

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Programa de Pós-Graduação em Estatística

FABRICIA VIEIRA DA FONSECA

**ANÁLISE DE REGRESSÃO LINEAR MÚLTIPLA PARA
DIMENSIONAMENTO DE CAMINHÕES RODOVIÁRIOS NA
MINERAÇÃO**

Belo Horizonte
2018

FABRICIA VIEIRA DA FONSECA

**ANÁLISE DE REGRESSÃO LINEAR MÚLTIPLA PARA
DIMENSIONAMENTO DE CAMINHÕES RODOVIÁRIOS NA
MINERAÇÃO**

Monografia para Especialização apresentada ao Programa de Pós-Graduação em Estatística com ênfase em Indústria e Mercado do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título do Grau de Especialista em Estatística.

Orientador: Prof. Roberto Quinino da Costa.

**Belo Horizonte
2018**

FABRICIA VIEIRA DA FONSECA

Análise de regressão linear múltipla para dimensionamento de caminhões rodoviários na mineração.

Monografia para Especialização apresentada ao Programa de Pós-Graduação em Estatística com ênfase em Indústria e Mercado do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título do Grau de Especialista em Estatística.

Aprovado em ____/____/____

BANCA EXAMINADORA

Prof. Roberto Quinino da Costa.– Orientador – UFMG

Prof. Lupércio França Bessegato – Avaliador – UFJF

Juliano Ribeiro Moreira – Avaliador – Instituto de Pesquisas Expertise

**Belo Horizonte
2018**

Dedico a minha família, em especial aos meus pais.

AGRADECIMENTOS

Agradeço a Deus por ter concedido a oportunidade de obter este aprendizado para minha formação profissional.

Aos meus pais, irmãos, sobrinhos que apoiam os meus estudos.

Ao Alberto pelo apoio e pela compreensão de minha ausência nesse período.

Ao orientador, professor Roberto Quinino da Costa.

Aos amigos pelo apoio e incentivo.

Agradeço a todos que de alguma forma incentivou e colaborou com essa pesquisa.

"Quando nada parece dar certo, vou ver o cortador de pedras
martelando sua rocha talvez 100 vezes, sem que uma única
rachadura apareça. Mas na centésima primeira martelada a rocha
se abre em duas, e eu sei que não foi aquela que conseguiu isso,
mas todas as que vieram antes."

Jacob Riis

RESUMO

O trabalho desenvolveu na necessidade de estimar a quantidade de minério a transportar através de caminhões rodoviários em uma empresa de minério de ferro utilizando-se regressão linear múltipla. A pesquisa foi baseada em dados históricos dos indicadores operacionais realizados do período de 2015 a 2017 podendo ser caracterizados a partir da análise exploratória com as principais estatísticas dos dados. O resultado final obteve um modelo estatístico útil o qual estima a quantidade de minério a transportar por esses caminhões de acordo com os indicadores operacionais programados. O trabalho possibilitou também auxiliar no planejamento estratégico de aumento de produtividade, qualidade e minimização de custos.

Palavras-chave: Capacidade de transporte. Indicadores Operacionais. Minério de Ferro. Regressão Linear Múltipla.

ABSTRACT

The work developed in the need to estimate the amount of ore to be transported by road trucks in an iron ore company using multiple linear regression. The research was based on historical data of the operational indicators performed from the period from 2015 to 2017 and can be characterized from the exploratory analysis with the main data statistics. The end result obtained a useful statistical model which estimates the amount of ore to be transported by these trucks according to the programmed operational indicators. The work also helped in the strategic planning of increase of productivity, quality and minimization of costs.

Keywords: Transport capacity. Operational Indicators. Iron ore. Multiple Linear Regression.

LISTA DE ILUSTRAÇÕES

Lista de Figuras

Figura 1 – Evolução da Produção Mineral Brasileira (bilhões de toneladas)	15
Figura 2 – Balança Mineral Brasileira	16
Figura 3 – Fluxograma simplificado do processo de mineração.	17
Figura 4 – Estratégica para elaboração do dimensionamento de frota de caminhões	21
Figura 5 – Desenho esquemático transporte de minério	29
Figura 6 – Estatísticas: disponibilidade	30
Figura 7 – Estatísticas: distância média percorrida	30
Figura 8 – Estatísticas: massa	31
Figura 9 – Estatísticas: produtividade efetiva	31
Figura 10 – Estatísticas: rendimento operacional	32
Figura 11 – Estatísticas: utilização efetiva	32
Figura 12 – Modelo Completo.....	34
Figura 13 – Análise de Resíduos	35
Figura 14 – Teste de normalidade para os resíduos	35
Figura 15 – Modelo Completo com transformação de Box-Cox	36
Figura 16 – Teste de normalidade para os resíduos após transformação	37
Figura 17 – Histograma Produtividade estimado	38
Figura 18 – Histograma Disponibilidade estimado	38
Figura 19 – Histograma Utilização estimado	39
Figura 20 – Histograma RO estimado.....	39
Figura 21 – Histograma DMT estimado	40
Figura 22 – Modelo sem a variável Disponibilidade	41
Figura 23 – Teste de normalidade para os resíduos do modelo sem variável disponibilidade	42
Figura 24 – Exemplos de valores estimados versus necessidade de transporte	44

Lista de Quadros

Quadro 1 – Análise de variância e estatística F	25
Quadro 2 – Descrição das variáveis.....	33

Lista de Tabelas

Tabela 1 – Correlação das variáveis	33
Tabela 2 – Limite Superior e Inferior intervalo de confiança <i>Bootstrap</i>	37
Tabela 3 – Limite Superior e Inferior intervalo de confiança <i>Bootstrap</i> (Modelo 02).....	42
Tabela 4 – Capacidade de transporte estimada.....	43

LISTA DE ABREVIATURAS E SIGLAS

DF	Disponibilidade física
DMT	Distância média de transporte
IBRAM	Instituto Brasileiro de Mineração
HM	Horas de manutenção
HP	Horas totais por período
HT	Horas totais
HTP	Horas totais programadas
MC	Manutenção corretiva
MP	Manutenção preventiva
Nequip	Número de equipamentos
PR	Produtividade
TP	Tempo perdido
RO	Rendimento Operacional
UFMG	Universidade Federal de Minas Gerais
UFJF	Universidade Federal de Juiz de Fora
UT	Utilização

SUMÁRIO

1 INTRODUÇÃO.....	12
1.1 Objetivos.....	13
1.1.1 Objetivo Geral.....	13
1.1.2 Objetivos Específicos	13
1.2 Justificativa	13
2 REVISÃO DA LITERATURA.....	15
2.1 Produção do minério de Ferro.....	15
2.2 O transporte de minério por caminhões.....	17
2.3 Planejamento de movimentação de produto	19
2.4 Regressão Linear	22
2.5 Regressão Linear Múltipla.....	23
2.6 Bootstrap	27
3 ESTUDO DE CASO.....	28
3.1 Delineamento da Pesquisa.....	28
3.2 Coleta das Amostras.....	28
3.3 Análise exploratória dos Dados	29
3.4 Análise de regressão	33
3.5 Modelo de regressão múltipla	34
4 CONSIDERAÇÕES FINAIS	45
ANEXO A: MACRO MATLAB PARA <i>BOOTSTRAP</i> POR PARES	46
REFERÊNCIAS BIBLIOGRÁFICAS	47

1 INTRODUÇÃO

A exploração dos metais fez o homem desvendar que para seu desenvolvimento esses bens minerais são essenciais para vida, pois origina em recursos físicos proporcionados pelos minerais depois de extraídos, tratados, utilizados nas siderurgias, disponibilizados para fabricação de manufaturados e bens-materiais e assim, usufruir do conforto oferecido por esses. Os minerais em geral são utilizados para fabricação de bens materiais essenciais a vida humana. O uso de minerais, em especial aos que contém ferro é responsável por grande parte da economia brasileira e mundial.

O cenário atual vivido pelas empresas é de competitividade, essas necessitam de obterem um diferencial a cada momento para tornar o produto como único ou que pelo menos tenha um diferencial positivo quanto ao seu uso pelo grupo de consumidores. Nesse contexto, as empresas têm optado por uma melhoria contínua dos produtos fornecidos, buscando inovar e de forma sustentável.

O transporte de minérios é uma das etapas que se tem um custo elevado na indústria mineral e para isso ser minimizado, deve ser planejado de tal caráter que reduza os custos e aumente a produtividade desse processo. Um dimensionamento de frota eficaz é de suma importância para as empresas.

Um dos métodos usuais de estimativa da produção de frota de caminhões é por meio da utilização dos indicadores operacionais que impactam na capacidade de transporte de massa de minério. Utilizando corretamente os dados históricos é possível obter resultados otimizados que venham a conceder lucros a empresa. É um método simples, mas que proporciona várias análises, pois permite analisar cenários diferentes de produção, a fim de verificar o melhor resultado que visa ter melhor custo benefício.

Com o auxílio de análises estatísticas as quais permitem extrair informações dos dados para obter uma melhor compreensão de situações cotidianas, é possível determinar o quanto o processo é capaz de produzir, verificar possíveis melhorias de desempenho e minimizar impactos negativos no processo em geral.

Realizar um modelo de Regressão Linear Múltipla que esteja adequado a realidade indicará previamente a empresa como cumprir a atividade e manter-se competitiva no mercado.

1.1 Objetivos

1.1.1 Objetivo Geral

O trabalho em estudo refere-se ao setor mineral que tem como principal objetivo desenvolver um modelo a fim de prever a massa de minério possível a ser transportada a partir um de dimensionamento de frota de caminhões de pequeno porte que realizam o transporte do produto da área de produção até ao local de escoamento, onde é realizado por meio do transporte ferroviário.

1.1.2 Objetivos Específicos

- Descrever o processo de mineração e equipamentos utilizados nesse processo com foco no transporte rodoviário;
- Analisar indicadores operacionais de produção dos recursos utilizados nessa atividade;
- Melhorar a alocação técnica e econômica dos recursos disponíveis para transporte de produto;
- Estudar os conceitos fundamentais para estimação de produtividade de frotas de caminhões.

1.2 Justificativa

Os empreendimentos, seja qual o ramo a que se refere, faz necessário tomar decisões, as quais podem levar ao crescimento ou a falência de um empreendedor. As ferramentas utilizadas para efetuar as análises e os estudos de viabilidade tem que ser extremamente fundamentadas, antes de efetuar qualquer investimento. Todos os parâmetros precisam ser analisados garantindo o esperado retorno do investimento.

A obtenção dos lucros empresariais depende de fatores favoráveis à produção, portanto é necessário planejamento e controle de cada processo para manter o produto no mercado com custo e qualidade adequados garantindo a satisfação do consumidor.

As crescentes pressões mundiais pela redução dos custos dos produtos siderúrgicos e aproveitamento mais racional dos recursos energéticos não renováveis, vêm incentivando os

produtores mundiais de matéria-prima a desenvolver produtos que possuam uma relação custo/benefício apropriado às demandas atuais.

As empresas necessitam de flexibilidade, qualidade, custos para manterem se no mercado e obterem lucros. Tompkins et al (1996) as novas tendências de mercado requerem os seguintes tipos de respostas, pelas empresas:

- Possuir capacidade de produzir produtos com qualidade superior, com custo (unitário) reduzido e, entrega no prazo exigido pelos clientes;
- O sistema deve ser projetado para ser flexível e confiável;
- Melhorar continuamente os processos de produção significa elaborar sistemas de manufatura mais sofisticados e competitivos.

As operações de carregamento e transporte representam aproximadamente 60% dos custos operacionais entre todos os processos relacionados, portanto são consideradas críticas e necessitam de análises.

O estudo apresentado é de grande importância sendo que o custo de exploração do minério é significativo. Quando começa a analisar o processo produtivo em comparação ao mercado e toma-se providências para a obtenção de um melhor resultado, o custo de operação tende a minimizar refletindo diretamente em todas as áreas de investimentos da empresa, compra de equipamentos, operação e quadro efetivo de funcionários.

É necessário utilizar de ferramentas teóricas para que a prática seja confiável, assertiva e construir modelos estatísticos é de grande valia a fim de melhorar o desempenho produtivo. Obtendo um dimensionamento de frota correto, os recursos financeiros e humanos são dispostos no conjunto da empresa com resultados lucrativos para empresa.

