

# **ESTIMAÇÃO SINTÉTICA DE MATRIZES ORIGEM/DESTINO A PARTIR DE CONTAGENS VOLUMÉTRICAS EM ÁREAS COM CONTROLE DO TRÁFEGO EM TEMPO REAL COM O AUXÍLIO DO QUEENSOD**

**Dante Diego de Moraes Rosado e Souza**

**Carlos Felipe Grangeiro Loureiro**

Programa de Mestrado em Engenharia de Transportes - PETRAN  
Universidade Federal do Ceará - UFC

## **RESUMO**

O objetivo geral deste trabalho foi estimar as demandas de deslocamento do transporte individual motorizado em uma área dentro da região controlada em tempo real pelo CTAFOR, utilizando o modelo de estimação sintética de matriz OD que se baseia em contagens volumétricas do tráfego, implantado no QueensOD. Após a calibração desse modelo, foi realizada uma análise de sensibilidade dos resultados de estimação frente à variação da qualidade dos dados de entrada (matriz OD semente e fluxos de tráfego). Como principal conclusão, tem-se que a obtenção de resultados satisfatórios de estimação depende fundamentalmente do fornecimento de uma matriz OD semente com qualidade.

## **ABSTRACT**

The main objective of this research work was to estimate motorized person trips patterns in an area within a region under control of Fortaleza's UTC system, using a synthetic OD matrix estimation model which is based in traffic flow volumes, running on QueensOD software. After calibrating the model, it was developed a sensitivity analysis on the estimation results due to the variation in the quality of input data (OD seed matrix and traffic volumes). As the main conclusion, it can be pointed out that obtaining satisfactory estimation results depends basically on a high quality OD seed matrix.

## **1. INTRODUÇÃO**

Atualmente, a cidade de Fortaleza possui 515 interseções semaforizadas, sendo 218 (42% do total) controladas a partir de uma central de tráfego em área (CTAFOR) por um sistema adaptativo em tempo real, que tem como base o pacote computacional SCOOT. Acredita-se que a partir de um entendimento mais apurado das características espaciais e temporais dos deslocamentos realizados pode-se aperfeiçoar o gerenciamento do tráfego urbano na área monitorada por esse sistema. Normalmente, os deslocamentos são representados por matrizes bidimensionais (matrizes OD), em que cada elemento da matriz representa o desejo de deslocamento entre uma origem e um destino. No caso particular do controle operacional do tráfego, é necessário que essas matrizes OD sejam confiáveis e de fácil e rápida obtenção (Paramahamsan, 1999). Nesse contexto, os métodos tradicionais de estimação de matrizes OD, que consistem na realização de entrevistas em domicílio ou no emprego, além de pesquisas de placa ao longo da rede viária, não são eficazes, devido às dificuldades no seu processo de coleta (altos custos e demora) e devido ao caráter estático de sua obtenção, que não permite representar a variabilidade do comportamento dos deslocamentos.

Diante do exposto, somando-se à facilidade de obtenção de dados de contagens volumétricas, já que, para garantir a operação em tempo real, o CTAFOR possui aproximadamente 630 laços detectores que registram durante todo o dia a demanda veicular que trafega pelo conjunto de vias da área controlada, acredita-se que o método de estimação sintética de matriz OD a partir de contagens volumétricas seja uma ferramenta viável e eficiente, capaz de prover subsídios para uma melhor compreensão dos deslocamentos realizados, em diferentes períodos de pico do tráfego, e conseqüentemente para um melhor gerenciamento operacional do tráfego. Deste modo, o objetivo geral deste trabalho foi estimar as demandas de deslocamento do transporte individual motorizado em uma área dentro da região controlada em tempo real pelo CTAFOR, utilizando o método de estimação sintética de matriz OD que se baseia em contagens volumétricas do tráfego, aproveitando a malha de laços detectores existentes nessa região. Como objetivos específicos, teve-se: i) delimitar e caracterizar uma região de estudo de caso quanto à oferta do sistema viário e demanda de deslocamento; ii) obter uma matriz OD da região de estudo, dos deslocamentos do transporte individual motorizado (apenas automóvel),

para uma determinada hora de pico, utilizando métodos tradicionais de coleta; iii) calibrar o modelo de estimação de matrizes OD a partir de contagens volumétricas, implantado no pacote computacional QueensOD, para a região do estudo de caso; e iv) fazer uma análise de sensibilidade dos dados de saída do modelo a partir da variação da qualidade dos dados de entrada, identificando a eficiência do modelo para as situações analisadas.

Este trabalho está estruturado em seis itens, incluindo esta introdução. No item 2, é apresentada uma breve revisão bibliográfica sobre os métodos sintéticos de estimação de matrizes a partir de contagens volumétricas, focando no desenvolvimento dos métodos que utilizam em sua formulação a teoria da maximização da entropia proposta por Wilson (1970). No item 3, é apresentada a área de estudo, enquanto no item 4 é descrito o processo de obtenção da matriz OD a partir de um método tradicional. No item 5, tem-se a aplicação do método de estimação de matrizes OD sintéticas com o pacote computacional QueensOD. Por fim, no item 6, são apresentadas as conclusões e recomendações para futuros trabalhos.

## **2. ESTIMATIÇÃO SINTÉTICA DE MATRIZES OD**

Os métodos sintéticos de estimação de matrizes OD a partir de contagens volumétricas vêm sendo desenvolvidos no decorrer das últimas quatro décadas (Abrahamsson, 1998; Van Aerde *et al.*, 2003) a partir do trabalho pioneiro de Robillard (1975). Estes métodos podem ser interpretados como o inverso das técnicas de alocação de tráfego (Bertoncini, 2007), pois, ao invés de obter os dados de fluxos ao longo da rede a partir da alocação de uma matriz OD, eles buscam obter (estimar) uma matriz OD que, quando alocada na rede, reproduza os dados de fluxo observados (Abrahamsson, 1998).

O desenvolvimento desses métodos ao longo dessas décadas sempre buscou complementar ou substituir a estimação realizada pelos métodos tradicionais, que apresentam uma série de desvantagens, como destacado anteriormente. A função da utilização dos métodos sintéticos é, portanto, suprimir essas desvantagens por meio da utilização de informações de fluxos de tráfego ao longo da rede viária, que é uma fonte de dados bastante atrativa, visto que sua obtenção é realizada com custos relativamente baixos, podendo ser feita de forma automática, no caso de redes monitoradas por uma central de controle de tráfego, requerendo pouca quantidade de recursos humanos e pouco interferindo nas realizações das viagens (Ortúzar e Willumsen, 1994). Apesar da existência de diversos estudos que abordam essas técnicas, Demarchi e Bertoncini (2004) citam que tais métodos não são conhecidos e nem amplamente utilizados por parte dos pesquisadores e analistas de sistemas de transporte, não só aqui no Brasil como também no exterior. Van Aerde *et al.* (2003) mencionam que os motivos que contribuem para isso são, além do desconhecimento mais detalhado da base teórica dos métodos, as dificuldades práticas encontradas no processo de estimativa das demandas entre os pares OD. Segundo Rakha *et al.* (1998), esses problemas estão relacionados principalmente, com: i) a possibilidade de obtenção de múltiplos resultados para a matriz OD, a partir de um mesmo conjunto de dados; ii) a inconsistência dos dados de fluxos de tráfego, que não permite a replicação dos fluxos observados em campo; e iii) ao desconhecimento das rotas utilizadas para os diversos deslocamentos realizados.

