000 ton / óleo ano a 50.000 ton / óleo ano. Não houve nenhuma indicação de ampliação de esmagadora. A Tabela 5 mostra a utilização das esmagadoras a cada período. A maioria das esmagadoras também teve utilização maior que 50% nos períodos analisados e no último período, todas tem uma alta utilização.

| Local\ Período | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 |
|----------------|------|------|------|------|------|------|------|------|------|------|
| 1 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| 2 | | | | | | 50 | 50 | 50 | 50 | 50 |
| 3 | | | | | | | | | | 10 |
| 4 | | | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 |
| 5 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| 6 | | | | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| 7 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 8 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 9 | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 |
| 10 | | | | | | | | 50 | 50 | 50 |

Tabela 4: Capacidade instalada nas esmagadoras a cada período (em mil ton / óleo ano)

| Local/ Período | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 |
|-------------------|------|------|------|------|------|------|------|------|------|------|
| 1 | 78% | 99% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% |
| 2 | | | | | | 35% | 58% | 58% | 75% | 84% |
| 3 | | | | | | | | | | 73% |
| 4 | | | 78% | 78% | 78% | 78% | 78% | 78% | 78% | 78% |
| 5 | 79% | 92% | 92% | 92% | 93% | 100% | 100% | 100% | 100% | 100% |
| 6 | | | | 44% | 81% | 81% | 100% | 100% | 100% | 100% |
| 7 | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% |
| 8 | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% |
| 9 | 58% | 75% | 75% | 75% | 86% | 86% | 86% | 86% | 86% | 86% |
| 10 | | | | | | | | 43% | 66% | 79% |

Tabela 5: Utilização das esmagadoras a cada período

Em comparação com a solução encontrada com o modelo de programação matemática, esta solução apresenta uma capacidade total instalada um pouco maior: apenas 5.000 ton/óleo por ano. De uma maneira geral, esta solução tem mais esmagadoras de pequeno porte (até 10.000 ton / óleo ano) além de não ter instalado nenhuma esmagadora com o maior porte possível (80.000 ton / óleo ano).

Dois dos locais selecionados pelo algoritmo genético foram iguais aos encontrados pelo outro modelo, e os demais são muito próximos, como pode-se verificar através da Figura 15, que compara os locais escolhidos pelos dois modelos.

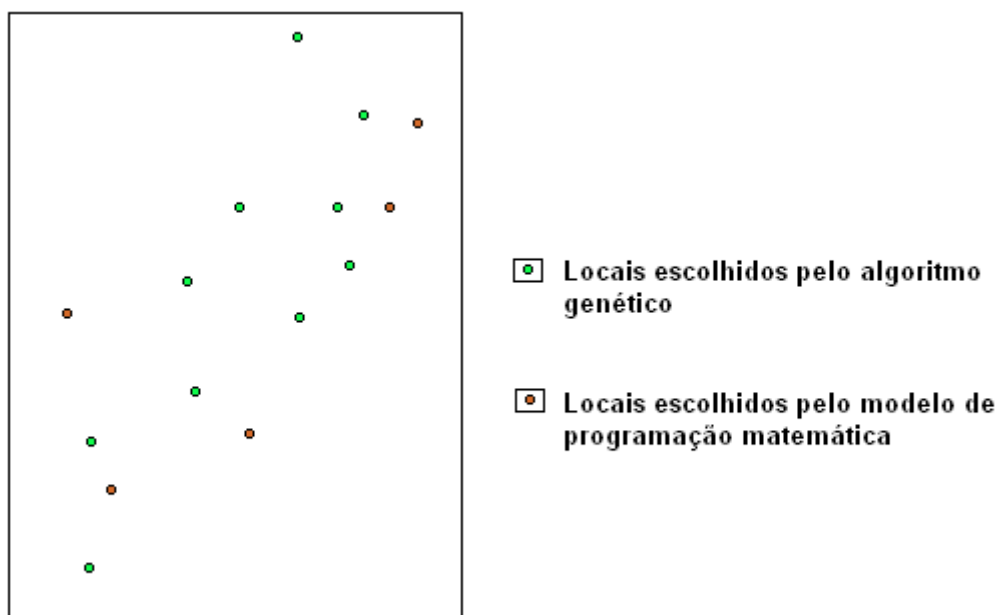


Figura 15: Comparação da distribuição espacial dos locais escolhidos para a instalação das esmagadoras nos dois modelos