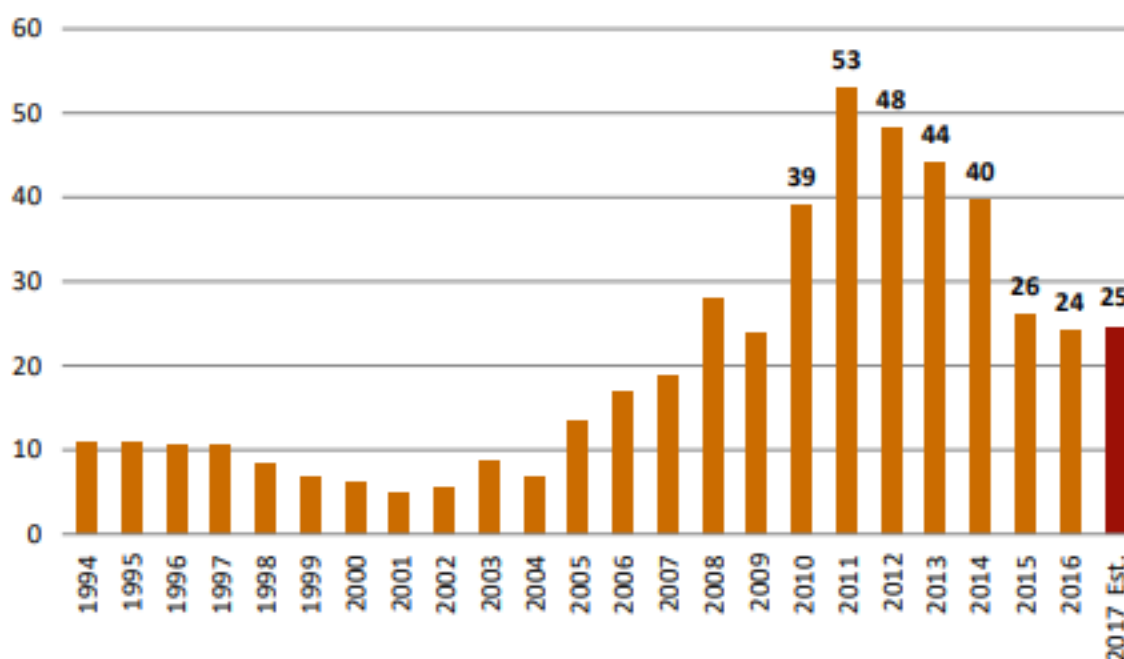
2 REVISÃO DA LITERATURA

A revisão da literatura apresenta um contexto a respeito do setor de minério de ferro, referente a produção na área de mineração, do processo de transporte rodoviário de minério a qual embasa este estudo e abrange a área de estatística quanto a Regressão Linear Múltipla.

2.1 Produção do minério de Ferro

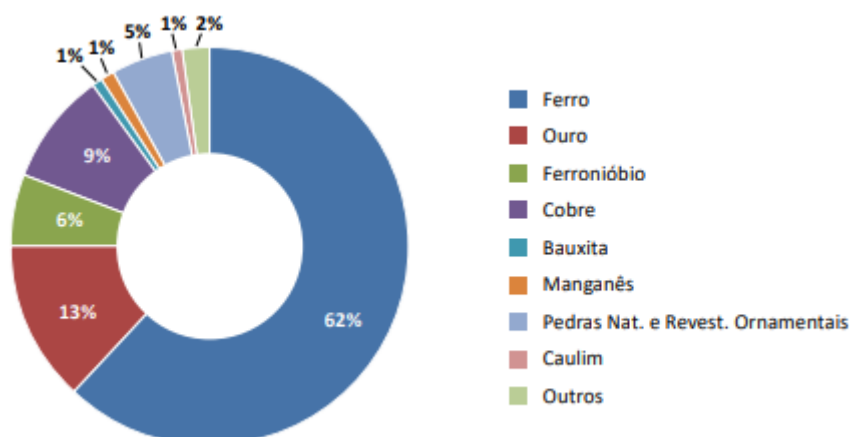
O Brasil é um país de crescente e intensa exploração de minerais (Fig. 1), em especial o minério de ferro (Fig. 2) sendo uns dos minerais mais utilizado na vida do homem.

Figura 1 – Evolução da Produção Mineral Brasileira (bilhões de toneladas)



Fonte: IBRAM (2017)

Figura 2 – Balança Mineral Brasileira

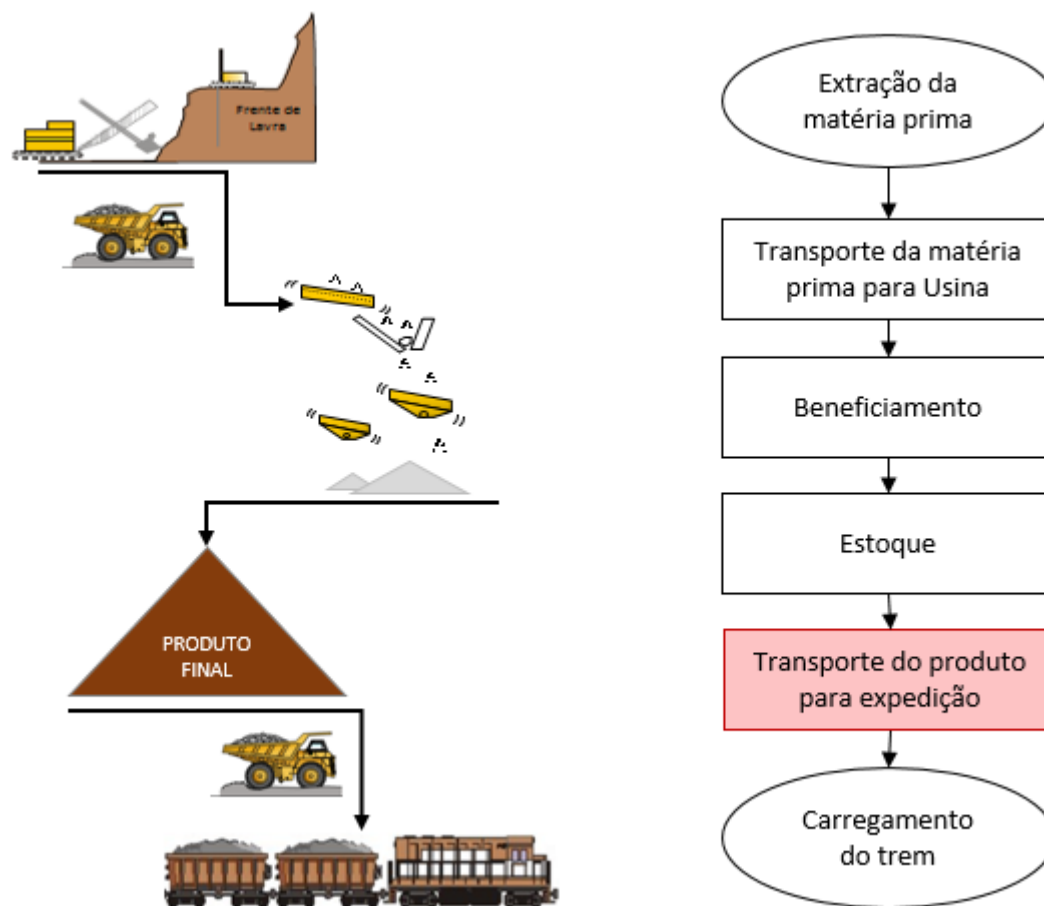


Fonte: IBRAM (2017)

A indústria mineral é um processo amplo que demanda áreas estratégicas, administrativas, apoio operacional, operacionais, comercial entre outras. Faz-se necessário estudo de abertura do empreendimento, tais como estudar as características comerciais do produto (área administrativa) e propriedades da área geografia onde encontra-se as rochas potenciais para explorar o mineral (operacional). Após esses estudos, o projeto de mineração avança com as premissas que darão suporte a escolher os equipamentos e mão de obra necessários, determinar a logística para o desenvolvimento do empreendimento.

De forma simplificada o setor de produção mineral tem como principal objetivo e consiste em extrair a matéria prima das rochas na forma natural em que se encontram na natureza, sendo essa a primeira etapa operacional. Uma segunda etapa denominada de beneficiamento requer em utilizar de processos industriais com equipamentos apropriados, que usando técnicas de separação de elementos químicos útil dos denominados contaminados (ruins que são prejudiciais no produto final) com melhor eficiência possível, disponibilizando o minério de ferro beneficiado como produto final para as indústrias siderúrgicas. Após o produto elaborado, ou seja, pronto para o consumo nas siderúrgicas os quais encontram-se nos estoques da empresa, deve-se ser entregue ao cliente, sendo a etapa fim do processo. Entre essas etapas existe o processo de transporte dos materiais em produção de uma etapa a outra, que pode ser realizado por correias transportadoras ou por caminhões. A Fig. 3 representa o processo da mineração descrito neste parágrafo.

Figura 3 – Fluxograma simplificado do processo de mineração.



Fonte: Elaborado pela autora (2018)

Este estudo de caso analisará o processo de transporte do produto para expedição a fim de obter resultados positivos para empresa.

2.2 O transporte de minério por caminhões

A indústria mineral, como visto anteriormente, é composta por várias etapas e uma das características desse setor é demandar de áreas extensas para suas instalações. Em cada área é realizada uma etapa do processo e entre estas encontra-se a necessidade do transporte seja da matéria-prima ou do produto final de uma a outra.

Um recurso muito utilizado pelas empresas até os dias atuais é o transporte rodoviário, ou seja, realizado por equipamentos móveis de carga denominado caminhões. Mesmo apresentando um custo benefício alto em relação ao outro meio de transporte como as correias transportadoras, as empresas ainda optam por esse transporte uma vez que existe vantagens relevantes em utilizar esse recurso, por exemplo a flexibilidade, uma vez instalado uma correia

transportadora, o processo não poderá sofrer alteração de local, o que não é inviável essa alteração, se fosse utilizado para transportar o produto por caminhões.

Algumas vantagens e desvantagens da mineração com transporte por caminhões principalmente nas rotas em que o material transportado é o produto final segundo LOPES (2010) são apresentadas a seguir.

As vantagens são:

- Alta flexibilidade operacional;
- O transporte pode ocorrer simultaneamente em várias frentes operacionais, facilitando a blendagem do material (mistura de dois ou mais produtos), garantindo a qualidade a ser entregue para o cliente;
- Os caminhões podem ser utilizados em outras frentes caso a frente que estiver ocorra alguma interdição;
- Alta mão de obra disponível para contratação no mercado de trabalho, devido ao fato de ser comum a utilização desse recurso nas mineradoras;
- Os equipamentos de pequeno porte são montados nas fábricas e ao chegar na mineração estão pronto para uso, se os equipamentos forem de grande porte, esses são montados em parte e ao chegar na mineração monta-se o conjunto;
- Os acessos e praças de trabalho podem ser realizados com os próprios equipamentos que foram adquiridos para realizar o transporte;
- As operações de carregamento dos caminhões pode sofrer alterações se necessário devido a capacidade de caminhões disponíveis;
- Em caso da área de trabalho apresentar algum risco iminente é fácil a evacuação dos equipamentos.

As desvantagens são:

- A eficiência energética é baixa, sendo que 50% são para o próprio deslocamento do equipamento até a frente de trabalho e 50% para o deslocamento das cargas;
- Devido a limitação de rampas nos acessos, as estradas tornam-se longas para que esse problema seja contornado;
- A abertura e conservação das vias de acesso dos caminhões possuem custo alto;
- Devido aos intemperes (chuvas e neblinas) a operação fica pouco produtiva ou até mesmo é paralisada devido à instabilidade de tração e baixa visibilidade;
- Necessidades de equipamentos auxiliares na operação tais como equipamentos de infraestrutura para conservação das vias e pipas para realizar a umectação das vias de acesso

a fim de reduzir a poeira (tornando as vias úmidas), melhorando a visibilidade dos operadores e a redução do impacto ambiental;

- A ocorrência de alteração dos indicadores operacionais pode causar a necessidade de mais caminhões para garantir a produção.

2.3 Planejamento de movimentação de produto

O planejamento da movimentação é de nível tático da organização, por integrar a estratégia com os níveis operacionais envolvidos na execução do plano.

