

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

Relatório de Pesquisa

Construção da Matriz Origem-Destino Observada de Transporte de Passageiros Inter-regional (Relatório 6)



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

Relatório de Pesquisa

**Construção da Matriz Origem-Destino Observada de Transporte de
Passageiros Inter-regional (Relatório 6)**

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

Claudio Hamilton Matos dos Santos

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

Chefe de Gabinete

Márcio Simão

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

**Construção da Matriz Origem-Destino Observada de Transporte de
Passageiros Inter-regional (Relatório 6)**

ipea

Brasília, 2017

EQUIPE DE PESQUISA

Coordenação-geral

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

EQUIPE TÉCNICA

Fabiano Mezadre Pompermayer (Diset/Ipea)
Erivelton Pires Guedes, Assessoria Técnica da Presidência (Astep) do Ipea
Akina Sakamori, bolsista do PNPD (Diset/Ipea)
Alan Ricardo da Silva, bolsista do PNPD (Diset/Ipea)
Carolina Andrade Silva, bolsista do PNPD (Diset/Ipea)
Daniel Alisson Feitosa Lopes, bolsista do PNPD (Diset/Ipea)
Diego Rosa Mambrin, bolsista do PNPD (Diset/Ipea)
Gabriel Gouveia Rabello, bolsista do PNPD (Diset/Ipea)
João Gabriel de Moraes Souza, bolsista do PNPD (Diset/Ipea)
Maircon Batista Ribeiro, bolsista do PNPD (Diset/Ipea)
Paulo Henrique Dourado da Silva, bolsista do PNPD (Diset/Ipea)
Pedro Veiga de Camargo, bolsista do PNPD (Diset/Ipea)
Priscila Nascimento de Alcântara Garcia, bolsista do PNPD (Diset/Ipea)
Rafaella Bandeira Cabral Cunha, bolsista do PNPD (Diset/Ipea)
Ramon de Almeida Bispo, bolsista do PNPD (Diset/Ipea)
Raquel Araujo de Almeida, bolsista do PNPD (Diset/Ipea)
Rennaly Patricio Sousa, bolsista do PNPD (Diset/Ipea)
Roberto Lazarte Kaqui, bolsista do PNPD (Diset/Ipea)
Thiago Guimarães Rodrigues, bolsista do PNPD (Diset/Ipea)

EQUIPE DE ELABORAÇÃO DO RELATÓRIO

Akina Sakamori, bolsista do PNPD (Diset/Ipea)
Raquel Araujo de Almeida, bolsista do PNPD (Diset/Ipea)
Priscila Nascimento de Alcântara Garcia, bolsista do PNPD (Diset/Ipea)
João Gabriel de Moraes Souza, bolsista do PNPD (Diset/Ipea)
Alan Ricardo da Silva, bolsista do PNPD (Diset/Ipea)
Rafaella Bandeira Cabral Cunha, bolsista do PNPD (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 DESCRIÇÃO DAS BASES DE DADOS DA PESQUISA DE CAMPO RODOVIÁRIA.....	9
3 TRATAMENTO DAS BASES DE DADOS DA PESQUISA DE CAMPO RODOVIÁRIA.....	11
4 TRATAMENTO DAS BASES DE DADOS DE OUTROS MODOS DE TRANSPORTE	17
5 JUNÇÃO DAS BASES DE DADOS (MODOS DE TRANSPORTE)	19

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 se obtém a matriz origem/destino 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 O/D 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 O/D 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 O/D envolve o levantamento e 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. Este trabalho envolve conhecimento técnico específico em modelagem de transportes e também em economia e estatística.

Com uma matriz O/D 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, beneficia-se das informações da matriz O/D ao cruzá-las com outras fontes de dados socioeconômicos, podendo realizar estudos diversos sobre o 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 faz a agregação dos 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 O/D, 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 rodoviária 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 (UF) 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 O/D 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 descreve os procedimentos para os tratamentos dos dados da pesquisa de campo, os quais foram realizados com objetivo de criar uma base que gerará a matriz origem/destino observada do transporte de passageiros no Brasil.

Inicialmente, para tratamento dos dados, apresenta-se a descrição das bases de dados de contagens manual e automática de veículos (Trafficcount e MetroCount), a base de entrevistas da pesquisa de campo rodoviária e as bases de ônibus e de transporte aeroviário. Adiante, será exposto o tratamento da estimação dos turnos faltantes na base rodoviária, ao qual são imputados valores estimados nos casos em que há falta de informação na contagem de veículos. Em seguida, apresenta-se o método de expansão das viagens para que se tenha uma representação significativa da população, visto que a pesquisa é amostral. Posteriormente, há uma explicação do balanceamento entre as quantidades de passageiros entre zonas e o tratamento das bases de dados de outros meios de transporte, em que se utilizam as bases de ônibus e de transporte aeroviário. Na parte final do relatório, apresenta-se o método de junção das bases para elaboração dos modelos de geração e atração de viagens no Brasil.

As etapas mencionadas no parágrafo anterior serão mais bem tratadas nos itens que seguem este relatório, que servirá como insumo do relatório seguinte, que apresenta a seleção e a demonstração dos modelos de regressão linear múltipla para estimação da geração e da atração de viagens.

