

UNIVERSIDADE FEDERAL DE UBERLÂNDIA - UFU
FACULDADE DE ENGENHARIA ELÉTRICA - FEELT
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Jean Carlos de Oliveira

**Obtenção do Modelo Matemático e
Mapa de Calor do Consumo de
Combustível de um Caminhão na
Mineração à Céu Aberto, Utilizando
Algoritmo Genético e Regressão Linear
Múltipla**

Uberlândia
2019

Jean Carlos de Oliveira

**Obtenção do Modelo Matemático e
Mapa de Calor do Consumo de
Combustível de um Caminhão na
Mineração à Céu Aberto, Utilizando
Algoritmo Genético e Regressão Linear
Múltipla**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Faculdade de Engenharia Elétrica como parte dos requisitos para a obtenção do título de Mestre em Engenharia Elétrica.

Área de concentração: Sistemas de Energia Elétrica

Orientador: Dr. Aniel Silva de Moraes

Uberlândia

2019

Trata-se da versão corrigida da dissertação. A versão original se encontra disponível na EESC/UFU que aloja o Programa de Pós-Graduação de Engenharia Elétrica.

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

O48o
2019 Oliveira, Jean Carlos de, 1985-
Obtenção do modelo matemático e mapa de calor do consumo de combustível de um caminhão na mineração à céu aberto, utilizando algoritmo genético e regressão linear múltipla [recurso eletrônico] / Jean Carlos de Oliveira. - 2019.

Orientador: Aniel Silva de Moraes.

Dissertação (mestrado) - Universidade Federal de Uberlândia, Programa de Pós-Graduação em Engenharia Elétrica.

Modo de acesso: Internet.

Disponível em: <http://dx.doi.org/10.14393/ufu.di.2019.39>

Inclui bibliografia.

Inclui ilustrações.

1. Engenharia elétrica. 2. Algoritmos genéticos. 3. Análise de regressão. 4. Redes elétricas - Redução de custos e tempo. 5. Combustível - Consumo. I. Moraes, Aniel Silva de, 1979-, (Orient.). II. Universidade Federal de Uberlândia. Programa de Pós-Graduação em Engenharia Elétrica. III. Título.

CDU: 621.3



UNIVERSIDADE FEDERAL DE UBERLÂNDIA

ATA DE DEFESA

Programa de Pós-Graduação em:	Engenharia Elétrica				
Defesa de:	Dissertação de Mestrado Acadêmico, 710, COPEL.				
Data:	27/05/2019	Hora de início:	13:00	Hora de encerramento:	15:00
Matrícula do Discente:	11712EEL007				
Nome do Discente:	Jean Carlos de Oliveira				
Título do Trabalho:	Obtenção do modelo matemático e mapa de calor do consumo de combustível de um caminhão na mineração à céu aberto, utilizando algoritmo genético e regressão linear múltipla.				
Área de concentração:	Sistemas de energia elétrica				
Linha de pesquisa:	Controle e automação				
Projeto de Pesquisa de vinculação:	Título: Técnicas de identificação de sistemas visando aplicabilidade em processos industriais Agência Financiadora: UNIVERSIDADE FEDERAL DE UBERLÂNDIA - (Programa Institucional de Apoio a Pesquisa - apoio financeiro). Início 02/01/2013. Término (em andamento).				

Reuniu-se no Anfiteatro do Bloco 1E da Faculdade de Engenharia Elétrica, Campus Santa Mônica, da Universidade Federal de Uberlândia, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Engenharia Elétrica, assim composta: Professores Doutores: Fábio Vincenzi Romualdo da Silva - FEELT/UFU; Márcio José da Cunha - FEELT/UFU; Kléber Lopes Fontoura - CEFET-MG; Aniel Silva de Moraes - UFU, orientador(a) do(a) candidato(a).

Iniciando os trabalhos o(a) presidente da mesa, Dr(a). Aniel Silva de Moraes, apresentou a Comissão Examinadora e o candidato(a), agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos(às) examinadores(as), que passaram a arguir o(a) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o(a) candidato(a):

Aprovado(a).

Esta defesa faz parte dos requisitos necessários à obtenção do título de **Mestre**.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Aniel Silva de Moraes, Professor(a) do Magistério Superior**, em 27/05/2019, às 16:42, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Marcio José da Cunha, Professor(a) do Magistério Superior**, em 27/05/2019, às 17:13, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Kleber Lopes Fontoura, Usuário Externo**, em 28/05/2019, às 00:16, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Fabio Vincenzi Romualdo da Silva, Professor(a) do Magistério Superior**, em 30/05/2019, às 15:40, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1215810** e o código CRC **9B7F323D**.

*Este trabalho é dedicado a minha família,
e especialmente a minha esposa Thamires
e aos meus pais
que são minhas motivações da vida.*

Agradecimentos

O autor agradece à UFU - Universidade Federal de Uberlândia e a Instale Tecnologia pela disponibilidade de pesquisas, desenvolvimento e testes da aplicação.

“A diferença entre o possível e o impossível está na vontade humana.”

(Louis Pasteur)

Resumo

Oliveira, J.C. **Obtenção do Modelo Matemático e Mapa de Calor do Consumo de Combustível de um Caminhão na Mineração à Céu Aberto, Utilizando Algoritmo Genético e Regressão Linear Múltipla.** 73 p. Dissertação de mestrado – Faculdade de Engenharia de Uberlândia, Universidade de Uberlândia, 2019.

O ambiente de negócios global força as organizações a melhorar seus processos e serviços como um meio para o sucesso e a sobrevivência. A busca por otimização tornou-se essencial para a melhor utilização dos recursos, redução dos custos e maximização dos resultados. Este trabalho propõe, através das ferramentas: Regressão Linear Múltipla (RLM) e Algoritmo Genético (AG), a obtenção de modelos matemáticos e um mapa de calor do consumo de combustível de um caminhão de mineração em operação, baseado em suas características físicas de rotas. Além disso, é desenvolvido um sistema responsável pelo processamento dos dados, parametrização do algoritmo genético e validação dos modelos obtidos por comparação entre eles e valores de referência. Através do modelo matemático e do mapa de calor, é possível criar rotinas, fornecer informações para sistemas de despacho de caminhões e alcançar a redução dos custos operacionais de mineração. O sistema desenvolvido pode desempenhar um papel importante durante a definição e criação de novas rotas, ajudando também na indicação de rotas econômicas. Além disso, pode ser útil ao revisar rotas existentes, suportando mudanças em sua topografia. A pesquisa usa dados reais e atuais coletados de um sistema de telemetria de uma mina a céu aberto. A Regressão Linear Múltipla foi realizada no ambiente MS Excel®, enquanto o Algoritmo Genético foi implementado no software Matlab®. Estudos que pretendem reduzir o consumo de combustível fornecem informações significativas para as empresas de mineração a céu aberto, uma vez que esse consumo representa uma grande parte dos custos operacionais. Também vale ressaltar o benefício de reduzir as emissões de gases de efeito estufa, cujo interesse e preocupação são gerais.

Palavras-chave: Algoritmo Genético, Regressão Linear Múltipla, Consumo de Combustível, Redução de Custos, Rota Econômica.

Abstract

Oliveira, J.C. **Obtaining the Mathematical Model and Heat Map of Fuel Consumption of a Truck in Open Pit Mines, through Multiple Linear Regression and Genetic Algorithm.** 73 p. Master Thesis – Uberlândia School of Engineering, University of Uberlândia, 2019.

The global business environment forces organizations to improve its processes and services as a means to success and survival. Seeking for optimization has become essential for better use of resources, reducing costs and maximizing results. This work proposes, with the help of both tools, Multiple Linear Regression (MLR) and Genetic Algorithm (GA), to obtain mathematical models and a heat map of fuel consumption, of a mining truck in operation based on its routes physical characteristics. In addition to, the developed system is responsible for data processing, the genetic algorithm parameterization and obtained models validation by comparison among them and reference values. Through the mathematical model and the heat map, it is possible to create routines, provide information for truck dispatch systems and achieve consequent mine-operating costs reduction. The developed system plays an important during the definition and creation of new routes, helping to indicate the economic one. Moreover, it can be useful when reviewing existing routes, supporting changes in their topography. The research uses real and current data collected from a telemetry system of an open pit mine. The Multiple Linear Regression was performed in MS Excel® environment, while the Genetic Algorithm was implemented at Matlab® software. Studies which intend to reduce fuel consumption, provide significant information for open-pit mining companies once this consumption represents a large part of operating costs. It is also worth emphasizing the benefit of reducing greenhouse gases emissions, which interest and concern are general.

Keywords: Genetic Algorithm, Multiple Linear Regression, Fuel Consumption, Cost Reduction, Economic Route.

Lista de ilustrações

Figura 1	Representação de rotas possíveis na operação de minas a céu aberto.	22
Figura 2	Entrada e saída de um VRP.	23
Figura 3	Sistema de coordenadas UTM.	25
Figura 4	Trajetória dos preços dos principais combustíveis.	26
Figura 5	Exemplo de rotas com distância e consumo.	27
Figura 6	Observações do modelo de regressão linear simples.	29
Figura 7	Gráficos de diferentes r.	32
Figura 8	Exemplo de dados para regressão.	35
Figura 9	Ferramenta de Análise de Dados do Excel.	35
Figura 10	Ferramenta de Regressão do Excel.	36
Figura 11	Relatório da Regressão do Excel.	37
Figura 12	Fluxograma do Algoritmo Genético.	40
Figura 13	Estrutura do cromossomo em representação binária.	40
Figura 14	População de cromossomos em representação binária.	41
Figura 15	Cruzamento de ponto único.	44
Figura 16	Cruzamento com dois pontos de corte.	44
Figura 17	Mutação aleatória no sexto gene.	45
Figura 18	Mutação por troca do segundo com o sétimo gene.	45
Figura 19	Estrutura do cromossomo em representação numérica.	46
Figura 20	Mutação Real do terceiro gene.	47
Figura 21	Mercedes-Benz Actros 4844.	48
Figura 22	Ilustração do trajeto do caminhão carregado.	50
Figura 23	Gráficos do consumo de combustível e distância percorrida por etapa de operação do caminhão.	52
Figura 24	Planilha com dados tratados.	54
Figura 25	Parametrização do RLM no Excel.	55

Figura 26	Combinações dos Métodos Implementados.	57
Figura 27	Interface do Sistema Desenvolvido.	58
Figura 28	Fluxograma de Utilização do Sistema Desenvolvido.	59
Figura 29	Relatório da Regressão Linear Múltipla no Excel®.	61
Figura 30	Gráficos da evolução de consumo real versus RLM.	64
Figura 31	Gráfico da evolução de consumo real versus AG configuração 1.	64
Figura 32	Gráfico da evolução de consumo real versus AG configuração 2.	64
Figura 33	Rota de Padrão para teste da previsão de consumo.	65
Figura 34	Mapa de calor do consumo da Rota de Padrão para o modelo do AG configuração 1.	66

Lista de tabelas

Tabela 1	Tabela ANOVA para RLM	37
Tabela 2	Representação Binária de Variáveis Reais	41
Tabela 3	Amostra dos dados coletados.	49
Tabela 4	Variáveis relativos ao deslocamento entre os pontos coletados.	50
Tabela 5	Variáveis relativos ao deslocamento com valores acumulados.	51
Tabela 6	Estrutura dos cromossomos.	54
Tabela 7	População de possíveis soluções geradas.	55
Tabela 8	Coefficientes do Modelo Obtido com RLM.	61
Tabela 9	Configurações do AG testadas.	62
Tabela 10	Coefficientes dos Modelos Obtidos com AG.	62
Tabela 11	Resultado dos Modelos Obtidos com AG.	63
Tabela 12	Comparativo entre Modelos Obtidos.	63
Tabela 13	Comparativo das Previsões de Consumo para Rota Padrão	65

Lista de siglas

AG Algoritmo Genético

ANOVA Analysis of Variance

ANP Agência Nacional do Petróleo, Gás e Biocombustíveis

CAN Controller Area Network

EPSG European Petroleum Survey Group

GPS Global Positioning System

hMOEA Hybrid Multiobjective Evolutionary Algorithm

MIT Massachusetts Institute of Technology

MVRP Multiobjective Vehicle Routing Problem

PIB Produto Interno Bruto

RLM Regressão Linear Múltipla

RLS Regressão Linear Simples

SIRGAS Sistema de Referência Geocêntrico para as Américas

UTM Universal Transversa de Mercator

WGS World Geodetic System

VRP Vehicle Routing Problem

Sumário

1	Introdução	17
1.1	Objetivo	18
1.2	Objetivos específicos	18
1.3	Organização da Dissertação	19
2	Operação de minas a céu aberto	21
2.1	Sistemas de despacho	22
2.1.1	Problema de Roteamento de Veículos (VRP)	22
2.1.2	Sistemas de despacho com Algoritmos Evolutivos	23
2.1.3	Sistema de Telemetria e Rede CAN	24
2.1.4	Sistema de posicionamento	24
2.2	Consumo de Combustíveis	25
3	Regressão Linear	28
3.1	Regressão Linear Simples	28
3.1.1	Determinação da equação de regressão linear simples	29
3.1.2	Estimador da variância	30
3.1.3	Coefficiente de determinação	31
3.2	Regressão Linear Múltipla	32
3.2.1	Determinação da equação de regressão linear múltipla	32
3.2.2	Estimador da variância	34
3.2.3	Coefficiente de determinação	34
3.2.4	Regressão Linear Múltipla com Excel	34
4	Algoritmo Genético	38
4.1	Conceitos Básicos	39
4.2	Representação Cromossômica	40
4.3	População Inicial	41

4.4	Função de Aptidão	42
4.5	Métodos de Seleção	42
4.5.1	Roleta	42
4.5.2	Torneio	43
4.5.3	Ranking	43
4.5.4	Elitismo	43
4.6	Cruzamento	43
4.7	Mutação	44
4.8	Critério de Parada	45
4.9	Representação Numérica	45
4.9.1	Métodos de Cruzamento Real	46
4.9.2	Mutação Real	47
5	Metodologia	48
5.1	Coleta dos dados	48
5.2	Tratamento dos dados	49
5.3	Análise dos dados	51
5.4	Formulação do Problema	52
5.5	Aplicação da Regressão Linear Múltipla	53
5.6	Algoritmo Genético Implementado	54
5.6.1	Função de Aptidão do AG Implementado	55
5.6.2	Operações Genéticas	57
5.7	Sistema Desenvolvido	57
6	Resultados	60
6.1	Obtenção do modelo com RLM	60
6.2	Obtenção dos modelos com AG	61
6.3	Comparação dos Modelos Obtidos	62
6.4	Aplicação dos Modelos em uma Rota Padrão	63
7	Conclusões	67
8	Trabalhos em Eventos	69
	Referências	70

Introdução

Historicamente a mineração tem importância relevante na economia brasileira, desde o tempo de colônia onde a extração do ouro destacava-se como uma das principais atividades. Atualmente no Brasil sua participação no PIB é de quase 5%, fornece produtos para uma diversidade de indústrias e contribui na criação de inúmeros empregos diretos e indiretos (VALE, 2019).

O mercado cada vez mais competitivo cria nas mineradoras a necessidade de aprimoramento de seus processos e serviços como forma de sobrevivência. A busca pela otimização torna-se essencial para o melhor aproveitamento de recursos e maximização dos resultados.

Na mineração, usualmente os processos se dão em grande escala, o que contribui para que melhorias, ainda que pequenas, possam gerar resultados relevantes. O aproveitamento eficiente de recursos requer o investimento em aplicações inteligentes, visando o desenvolvimento do setor que é essencial para o crescimento econômico do país (ALEXANDRE, 2010).

As mineradoras realizam as atividades em minas subterrâneas ou a céu aberto. O setor responsável pela operação da mina possui um alto custo operacional, sendo o transporte de material extraído na frente de lavra até o ponto de descarga (britador, patio de estoque, etc.) um dos principais componentes.

O transporte de material em mina a céu aberto, comumente realizado por caminhões, pode atingir até 60% dos custos totais de operação (ERCELEBI; BASCETIN, 2009), possui assim, alto impacto econômico, por envolver elevados custos operacionais e de capital (QUEVEDO; DIALLO; LUSTOSA, 2009). Este custo elevado se deve ao grande investimento que se requer para aquisição de equipamentos de carga e transporte, assim como outros recursos, tais como pessoal, combustível e custo de manutenção dos equipamentos (CETIN, 2004). O problema do transporte é mover o material retirado da mina de modo que o custo seja minimizado (GERSHON, 1982). Logo, uma pequena redução percentual nos custos de transporte resulta em economias significativas (ALARIE; GAMACHE, 2002).

O consumo de combustível por caminhões e equipamentos de lavra, por sua vez, pode representar até 60% dos custos com transporte (SAHIN et al., 2009), e conseqüente parcela

notória nos custos totais de operação de uma mina a céu aberto. Logo, é razoável afirmar que a otimização no consumo de combustível pode trazer impacto significativo na redução dos custos de operação, reforçado pela crescente alta nos preços deste insumo.

Ainda temos uma questão social que demanda a redução das emissões de carbono, fazendo com que as tecnologias e estudos que propiciam uma redução no uso de combustível sejam bem vistas pelas mineradoras (LOPES, 2010).

1.1 Objetivo

O objetivo deste trabalho é fornecer o modelo matemático e mapa de calor do consumo de combustível de um caminhão de mineração em operação em uma mina à céu aberto, a partir de dados reais coletados do sistema de telemetria. A ideia é relacionar o consumo com características físicas das rotas. Fornecendo parâmetros para otimização da operação de minas, redução de custos e maximização de resultados operacionais.

1.2 Objetivos específicos

Além do objetivo principal, também buscou-se:

- ❑ Propor a utilização de Regressão Linear Múltipla (RLM) em problema de consumo de combustível na mineração;
- ❑ Propor a utilização de Algoritmos Genéticos (AG) em problema de consumo de combustível na mineração;
- ❑ Caracterizar o consumo de combustível de um caminhão na mineração;
- ❑ Comparar os modelos matemáticos do consumo obtidos pelo RLM e AG;
- ❑ Desenvolver um sistema para o tratamento dos dados, parametrização e operação do AG, validação dos modelos obtidos e criação do mapa de calor do consumo de combustível;
- ❑ Apresentar função aptidão que pode ser incorporada à algoritmos evolutivos usados em sistemas de despacho;
- ❑ Fornecer informações relevante para o sistema de despacho.

O trabalho tem com destaque a proposta da utilização de Regressão Linear Múltipla (RLM) e Algoritmo Genético na modelagem matemática e mapeamento do consumo de combustível do problema, assim como o desenvolvimento de um sistema para auxiliar o uso destas ferramentas na busca de soluções otimizadas para o mesmo.

