

# CONSTRUÇÃO DA MATRIZ ORIGEM-DESTINO DE TRANSPORTE INTER-REGIONAL DE CARGAS E PASSAGEIROS PARA O PLANO NACIONAL DE LOGÍSTICA INTEGRADA

## Relatório de Pesquisa

Modelos de Distribuição para Matriz Origem-Destino de Transporte de Cargas Inter-regional: desenvolvimento de um conjunto de ferramentas e calibração inicial (Relatório 4)





# **CONSTRUÇÃO DA MATRIZ ORIGEM-DESTINO DE TRANSPORTE INTER-REGIONAL DE CARGAS E PASSAGEIROS PARA O PLANO NACIONAL DE LOGÍSTICA INTEGRADA**

## **Relatório de Pesquisa**

**Modelos de Distribuição para Matriz Origem-Destino de Transporte de Cargas Inter-regional: desenvolvimento de um conjunto de ferramentas e calibração inicial (Relatório 4)**

**ipea**

## **Governo Federal**

### **Ministério do Planejamento, Desenvolvimento e Gestão**

**Ministro interino** Dyogo Henrique de Oliveira

## **ipea** Instituto de Pesquisa Econômica Aplicada

Fundação pública vinculada ao Ministério do Planejamento, Desenvolvimento e Gestão, o Ipea fornece suporte técnico e institucional às ações governamentais – possibilitando a formulação de inúmeras políticas públicas e programas de desenvolvimento brasileiro – e disponibiliza, para a sociedade, pesquisas e estudos realizados por seus técnicos.

#### **Presidente**

Ernesto Lozardo

#### **Diretor de Desenvolvimento Institucional, Substituto**

Carlos Roberto Paiva da Silva

#### **Diretor de Estudos e Políticas do Estado, das Instituições e da Democracia**

João Alberto De Negri

#### **Diretor de Estudos e Políticas Macroeconômicas**

José Ronaldo de Castro Souza Júnior

#### **Diretor de Estudos e Políticas Regionais, Urbanas e Ambientais**

Alexandre Xavier Ywata de Carvalho

#### **Diretora de Estudos e Políticas Setoriais de Inovação e Infraestrutura**

Fernanda De Negri

#### **Diretora de Estudos e Políticas Sociais**

Lenita Maria Turchi

#### **Diretor de Estudos e Relações Econômicas e Políticas Internacionais**

Sérgio Augusto de Abreu e Lima Florêncio Sobrinho

#### **Assessora-chefe de Imprensa e Comunicação**

Regina Alvarez

Ouvidoria: <http://www.ipea.gov.br/ouvidoria>

URL: <http://www.ipea.gov.br>

# **CONSTRUÇÃO DA MATRIZ ORIGEM-DESTINO DE TRANSPORTE INTER-REGIONAL DE CARGAS E PASSAGEIROS PARA O PLANO NACIONAL DE LOGÍSTICA INTEGRADA**

## **Relatório de Pesquisa**

**Modelos de Distribuição para Matriz Origem-Destino de Transporte de Cargas Inter-regional: desenvolvimento de um conjunto de ferramentas e calibração inicial (Relatório 4)**

**ipea**

Rio de Janeiro, 2017

## **EQUIPE TÉCNICA**

### **Coordenação-geral**

Fabiano Mezadre Pompermayer – Diretoria de Estudos e Políticas Setoriais de Inovação, Regulação e Infraestrutura (Diset) do Ipea

### **Equipe de pesquisa**

Fabiano Mezadre Pompermayer – Diset/Ipea

Erivelton Pires Guedes – Assessoria Técnica da Presidência (Astep) do Ipea

Akina Sakamori – pesquisadora do Programa de Pesquisa para o Desenvolvimento Nacional (PNPD) na Diset/Ipea

Alan Ricardo da Silva – pesquisador do PNPD na Diset/Ipea

Carolina Andrade Silva – pesquisadora do PNPD na Diset/Ipea

Daniel Alisson Feitosa Lopes – pesquisador do PNPD na Diset/Ipea

Diego Rosa Mambrin – pesquisador do PNPD na Diset/Ipea

Gabriel Gouveia Rabello – pesquisador do PNPD na Diset/Ipea

João Gabriel de Moraes Souza – pesquisador do PNPD na Diset/Ipea

Maircon Batista Ribeiro – pesquisador do PNPD na Diset/Ipea

Paulo Henrique Dourado da Silva – pesquisador do PNPD na Diset/Ipea

Pedro Veiga de Camargo – pesquisador do PNPD na Diset/Ipea

Priscila Nascimento de Alcântara Garcia – pesquisadora do PNPD na Diset/Ipea

Rafaella Bandeira Cabral Cunha – pesquisadora do PNPD na Diset/Ipea

Ramon de Almeida Bispo – pesquisador do PNPD na Diset/Ipea

Raquel Araújo de Almeida – pesquisadora do PNPD na Diset/Ipea

Rennaly Patrício Sousa – pesquisadora do PNPD na Diset/Ipea

Roberto Lazarte Kaqui – pesquisador do PNPD na Diset/Ipea

Thiago Guimarães Rodrigues – pesquisador do PNPD na Diset/Ipea

### **Elaboração do relatório**

Pedro Veiga de Camargo – pesquisador do PNPD na Diset/Ipea

Fabiano Mezadre Pompermayer – Diset/Ipea

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do Instituto de Pesquisa Econômica Aplicada ou do Ministério do Planejamento, Desenvolvimento e Gestão.

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

## SUMÁRIO

APRESENTAÇÃO .....	7
1 INTRODUÇÃO .....	9
2 IMPLEMENTAÇÃO COMPUTACIONAL.....	9
3 ESTRUTURA DE DADOS E INSTRUÇÕES GERAIS DE USO .....	11
4 RESULTADOS DA CALIBRAÇÃO DOS MODELOS .....	15
REFERÊNCIAS .....	17
BIBLIOGRAFIA COMPLEMENTAR .....	17
APÊNDICE.....	18





## APRESENTAÇÃO

Este relatório faz parte de uma série de relatórios de pesquisa do Ipea, cujo objetivo é documentar o processo de construção das matrizes origem/destino (O/D) de transporte inter-regional de cargas e passageiros que subsidiaram a elaboração do Plano Nacional de Logística Integrada (PNLI), fruto da parceria entre a Empresa de Planejamento e Logística (EPL), empresa pública do governo federal, e o Ipea, formalizada pelo Termo de Cooperação para Descentralização de Crédito nº 2, de 29 de maio de 2013.

A EPL foi criada em 2012 para apoiar o planejamento da infraestrutura de transportes do país. Uma de suas tarefas é elaborar o Plano Nacional de Logística Integrada. Em certa medida, ela cumpre missão semelhante à do Grupo de Estudos para Integração da Política de Transportes (Geipot) nas décadas de 1970 e 1980. A empresa vem se capacitando para a modelagem da oferta de infraestrutura de transportes, mas para a modelagem da demanda por transportes, que envolve mais conhecimento sobre as atividades econômicas no país e sua distribuição regional do que sobre transportes, havia a necessidade de execução externa. Algumas consultorias foram abordadas, mas surgiu a possibilidade de o estudo ser realizado pelo Ipea, que prontamente atendeu.

Nesse processo, um dos passos iniciais é o levantamento do padrão de viagens inter-regionais de cargas e passageiros, identificando os potenciais de geração e atração de viagens de cada localidade, bem como sua distribuição entre elas. Os modos de transporte usados também são levantados. Assim obtém-se a matriz OD das viagens de carga e passageiro para um ano-base, em cada modo. A partir daí, são projetados os crescimentos de demanda por transporte para os anos futuros, isto é, as matrizes OD futuras, que são confrontadas com a oferta de infraestrutura a fim de identificar seus gargalos e projetar a necessidade de melhorias, como construção ou ampliação de rodovias, ferrovias, hidrovias e portos. A construção da matriz OD é essencial para o planejamento da infraestrutura de transportes do país, uma vez que permite a construção de cenários e a otimização dos investimentos.

A elaboração de uma matriz OD envolve o levantamento e o cruzamento de uma série de dados socioeconômicos, que são ajustados e calibrados com pesquisas de campo, em que viajantes são entrevistados em locais predeterminados para identificar diversos atributos de sua viagem, como origem, destino, motivo, tipo e valor da carga, modo(s) de transporte utilizado(s) etc. Esse trabalho envolve conhecimento técnico específico em modelagem de transportes e também em economia e estatística.

Com uma matriz OD consolidada e abrangente, a EPL poderá realizar suas avaliações de necessidade de expansão da oferta de infraestrutura de transporte. O Ipea, além de apoiar a EPL nestas avaliações, se beneficia das informações da matriz OD ao cruzá-las com outras fontes de dados socioeconômicos, podendo realizar estudos diversos sobre desenvolvimento regional, matrizes insumo-produto e fomento produtivo, por exemplo. Adicionalmente, a disponibilização dessa base de dados à sociedade permite o desenvolvimento do próprio setor de transporte e logística.

Oito relatórios descrevem as metodologias usadas em todo o processo de construção da base de dados e suas projeções. O último desses relatórios reúne os procedimentos intermediários para a construção das matrizes e apresenta alguns dos resultados agregados. Há, ainda, três relatórios que avaliaram o estado da arte de modelos de construção da

matriz OD, com proposições para as novas versões do PNLI. A seguir, a lista com os títulos de cada relatório.

- 1) *Desenho da Pesquisa Origem-Destino do Transporte Rodoviário no Brasil.*
- 2) *Construção da Matriz Origem-Destino Observada para Transporte de Cargas Inter-regional.*
- 3) *Modelos de Regressão para Geração e Atração de Viagens do Transporte de Cargas Inter-regional.*
- 4) *Modelos de Distribuição para Matriz Origem-Destino de Transporte de Cargas Inter-regional: desenvolvimento de um conjunto de ferramentas e calibração inicial.*
- 5) *Cenários de Projeção das Atividades Econômicas por Unidade da Federação para a Projeção da Matriz Origem-Destino de Transporte de Cargas Inter-regional.*
- 6) *Construção da Matriz Origem-Destino Observada de Transporte de Passageiros Inter-regional.*
- 7) *Modelos de Regressão para Geração e Atração de Viagens de Passageiros.*
- 8) *Projeções das Matrizes Origem-Destino de Carga e Passageiros: aplicando os modelos de geração e distribuição de viagens com as projeções socioeconômicas.*
- 9) *Estado da Arte em Métodos de Construção de Matrizes Origem-Destino para o Transporte de Cargas Inter-regional.*
- 10) *Estado da Arte de Métodos de Estimação de Matrizes Origem-Destino para Passageiros à Longa Distância.*
- 11) *Métodos de Estimação da Matriz Origem-Destino para o Transporte Aéreo de Passageiros.*

## 1 INTRODUÇÃO

Este relatório apresenta o processo de desenvolvimento de um conjunto de ferramentas computacionais para a calibração de modelos de distribuição de viagens do tipo gravitacional e dos resultados preliminares dessa calibração para a atualização do Plano Nacional de Logística Integrada (PNLI), conduzido pela Empresa de Planejamento e Logística (EPL) e pelo Ipea.

Além de uma breve explicação dos princípios e algoritmos utilizados na implementação da ferramenta computacional e dos parâmetros que podem ser alterados pelo usuário durante a sua utilização, também serão apresentados os resultados da calibração de tais modelos para as matrizes origem-destino (OD) do PNLI, os quais poderão ser atualizados com matrizes de impedância específicas para cada produto analisado, a serem fornecidas pela EPL. Este relatório não pretende, entretanto, descrever em detalhe o código computacional desenvolvido, mas apenas permitir ao usuário que o utilize de forma apropriada. Tanto o código quanto a estrutura de diretórios e alguns arquivos de exemplos serão disponibilizados pelos autores no sítio eletrônico do Ipea.

Modelos de distribuição do tipo gravitacional têm sido formalmente utilizados desde meados da década de 1950 com a publicação do trabalho de Casey (1955), como é bem apontado por Ortúzar e Willumsen (2011). Foi apenas na década de 1970 que Wilson (1970; 1974) apresentou uma derivação teórica formal para tal modelo com base no princípio de maximização de entropia. A longevidade dessa técnica, no entanto, não advém da sua derivação teórica, mas da facilidade de sua utilização e da qualidade de seus resultados.

## 2 IMPLEMENTAÇÃO COMPUTACIONAL

Apesar de modelos gravitacionais sintéticos serem utilizados há várias décadas, e estarem disponíveis virtualmente em todos os pacotes computacionais de planejamento de transportes à venda no mercado, nenhum desses pacotes estava disponível para calibração e aplicação de tais modelos de distribuição durante o desenvolvimento dos modelos. Procedeu-se, então, à implementação dos procedimentos de calibração e à aplicação de modelos gravitacionais em linguagem computacional.