A realização de um plano de movimentação com caminhões de pequeno porte deve considerar premissas operacionais que são características desse processo. São estas segundo SILVA (2009):

- Disponibilidade do equipamento: representar o fator em que o equipamento está disponível para operar ou seja não apresenta problemas mecânicos, existe mão de obra para operá-lo, não há impedimentos nas vias para que o transporte ocorra como obras, chuvas fortes. Como estes fatos existirão sempre, faz-se necessário estimar qual percentual implicará devido aos fatores citados, nesse caso a disponibilidade é dividida em disponibilidade mecânica e em disponibilidade física. A disponibilidade mecânica são os fatores que implicam nesse índice está relacionado com as condições mecânicas do caminhão em relação ao tempo teórico possível do período que deseja executar a atividade e é calculada conforme equação 1. A disponibilidade física corresponde ao percentual do tempo que os caminhões estão em condições de executar o transporte de material em relação ao tempo teórico possível do período, é calculada pela equação 2.

$$DM = (HT - (MP + MC + TP)) / HT \times 100\% \quad (1)$$

$$DF = (HP - HM) / HT \times 100\% \quad (2)$$

Onde:

DM = disponibilidade mecânica;

HT = representa o total de horas teóricas possíveis no período analisado;

MP = manutenção preventiva, representa todo o tempo de serviço de manutenção programada para o equipamento, qual objetivo é a conservação e inspeção dele, a fim de mantê-lo em condições satisfatórias de operação;

MC = manutenção corretiva, representa todo o tempo de serviço de manutenção não programada do equipamento cuja finalidade é corrigir deficiências que possam paralisá-lo;

TP = tempo perdido correspondente à locomoção do equipamento por motivos de intervalos do operador (almoço, café, troca de turno, entre outros);

DF = disponibilidade física;

HP = representa às horas totais calculadas para o período, na base dos turnos previstos e considerando a disponibilidade mecânica;

HM = representa às horas de reparos na oficina ou em campo, incluindo a falta de peças no estoque ou falta de equipamentos auxiliares.

- Utilização (física) do equipamento: corresponde à parcela em que o equipamento está em operação, portanto trata-se de fator aplicável sobre as horas disponíveis do equipamento, ou seja, é o percentual do tempo que os caminhões esteve em operação em relação do tempo disponível do equipamento. Os fatores como falta de operadores, condições climáticas, preparação das frentes de trabalho, paralização de equipamentos auxiliares impactarão na utilização. O fator utilização é calculado a partir da equação 3.

$$UT = HT/(HP-HM) \times 100\% \quad (3)$$

HT = total de horas efetivamente trabalhadas;

HP = representa às horas totais calculadas para o período, na base dos turnos previstos e considerando a disponibilidade mecânica;

HM = representa às horas de reparos na oficina ou em campo, incluindo a falta de peças no estoque ou falta de equipamentos auxiliares.

- Rendimento operacional: expressa, em percentual, o tempo de operação efetiva dos caminhões, ou seja, é a relação entre as horas efetivamente trabalhadas e as horas programadas. Esse fator é calculado a partir do produto da disponibilidade física pela utilização.

- Tempo de ciclo: tempo total gasto para executar um conjunto de determinadas operações como: manobra, carga, descarga, basculamento, deslocamento, etc. É calculado através do somatório de tempo de todas as atividades que compõem essa operação.

- Produtividade efetiva dos equipamentos: eficiência do processo em transportar em relação ao tempo.

No sistema de gestão de produção, a atividade inicial trata-se do planejamento de produção, a qual é uma atividade difícil, dado o número de variáveis existentes e à dinâmica desses sistemas, marcada por constantes mudanças (RODRIGUES, 1994).

Usando os indicadores operacionais de produção é possível estimar a produção dos equipamentos segundo CLARKE et al (1990), conforme equação 4.

$$PR = DF \times UT \times PR \times HTP \times Nequip \quad (4)$$

Onde:

P = produção da frota;

DF = disponibilidade física dos equipamentos;

UT = fator de utilização dos equipamentos;

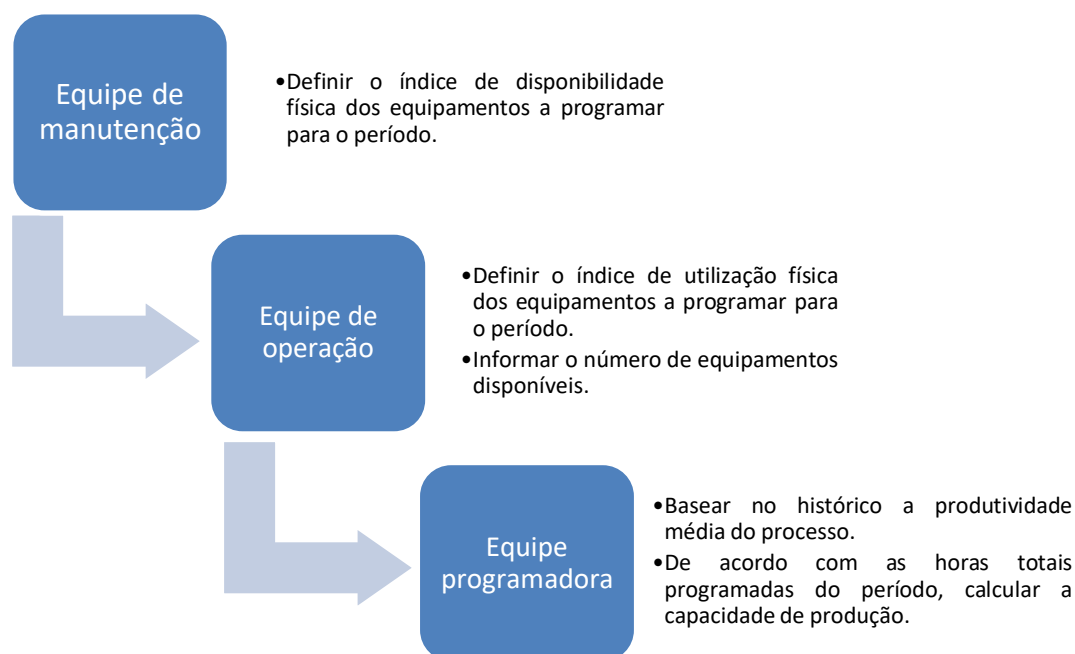
PR = produtividade efetiva dos equipamentos;

HTP = horas totais programadas;

Nequip = número de equipamentos na frota.

O dimensionamento (Fig. 4) torna-se mais adequado quanto maior a aderência dos indicadores de produção programados aos praticados. Portanto os indicadores para os estudos a realizar, devem ser fundamentos em estudos de campos, aderentes as médias históricas praticada na operação evitando a subestimação ou superestimação desses ocasionando um mau dimensionamento da frota, perda de produtividade e elevação do custo operacional.

Figura 4 – Estratégica para elaboração do dimensionamento de frota de caminhões



Fonte: Elaborado pela autora (2018)

2.4 Regressão Linear

A regressão linear é um método estatístico para se estimar o valor esperado de uma variável y (resposta), dados os valores de algumas outras variáveis x (preditoras), as quais estão relacionadas de maneira não determinísticas.

Assim, o termo regressão foi introduzida por Francis Galton (1822-1911) em um estudo da relação entre as alturas dos pais e filhos, em 1885, no qual concluiu que os pais altos tendem a ter filhos altos e os pais baixos tendem a ter filhos baixos. Percebeu também que muitos dos pais de grande estatura têm filhos menores e muitos dos pais de pequena estatura têm filhos mais altos do que eles próprios.

Durante seis anos, Galton “conseguiu 9000 registros familiares, muitos deles completos, que levaram dez anos para serem analisados” (CONT, 2008, p.204). Ele possuía as frequências observadas como provas de sua teoria, mas ainda necessitava de “descrever os mecanismos de transmissão tanto dos caracteres quanto dos talentos” (CONT, 2008, p.206). Assim, Galton descobriu a regressão à média. Se a regressão à média não ocorre, os filhos de pais altos seriam ainda mais altos; os netos mais altos ainda e assim os seres humanos mais altos seriam cada vez mais altos (MLODINOW, 2009).

A regressão diz-se linear porque a relação entre as variáveis pode ser representada por uma função linear. Quando se dispõe de apenas uma variável controlada o diagrama de dispersão é uma representação gráfica que permite visualizar a nuvem de pontos que caracteriza a Relação funcional entre as duas variáveis.

Caso o modelo seja composto por apenas uma variável independente (ou único regressor ou preditor), o modelo é denominado como regressão linear simples, em que a média da variável aleatória Y esteja relacionada a X pela seguinte forma generalizada.

$$E(Y|X_i) = \beta_1 + \beta_2 X_i \quad (5)$$

em que Y é a variável dependente, X_i é a variável independente, β são os coeficientes de regressão, em que β_1 e β_2 são conhecidos também como intercepto e coeficiente de inclinação, respectivamente (GUJARATI, 2011).

Generalizando em um modelo linear probabilístico, para um valor fixo de X , o valor real de Y seja determinado pela função do valor médio mais um termo de erro aleatório,

$$Y = \beta_1 + \beta_2 X_i + \epsilon \quad (6)$$

em que ϵ é o termo de erro aleatório (MONTGOMERY, 2015).

A maioria dos problemas práticos envolve mais de uma variável independente para fins de previsão, ou seja, contém mais de um regressor, dessa forma o modelo com k variáveis regressoras é denominado como regressão linear múltipla conforme a equação 7.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \dots + \beta_k X_k + \epsilon \quad (7)$$

em que Y é a variável dependente, X são as variáveis independentes, β são os coeficientes de regressão, em que β_0 é a interseção do plano, $\beta_j, j=0,1, \dots, k$, são os coeficientes parciais de regressão e ϵ refere-se ao termo de erro aleatório (MONTGOMERY, 2015).

2.5 Regressão Linear Múltipla

O problema apresentado neste trabalho possui mais de uma variável independente (preditora), o que justifica a necessidade de utilizar da metodologia de regressão linear múltipla.

A análise de regressão múltipla é a técnica mais adequada quando se deseja investigar simultaneamente os efeitos que várias variáveis independentes poderão causar na variável dependente. Mesmo quando o interesse está voltado para uma única variável independente é aconselhável incluir outras variáveis capazes de afetar a variável dependente. Pois em regressão linear múltipla todas as mudanças ocorridas em uma variável podem ser explicadas por uma referência a mudanças em várias outras variáveis.

No ajuste de um modelo de regressão múltipla, podem ser compreendidas utilizando a notação matricial, em que, haja k variáveis regressoras, n observações e que o modelo seja conforme equação 7 citada anteriormente. Trata-se de sistema de n equações, que pode expresso na notação matricial como,

$$y = X\beta + \epsilon \quad (8)$$

Sendo:

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, X = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1k} \\ 1 & X_{21} & X_{22} & \dots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \dots & X_{nk} \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} \text{ e } \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix},$$

Em que, Y é o vetor das observações (n x 1), X é a matriz das variáveis independentes (n x p), β é o vetor dos coeficientes de regressão (n x 1) e ϵ é vetor dos erros aleatórios (MONTGOMERY, 2015).

Assim a estimação dos coeficientes de regressão múltipla, é possível recorrer ao método dos mínimos quadrados, o qual tem a finalidade de encontrar a reta que minimiza a distância entre os pontos observados e a reta.

Minimizando a equação tem-se:

$$L = \sum_{i=1}^n \epsilon_i^2 = \epsilon' \epsilon = (Y - X\beta)'(Y - X\beta) \quad (9)$$

O objetivo é minimizar L com relação a $\beta_0, \beta_1, \dots, \beta_k$. Derivando em função de β e igualando a zero, tem-se o estimador de mínimos quadrados de β :

$$\hat{\beta} = (X'X)^{-1}X'Y \quad (10)$$

Obtido o vetor de parâmetros, deve se observar a significância dos resultados obtidos. Portanto deve estimar a variância do termo do erro conforme a seguir (MONTGOMERY, 2015).

A soma de quadrados dos resíduos é dada pela expressão:

$$SQ_E = Y'Y - \hat{\beta}X'Y \quad (11)$$

A soma de quadrados da regressão é dada pela expressão:

$$SQ_R = \hat{\beta}X'Y - \frac{\sum_{i=1}^n (Y_i)^2}{n} \quad (12)$$

E, por fim, a soma de quadrados total é dada pela expressão:

$$SQ_T = Y'Y - \frac{\sum_{i=1}^n (Y_i)^2}{n} \quad (13)$$

Dessa forma a soma de quadrados dos resíduos é conhecida como a decomposição da soma de quadrados total e pode ser escrita como:

$$SQ_E = SQ_T - SQ_R \quad (14)$$

Estatisticamente essas somas de quadrados seguem distribuição de qui-quadrado, caso haja normalidade, sendo que SQ_E , SQ_R e SQ_T possuem $n-p$, $p-1$ e $n-1$ grau de liberdade respectivamente.