2 DESCRIÇÃO DAS BASES DE DADOS DA PESQUISA DE CAMPO RODOVIÁRIA

A pesquisa de campo tratada neste relatório foi realizada em duas etapas. A primeira etapa foi executada nos meses de novembro e dezembro de 2013 e a segunda etapa nos meses de abril e maio de 2014. Como a primeira etapa não teve seus dados validados, foram usados apenas os obtidos na segunda etapa. A seguir, serão descritas as principais bases utilizadas na construção das matrizes O/D. Vale ressaltar, ainda, que a mesma pesquisa de campo levantou os dados de movimentação de cargas nas rodovias, usados na matriz O/D observada de cargas, discutida em outro relatório.

2.1 Definições

A base MetroCount contém os dados coletados por meio de um equipamento de mesmo nome, com o objetivo de contar automaticamente a quantidade de veículos por tipo, segundo seu comprimento e a quantidade de eixos. Para a coleta dos dados, estes equipamentos mantiveram-se ligados 24 horas por dia, durante o período de uma semana, e cobriam os dois sentidos da rodovia.

A base de dados Trafficcount, contém os dados coletados por meio de um contador manual, que permite classificar os veículos conforme tamanho e uso (motos, automóveis, micro-ônibus, camionetes, ônibus, caminhões, caminhões articulados etc.). Tal contagem ocorreu em uma escala de turnos de acordo com os dias da semana, apresentada no quadro 1.

O MetroCount se distingue do contador manual por ser automático e não diferenciar ônibus em relação a caminhões de pequeno porte, devido à semelhança do desenho dos eixos. Além disso, a existência de duas bases com a mesma contagem traz a possibilidade da checagem e escolha do dado mais fidedigno. A base de dados Trafficcount apresentava inicialmente o nome *Contagens5mvo*.

QUADRO 1

Escala de turnos de contagem manual de acordo com o dia da semana

Dia	Domingo	Segunda-feira	Terça-feira	Quarta-feira	Quinta-feira	Sexta-feira	Sábado
0-6h		X		X			X
6-12h		X	X	X	X		X
12-18h	X		X	X	X	X	
18-24h	X			X		X	

Fonte: Termo de Referência da Pesquisa de Campo Rodoviária da Empresa de Planejamento e Logística.

A base de dados *Entrevistas* refere-se à pesquisa de campo e é formada por meio da junção de três bases disponibilizadas pela Empresa de Planejamento e Logística, cujos nomes são: *RespostasOD*, *Perguntas* e *QuestionáriosOD*. Os dados levantados foram:

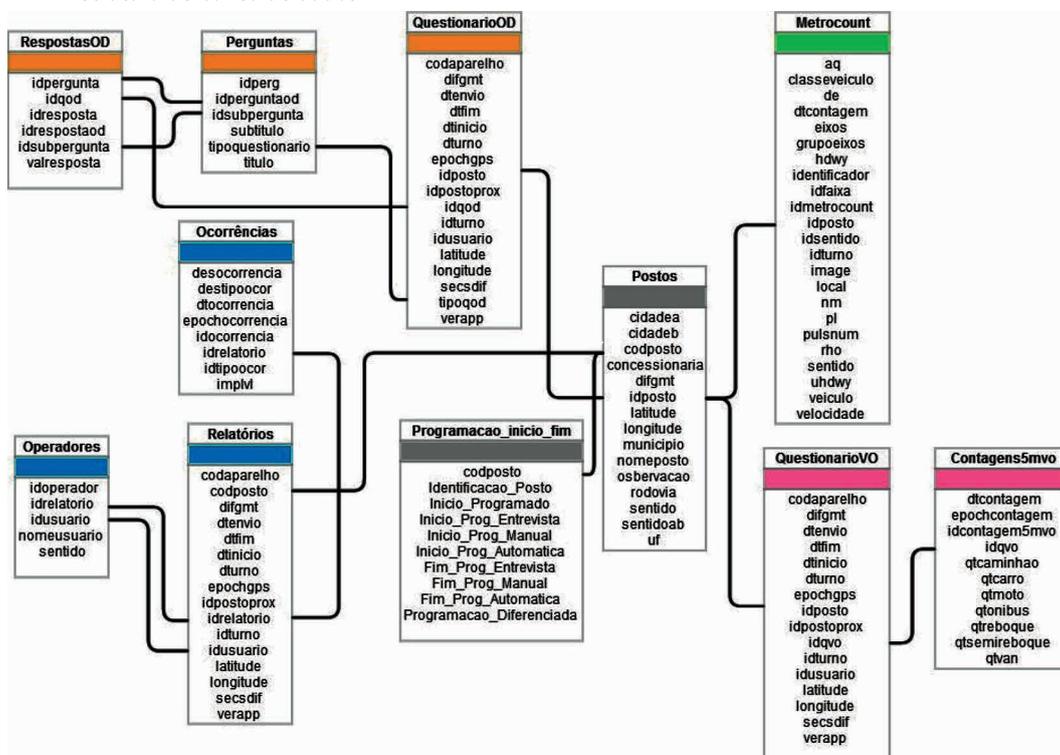
- locais de origem, destino e residência (país, estado, município) do viajante entrevistado;
- tipologia do veículo (classificação, número de eixos totais, número de eixos rodantes, propriedade do veículo);
- dados da carga transportada (valor do frete, valor da carga, peso da carga);
- motivo (trabalho, passeio, estudo, saúde) e frequência da viagem de passageiro (diária, semanal, quinzenal, mensal); e
- dados da rota (influência de postos de pedágio ou qualidade da rodovia na escolha da rota).