A linguagem computacional escolhida para essa atividade foi Python (Van Rossum e Drake, 2011), devido a alguns fatores principais. O primeiro deles é o caráter gratuito dessa linguagem e da ampla disponibilidade de bibliotecas para computação científica, particularmente NumPy (Van Der Walt, Colbert e Varoquaux, 2011) e Matplotlib (Hunter, 2007).

O segundo fator determinante foi a facilidade de implementação e portabilidade do código, o que permitiu a sua transferência para o Ipea com relativa facilidade. O terceiro e último fator foi a familiaridade do programador com essa linguagem, o que permitiu a finalização do código em um intervalo de tempo relativamente curto.

## 2.1 Modelo matemático

Das diversas formulações utilizadas para modelos de distribuição, decidiu-se pela implementação do modelo gravitacional, com forma funcional explícita na equação (1).

$$T_{ij} = \alpha O_i D_j f(c_{ij}) \quad (1)$$

Onde  $T_{ij}$  representa a quantidade de viagens entre a zona de origem  $i$  e a de destino  $j$ ;  $O_i$  é o total de viagens geradas a partir da zona  $i$ ;  $D_j$  é o total de viagens atraídas para a zona  $j$ ;  $f(c_{ij})$  é a função de impedância entre cada par de zonas  $i-j$ , que tem como parâmetro-chave o custo de transporte  $c_{ij}$ , entre a zona  $i$  e a zona  $j$ ; e  $\alpha$  é um parâmetro a ser calibrado.

Das três formas funcionais mais utilizadas em planejamento de transporte de indivíduos, duas se aplicam bem ao caso do transporte de carga. São elas a função exponencial e a função potência negativa, representadas nas equações (2) e (3).

$$f(c_{ij}) = \exp(-\beta c_{ij}) \quad (2)$$

$$f(c_{ij}) = c_{ij}^{-n} \quad (3)$$

A escolha de tais formas funcionais, ainda que arbitrária, foi pautada pela inexistência de evidência prática ou teórica contundente de que a utilização de formas funcionais alternativas ou mesmo de fatores de fricção poderiam oferecer quaisquer benefícios. A existência do algoritmo desenvolvido por Hyman (1969), para a calibração de funções com as formas funcionais propostas, reforça tais escolhas. Os parâmetros  $\beta$  e  $n$  precisam ser calibrados com base nas matrizes OD observadas.

Em linhas gerais, o modelo gravitacional preconiza que a quantidade de viagens para um par de zonas  $i-j$  é diretamente proporcional ao total de viagens originadas em  $i$ , ao total de viagens destinadas a  $j$ , e inversamente proporcional à impedância entre as duas zonas  $i$  e  $j$ . O sinal negativo de  $\beta$  e  $n$  nas equações (2) e (3) indica a proporcionalidade de  $T_{ij}$  inversa à impedância.

## 2.2 Requisitos para funcionamento em série

Uma vez que a utilização do *software* se dará com a calibração de modelos para múltiplos produtos em cada nova atualização do modelo, e que a aplicação de tais modelos resultará em um número ainda maior de matrizes projetadas, busca-se definir uma estrutura de dados que permita o fornecimento de grandes volumes de dados para o *software*, tanto na fase de calibração quanto na fase de aplicação. Além disso, a EPL solicitou que as matrizes resultantes fossem apresentadas em um único arquivo, de forma a facilitar sua importação por outros *softwares*.

### 2.3 Software necessário e configurações mínimas

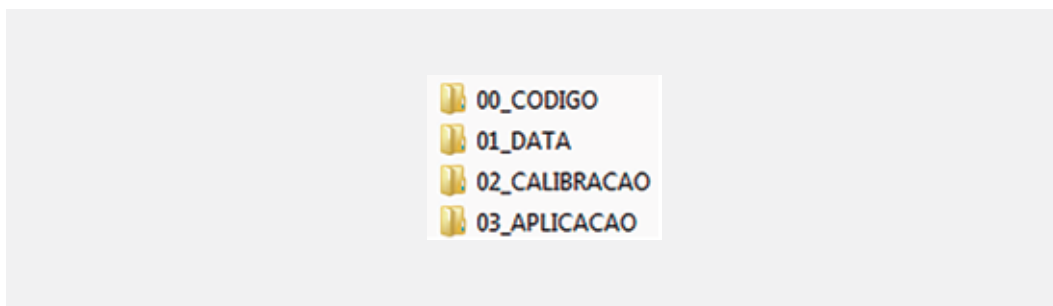
Os requisitos de *software* são bastante reduzidos, e se limitam a Python 2.7.x e às duas bibliotecas mais comumente usadas para computação científica em Python: NumPy (1.8 ou mais recente) e Matplotlib (1.4 ou mais recente).

O tempo de processamento será compatível com o desempenho do disco rígido, do processador e da memória do computador, sendo que o processo de calibração dos modelos deverá tomar de 50 Mb a 100 Mb da memória disponível. Para a aplicação do modelo de calibração, no entanto, a utilização de memória é proporcional ao número de matrizes sendo distribuídas na mesma rodada. Assim sendo, pode-se esperar uma utilização de memória superior a 1 Gb para o caso de um grande número de matrizes.

## 3 ESTRUTURA DE DADOS E INSTRUÇÕES GERAIS DE USO

A estrutura de diretórios concebida para a calibração e aplicação dos modelos gravitacionais, apresentada na figura 1, foi preparada para armazenar todos os dados necessários para se trabalhar com o modelo de calibração. Ela foi considerada fixa em todo o código e, portanto, não pode ser alterada sem que se altere também as linhas de código que se referem a tal estrutura.

FIGURA 1  
Estrutura de diretórios



Elaboração dos autores.

Obs.: figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

### 3.1 Calibração de modelos

O processo de calibração de modelos depende do programa denominado *Calibra\_gravitacional.py*, que se encontra na pasta 00\_CODIGO. Esse programa utiliza duas fontes de dados, que devem ser colocados na pasta 01\_DATA, e mais um parâmetro relacionado à forma funcional da calibração.

A primeira fonte de dados necessária é o arquivo com as matrizes dos produtos para os quais os modelos devem ser calibrados. Conforme mostrado na figura 2, esse arquivo deve ser do tipo texto separado por vírgulas (*comma separated values* – CSV), ter decimais separados por ponto e conter os seguintes campos, nessa ordem: produto, zona de origem, zona de destino e volume.

FIGURA 2  
Estrutura do arquivo com as matrizes de entrada para calibração

```
1 PRODUTO,Zonas_origem,Zonas_destino,Volumes
2 1,1,7,135.5912
```

Elaboração dos autores.

Obs.: figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

A segunda fonte de dados é a matriz de impedância a ser utilizada (a mesma matriz será utilizada para a calibração de todos os modelos), a qual também deve ser fornecida em formato CSV, ter decimais separados por ponto e conter os seguintes campos: zona de origem, zona de destino e impedância, conforme ilustrado na figura 3.

FIGURA 3  
Estrutura do arquivo com a matriz de impedância para calibração

```
1 zona_o,zona_d,distancia_OD
2 1,1,15.00000
```

Elaboração dos autores.

Obs.: figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

É importante ressaltar que a restrição ao uso de uma única matriz de impedância resume-se ao processo de calibração apenas. Caso o usuário queira calibrar modelos utilizando matrizes de impedância distintas para cada produto, basta calibrar tais modelos em rodadas distintas, substituindo a matriz de impedância adequadamente.

O parâmetro necessário consiste na definição da forma funcional que será calibrada (“EXPO” para a função exponencial, e “POWER” para a função de potência negativa). Tanto o parâmetro de definição do modelo quanto os nomes dos arquivos com as matrizes de entrada e de impedância podem ser alterados em qualquer editor de texto no corpo do código *Calibra\_gravitacional.py* nas linhas 25, 26 e 27, respectivamente, conforme ilustrado na figura 4.

FIGURA 4  
Parâmetros a serem alterados para calibração de modelos

```
25 function="EXPO" #"EXPO" ou "POWER"
26 input_matrix='input_matrix.csv' #Nome do arquivo com as matrizes dos produtos
27 cost='cost.csv' #Nome do arquivo com a matriz de impedância
```

Elaboração dos autores.

Obs.: figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Após a calibração dos modelos ser processada com sucesso, todos os resultados poderão ser encontrados no diretório 02\_CALIBRACAO. O relatório da calibração será encontrado no arquivo log.txt, e conterá todas as mensagens que foram exibidas para o usuário durante o processo de calibração, incluindo os valores dos parâmetros calibrados, e os resultados finais com os histogramas para cada um dos produtos. As matrizes geradas pela aplicação dos modelos calibrados poderão ser encontradas nas pastas EXPO ou POWER, dependendo da forma funcional escolhida para a calibração.

### 3.2 Aplicação de modelos

A aplicação de modelos depende de um número maior de parâmetros. Essa necessidade, que advém tanto do maior número de parâmetros para a geração de cada matriz quanto da existência de múltiplos cenários e horizontes de planejamento, exigiu uma estrutura de dados mais complexa em sua construção.

São quatro os arquivos de entrada necessários para a aplicação dos modelos, os quais se encontram em duas pastas distintas, de acordo com a sua natureza. O primeiro arquivo necessário é o que contém os vetores de produção e atração de cada matriz e que deve ser colocado no diretório 01\_DATA. A estrutura de dados desse arquivo é variável de acordo com o número de matrizes sendo geradas, e contém um campo para produção e um para atração de cada produto, além dos campos de produto e zona, os quais se aplicam para todos os vetores, conforme ilustrado na figura 5. Assim como os arquivos anteriores, os vetores também devem ser fornecidos em formato CSV e ter decimais separados por ponto. Cada um deles é fornecido em uma coluna, cujo nome indica se é origem ou destino, e o ano/cenário. Por exemplo, O\_2\_medio indica o vetor de geração (origem) no ano 2 no cenário de médio crescimento.

FIGURA 5

#### Estrutura de dados para arquivo com vetores de produção e atração

1	Produto,Zona,O_2_pessimista,O_2_medio,O_2_otimista,O_7_pessimista,O_7_medio,O_7_otimista,O_12_pessimista,...
2	6,1,5834.41,4134.70,4239.32,5997.56,7071.87,7511.40,4147.01,8014.52,8901.29,4269.98,8946.32,10422.13,4385.34,9938.93,...
3	6,2,3878.12,4022.92,4104.45,3779.25,4282.27,4597.34,3491.71,4543.13,5135.87,3614.15,4806.52,5725.14,3544.23,5075.67,...

Elaboração dos autores.

Obs.: figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

É preciso frisar, no entanto, que os vetores para diferentes produtos, mas para o mesmo ano/cenário, devem ser apresentados empilhados, e títulos repetidos para vetores de produção e atração não são permitidos.

O segundo arquivo necessário para a aplicação dos modelos é o arquivo de impedância, que também deve ser inserido no diretório 01\_DATA. A estrutura e o formato desse arquivo são os mesmos descritos para o caso da calibração de modelos e ilustrados na figura 3. No caso da aplicação, múltiplos arquivos contendo diferentes matrizes de impedância (mas com a mesma estrutura e no mesmo diretório) podem ser utilizados, conforme será descrito a seguir.

O arquivo de controle da aplicação de modelos, que deve ser colocado no diretório 03\_APLICACAO, é aquele que garante a correspondência entre os vetores de produção e atração e a matriz de impedância adequada, bem como aquele que define o nome do campo no qual a matriz resultante será escrita, o que por sua vez garante a correta importação desta quando utilizada em outros pacotes computacionais. Mais uma vez, o arquivo deve ser fornecido no formato CSV, com uma entrada para cada matriz a ser gerada e campos para o produto, vetor de produção, vetor de atração, arquivo de impedância, ano do vetor, cenário de crescimento e nome da matriz resultante, conforme apresentado na figura 6.

FIGURA 6  
Estrutura de dados para arquivo de controle da aplicação de modelos

```

1 Índice de produto,Producao,Atracao,Arquivo de impedancia,Ano,Cenario,NomeColunaMatriz
2 1,O_2_pessimista,D_2_pessimista,cost2002.csv,2002,PESS,1_2002_PESS
3 2,O_2_pessimista,D_2_pessimista,cost2002.csv,2002,PESS,2_2002_PESS
4 3,O_2_pessimista,D_2_pessimista,cost2002.csv,2002,PESS,3_2002_PESS
5 5,O_2_pessimista,D_2_pessimista,cost2002.csv,2002,PESS,5_2002_PESS
6 1,O_2_medio,D_2_medio,cost2002.csv,2002,MED,1_2002_MED

```

Elaboração dos autores.