Diante da quantidade reduzida de trabalhos com a aplicação das ferramentas de Regressão Linear Múltipla e Algoritmo Genético na mineração, especificamente na modelagem e mapeamento do consumo de combustível de caminhões, o presente estudo faz-se relevante.

1.3 Organização da Dissertação

O presente trabalho é dividido em oito capítulos e referências bibliográficas.

O Capítulo 1 é destinado à Introdução, na qual tem-se as motivações que deram origem ao trabalho e os principais objetivos propostos.

O Capítulo 2 trata da Operação de Minas a Céu Aberto, descrevendo os principais pontos do processo, destaca-se os sistemas de despacho e telemetria. Ainda é apresentado tópicos acerca do consumo de combustíveis.

O Capítulo 3 traz os conceitos da Regressão Linear, estabelece os modelos das regressões simples e múltipla, demonstra os cálculos de estimação dos coeficientes dos modelos, dos estimadores de variância e coeficientes de determinação. Ainda neste encontram-se orientações do uso do Excel na estimação do modelo de regressão linear múltipla.

O Capítulo 4 aborda a teoria dos Algoritmos Genéticos, são definidos primeiramente conceitos básicos, termos usados e um fluxograma do algoritmo, então é tratado a representação cromossômica, geração da população inicial e da função de aptidão. Na sequência tem-se os principais métodos de seleção (roleta, torneio, ranking e elitismo) e outras operações genéticas como cruzamento e mutação. Neste capítulo ainda destaca-se os critérios de parada, a representação cromossômica numérica e seus operadores numéricos.

O Capítulo 5 contém a metodologia do trabalho proposto, trazendo inicialmente os aspectos da coleta e tratamentos dados, apontando os pontos levantados e considerações assumidas com a análise do dados. Logo após tem-se a formulação do problema e sua aplicação na regressão linear múltipla. É apresentado ainda o algoritmo genético implementado e a definição da sua função de aptidão. O capítulo é encerrado com o sistema desenvolvido, o qual se destina para o tratamento dos dados, parametrização e operação do AG, criação da planilha para regressão, validação dos modelos obtido e criação do mapa de calor do consumo de combustível.

O Capítulo 6 apresenta os resultados, começando com a obtenção e análise de modelos para o problema proposto através de regressão linear múltipla e algoritmo genético. Logo após é mostrado a comparação dos modelos, que foram testados e validados para avaliação do desempenho na solução do problema. Finalizando o capítulo tem-se os resultados da aplicação dos modelos em um rota padrão para realização da previsão de consumo e também a geração do mapa de calor do consumo no trajeto da rota.

O Capítulo 7 dedica-se as conclusões e discussões do trabalho proposto, além de sugestões para trabalhos futuros.

O Capítulo 8 traz os trabalhos apresentados em eventos.

Por fim é apresentado as referências bibliográficas que foram utilizados para realização deste trabalho.

Operação de minas a céu aberto

As minas a céu aberto representam a grande maioria da superfície de trabalho (HARTMAN; MUTMANSKY, 2002) em oposição às minas subterrâneas. Sua operação consiste na organização de equipamentos de carga e transporte no processo de lavra. As frentes, também chamadas de ponto de extração, são os locais onde é realizado a retirada de material pelos equipamentos de carga, como por exemplo, as escavadeiras, que retiram e transferem o material para os equipamentos de transporte, normalmente caminhões, estes levam o material até o ponto de basculamento, onde o material é depositado, podendo ser uma pilha de estéril, quando o material não é adequado ao processo; pilha de homogeneização, quando deseja-se misturar o minério antes do começar o beneficiamento ou a quantidade extraída é maior que a capacidade de processamento; e unidade de beneficiamento, quando o minério será processado (RODRIGUES, 2016).

A lavra com transporte por caminhões em mina de céu aberto, segundo Lopes (LOPES, 2010), consiste de desmonte do material “in situ” que pode ser executado por meio de perfuração, explosivos ou mecanicamente. Dependendo da geologia do minério lavrado é a própria escavadeira ou carregadeira a responsável por escavar e carregar os caminhões para o transporte até a estação de britagem e descarga no silo.

Em minas a céu aberto a alocação de caminhões se dá basicamente de duas formas: estática e dinâmica. Os caminhões na alocação estática possuem rota fixa durante um determinado período, realizam assim, o processo de carga e descarga em pontos pré-estabelecidos. Na alocação dinâmica os caminhões não recebem rotas fixas, assim depois de cada viagem, eles recebem orientações dos pontos onde devem realizar a próxima carga e descarga (RODRIGUES, 2016).

Segundo Kolonja (KOLONJA; KALASKY; MUTMANSKY, 1993), a operação de minas historicamente se dava com a alocação estática de um caminhão à um determinado equipamento de carga, de acordo com sua capacidade de contribuição nos objetivos de produção a curto prazo.

De acordo com Rodrigues (RODRIGUES, 2016) a alocação estática ainda é o método mais utilizado nas mineradoras, porém apresenta menor produtividade devido a formação

de filas e ociosidade dos equipamentos de cargas, sendo mais comum em mina de médio e pequeno porte.

2.1 Sistemas de despacho

As minas a céu aberto possuem diversos pontos de extração e também de basculamento, criando então uma grande possibilidade de rotas. A Figura 1 apresenta a representação de parte das rotas possíveis em uma determinada mina a céu aberto.

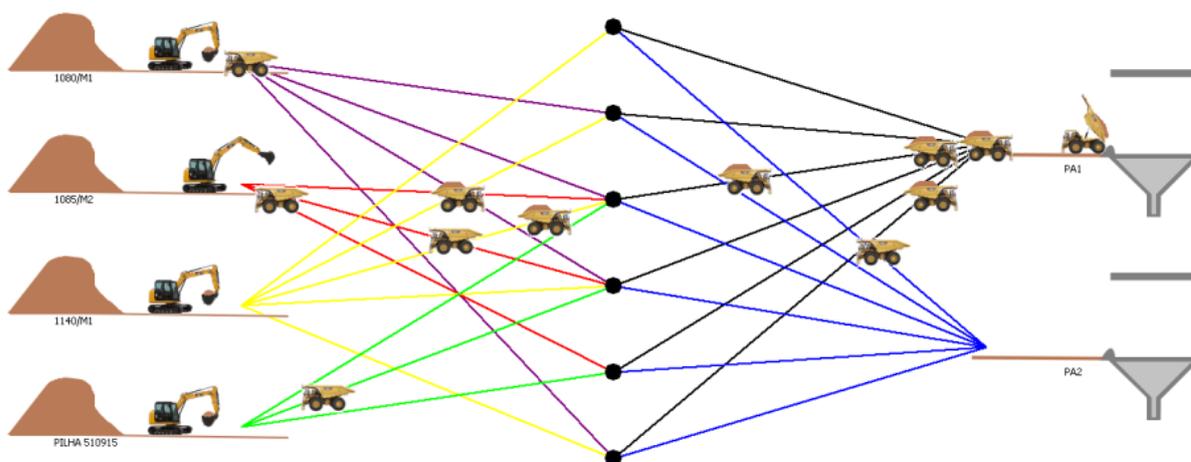


Figura 1: Representação de rotas possíveis na operação de minas a céu aberto.

Fonte: Torres (TORRES, 2017)

A operação de equipamentos de carga e caminhões possuem uma contribuição importante nos custos operacionais, assim a otimização da alocação pode representar resultados significativos.

Sistemas de despacho podem ser utilizados com objetivo de otimizar a operação e a alocação de equipamentos. Sendo responsáveis por gerar a alocação dinâmica de caminhões, utilizando critérios de operação que buscam maximizar a utilização de equipamentos para aumentar a produtividade.

2.1.1 Problema de Roteamento de Veículos (VRP)

Os sistemas de despachos buscam resolver problemas do tipo VRPs (Vehicle routing problem). O Problema de Roteamento de Veículos é amplamente estudado por pesquisadores na busca por rotas que otimizem o transporte e distribuição de produtos. Trata-se de um problema de otimização combinatória, que busca encontrar um conjunto de rotas para atender um certo número de clientes a partir de um ou mais depósitos. Na Figura 2, (NEO, 2019), pode-se ver uma entrada (à esquerda) e a possível saída (à direita) de um VRP. Na entrada temos na parte central um depósito e em volta deste, os clientes

que devem ser atendidos. Na saída temos os mesmos componentes, porém agora com as possíveis rotas de atendimento.

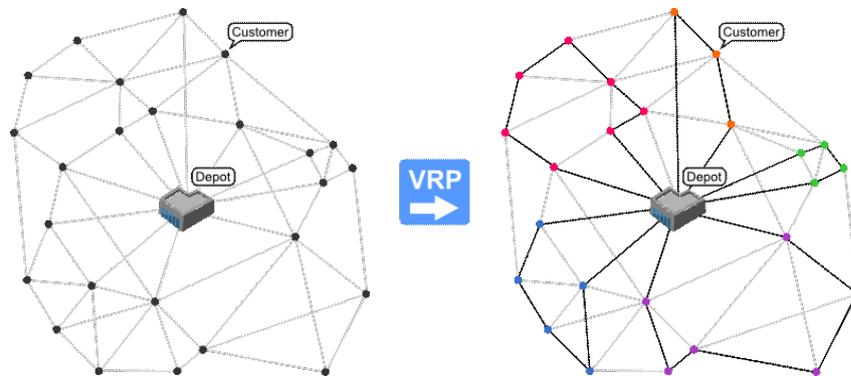


Figura 2: Entrada e saída de um VRP.

Fonte: NEO (NEO, 2019)

O primeiro trabalho que registra este estudo é o de Dantzing (DANTZIG; RAMSER, 1959), que aplicaram o conceito na distribuição de combustível. Desde então, diversos trabalhos, com diferentes metodologias e abordagens buscam encontrar soluções que minimizem o custo de transporte com base na redução da distância global percorrida, bem como a redução do número de veículos destinados a determinada tarefa e também do tempo de entrega.

Hjorring e Holt (HJORRING; HOLT, 1999) afirmam que a solução de roteamento de veículos se situa na classe dos problemas difíceis de serem resolvidos e por isso demandam um elevado custo computacional, isto é, muito processamento.

2.1.2 Sistemas de despacho com Algoritmos Evolutivos

Uma boa estratégia é a utilização de algoritmos evolutivos para a alocação dinâmica nestes sistemas e solução dos VRPs. Um dos motivos do uso de AG é que eles são robustos, eficientes e podem trabalhar com uma ampla faixa de funções objetivo e encontrar múltiplas rotas de modo simultâneo (WARDLAW; BHAKTIKUL, 2004).

Alexandre (ALEXANDRE, 2010) buscou desenvolver uma ferramenta com interface gráfica para simulação de uma mina genérica, além da possibilidade do acoplamento de algoritmos responsáveis pelo despacho, baseados em algoritmos evolutivos multiobjetivo e metodologia MVRP (Multiobjective Vehicle Routing Problem).

Mendes (MENDES et al., 2016) utilizou uma abordagem multiobjetivo para o problema de despacho de veículos em minas a céu aberto. Foi proposto um algoritmo evolucionário multiobjetivo, chamado de Hybrid MOEA (hMOEA) para resolução do problema com alocação dinâmica dos veículos.

Torres (TORRES, 2017) apresentou uma rede de sistemas embarcados em caminhões para otimização de despacho em mineração com múltiplas rotas. Utiliza-se algoritmo

genético no despacho de caminhões em mina a céu aberto, de modo a orientar o operador para que o mesmo siga a melhor rota, visando à redução de perdas e o aumento da produção.

2.1.3 Sistema de Telemetria e Rede CAN

Os sistemas de despacho obtêm informações dos equipamentos em operação geralmente por meio de sistemas de telemetria, os quais permitem o monitoramento contínuo, onde dados em tempo real são enviados para sistemas de controle e operação (PORTO et al., 2015). A telemetria recebe as informações de módulos de sensores instalados nos equipamentos da mina e de dados da rede CAN (Controller Area Network) dos próprios caminhões.

A rede CAN é um barramento serial, desenvolvido para o monitoramento de sensores e atuadores em sistema automotivos (FERNANDES, 2016). Esta rede garante uma comunicação segura e eficiente entre sensores, atuadores e controladores das aplicações fundamentais de veículos como controle de freios, motor, transmissão, dentre outras, mesmo em ambientes hostil a transmissão de dados (SOUZA; CAMPOS, 2017).

2.1.4 Sistema de posicionamento

Para o controle da mina em tempo real os sistemas de despachos precisam saber a localização dos equipamentos em operação, assim sistemas de posicionamento são utilizados para que a alocação dinâmica seja eficiente.

Estes sistemas possibilitam a vinculação das medições topográficas com pontos obtidos por GPS, assim obtêm-se pontos georreferenciados, normalmente, em coordenadas UTM (Universal Transversa de Mercator), vinculadas ao Sistema Geodésico Brasileiro (atualmente SIRGAS2000 e SAD-69), (MUNDOGEO, 2019).

A projeção UTM é um sistema analítico que visa minimizar as deformações de um mapa a níveis toleráveis representando um sistema ortogonal. Por definição o sistema de projeção UTM, é um sistema plano o qual divide o globo por fusos, Figura 3.

Dentro desse sistema com a finalidade de organização o Grupo de Pesquisa Petrolífera Européia (EPSG) classificou todos os sistemas de referências por códigos (Código EPSG) ou seja qualquer lugar do mundo pode ser identificado através desse padrão.

Datum é um conjunto de elementos que permite dimensionar o elipsoide e posicioná-lo relativamente ao Geóide. Visto que ao se representar a terra temos as seguintes formas: Elipse, Elipsóide e Geóide, este último é o que mais se aproxima com a realidade da terra, e os dois primeiros são formas de representações matemáticas para facilitar os cálculos realizados em cartografia.

O datum é classificado de acordo com a região abrangida, quanto ao tipo de coordenada e quanto as técnicas utilizadas. Em suma temos os datums locais que são definidos de

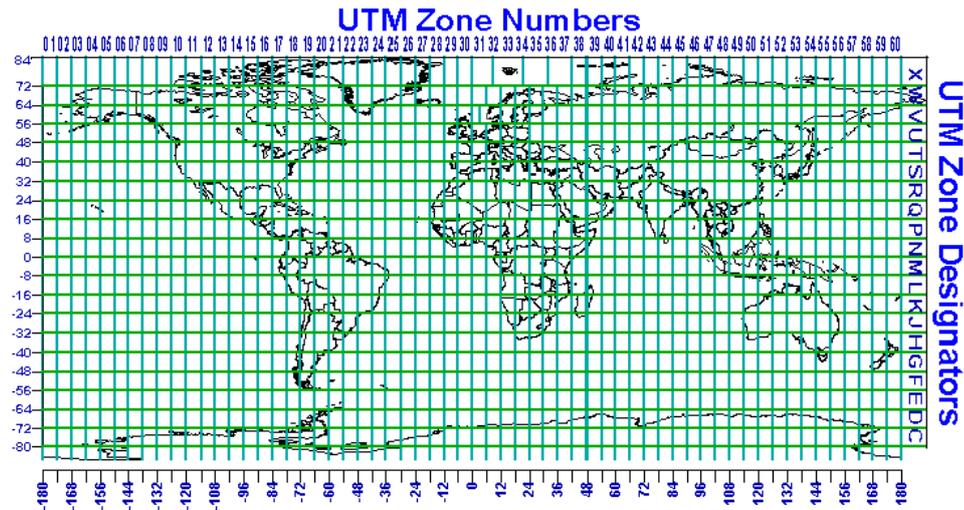


Figura 3: Sistema de coordenadas UTM.

Fonte: Ferreira (FERREIRA, 2017)

modo tal que a superfície de referência (no caso o Elipsóide) se ajuste bem a uma dada região da terra, exemplo SIRGAS 2000. E o datum global que são definidos de modo a garantir um ajuste de toda a superfície da Terra, exemplo WGS 84.

2.2 Consumo de Combustíveis

Dentre os vários estudos sobre VRPs e alocação dinâmica, a grande maioria está direcionado na busca pela a otimização de rotas através da redução distância total percorrida e redução do número de veículos.

Entretanto outro aspecto que as mineradoras estão realmente interessadas é a redução dos custos com transporte. Deste modo, deve-se direcionar a busca pela otimização dos principais componentes dos custos, como por exemplo o consumo de combustível.

Apesar do consumo de combustível estar muito ligado a distância percorrida, outros fatores também podem influenciá-lo, como a peso da carga transportada, modo de condução do veículo, velocidade, aspectos físicos da rota, como inclinações da pista e resistência do ar.

Segundo o trabalho de Sahin (SAHIN et al., 2009), o custo com combustível representa 60% dos custos totais para o transporte. Logo, podemos afirmar que qualquer redução deste, representa um impacto significativo. Principalmente se considerarmos a crescente alta nos preços do combustível conforme mostrado pela Figura 4, que traz a trajetória dos preços nos últimos anos, segundo Agência Nacional do Petróleo (ANP) (ANP, 2019).

Além disso, devemos ressaltar o benefício para toda sociedade e meio ambiente, com a redução da emissão de dióxido de carbono (gás do efeito estufa).

O trabalho de Xiao (XIAO et al., 2012), apresenta o desenvolvimento de um modelo

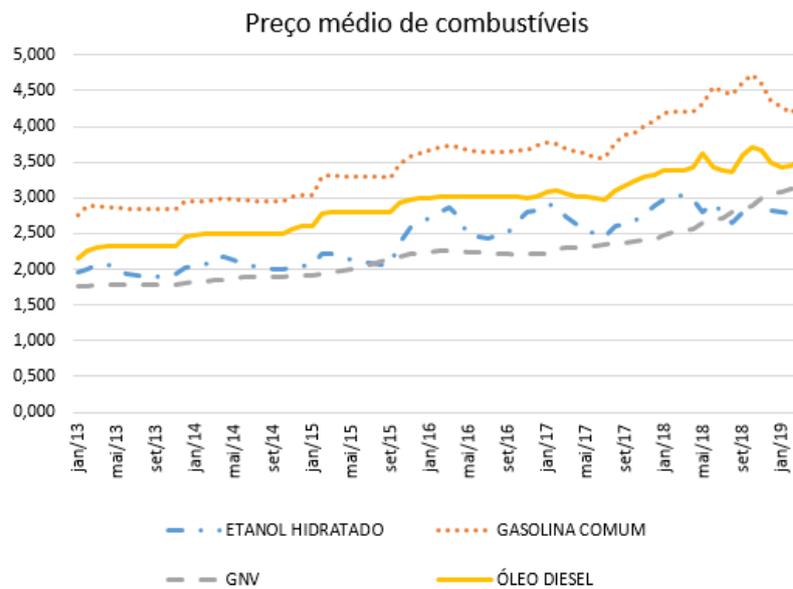


Figura 4: Trajetória dos preços dos principais combustíveis.