2.2 Estrutura do banco de dados

A figura 1 ilustra quais bases de dados são interligadas, e por meio de quais variáveis isso ocorre.

FIGURA 1

Estrutura de banco de dados



Elaboração dos autores.

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

3 TRATAMENTO DAS BASES DE DADOS DA PESQUISA DE CAMPO RODOVIÁRIA

Inicialmente as bases foram tratadas separadamente para que, em seguida, fosse possível a construção de uma única base com as variáveis necessárias para análise. Neste relatório são apresentadas somente as variáveis usadas na matriz de passageiros.

3.1 Tratamento inicial

3.1.1 Trafficcount

Para a construção da base de dados Trafficcount, o primeiro passo foi a importação do arquivo *Contagens5mvo*. Em seguida, foram retiradas as informações duplicadas oriundas de falhas nos equipamentos de contagem.

Para acrescentar a localização dos postos de coleta na base Trafficcount, utilizou-se a base de dados *Postos* que possui informações como cidade, município e estado do posto. Esses dados foram utilizados para verificar se um veículo com determinado par origem-destino deveria passar pelo posto em questão.

Em seguida, foram adicionadas as informações do cronograma de coleta dos dados. A EPL disponibilizou uma programação informando as datas de início e fim das contagens de tráfego, tanto para o contador manual quanto para o automático. Essa base se chama *Programação dos postos*. Foram desconsideradas contagens realizadas fora do período informado no cronograma.

Para determinar a quantidade de cada tipo de veículo que trafegou por turno em cada posto, foram somadas as contagens de acordo com as variáveis referentes ao turno (*diamesano, turno, hora*), ao posto e seu respectivo sentido (*codposto, idposto, sentido, sentidoab*). Em seguida, foram somados os veículos de passeio e os de carga separadamente.

3.1.2 MetroCount

Para a construção da base de dados MetroCount foi realizada a importação do arquivo de mesmo nome, *Metrocount*. Essa base possuía uma variável de nome *hdwy* que representa o tempo entre dois veículos que trafegavam em uma mesma faixa. Quando essa variável apresentava zero como valor, era possível identificar contagens duplicadas decorrentes de falhas do aparelho na apuração automática. Após verificar que esses casos estavam realmente duplicados, uma dessas observações foi desconsiderada.

Para validação dos dias de coleta, também foi utilizada a base com o cronograma disponibilizado pela EPL. Contagens realizadas fora do período informado no cronograma foram desconsideradas.

Na base MetroCount, foram acrescentadas as informações dos postos de coleta por meio da mesma metodologia utilizada na base Trafficcount.

Com isso, os veículos do contador MetroCount foram classificados como veículos de carga ou de passeio, de acordo com a descrição do tipo de veículo, demonstrado na tabela 1.

TABELA 1
Classificação dos veículos (MetroCount)

Classificação	Tipo de veículo	Total de veículos
	2C/2CB	2.182.533
	2C2	25.468
	2C3	5.584
	2I3	5.303
	2S1	107.694
	2S2	330.047
	2S3	214.65
	3C/3CB	971.765
Carga	3C2	35.191
	3C3	12.557
	3I3	53.515
	3S1	8.379
	3S2	57.391
	3S3	324.535
	4C	13.951
	4CD	141.545
	Bit	157.08
	Util	479.72
Desconsiderado	???	1.822.212
	Moto	878.636
Passeio	Pass	13.651.705
	Rod/Trit	24.406
Total final		21.503.867

Elaboração dos autores.

Obs.: ??? = veículos não identificados pelo aparelho de contagem.

Neste relatório foram analisados apenas os veículos de passeio.

3.1.3 Entrevistas

Para a construção da base de dados *Entrevistas*, foram necessárias algumas transformações nas bases que a compõem. A base de *Respostas* foi disponibilizada no formato vertical, ou seja, cada linha correspondia a uma resposta. A base em questão foi transposta para o mesmo formato da base *Entrevistas*, ou seja, cada linha correspondendo a uma entrevista e cada coluna uma variável/pergunta.

A base *QuestionárioOD*, composta de informações técnicas de cada questionário, como *id do entrevistador*, *id do posto*, *data de início* e *fim da entrevista*, foi acrescida à base *Perguntas* junto com a base *Respostas* já transposta. Para o cruzamento dessas bases, foi utilizada as seguintes variáveis: *idpergunta* (identifica qual pergunta o entrevistado está respondendo), *idsubpergunta* (identifica qual item, da pergunta o entrevistado está respondendo) e *idresposta* (identificação da resposta).

A base *Entrevistas* foi dividida em duas, de acordo com o tipo de questionário: carga ou passageiro. Para fins da análise deste relatório, a base foi filtrada por tipo de questionário que contemplava as questões referentes a passageiros.