Em problemas de regressão linear múltipla, certos testes de hipóteses sobre os parâmetros do modelo são úteis para verificar a "adequabilidade" do modelo.

Assim, através da estatística F pode-se determinar se há uma relação linear entre a variável resposta e algumas das variáveis preditoras. Considere-se testar as hipóteses:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1: \text{pelo menos um } \beta_i \neq 0, i = 1, 2, \dots, k.$$

Se rejeitar H_0 , tem-se que ao menos uma variável explicativa contribui significativamente para o modelo.

A partir dos resultados descritos até o momento é possível construir o quadro de análise da variância com a finalidade de calcular a estatística F .

Quadro 1 – Análise de variância e estatística F

Curvas de variação	Graus de Liberdade	Soma dos quadrados	Quadrados mínimos	Teste F
Regressão Linear	k= p-1	SQReg	$\frac{SQ_R}{p-1}$	$\frac{SQ_R}{SQ_E}$
Resíduo	n - p	SQRes	$\frac{SQ_R}{n-p}$	-
Total	n - 1	SQTot	-	-

Fonte: Elaborado pela autora (2018)

A proporção da variabilidade de Y explicada pelas variáveis preditoras é verificado pelo coeficiente de determinação (MONTGOMERY, 2015). Assim, quanto mais próximo de 1, maior é a explicação da variável resposta pelo modelo ajustado.

O coeficiente de determinação é dado por:

$$R^2 = \frac{SQ_R}{SQ_T} \quad (15)$$

O coeficiente de determinação ajustado é definido como:

$$R_{ajus}^2 = 1 - \frac{SQ_E/(n-p)}{SQ_T/(n-1)} \quad (16)$$

Na maioria dos casos usam da estatística do coeficiente de determinação ajustado, uma vez que $SQ_E/(n-p)$ é a média quadrática do resíduo e $SQ_T/(n-1)$ é uma constante, assim somente aumentará quando uma varável for adicionada no modelo, se a média quadrática do resíduo diminuir (MONTGOMERY, 2015).

Verificar se as variáveis preditoras são importantes e necessárias para o modelo é fundamental, para isso é necessário realizar os testes de hipóteses para o modelo e hipóteses individuais para os coeficientes da regressão, podendo o modelo ser eficaz e melhor com a inclusão ou com a exclusão de novas variáveis.

As hipóteses para testar a significância são dadas por:

$$H_0: \beta_0 = \beta_1 + \beta_2 + \dots + \beta_k = 0$$

$$H_1: \text{pelo menos um } \beta_j \neq 0, j = 1, 2, \dots, k$$

Pode se utilizar da estatística F conforme Tab. 01, a fim de testar a existência de regressão no modelo, assim se $f_0 > f_{\alpha, p-1, n-p}$, rejeita-se a hipótese nula ao nível de significância (α) em favor da hipótese alternativa (H_1), que no mínimo um dos coeficientes é diferente de zero ($\beta_j \neq 0$), verificando-se a existência de regressão linear entre as variáveis do modelo (MONTGOMERY, 2015).

Para testar a influência de uma variável explicativa m dos testes mais utilizado é o teste t, as hipóteses para testar a significância dos coeficientes de regressão de forma individual são dadas por:

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0, j = 1, 2, \dots, k$$

Se H_0 não é rejeitada, pode-se afirmar que, ao nível de significância (α), essa variável X_j pode ser retirada do modelo pois não é significativa no modelo.

A fórmula para realizar o teste é:

$$T_j = \frac{\hat{\beta}_j}{s(\hat{\beta}_j)}, j = 1, 2, \dots, k \quad (17)$$

No qual:

$S(\hat{\beta})$ – estimador do erro padrão de β_i

T_j – estatística t de Student com n-p graus de liberdade.

Se o p-valor menor que ao nível de significância desejado, rejeita-se a hipótese nula ao nível de significância (α) em favor da hipótese alternativa (H_1), verificando-se a variável é importante para o modelo.

Um fator importante na regressão é a análise dos resíduos, que representa um conjunto de análises para investigar se o modelo de regressão é adequado. Se o modelo for adequado, as suposições a seguir são observadas na análise dos resíduos concluindo que o modelo está adequado. Sendo:

- São independentes;
- Possuem variância constante;
- Possuem média zero;
- Seguem distribuição normal.

As técnicas mais utilizadas para realizar a análise são gráficas podendo também fazer o uso de outras técnicas como testes estatísticos a partir dos resíduos gerados no modelo. As técnicas gráficas, por serem visuais, podem ser subjetivas e por isso técnicas formais são mais indicadas para a tomada de decisão.

2.6 Bootstrap

Em 1979 Bradley Efron propôs o método *Bootstrap* pela primeira vez. Esta técnica usualmente é utilizada como um procedimento não paramétrico de reamostragem computacional intensivo usado para estimação de distribuições amostrais (EFRON, 1979).

Para a resolução de problemas de regressão em presença de não normalidade dos erros, e heterocedasticidade, o *bootstrapping* poderá ser aplicado. Neste trabalho utilizaremos o *Bootstrap* de Pares.

No *Bootstrap* de Pares é realizado uma amostragem dos pares originais de valores (Y_i, x'_i) , onde Y_i é a i -ésima observação e x'_i é a i -ésima linha da matriz X . Os principais passos são descritos abaixo.

1. Retirar uma amostra aleatória com reposição de tamanho n dos dados originais (Y_i, x'_i) , tem-se os \hat{Y}_i^{BP}, X_i^{BP} ;

2. Ajusta-se um modelo de MQO, considerando a amostra coletada anteriormente. Isto é,

$$\hat{Y}^{BP} = X^{BP} \hat{\beta}^* + \hat{\varepsilon} \quad (18)$$

Ajusta-se este modelo para obter as estimativas de $\hat{\beta}^*$

3. Repetir os passos de 1 e 2 R vezes.

Depois de R reamostragens, obtém-se a distribuição empírica de $\hat{\beta}^*$ e todas as estatísticas empíricas relativas a ela. Um intervalo de confiança percentil *Bootstrap* de pares e usar $\hat{\beta}_{(R+1)(\alpha^*/2)}^*$ e $\hat{\beta}_{(R-1)(\alpha^*/2)}^*$ como limites inferior e superior do intervalo $100(1-\alpha^*)\%$ de confiança, é possível verificar a significância dos parâmetros do modelo. Se o intervalo conter o valor zero então não podemos rejeitar que o parâmetro seja nulo ao nível de confiança $100(1-\alpha^*)\%$.

3 ESTUDO DE CASO

Este capítulo apresenta os estudos desenvolvidos, a qual desenvolveu de forma exploratória descritiva de natureza quantitativa, realizando uma análise referente ao sistema de transporte do produto final da produção até a área em que será expedido.

3.1 Delineamento da Pesquisa

O trabalho abordou o estudo de caso, assim a pesquisa juntou-se de fontes tais como dissertações e livros, para fundamentar-se na parte teórica, realizando uma revisão bibliográfica do conceito sobre o tema proposto e da coleta de dados, as quais se desenvolveram simultaneamente, uma auxiliando a outra, no sentido de atingir os objetivos propostos.

Os dados de interesse para esse trabalho são indicadores operacionais, sendo estes: disponibilidade física, utilização efetiva, produtividade efetiva, rendimento operacional e DMT, necessários para realizar o dimensionamento de frota de caminhões de pequeno porte, a fim de obter dados analíticos para tomada de decisão na empresa, melhor custo benefício para uma empresa de mineração ferrífera.

Para se realizar uma pesquisa é preciso promover o confronto entre os dados, as evidências, as informações coletadas sobre determinado assunto e o conhecimento teórico acumulado a respeito dele (LUDKE, 1986).

A empresa não se disponibilizou a divulgação do nome, sendo assim considerada com uma empresa situada no Quadrilátero Ferrífero.

3.2 Coleta das Amostras

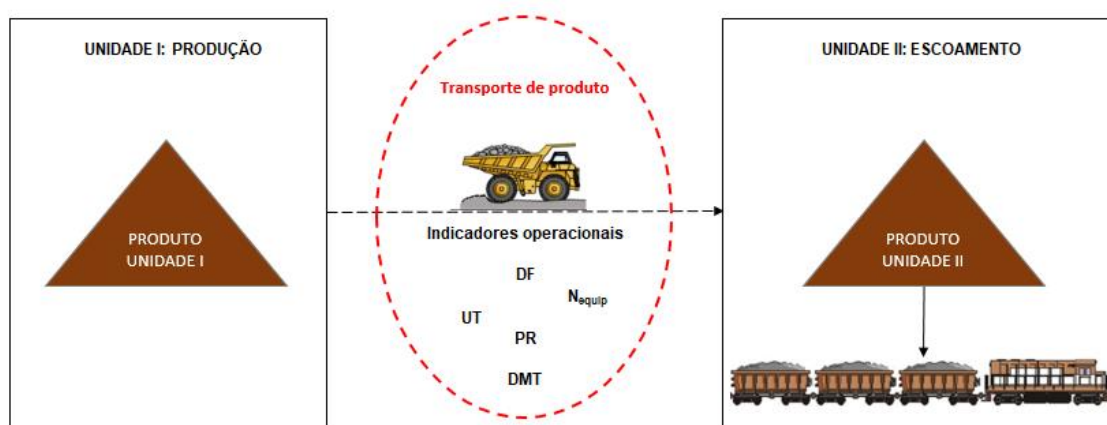
As amostras são representadas por dados referente aos indicadores operacionais diários do período de janeiro de 2015 a agosto de 2017 nos dias em que ocorreu movimentação de produto, totalizando em 959 dias, ou seja, para cada indicador foram coletados 959 dados referente a uma empresa situada no Quadrilátero Ferrífero, região localizada no centro-sul do estado de Minas Gerais que abrange os municípios de Caeté, Congonhas, Itabira, João Monlevade, Mariana, Ouro Preto e Santa Bárbara, entre outros.

Os indicadores operacionais foram coletados a partir do sistema computacional utilizado pela empresa, o qual reuni as informações dos indicadores a cada viagem do caminhão, pois esse permite gravar o tempo percorrido, distância média, peso da massa carregada, entre

outros dados necessários para compor esses indicadores. O sistema permite até mesmo que todas viagens realizadas sejam agrupadas referente a todos caminhões utilizados no dia, apresentando assim, o resultado de cada indicador diariamente. As informações foram organizadas em um banco de dados para conter os dados necessários para a realização da análise dos mesmos.

O banco de dados está composto pelos indicadores: disponibilidade física, distância média percorrida, massa transportada, produtividade efetiva, rendimento operacional e utilização efetiva, referente ao realizado no dia. Os equipamentos que são responsáveis por carregar os caminhões, denominados equipamento de carga, podem também interferir no dimensionamento, porém, para esse estudo não foi considerado esses pois, nesta empresa existe uma alta disponibilidade dos mesmos, não afetando o processo para o desenvolvimento desse trabalho uma vez que ocorre a parte.