Fonte: Elaborado pelo Autor

de otimização do consumo de combustível, ele apresenta um exemplo, Figura 5, onde o modelo obtido em seu trabalho é aplicado em um problema de rotas simples. Podemos ver que: a menor distância não garantiu o menor consumo de combustível e que mesmo em rotas com distâncias iguais, o consumo pode ser diferente.

Neste sentido, alguns pesquisadores tem destacado a importância do estudo do consumo de combustíveis.

Gaur (GAUR; MUDGAL; SINGH, 2013) propõe o estudo do problema de roteamento de veículos, focado na minimização do consumo de combustível, como forma de redução do custo de operação de transporte.

Xavier (XAVIER et al., 2010) procurou realizar minimização do consumo de combustível para rotas de coleta de lixo, utilizando técnicas de programação linear associada a métodos heurísticos, além do algoritmo apresentado por (DANTZIG; RAMSER, 1959) para encontrar o menor caminho para o aterro.

Campos Júnior (JÚNIOR et al., 2013) buscou a redução do consumo de diesel em caminhões fora de estrada, através do desenvolvimento de uma metodologia de monitoramento e controle para a sistematização da gestão de desempenho, através dos métodos “Top 10” e PDCA, além de rotinas de manutenção.

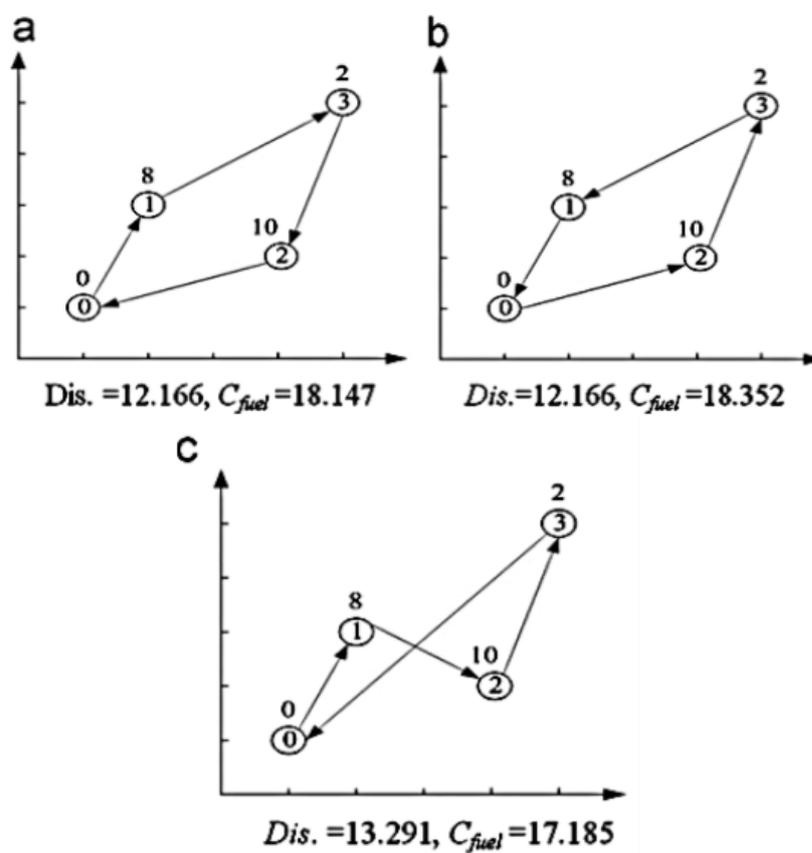


Figura 5: Exemplo de rotas com distância e consumo.

Fonte: Xiao (XIAO et al., 2012)

Regressão Linear

Quando trabalha-se com dados de diversas variáveis, uma questão importante é saber se existe uma relação entre elas. Por exemplo, o grau de interferência do tempo de estudo e da condição financeira no desempenho acadêmico de um aluno. A regressão linear é um conjunto de ferramentas que busca encontrar o relacionamento entre variáveis estudadas.

A regressão simples nos permite a análise da relação entre duas variáveis, uma variável dependente e outra independente. Já a regressão múltipla, que é uma extensão da regressão simples, nos possibilita a investigação da relação de diversas variáveis independentes com uma dependente, o que a torna uma ferramenta mais flexível, versátil e complexa, possibilitando sua utilização em uma enorme variedade de problemas (NEUFELD; CELESTE, 2003).

3.1 Regressão Linear Simples

Quando o conjunto de dados é submetida a uma regressão linear simples (RLS), podemos encontrar a relação entre duas variáveis, sendo x chamado de variável independente ou preditora e y como variável dependente ou de resposta. A regressão resulta numa equação que fornece a previsão do valor y para um certo valor de x . Essa equação é chamada de linear porque a relação apresentada entre as duas variáveis resulta em uma linha reta.

O modelo da regressão linear simples pode então ser definido da seguinte forma (BARROSO et al., 1987):

$$Y = B_0 + B_1x + \epsilon \quad (1)$$

Onde:

B_0 - intercepto da reta;

B_1 - inclinação da reta;

ϵ - erro aleatório do modelo.

A inclinação B_1 (coeficiente angular da reta) é uma das grandezas mais importantes em regressões lineares, representa a mudança esperada em Y para cada unidade de x , um valor muito próximo de zero indica pouca ou nenhuma relação entre as variáveis analisadas, altos valores positivos ou negativos indicam grande relação. Já o intercepto B_0 representa o valor de Y quando $x = 0$. A variável ϵ é chamada de erro aleatório do modelo, sem a qual o par observado (x, y) corresponde a um ponto disposto exatamente na reta de regressão, a inclusão deste termo no modelo permite que (x, y) esteja acima ou abaixo da reta, conforme Figura 6.

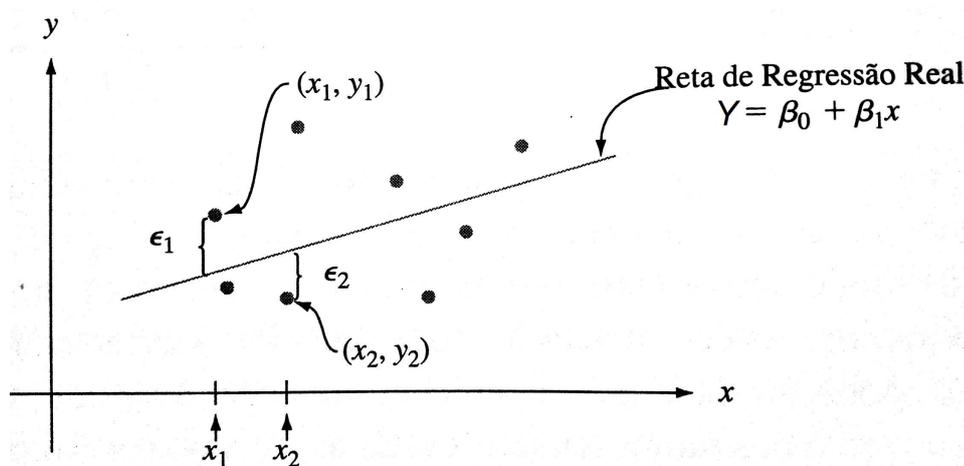


Figura 6: Observações do modelo de regressão linear simples.

Fonte: Devore (DEVORE, 2006)

3.1.1 Determinação da equação de regressão linear simples

Com base numa amostra de dados é preciso encontrar a reta da regressão linear simples que ofereça a melhor aderência aos pontos dos dados observados:

$$\hat{Y} = b_0 + b_1 x_i \quad (2)$$

Onde:

\hat{Y} - previsão de Y para uma observação x_i ;

x_i - valor de x para uma observação i ;

b_0 - estimador de B_0 ;

b_1 - estimador de B_1 .

O problema é determinar os valores dos parâmetros b_0 e b_1 de modo que a reta se ajuste satisfatoriamente aos pontos relativos aos dados. A reta oferece uma boa aderência, se os desvios dos pontos observados em relação à reta forem pequenos ($Y - \hat{Y}$). Assim a melhor reta é aquela que tem a menor soma possível do quadrados dos desvios. Há diversos métodos para se determinar os coeficientes, sendo o mais eficaz o Método dos

Mínimos Quadrados, que tem com objetivo minimizar a função M que é o somatório dos quadrados dos desvios, (MARTINS; DOMINGUES, 2017):

$$M = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (3)$$

Reescrevendo temos:

$$M = \sum_{i=1}^n [Y_i - (b_0 + b_1 x_i)]^2 \quad (4)$$

Os valores de minimização de b_0 e b_1 são identificados tomando-se as derivadas parciais e igualando-as a zero:

$$\frac{\delta M}{\delta b_0} = 0; \quad \frac{\delta M}{\delta b_1} = 0 \quad (5)$$

$$\begin{aligned} \frac{\delta M}{\delta b_0} &= -2 \sum (Y - b_0 - b_1 x_i) = 0 \\ \frac{\delta M}{\delta b_1} &= -2 \sum x_i (Y - b_0 - b_1 x_i) = 0 \end{aligned} \quad (6)$$

Obtendo assim as equações normais para determinação de b_0 e b_1 :

$$\begin{aligned} \sum Y_i &= n b_0 + b_1 \sum x_i \\ \sum x_i Y_i &= b_0 \sum x_i + b_1 \sum x_i^2 \end{aligned} \quad (7)$$

Os valores de $\sum Y_i$, $\sum x_i$, $\sum x_i Y_i$ e $\sum x_i^2$ são facilmente obtidos a partir da amostra de pares ordenados (x, y) . Assim pode-se estimar os valores de b_0 e b_1 que minimizam os desvios do valor estimado com a reta de regressão por:

$$\begin{aligned} b_1 &= \frac{S_{XY}}{S_{XX}} \\ b_0 &= \bar{y} - b_1 \bar{x} \end{aligned} \quad (8)$$

Onde:

$$\begin{aligned} S_{XY} &= \sum x_i Y_i - \frac{\sum x_i \sum Y_i}{n} \\ S_{XX} &= \sum x_i^2 - \frac{(\sum x_i)^2}{n} \\ \bar{y} &= \frac{\sum Y_i}{n} \\ \bar{x} &= \frac{\sum x_i}{n} \end{aligned} \quad (9)$$

3.1.2 Estimador da variância

O parâmetro σ^2 determina a variabilidade inerente no modelo de regressão, sendo que valores altos indicam grande dispersão dos pontos observados em relação a reta de regressão, já valores menores indicam uma boa proximidade. Uma maneira de determinar

o grau de variabilidade dos valores é dado pela variância S^2 que é um estimador não tendencioso de σ^2 , (DEVORE, 2006):

$$S^2 = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n - 2} \quad (10)$$

A variância então é dada pela divisão da soma dos quadrados dos erros pelo divisor $n - 2$, que é o grau de liberdade associada a estimativa. Como deve-se estimar b_0 e b_1 previamente, grau de liberdade sofre uma perda de 2. O desvio padrão da estimativa é a raiz quadrada de S^2 .

A soma dos quadrados dos erros denomina-se Variação Residual:

$$VR = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = S_{YY} - b_1 S_{XY} \quad (11)$$

Tem-se também a Variação total:

$$VT = S_{YY} = \sum_{i=1}^n (Y_i - \bar{Y})^2 = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n} \quad (12)$$

E ainda a Variação Explicada pela Variável Independente:

$$VE = b_1 S_{XY} = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 \quad (13)$$

3.1.3 Coeficiente de determinação

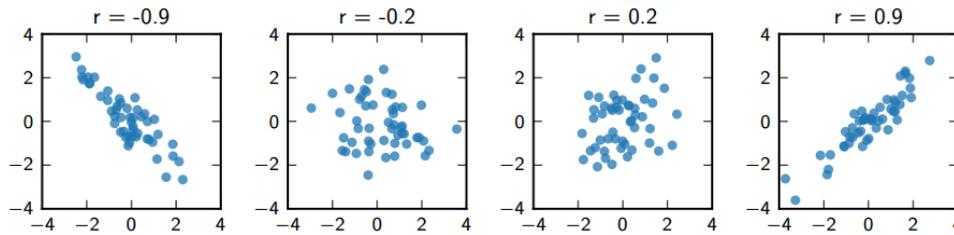
O coeficiente de determinação, também chamado de coeficiente de explicação, é interpretado como a proporção da variação de Y que é devida ao modelo de regressão:

$$r^2 = \frac{VE}{VT} = \frac{VT - VR}{VT} \quad (14)$$

Assim percebe-se que $0 \leq r^2 \leq 1$, sendo que quanto maior seu valor, mais o modelo de regressão consegue explicar as variações de Y , (MARTINS; DOMINGUES, 2017):

- ❑ Se $0.9 \leq r^2 < 1.0$ - Alto poder explicação do modelo.
- ❑ Se $0.8 \leq r^2 < 0.9$ - Bom poder explicação do modelo.
- ❑ Se $0.6 \leq r^2 < 0.8$ - Médio poder explicação do modelo.
- ❑ Se $r^2 < 0.6$ - Poder de explicação ruim do modelo.

O coeficiente de determinação (r^2) é o quadrado do coeficiente de correlação linear (r), que indica o grau de relação linear entre as variáveis $-1 \leq r \leq 1$. A Figura 7 apresenta gráficos de dados com determinados coeficientes de correlação (r), valores afastados de 0 indicam um relacionamento mais forte do que valores mais próximos de 0 e valores negativos indicam uma relação inversa, enquanto valores positivos indicam uma relação direta.

Figura 7: Gráficos de diferentes r .

Fonte: MIT (MIT, 2019)

3.2 Regressão Linear Múltipla

Muitos problemas na prática não conseguem ser explicados com o uso da Regressão Linear Simples, pois normalmente envolvem mais de duas variáveis. Uma alternativa é o uso da Regressão Linear Múltipla (RLM), que nada mais é do que uma extensão da anterior. Na RLM porém não temos a restrição de duas variáveis, pois permite mais de uma variável independente, tornando assim uma ferramenta mais versátil e aplicável a uma gama maior de problemas (NEUFELD; CELESTE, 2003).

O modelo de Regressão Linear Múltipla busca relacionar uma variável dependente Y com mais de uma variável independente, seja k o número de variáveis independentes e assumindo que $k \geq 2$, o modelo pode ser apresentado da seguinte forma:

$$Y = B_0 + B_1x_1 + B_2x_2 + \dots + B_kx_k + \epsilon \quad (15)$$

Onde:

Y - Variável dependente;

x_1, x_2, \dots, x_k - Variáveis independentes;

$B_0, B_1, B_2, \dots, B_k$ - coeficiente de contribuição das variáveis em relação à Y ;

ϵ - erro aleatório do modelo.

3.2.1 Determinação da equação de regressão linear múltipla

A medida que o número de variáveis independentes aumenta, os cálculos tornam-se extremamente trabalhosos, tornando assim o uso de softwares específicos para determinação da regressão linear vantajoso. Novamente, utiliza-se o Métodos dos Mínimos Quadrados para determinação dos coeficientes do modelo. Sendo a função de regressão linear múltipla expressa por:

$$\hat{Y} = b_0 + b_1x_{1i} + b_2x_{2i} + \dots + b_{ki}x_{ki} \quad (16)$$

Onde:

k - número de variáveis independentes;

\hat{Y} - previsão de Y para uma observação i ;

$x_{1i}, x_{2i}, \dots, x_{ki}$ - valor de x_1, x_2, \dots, x_k para uma observação i ;

$b_0, b_1, b_2, \dots, b_k$ - estimadores de $B_0, B_1, B_2, \dots, B_k$;

De modo similar ao apresentado na regressão linear simples, o objetivo do método é determinar os valores de $b_0, b_1, b_2, \dots, b_k$ que minimizem a função M , onde n representa o número de observações:

$$M = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (17)$$

Reescrevendo temos:

$$M = \sum_{i=1}^n [Y_i - (b_0 + b_1 x_{1i} + b_2 x_{2i} + \dots + b_{ki} x_{ki})]^2 \quad (18)$$

Os valores de minimização dos estimadores ($b_0, b_1, b_2, \dots, b_k$) são identificados tomando-se as derivadas parciais e igualando-as a zero:

$$\frac{\delta M}{\delta b_0} = 0; \quad \frac{\delta M}{\delta b_1} = 0; \quad \frac{\delta M}{\delta b_2} = 0; \quad \dots; \quad \frac{\delta M}{\delta b_k} = 0 \quad (19)$$

$$\begin{aligned} \frac{\delta M}{\delta b_0} &= -2 \sum (Y - b_0 - b_1 x_1 - b_2 x_2 - \dots - b_k x_k) = 0 \\ \frac{\delta M}{\delta b_1} &= -2 \sum x_1 (Y - b_0 - b_1 x_1 - b_2 x_2 - \dots - b_k x_k) = 0 \\ \frac{\delta M}{\delta b_2} &= -2 \sum x_2 (Y - b_0 - b_1 x_1 - b_2 x_2 - \dots - b_k x_k) = 0 \\ &\vdots \\ \frac{\delta M}{\delta b_k} &= -2 \sum x_k (Y - b_0 - b_1 x_1 - b_2 x_2 - \dots - b_k x_k) = 0 \end{aligned} \quad (20)$$

Fornecendo-nos o sistema de equações normais para determinação dos estimadores:

$$\begin{cases} \sum Y = n b_0 + b_1 \sum x_1 + b_2 \sum x_2 + \dots + b_k \sum x_k \\ \sum x_1 Y = b_0 \sum x_1 + b_1 \sum x_1^2 + b_2 \sum x_1 x_2 + \dots + b_k \sum x_1 x_k \\ \sum x_2 Y = b_0 \sum x_2 + b_1 \sum x_2 x_1 + b_2 \sum x_2^2 + \dots + b_k \sum x_2 x_k \\ \vdots \\ \sum x_k Y = b_0 \sum x_k + b_1 \sum x_k x_1 + b_2 \sum x_k x_2 + \dots + b_k \sum x_k^2 \end{cases} \quad (21)$$

Este sistema de equações pode ser apresentado de maneira mais conveniente se adotarmos matrizes e vetores, sendo que:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{kn} \end{bmatrix}, \quad B = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{bmatrix}.$$

Assim o sistema pode ser reescrito como:

$$X^T y = X^T X B \quad (22)$$

Portanto os estimadores são dados por:

$$B = (X^T X)^{-1} X^T y \quad (23)$$

3.2.2 Estimador da variância

Como na regressão linear simples, a variância σ^2 do erro do modelo de regressão linear múltipla é um indicador da qualidade de ajustamento da amostra. A variância amostral é dada por:

$$S^2 = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n - (k + 1)} \quad (24)$$

A variância então é dada pela divisão da soma dos quadrados dos erros pelo divisor $n - (k + 1)$, que é o grau de liberdade associada a estimativa. Como deve-se estimar b_0, b_1, \dots, b_k previamente, o grau de liberdade sofre uma perda de $k + 1$, k é relativo ao número variáveis independentes e o $+1$ é devido ao b_0 . O desvio padrão da estimativa é a raiz quadrada de S^2 .