Foram encontrados alguns problemas na base *Entrevistas*, como casos em que os nomes dos municípios possuíam erros ortográficos. Para isso, foram corrigidos os nomes dos municípios a fim de poder realizar as análises compostas neste relatório.

A grafia dos municípios foi padronizada de forma manual. A informação do nome dos municípios de origem e de destino da viagem precisava estar de acordo com os nomes oficiais disponibilizados pelo Instituto Brasileiro de Geografia e Estatística (IBGE).

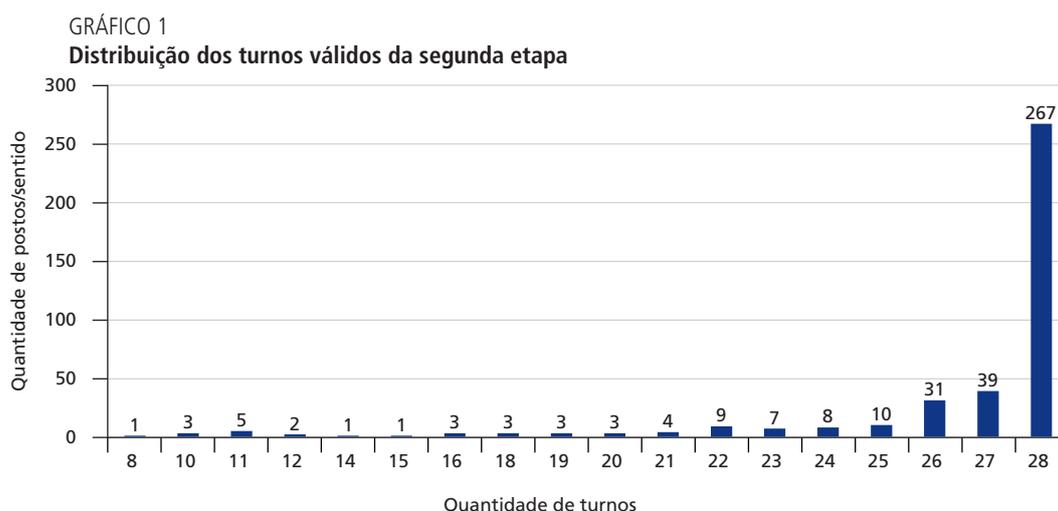
Havia também nomes de distritos e bairros, que foram alocados aos respectivos municípios. Porém, não foi possível a identificação de alguns nomes devido à grafia díspar dos nomes exatos dos municípios, sendo estes casos alocados à capital do estado informado.

Foram consideradas na base *Entrevistas* apenas:

- entrevistas coletadas nos dias programados, conforme cronograma disponibilizado pela EPL; e
- entrevistas com descrição dos municípios de origem e de destino passíveis de codificação, de acordo com os nomes oficiais disponibilizados pelo IBGE.

3.2 Estimação dos turnos faltantes

Na segunda etapa, para a análise de dados dos passageiros, não foi realizada a análise gráfica inicial da primeira etapa, conforme feita na análise dos dados de carga,¹ a qual tinha o intuito de comparar as contagens dos equipamentos utilizados na pesquisa de campo, pois a EPL assegurou uma boa qualidade dos dados oriundos dos equipamentos Metrocount e Trafficcount, sendo necessária somente a estimação de turnos faltantes. Uma análise descritiva dos dados da segunda etapa pode ser visualizada no gráfico 1, em que foram obtidos 267 casos (66,75%) de postos/sentidos totalmente completos (com sete dias da semana e com quatro turnos por dia) e 133 casos (33,25%) com postos/sentidos incompletos.



Fonte: ANAC (2012).
Elaboração dos autores.

Para se inferir a sazonalidade das viagens de passageiros a partir apenas da segunda etapa da pesquisa de campo, foi feito um ajuste usando a base de dados de contagens de veículos em radares fixos do Departamento Nacional de Infraestrutura de Transportes (DNIT), a qual foi cedida pela EPL. A base do DNIT continha informações sobre as quantidades de veículos que passavam na rodovia, as quais eram inferidas por meio dos radares de controle de velocidade.