Figura 5 – Desenho esquemático transporte de minério



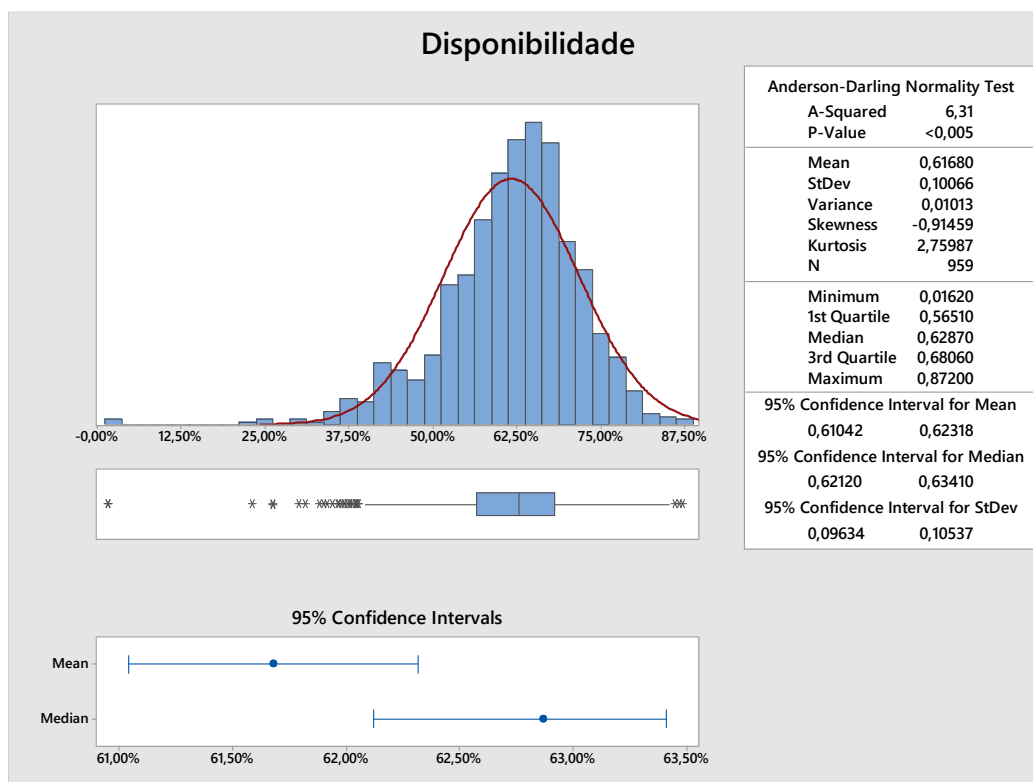
Fonte: Dados da pesquisa (2018)

A partir desses indicadores foi possível obter o indicador rendimento operacional por meio do produto da disponibilidade física e utilização.

3.3 Análise exploratória dos Dados

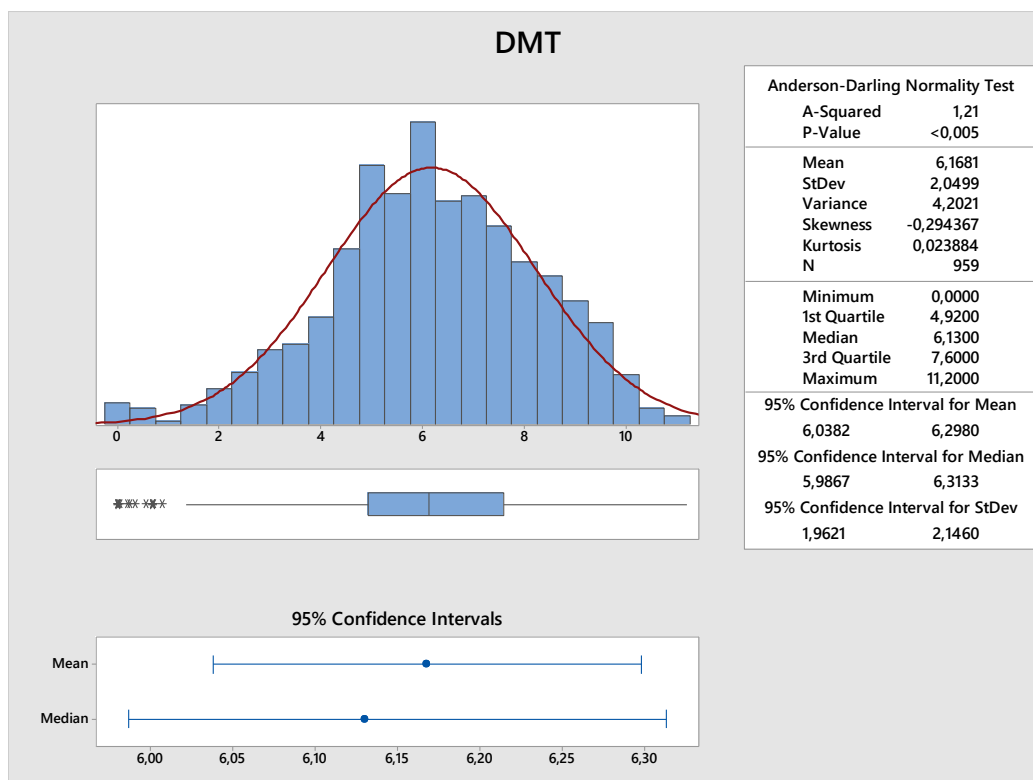
Em posse das informações coletadas no período descrito anteriormente, tornou-se possível conhecer, determinar as características dos conjuntos de dados dos indicadores operacionais usando as estatísticas de análise exploratório dos dados, representadas nas Fig. 6-11.

Figura 6 – Estatísticas: disponibilidade



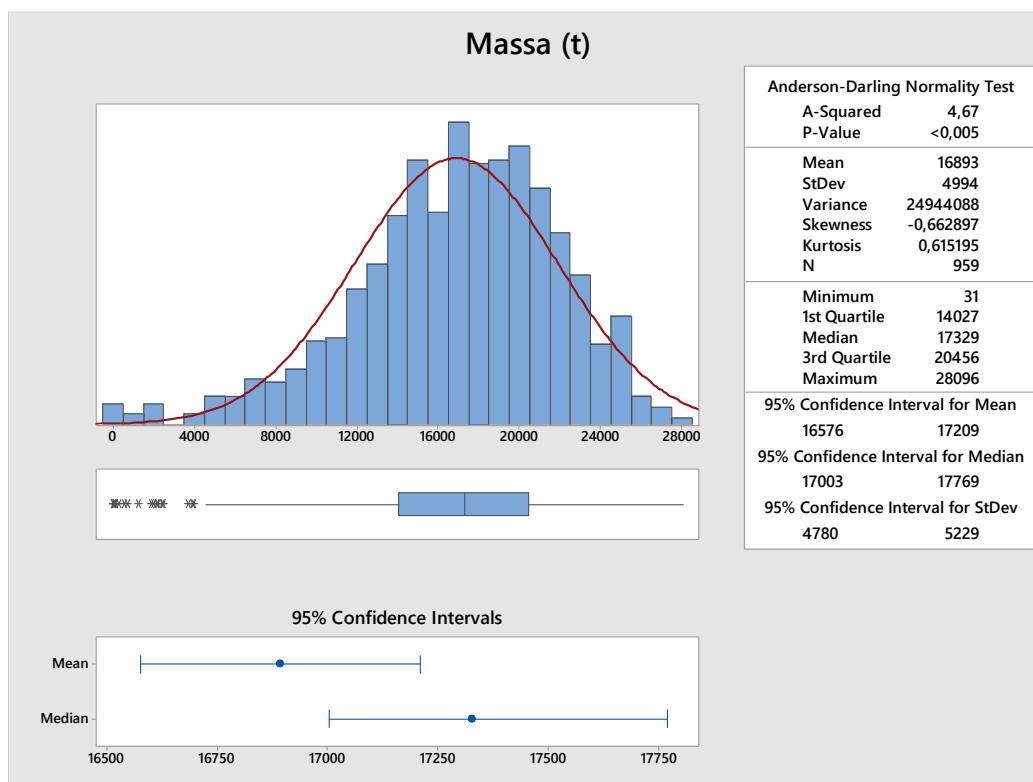
Fonte: Saída do Minitab (2018)

Figura 7 – Estatísticas: distância média percorrida



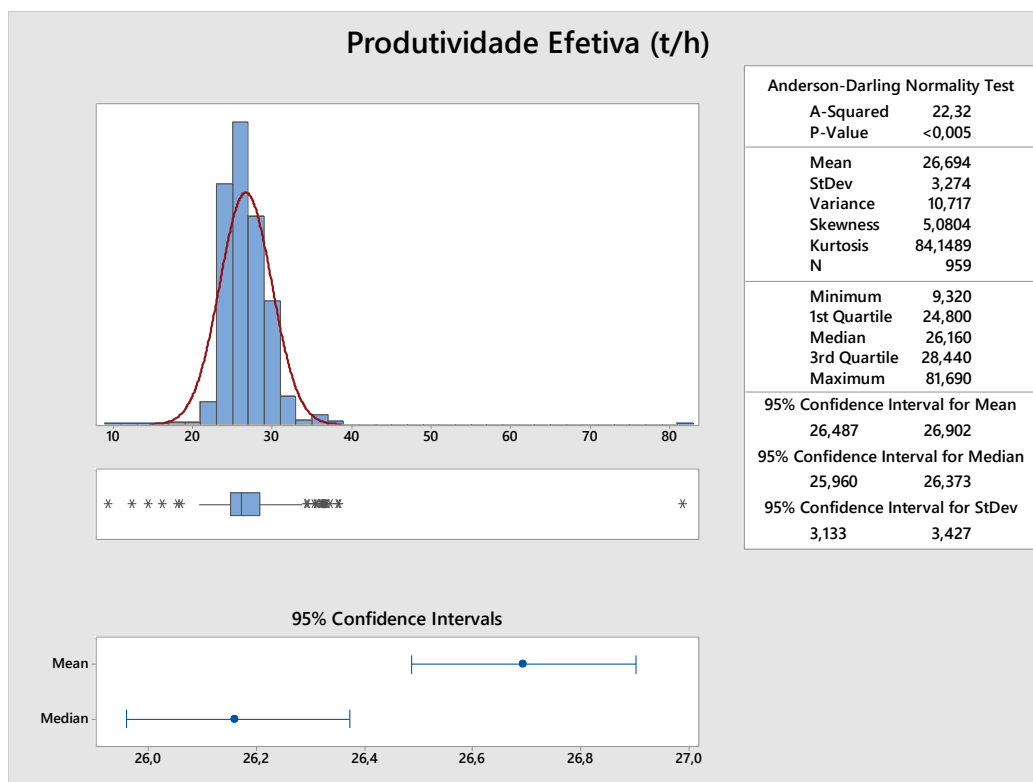
Fonte: Saída do Minitab (2018)

Figura 8 – Estatísticas: massa



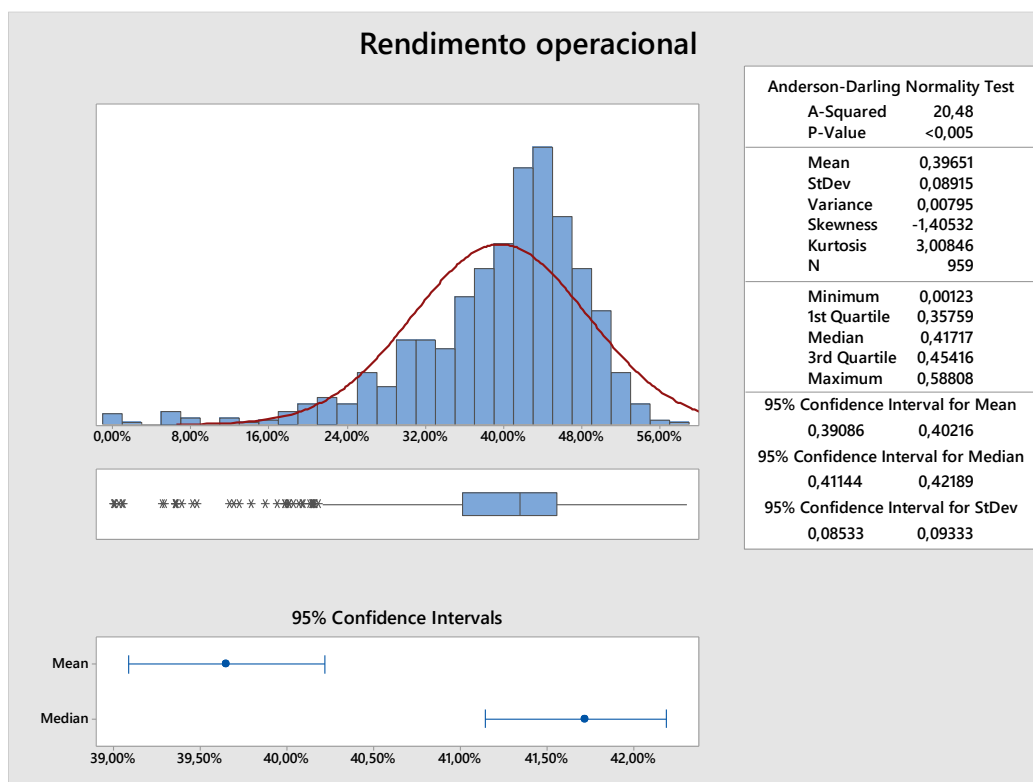
Fonte: Saída do Minitab (2018)