Variação Residual que é a soma dos quadrados dos erros::

$$VR = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (25)$$

Variação total:

$$VT = \sum_{i=1}^n (Y_i - \bar{Y}_i)^2 = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n} \quad (26)$$

Variação Explicada pela Variável Independente:

$$VE = \sum_{i=1}^n (\hat{Y}_i - \bar{Y}_i)^2 \quad (27)$$

3.2.3 Coeficiente de determinação

Este coeficiente mede o quanto as variações de Y são explicadas pelo modelo obtido, quanto maior o seu valor mais as variações são explicadas ($0 \leq R^2 \leq 1$):

$$R^2 = \frac{VE}{VT} = \frac{VT - VR}{VT} \quad (28)$$

3.2.4 Regressão Linear Múltipla com Excel

O software Microsoft® Excel possui ferramentas que permitem a análise de regressão simples ou múltipla para um conjunto de dados, além de funções internas e assistente para criação de gráficos.

Primeiramente os dados devem ser colocados numa planilha do Excel, e dispostos de maneira similar ao mostrado na Figura 8, que traz um exemplo de dados para regressão, onde temos duas variáveis independentes (X_1, X_2), uma variável dependente (Y) e um total de doze observações.

	A	B	C	D
	Mês	Despesa Treinamento (X_1)	Despesa Marketing (X_2)	Lucro (Y)
1				
2	Janeiro	5	3	16
3	Fevereiro	5	6	16
4	Março	5	12	27
5	Abril	10	3	18
6	Mai	10	6	20
7	Junho	10	12	28
8	Julho	15	3	26
9	Agosto	15	6	27
10	Setembro	15	12	32
11	Outubro	10	5	19
12	Novembro	10	10	27
13	Dezembro	10	15	30

Figura 8: Exemplo de dados para regressão.

Fonte: Elaborado pelo autor

Com os dados preparados executa-se a ferramenta Análise de Dados, localizada na aba “Dados”, Figura 9, e depois seleciona-se a opção de Regressão.

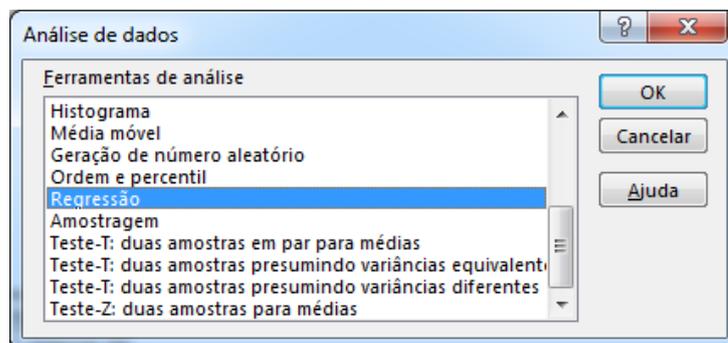


Figura 9: Ferramenta de Análise de Dados do Excel.

Fonte: Elaborado pelo autor

Na caixa de diálogo da ferramenta de Regressão é preciso informar os campos relativos aos dados na planilha:

- Intervalo Y de entrada: informar a faixa de dados referente a variável dependente Y (incluindo o nome da variável);
- Intervalo X de entrada: informar a faixa de dados referente as variáveis independentes X_1 até X_k (incluindo os nomes das variáveis);

- Habilitar a opção Rótulos.

Na Figura 10 temos a ferramenta de Regressão do Excel com as informações dos dados apresentados na Figura 8.

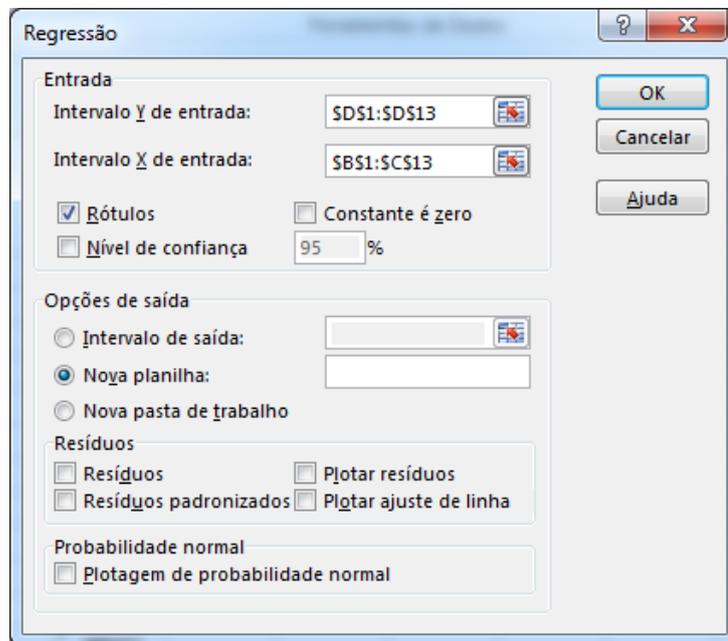


Figura 10: Ferramenta de Regressão do Excel.

Fonte: Elaborado pelo autor

Com essas configurações é gerado um relatório do resumo dos resultados da regressão em uma nova planilha, Figura 11. Esse relatório é composto por três tabelas: Estatística de regressão, ANOVA e Coeficientes.

A tabela Estatística de regressão apresenta:

- R múltiplo - é a raiz quadrada de R^2 (coeficiente de determinação). Esse valor é frequentemente chamado de coeficiente de correlação e é medida alternativa do relacionamento entre as variáveis independentes com a dependente;
- R-Quadrado - é coeficiente de determinação, R^2 ;
- R-Quadrado ajustado - uma variação do coeficiente de determinação que leva em conta o número de variáveis independentes consideradas;
- Erro padrão - é a raiz quadrada de s^2 , é uma medida de aderência;
- Observações - número de observações de dados usado para estimar o modelo de regressão.

A tabela ANOVA fornece os graus de liberdade (gl), a soma dos desvios quadráticos (SQ), a média dos quadrados (MQ), a estatística f (F) e o valor p para essa estatística f

	A	B	C	D	E	F	G
1	RESUMO DOS RESULTADOS						
2							
3	<i>Estadística de regressão</i>						
4	R múltiplo	0,9501939					
5	R-Quadrado	0,9028684					
6	R-quadrado ajustado	0,8812835					
7	Erro padrão	1,9481604					
8	Observações	12					
9							
10	ANOVA						
11		<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>	
12	Regressão	2	317,508705	158,75435	41,828877	2,77411E-05	
13	Resíduo	9	34,1579618	3,7953291			
14	Total	11	351,666667				
15							
16		<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>
17	Interseção	7,2488323	2,00201421	3,6207696	0,0055651	2,71996148	11,77770306
18	Despesa Treinamento	0,8666667	0,15906663	5,4484504	0,0004065	0,506832942	1,226500391
19	Despesa Marketing	1,0216561	0,13906566	7,3465734	4,343E-05	0,707067683	1,336244419

Figura 11: Relatório da Regressão do Excel.

Fonte: Elaborado pelo autor

(F de significação), que é o menor nível de significância com que se rejeitaria a hipótese nula.

A Tabela 1 apresenta a ANOVA com o preenchimento baseado nas variáveis apresentadas anteriormente para a Regressão Linear Múltipla.

Tabela 1: Tabela ANOVA para RLM

Fonte de variação	gl	Soma dos Quadrados	Quadrado Médio	f
Regressão	k	VR	$\frac{VR}{k}$	$\frac{VR}{k s^2}$
Erro	$n - (k + 1)$	VE	$s^2 = \frac{VE}{n - (k + 1)}$	
Total	$n - 1$	VT		

A tabela Coeficientes apresenta o valores dos coeficientes do modelo de regressão, o erro padrão, seus parâmetros estatísticos e o intervalo de valores dentro dos limites de confiança do modelo. A primeira linha é relativa ao coeficiente de intercepto (b_0) e as demais são os coeficientes das variáveis independentes.

Para os dados da Figura 8 e com base no relatório da Figura 11, temos o seguinte modelo de regressão linear múltipla:

$$Y = 7,2488323 + 0,8666667X_1 + 1,0216561X_2 \quad (29)$$

Algoritmo Genético

Charles Darwin apresentou em 1859 a teoria da evolução natural em seu trabalho origem das espécies, onde organismos biológicos se evoluem ao longo das gerações pelo princípio da seleção natural, (DARWIN, 1859). Os indivíduos em uma população competem por recursos e os da mesma espécie competem entre si para conquistar parceiros para reprodução. Assim indivíduos com baixa performance tendem a fracassar neste processo de seleção e possuem poucas chances de sobrevivência. Logo indivíduos mais preparados possuem maior probabilidade de sucesso, podendo sobreviver e deixar mais descendentes.

O processo de reprodução também pode gerar filhos com melhores habilidades do que os pais, devido à combinação de boas características dos progenitores, isto faz que a cada geração surjam indivíduos mais adaptados ao meio onde vivem.

Holland em seu trabalho “Adaptation in natural and artificial systems” em 1975, descreveu como aplicar o princípio da evolução natural em problemas de otimização, dando origem ao primeiro Algoritmo Genético (AG), (HOLLAND et al., 1992). Desde então estes foram fortemente estudados, tornando-se uma poderosa ferramenta para resolução de problemas.

Os AGs compõem a classe dos algoritmos evolucionários e podem ser definidos com uma técnica de busca baseada no processo biológico de seleção natural e reprodução genética (MITCHELL, 1998). As populações de indivíduos ou cromossomos são submetidas aos operadores genéticos para criar um processo de evolução natural. Deste modo, os indivíduos competem entre si, sendo que os mais aptos para solução do problema, tendem a sobreviver e deixar mais descendentes.

Os algoritmos evolucionários são técnicas heurísticas, não garante a qualidade da resposta, mas normalmente tendem a encontrar o solução ótima ou ficam próximos dela. Também fazem parte de técnicas de busca, classificada como aleatória-guiada, pois possui componentes aleatórios, mas dependem do estado atual para determinação do estado subsequente. O cenário conhecido guia a busca do algoritmo, o que diferencia de método puramente aleatórios (LINDEN, 2012).

4.1 Conceitos Básicos

A seguir temos alguns conceitos básicos:

- ❑ Cromossomo: cadeia de caracteres representando alguma informação relativa às variáveis do problema. O cromossomo representa uma possível solução.
- ❑ Gene: unidade básica do cromossomo.
- ❑ População: conjunto de cromossomos.
- ❑ Geração: número de iterações que o AG executa.
- ❑ Operações genéticas: operações que o AG realiza sobre os cromossomos.
- ❑ Espaço de busca: região que compreende as soluções possíveis.
- ❑ Função objetivo (aptidão): função que se quer otimizar e que usada para avaliar o desempenho de cada cromossomo.
- ❑ Método de seleção: método pelo qual os membros da população são selecionados para serem submetidos as operações genéticas.
- ❑ Mutação: função que altera conteúdo de um ou mais genes do cromossomo.
- ❑ Taxa de mutação: é a taxa que determina a probabilidade de mutação do indivíduo.
- ❑ Cruzamento: operador genético utilizado para cruzar dois indivíduos e gerar novos filhos.
- ❑ Taxa de cruzamento: é a taxa que determina a probabilidade de cruzamento entre dois indivíduos.

Na Figura 12, podemos ver o fluxograma de um algoritmo genético, onde iniciamos com a criação de uma população de indivíduos, logo após é realizado o teste por uma função de aptidão e os indivíduos são classificados pelo seu desempenho, em seguida é verificado se o critério de parada foi satisfeito, esse pode ser um desempenho específico, número de gerações (combinação dos indivíduos formando nova população), tempo computacional ou outro critério pré-estabelecido. Caso não satisfeito, dar-se início a criação de uma nova geração a partir da população atual, sendo os indivíduos selecionados, de acordo com o desempenho anterior, e recombinados “cruzamento”. Estes são submetidos ao processo de mutação, gerando novos indivíduos, que serão avaliados pela função de aptidão e o processo se repete até que o critério de parada seja satisfeito.

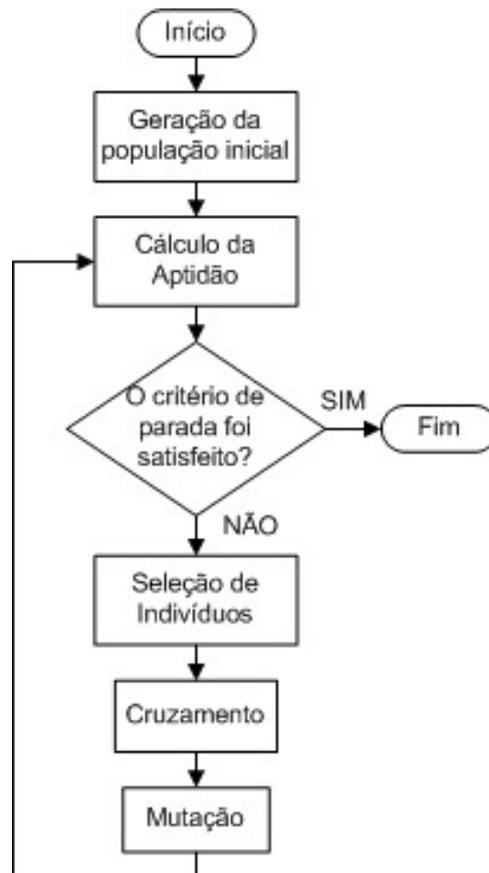


Figura 12: Fluxograma do Algoritmo Genético.

Fonte: Elaborado pelo Autor

4.2 Representação Cromossômica

A representação cromossômica é uma forma de padronizar as informações relativa ao problema de uma maneira a ser tratada computacionalmente pelo algoritmo. Assim os cromossomos carregam as informações relativas às características de possíveis soluções para o problema a ser resolvido. Os mesmos são formados por genes, que são a unidade básica, cada gene determina uma característica do indivíduo. Vale ressaltar que a representação cromossômica é completamente arbitrária, sua definição é determinada pela adequação ao problema .

A representação binária, Figura 13, é a mais simples e usada pelos desenvolvedores de AG, nela os cromossomos são uma sequência de bits, assim cada bit ou conjunto de bits representam uma características que depende do problema em questão.

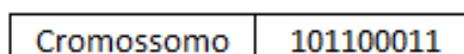


Figura 13: Estrutura do cromossomo em representação binária.

Fonte: Elaborado pelo autor

Para utilizar a representação binária num problema que trabalha com números reais deve-se primeiramente definir o intervalo de busca das variáveis e depois o número de bits de acordo com a precisão necessária. Na Tabela 2 tem-se o valor binário e o valor correspondente para um caso de três variáveis com intervalo de dados distintos.

Tabela 2: Representação Binária de Variáveis Reais

Código Binário	Intervalo 1	Intervalo 2	Intervalo 3
000	10,0	2,0	1,0
001	20,0	6,0	1,5
010	30,0	10,0	2,0
011	40,0	14,0	2,5
100	50,0	18,0	3,0
101	60,0	22,0	3,5
110	70,0	26,0	4,0
111	80,0	30,0	4,5

Assim o cromossomo com código 100 representaria no Intervalo 1 um indivíduo de valor 50,0, no Intervalo 2 o valor 18,0 e no Intervalo 3 o valor 3,0. Nota-se que, o primeiro intervalo é maior que os demais, logo a diferença entre um valor e o próximo é grande, o que lhe confere uma baixa precisão. Neste caso o valor ótimo para resolver o problema pode não ter sido representado. Para aumentar a precisão é necessário usar uma quantidade maior de bits de representação, deste modo o intervalo seria mais dividido e teríamos uma quantidade maior de valores representados, aumentando as chances de aproximar do valor ótimo.

4.3 População Inicial

A população é o conjunto de indivíduos candidatos a solução do problema, Figura 14. A população inicial na maioria dos trabalhos é gerada de forma aleatória dentro do espaço de busca, pode-se utilizar uma heurística para garantir uma melhor exploração do espaço. O tamanho da população é um ponto importante, trabalhar com grandes populações possibilita uma cobertura maior do espaço de busca, mas isto tem um grande custo computacional e aumenta consideravelmente o tempo de execução do AG.

Cromossomo 1	101100011
Cromossomo 2	001101111
Cromossomo 3	101011001
Cromossomo 4	011010110

Figura 14: População de cromossomos em representação binária.

Fonte: Elaborado pelo autor

4.4 Função de Aptidão

A função de aptidão ou função fitness é responsável por determinar o grau de adequação de um indivíduo como solução do problema, ela deve conter os objetivos, critérios e restrições.

Todos os indivíduos da população devem ser submetidos à função de aptidão, assim cada um recebe um valor numérico que expressa o grau de qualidade deste no problema em questão. Este valor, chamado de fitness, é utilizado posteriormente pelos métodos de seleção para determinar os progenitores da próxima geração. Desde modo, se o cromossomo 1 aproxima-se mais da solução do problema que o cromossomo 2, o fitness atribuído a ele deve ser maior, o que vai lhe conceder uma maior probabilidade de seleção para realizar o cruzamento e deixar descendentes.

4.5 Métodos de Seleção

A seleção consiste no processo que determina quais indivíduos da população atual irão ser submetidos as operações genéticas para formar novos indivíduos. O processo de seleção visa identificar os indivíduos mais aptos a realizarem os cruzamentos, de forma a garantir que seus genes permaneçam nas próximas gerações, e a população continue a evoluir na busca do ponto ótimo.

Um problema que aparece com frequência nos processos de seleção é o superindivíduo. Algumas vezes quando o processo de seleção é muito forte, um indivíduo com uma aptidão muito grande pode dominar a população, fazendo com que a diversidade seja reduzida e impedindo que a população evolua, assim pode-se encontrar um ponto próximo do ótimo, mas que não seja o ótimo. No sentido oposto, se o processo de seleção for fraco, a população terá uma aptidão média baixa, o que fará com que a evolução seja pequena (MITCHELL, 1998).