Dentre as linhas de pesquisa desenvolvidas, destacam-se as que se baseiam no princípio da maximização da entropia proposto por Wilson (1970) e que foi incorporado pela primeira vez no processo de estimação de matrizes OD por Willumsen (1978). A premissa dessa teoria é que a matriz OD com maior entropia será também a matriz OD com maior probabilidade de ocorrer em campo. Deste modo, Willumsen (1978) estimou a matriz OD mais provável a partir da maximização da entropia e de um conjunto de restrições de fluxo do tráfego como apresentado a seguir nas Equações 1, 2, 3 e 4. Na utilização dessa formulação, alguns pré-requisitos ainda foram necessários, como a existência da continuidade dos fluxos de tráfego e o conhecimento prévio da utilização das rotas entre um par OD.

$$\text{MAXIMIZAR } S(T_{ij}) = \frac{T!}{\prod_{ij} T_{ij}!} \quad (1)$$

$$\text{sujeito a: } V_a = \sum_{ij} T_{ij} \cdot P_{ij}^a \quad (2)$$

$$T_{ij} \geq 0 \quad (3)$$

$$T = \sum_{ij} T_{ij} \quad (4)$$

em que:  $T$ : total de viagens na matriz;  
 $T_{ij}$ : quantidade de viagens de  $i$  para  $j$ ;  
 $V_a$ : volume observado (coletado) no arco  $a$  da rede;  
 $P_{ij}^a$ : proporção das viagens de  $i$  para  $j$  que utiliza o arco  $a$  da rede.

Ao longo dos anos, diversas outras formulações foram propostas baseadas na linha de pesquisa iniciada por Willumsen (1978). Van Zuylen e Willumsen (1980) buscaram minimizar os problemas referentes à falta de continuidade dos fluxos de tráfego considerando o comportamento estocástico dessa variável, a partir da técnica de maximização da verossimilhança. Willumsen (1984) propôs, em seguida, uma alteração em sua formulação inicial (Equação 5) incorporando informações de uma matriz OD existente (matriz semente). Neste caso, a utilização de uma matriz semente garantiu que as demandas fossem estimadas mantendo a proporcionalidade das demandas-semente em relação ao total de viagens da matriz semente.

$$\text{MAXIMIZAR } S(T_{ij}) = \frac{T!}{\prod_{ij} T_{ij}!} \prod_{ij} \left( \frac{t_{ij}}{\sum_{ij} t_{ij}} \right)^{T_{ij}} \quad (5)$$

em que  $t_{ij}$ : quantidade de viagens de  $i$  para  $j$  na matriz semente.

Até então, os algoritmos de solução desenvolvidos eram fundamentados na premissa de que a utilização das rotas para um determinado deslocamento deveria ser definida *a priori*, independentemente do processo de estimação da matriz. Essa premissa limitava, portanto, a utilização dos métodos de estimação em situações de não congestionamento da rede. Segundo Ortúzar e Willumsen (1994), em situações nas quais a existência de congestionamento influencia a escolha das rotas, essa premissa é questionável, visto que a matriz O-D e o processo de escolha das rotas tornam-se interdependentes.

Embora estudos anteriores tivessem proposto a solução do problema de estimação da matriz incorporando o conceito de equilíbrio do usuário (Nguyen, 1977; Willumsen, 1984), foi Fisk (1988) que forneceu a solução mais robusta, estendendo o modelo de Van Zuylen e Willumsen (1980) para o caso de redes congestionadas, a partir da introdução do conceito de equilíbrio do usuário como restrição. Na formulação de Fisk, a teoria da maximização da entropia e a alocação de equilíbrio do usuário foram combinadas em um único pacote computacional, com as rotas e as demandas nos pares OD sendo definidas em um processo iterativo em dois níveis, repetido até a obtenção de valores consistentes para ambos. A maioria das formulações que tratam o congestionamento no processo de estimação de matrizes a partir de contagens volumétricas possui uma estrutura em dois níveis, como a apresentada por Fisk, na qual, em um nível superior, a matriz OD é estimada e, em um nível inferior, a utilização das rotas é definida em função da última matriz estimada (Abrahamsson, 1998).

Na tentativa de tornar disponível de forma prática o método de estimação de matrizes OD a partir de dados de fluxos de tráfego, diversos modelos foram implantados computacionalmente. Com destaque, podemos citar os modelos implementados nos pacotes computacionais

QueensOD (M. Van Aerde & Assoc., 2002) e TransCAD (Caliper, 2002). A formulação implantada no QueensOD, que foi utilizada neste trabalho para estimação de matrizes OD, baseia-se na linha de pesquisa iniciada por Willumsen (1978). Segundo M. Van Aerde & Assoc. (2002), o objetivo foi criar um pacote computacional que desse suporte ao microssimulador Integration, fornecendo uma matriz OD capaz de ser alocada na rede microssimulada. Para tanto, os dois pacotes computacionais compartilham a mesma estrutura de dados e tipo de arquivos de entrada, não sendo necessário, entretanto, possuir o Integration para utilizar o QueensOD na estimação de matrizes OD. A formulação implantada no QueensOD considera o comportamento estocástico da variável fluxo de tráfego e a interdependência entre a estimação da matriz OD e o processo de escolha das rotas em cada par OD. A seguir, é apresentada a Equação 6, a ser maximizada pelo QueensOD no processo de estimação, em que a primeira parcela representa a verossimilhança entre a matriz estimada e a matriz semente e a segunda representa o erro entre os fluxos observados e estimados (Van Aerde *et al.*, 2003). Nesta formulação, a restrição de fluxo definida pela Equação 2 foi incorporada na Equação 6, aplicando-se o multiplicador Lagrangeano ( $\lambda_{ij}$ ). Além de maximizar a Equação 6, a formulação implantada no QueensOD busca também minimizar os erros entre as demandas sementes e demandas estimadas e entre os fluxos observados e os fluxos estimados como apresentado na Equação 7. O QueensOD estabelece três formas para calcular o erro em que  $w$  será igual a: 1 (erro quadrático);  $\sqrt{X_i^{obs}}$  (erro poissoniano); e  $X_i^{obs}$  (erro relativo). A dissertação de Paramahansan (1999) descreve em detalhes a formulação matemática e o algoritmo de solução do QueensOD.