Figura 9 – Estatísticas: produtividade efetiva



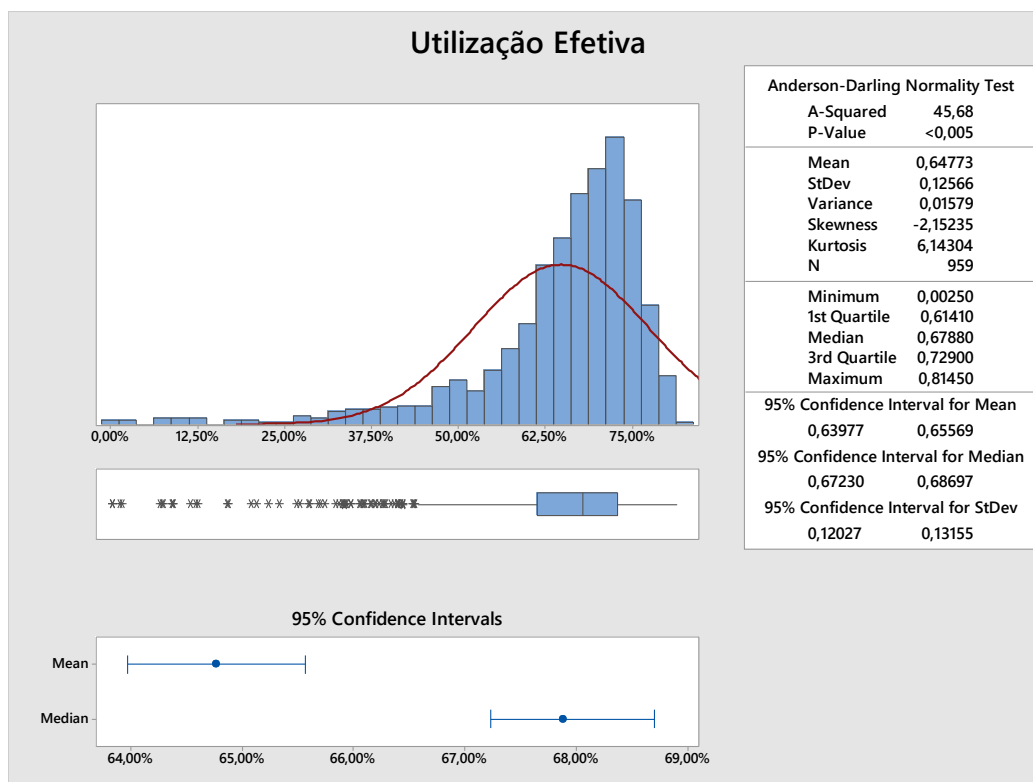
Fonte: Saída do Minitab (2018)

Figura 10 – Estatísticas: rendimento operacional



Fonte: Saída do Minitab (2018)

Figura 11 – Estatísticas: utilização efetiva



Fonte: Saída do Minitab (2018)

3.4 Análise de regressão

O objetivo da criação do modelo é prever a capacidade de massa a transportar e conforme a revisão literária, faz-se necessário a definição da variável resposta e as várias preditoras as quais serão testadas no modelo. Nesse trabalho as variáveis foram classificadas em variável resposta e preditora conforme o quadro a seguir:

Quadro 2 – Descrição das variáveis

Variáveis:	Descrição:	
Massa (t)	Variável dependente / resposta	Quantidade de material em toneladas transportado entre uma unidade e outra durante o dia.
Produtividade Efetiva (t/h)	Variável independente / preditora	Eficiência do processo em transportar o material por hora.
Disponibilidade Física (%)	Variável independente / preditora	Percentual do tempo que os caminhões estão em condições de executar o transporte de material em relação a hora total do dia.
Utilização Efetiva (%)	Variável independente / preditora	Percentual do tempo que os caminhões esteve em operação em relação do tempo disponível (DF).
Rendimento Operacional (%)	Variável independente / preditora	Expressa, em percentual, o tempo de operação efetiva dos caminhões, sendo o produto dos indicadores Disponibilidade Física e Utilização Efetiva
DMT (km)	Variável independente / preditora	Distância média percorrida das viagens.

Fonte: Elaborado pela autora (2018)

A fim de verificar a relação entre as variáveis citadas, foi construída a Tab. 1 de correlação:

Tabela 1 – Correlação das variáveis

	Massa (t)	Produtividade Efetiva (t/h)	Disponibilidade	Utilização Efetiva	RO	DMT
Massa (t)	1					
Produtividade Efetiva (t/h)	0,1928	1				
Disponibilidade	0,3059	-0,3624	1			
Utilização Efetiva	0,5556	0,1244	-0,2385	1		
RO	0,7273	-0,0728	0,4569	0,7348	1	
DMT	0,9713	0,2473	0,2082	0,5090	0,6127	1

Fonte: Elaborado pela autora (2018)

É possível avaliar através dos resultados obtidos o quanto as variáveis preditoras relacionam com a variável resposta.

3.5 Modelo de regressão múltipla

O modelo de regressão múltipla foi elaborado no software Minitab, o qual determina os coeficientes, as análises dos testes como o cálculo da estatística F e T e o valor-p para que seja avaliado o modelo e entre testes foram realizados no Matlab.

Inicialmente relacionamos a variável massa em função das variáveis preditoras conforme Fig. 12-14.

Figura 12 – Modelo Completo

Analysis of Variance					
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	5	23259677499	4651935500	6963,39	0,000
Produtividade Efetiva (t/h)	1	2039310	2039310	3,05	0,081
Disponibilidade	1	3266	3266	0,00	0,944
Utilização Efetiva	1	4190603	4190603	6,27	0,012
RO	1	53440691	53440691	79,99	0,000
DMT	1	8846979741	8846979741	13242,86	0,000
Error	953	636657837	668056		
Total	958	23896335336			

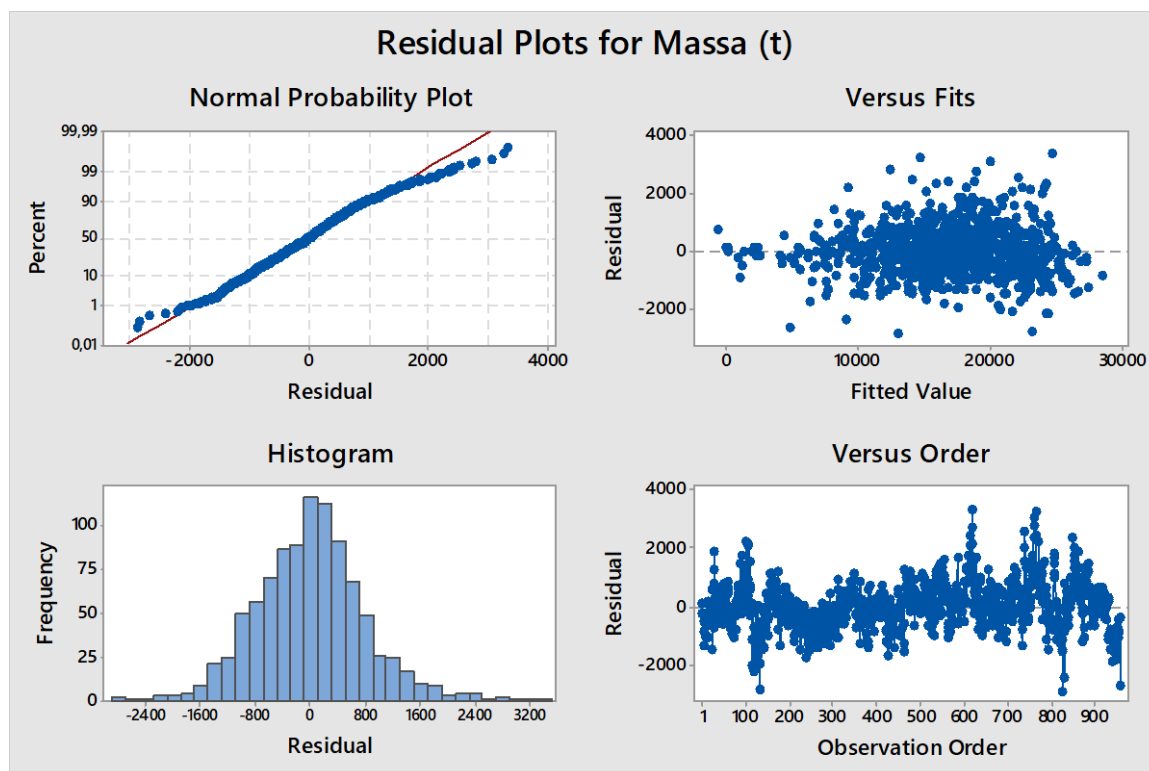
Model Summary			
S	R-sq	R-sq(adj)	R-sq(pred)
817,347	97,34%	97,32%	97,29%

Coefficients					
Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-270	848	-0,32	0,750	
Produtividade Efetiva (t/h)	17,45	9,99	1,75	0,081	1,53
Disponibilidade	-75	1069	-0,07	0,944	16,59
Utilização Efetiva	-2704	1080	-2,50	0,012	26,39
RO	14711	1645	8,94	0,000	30,84
DMT	2052,8	17,8	115,08	0,000	1,92

Regression Equation	
Massa (t) = -270 + 17,45 Produtividade Efetiva (t/h) - 75 Disponibilidade - 2704 Utilização Efetiva + 14711 RO + 2052,8 DMT	

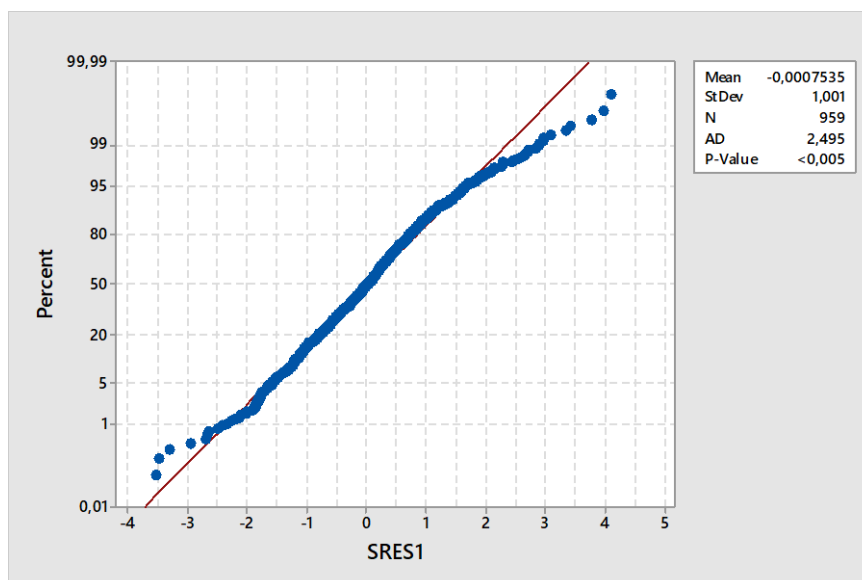
Fonte: Saída do Minitab (2018)

Figura 13 – Análise de Resíduos



Fonte: Saída do Minitab (2018)

Figura 14 – Teste de normalidade para os resíduos



Fonte: Saída do Minitab (2018)

Apesar do modelo apresentar um coeficiente de determinação alto (97,29%) os testes de hipóteses ficam comprometidos pois existe forte evidência para rejeitar que os resíduos são

normais ($p\text{-value} < 0,005$). Assim, optamos por utilizar transformação de Box-Cox. O resultado obtido é apresentado na Fig. 15-16.