4.5.1 Roleta

O método da roleta é o mais comum para seleção de indivíduos. Consiste em definir uma probabilidade individual com base no fitness do indivíduo e da soma total dos fitness, e posteriormente definir uma probabilidade acumulada com base na probabilidade individual. Em cada intervalo de probabilidade acumulada está associado um indivíduo, assim, um número é sorteado, e o intervalo de probabilidade acumulada que contiver este número é escolhido (MITCHELL, 1998).

Este método apresenta o problema do superindivíduo, isto é, como a probabilidade acumulada é baseada na aptidão, indivíduos com aptidão alta terão uma probabilidade maior de serem selecionados, geralmente muito maior.

4.5.2 Torneio

O torneio consiste em escolher aleatoriamente um grupo de indivíduos e testar suas aptidões, aquele que apresentar a melhor aptidão dentre os escolhidos, será selecionado para realizar o crossover. Este método diminui a probabilidade de se selecionar um indivíduo com baixa aptidão, porém aumenta a chance de uma convergência prematura da população para um superindivíduo (SOARES, 1997).

4.5.3 Ranking

Na seleção por ranking os indivíduos são ordenados de acordo com sua aptidão, e a probabilidade de seleção é então baseada na sua posição relativa e não na aptidão. Este método diminui a probabilidade de ocorrência de superindivíduos, visto que a probabilidade de seleção dos indivíduos não se diferencia muito um em relação ao outro (LINDEN, 2012).

4.5.4 Elitismo

O elitismo transfere automaticamente os indivíduos com melhor desempenho para a próxima geração, isto garante que eles não se percam com os processos de cruzamento e mutação. Deve-se usar esse tipo de estratégia com uma baixa percentagem, para não forçar uma convergência no algoritmo, impedindo-o de encontrar soluções ainda melhores (JONG, 1975).

4.6 Cruzamento

O cruzamento é o processo onde dois indivíduos selecionados, chamados de pais, são recombinados para formar novos candidatos a solução do problema, chamados de filhos. A ideia é trocar as informações genéticas entre os pais, de modo a transmitir boas características para os filhos, criando assim uma geração melhor que a anterior.

Neste processo os filhos recebem pedaços da sequência de genes dos pais, como os indivíduos de melhor fitness tendem a ser selecionados para o cruzamento, a nova geração tendem a se aproximar mais da solução ótima.

Os métodos de cruzamento mais comuns são: o ponto único de corte e o de dois pontos. Entende-se como pontos de cortes possíveis, as posições entre dois genes, assim o número de possibilidades é sempre $n - 1$, onde n é o número de genes do cromossomo (BEASLEY; BULL; MARTIN, 1993).

No cruzamento de ponto único é realizado um corte num ponto escolhido aleatoriamente nos cromossomos pais. O primeiro filho é formado com a junção da sequência de genes à esquerda do corte no Pai 1, com a sequência à direita do corte no Pai 2. O

segundo filho é criado com a sequência à esquerda do Pai 2 e os genes à direita do Pai 1, conforme ilustrado na Figura 15.

Pai 1	1 0 1 1 0	0 0 1 1
Pai 2	0 0 1 1 0	1 1 1 1
Filho 1	1 0 1 1 0	1 1 1 1
Filho 2	0 0 1 1 0	0 0 1 1

Figura 15: Cruzamento de ponto único.

Fonte: Elaborado pelo autor

O cruzamento com dois pontos de corte tem a mesma ideia do ponto único, porém é realizado dois cortes em pontos aleatórios nos cromossomos pais, dividindo-os assim em três partes. Obtém-se os filhos fazendo uma troca da sequência de genes entre os dois pontos de cortes, conforme a Figura 16. O primeiro filho é formado com a junção das extremidades do Pai 1 e a parte central do Pai 2. O segundo filho é obtido com as extremidades do Pai 2 com a parte do meio do Pai 1.

Pai 1	1 0 1	1 0 0	0 1 1
Pai 2	0 0 1	1 0 1	1 1 1
Filho 1	1 0 1	1 0 1	0 1 1
Filho 2	0 0 1	1 0 0	1 1 1

Figura 16: Cruzamento com dois pontos de corte.

Fonte: Elaborado pelo autor

4.7 Mutação

A mutação é um operador genético de extrema importância, pois gera a biodiversidade dos indivíduos. Mesmo sendo normalmente aplicado com baixa probabilidade de ocorrência, possibilita que o algoritmo explore regiões além da área de convergência. Dentre as estratégias mais comuns estão a mutação aleatória e por troca.

Na mutação aleatória um gene é escolhido aleatoriamente e seu valor é alterado, o resultado deve produzir um indivíduo válido, dentro do espaço de busca, Figura 17.

A mutação por troca é realizado com o sorteio de duas posições dos genes e se dá através da permuta dos seus valores, Figura 18. Vale observar que no caso dos valores dos

Cromosso	1	0	1	1	0	1	1	1	1
Cromosso Mutado	1	0	1	1	0	0	1	1	1

Figura 17: Mutaç o aleat ria no sexto gene.

Fonte: Elaborado pelo autor

genes serem os mesmos, o resultado ser  um cromossomo sem altera o na sua sequ ncia de gen tica.

Cromosso	1	0	1	1	0	1	1	1	1
Cromosso Mutado	1	1	1	1	0	0	0	1	1

Figura 18: Muta o por troca do segundo com o s timo gene.

Fonte: Elaborado pelo autor

4.8 Crit rio de Parada

O AG realiza as opera es descritas continuamente, conforme o fluxograma da Figura 12, at  que o crit rio de parada seja satisfeito. Este deve ser estipulado de acordo com as caracter sticas o problema, condi es aceit veis e desempenho desejado (SIVANANDAM; DEEPA, 2007). Dentre os crit rios de parada comumente utilizados temos:

- ❑ N mero M ximo de Gera es: O algoritmo para quando um n mero espec fico de gera es foi atingido.
- ❑ Tempo de Execu o: O processo chega ao fim quando um intervalo de tempo espec fico foi decorrido.
- ❑ Estagna o do fitness: O AG para quando n o h  mudan a no melhor fitness da popula o, os indiv duos param de evoluir por um n mero determinado de gera es.
- ❑ Qualidade Atingida: O algoritmo para quando algum indiv duo atinge um valor de fitness dentro da faixa de desempenho aceit vel para o problema em quest o.

4.9 Representa o Num rica

Em muitos problemas a codifica o bin ria do cromossomo, que   um conjunto de informa es que representam uma poss vel solu o, se ad qua satisfatoriamente, por m

podemos trabalhar com variáveis reais, o que traz como benefício um menor esforço computacional e fato da precisão não depender do número de bits (JANIKOW; MICHALEWICZ, 1991).

A Figura 19 apresenta a codificação de um cromossomo em variáveis reais.

Cromossomo	3.15	0.16	2.45
------------	------	------	------

Figura 19: Estrutura do cromossomo em representação numérica.

Fonte: Elaborado pelo autor

4.9.1 Métodos de Cruzamento Real

Para realizar a combinação de indivíduos codificados por variáveis reais tem-se a alternativa da aplicação do operador aritmético (LINDEN, 2012). Nele os filhos são gerados a partir de parâmetros que multiplicam os genes dos pais. Destacam-se os métodos: operador Radcliffe e operador Wright:

- Radcliffe: baseado em um parâmetro fixo $\lambda \in [0, 1]$, que indica a carga genética de cada indivíduo selecionado na formação de novos indivíduos, (RADCLIFFE; SURRY, 1994). A seguir tem-se o algoritmo do operador Radcliffe:

$$Pai1 = [p_{m1}, p_{m2}, \dots, p_{mn}]$$

$$Pai2 = [p_{d1}, p_{d2}, \dots, p_{dn}]$$

$$Filho1 = [(\lambda) * p_{m1} + (1 - \lambda) * p_{d1}, \dots, (\lambda) * p_{mn} + (1 - \lambda) * p_{dn}]$$

$$Filho2 = [(1 - \lambda) * p_{m1} + (\lambda) * p_{d1}, \dots, (1 - \lambda) * p_{mn} + (\lambda) * p_{dn}]$$

- Wright: neste método são gerados três novos indivíduos que devem ser validados. Se todos estiverem no espaço de busca, dois devem ser selecionados para a próxima geração. Essa seleção pode ser feita por sorteio ou por avaliação dos dois melhores filhos (WRIGHT, 1991). Abaixo temos o algoritmo do operador Wright:

$$Pai1 = [p_{m1}, p_{m2}, \dots, p_{mn}]$$

$$Pai2 = [p_{d1}, p_{d2}, \dots, p_{dn}]$$

$$Filho1 = [(0, 5) * p_{m1} + (0, 5) * p_{d1}, \dots, (0, 5) * p_{mn} + (0, 5) * p_{dn}]$$

$$Filho2 = [(1, 5) * p_{m1} + (-0, 5) * p_{d1}, \dots, (1, 5) * p_{mn} + (-0, 5) * p_{dn}]$$

$$Filho3 = [(-0, 5) * p_{m1} + (1, 5) * p_{d1}, \dots, (-0, 5) * p_{mn} + (1, 5) * p_{dn}]$$

4.9.2 Mutação Real

O processo de mutação real, ocorre através do sorteio do gene a ser mutado e a troca do seu conteúdo por um valor gerado aleatoriamente dentro do espaço de busca possível. Na Figura 20, temos uma mutação ocorrendo no terceiro gene do cromossomo.



Cromossomo	<i>3.15</i>	<i>0.16</i>	<i>2.45</i>
Cromossomo Mutado	<i>3.15</i>	<i>0.16</i>	<i>1.13</i>

Figura 20: Mutação Real do terceiro gene.

Fonte: Elaborado pelo autor

Metodologia

5.1 Coleta dos dados

O trabalho foi desenvolvido a partir da coleta de dados reais da operação de um caminhão da marca Mercedes-Benz, modelo Actros 4844, Figura 21, basculante com tração 8x4, motor BlueTec 5 V6 de 435 cv de potência e transmissão totalmente automatizada de 12 marchas, o caminhão é indicado para aplicação fora-de-estrada e em canteiros de obras.



Figura 21: Mercedes-Benz Actros 4844.

Fonte: Mercedes-Benz (MERCEDES-BENZ, 2019)

Os dados foram coletados do banco de dados do sistema de despacho que obtêm as

informações a partir do sistema de telemetria e da Rede CAN dos equipamentos da mina (caminhões e escavadeiras). A mina é do tipo céu aberto e contém dezenas de caminhões como o utilizado neste trabalho.

A massa de dados contém informações relativas a um mês, Julho de 2018, da operação do caminhão, sendo o intervalo entre as amostras de 10 (dez) segundos, o que resultou em 181.862 amostras realizadas no período.

Temos informações sobre a localização, consumo acumulado, data, hora, velocidade, frente de trabalho, odômetro, operador e status do caminhão. A Tabela 3 mostra parte dos dados coletados, onde cada linha representa um ponto no deslocamento do caminhão em uma rota de operação.

Tabela 3: Amostra dos dados coletados.

Hora	Latitude	Longitude	Altitude (m)	Ocorrência	Odômetro (m)	Combustível Consumido (L)	Vel. (km/h)
07:22:30	296505,4186	7824043,5409	1122,1	Processo de Carga	193722395	128505,5	0
07:22:40	296507,1804	7824042,6420	1122,2	Processo de Carga	193722405	128505,5	9
07:22:50	296531,0953	7824030,5474	1123,6	Operando Cheio	193722445	128505,5	17
07:23:00	296588,0650	7824016,0617	1124,5	Operando Cheio	193722510	128506,0	27
07:23:10	296670,4798	7824008,8855	1124,9	Operando Cheio	193722605	128506,0	32
07:23:20	296764,5095	7824010,5106	1125,8	Operando Cheio	193722685	128506,0	35
07:23:30	296864,4853	7824011,2839	1126,0	Operando Cheio	193722785	128506,0	37
07:23:40	296969,0270	7824010,8132	1125,1	Operando Cheio	193722890	128506,0	39
07:23:50	297078,7996	7824011,5086	1126,1	Operando Cheio	193723000	128506,0	39
07:24:00	297188,0183	7824013,8578	1127,5	Operando Cheio	193723115	128506,5	39

Esta tabela apresenta na coluna “Hora” o horário que foi registrado a coleta de dados em horas, minutos e segundos. As colunas “Latitude”, “Longitude” e “Altitude” fornecem a localização do caminhão, sendo estas coordenadas dadas no sistema de referência EPSG:31983 (SIRGAS 2000 / UTM zone 23). Na coluna “Ocorrência” tem-se a operação que está sendo realizada. “Odômetro” mostra a distância percorrida acumulada em metros e “Combustível Consumido” é o valor acumulado do consumo do caminhão. A velocidade registra da amostra dada em km/h é apresentada na coluna “Vel”.

5.2 Tratamento dos dados

Após a coleta trabalhou-se da busca pelas informações relativas ao deslocamento do caminhão entre os pontos de amostragem, como por exemplo: distância percorrida, tempo gasto, elevação do deslocamento e consumo de combustível.

Tomando como base a Figura 22 onde temos o deslocamento do caminhão entre dois pontos de amostragem A e B, pode-se calcular as variáveis d_{AB} , t_{AB} , h_{AB} e $cons_{AB}$ a partir da Tabela 3:

$$d_{AB} = Odômetro_B - Odômetro_A \quad (30)$$

$$t_{AB} = Hora_B - Hora_A \quad (31)$$

$$h_{AB} = Altitude_B - Altitude_A \quad (32)$$

$$cons_{AB} = CombustívelConsumido_B - CombustívelConsumido_A \quad (33)$$

Onde:

d_{AB} - Distância percorrida entre os pontos A e B (m);

t_{AB} - Tempo de gasto no deslocamento entre os pontos A e B (s);

h_{AB} - A elevação entre os pontos A e B (m);

$cons_{AB}$ - Consumo de combustível entre A e B (l).

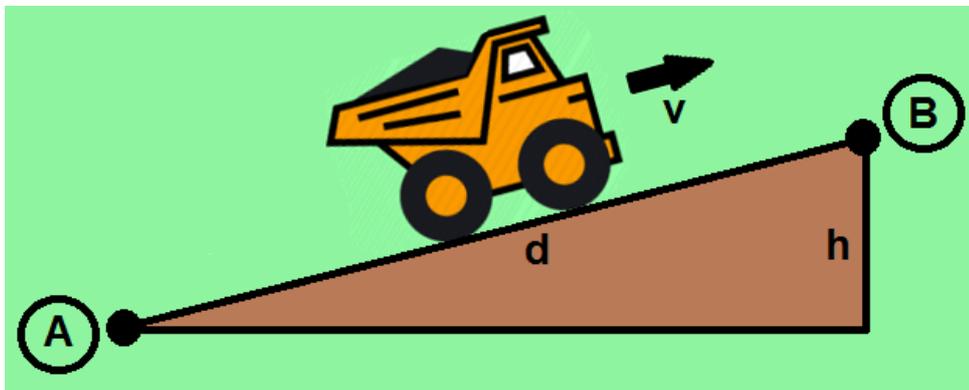


Figura 22: Ilustração do trajeto do caminhão carregado.

Fonte: Elaborado pelo autor

A Tabela 4 apresenta os valores das variáveis do deslocamento do caminhão calculadas com base nos dados coletados presentes na Tabela 3.

Tabela 4: Variáveis relativos ao deslocamento entre os pontos coletados.

Pontos	d(m)	t(s)	h(m)	cons(l)
1-2	10	10	0,1	0,0
2-3	40	10	1,4	0,0
3-4	65	10	0,9	0,5
4-5	95	10	0,4	0,0
5-6	80	10	0,9	0,0
6-7	100	10	0,2	0,0
7-8	105	10	-0,9	0,0
8-9	110	10	1,0	0,0
9-10	115	10	1,4	0,5

Os resultados da Tabela 4 podem nos levar a uma interpretação errônea, pois mostram deslocamentos que supostamente não tiveram consumo de combustível, o que não é real.

Isto ocorre pelo fato da medição do consumo ser acumulativa e ter uma precisão de 0,5 litros, assim quando verificamos uma mudança no valor do consumo, o acréscimo de 0,5 é relativo a todo percurso realizado desde a mudança de valor anterior.

Assim, o consumo apresentado pelo deslocamento entre os Pontos 9-10 na verdade é referente ao deslocamento entre os Pontos 4-10. Portanto para termos uma relação das outras variáveis com o consumo é preciso acumular os valores calculados da distância percorrida, do tempo gasto e da elevação, ponto à ponto, desde o ponto 4 até ao ponto 10, Tabela 5 .

Tabela 5: Variáveis relativos ao deslocamento com valores acumulados.

Pontos	d_acumulado (m)	t_acumulado (s)	h_acumulado (m)	cons_acumulado (l)
4-10	605	60	3,0	0,5

Na coluna Ocorrência da Tabela 3 temos as várias etapas de operação do caminhão (carregamento, operando vazio, báscula, operando cheio e outros), sendo que cada uma possui a taxa de consumo de combustível diferente das outras, ou seja determinadas etapas apresentam uma relação litros/km ou litros/hora maior ou menor que as demais.

Com o objetivo de caracterizar o consumo de combustível de acordo com a etapa de operação, buscou-se relacionar as informações de deslocamento com a informação do coluna Ocorrência. Porém outra interpretação errada pode acontecer, novamente relacionado com fato da medição do consumo ser acumulativa e registrada a cada 0,5 litros. Uma medida do consumo dentro da faixa de precisão pode começar em uma etapa e terminar em outra, isso faz com que o consumo seja associado de forma equivocada a última etapa, mais porém ocorreu em grande parcela noutra.

Assim, para termos uma caracterização mais fiel da taxa de consumo de cada etapa, os consumos que começaram em uma etapa e terminaram em outra, foram descartados, ou seja, consideramos apenas os consumos que começaram e terminaram dentro de uma mesma etapa.

5.3 Análise dos dados

Uma análise preliminar, realizada para mapear as características do consumo de combustível, demonstra o deslocamento do caminhão carregado como o responsável pela maior parcela no consumo de combustível. A Figura 23 apresenta dois gráficos, o primeiro representa a consumo de combustível e o segundo a distância percorrida. A etapa “operando cheio” representa a parte da operação na mina que o caminhão realiza carregado de minério, ou seja, do ponto de carga até o ponto de descarga, a etapa “Operando vazio” representa o trajeto contrário, nele o caminhão está descarregado e se desloca do ponto de descarga até o ponto de carga. “Processo de carga” representa as etapas de carrega-

mento: deslocamento na fila de carregamento e manobras para realização de carga, já “Processo de descarga” representa as etapas do descarregamento: deslocamento na fila de descarga, manobras para a realização de descarga e báscula. E o status “Outros” representa as demais etapas dos deslocamentos, como abastecimentos, trocas de turnos, oficina, etc.