$$\text{MAXIMIZAR } S_{ij} = TLn \frac{T}{t} - T - \sum_{ij} \left( T_{ij} Ln \left( \frac{T_{ij}}{t_{ij}} \right) - T_{ij} \right) + \sum_{ij} \lambda_{ij} \cdot 2 \left( \sum_a (V_a p_{ij}^a) - \left( \sum_a p_{ij}^a \left( \sum_{ij} T_{ij} p_{ij}^a \right) \right) \right) \quad (6)$$

$$E = \sum_i \left( \frac{X_i^{obs} - X_i^{est}}{w} \right)^2 \quad (7)$$

### 3. ÁREA DO ESTUDO DE CASO

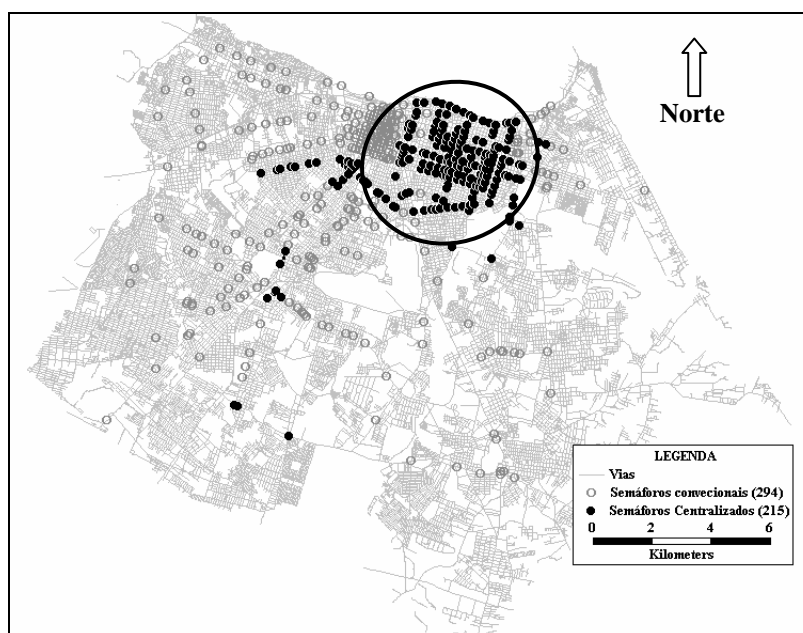
A primeira etapa buscando alcançar os objetivos estabelecidos foi delimitar uma área de estudo de caso que fosse representativa da região controlada em tempo real. A idéia ao aplicar o método de estimação de matrizes OD implantado no QueensOD, para essa área de estudo, foi permitir uma maior compreensão desse método e obter conclusões e recomendações para estimativas posteriores de matrizes OD para toda a região controlada em tempo real, considerando diferentes períodos de pico do tráfego.

Na Figura 1, é apresentado o mapa do município de Fortaleza com a disposição espacial das interseções e o tipo de controle semafórico. Optou-se por determinar a área do estudo de caso dentro da região em destaque na Figura 1 e detalhada na Figura 2a. Essa região, além de possuir o mesmo tipo de controle semafórico, caracteriza-se também pela presença de congestionamento e, conseqüentemente, pelo uso de múltiplas rotas entre dois pontos da rede. Na Figura 2b, é apresentada a área do estudo de caso, que compreende uma rede de 15 interseções, sendo oito semaforizadas e sete não-semaforizadas, e é limitada pela Rua Pereira Filgueiras (ao norte), Avenida Santos Dumont (ao sul), Rua Carlos Vasconcelos (ao oeste) e Avenida Barão de Studart (ao leste). Com exceção da Avenida Barão de Studart, que é uma via de sentido duplo, com duas faixas por sentido segregadas por canteiro central, as demais vias da rede são de sentido único com duas faixas de tráfego. Em todas as vias, o transporte coletivo compartilha as faixas de tráfego com o transporte individual.

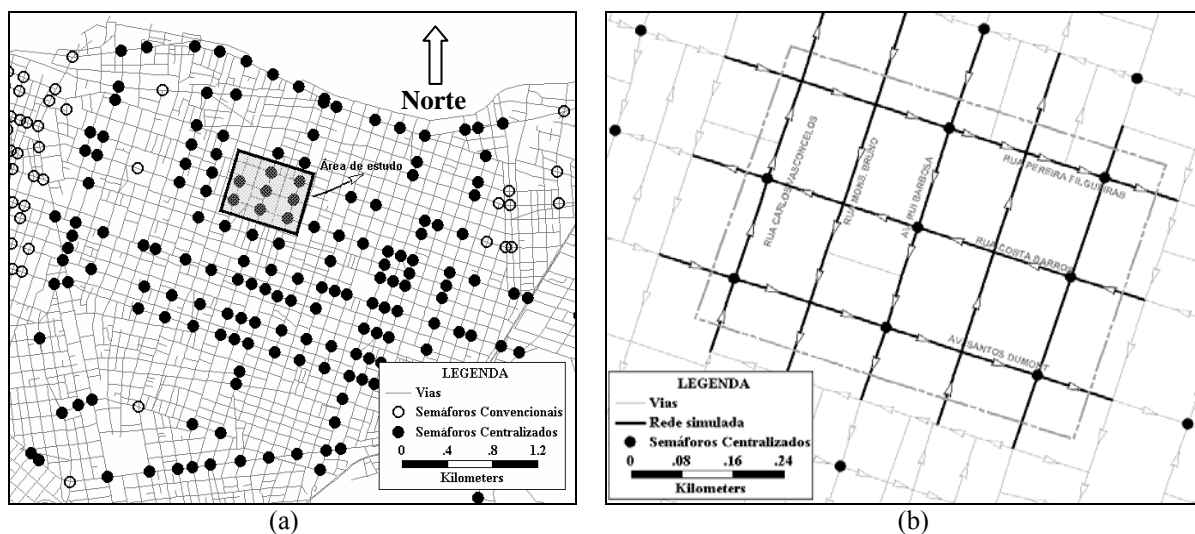
### 4. MATRIZ OD OBSERVADA

A etapa seguinte foi planejar e executar a coleta da matriz OD dessa área de estudo a partir de algum método tradicional. O objetivo da obtenção dessa matriz foi poder confrontar os resultados de estimação realizados pelo QueensOD com uma matriz observada em campo,

permitindo assim uma avaliação da eficiência dessa estimação. O método tradicional escolhido para obter a matriz OD foi o método de leitura de placa. Em uma etapa anterior ao planejamento da pesquisa, foi necessário realizar a codificação da rede, identificando os pontos de origem e destino, já que os pontos de observação da pesquisa de placa devem ser definidos em função do zoneamento proposto. Por ser uma região relativamente pequena, optou-se por desagregar ao máximo a rede. Deste modo, cada ponto de entrada e saída foi classificado como ponto de origem e destino, respectivamente, e cada quadra foi classificada como um ponto de origem e destino. A Figura 3a apresenta a rede codificada. Observa-se que em todas as quadras existe um centróide ligado à rede a partir de um conector. Após a codificação da rede, foram identificados oito pontos de entrada na rede, classificados como origem, oito pontos de saída classificados destinos, e 25 pontos representando as quadras, classificados como origem e destino, totalizando, no final, 33 origens e 33 destinos.