Obs.: figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Vale a pena frisar que o ano horizonte, o cenário de crescimento e o nome da matriz resultante não têm formato obrigatório e são apenas para referência do próprio analista, enquanto os demais campos têm formatos e valores que devem ser consistentes com outras fontes de dados utilizadas no modelo. Os códigos dos produtos devem ser os mesmos utilizados no arquivo de vetores no campo de produto, e os nomes dos vetores de produção e atração devem, por sua vez, coincidir com os nomes dos campos de produção e atração fornecidos no arquivo de vetores. Os nomes das matrizes de impedância também devem ser os mesmos existentes no diretório 01\_DATA, e a falha em garantir consistência para qualquer uma dessas informações resultará em erro no programa. Além disso, os vetores de produção e atração escolhidos em cada aplicação devem estar balanceados, isto é, a soma de viagens produzidas deve ser igual a de viagens atraídas, para cada produto.

O último arquivo necessário para a aplicação dos modelos é aquele que contém a informação da forma funcional e o valor de parâmetro para cada um dos produtos analisados. A estrutura desse arquivo é extremamente simples e consiste em um campo para o código do produto, um para a forma funcional a ser utilizada (“EXPO” para a função exponencial, e “POWER” para a função de potência negativa) e outro para o valor do parâmetro, conforme ilustrado na figura 7. O arquivo deve ser do tipo CSV e ter decimais separados por ponto. Esse arquivo também deve ser colocado no diretório 03\_APLICACAO.

FIGURA 7  
Estrutura de dados para arquivo com parâmetros de modelos

```

1 Produto,Modelo,Beta
2 1,POWER,0.80024800
3 2,POWER,1.64946600
4 3,POWER,0.99068300
5 4,EXPO,0.00156100

```

Elaboração dos autores.

Obs.: figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Para se proceder à aplicação dos modelos, é necessário rodar o programa *Aplica\_Gravitacional.py* após a preparação dos quatro arquivos de entrada descritos e a alteração de seis parâmetros no texto do código (o que pode ser feito em qualquer editor de texto) entre as linhas 20 e 25, conforme ilustrado na figura 8.



FIGURA 8  
Parâmetros a serem alterados para aplicação de modelos

```
20 file_modelos='modelos.csv'  
21 instrucoes='controle_aplicacao_modelo.csv'  
22 vet_file='VETORES.csv'  
23 REMOVE_PORTOS=True  
24 REMOVE_INTRAZONAL=True  
25 MATRIZ_COMPLETA=False
```

Elaboração dos autores.  
Obs.: figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Os três primeiros parâmetros a serem modificados são os que correspondem aos arquivos de entrada: modelos da função de impedância (contendo a forma funcional e seu parâmetro); controle da aplicação; e vetores de produção e atração, respectivamente. Os três parâmetros complementares consistem em instruções específicas para o tratamento de fluxos relacionados a portos, fluxos intrazonais e formato das matrizes de saída.

Os valores possíveis para esses parâmetros são verdadeiro e falso (*True* ou *False*) e têm os seguintes significados:

- REMOVE\_PORTOS=True: para zerar os fluxos da matriz para o exterior, zonas 570 a 577;
- REMOVE\_PORTOS=False: para manter os fluxos para o exterior;
- REMOVE\_INTRAZONAL=True: para zerar o fluxo intrazonal;
- REMOVE\_INTRAZONAL=False: para manter o fluxo intrazonal;
- MATRIZ\_COMPLETA=True: para escrever todas as células da matriz, incluindo zeros); e
- MATRIZ\_COMPLETA=False: para escrever somente células diferentes de zero.

Três resultados são escritos no diretório 03\_APLICACAO com os resultados da rodada. O primeiro deles é o arquivo log.txt, que contém todas as mensagens mostradas ao usuário na tela. O segundo se chama results.csv e contém a mesma informação do arquivo log.txt, embora sua informação seja em formato de banco de dados e possa ser analisada mais facilmente quando se aplicam modelos para um grande número de matrizes. Essas informações incluem produto, ano, cenário, forma funcional e parâmetro do modelo, além do total da matriz distribuída, do nível de convergência obtido, do número de iterações realizadas e de eventuais erros encontrados. O último arquivo gerado é aquele que contém as matrizes distribuídas propriamente ditas e se chama matrizes\_finais.csv.

## 4 RESULTADOS DA CALIBRAÇÃO DOS MODELOS

Os resultados de calibração dos modelos apresentados neste relatório servem para demonstrar o uso da ferramenta e o processo de calibração. Os resultados são apresentados na tabela 1 e, apesar de ser possível analisar cada um dos resultados, é mais interessante avaliá-los de forma comparativa.

TABELA 1  
Parâmetros dos modelos calibrados para a matriz final

Produto	Modelo		Nome do produto	Classe de custo
	Função	Parâmetro		
1	POWER	0.48107562	Alimentos e bebidas (processados)	Carga geral
2	POWER	0.19424396	Carvão mineral	Granel sólido
3	POWER	0.32130040	Celulose e papel	Carga geral
4	EXPO	0.00234831	Cimento	Carga geral
5	EXPO	0.00079459	Combustíveis	Combustíveis
6	EXPO	0.00084011	Farelo de soja	Granel sólido
7	POWER	0.29049531	Manufaturados	Carga geral
8	POWER	0.62232454	Milho em grão	Granel sólido
9	POWER	0.39044392	Minério de ferro	Granel sólido
10	POWER	0.70148905	Outros da lavoura e pecuária	Carga geral
11	EXPO	0.00134135	Outros minerais	Granel sólido
12	EXPO	0.00036783	Petroquímicos e químicos	Carga geral
13	POWER	0.45268880	Produtos básicos de borracha, plástico e não metálicos	Carga geral
14	POWER	0.69319770	Produtos da exploração florestal e da silvicultura	Carga geral
15	EXPO	0.00132825	Soja em grão	Granel sólido

Elaboração dos autores.

Em tempo, o processo de avaliação utilizado para a escolha do modelo final para cada produto é puramente qualitativo e pode ser resumido como a avaliação dos histogramas de viagem pelo analista responsável pela seleção do modelo. Em certos casos, como o do produto 1, alimentos e bebidas (processados), o modelo de potência inversa apresenta uma similaridade muito maior em relação ao histograma da matriz observada do que o modelo exponencial, o que torna a escolha desse modelo a mais coerente.

Em alguns casos, como com carvão mineral e celulose e papel, as diferenças entre as aderências de ambos os modelos são mínimas, o que faz a escolha mais difícil, mas também menos relevante, uma vez que eles apresentam desempenhos similares. Como regra geral, o modelo de potência negativa foi preferido em relação ao exponencial, mesmo que aparentemente aquele tenha a tendência de superestimar o número de viagens de curta distância (histogramas de celulose e papel e cimento, na tabela 1, por exemplo). Considerando a localização dos postos de coleta da pesquisa de OD nas rodovias, que priorizaram os segmentos rodoviários com grande volume de tráfego (em geral no entorno de grandes cidades), estima-se que muitas viagens de curta distância entre pares OD de zonas pequenas (cidades do interior em geral) não tenham sido devidamente amostradas. A escolha do modelo de potência negativa tenta suprir essa limitação do levantamento de campo realizado.

Analisando-se os resultados apresentados na tabela 1, é possível perceber que eles são razoáveis. Um dos exemplos dessa razoabilidade é o fato de que o parâmetro escolhido para o modelo para a matriz de produtos da exploração florestal e da silvicultura é cerca de o dobro do parâmetro calibrado para manufaturados, indicando viagens com distâncias maiores para produtos deste último setor.

Uma modificação importante na calibração para a matriz final foi o tratamento das distâncias entre cada par de zonas de OD, conforme o uso ou o não uso de modos de transporte de menor custo – ferroviário, aquaviário e dutoviário. Na classe carga geral, apenas o modo rodoviário é considerado, sendo mantidas as distâncias originais.

Para a classe granel sólido, as distâncias entre pares de zonas que compreendem rotas ferroviárias ou aquaviárias (e também minerodutos) foram reduzidas pela metade. E na classe de combustíveis, o mesmo procedimento foi adotado para os pares de zonas em que há oleodutos disponíveis.

As agregações dos produtos, conforme lista apresentada na tabela 1, foram necessárias devido à classificação não coerente de produtos na pesquisa de campo. As principais limitações para a alocação ocorrem devido ao açúcar estar incluso em alimentos e bebidas (processados), e os produtos siderúrgicos em manufaturados, o que limita uma adequada calibração dos custos de transporte ferroviário e aquaviário para estes produtos. Para o açúcar, o fato de a pesquisa ter ocorrido na entressafra agravou os resultados.

Esses resultados, aliados aos histogramas de distância de viagem, demonstram a adequabilidade do modelo gravitacional ao problema de distribuição de viagens proposto. Os histogramas para todos os parâmetros calibrados para a matriz final estão no apêndice deste relatório. Apesar de a calibração indicar boa aderência, a aplicação do modelo gravitacional a produtos mais homogêneos e de baixo valor agregado, como *commodities* agrícolas e minerais, pode não ser adequada. Para tais produtos, é mais comum o padrão de viagens representar o menor custo total de transporte na rede, isto é, o volume movimentado a partir de uma origem segue pela rota de menor custo, o que tende a concentrar os volumes em poucos pares OD. Já o modelo gravitacional tende a dispersar as viagens entre vários pares OD, ainda que concentrando os volumes nos pares com menor custo.

## REFERÊNCIAS

- CASEY, H. J. Applications to traffic engineering of the law of retail gravitation. **Traffic Quarterly**, v. 9, p. 23-35, 1955.
- HUNTER, J. D. Matplotlib: a 2D graphics environment. **Computing in Science and Engineering**, v. 9, n. 3, p. 90-95, 2007.
- HYMAN, G. M. The calibration of distribution models. **Environment and Planning**, v. 1, p. 105-112, 1969.
- ORTÚZAR, J. D.; WILLUMSEN, L. G. **Modelling transport**. London: John Wiley & Sons, 2011.
- VAN DER WALT, S.; COLBERT, S. C.; VAROQUAUX, G. The number py array: a structure for efficient numerical computation, computing in science and engineering. **Computing in Science and Engineering**, v. 13, n. 2, p. 22-30, 2011.
- VAN ROSSUM, G.; DRAKE, F. L. (Ed.). **Python reference manual**. Virginia: Python Labs, 2011.
- WILSON, A. G. **Entropy in urban and regional modelling**. London: Pion, 1970.
- \_\_\_\_\_. **Urban and regional models in geography and planning**. New York: John Wiley & Sons, 1974.

## BIBLIOGRAFIA COMPLEMENTAR

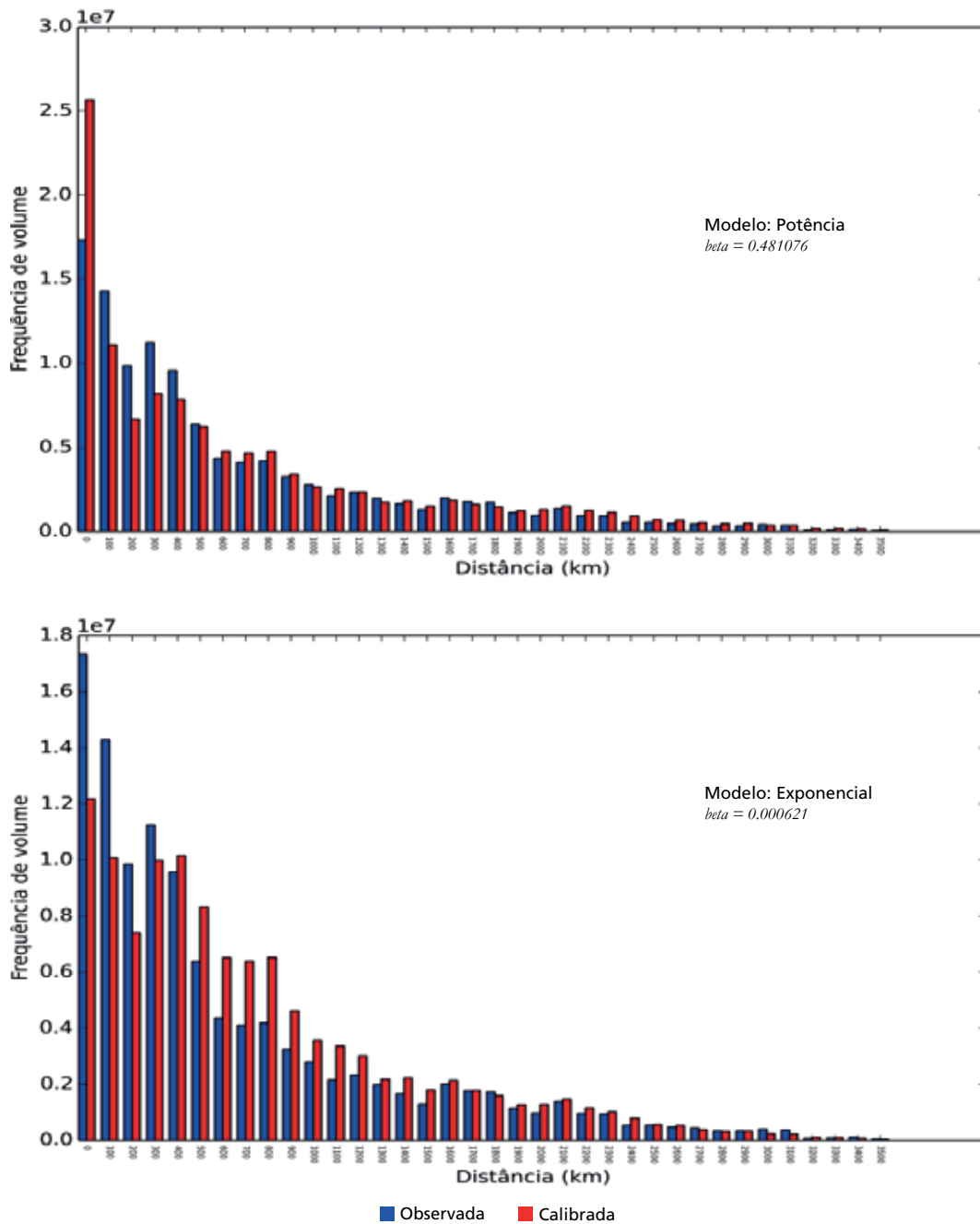
- WILLIAMS, I. A comparison of some calibration techniques for doubly constrained models with an exponential cost function. **Transportation Research**, n. 10, p. 1-104, 1976.