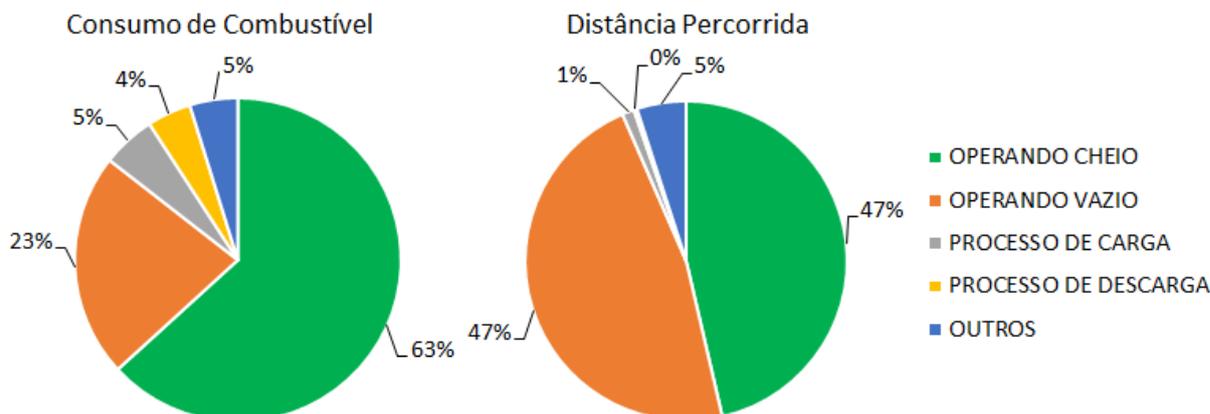


Figura 23: Gráficos do consumo de combustível e distância percorrida por etapa de operação do caminhão.

Fonte: Elaborado pelo autor

Nota-se na Figura 23 que a distância percorrida pelo caminhão na etapa operando cheio é praticamente a mesma do operando vazio, 47% para cada (no gráfico estão com a mesma porcentagem, mas possuem uma pequena diferença) e as demais etapas apresentaram valores igual ou inferiores à 5%. Quanto ao consumo de combustível, a etapa operando cheio representa 63% do total, contra 23% do status operando vazio, sendo os outros status juntos, responsáveis por 14%.

Com isso o trabalho se direcionou ao estudo e obtenção do modelo de consumo de combustível do caminhão carregado, realizando o trajeto do ponto de carga ao ponto de descarga. Como o ponto de descarga é normalmente em um local com a elevação maior que o ponto de carga, esse trajeto é um aclave, como ilustrado na Figura 22, onde o caminhão realiza o deslocamento do ponto A (menor altitude) para o ponto B (maior altitude).

5.4 Formulação do Problema

Algumas combinações de variáveis foram testadas, porém as que se mostraram mais relevantes para a obtenção do modelo do consumo de combustível foram: distância percorrida, elevação e tempo gasto entre os pontos (informações retiradas da Tabela 3), logo estas são nossas variáveis de entrada e o consumo nossa variável de saída. Assim chegou-se

a formulação base para o problema, Equação 34. Onde busca-se o consumo de combustível entre dois pontos A e B da rota do caminhão:

$$cm_{AB} = \alpha * h_{BA} + \beta * d_{BA} + \gamma * t_{BA} \quad (34)$$

Onde:

cm_{AB} - Consumo modelado entre os pontos A e B (l);

h - Elevação entre os pontos de A e B (m);

d - Distância linear entre os pontos de A e B (km);

t - Tempo gasto no percurso entre A e B (min);

α - Coeficiente de relação consumo/elevação (l/m);

β - Coeficiente de relação consumo/distância (l/km);

γ - Coeficiente de relação consumo/tempo (l/min).

Os valores de h , d e t são obtidos dos dados coletados e os valores de α , β e γ são os coeficientes de relação a serem buscados para o modelo do problema.

5.5 Aplicação da Regressão Linear Múltipla

A regressão linear múltipla (RLM) nos possibilita a investigação da relação de diversas variáveis independentes com uma dependente. Aplicando a RLM no problema proposto neste trabalho temos através da Equação 34:

- Variáveis independentes: h , d e t ;
- Variável dependente: cm_{AB} .
- E os valores de α , β e γ são os coeficientes de relação a serem encontrados através pela regressão.

Para realizar a RLM utilizou-se o software Microsoft® Excel, assim foi necessário colocar as informações tratadas a partir dos dados coletados em uma planilha, Figura 24.

A Figura 25 traz a configuração usada na ferramenta de regressão do Excel. O “Intervalo Y de entrada” recebeu os valores da variável dependente cm_{AB} , já o “Intervalo X de entrada” recebeu os valores relativos as variáveis independentes h , d e t . O nível de confiança utilizado foi de 95%. O opção “Rótulos” foi marcada pois a planilha criada possui uma linha de cabeçalho.

A opção “Constante é zero” foi marcada pois a formulação não prevê a presença de uma constante. Uma vez que a estratégia adotada para tratamento dos dados é focada na alteração do valor do consumo e a precisão da medição é igual à 0,5 litros, todos os valores de consumo registrados nesta planilha serão igual ao valor da precisão. Assim a RLM não retornaria coeficientes de relação adequados, todos seriam iguais à zero e a

	A	B	C	D	E	F
1	Ponto Final	h(m)	d(km)	t(min)	cons(l)	Ocorrência
2	7389	10,2	0,265	0,50	0,5	Operando Cheio
3	7441	9,7	0,335	0,67	0,5	Operando Cheio
4	7443	10,1	0,170	0,33	0,5	Operando Cheio
5	7448	2,5	0,435	0,83	0,5	Operando Cheio
6	7502	14,4	0,280	0,50	0,5	Operando Cheio
7	7504	5,8	0,190	0,33	0,5	Operando Cheio
8	7513	3,4	0,715	1,50	0,5	Operando Cheio
9	7559	11,6	0,250	0,50	0,5	Operando Cheio
10	7561	6,2	0,190	0,33	0,5	Operando Cheio
11	7613	8,7	0,260	0,50	0,5	Operando Cheio

Figura 24: Planilha com dados tratados.

Fonte: Elaborado pelo autor

constante seria igual 0,5 (neste cenário mesmo dados relativos ao caminhão parado, como todas variáveis independentes iguais à zero, resultaria em um consumo de 0,5 litros).

Os demais itens da Figura 25 não marcados por não impactarem no modelo de regressão linear proposto para o problema apresentado neste trabalho.

5.6 Algoritmo Genético Implementado

O algoritmo genético foi implementado no software Matlab®, a escolha deste se deu pelos seus recursos matemáticos.

Adotou-se a representação numérica dos cromossomos, com o intuito de se trabalhar com a precisão máxima do software e ter ao mesmo tempo um tamanho reduzido do cromossomo. Cada gene representa então, uma variável de interesse, ou seja, os coeficientes de relação entre o consumo de combustível e as variáveis coletadas, apresentados na Equação 34.

A Tabela 6 apresenta a estrutura cromossômica dos indivíduos candidatos à solução do problema.

Tabela 6: Estrutura dos cromossomos.

	Gene 1	Gene 2	Gene 3
Cromossomo	α	β	γ

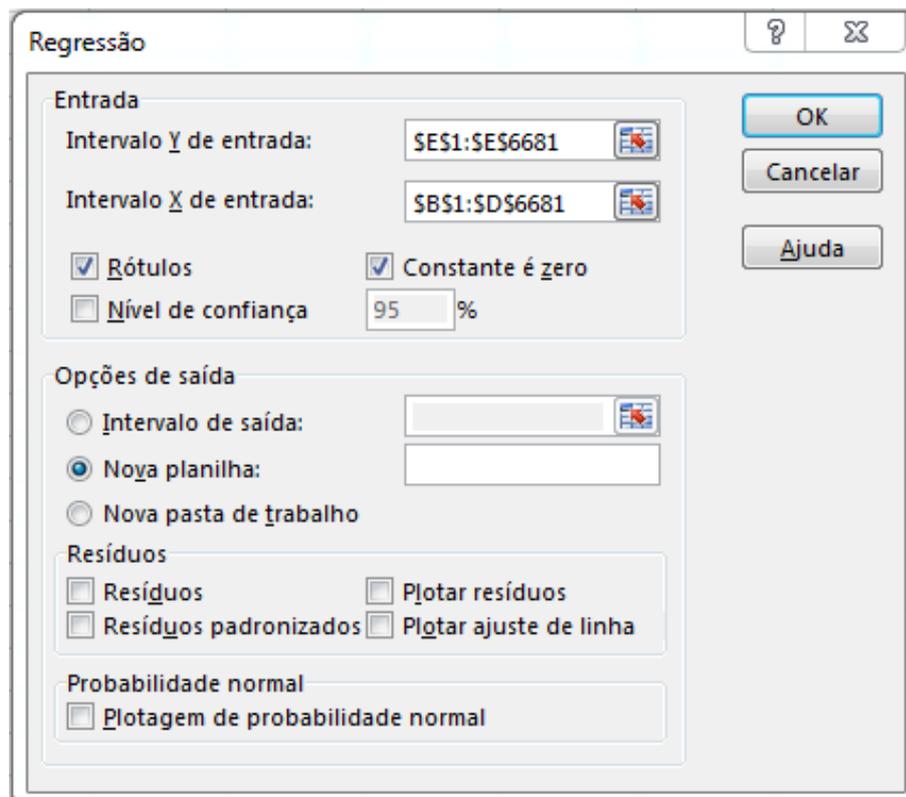


Figura 25: Parametrização do RLM no Excel.

Fonte: Elaborado pelo autor

Na Tabela 7 temos uma população de possíveis soluções, baseadas na estrutura apresentada e gerada pelo AG, o tamanho desta pode ser configurado.

Tabela 7: População de possíveis soluções geradas.

	α	β	γ
Cromossomo 1	0,355730517	0,055375002	0,971185383
Cromossomo 2	0,047402003	0,246893222	0,560516674
Cromossomo 3	0,758701311	0,362389118	0,499499043
Cromossomo 4	0,993416725	0,343927665	0,838651495
Cromossomo 5	0,768445652	0,401063941	0,274384835
Cromossomo 6	0,174117408	0,962800091	0,835857662
Cromossomo 7	0,114799752	0,283432469	0,156051957
Cromossomo 8	0,211004090	0,833436862	0,063520824
Cromossomo 9	0,561940911	0,477252921	0,035371726
Cromossomo 10	0,867818910	0,559844423	0,336919170

5.6.1 Função de Aptidão do AG Implementado

A função de aptidão é responsável por avaliar os indivíduos da população gerada. No AG implementado esta função ocorre em duas etapas, a primeira tem como base a Equação 34 e as variáveis coletadas e nela calcula-se o consumo de todos dados tratados

dos pontos de rotas para cada cromossomo gerado, na segunda os valores calculados são comparados com os consumos reais medidos.

Na primeira etapa cria-se uma matriz de consumo, onde cada linha é relativa à um cromossomo da população gerada e é o vetor dos consumos calculados para cada ponto de dados coletados e tratados, Equação 35.

```

for j=1:m
  for i=1:n
    
$$cm_{ji} = \alpha_j * h_i + \beta_j * d_i + \gamma_j * t_i$$

  end
end

```

Onde:

j - índice dos cromossomos

i - índice dos dados coletados e tratados

m - tamanho da população dos cromossomos

n - número total de pontos dos dados coletados e tratados

α_j , β_j e γ_j - coeficientes do cromossomo de índice j

h_i , d_i e t_i - variáveis coletadas e tratadas do ponto de índice i

Na segunda etapa os valores calculados anteriormente são comparados, através da Equação 36, com os consumos reais registrado (dados coletados).

```

for j=1:m
    
$$EQM_j = \frac{\sum_{i=1}^n (cr_i - cm_{ji})^2}{n}$$

end

```

Onde:

j - índice dos cromossomos

i - índice dos dados coletados e tratados

m - tamanho da população dos cromossomos

n - número total de pontos dos dados coletados e tratados

EQM_j - erro quadrático médio entre consumo modelo e real para o cromossomo j

cr_i - consumo real do ponto i (1);

cm_i - consumo modelado do ponto i (1);

O objetivo do AG é minimizar o valor do erro quadrático médio (EQM) entre o consumo modelado e o real. Quanto menor esse valor, mais o sistema modelado se aproxima do real.

5.6.2 Operações Genéticas

Para realizar as operações genéticas o AG foi desenvolvido com a possibilidade da utilização de dois métodos de cruzamento: Wright e Radcliffe, e três métodos de seleção: Roleta, Torneio e Ranking, além da opção de uso do elitismo. O intuito foi permitir o trabalho com diversas configurações e métodos, possibilitando a avaliação do desempenho de diferentes operadores genéticos para o problema em questão, na busca por modelos que consigam representar satisfatoriamente o consumo de combustível do caminhão.

A Figura 26 mostra seis combinações possíveis dos métodos de seleção e cruzamento, sendo que cada uma ainda pode utilizar ou não o elitismo.

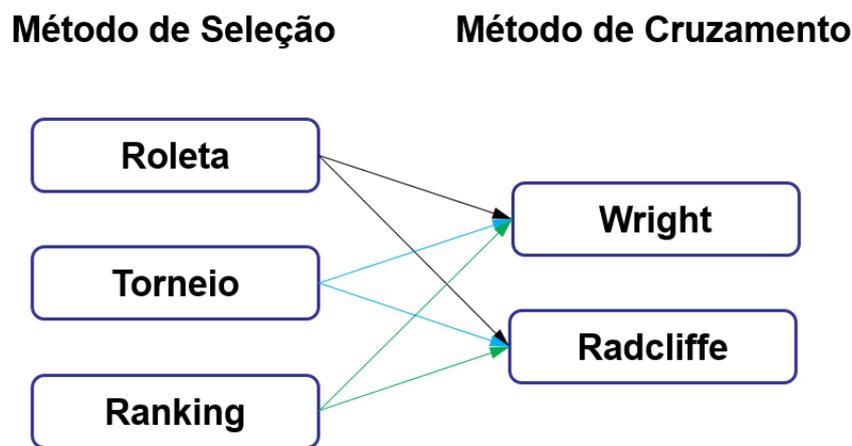


Figura 26: Combinações dos Métodos Implementados.

Fonte: Elaborado pelo autor

Quanto a mutação, o operador utilizado foi o real aleatório. Neste, um gene é escolhido aleatoriamente e seu conteúdo é alterado, também de forma aleatória, sempre dentro do espaço de busca. A mutação ocorre de acordo com uma taxa previamente configurada.

5.7 Sistema Desenvolvido

O sistema desenvolvido no Matlab®, cuja interface é vista na Figura 27, realiza o tratamento, filtro e formatação dos dados reais coletados para serem utilizados na regressão linear múltipla e no algoritmo genético. O sistema permite também a utilização e parametrização do algoritmo genético implementado, validação dos modelos obtidos, previsão e geração do mapa de consumo de combustível em uma rota existente ou a ser criada.

O sistema é composto basicamente de duas ferramentas: Algoritmo Genético e Ferramenta de Análise, sendo cada uma dividida em duas partes:

- A Ferramenta do Algoritmo Genético está localizada na parte esquerda da interface. A parte superior desta é responsável por realizar a importação, tratamento e filtro dos dados coletados, além de gerar as informações para aplicação da Regressão

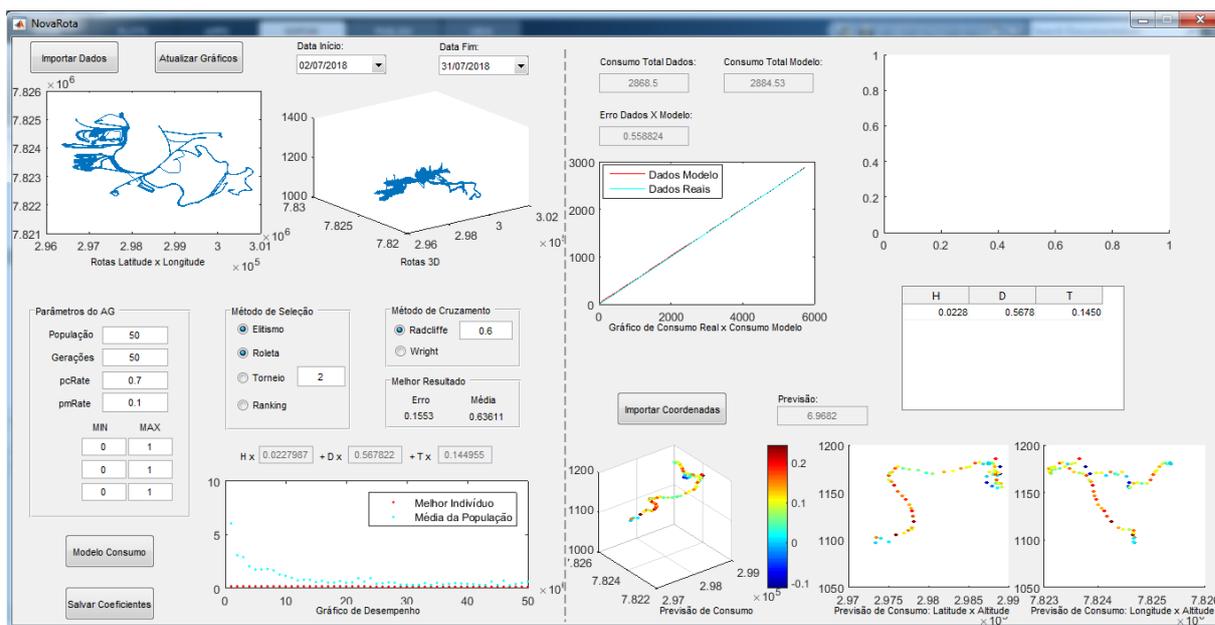


Figura 27: Interface do Sistema Desenvolvido.

Fonte: Elaborado pelo autor

Linear Múltipla. Nela temos dois botões, um para realizar a importação do arquivo com os dados e o outro para atualização dos gráficos, duas caixas de seleção para especificarmos o período a ser avaliado e dois gráficos para representar os percursos realizados pelo caminhão no período avaliado, sendo um gráfico 2D (latitude x longitude) e outro 3D (latitude x longitude x altitude). Na parte inferior encontra-se o AG propriamente dito, onde o usuário pode configurar seus parâmetros: tamanho da população, número de gerações, taxa de cruzamento, taxa de mutação, intervalo de busca dos coeficientes, selecionar o método de seleção, cruzamento e habilitar o elitismo. Para operar o AG tem-se dois botões, um para iniciar a execução e outro salvar os coeficientes encontrados. Encontra-se ainda, nesta parte, informações sobre o desempenho do AG, com caixas de texto com o resultado dos coeficientes, do erro médio quadrático e da média da população. Ainda tem-se um gráfico com a evolução do AG durante as gerações processadas.