Outro teste realizado foi utilizar no algoritmo genético o mesmo conjunto de candidatos reduzido, resultante da heurística utilizada no modelo de programação matemática. A Figura 16 mostra a evolução da solução com 13 candidatos, comparada com o resultado do modelo de programação matemática e do modelo de algoritmo genético com 96 candidatos. Este problema reduzido, com 100 gerações, foi executado em 12 horas pelo modelo proposto e encontrou uma solução 2,6% abaixo do modelo de programação matemática.

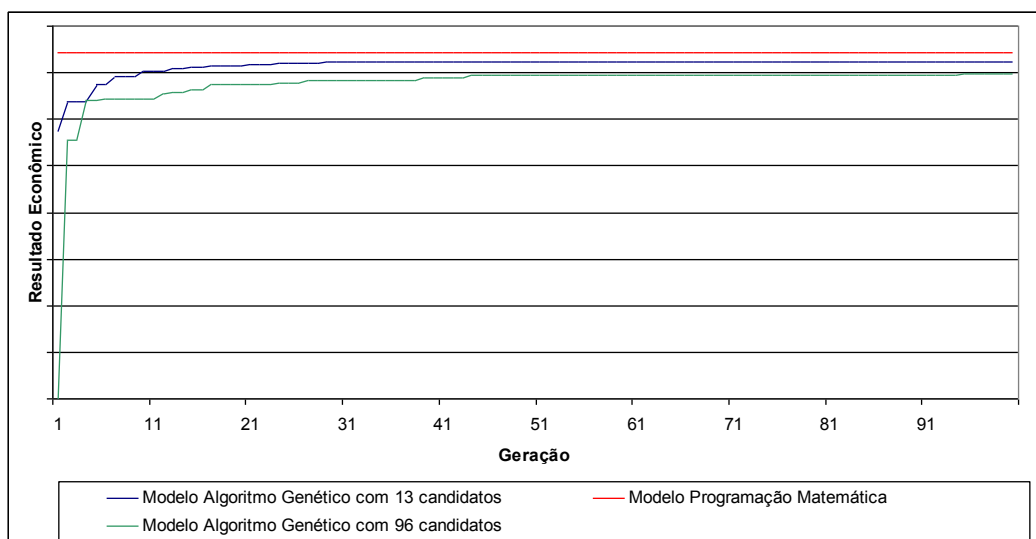


Figura 16: Comparação da evolução da solução no modelo que utiliza algoritmo genético com 13 e 96 candidatos com a melhor solução encontrada com o modelo de programação matemática

Pode-se verificar que a solução melhorou, e que, antes de 10ª geração, supera a solução do modelo de algoritmo genético com 96 candidatos. Isto ocorre devido a significativa redução do espaço de busca, com a redução do número de candidatos. Nesta solução foram escolhidos 7 locais para a instalação de esmagadoras. A Figura 17 apresenta a distribuição destes locais comparados com os locais escolhidos pelo modelo de programação matemática, e os locais são praticamente os mesmos. Na Figura só é possível identificar 1 ponto indicado pelo modelo de programação matemática, pois os demais são os mesmos indicados pelo algoritmo genético. O local alternativo, escolhido pelo modelo proposto, está envolvido em vermelho junto com o local selecionado pelo modelo de programação matemática. Pode-se verificar que estes locais são muito próximos.