## APÊNDICE

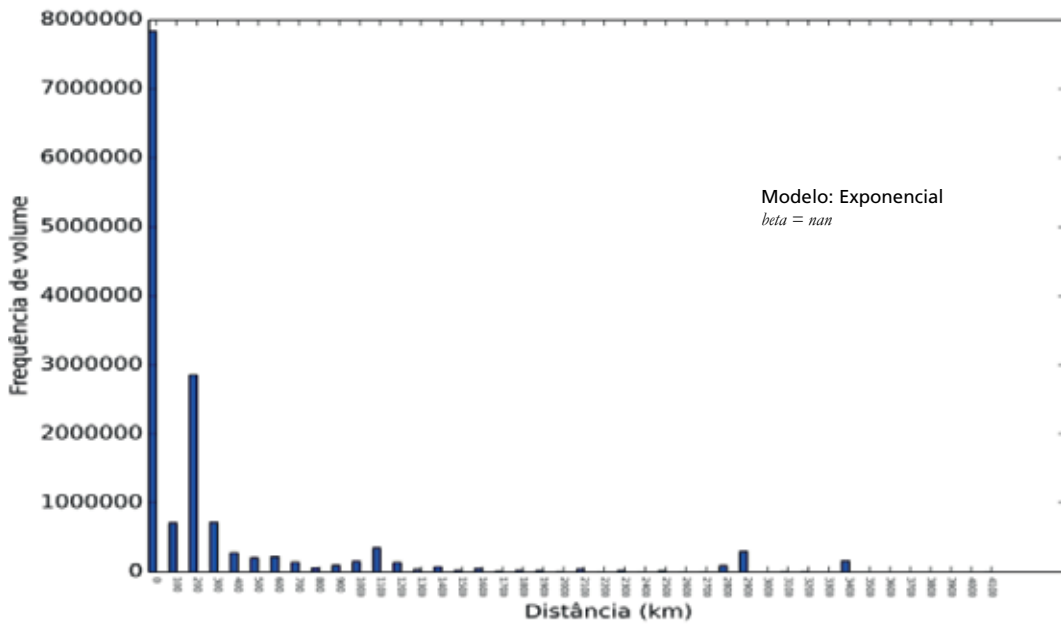
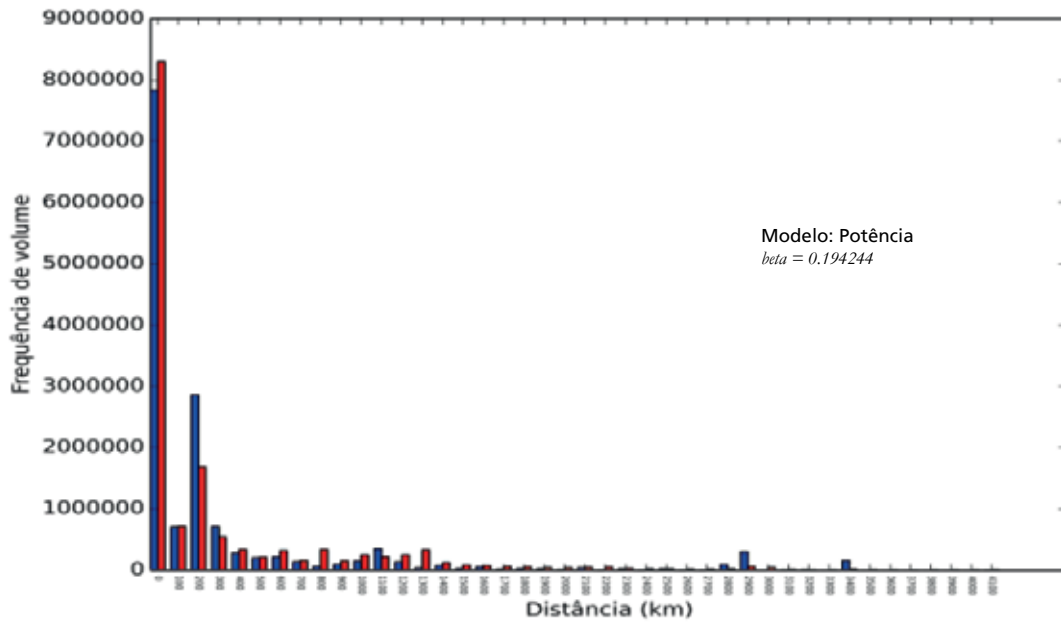
GRÁFICO A

Distribuição das distâncias de viagem por produto para a matriz final

A.1 – Alimentos e bebidas (processados)

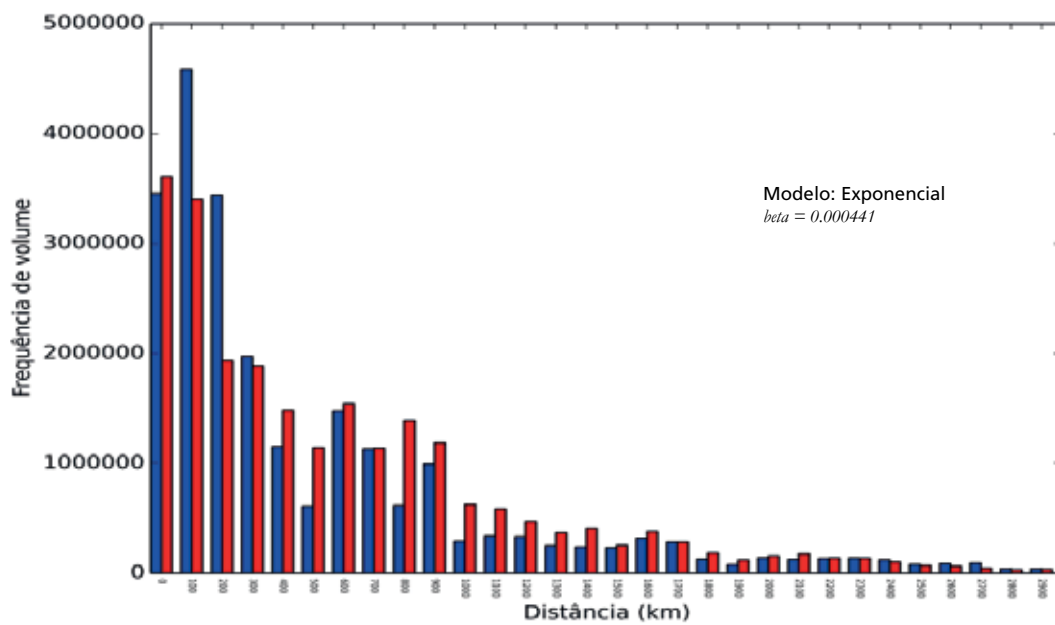
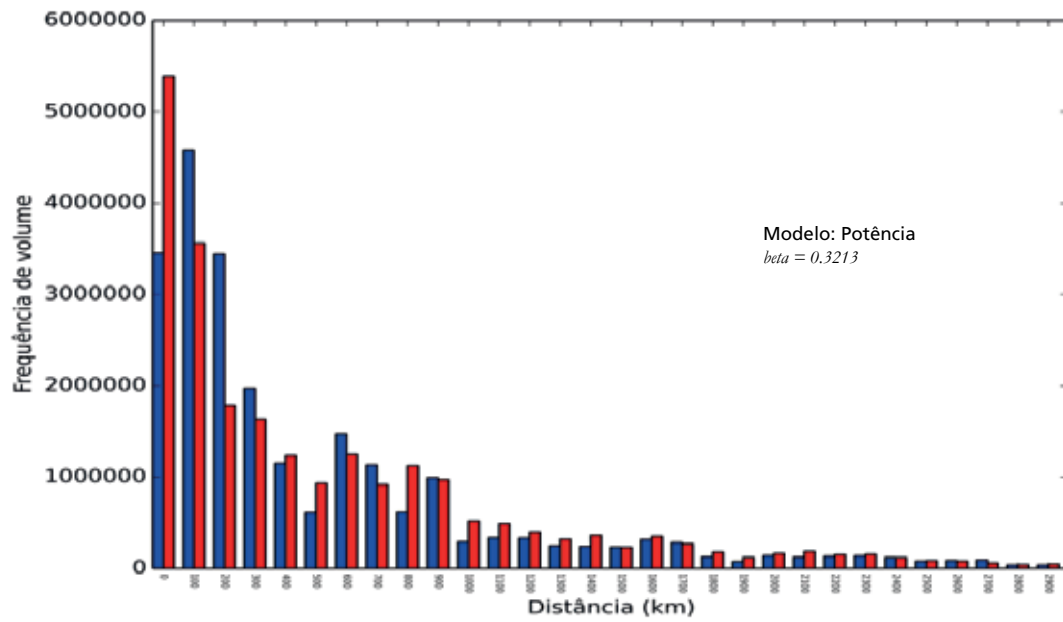


A.2 – Carvão mineral



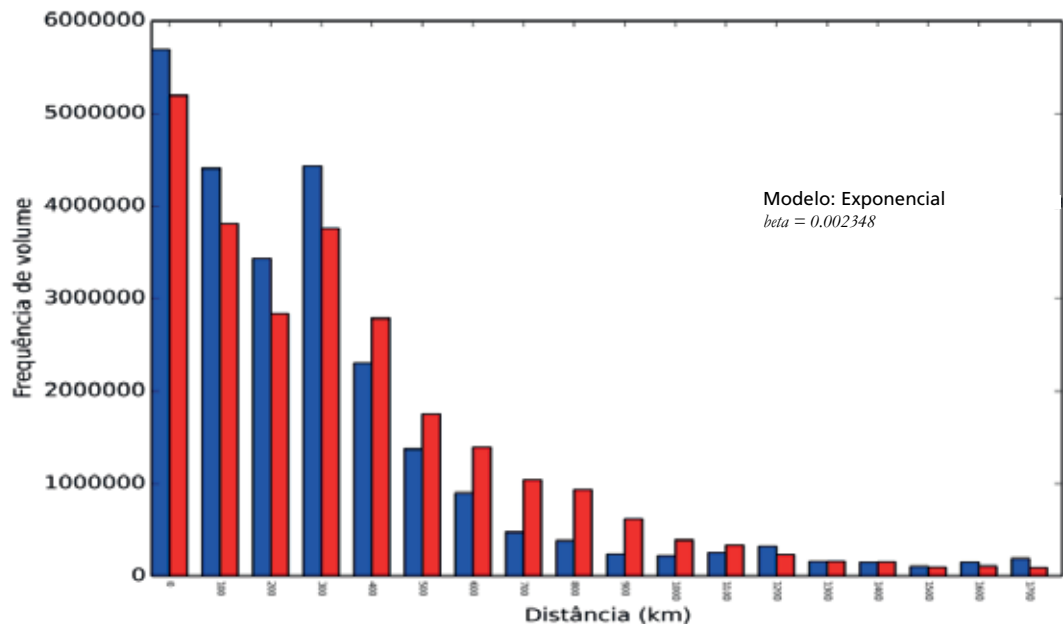
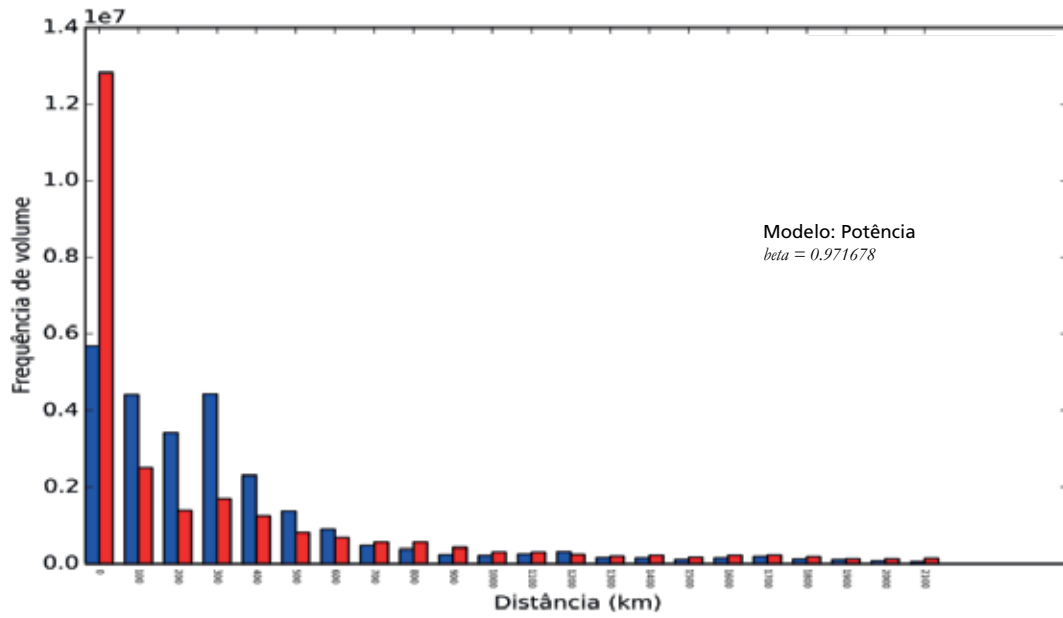
■ Observada ■ Calibrada