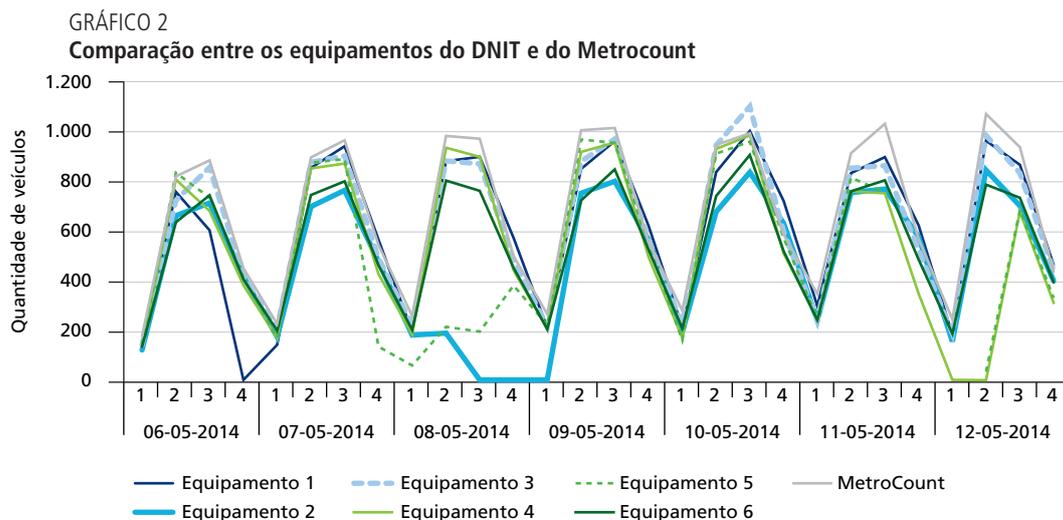
Entretanto, no decorrer do ajuste, surgiram alguns problemas considerados relevantes no cruzamento da base de dados do DNIT com a base de dados da pesquisa de campo. Um dos problemas identificado foi o fato de que nem sempre os equipamentos do DNIT poderiam estar no mesmo lugar em que o equipamento Metrocount havia sido instalado.

1. Ver relatório 2, *Construção da matriz origem/destino observada para transporte de cargas inter-regional*, desta série de relatórios de pesquisa.

Pode-se citar, por exemplo, uma situação em que o equipamento do DNIT possa ter sido instalado antes de uma bifurcação e o equipamento do Metrocount em um local depois desta bifurcação, ocasionando um menor fluxo após a bifurcação e conseqüentemente, uma diferença de contagens de veículos entre os equipamentos.

Outro problema foi a constatação de que cada equipamento pudesse captar quantidades diferentes de faixas da rodovia. Ou seja, poderia haver um equipamento que captava as informações dos veículos em somente uma faixa da rodovia, enquanto outro equipamento conseguia captar as quantidades de veículos de duas ou mais faixas da rodovia. Assim, estes tipos de situações poderiam causar divergência e confusão em relação a qual informação utilizar.

Para se conseguir utilizar os dados do DNIT nessas análises, foi feita uma comparação em cada posto/sentido entre os valores do Metrocount e dos vários equipamentos do DNIT que haviam na rodovia próximos ao posto de coleta. Os valores do Trafficcount não foram considerados na comparação, visto que os seus respectivos valores para posto/sentido estavam de acordo com o Metrocount. No gráfico 2, está representado um exemplo dos gráficos que foram analisados na comparação dos valores do DNIT e Metrocount. No eixo X (eixo horizontal), estão apresentados os dias da semana com os seus respectivos turnos e, no eixo Y (eixo vertical), as quantidades de veículos totais (carga e passeio). Cada linha do gráfico representa um equipamento da base do DNIT ou do Metrocount, de acordo com as cores da legenda.



Elaboração dos autores.

Após os ajustes das quantidades de veículos feitas por meio das escolhas entre os valores oriundos do DNIT ou do Metrocount para os postos/sentidos com turnos completos, foram realizadas as imputações dos postos/sentidos com os turnos incompletos (menores que 28 turnos).

Essa estimativa foi realizada da seguinte forma: primeiro, gerou-se as médias, por turno e dia da semana, das quantidades de veículos dos sentidos totalmente completos (sete dias e 28 turnos); em seguida, gerou-se uma distribuição com o percentual médio de carros que passaram em cada turno e dia da semana, por meio da divisão das médias calculadas para cada turno em relação pela soma total das médias.

A partir deste percentual, estimaram-se os turnos faltantes, com o intuito de obter os dados completos. Esta estimação dos turnos faltantes foi feita por meio da divisão da soma da quantidade de veículos nos turnos validados pela soma do percentual que esses turnos validados representam, por posto de coleta. Portanto, no final deste cálculo, obteve-se uma base completa com as quantidades de veículos por posto/sentido para todos os turnos.

Sabe-se que a quantidade de veículos que passa em um posto em determinado sentido, deve ser aproximadamente a mesma que passa no sentido oposto. A partir disso, foram analisadas as quantidades de veículos por sentido em cada um dos postos.

Assim, foi identificado, em apenas quatro postos, um desbalanceamento relevante, ou seja, havia diferenças significativas entre as quantidades de veículos que passavam em cada um dos sentidos destes postos. Assim, as quantidades de veículos dos sentidos dos respectivos postos foram ajustadas com o intuito de resolver este problema.

3.3 Fator de expansão

O objetivo dessa etapa é estimar a quantidade total de pessoas em veículos de passeio por par origem/destino por meio dos contadores automáticos e os resultados das entrevistas de campo. Para isso foi calculado o fator de expansão em três fases.

A primeira fase equivale ao cálculo do $Fator1_p$, que corresponde à razão entre a quantidade de veículos de passageiros que foram contados (V_c), após ajuste de turnos faltantes, e a quantidade de veículos entrevistados na pesquisa (V_p), e multiplicado pela quantidade de pessoas P_E dentro dos carros entrevistados, por posto (p).

$$Fator1_p = \frac{V_c}{V_p} \cdot P_E \quad (1)$$

Essa abordagem pode levar a estimativas discrepantes para pares O/D que tenham sido detectados em mais de um posto de pesquisa. A forma mais adequada de tratar estas estimativas é considerando a participação do posto de coleta nas possíveis rotas (caminhos) entre uma origem e um destino. O método mais empregado para se estimar essas participações é por meio de alocação de fluxos em rede, que dependem de uma boa calibração dos custos de cada caminho. Como estes custos ainda estavam sendo ajustados pela EPL, não foi possível usar tal abordagem.