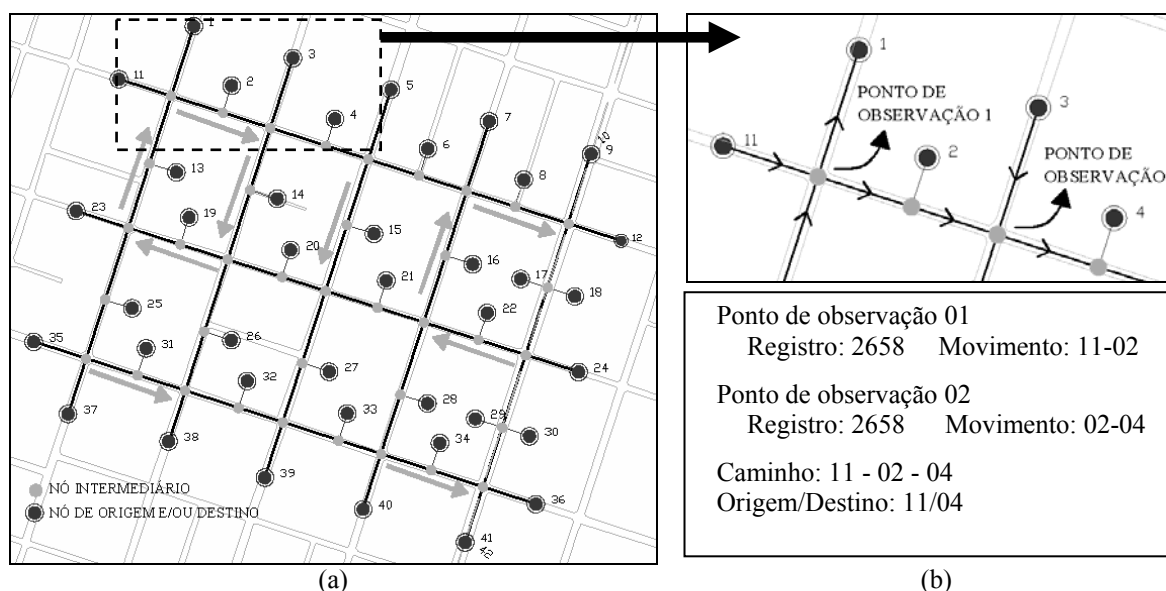
**Figura 1:** Tipos de controle das interseções semaforizadas do município de Fortaleza



**Figura 2:** Área do estudo de caso

A pesquisa de placa foi realizada no dia 07/06/06 (quarta-feira), durante o pico da tarde, entre 17h00 e 18h30, sendo dividido em seis intervalos de 15 minutos. Devido à dificuldade em se registrar manualmente todas as placas dos veículos, o registro das placas foi feito de forma

amostral. Foram registrados apenas os automóveis de cor prata, que, como identificado no banco de dados do Detran-CE, representam cerca de 20% de toda a população. É importante destacar ainda que o registro das placas foi feito de forma parcial, sendo anotados apenas os dígitos numéricos. Para a realização da pesquisa de leitura de placa foram necessários 28 pesquisadores, sendo definidas, como ponto de observação, todas as 15 interseções da área de estudo. Após a realização da pesquisa, os dados foram tabulados. Nessa tabulação, cada registro foi identificado com o ponto de observação onde o registro foi realizado, com o intervalo de coleta e com o movimento de conversão realizado na interseção (aproximação de origem e de destino). Após a tabulação dos dados, estes foram processados com o auxílio de uma macro desenvolvida no ambiente VBA (*Visual Basic for Application*), compondo o caminho realizado pelos veículos e determinando conseqüentemente as suas origens e destinos. Na Figura 3b, é apresentado um exemplo da obtenção da origem e destino de uma determinada placa. Neste caso, o veículo cruza o ponto de observação 1, com origem no nó 11 e destino ao nó 2. Em seguida, esse mesmo veículo cruza o ponto de observação 2, com origem no nó 11 e destino ao nó 4, não sendo mais identificado em nenhum outro ponto. Compondo o caminho desse veículo, determina-se a sua origem, nó 11, e destino, nó 4.



**Figura 3:** Codificação da rede (a) e composição dos caminhos das placas registradas (b)

Após o processamento dos dados, foram obtidas as matrizes OD amostrais dos seis intervalos de 15 minutos definidos anteriormente, que apresentaram demanda amostral horária variando entre 2.170 e 2.670 viagens. Nessas matrizes, de um total de 1089 pares OD possíveis (33 origens x 33 destinos), foram observadas viagens apenas em 312 pares OD (29% do total). Para análise posterior da estimação da matriz OD, foi considerada somente a faixa horária entre 17h15 e 18h15, sendo excluídos o primeiro e o último intervalo. Esses intervalos desconsiderados foram incluídos na pesquisa de placa apenas para identificar possíveis viagens que tiveram início ou término fora do período estabelecido para análise. Obtida a matriz OD amostral, foi realizada em seguida a expansão dessa matriz a partir do método de Furness (Ortúzar e Willumsen, 1994), com dupla restrição (origens e destinos). No final do processo de expansão, foi obtida uma matriz OD populacional com demanda total horária de aproximadamente 11.000 viagens.

## 5. ESTIMAÇÃO DE MATRIZES OD SINTÉTICAS COM O QUEENSOD

Além dos dados de fluxos observados nos arcos da rede, o programa QueensOD necessita para realizar a estimação de matrizes OD, de uma série de informações relativas às características geométricas, de tráfego e de controle da rede a ser simulada, sendo todas essas informações

codificadas em arquivos no formato texto (Demarchi *et al.*, 2004). Em conjunto com os dados de fluxos observados, são fornecidos também os dados de capacidade e de tempos de viagem, utilizados para compor as rotas utilizadas entre os diversos pares OD. É necessário informar para o QueensOD tanto o tempo de viagem de fluxo livre, obtido a partir da velocidade de fluxo livre do arco, quanto o tempo de viagem observado, obtido a partir da velocidade média do tráfego no arco durante a coleta dos dados de fluxo. Para os arcos da rede controlados por semáforos (24 de um total de 117), os dados de fluxo de tráfego e tempo de viagem foram obtidos de forma automática a partir de consultas ao banco de dados dinâmico do CTAFOR, com o auxílio da interface TRANSCOOT (Meneses, 2003). Nos demais arcos da rede, não foram coletadas informações adicionais de fluxo de tráfego, sendo necessário fornecer apenas o tempo de viagem de fluxo livre, que foi estimado em função da velocidade máxima da via e da sua capacidade. Para iniciar as estimações das matrizes, foi estipulado, ainda, o número máximo de iterações possíveis (1.000 iterações), o critério de convergência para os erros de demanda e fluxos (0,01) e a quantidade máxima de rotas entre cada par OD (cinco rotas).