## A.3 – Celulose e papel



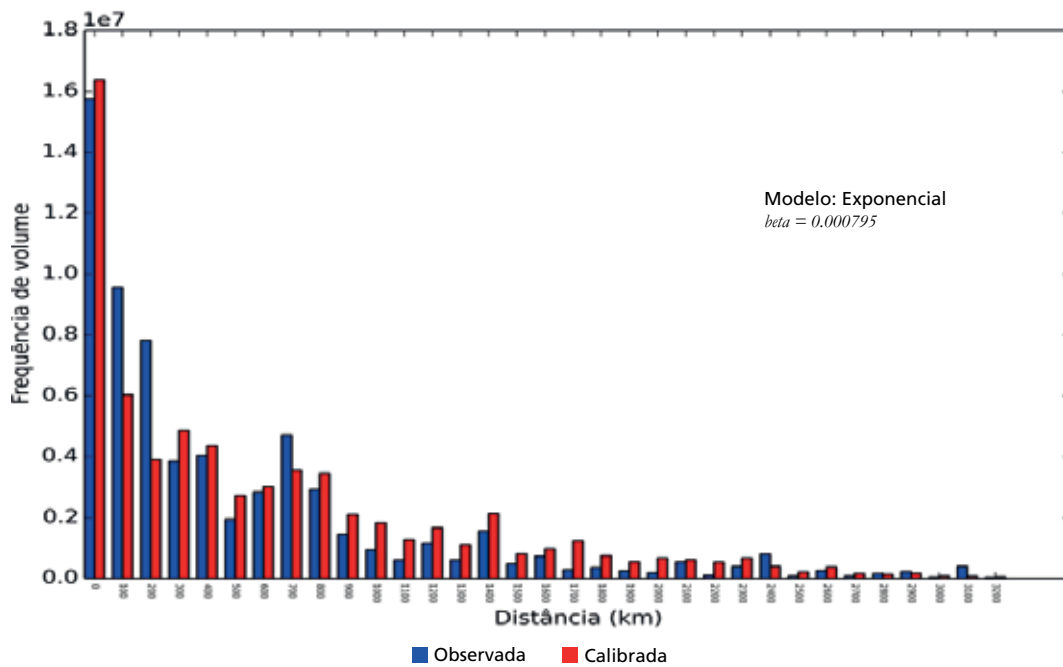
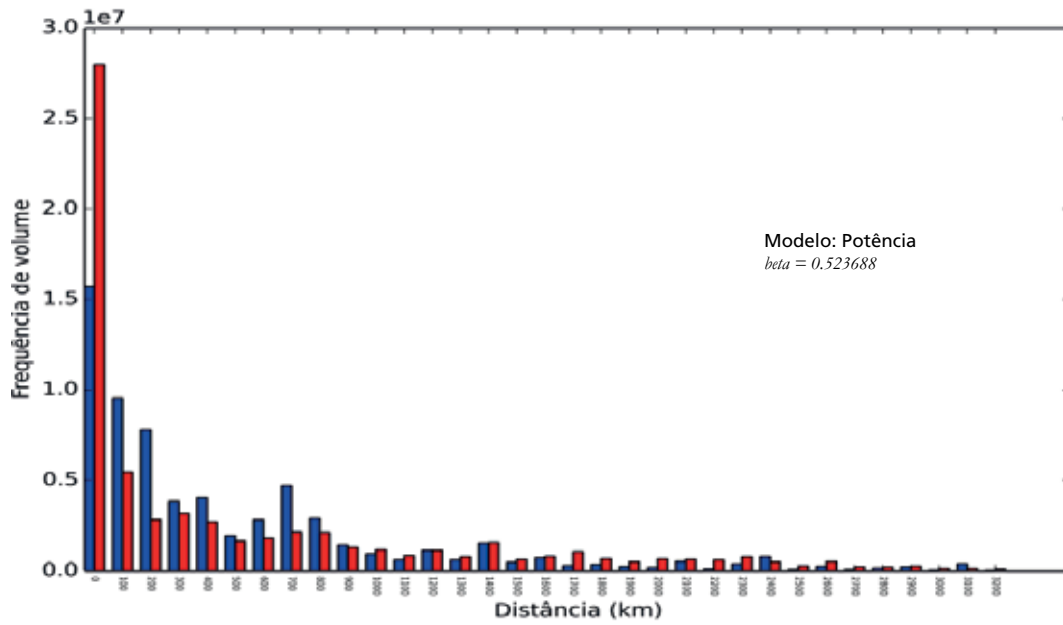
■ Observada ■ Calibrada

A.4 – Cimento



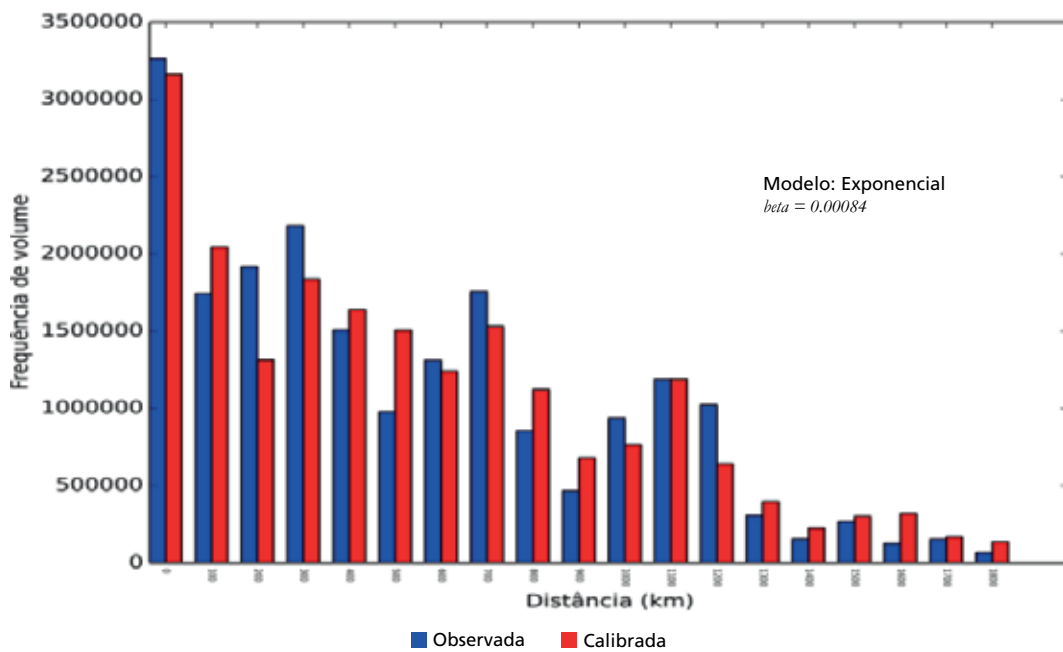
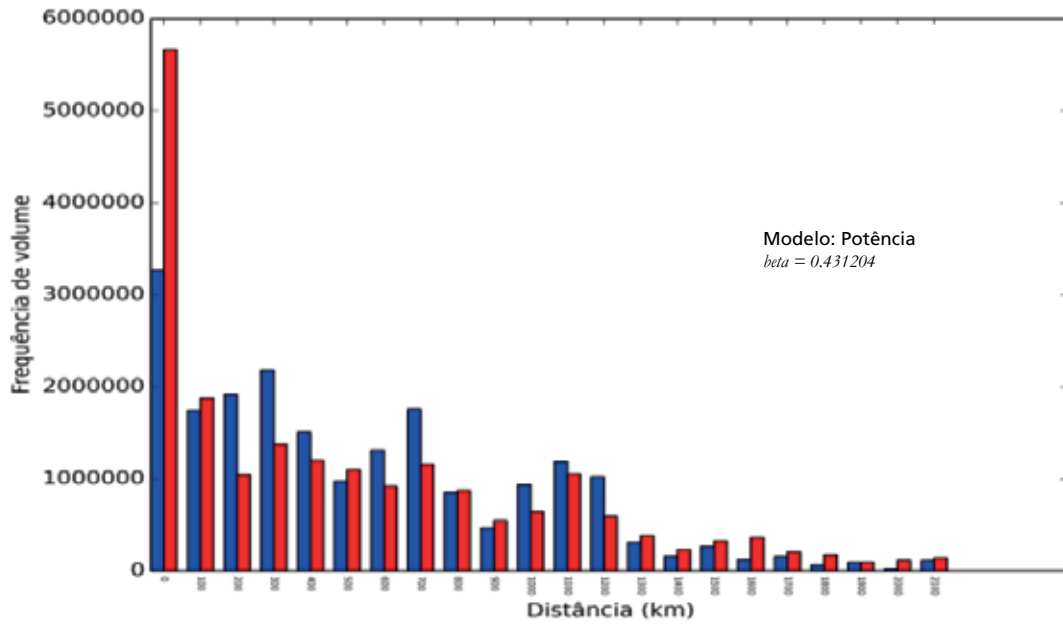
■ Observada ■ Calibrada

## A.5 – Combustíveis



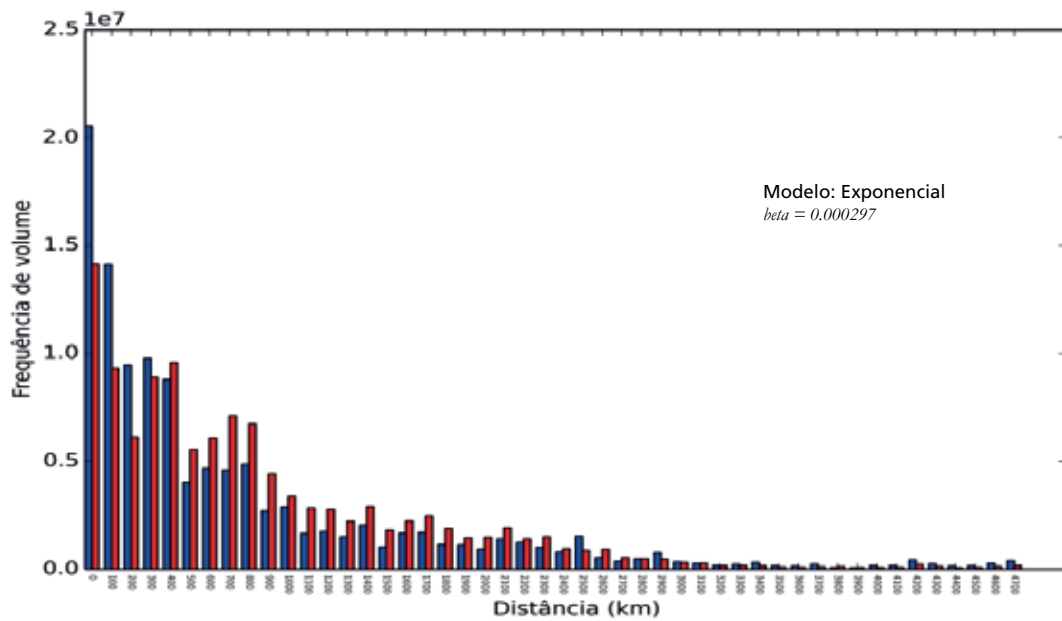
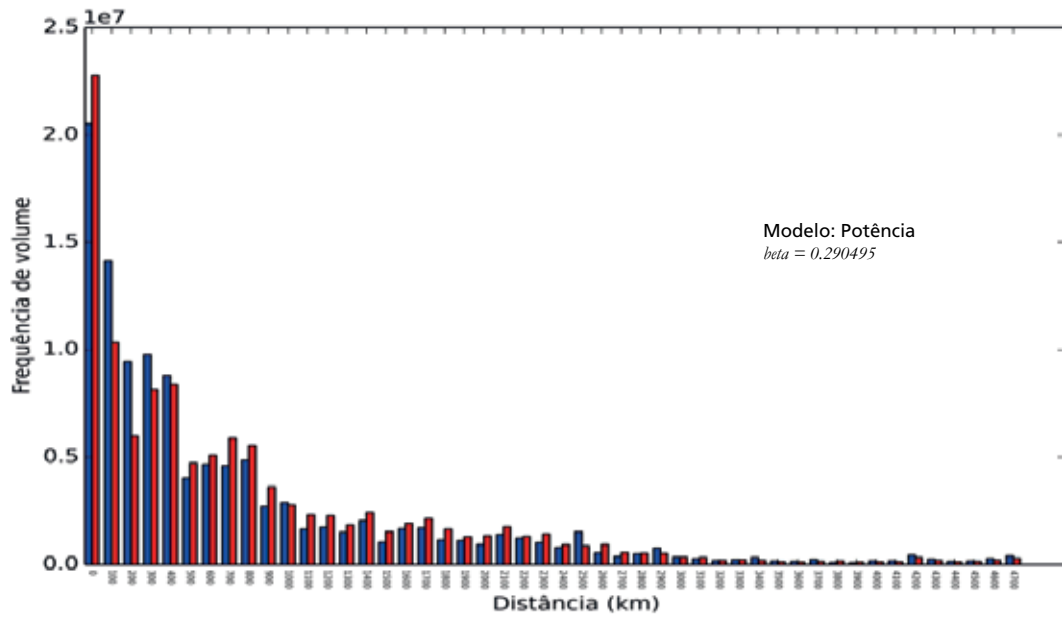