- A Ferramenta de Análise, localizada na parte direita da interface, é destinada a validação dos modelos obtidos na ferramenta do AG e geração do mapa de calor de consumo de combustível em uma rota padrão. Na parte superior temos valores comparativos e o erro percentual do consumo modelado em relação ao real registrado nos dados. Ainda encontra-se gráficos comparativos da evolução do consumo por ponto de rota e uma tabela de coeficientes encontrados. Na parte inferior tem-se uma “sub-ferramenta” de previsão de consumo, onde uma rota padrão é testada a partir do conjunto de coordenadas do seu trajeto e a velocidade planejada pela equipe de operação da mina, essas informações são importadas através do botão

“Importar Coordenadas”. Assim é realizada uma previsão do consumo e é gerado o mapa de calor do consumo da rota, que permite ao usuário verificar os pontos que apresentam o maior índice de consumo. Esse mapa é composto de três gráficos, sendo um 3D (latitude x longitude x altitude) e dois 2D (latitude x altitude e longitude x altitude).

Na Figura 28 temos um fluxograma de utilização da Sistema Desenvolvido.

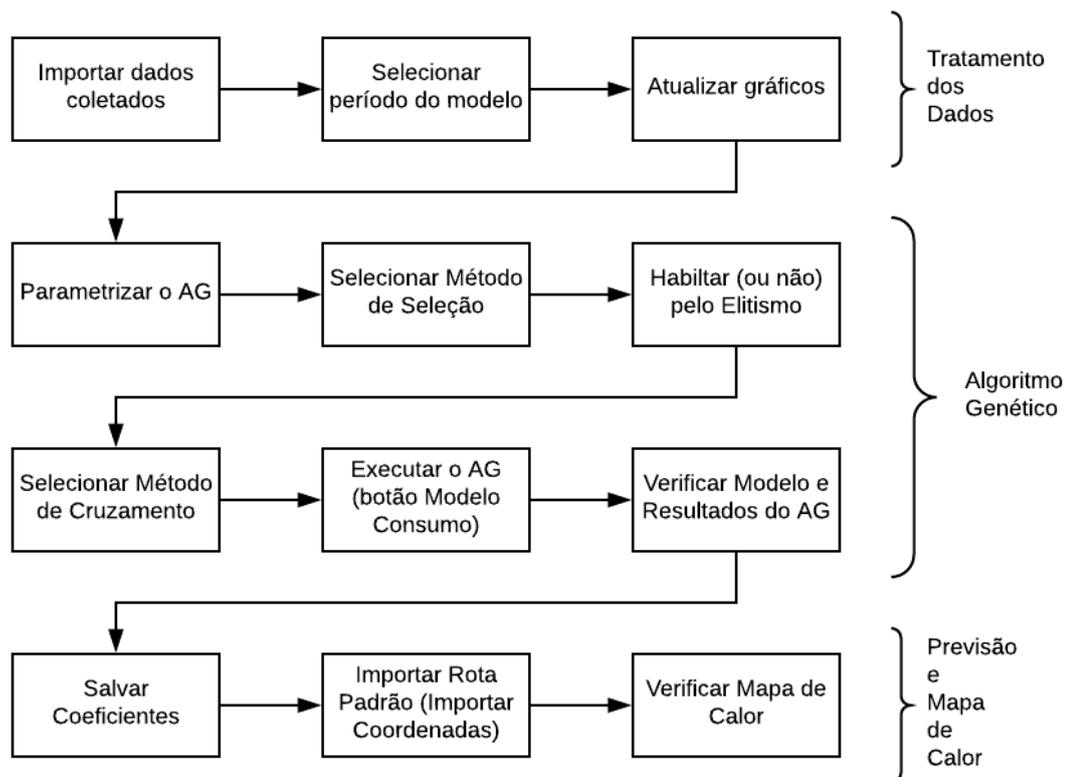


Figura 28: Fluxograma de Utilização do Sistema Desenvolvido.

Fonte: Elaborado pelo autor

Resultados

A massa de dados foi submetida ao sistema desenvolvido, no qual seus dados foram tratados, filtrados e formatados para serem usados na Regressão Linear Múltipla e no Algoritmo Genético. O sistema fornece um arquivo xls para utilização da RLM no Excel® e permite a manipulação do AG para busca de possíveis soluções do problema proposto.

6.1 Obtenção do modelo com RLM

A RLM foi realizada com nível de confiança de 95% e com a constante de intercepto do eixo Y igual a zero. Esse processo originou o relatório da Figura 29.

O coeficiente de explicação ou determinação (R-Quadrado) fornecido pelo Excel® para o modelo é igual a 0,9211017, o que representa um bom poder de explicação, ou seja, as variáveis elevação (h), distância (d) e tempo (t), em conjunto, por intermédio do modelo explicam 92,11% das alterações ou modificações ocorridas no consumo de combustível. O restante 7,89% representa as variáveis aleatórias (erro) não consideradas na formulação do modelo.

Verifica-se na Tabela ANOVA que o F de significação é igual a 0, sendo inferior ao nível de significância de 5% (0,05), podendo então admitir a existência da regressão linear do consumo de combustível sobre pelo menos uma das variáveis elevação, distância e tempo.

Com relação aos coeficientes obtidos tem-se:

- Para a variável elevação (h): o coeficiente encontrado foi $\alpha = 0,0238069$ com intervalo de confiança $[0,023477104 : 0,02413669]$. Como o valor-P é igual a zero e esse valor é menor do que 0,025 (0,05/2), admite-se que existe regressão do consumo de combustível sobre a variável elevação.
- Para a variável distância (d): o coeficiente encontrado foi $\beta = 0,6324224$ com intervalo de confiança $[0,612409447 : 0,65243533]$. Como o valor-P é igual a zero e esse valor é menor do que 0,025 (0,05/2), admite-se que existe regressão do consumo de combustível sobre a variável distância.

	A	B	C	D	E	F	G
1	RESUMO DOS RESULTADOS						
2							
3	<i>Estadística de regressão</i>						
4	R múltiplo	0,9597404					
5	R-Quadrado	0,9211017					
6	R-quadrado ajustado	0,9208998					
7	Erro padrão	0,1404809					
8	Observações	5737					
9							
10	ANOVA						
11		<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>	
12	Regressão	3	1321,09011	440,363	22313,95	0	
13	Resíduo	5734	113,159887	0,01973			
14	Total	5737	1434,25				
15							
16		<i>Coefficient</i>	<i>Erro</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>
17	Interseção	0	#N/D	#N/D	#N/D	#N/D	#N/D
18	H	0,0238069	0,00016823	141,514	0	0,023477104	0,02413669
19	D	0,6324224	0,01020872	61,9493	0	0,612409447	0,65243533
20	T	0,0591108	0,00488913	12,0903	3,01E-33	0,049526259	0,06869533

Figura 29: Relatório da Regressão Linear Múltipla no Excel®.

Fonte: Elaborado pelo autor

- Para a variável tempo (t): o coeficiente encontrado foi $\gamma = 0,0591108$ com intervalo de confiança $[0,049526259 : 0,06869533]$. Como o valor-P é praticamente zero ($3,01 * 10^{-33}$) e esse valor é menor do que 0,025 ($0,05/2$), admite-se que existe regressão do consumo de combustível sobre a variável tempo.

A Tabela 8 apresenta os coeficientes do modelo obtido com a RLM.

Tabela 8: Coeficientes do Modelo Obtido com RLM.

	α	β	γ
Modelo RLM	0,0238069	0,6324224	0,0591108

6.2 Obtenção dos modelos com AG

O AG implementado no sistema foi então executado diversas vezes. Como o mesmo permite um grande variedade de configurações, várias foram testadas na busca por modelos que conseguissem representar uma solução satisfatória para o problema do trabalho.

A Tabela 9, traz duas das configurações testadas e que obtiveram modelos com resultados próximos aos valores de referências.

Tabela 9: Configurações do AG testadas.

Parâmetros	Configuração	
	1	2
Tamanho da População:	50	50
Número de Gerações:	200	200
Método de Seleção:	Torneio	Roleta
Método de Cruzamento:	Wright	Radcliffe
Parâmetro Radcliffe:	-	0.6
Elitismo:	Sim	Sim
Taxa de Cruzamento:	0.7	0.7
Taxa de Mutação:	0.1	0.1

A Tabela 10 apresenta dez modelos obtidos, sendo cinco para cada uma das configurações apresentadas na Tabela 9.

Tabela 10: Coeficientes dos Modelos Obtidos com AG.

Modelo	Configuração	α	β	γ
1	1	0,023807046	0,632416432	0,059102790
2	1	0,023810420	0,633185986	0,058598121
3	1	0,023806898	0,632421045	0,059111063
4	1	0,023806896	0,632422375	0,059110799
5	1	0,023806972	0,632390818	0,059134951
6	2	0,028183206	0,465270125	0,151779231
7	2	0,026848907	0,411392251	0,166226657
8	2	0,025384641	0,468256865	0,158680035
9	2	0,030793312	0,427870313	0,155105522
10	2	0,027995145	0,442657181	0,160054486

Os modelos testados e comparados com os valores de referência pela Ferramenta de Análise do Sistema Desenvolvido apresentaram os resultados da Tabela 11, que mostra o valor do consumo total modelado e o erro em relação ao consumo total real da massa de dados (que é de 2868,5l), assim como o desvio padrão dos pontos de rotas.

6.3 Comparação dos Modelos Obtidos

A Tabela 12 apresenta o comparativo entre o modelo da RLM com os obtidos pelo AG configuração 1 (modelo 3 da Tabela 11) e o AG configuração 2 (modelo 8 da Tabela 11). As colunas mostram os desempenhos gerados pela Ferramenta de Análise, onde temos o valor do consumo total modelado, o erro em relação ao consumo total da massa de dados (2868,5l) e o desvio padrão para os pontos de rotas. Os resultados da RLM são próximos

Tabela 11: Resultado dos Modelos Obtidos com AG.

Modelo	Consumo(l)	Erro(%)	Desvio Padrão
1	2642,13	7,89	0,1347991
2	2641,60	7,91	0,1347721
3	2642,18	7,89	0,1348016
4	2642,18	7,89	0,1348017
5	2642,23	7,89	0,1348041
6	2836,07	1,13	0,1534132
7	2895,83	0,95	0,1523799
8	2788,76	2,78	0,1470858
9	2843,39	0,88	0,1657318
10	2815,63	1,84	0,1532699

ao do AG configuração 1, já o AG configuração 2 demonstra um erro total menor, com um desvio padrão ligeiramente maior que os outros.

Tabela 12: Comparativo entre Modelos Obtidos.

Método	α	β	γ	Consumo (L)	Erro (%)	Desvio Padrão
RLM	0,0238069	0,6324223	0,0591108	2642,18	7,89	0,1348017
AG Configuração 1	0,0238069	0,6324210	0,0591110	2642,18	7,89	0,1348016
AG Configuração 2	0,0253846	0,4682568	0,1586800	2788,76	2,78	0,1470858

As Figuras 30, 31 e 32 mostram gráficos comparativos da evolução do consumo real e dos consumos modelados, o eixo das ordenadas (Y) representa o consumo de combustível acumulado em litros e no eixo das abscissas (X) temos os pontos de coleta de dados (número da amostra) das rotas percorridas pelo caminhão durante o período avaliado. Os dois primeiros gráficos, Figuras 30 e 31, demonstram respectivamente os desempenhos dos modelos da RLM e do AG configuração 1, nota-se que apresentam uma evolução similar. Já o gráfico da Figura 32, que representa o AG configuração 2, diferiu-se dos demais, pois sua resposta aproximou-se mais dos valores reais, logo a diferença entre o real e modelado é menor que os anteriores.

6.4 Aplicação dos Modelos em uma Rota Padrão

Estes modelos quando testados em uma rota real padrão, mostrada na Figura 33 através de três gráficos, onde tem-se a representação dos seus pontos em 3D, da Latitude por Altitude e a Longitude por Altitude. Para esta rota tem-se o histórico de operação e média de consumo conhecidos, usada aqui para validação e comparação, apresentaram os resultados da Tabela 13. O erro foi calculado com base no consumo médio registrado nesta rota (6,5l). Novamente o desempenho do modelo da RLM e do AG configuração 1

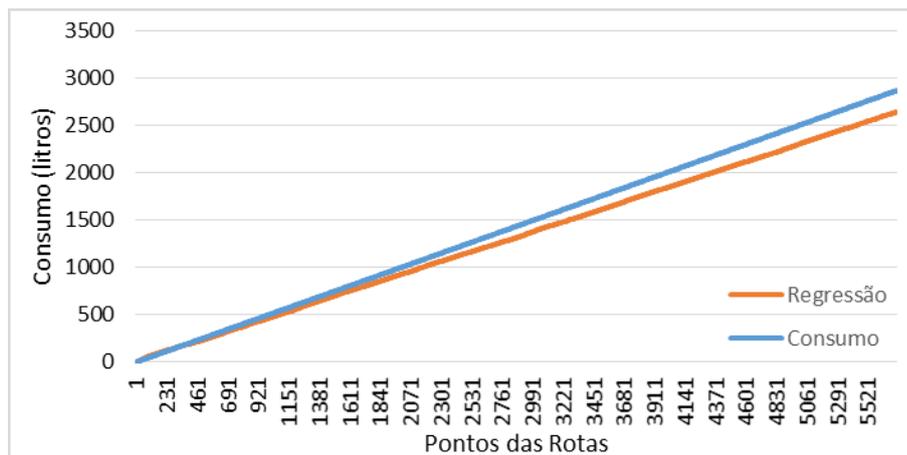


Figura 30: Gráficos da evolução de consumo real versus RLM.

Fonte: Elaborado pelo autor

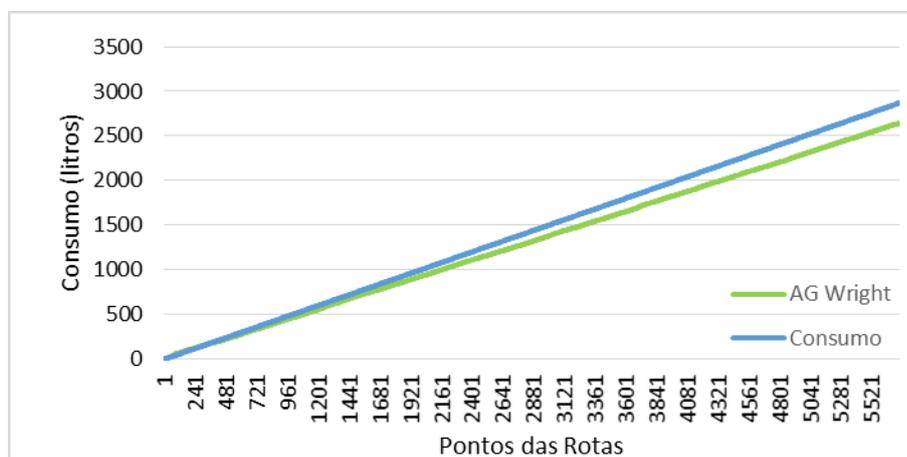


Figura 31: Gráfico da evolução de consumo real versus AG configuração 1.

Fonte: Elaborado pelo autor

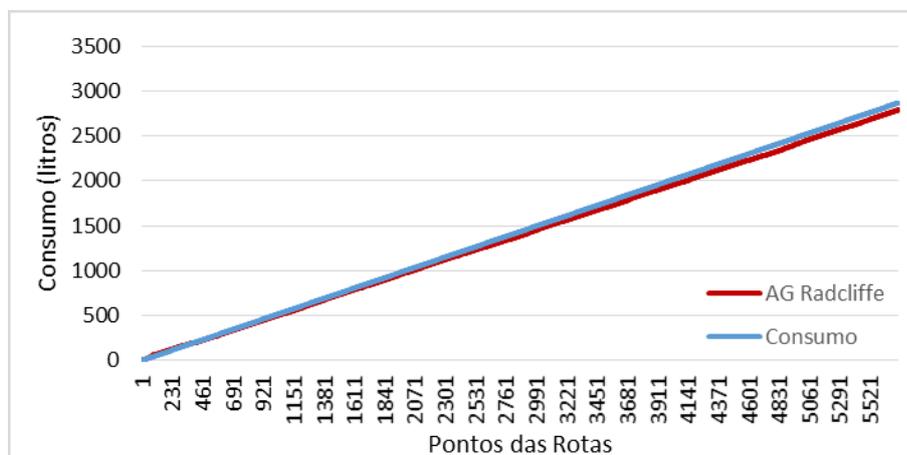


Figura 32: Gráfico da evolução de consumo real versus AG configuração 2.

Fonte: Elaborado pelo autor

foram similares e desta vez se aproximaram mais do valor real do que o AG configuração 2.

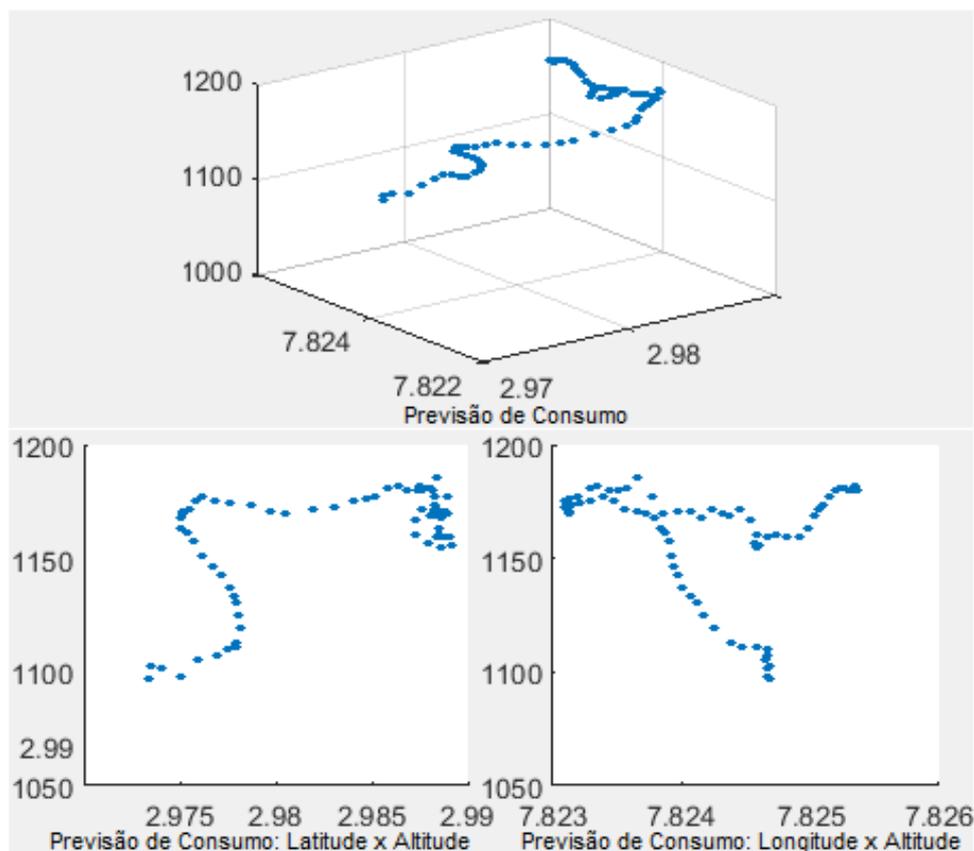


Figura 33: Rota de Padrão para teste da previsão de consumo.