### 5.1. Calibração do modelo de estimação

Com todos os dados obtidos e com os arquivos de entrada codificados, foi iniciado o processo de calibração do modelo de estimação do QueensOD. Esta etapa de calibração consistiu em verificar apenas quais critérios de erros (Equação 7) são os mais adequados para a rede simulada, já que os demais parâmetros do modelo são fixos. Apesar de Demarchi *et al.* (2004) terem recomendado a utilização do erro relativo para minimizar as diferenças entre as viagens das matrizes OD semente e estimada, e o erro quadrático para minimizar as diferenças entre fluxos observados e estimados, esta calibração se justifica, pois no trabalho de Demarchi *et al.* a estimação da matriz OD foi realizada sem o fornecimento de uma matriz OD semente, sendo a eficiência da estimação analisada apenas em função da replicação dos fluxos observados. Para a calibração do modelo foram definidos, inicialmente, três cenários distintos quanto à matriz OD semente a ser fornecida ao QueensOD. O primeiro cenário consistiu em fornecer uma matriz OD semente binária, com demanda igual a um, entre os pares OD que tiveram viagens identificadas na pesquisa de placa, e demanda igual a zero, caso contrário. No segundo cenário, a matriz OD semente foi a matriz OD amostral, obtida durante a pesquisa de placa. Já no terceiro cenário, foi fornecida a matriz OD expandida como matriz semente. Em cada cenário definido foram realizadas nove estimações, com combinações diferentes de critérios de erro (quadrático, poissoniano e relativo), utilizados para avaliar as estimações das demandas e dos fluxos. Em seguida, os resultados dessas estimações foram avaliados a partir da técnica de regressão linear, sendo as matrizes OD estimadas confrontadas com a matriz OD expandida, considerada como matriz OD observada na população dos deslocamentos realizados.

Analisando os resultados da estimação, quanto ao critério de erro utilizado para avaliar a diferença entre fluxos, observou-se que os resultados divergem pouco entre si, mostrando que para a rede de estudo esse parâmetro não influencia significativamente na qualidade da matriz estimada. Observou-se ainda que, em todas as estimações realizadas, o erro médio entre os fluxos estimados e observados foi próximo de zero. A Tabela 1 mostra os resultados do coeficiente de determinação ( $R^2$ ) e, entre parênteses, o coeficiente da reta de regressão para esses testes, considerando a utilização do erro quadrático para o cálculo do erro entre os fluxos estimados e observados. Em todos os testes, foi utilizado um grau de confiança de 95% e foi avaliada a nulidade do valor do intercepto da reta de regressão.

**Tabela 1:** Resultados dos testes de regressão para calibração do tipo de erro a ser utilizado

Tipo de erro	Matriz semente		
	Binária	Amostral	Expandida
<b>Quadrático</b>	0,20 (0,31)	0,50 (0,49)	0,97 (1,01)
<b>Poissoniano</b>	0,21 (0,31)	0,78 (0,74)	0,96 (1,01)
<b>Relativo</b>	0,24 (0,40)	0,88 (1,01)	0,93 (1,01)

Percebe-se, analisando a Tabela 1, que no cenário em que é fornecida uma matriz OD semente binária, os resultados dos testes de regressão foram semelhantes. Comparando ainda as matrizes OD estimadas, não se percebe diferenças significativas, mostrando que, neste caso, o tipo de erro utilizado pouco interfere no resultado de estimação. Isto ocorre pois o fornecimento de uma matriz OD semente constante faz com que o QueensOD maximize a Equação 6, dando mais importância à segunda parcela da equação, que garante a continuidade dos fluxos de tráfego, e minimize apenas os erros entre os fluxos estimados e observados. Já no cenário da matriz OD semente amostral, os resultados diferiram entre si, sendo a estimação que utilizou o erro relativo a única que apresentou um resultado satisfatório, com um coeficiente de determinação igual a 0,88, confirmando para este caso a recomendação feita por Dermachi *et al.* (2004). No último cenário, e de forma semelhante ao primeiro, as matrizes OD estimadas, comparadas entre si, apresentaram pequenas diferenças, mostrando que o tipo de erro utilizado pouco interfere no resultado de estimação.

## **5.2. Análise de sensibilidade dos resultados quanto à qualidade da matriz OD semente**

Avaliando os resultados das estimações realizadas quanto à variação da matriz semente, ainda com base na Tabela 1, percebe-se que no cenário em que é fornecida uma matriz semente binária, as demandas da matriz OD estimada diferem consideravelmente das da matriz OD expandida. Os coeficientes de determinação encontrados para este cenário não foram satisfatórios, apresentando valores próximos de 0,20.