Figura 15 – Modelo Completo com transformação de Box-Cox

Rounded λ	1
Estimated λ	1,03382
95% CI for λ	(0,996323; 1,07232)

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	5	23259776537	4651955307	6963,40	0,000
Produtividade Efetiva (t/h)	1	2039683	2039683	3,05	0,081
Disponibilidade	1	3311	3311	0,00	0,944
Utilização Efetiva	1	4192374	4192374	6,28	0,012
RO	1	53444025	53444025	80,00	0,000
DMT	1	8847132748	8847132748	13243,05	0,000
Error	953	636659562	668058		
Total	958	23896436099			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
817,348	97,34%	97,32%	97,29%

Coefficients

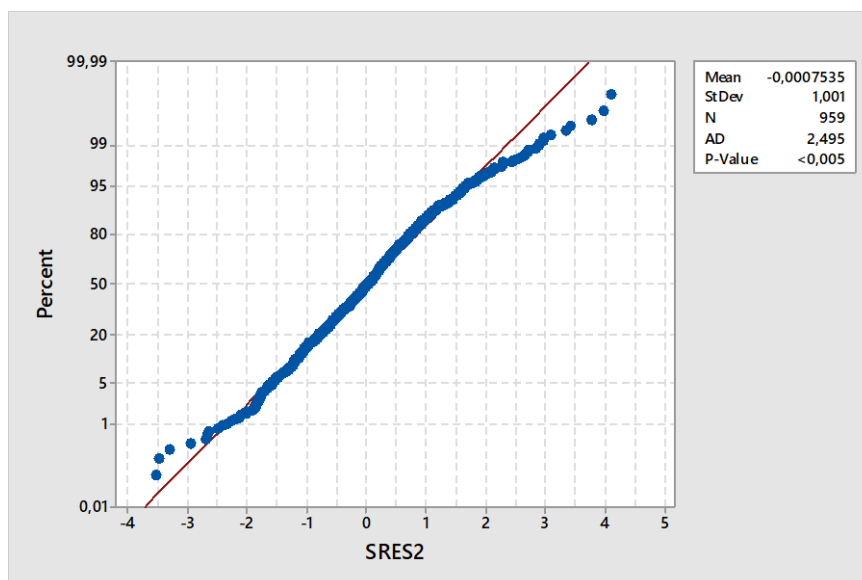
Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-270	848	-0,32	0,750	
Produtividade Efetiva (t/h)	17,45	9,99	1,75	0,081	1,53
Disponibilidade	-75	1069	-0,07	0,944	16,59
Utilização Efetiva	-2705	1080	-2,51	0,012	26,39
RO	14711	1645	8,94	0,000	30,84
DMT	2052,9	17,8	115,08	0,000	1,92

Regression Equation

Massa (t) = -270 + 17,45 Produtividade Efetiva (t/h) - 75 Disponibilidade
- 2705 Utilização Efetiva + 14711 RO + 2052,9 DMT

Fonte: Saída do Minitab (2018)

Figura 16 – Teste de normalidade para os resíduos após transformação



Fonte: Saída do Minitab (2018)

Porém a evidência de não normalidade dos resíduos continuou similar ($p\text{-value} < 0,005$). Assim, optamos em utilizar a metodologia *Bootstrap* não paramétrico por pares. Utilizando a macro do Anexo A, desenvolvida em Matlab, a variável Disponibilidade Física (%) foi considerada não significativa. A Tab. 2 apresenta os percentis 2,5% e 97,5% e mostra que o intervalo para variável Disponibilidade Física (%) contém o zero, indicando que existe evidências que seja nula ao nível de confiança 95% (β_2). Analisando as estatísticas dos coeficientes calculadas pelo software Minitab anteriormente, observa-se que a variável disponibilidade devido o p-valor do teste é 0,944, ao nível de significância de 5%, não possuía influência significativa no modelo.

O mesmo ocorrera para a variável produtividade efetiva, que apresentou o p-valor do teste de 0,081, que ao nível de significância de 5%, não possuía influência significativa no modelo, como mostra que o intervalo de confiança *Bootstrap* para variável não contém o zero, essa permanecerá no modelo.

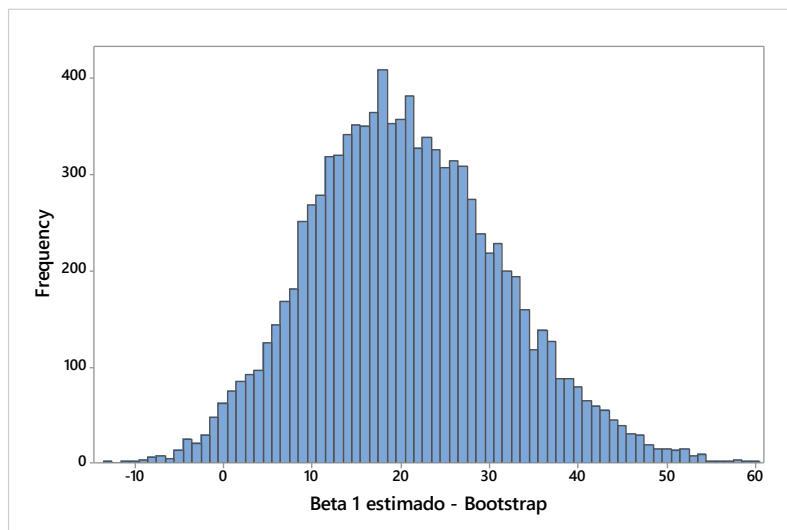
Tabela 2 – Limite Superior e Inferior intervalo de confiança *Bootstrap*

	Produtividade	Disponibilidade	Utilização	RO	DMT
Percentil 97,5%	43,40073	1707,617	-950,955	18980,05	2091,621
Percentil 2,5%	0,94649	-2599,38	-5542,49	12096,7	2008,21

Fonte: Saída do Matlab (2018)

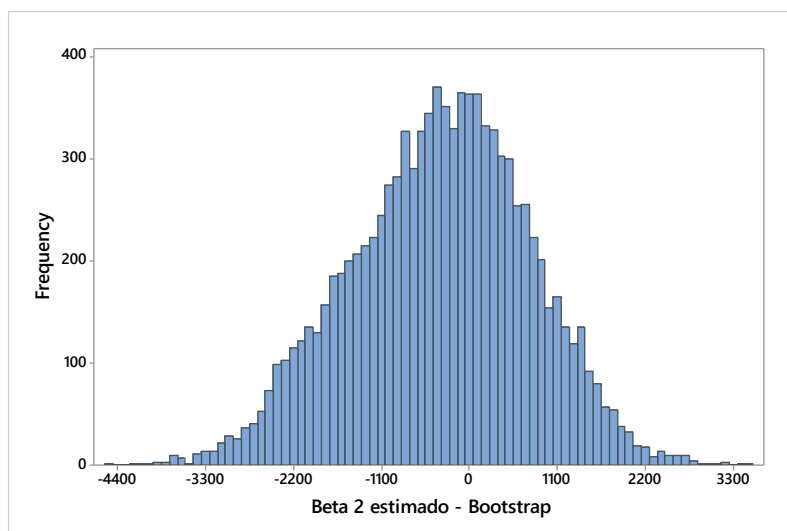
As Fig. 17-21 representam os histogramas dos estimadores.

Figura 17 – Histograma Produtividade estimado



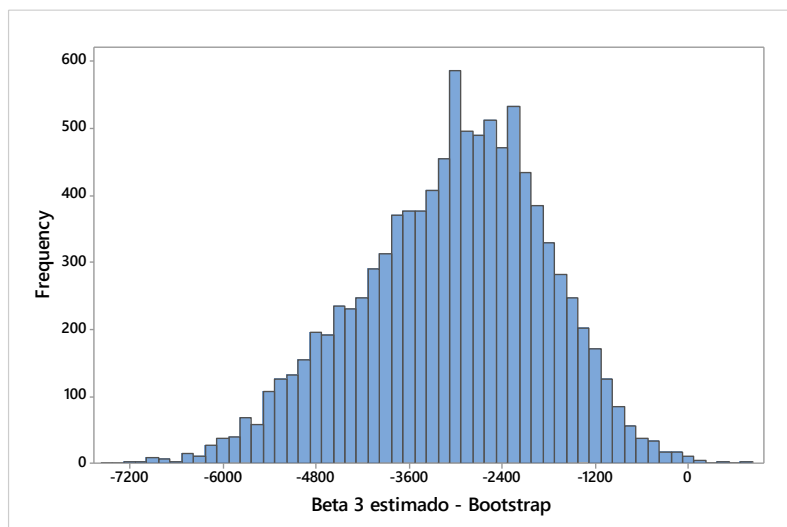
Fonte: Saída do Matlab (2018)

Figura 18 – Histograma Disponibilidade estimado



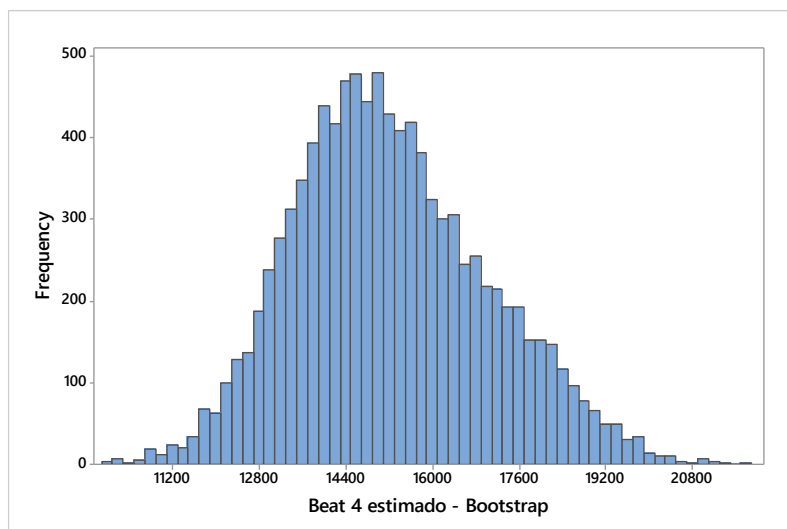
Fonte: Saída do Matlab (2018)

Figura 19 – Histograma Utilização estimado



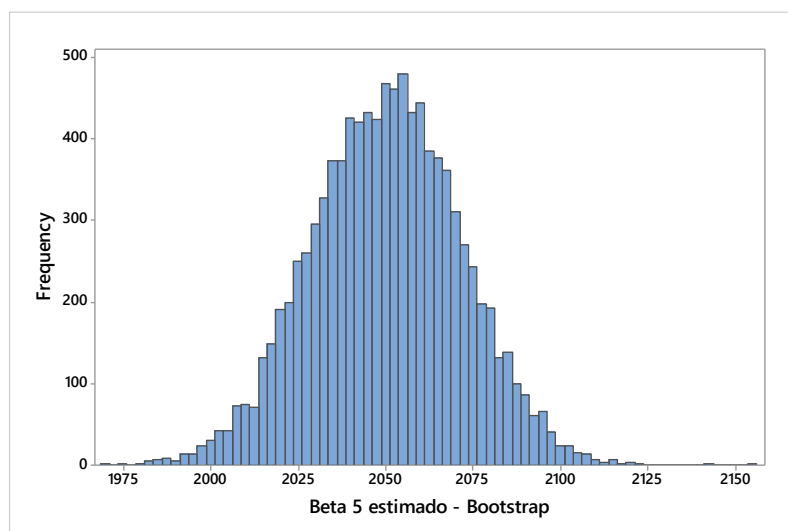
Fonte: Saída do Matlab (2018)

Figura 20 – Histograma RO estimado



Fonte: Saída do Matlab (2018)

Figura 21 – Histograma DMT estimado



Fonte: Saída do Matlab (2018)

Ao retirar a variável disponibilidade, obtém se um novo modelo descrito na Fig. 22-23.