Alternativamente, adotou-se a estratégia de considerar a maior estimativa encontrada para cada par O/D como sendo a informação mais plausível. Por exemplo, se para um par O/D obteve-se a estimativa de quatrocentos passageiros a partir de um posto e trezentos em outro posto, utiliza-se a estimativa de quatrocentos passageiros, pois o segundo posto poderia estar em uma rota que possuísse caminhos alternativos, dividindo o tráfego do par O/D. Assim, a segunda fase equivale ao cálculo do $Fator2_{ijp}$. Esse corresponde à multiplicação entre o $Fator1_p$ e a razão entre o máximo e a soma da estimativa total (Est_{ijp}) no par origem i e destino j do posto p . A estimativa total (Est_{ijp}) é o produto do primeiro fator de expansão pela quantidade de passageiros que foram entrevistados em cada posto (p) e par origem/destino (ij).

$$Fator2_{ijp} = Fator1_p \frac{Max(Est_{ijp})}{\Sigma(Est_{ijp})} \quad (2)$$

A terceira fase consiste na expansão das quantidades de passageiros da semana de entrevista para um ano completo. Este cálculo foi realizado por meio da análise das informações provenientes de radares do DNIT do período de janeiro de 2014 a maio de 2014. A análise foi baseada na frequência de veículos de passeio durante as semanas deste período em relação à semana da pesquisa por postos. Ou seja, a partir das informações de radares do DNIT, é possível definir o padrão de fluxo semanal de carros de passeio nos postos, e assim verificar se a semana que foi realizada a pesquisa foi uma semana atípica para expansão anual. Em casos que tenham sido verificados que na semana da pesquisa o fluxo foi maior ou menor que o usual, foi realizado uma correção na proporção de semanas a serem utilizadas para expansão anual. O cálculo deste ajuste é baseado nas médias das razões do fluxo de carros na semana capturados pelos radares do DNIT associados ao posto de pesquisa e o fluxo de carros obtidos pela pesquisa nos postos. A média dessas razões foi calculada por UF e assim foram multiplicadas pela quantidade de semanas do ano (aproximadamente 52,14 semanas) para obter a razão média anual ajustada.

Como alguns radares do DNIT não capturaram continuamente o fluxo de carros e as informações do DNIT utilizadas para análise estavam restritas apenas aos meses de janeiro de 2014 a maio de 2014, o ajuste de semanas no ano foi realizado apenas nas quantidades de semanas disponíveis, e, para as demais semanas, não houve ajustes. Dessa forma, para as semanas sem ajuste, foi considerado o número de semanas no ano usual (52,14 semanas no ano). Por exemplo, no estado de Sergipe, se na semana da pesquisa o fluxo de carros foi atípico com menos carros em média devido a um feriado, em uma razão média de 1,21 (valor médio de veículos semanal/valor de veículos na semana da pesquisa) e para o cálculo desta razão média foram utilizados catorze semanas de informações do DNIT, a quantidade de semanas no ano, neste caso, será de 55,14 semanas. Este valor é obtido por meio da soma das quantidades de semanas corrigidas e as demais semanas do ano (52,14 semanas menos a quantidade de semanas utilizadas). A correção das semanas utilizadas (DNIT) é calculada por meio multiplicação da razão média pela quantidade de semanas utilizadas (DNIT). Assim, o número ajustado de semanas no ano foi dado por $(1,2142 \cdot 14 + (52,14 - 14)) = 55,14$ semanas.

A tabela 2 mostra a média de semanas no ano por UF após correção da razão de veículos de passeio semanal do DNIT por veículos de passeio na semana da entrevista.

TABELA 2
Média de semanas no ano por UF

UF	Razão média semanal	Quantidade de semanas utilizadas (DNIT)	Fator expansão semana para ano ajustado
AC	1.03017	13	52.54
AL	1.05438	21	53.28
BA	1.09818	22	54.30
CE	1.07725	11	52.99
ES	1.08949	20	53.93
GO	1.08120	22	53.93
MA	1.08862	21	54.00
MG	1.08603	22	54.04
MT	1.06982	22	53.68
PA	0.99155	21	51.97
PB	1.08605	21	53.95
PE	1.11266	22	54.62
RN	0.98818	22	51.88
RO	1.07843	22	53.87
RS	1.03176	22	52.84
SC	1.01210	22	52.41
SE	1.21420	14	55.14
TO	0.99776	21	52.10

Elaboração dos autores.