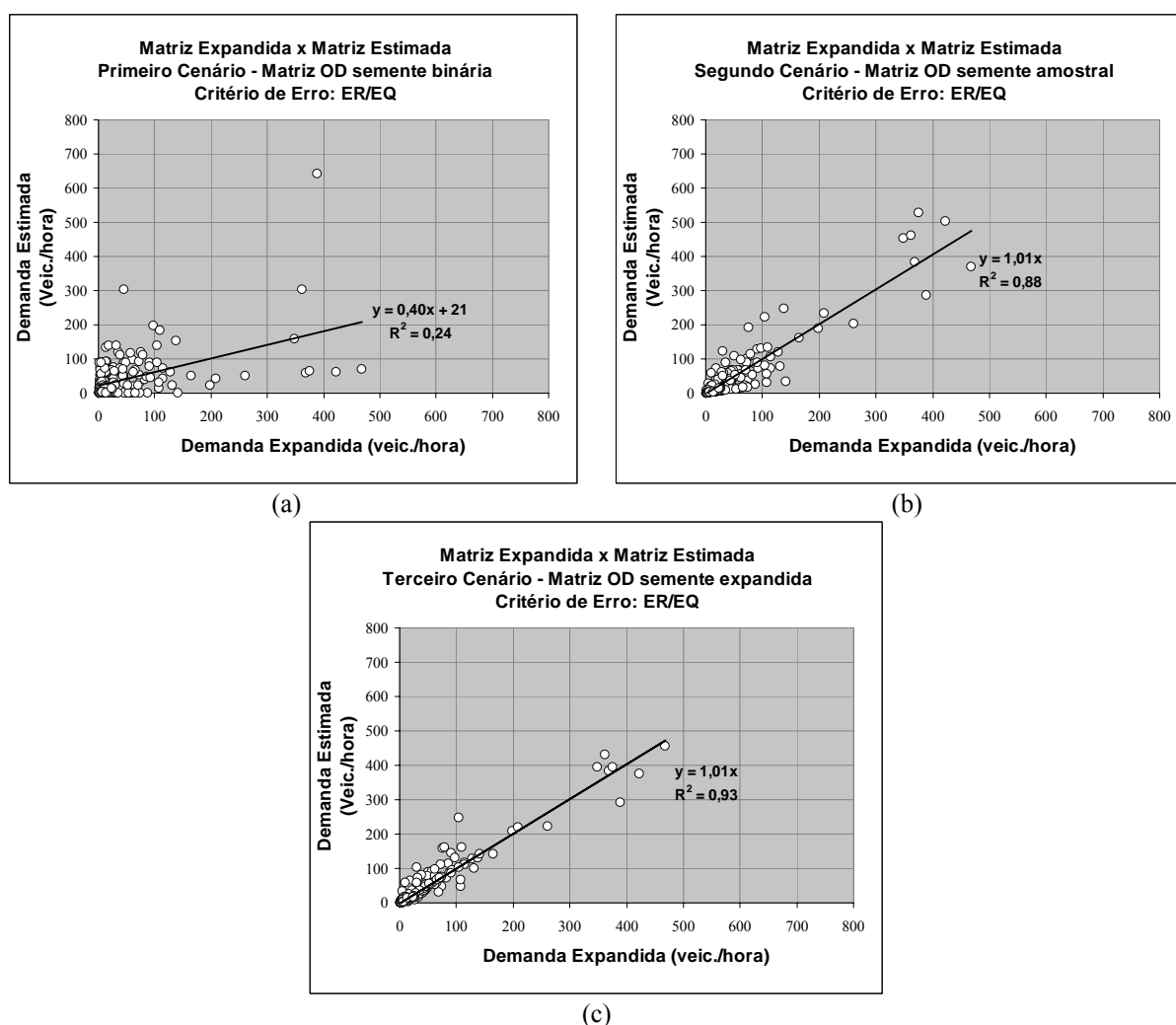
No cenário em que é fornecida uma matriz semente amostral, percebe-se a melhora na estimação da matriz OD em função do aumento da qualidade da matriz OD semente; entretanto, como exposto anteriormente, apenas na situação em que foi utilizado o erro relativo para avaliar as diferenças entre as demandas, pode-se considerar o resultado como satisfatório. O último cenário, que na realidade é um cenário fictício, pois não há interesse em se estimar uma matriz OD a partir da própria matriz observada em campo, apresentou os melhores resultados com os coeficientes de determinação e coeficiente da reta de regressão próximos de 1,0. Analisando-se ainda o total de viagens das matrizes OD estimadas, verificou-se que em todas as matrizes esse valor ficou em torno de 11.000, que é um valor semelhante ao encontrado na matriz OD expandida.

A seguir, na Figura 4, são mostrados os gráficos de dispersão, com as respectivas retas de regressão, entre as demandas da matriz OD estimada e as da matriz OD expandida (observada), considerando o cálculo do erro quadrático entre os fluxos estimado e observado, para os três cenários distintos quanto a matriz semente fornecida. Observando a Figura 4a, que representa o cenário em que é fornecida uma matriz semente binária, é possível perceber a divergência entre as demandas estimadas e observadas. Nas estimações deste cenário, foi observado um erro médio absoluto entre as demandas variando em torno de 30 viagens e erro máximo absoluto igual a 400 viagens. Na Figura 4b, percebe-se a melhora na estimação da matriz OD. Neste caso, foi fornecida uma matriz OD semente amostral e os erros médio e máximo absoluto encontrados foram, respectivamente, iguais a 13 e 150 viagens. Na Figura 4c, que representa o cenário em que é fornecida a própria matriz OD observada como matriz semente, tem-se o melhor resultado de estimação, com erro médio absoluto igual a oito viagens e erro máximo absoluto igual a 140 viagens.

Conclui-se que para uma estimação satisfatória de uma matriz OD é necessário o fornecimento de uma matriz semente de qualidade, que forneça informações sobre o padrão de deslocamento da população. Apesar de uma determinada matriz OD ser a mais provável de ocorrer em campo, de acordo com o princípio da entropia proposto por Wilson (1970), não necessariamente ela será a matriz que melhor representará, na realidade, o comportamento de deslocamento de uma população, mesmo replicando os fluxos observados. Não há nenhuma



garantia de que essa matriz estimada será a matriz que ocorrerá realmente em campo, como ocorreu no resultado de estimação quando foi fornecida uma matriz OD semente binária. Já na estimação realizada quando é fornecida uma matriz OD semente amostral, apesar de as demandas-sementes possuírem magnitudes diferentes das da matriz OD expandida, elas possuem a mesma estrutura, ou seja, as proporções de viagens entre cada demanda e o total de viagens ( $t_{ij}/\sum t_{ij}$ ) são semelhantes. A partir da Equação 5, que originou a função a ser maximizada pelo QueensOD, é possível explicar essa constatação. No fornecimento de uma matriz semente, considera-se que a probabilidade de ocorrência de uma viagem nessa matriz será a mesma que na matriz a ser estimada. Deste modo, uma determinada demanda entre uma origem  $i$  e um destino  $j$  que tenha uma pequena proporção na matriz semente ( $t_{ij}/\sum t_{ij}$ ), ou seja, baixa probabilidade de ocorrer, tenderá a assumir valores pequenos na matriz estimada. Isto acontece porque, quando uma demanda com baixa probabilidade de ocorrer aumenta na matriz OD a ser estimada, o valor da função objetivo (Equação 5) tende a assumir valores menores. A matriz que maximiza a função objetivo tenderá a assumir valores menores para demandas com baixa probabilidade de ocorrer e valores maiores para demandas com alta probabilidade.