Figura 22 – Modelo sem a variável Disponibilidade

Box-Cox transformation

Rounded λ	1
Estimated λ	1,03124
95% CI for λ	(0,994737; 1,06874)

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	4	23259773226	5814943307	8713,33	0,000
Produtividade Efetiva (t/h)	1	2617242	2617242	3,92	0,048
Utilização Efetiva	1	44875540	44875540	67,24	0,000
RO	1	535820510	535820510	802,89	0,000
DMT	1	9139894996	9139894996	13695,57	0,000
Error	954	636662872	667362		
Total	958	23896436099			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
816,922	97,34%	97,32%	97,31%

Coefficients

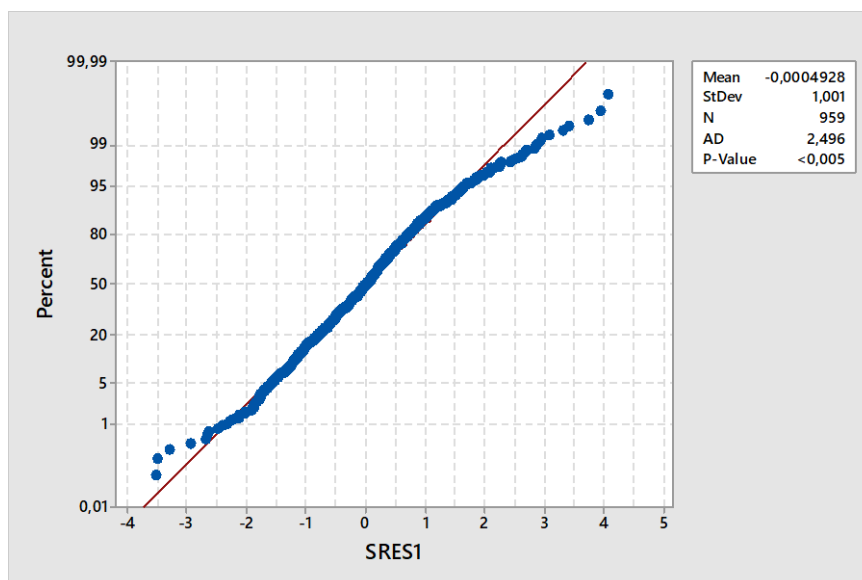
Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-327	267	-1,22	0,222	
Produtividade Efetiva (t/h)	17,76	8,97	1,98	0,048	1,24
Utilização Efetiva	-2632	321	-8,20	0,000	2,34
RO	14601	515	28,34	0,000	3,03
DMT	2052,6	17,5	117,03	0,000	1,86

Regression Equation

Massa (t) = -327 + 17,76 Produtividade Efetiva (t/h) - 2632 Utilização Efetiva + 14601 RO + 2052,6 DMT

Fonte: Saída do Minitab (2018)

Figura 23 – Teste de normalidade para os resíduos do modelo sem variável disponibilidade



Fonte: Saída do Minitab (2018)

Verificando o novo modelo que não possui mais a variável disponibilidade, o coeficiente de determinação permanece alto (97,31%) e não apresentando mudanças expressivas devido ao número de observações e variáveis do modelo. Reflexo, do elevadíssimo valor encontrado para o R^2 .

As variáveis DMT, produtividade e RO, aumentam a massa possível de ser transportada enquanto a variável utilização reduz. A variável RO possui maior impacto no modelo, uma vez que o coeficiente de determinação é alto (14601) em relação as demais variáveis.

A normalidade dos resíduos não apresentou mudança significativa, o modelo apresentado na Fig. 21-22 possui evidência de não normalidade dos resíduos. Assim, realizando os testes de significância dos parâmetros utilizando o procedimento *Bootstrap* obtém os limites de confiança descritos na Tab. 03.

Tabela 3 – Limite Superior e Inferior intervalo de confiança *Bootstrap* (Modelo 02)

	Produtividade	Utilização	RO	DMT
Percentil 97,5%	44,34246	-2063,218069	15563,32	2091,049
Percentil 2,5%	4,751279	-3280,621433	13763,5	2008,625

Fonte: Saída do Matlab (2018)

Logo observa que nenhum intervalo contém o valor zero indicando que as variáveis Produtividade Efetiva (t/h), Utilização Efetiva, RO e DMT são significativas (ao nível 5%) para

explicar a Massa (t). O intervalo de confiança para variável utilização sofreu alteração significativa em relação ao teste para o modelo com a variável disponibilidade, podendo este intervalo ter sido influenciado pela retirada da variável.

Neste sentido, o modelo final da equação de regressão ficou estabelecido como:

$$\text{Massa (t)} = -327 + 17,76 \text{ Produtividade Efetiva (t/h)} - 2632 \text{ Utilização Efetiva} + 14601 \text{ RO} + 2052,6 \text{ DMT} \quad (19)$$

Este modelo permite estimar a massa possível de ser transportada de produto de acordo com os indicadores operacionais, que no modelo são representados pelas variáveis. A Tab. 4 exemplifica três situações típicas em que é necessário estimar a Massa (t), quando os indicadores possuem os valores apresentados na tabela.

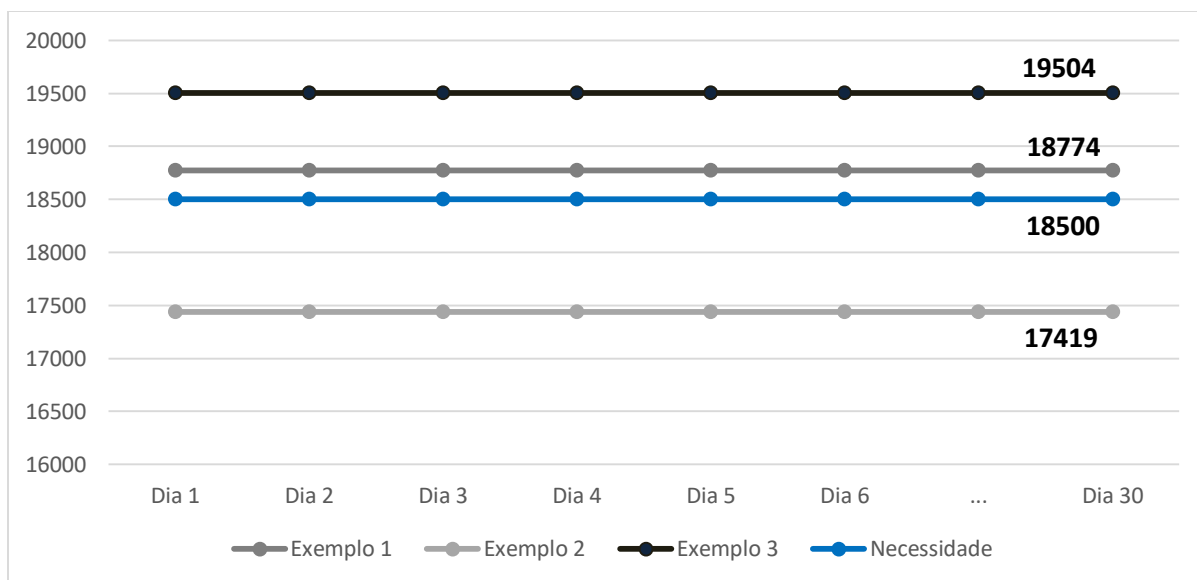
Tabela 4 – Capacidade de transporte estimada

Exemplo	Massa (t)	Produtividade efetiva (t/h)	Utilização efetiva	RO	DMT
1	18773,69	26,5	0,65	0,55	6
2	17418,85	25,0	0,60	0,45	6
3	19503,74	26,5	0,65	0,60	6

Fonte: Elaborado pela autora (2018)

Em um caso hipotético em que a necessidade de transporte seja 18500 t/dia, verificamos que os índices operacionais podem ainda ser alterados aumentando o RO no exemplo 1, mas não necessariamente tão elevado quanto no exemplo 3. Assim, a empresa consegue verificar qual índice operacional a alterar será a melhor estratégia de acordo com custo, gestão, mão de obra e outros fatores que impactam indiretamente em seus recursos. Na Fig. 24 observa que a massa necessária a ser transportada diariamente é de 18500t, portanto os valores que as variáveis devem assumir para melhor atender o dado proposto, são as do exemplo 1, que em comparação com o exemplo 3, ocorreu uma redução do rendimento operacional, concluindo a não necessidade de valores tão altos, ou seja, os recursos estariam sendo superestimados.

Figura 24 – Exemplos de valores estimados versus necessidade de transporte



Fonte: Elaborado pela autora (2018)

4 CONSIDERAÇÕES FINAIS

Os modelos de regressão são comumente utilizados para associar variáveis o qual explica a influência de cada variável nessa associação e para prever resultados. Nesse trabalho realizado demonstra o quanto útil é análise de regressão linear para resolver as questões cotidianas que necessitam de uma modelagem.

No estudo realizado o qual o objetivo era elaborar um modelo de previsão da massa possível a ser transportada entre duas unidades, obteve o modelo a ser utilizado para prever a massa de produto possível a transportar com caminhões.

O coeficiente de determinação ajustado o qual a proporção da variável resposta é explicada pelas variáveis preditoras é elevado, esse apresentou um resultado de 97,32%.

A princípio os testes de hipóteses ficam comprometidos pois existe forte evidência para rejeitar que os resíduos são normais ($p\text{-value} < 0,005$). A metodologia *Bootstrap* não paramétrico por pares utilizada para testar a significância dos estimadores, descartou uma variável e foi possível verificar que as demais são significantes.

Abordar todas as variáveis que devem ser levadas em consideração para o cálculo da movimentação de massa possível, contribuirá para um planejamento mais eficaz de vendas e uma melhor distribuição e cálculo de necessidade de frota para execução da atividade de transporte de minério.

Um estudo semelhante a este poderá ser realizado com os equipamentos de carga, verificando a influência de cada indicador operacional que serão denominados como variáveis preditoras, analisando as possibilidades de desenvolvimento de melhorias no processo.

ANEXO A: MACRO MATLAB PARA *BOOTSTRAP* POR PARES

```

clear all
%Testando H0: B=0 versus H1:B~0 por Bootstrap não paramétrico
(inclinações)

load Dados %Carrega os dados obtidos. Última Coluna é o Y.

corridas=10000;
Dados=[ones(959,1) Dados]; %Colocando Vetor de 1's na Matriz XX. Demanda do
Matlab.

R=regress(Dados(:,7),Dados(:,1:6)); %Estimativas Pontuais

for i=1:corridas
B=randsample([1:959]',959,true);
DD=[];
for z=1:959
z1=B(z,1);
DD=[DD;Dados(z1,:)];
end

RR=regress(DD(:,7),DD(:,1:6)); %Estimativas Pontuais via Bootstrap por
pares

Resultado(i,1)=RR(2,1);
Resultado(i,2)=RR(3,1);
Resultado(i,3)=RR(4,1);
Resultado(i,4)=RR(5,1);
Resultado(i,5)=RR(6,1);

end

LS = prctile(Resultado,97.5)
LI = prctile(Resultado,2.5)

```

REFERÊNCIAS BIBLIOGRÁFICAS

- CONT, V. D. **Francis Galton: eugenia e hereditariedade**. Scientiae Studia, SciELO Brasil, v. 6, n. 2, p. 201–218, 2008. Disponível em: <<http://www.scielo.br/pdf/ss/v6n2/04.pdf>>. Acesso em: 01 jun 2018.
- CLARKE, M.P. DENBY, B. SCHOFIELD, D. **Decision making tools for surface mine equipment selection**. Mining Science and Technology, n.10. 1990. Pg 323-335.
- EFRON, B. (1979). **Bootstrap Methods: Another Look at the Jackknife**. The Annals of Statistics. 7 (1): 1–26
- EFRON, B. & HASTIE, T. (2016). **Computer Age Statistical Inference: Algorithms, Evidence, and Data Science**. New York, NY: Cambridge University Press.
- IBRAM **Relatório Anual de Atividades Jun/16 a Jun/17**, Brasília: 2017
- GUJARATI, D. N.; PORTER, D. C. **Econometria básica**. 5. ed. Porto Alegre: AMGH, 2011. 924 p.
- 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**. Ouro Preto: Escola de Minas da Universidade Federal de Ouro Preto. Dissertação de Mestrado. Programa de Pós-Graduação em Engenharia Mineral. 2010. 105p.
- LUDKE, M. **Pesquisa em educação: abordagens qualitativas**. São Paulo: EPU, 1986. 96p.
- MLODINOW, L. **O andar do bêbado: como o acaso determina nossas vidas**. Rio de Janeiro: Zahar, 2009. Tradução: Diego Alfaro.
- MONTGOMERY, Douglas C. **Estatística Aplicada e Probabilidade para Engenheiros**. Tradução e revisão técnica Verônica Calado. Rio de Janeiro, LTC, 2015.
- RODRIGUES, L. F. **Análise comparativa de metodologias utilizadas no despacho de caminhões em minas a céu aberto**. Universidade Federal de Minas Gerais. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia de Produção. 2006. 87p.
- SILVA, V. C. **Apostila de Carregamento e transporte de rochas**. Escola de Minas da Universidade Federal de Ouro Preto. Ouro Preto. 2009. Pg. 32-89
- TOMPKINS, J. A.; WHITE, J. A.; BOZER, Y. A.; FRAZELLE, E. H.; TANCHOCO, J. M. A.; TREVINO, J. **Facilities planning**. 2ª ed., John Wiley. New York, 1996.