A.6 – Farelo de soja



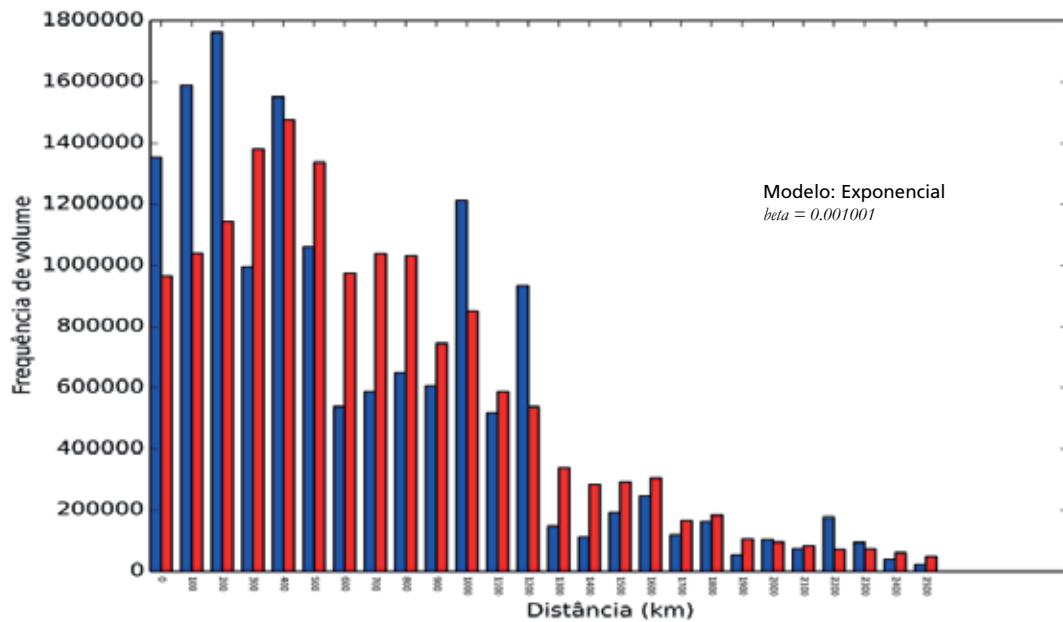
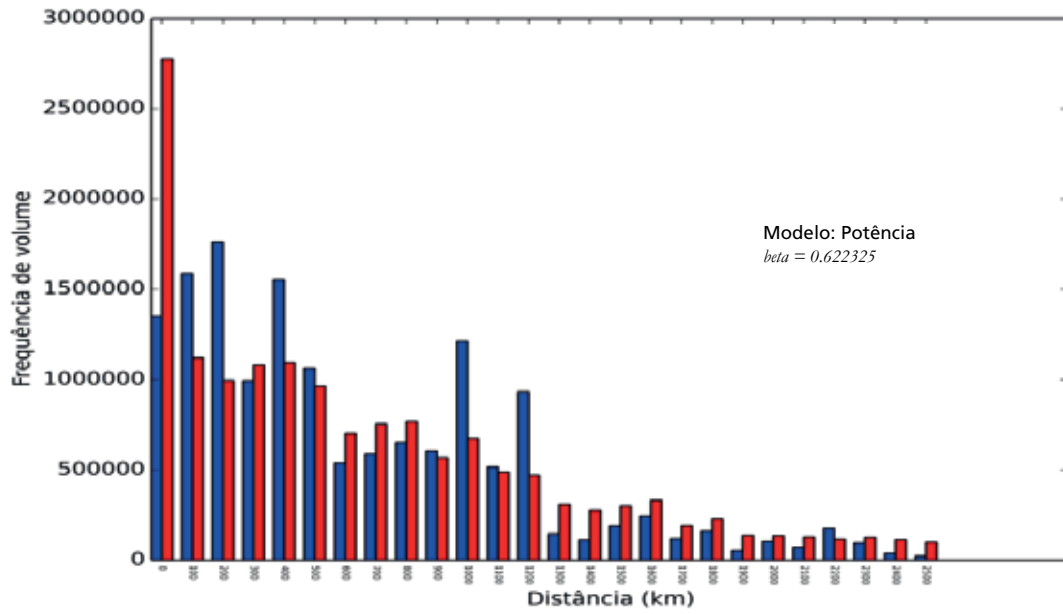
■ Observada ■ Calibrada

## A.7 – Manufaturados



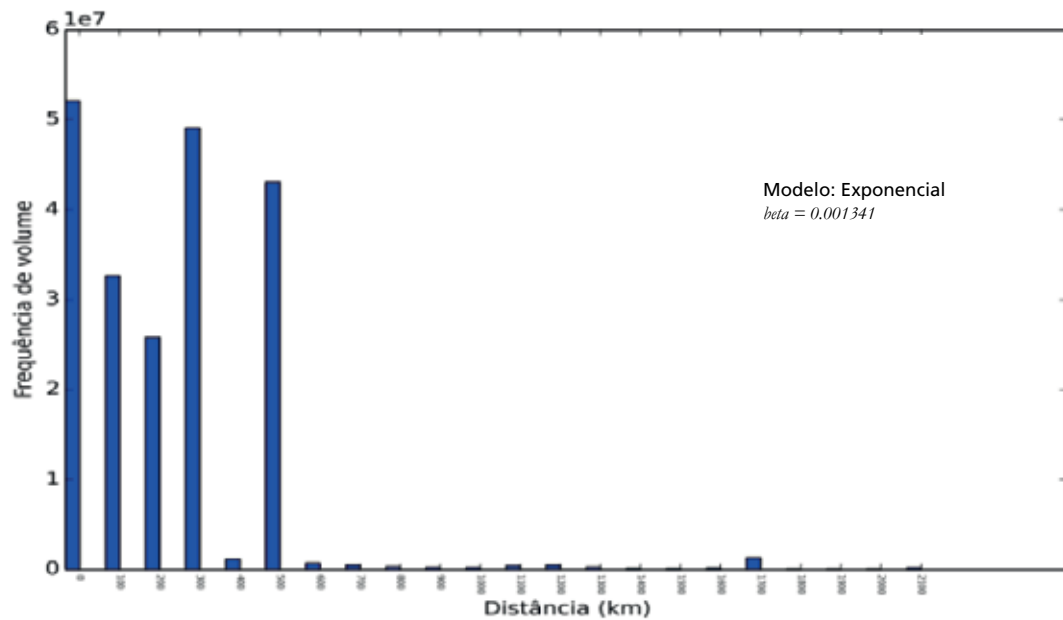
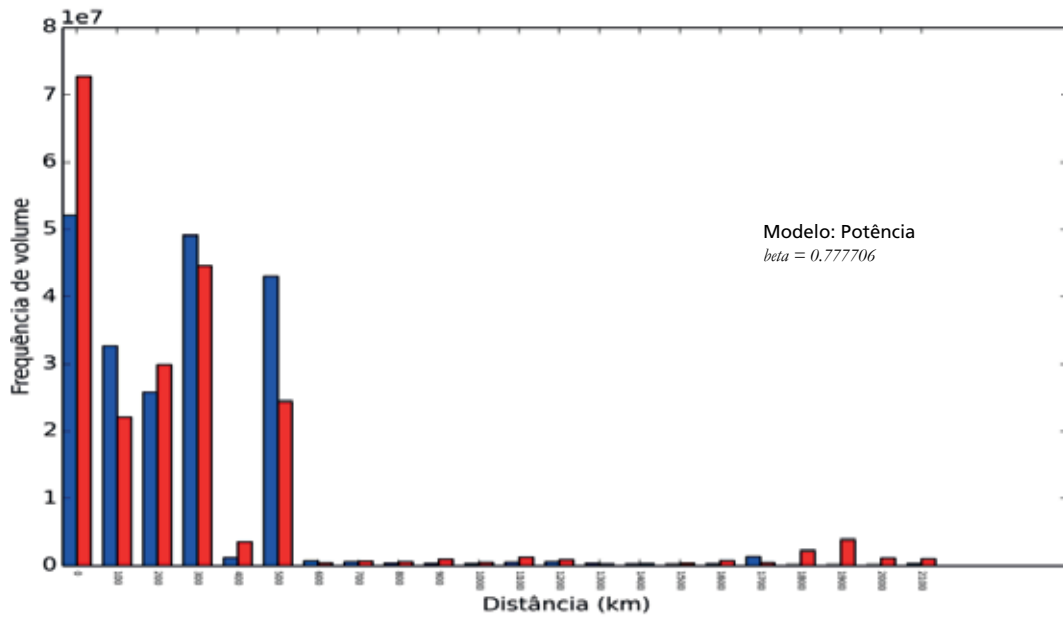
■ Observada ■ Calibrada

A.8 – Milho em grão



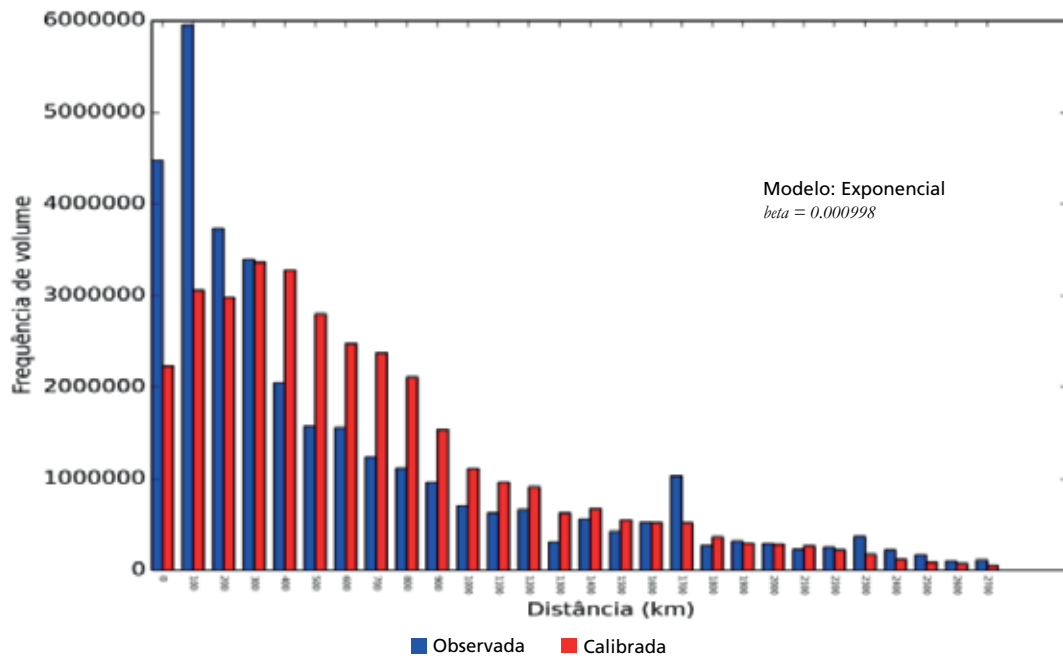
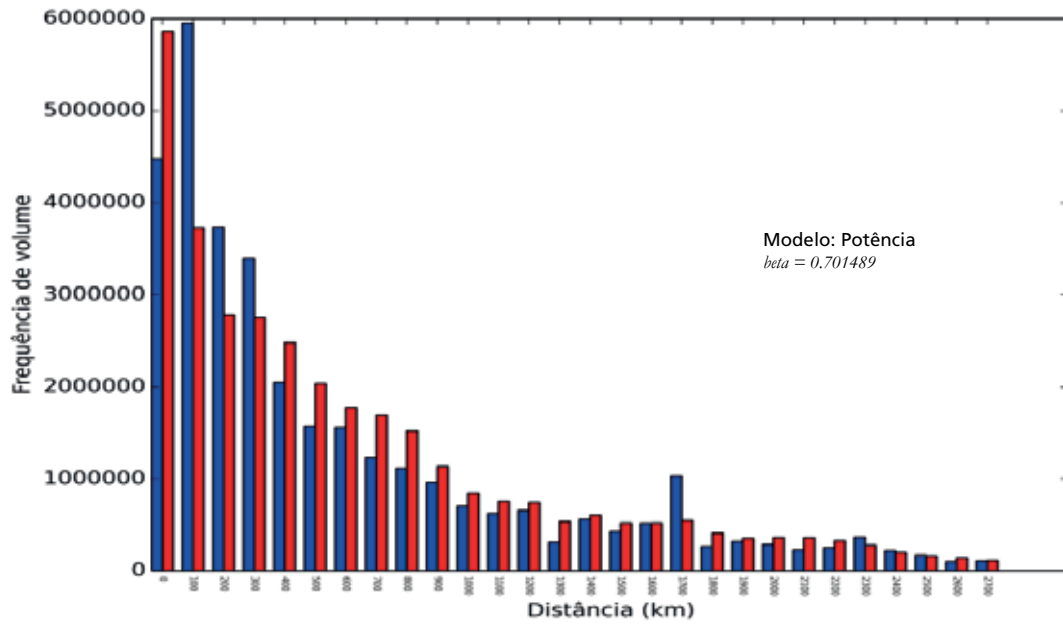
■ Observada ■ Calibrada