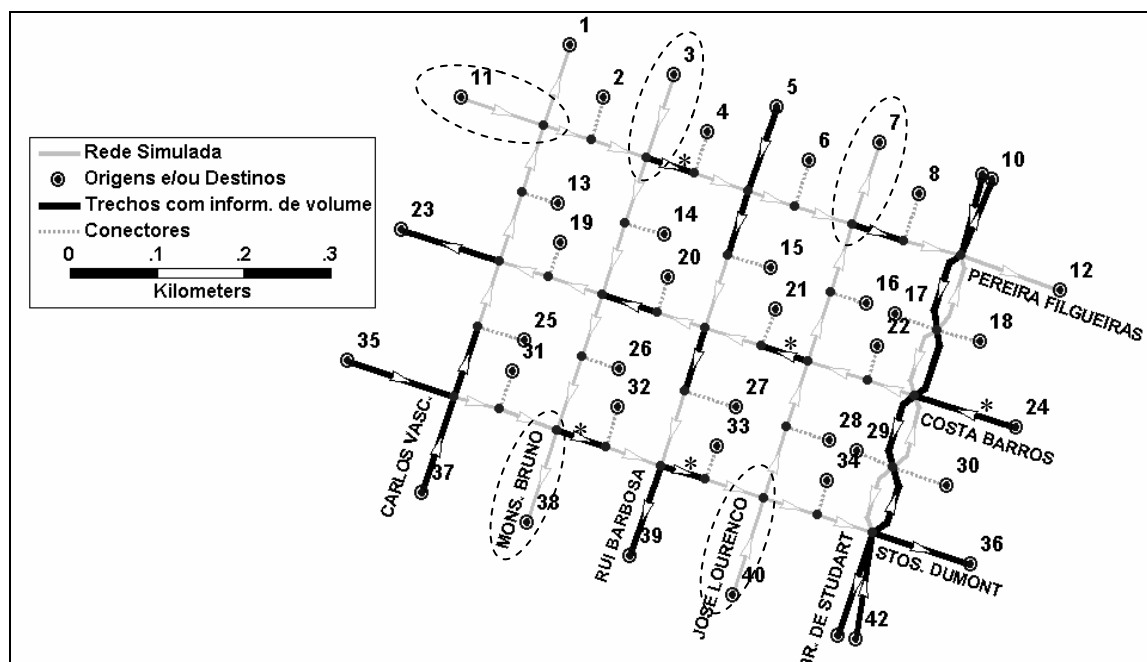


**Figura 4:** Resultados de estimação para os três cenários de matriz semente

### 5.3. Análise de sensibilidade dos resultados quanto à qualidade dos dados de fluxos

A etapa seguinte da análise de sensibilidade foi a verificação dos resultados quanto à quantidade de informações de fluxos de tráfego. As estimações realizadas até então tinham utilizado como informação de fluxo do tráfego todos os dados dos laços de detecção existentes na área de estudo (24 no total – ver Figura 5). A partir desse cenário, foram definidos mais dois cenários, um aumentando a quantidade de arcos com dados de fluxo (segundo cenário) e outro

diminuindo essa quantidade (terceiro cenário). Nas estimações realizadas nesta análise de sensibilidade, foi fornecida a matriz OD semente amostral, obtida na pesquisa de placa, e foram utilizados os critérios de erro que apresentaram resultados satisfatórios na etapa de calibração (erro quadrático para fluxo e erro relativo para demanda).



**Figura 5:** Arcos da rede simulada que possuem laços de detecção

Na definição do segundo cenário, o conjunto de informações de fluxos existentes foi complementado com dados coletados a partir de pesquisas volumétricas, em arcos da rede que pertencessem às rotas entre os pares OD que não possuíam nenhuma informação. Esses arcos estão destacados na Figura 5 por elipses. Na etapa de calibração, foi observado que o QueensOD desconsidera, no processo de estimação, os pares OD que não possuem informação de fluxos nas rotas utilizadas, ou seja, para estes pares, a demanda na matriz estimada é igual à na matriz semente. Deste modo, era esperado que essas novas informações contribuíssem para aprimorar a qualidade da matriz estimada. O terceiro cenário é fictício e serviu apenas para avaliação do comportamento do modelo de estimação das demandas frente a essas configurações de dados de fluxos. Foram definidas três configurações de vias distintas quanto à quantidade de arcos com informações de fluxos de tráfego. A primeira configuração foi definida para a Rua Pereira Filgueiras, que possuía dois arcos com dados de fluxo e passou a ter apenas um arco com essa informação. Uma outra configuração foi definida para a Rua Costa Barros, que possuía informação em quatro arcos, incluídos os arco de entrada e saída da via, e passou a ter informações em apenas dois arcos. Neste caso, foi retirada a informação de fluxo de um arco de saída, com o fim de avaliar a importância dessa informação no processo de estimação. A última configuração foi definida para a Avenida Santos Dumont, em que as informações de dois arcos intermediários foram excluídas, ficando apenas informações de fluxo nos arcos de entrada e saída dessa via. Os arcos que tiveram a informação extraída do processo de estimação estão identificados na Figura 5 com um asterisco (\*).

Os resultados de estimação das demandas para o segundo cenário apresentaram no geral uma pequena melhora. Enquanto o coeficiente de determinação da regressão entre a demanda estimada e a expandida foi igual a 0,88 no primeiro cenário, no segundo cenário esse coeficiente foi igual a 0,91. Os resultados mostraram ainda que o erro médio absoluto para o segundo cenário baixou de 13 para 10 viagens e o erro absoluto aumentou de 151 para 161 viagens. Verifica-se que os ganhos no resultado final foram muito pequenos. Entretanto,

quando se analisa separadamente as demandas dos nós que são diretamente afetados pela falta de informação de fluxo, percebe-se um ganho mais significativo. Nas demandas com origem no nó 40, o erro médio absoluto e o erro máximo reduziram-se, respectivamente, de 17 para quatro viagens e de 57 para nove viagens. Nas demandas com origem nas zonas 3 e 11, ganhos significativos também foram observados.

Os resultados de estimação das demandas para o terceiro cenário também não apresentaram, no geral, diferenças significativas em relação aos resultados do primeiro cenário. Os coeficientes de determinação da regressão entre as demandas estimadas e expandidas para este cenário foi igual ao do primeiro cenário. O erro absoluto médio encontrado foi o mesmo sendo igual a 13 viagens e o erro máximo aumentou de 151 para 161 viagens. Destaca-se aqui que as demandas que deixaram de ter informações de fluxos nas rotas utilizadas foram desconsideradas pelo modelo de estimação e tiveram seus resultados de estimação prejudicados. Analisando, ainda, algumas demandas separadamente, como no caso das demandas originadas nos nós 11, 24 e 35, zonas mais afetadas pela retirada de informações de fluxos de tráfego, observou-se que os resultados de estimação para as zonas 11 e 35 obtiveram pequenas variações, diferente dos resultados de estimação encontrados na zona 24. No caso das zonas 11 e 35, os erros médios absolutos entre as demandas estimadas e expandidas se mantiveram: na zona 11 esse erro foi igual a 11 viagens e na zona 35 igual a 15 viagens. Os erros máximos também mantiveram os mesmos valores. Já para a zona 24, o erro médio absoluto aumentou de 12 para 21 viagens e o erro máximo aumentou de 81 para 134 viagens. A justificativa para uma diferença significativa nos resultados de estimação das demandas na zona 24, está no fato de o arco, que teve a informação de fluxo retirada, ser um arco de saída da zona. Deste modo, os dados de fluxo desse arco já informavam ao modelo de estimação a quantidade de viagens originadas nessa zona, facilitando o processo de estimação. Com a retirada dessa informação de fluxo, o modelo deixou de possuir essa informação e, conseqüentemente, estimou com menos qualidade as demandas originadas nessa zona.