Fonte: Elaborado pelo autor

Tabela 13: Comparativo das Previsões de Consumo para Rota Padrão

Método	Consumo Médio (L)	Previsão (L)	Erro (%)
RLM	6,5	6,29	3,23
AG Configuração 1	6,5	6,29	3,23
AG Configuração 2	6,5	6,81	4,77

O sistema desenvolvido forneceu também um mapa calor do consumo de combustível para os modelos obtidos, com a indicação do nível de consumo por pontos da rota através de cores. Como os modelos obtidos pelos métodos apresentados anteriormente apresentaram resultados próximos, os mapas de calor se assemelham-se muito. A Figura 34 traz a rota padrão submetida à um modelo obtido com o AG configuração 1, onde verifica-se a vista 3D, a vista da Latitude pela Altitude e a vista da Longitude pela Altitude. Permitindo assim a indicação dos trechos com maior índice de consumo de combustível, através das cores quentes atribuídas a estes pontos.

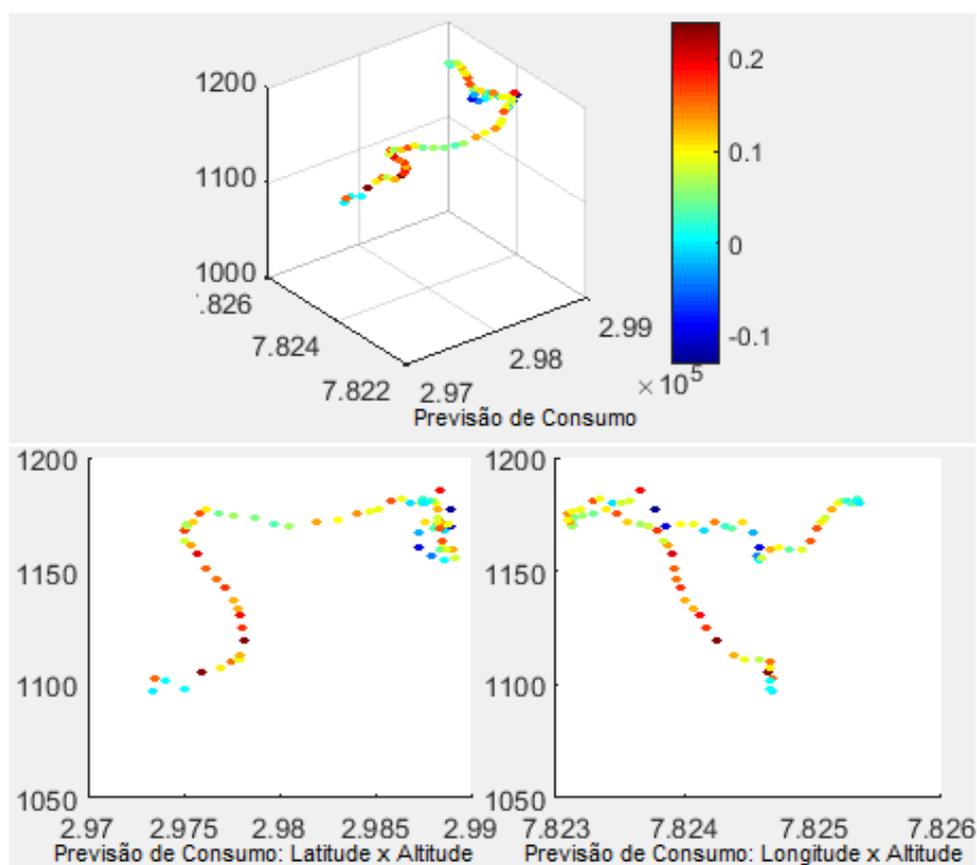


Figura 34: Mapa de calor do consumo da Rota de Padrão para o modelo do AG configuração 1.

Fonte: Elaborado pelo autor

Conclusões

O presente trabalho apresentou o desenvolvimento de um sistema e o uso de ferramentas como Algoritmo Genético e Regressão Linear Múltipla para obtenção do modelo matemático e o mapa de calor do consumo de combustível de um caminhão na mineração a partir de dados reais coletados do sistema de telemetria em uma mina a céu aberto.

A aplicação das ferramentas AG e RLM se mostraram eficiente para o problema proposto, os modelos obtidos se aproximaram aos valores de referência. O AG com método de cruzamento Wright teve praticamente o mesmo desempenho da RLM, já o AG com cruzamento Radcliffe conseguiu obter um modelo que aproximou-se mais dos valores de referência na evolução do consumo de combustível.

Alguns fatores como: o peso da carga, que normalmente fica em torno 35 toneladas, mas dependendo do operador do equipamento de carga pode variar para mais ou menos; diferentes motoristas, cada um com sua forma conduzir o veículo; e até condições climas como chuva, vento e resistência do ar podem ser responsáveis pela pequena diferença entre o valor real e o modelado.

O sistema desenvolvido facilitou o uso das ferramentas propostas, possibilitou a validação dos modelos obtidos e geração do mapa de calor.

O mapa de calor gerado possibilitou a visualização de forma clara e fácil dos pontos e trechos da rota com maior índice de consumo de combustível, o que pode criar a possibilidade da realização de mudanças afim redução do consumo.

Através do modelo matemático e do mapa de calor é possível criar rotinas e fornecer informações ao sistema de despacho na busca pela redução dos custos de operação da mina. O sistema desenvolvido se apresenta como uma importante ferramenta na definição e criação de novas rotas, assim como na alteração de rotas existentes.

O uso do presente trabalho na otimização da definição, criação de novas rota e alterações na rotas existentes pode criar uma redução no consumo de combustível não apenas do caminhão estudado, mas de toda a frota em operação. No Brasil existem minas à céu aberto que operam com mais de cem caminhões similares ao estudado, rodando 24 horas por dia simultaneamente. Assim uma pequena melhoria alcançada pode ser replicada e

bons resultados obtidos.

Como possibilidade de trabalhos posteriores temos o estudo de um número maior de caminhões, com o objetivo de caracterizar a frota em operação, o que poderia potencializar a aplicação da sistema proposto e maximizar os resultados com sua utilização. Além disso pode-se considerar a aplicação de outras ferramentas e algoritmos para obtenção dos modelos e caracterização do consumo, como por exemplo as Redes Neurais Artificiais e os Algoritmos Diferenciais Evolutivos.

Estudos que visam a redução do consumo de combustível se mostram importantes para empresas que realizam transporte por caminhões, pois este representa grande parcela dos custos de operação. Vale ressaltar também o benefício da redução de emissão de gases do efeito estufas, o que é de interesse de toda a sociedade.

Trabalhos em Eventos

Os resultados alcançados foram expostos nos seguintes trabalhos:

- ❑ Trabalho intitulado “Ferramenta que Fornece o Modelo Matemático de um Caminhão de Mineração Utilizando Algoritmo Genético” apresentado em comunicação oral na XVI Conferência de Estudos em Engenharia Elétrica - XVI CEEL - realizada no período de 05 a 09 de novembro de 2018 no campus Santa Mônica da Universidade Federal de Uberlândia (UFU).
- ❑ Trabalho intitulado “Obtenção do Modelo Matemático e Mapa de Calor do Consumo de Combustível de um Caminhão em Mineração à céu aberto, através de Regressão Linear Múltipla e Algoritmo Genético” aprovado para ser apresentado em comunicação oral no Congresso Brasileiro de Instrumentação, Sistemas e Automação - COBISA - período de 14 a 16 de maio de 2019 no Centro de Convenções da Universidade Estadual de Campinas (UNICAMP).

Referências

ALARIE, S.; GAMACHE, M. Overview of solution strategies used in truck dispatching systems for open pit mines. **International Journal of Surface Mining, Reclamation and Environment**, Taylor & Francis, v. 16, n. 1, p. 59–76, 2002. Disponível em: <<https://doi.org/10.1076/ijsm.16.1.59.3408>>.

ALEXANDRE, R. F. **Modelagem, simulação da operação e otimização multiobjetivo aplicada ao problema de despacho de veículos em minas a céu aberto**. Tese (Thesis), 2010.

ANP. Web Page, **Levantamento de Preços e de Margens de Comercialização de Combustíveis**. 2019. Disponível em: <<http://www.anp.gov.br/wwwanp/precos-e-defesa-da-concorrenca/precos/levantamento-de-precos>>.

BARROSO, L. C. et al. **Cálculo Numérico (com aplicações)**. [S.l.]: São Paulo: Harbra, 1987.

BEASLEY, D.; BULL, D. R.; MARTIN, R. R. An overview of genetic algorithms: Part 1, fundamentals. **University computing**, v. 15, n. 2, p. 56–69, 1993.

CETIN, N. Open-pit truck/shovel haulage system simulation. **A thesis of the graduate School Of Natural And Applied Sciences Of Middle East Technical Universtity. Turkey**, 2004.

DANTZIG, G. B.; RAMSER, J. H. The truck dispatching problem. **Management science**, Informs, v. 6, n. 1, p. 80–91, 1959. Disponível em: <<https://doi.org/10.1287/mnsc.6.1.80>>.

DARWIN, C. **On the Origin of Species by Means of Natural Selection**. London, 1859. Disponível em: <<https://doi.org/10.5962/bhl.title.68064>>.

DEVORE, J. L. **Probabilidade e estatística: para engenharia e ciências**. [S.l.]: Pioneira Thomson Learning, 2006.

ERCELEBI, S.; BASCETIN, A. Optimization of shovel-truck system for surface mining. **Journal of the Southern African Institute of Mining and Metallurgy**, v. 109, n. 7, p. 433–439, 2009. ISSN 2225-6253.

FERNANDES, J. N. O. A real-time embedded system for monitoring of cargo vehicles, using controller area network (can). **IEEE Latin America Transactions**, v. 14, n. 3,

p. 1086–1092, 2016. ISSN 1548-0992. Disponível em: <<https://doi.org/10.1109/TLA-2016.7459583>>.

FERREIRA, N. R. A. Web Page, **O que são sistemas de referência e de coordenadas?** 2017. Disponível em: <<http://blog.droneng.com.br/geodesia-sistemas-de-referencia-e-de-coordenadas/>>.

GAUR, D. R.; MUDGAL, A.; SINGH, R. R. Routing vehicles to minimize fuel consumption. **Operations Research Letters**, v. 41, n. 6, p. 576–580, 2013. ISSN 0167-6377. Disponível em: <<https://doi.org/10.1016/j.orl.2013.07.007>>.

GERSHON, M. A linear programming approach to mine scheduling optimization. In: **17th APCOM**. [S.l.: s.n.], 1982. p. 483–493.

HARTMAN, H. L.; MUTMANSKY, J. M. **Introductory mining engineering**. [S.l.]: John Wiley & Sons, 2002.

HJORRING, C.; HOLT, J. New optimality cuts for a single-vehicle stochastic routing problem. **Annals of Operations Research**, Springer, v. 86, p. 569–584, 1999. Disponível em: <<https://doi.org/10.1023/A:1018995927636>>.

HOLLAND, J. H. et al. **Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence**. MIT press, 1992. Disponível em: <<https://doi.org/10.7551/mitpress/1090.001.0001>>.

JANIKOW, C. Z.; MICHALEWICZ, Z. An experimental comparison of binary and floating point representations in genetic algorithms. In: **ICGA**. [S.l.: s.n.], 1991. p. 31–36.

JONG, K. A. D. Analysis of the behavior of a class of genetic adaptive systems. 1975.

JÚNIOR, C. R. C. et al. Desenvolvimento de uma metodologia para redução do consumo específico de diesel em caminhões fora de estrada em uma empresa do setor de mineração. **UFOP**, 2013.

KOLONJA, B.; KALASKY, D. R.; MUTMANSKY, J. M. Optimization of dispatching criteria for open-pit truck haulage system design using multiple comparisons with the best and common random numbers. In: IEEE. **Proceedings of 1993 Winter Simulation Conference-(WSC'93)**. 1993. p. 393–401. Disponível em: <<https://doi.org/10.1145/256563.256671>>.

LINDEN, R. **Algoritmos Genéticos (3a edição)**. [S.l.]: Ciência Moderna Ltda, 2012. ISBN 9788539901951.

LOPES, J. R. Viabilização técnica e econômica da lavra contínua de minério de ferro com o uso de sistema de britagem móvel “in pit” auto propelido. Programa de Pós-Graduação em Engenharia Mineral. Departamento de Engenharia ?, 2010.

MARTINS, G. d. A.; DOMINGUES, O. **Estatística geral e aplicada**. [S.l.]: Altas, 2017.

- MENDES, J. B. et al. A hybrid multiobjective evolutionary algorithm for truck dispatching in open-pit-mining. **IEEE Latin America Transactions**, v. 14, n. 3, p. 1329–1334, 2016. ISSN 1548-0992. Disponível em: <<https://doi.org/10.1109/TLA.2016.7459617>>.
- MERCEDES-BENZ. Web Page, **Actros 4844 8x4 Basculante Fora de Estrada**. 2019. Disponível em: <<https://www.mercedes-benz.com.br/caminhoes/actros/actros-4844-8x4-basculante-fora-de-estrada>>.
- MIT. Web Page, **Chapter 3 - Linear Regression**. 2019. Disponível em: <<http://www.mit.edu/~6.s085/notes/lecture3.pdf>>.
- MITCHELL, M. **An introduction to genetic algorithms**. [S.l.]: MIT press, 1998.
- MUNDOGEO. Web Page, **Coordenadas topográficas X Coordenadas UTM**. 2019. Disponível em: <<https://mundogeo.com/blog/2013/06/05/coordenadas-topograficas-x-coordenadas-utm/>>.
- NEO. Web Page, **Vehicle Routing Problem**. 2019. Disponível em: <<http://neo.lcc-uma.es/vrp/vehicle-routing-problem/>>.
- NEUFELD, J. L.; CELESTE, J. L. **Estatística aplicada à administração usando Excel**. [S.l.]: Pearson Prentice Hall, 2003.
- PORTO, F. S. d. A. et al. Conference Paper, **Telemetria Aplicada a Manutenção de Caminhões: Um Estudo Sobre seus Efeitos no Setor de Manutenção de uma Empresa Mineradora**. 2015. 1-15 p.
- QUEVEDO, J.; DIALLO, M.; LUSTOSA, L. **Modelo de simulação para o sistema de carregamento e transporte em mina a céu aberto**. Pontifícia Universidade Católica do Rio de Janeiro. Tese (Thesis), 2009.
- RADCLIFFE, N. J.; SURRY, P. D. Formal memetic algorithms. In: SPRINGER. **AISB Workshop on Evolutionary Computing**. 1994. p. 1–16. Disponível em: <https://doi.org/10.1007/3-540-58483-8_1>.
- RODRIGUES, L. F. **Análise comparativa de metodologias utilizadas no despacho de caminhões em minas a céu aberto**. Dissertação (Mestrado) — UFMG - Universidade Federal de Minas Gerais, 2016.
- SAHIN, B. et al. An approach for analysing transportation costs and a case study. **European Journal of Operational Research**, v. 193, n. 1, p. 1–11, 2009. ISSN 0377-2217. Disponível em: <<https://doi.org/10.1016/j.ejor.2007.10.030>>.
- SIVANANDAM, S.; DEEPA, S. **Introduction to genetic algorithms**. [S.l.]: Springer Science & Business Media, 2007.
- SOARES, G. L. Algoritmos genéticos: estudo, novas técnicas e aplicações. **Belo Horizonte: Universidade Federal de Minas Gerais**, p. 40, 1997.
- SOUZA, A. G. de; CAMPOS, G. L. Rede can veicular: levantamento bibliográfico e apresentação de conceitos iniciais. **ForScience**, v. 5, n. 1, 2017. Disponível em: <<https://doi.org/10.29069/forscience.2017v5n1.e234>>.

- TORRES, M. F. D. **Rede de Sistemas Embarcados que utilizam Algoritmos Genéticos aplicado na otimização de Despacho em Mineração com Múltiplas Rotas**. Dissertação (Mestrado) — UFU - Universidade Federal de Uberlândia, 2017.
- VALE. Web Page, **Qual a importância da mineração para a economia do país?** 2019. Disponível em: <<http://http://www.vale.com/brasil/pt/aboutvale/news/paginas/qual-a-importancia-da-mineracao-para-a-economia-do-pais.aspx>>.
- WARDLAW, R.; BHAKTIKUL, K. Comparison of genetic algorithm and linear programming approaches for lateral canal scheduling. **Journal of irrigation and drainage engineering**, American Society of Civil Engineers, v. 130, n. 4, p. 311–317, 2004. Disponível em: <[https://doi.org/10.1061/\(ASCE\)0733-9437\(2004\)130:4\(311\)](https://doi.org/10.1061/(ASCE)0733-9437(2004)130:4(311))>.
- WRIGHT, A. H. Genetic algorithms for real parameter optimization. In: **Foundations of genetic algorithms**. Elsevier, 1991. v. 1, p. 205–218. Disponível em: <<https://doi.org/10.1016/B978-0-08-050684-5.50016-1>>.
- XAVIER, R. S. et al. Heurística para modelagem e minimização do consumo de combustível para rotas de coleta de lixo. **Bento Gonçalves**, p. 12, 2010.
- XIAO, Y. et al. Development of a fuel consumption optimization model for the capacitated vehicle routing problem. **Computers & Operations Research**, Elsevier, v. 39, n. 7, p. 1419–1431, 2012. Disponível em: <<https://doi.org/10.1016/j.cor.2011.08.013>>.