Assim, o $Fator3_{ijp}$ é o produto entre o $Fator2_{ijp}$ e a média de semanas no ano (tabela 2), resultando na expansão anual da quantidade de passageiros por par O/D.

$$Fator3_{ijp} = Fator2_{ijp} (Fator\ expansão\ semana\ para\ ano\ ajustado) \quad (3)$$

O fator de expansão final, $Fator3_{ijp}$, resume uma expansão da amostra para a população total corrigida pelo valor máximo de passageiros entre os postos de cada par O/D que foram identificados em mais de um posto de pesquisa. Este fator representa também a expansão da semana da pesquisa para um ano completo, considerando as variações no tráfego de veículos devido a questões de sazonalidade.

3.4 Balanceamento dos pares O/D

Sabe-se que a quantidade de passageiros que saem de uma origem e vão para um determinado destino, por qualquer que seja o motivo – viagens a trabalho, para o trabalho, a passeio, entre outros –, retornam ao seu local de origem. Assim, a fim de assegurar essa suposição, foi realizado um balanceamento entre as zonas de origem e as zonas de destino em relação à quantidade de passageiros que fizeram a viagem.

Esse balanceamento foi realizado da seguinte forma: foi calculado o valor máximo entre cada par origem/destino e destino/origem e, em seguida, foi replicado o valor máximo para ambos os pares. Por exemplo, se a quantidade de passageiros com origem A para o destino B foi 100, e com origem B e destino A foi 300, então, o valor máximo entre eles é 300. Assim, para o par com origem A e destino B, a quantidade de passageiros foi substituída pelo valor 300.

Em seguida, foram analisados se existiam pares de zonas que só tinham *ida* ou só tinham *volta*, e foi detectado que nem todas as zonas estavam completas com os seus respectivos *pares*. Para os casos que só tinham a contagem de um sentido, o valor foi replicado para o outro sentido. Em outras palavras, se tinha somente o sentido O/D com o valor 20 e não tinha o sentido D/O, foi criado o sentido D/O e replicado o valor 20 de quantidade de passageiros que fizeram a viagem neste sentido. Assim, com esse ajuste, foi garantida a validação da suposição citada anteriormente de que, aproximadamente, a mesma quantidade de passageiros que saem de uma origem para um determinado destino, retorna a sua origem.

Ainda na linha de ajuste dos dados, com intuito de reduzir ao máximo o viés dos resultados que normalmente há em pesquisas, foi analisada a presença de pares O/D únicos e quais consequências negativas haveriam ao mantê-los na análise. Definiu-se como conceito de par O/D único aquele par com origem O e destino D com quantidade de apenas um (1) veículo entrevistado na pesquisa de campo rodoviária. Foi verificado que a presença destes pares O/D causam distorções nos resultados das análises. Um dos prejuízos seria no momento da expansão dos dados da amostra para representar a população, em que esse único veículo de passeio poderia ser multiplicado por um fator muito alto, gerando, assim, uma quantidade de carros para aquele par O/D que não condiz com a realidade. Desse modo, foram excluídos os pares O/Ds únicos da base de dados, ficando somente os pares O/Ds que tiveram mais de um (1) veículo entrevistado.

4 TRATAMENTO DAS BASES DE DADOS DE OUTROS MODOS DE TRANSPORTE

Além das informações de rodovia, descritas até o momento, para a matriz origem/destino de viagens de passageiros completa, foram consideradas também viagens de ônibus e aeroviárias para a elaboração da matriz origem/destino de viagens de passageiros.

O transporte hidroviário de passageiros, muito utilizado na região Norte, não foi considerado neste estudo por inviabilidade de se executar pesquisa de campo no mesmo período das demais. Este modo de transporte será objeto de pesquisa futura.

A matriz de passageiros de ônibus foi construída com dados secundários e dados da pesquisa realizada pela Agência Nacional de Transportes Terrestres (ANTT) anteriormente. A metodologia para gerar esta matriz é diferente da utilizada para passageiros de veículo de passeio, portanto não foi possível utilizar a pesquisa rodoviária para suprir as viagens de ônibus. As pesquisas de ônibus foram realizadas em terminais rodoviários, enquanto que as de passageiros de veículos de passeio foram realizadas nas margens das rodovias. A quantidade de passageiros de ônibus foi estimada por um modelo estatístico, no qual considerou viagens interestaduais e intermunicipais.

A matriz origem/destino de passageiros de transporte aéreo foi gerada por pesquisa em aeroportos realizada pela EPL. As viagens de ônibus e aeroviárias serão descritas com mais detalhe nos próximos itens.

4.1 Dados de ônibus

Para obter a matriz origem/destino de passageiros de ônibus foi necessário considerar passageiros tanto em viagens de ônibus interestadual quanto em viagens intermunicipais (intraestadual). Sendo assim, a quantidade de passageiros de viagens interestaduais foi extraída da pesquisa realizada pela ANTT, e parte das viagens intermunicipais foram compiladas por dados obtidos pela EPL junto às secretarias estaduais de transporte e órgãos correlatos. Como os dados de viagens intermunicipais foram disponibilizados para apenas algumas Unidades de Federação (UF), para as demais foi construído um modelo estatístico que estimou as viagens intermunicipais baseada nas viagens observadas das UFs disponíveis.

Segue descrição das informações das viagens intraestaduais e interestaduais utilizadas neste estudo.