## **6. CONCLUSÕES**

Este trabalho apresentou o processo de estimação de matrizes OD com o auxílio do modelo implantado no QueensOD, para uma área de estudo controlada em tempo real pelo CTAFOR. Na etapa de calibração desse modelo, foi confirmada a recomendação feita por Dermachi *et al.* (2004) quanto à utilização do erro relativo para tratar das diferenças entre as demandas estimadas e observadas. Destaca-se, entretanto, que no caso em que é fornecida uma matriz OD semente constante, a definição deste critério torna-se irrelevante, visto que o modelo de estimação passa a desconsiderar esse erro em seu processo de estimação. Os resultados desta calibração indicaram ainda que a definição de qual critério de erro a ser utilizado para avaliar as diferenças entre os fluxos estimados e observados é irrelevante para o caso da rede da área de estudo. Constatou-se que, neste caso, variando-se o critério de erro utilizado, não era possível observar variações significativas nos resultados de estimação. Esta constatação pode ser explicada pelo fato de o modelo de alocação do QueensOD ter conseguido representar bem as escolhas de rotas, visto que os resultados de estimação dos fluxos para todas as simulações realizadas apresentaram um erro médio absoluto próximo de zero.

Quanto à qualidade das estimações, os resultados mostraram que o não fornecimento de uma boa matriz semente, mesmo com uma grande quantidade de arcos com dados de fluxos, não permite o modelo estimar uma matriz adequada. Neste caso, apesar de a matriz OD estimada ser a matriz com maior entropia, ela não será necessariamente a matriz observada em campo. Por outro lado, o fornecimento de uma matriz amostral, com magnitudes diferentes da matriz OD observada, entretanto com uma estrutura semelhante, apresentou resultados satisfatórios, mostrando que o modelo é capaz de expandir uma matriz OD amostral obtida por algum outro método.

Outra conclusão importante refere-se ao fato de o QueensOD não considerar no processo de estimação os pares OD que não possuem informação de fluxo de tráfego nas rotas utilizadas. Por fim, os resultados mostraram que, para uma boa estimação de todas as demandas, é necessário obter informações de fluxo em pelo menos um arco de todas as rotas utilizadas. Uma outra observação importante refere-se às informações em arcos de entrada e saída da rede, que, como verificado, contribuem significativamente com a melhora dos resultados de estimação, visto que são informações diretas sobre a quantidade de viagens originadas ou destinadas para uma determinada zona.

Como recomendações para estudos futuros têm-se: i) avaliar os resultados de estimação do modelo de estimação do QueensOD para uma rede mais complexa do que a utilizada nesta pesquisa, em que os efeitos de congestionamento sejam mais evidentes e que, conseqüentemente, o modelo de alocação do tráfego tenha mais dificuldades em estimar as rotas utilizadas; ii) verificar com mais detalhe a correlação entre a quantidade de informações de fluxos de tráfego nos arcos das rotas e a qualidade da estimação das demandas; e iii) identificar e implantar algum método eficiente na coleta de matrizes OD amostrais que permita, associado às vantagens em possuir uma central de controle do tráfego em tempo real, estimar matrizes OD para diferentes períodos e, conseqüentemente, avaliar a variabilidade da demanda.

#### REFERÊNCIAS BIBLIOGRÁFICAS

- Abrahamsson, T. (1998) *Estimation of Origin-Destination Matrices Using Traffic Counts – A Literature Survey*. International Institute for Applied Systems Analysis, Interim Report IR- 98-021, Austria.
- CALIPER (2002) *Travel Demand Modeling With TransCAD 4.5*. Caliper Corporation, USA.
- Bertoncini, B.V. (2007) *Uma Proposta de Carregamento Incremental de Fluxos Veiculares para a Estimação de Matriz O-D Sintética*. Dissertação de Mestrado, EESC/USP.
- Demarchi, S. H. e Bertocini, B. V. (2004) Determinação de Matrizes O-D Sintéticas a partir de Contagens Volumétricas. *Anais do XVIII Congresso de Pesquisa e Ensino em Transportes*, ANPET, Florianópolis.
- Demarchi, S. H.; Bertocini, B. V. e Lima, E. P. (2004) Estimativa de Demandas Sintéticas para a Região Central de Maringá Utilizando o QueensOD. *Anais do XVIII Congresso de Pesquisa e Ensino em Transportes*, ANPET, Florianópolis.
- Fisk C. S. (1988) On Combining Maximum Entropy Trip Matrix Estimation with User Optimal Assignment. *Transport Research*, 22B.
- Nguyen, S. (1977) *Estimating an OD Matrix from Network Data: A Network Equilibrium Approach*. Publication 87. Centre de Recherche sur les Transports, Université de Montreal.
- Meneses, H. B. (2003) Interface Lógica em Ambiente SIG para Bases de Dados de Sistemas Centralizados de Controle de Tráfego Urbano em Tempo Real. *Dissertação de Mestrado, Programa de Mestrado em Engenharia de Transportes*, Universidade Federal do Ceará.
- M. VAN AERDE & ASSOC.,LTD. (2002) *QueensOD Rel. 2.10 User's Guide: Estimating Origin Destination Traffic Demands from Links Flow*. M. Van Aerde & Assoc.,Ltd, Ontario, Canada.
- Ortúzar, J. de D.; Willumsen, L. G. (1994) *Modelling Transport*. Jonh Wiley & Sons, Chichester.
- Paramahamsan, H. (1999) *Fundamental Properties of Synthetic O-D Generation Formulations and Solutions*. M.Sc. Thesis, Department of Civil and Environmental Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA.
- Rakha, H.; Paramahamsan, H.; Van Aerde, M. (1998) Static Synthetic O-D Estimation Formulations, Assumptions and Limitations. *Transportation Research Board*. The 77<sup>th</sup> Annual Meeting, Washington, DC, USA.
- Robillard, P. (1975) Estimating the O-D Matrix from Observed Link Volumes. *Transportation Research Part B*, n.9, p.123-128.
- Van Aerde, M.; Rakha, H. e Paramahamsan, H. (2003) Estimation of OD matrices: the Relationship between Practical and Theoretical Considerations. *Transportation Research Board*. The 82<sup>nd</sup> Annual Meeting, Washington, DC, USA.
- Van Zuylen, H. e Willumssen L. G. (1980) The Most Likely Trip Matriz Estimated from Traffic Counts. *Transportation Research*, 14B.
- Willumssen L. G. (1978) Estimation of an O-D Matrix from Traffic Counts: A review. *Institute for Transport Studies*, Working paper no. 99, Leeds University.
- Willumssen L. G. (1984) Estimating Time- Dependent Trip Matrices from Traffic Counts. *Ninth International Symposium on Transportation and Traffic Theory*.
- Wilson A. G. (1970) *Entropy in Urban and Regional Modelling*. Pion, London, England.

**Endereço:** Dante Diego de M. R. e Sousa (dante@det.ufc.br) e Carlos Felipe G. Loureiro (felipe@det.ufc.br)