A.9 – Minério de ferro

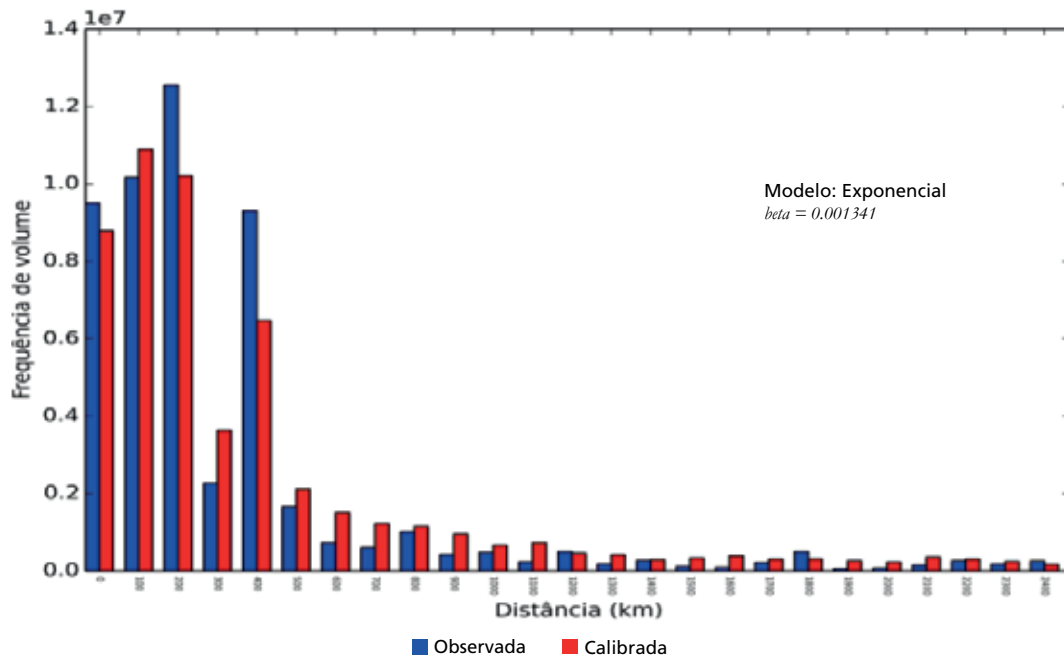
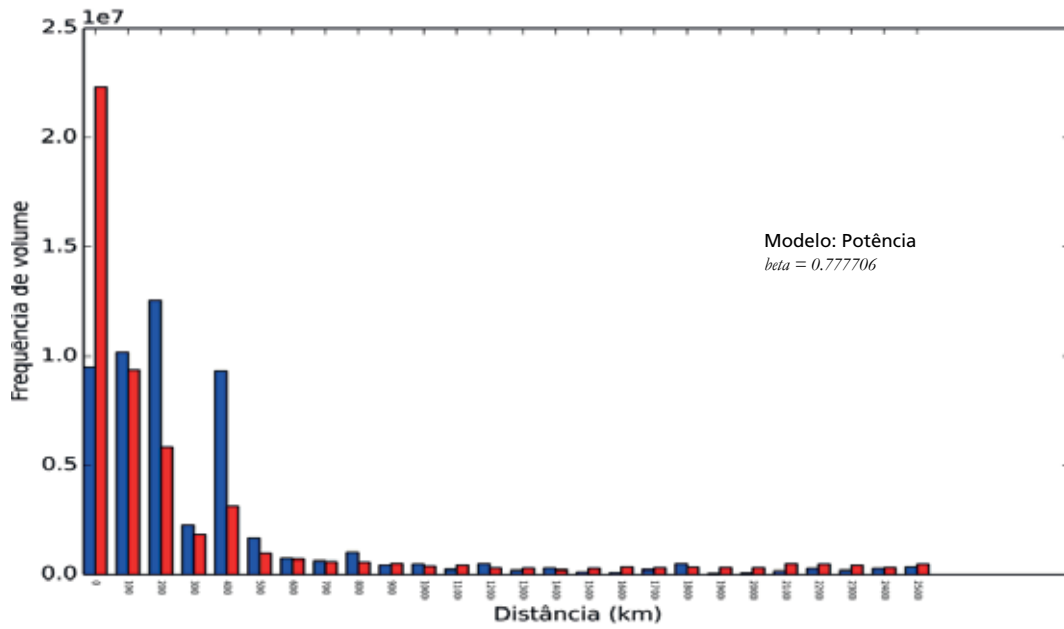


■ Observada ■ Calibrada

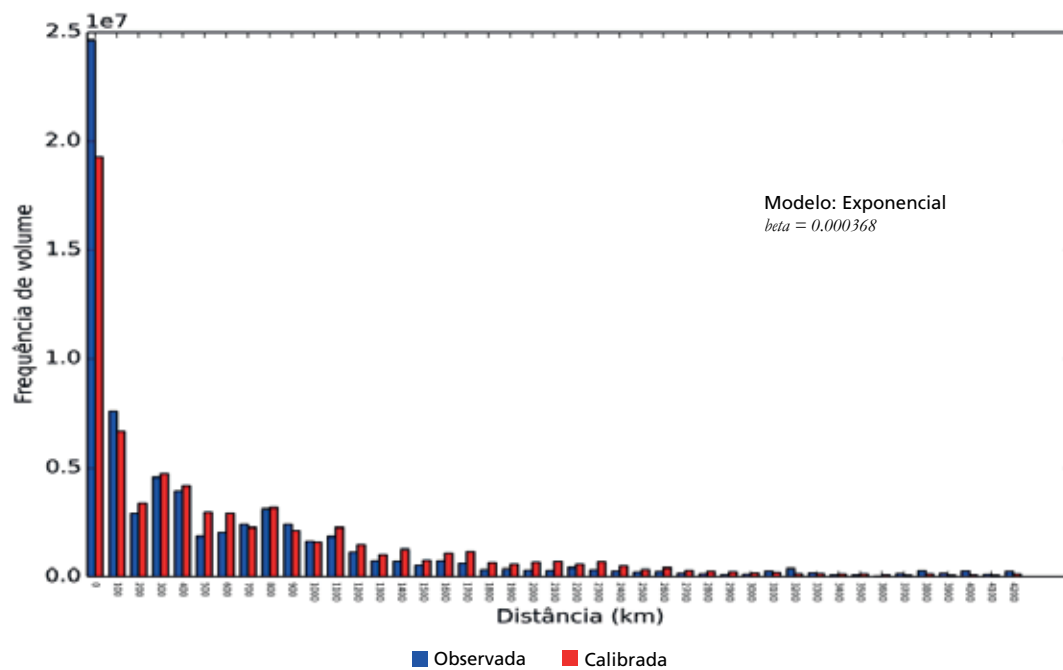
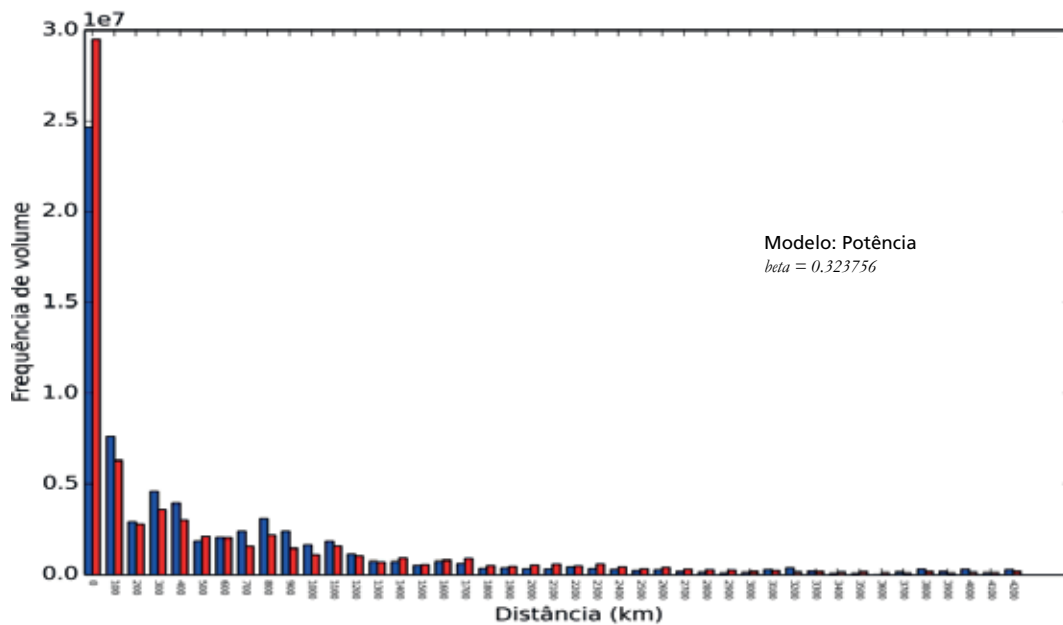
A.10 – Outros da lavoura e pecuária