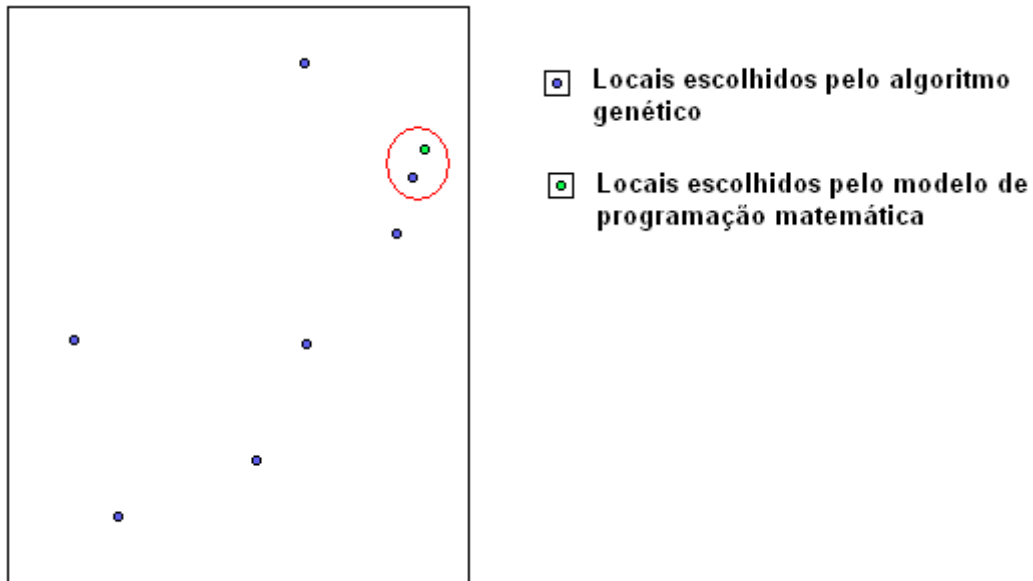


Figura 17: Comparação da distribuição espacial dos locais escolhidos para a instalação das esmagadoras nos dois modelos

Conclui-se então que o modelo proposto fornece soluções satisfatórias para este problema, com pouco processamento computacional e sem a necessidade de realizar tratamento nos dados de entrada. E, ainda, que se for realizada a redução do número de candidatos, o modelo proposto apresenta resultados tão bons quanto o modelo de programação matemática, com um esforço computacional menor e sem a necessidade de realizar o agrupamento dos pontos de oferta.

6. CONCLUSÕES E TRABALHOS FUTUROS

O presente trabalho apresentou uma solução para a localização ótima de unidades de extração de óleos vegetais para suprimento a usinas produtoras de biodiesel da Petrobras Biocombustível. O problema representado consiste em determinar, a partir de um conjunto de municípios candidatos, quais os melhores locais para a instalação de unidades de extração de óleos vegetais, além dos tipos de oleaginosas a serem esmagadas, as tecnologias de esmagamento adotadas, o porte das instalações e o ano de realização do investimento.

O problema tratado havia sido solucionado pela equipe de Pesquisa Operacional da Petrobras, através de um modelo que utiliza técnica de programação matemática. Este trabalho apresenta um modelo alternativo ao atualmente utilizado baseado em algoritmo genético.

O modelo atualmente utilizado apresentou boas soluções com a realização de um pré-processamento nos dados de entrada que: agrupavam os pontos de oferta para reduzir o número de variáveis de fluxo e eliminavam a maior parte dos locais candidatos através de uma heurística.

O modelo proposto apresentou bons resultados em tempo satisfatório, sem a necessidade de realizar pré-processamento nos dados de entrada. O resultado obtido foi apenas 6% menor do que o melhor resultado obtido com o modelo atualmente utilizado. Ao utilizar a heurística de eliminação de candidatos nos dados de entrada do algoritmo genético, o mesmo consegue obter soluções tão boas quanto o modelo de programação matemática. E, ainda, com um esforço computacional menor e sem a necessidade de reduzir os fluxos entre os pontos de ofertas e os candidatos, através do agrupamento dos pontos de oferta.

Outro ponto importante é que o modelo proposto pode ser utilizado como ferramenta de análise diretamente pelo usuário final de maneira mais fácil que o modelo de programação matemática, onde é mandatório que sejam realizados tratamentos nos dados de entrada. Por outro lado, uma desvantagem do modelo proposto é que podem ser necessários ajustes nos parâmetros do algoritmo genético, quando for utilizado outro conjunto de dados, o que pode dificultar o uso pelo usuário final.