4.1.1 Viagens intraestaduais

Para viagens intraestaduais de ônibus, ou seja, viagens entre municípios de mesma UF, foi possível capturar informações de cinco estados: Espírito Santo (ES); Goiás (GO); Minas Gerais (MG); Paraná (PR) e São Paulo (SP). Para cada um desses estados, foram utilizados dados de fontes diferentes. No Espírito Santo, o número de passageiros transportados de ônibus intermunicipal dentro do estado, no ano de 2013, foi fornecido pelo estudo de tráfego disponibilizado pelo Departamento de Estradas de Rodagem do Espírito Santo (DER/ES).

No caso do estado de Goiás, os dados foram fornecidos pela Agência Goiana de Regulação (AGR/GO) que são referentes ao transporte de passageiros intermunicipal relativo ao ano de 2012. Esses dados apresentam a frequência de ônibus em cada linha, o total de passageiros por linha e o total de número de viagens, por sentido – origem e destino – em cada linha.

Com respeito ao estado de Minas Gerais, os dados fornecidos são advindos do Plano Mineiro de Desenvolvimento Integrado do Estado de Minas Gerais. Estes dados apresentam o número de passageiros por par origem/destino no transporte intermunicipal do estado de Minas Gerais no ano de 2011.

Assim como no caso do Espírito Santo, no Paraná os dados de passageiros transportados dentro do estado foram fornecidos pelo Departamento de Estrada de Rodagem (DER/PR). Os dados apresentam o número de passageiros por par origem/destino do transporte intermunicipal do estado do Paraná no período de novembro de 2012 a outubro de 2013.

Para o estado de São Paulo, foram fornecidos pela Agência de Transporte do Estado de São Paulo (Artesp) os dados de passageiros no transporte intermunicipal, com o número de viagens realizadas no estado de São Paulo no ano de 2012.

As informações de origem e destino dos passageiros de ônibus intermunicipal desses cinco estados foram agregadas numa base de dados única e o número de viagens por par origem/destino foi balanceado da mesma forma que na base de rodovia na subseção 3.4 deste relatório, com a diferença do critério de escolha, que neste caso, foi a média da quantidade de ida e volta para cada par origem/destino.

4.1.2 Viagens interestaduais

Foi utilizada a matriz origem/destino de passageiros de ônibus interestadual produzida pela ANTT para o ano base de 2009, com base em entrevistas em terminais rodoviários de ônibus, fornecida pela EPL. O número de passageiros de ônibus interestadual por par origem/destino foi balanceado da mesma forma que na base de rodovia na subseção 3.4 deste relatório.

O modelo estatístico para estimar as viagens intermunicipais (exceto das UFs ES, GO, MG, PR e SP) foi realizado por nível hierárquico da pesquisa Região de Influência das Cidades (Regic), do IBGE. Os resultados deste modelo estão detalhados no relatório seguinte. A construção da matriz origem/destino de passageiros de ônibus observada foi finalizada somando as viagens interestaduais e as viagens intraestaduais estimadas.

4.2 Dados aeroviários

A matriz origem/destino de passageiros de transporte aéreo considerada neste estudo foi gerada por pesquisa em aeroportos realizada pela EPL no ano base 2014. Assim como os passageiros de ônibus e de veículos de passeio, o número de passageiros de avião por par origem/destino foi balanceado (como no item 3.4), utilizando o critério de escolha do valor médio das quantidades de viagens entre ida e volta para cada par origem/destino.

5 JUNÇÃO DAS BASES DE DADOS (MODOS DE TRANSPORTE)

A partir da realização das fases apresentadas anteriormente, foi elaborada a base final de passageiros, na qual foram somados os passageiros de veículos de passeio, ônibus e avião por origem. Como as viagens por cada modo de transporte foi balanceada, ou seja, a quantidade de passageiros na ida (zona X para zona Y) foi igual à volta (zona Y para zona X), a base final por destino é idêntica à por origem, portanto não houve a necessidade de serem trabalhadas as zonas de destino.

Utilizando a base final de passageiros por origem, foram testados modelos de regressão linear múltipla para projetar viagens futuras. Explicações sobre o modelo de regressão e resultados podem ser encontradas no próximo relatório.

Ipea – Instituto de Pesquisa Econômica Aplicada

EDITORIAL

Coordenação

Cláudio Passos de Oliveira

Supervisão

Everson da Silva Moura

Reginaldo da Silva Domingos

Revisão

Clícia Silveira Rodrigues

Idalina Barbara de Castro

Leonardo Moreira Vallejo

Marcelo Araujo de Sales Aguiar

Marco Aurélio Dias Pires

Olavo Mesquita de Carvalho

Regina Marta de Aguiar

Alessandra Farias da Silva (estagiária)

Lilian de Lima Gonçalves (estagiária)

Luiz Gustavo Campos de Araújo Souza (estagiário)

Paulo Ubiratan Araujo Sobrinho (estagiário)

Pedro Henrique Ximendes Aragão (estagiário)

Editoração

Bernar José Vieira

Cristiano Ferreira de Araújo

Daniella Silva Nogueira

Danilo Leite de Macedo Tavares

Jeovah Herculano Szervinsk Junior

Leonardo Hideki Higa

Herllyson da Silva Souza (estagiário)

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

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**