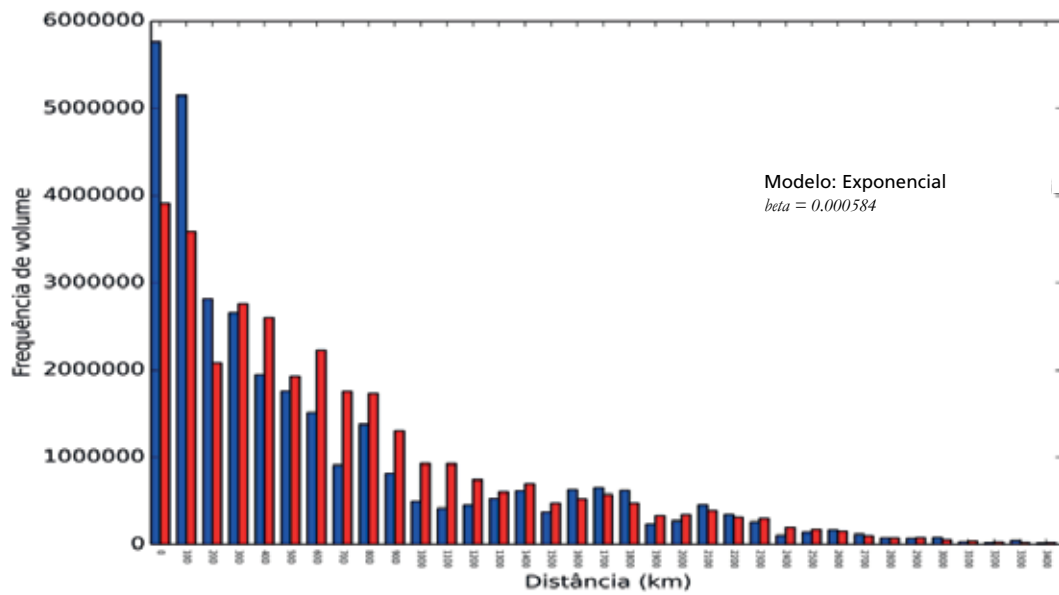
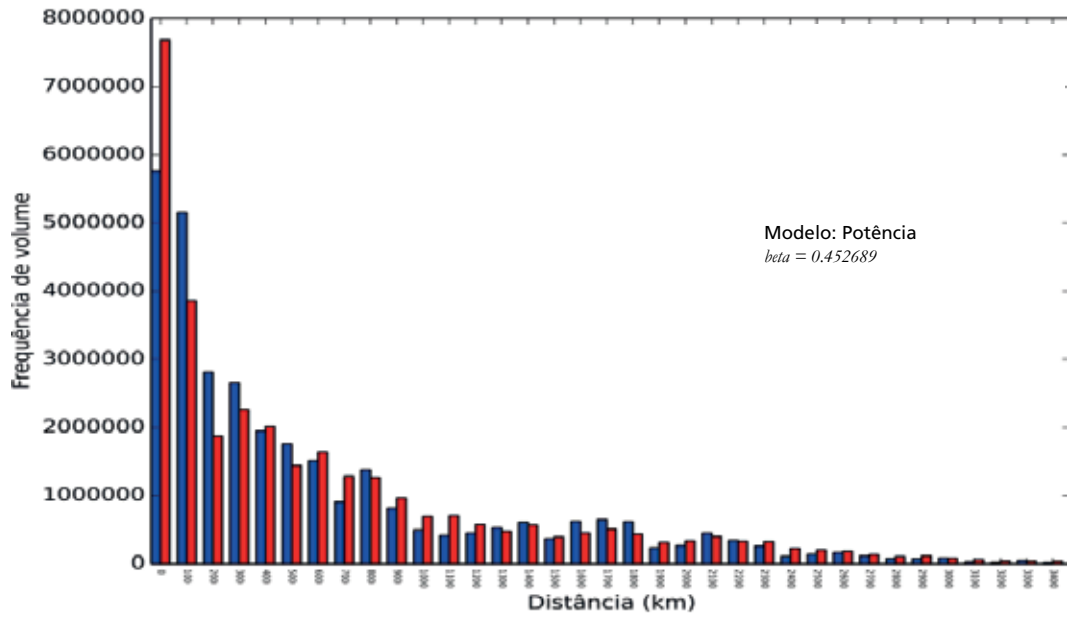
## A.11 – Outros minerais



A.12 – Petroquímicos e químicos



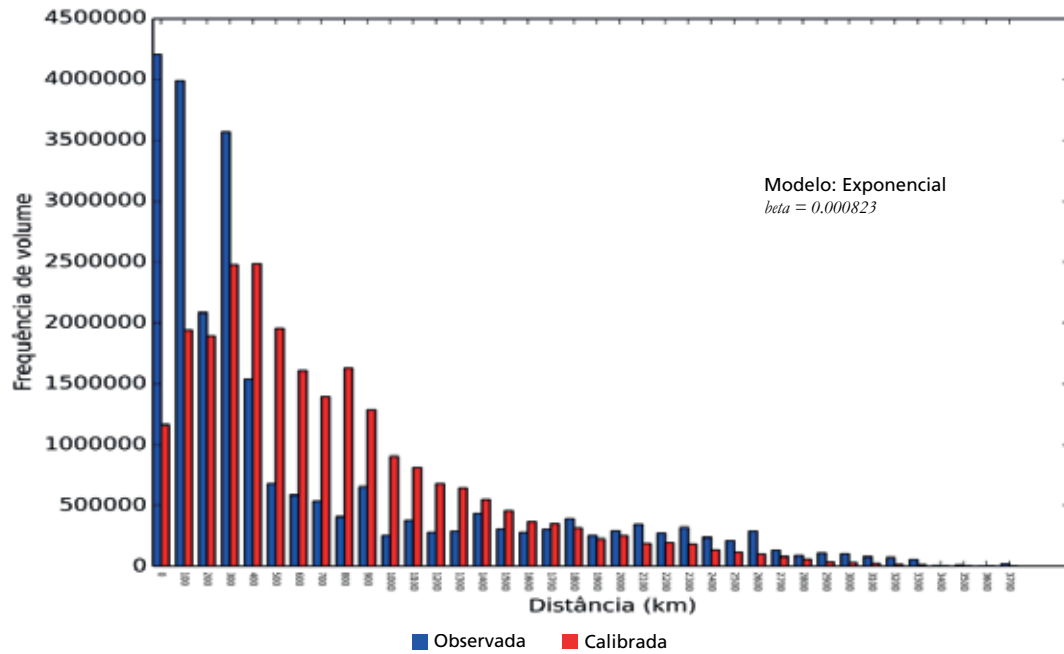
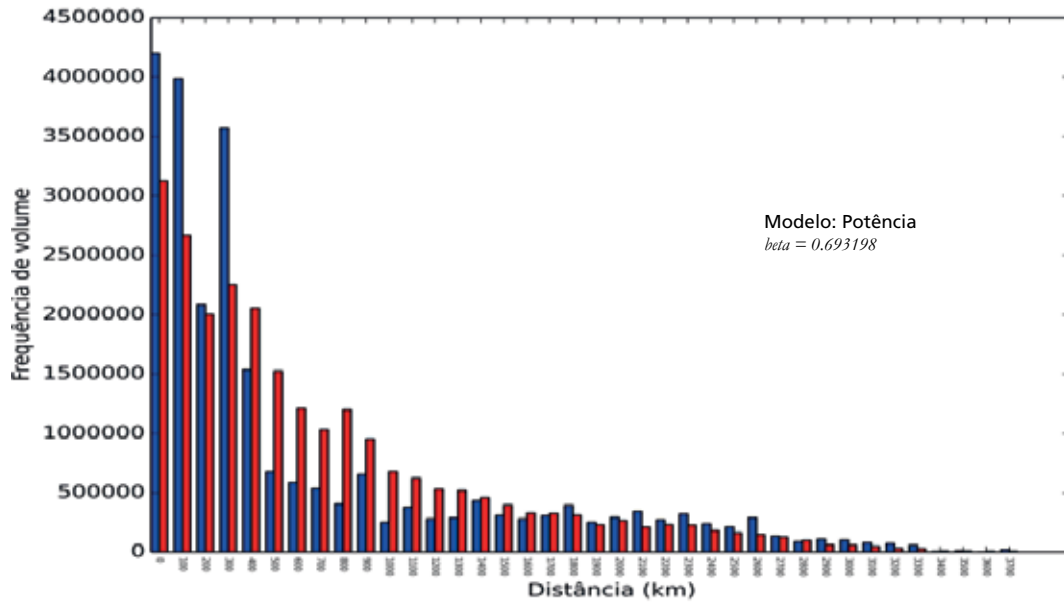
A.13 – Produtos básicos de borracha, plástico e não metálicos



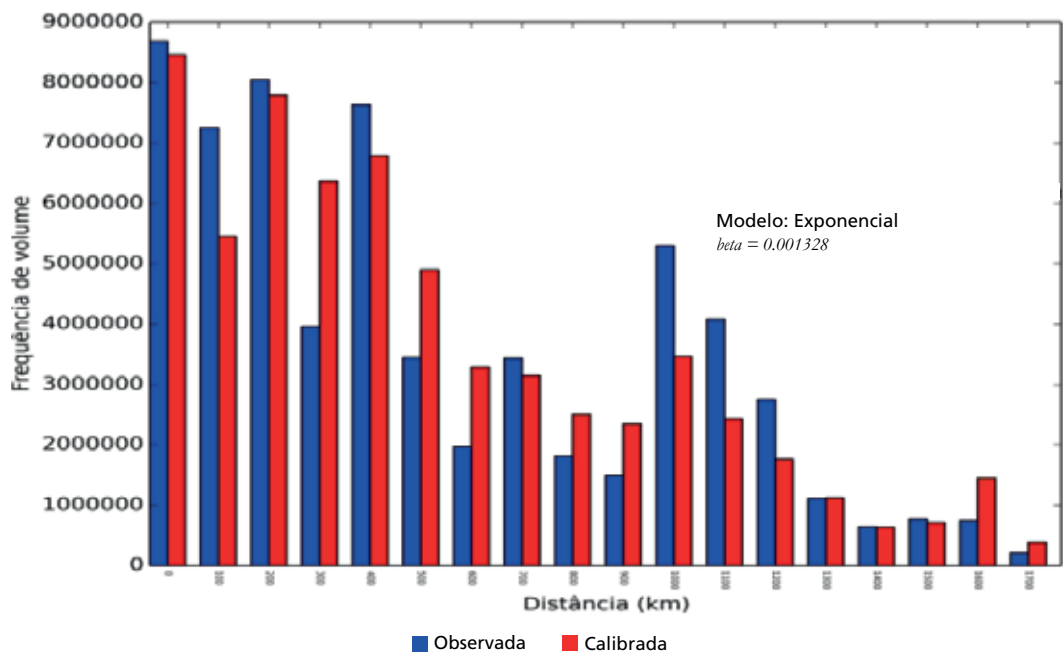
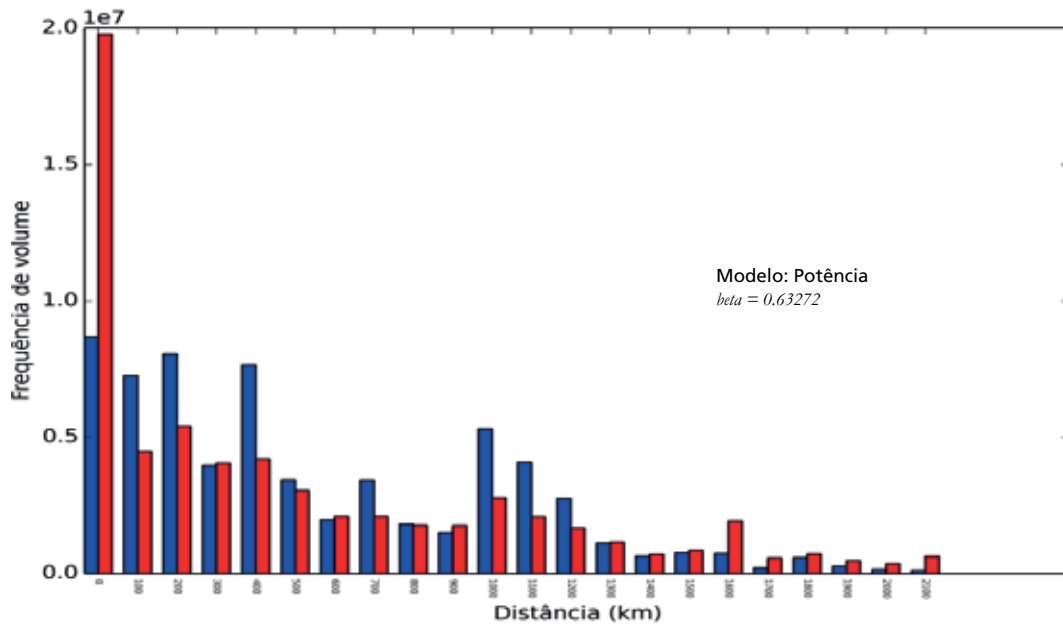
■ Observada ■ Calibrada



A.14 – Produtos da exploração florestal e da silvicultura



## A.15 – Soja em grão



Elaboração dos autores.

Obs.: figura reproduzida em baixa resolução e cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).



## **Ipea – Instituto de Pesquisa Econômica Aplicada**

### **EDITORIAL**

#### **Coordenação**

Cláudio Passos de Oliveira

#### **Supervisão**

Andrea Bossle de Abreu

#### **Revisão**

Carlos Eduardo Gonçalves de Melo

Elaine Oliveira Couto

Laura Vianna Vasconcellos

Luciana Nogueira Duarte

Mariana Silva de Lima

Vivian Barros Volotão Santos

Thais da Conceição Santos Alves (estagiária)

#### **Editoração eletrônica**

Aeromilson Mesquita

Aline Cristine Torres da Silva Martins

Carlos Henrique Santos Vianna

Gláucia Soares Nascimento (estagiária)

#### **Capa**

Andrey Tomimatsu

*The manuscripts in languages other than  
Portuguese published herein have not been proofread.*

#### **Livraria Ipea**

SBS – Quadra 1 – Bloco J – Ed. BNDES, Térreo

70076-900 – Brasília – DF

Tel.: (61) 2026 5336

Correio eletrônico: [livraria@ipea.gov.br](mailto:livraria@ipea.gov.br)







## **Missão do Ipea**

Aprimorar as políticas públicas essenciais ao desenvolvimento brasileiro por meio da produção e disseminação de conhecimentos e da assessoria ao Estado nas suas decisões estratégicas.

**ipea** Instituto de Pesquisa  
Econômica Aplicada

MINISTÉRIO DO  
**PLANEJAMENTO,  
DESENVOLVIMENTO E GESTÃO**