Como trabalhos futuros, podemos citar: necessidade de utilização de software mais eficiente para cálculo da função de avaliação, para a diminuição do tempo da rodada; necessidade de refinamento e mais testes para a

configuração do modelo de algoritmo genético; estudo de utilização de heurística para eliminar candidatos, assim como a utilizada no modelo de programação matemática, com o objetivo de reduzir o espaço de busca da solução; com a redução do tempo de cada rodada, devem-se realizar mais experimentos para obter uma solução mais robusta.

REFERÊNCIAS BIBLIOGRÁFICAS

- AROSTEGUI, M. A., KADIPASAOGLU, S. N., KHUMAWALA, B. M. **An empirical comparison of Tabu Search, Simulated Annealing, and Genetic Algorithms for facilities location problems**, International Journal of Production Economics, 2005.
- BEASLEY, D., BULL, D.R., MARTIN R.R. **An overview of genetic algorithms: Part I. Fundamentals**, University Computing 15, p. 58–69, 1993
- BRUNDEAU, M. L., CHIU, S. S. **An overview of representative problems in location research**, Management Science, 1989.
- CHAUDRY, S.S., HE, S., CHAUDRY, P.E., **Solving a class of facility problem using genetic algorithms**, Expert Systems 20 (2), p. 86-91, 2003.
- DARWIN, C., **The Origin of Species**, John Murray, Albermale Street, London, 1859.
- DAVIS, L. D., MITCHELL, M. **Handbook of Genetic Algorithms**, Van Nostrand Reinhold, New York, 1991.
- GOLDBERG, D.E. **Genetic algorithms in search, optimization and machine learning**, Addison-Wesley, 1989.
- HOLLAND, J., **Adaptation in Natural and Artificial Systems**, MIT Press, 1992.
- IBM, **IBM ILOG CPLEX**, Versão 11.2, Disponível em: <<http://www-01.ibm.com/software/integration/optimization/cplex/>>, 2009.
- ICA - LABORATÓRIO DE INTELIGÊNCIA COMPUTACIONAL APLICADA, **GACOM – Genetic Algorithms Components**, Disponível em: <<http://www.ica.ele.puc-rio.br/>>, 2009.
- JARAMILLO, J. H., BHADURY, J., BATTÀ, R. **On the use of genetic algorithms to solve location problems**, 2002.
- KRARUP, J., PRUZAN, M. **The simple plant location problem: Survey and synthesis**, 1981.
- KRATICA, J., TOSIC, D., FILIPOVIC, V., LJUBIC, I. **Solving the simple plant location problem by genetic algorithm**, Operations Research, 2001.
- MENDEL, G., **Ensaio com Plantas Híbridas**, Sociedade de História Natural de Brno, 1865.

MICHALEWICZ, Z., **Genetic Algorithm + Data Structure = Evolution Programs**, Springer-Verlag, NEW YORK, 1996.

PACHECO, M. A. C, **Algoritmos Genéticos: Princípios e Aplicações, Apostila de Computação Evolucionária**, ICA: Laboratório de Inteligência Computacional Aplicada, 1999.

PACHECO, M. A. C, **Curso de Otimização por Algoritmos Genéticos**, Curso de Especialização em Business Intelligence, PUC-RIO, 2008.

PARAGON DECISION TECHNOLOGY, **AIMMS optimization software**, Disponível em: <<http://www.aimms.com/>>, 2009.

SRIDHARAN, R. The **capacitated plant location problem**, European Journal of Operational Research, 1995.

SYARIF, A., YUN, Y., GEN, M. **Study on multi-stage logistic chain network; a spanning tree-based genetic algorithm approach**, Computers & Industrial Engineering 43 (1-2); p. 299-315, 2002.

TAHA, H. A. **Pesquisa Operacional**, 8a edição, Pearson Education do Brasil, 2007.

WEBER, A. Uber den Standort der Industrien, 1909; translated as Alfred Weber's **Theory of the Location of Industries**, University of Chicago, 1929.