Discussão conceitual sobre métodos de reconstrução de matrizes origem-destino estáticas em redes de transportes

Francisco Moraes de Oliveira Neto[1], Anselmo Ramalho Pitombeira Neto[2],
Carlos Felipe Grangeiro Loureiro[3] e Bruno Vieira Bertoncini[4]

Resumo: Matrizes origem-destino (OD) quantificam a demanda por transporte em uma região geográfica e constituem peças-chave em estudos de planejamento, projeto e gerenciamento de sistemas de transportes urbanos e regionais. Tradicionalmente, matrizes OD são obtidas por meio da realização de levantamentos amostrais diretos, tais como: pesquisas domiciliares, contagem de placas de veículos e entrevistas em terminais de transporte público. Uma alternativa de menor custo consiste em *sintetizar* uma matriz OD por meio de métodos matemáticos com o uso de dados provenientes de contagens de tráfego feitas na rede de transportes em estudo, os quais podem ser classificados em métodos de *reconstrução* e de *estimação*. Este artigo apresenta uma discussão conceitual com foco nas premissas e limitações dos principais métodos de reconstrução baseados em maximização de entropia, mínimos quadrados generalizados e inferência bayesiana. Descreve-se detalhadamente a fundamentação matemática dos métodos e são feitas recomendações para o avanço e aplicação efetiva dos mesmos.
Palavras-chave: Modelagem da Demanda por Transportes, Matrizes Origem-Destino Sintéticas, Reconstrução de Matrizes Origem-Destino, Estimação de Matrizes Origem-Destino.

Abstract: Origin-destination (OD) matrices quantify the demand for transport in a geographic region, and play a key role in planning studies, design and management of urban and regional transport systems. Traditionally, the estimation of OD matrices consists of making direct sample surveys, such as: household interviews, vehicle plate counting and interviews in public transport terminals. A lower cost alternative is to synthesize an OD matrix by means of mathematical methods using traffic volumes observed in a transport network, which can be classified into methods of reconstruction and estimation. This paper presents a conceptual discussion focused on the assumptions and limitations of the main reconstruction methods based on maximizing entropy, generalized least squares and Bayesian inference. We describe in detail the mathematical foundation of the methods and make recommendations to their improvement and effective application.
Keywords: Transport Demand Modeling, Synthetic Origin-Destination Matrices, Reconstruction of Origin-Destination Matrices, Estimation of Origin-Destination Matrices.

1. INTRODUÇÃO

A obtenção da matriz origem-destino (OD) dos fluxos de viagens por meio de métodos diretos, como pesquisa domiciliar, é onerosa, demorada e frequentemente não resulta em um padrão representativo da população dos deslocamentos na área em estudo. Nestes métodos, apenas uma pequena proporção de viagens é observada para um dia típico do tráfego, o que resulta em muitos pares OD nulos ou com poucas viagens na amostra, composta por uma única realização dos fluxos OD. Uma alternativa para se obter uma matriz OD representativa do comportamento das viagens, em princípio de menor custo e mais eficiente, é a sua *sintetização* a partir de dados provenientes de contagens volumétricas de tráfego. Embora este problema já venha sendo estudado há mais de quatro décadas, os modelos ou abordagens sintéticas, baseadas em volumes nos arcos ou nós da rede de transporte, ainda não são devidamente empregados pela comunidade técnica. Uma das principais barreiras à aplicação efetiva desses modelos é a deficiência na compreensão dos objetivos, premissas e limitações das suas diversas formulações propostas na literatura.

Os primeiros métodos de obtenção da matriz OD sintética (Robillard, 1975) tiveram por base unicamente a relação entre os volumes de tráfego nos arcos da rede de transportes e os respectivos fluxos OD que os geraram. Esta relação forma um sistema de equações que em geral é indeterminado, ou seja, existem múltiplas soluções (matrizes OD) possíveis. Por este motivo, buscou-se desenvolver modelos de otimização que encontrassem a matriz *mais provável* dentre aquelas consistentes com os volumes de tráfego observados, podendo ou não utilizar informação adicional (Van Zuylen e Willumsen, 1980). Neste caso, a função-objetivo é derivada de algum princípio sobre o fenômeno, ou adota-se alguma função pragmática para encontrar a matriz única.

Embora a literatura destaque que os modelos de otimização são parte integrante dos métodos de estimação da matriz OD sintética, não fica claro o que estes métodos buscam estimar: se a matriz dos fluxos médios; ou se uma matriz para um dado período de tempo em que volumes nos arcos da rede foram observados. Quando os volumes são observados em um único período, pode-se argumentar que os métodos tentam encontrar uma aproximação para a matriz que gerou aqueles volumes de tráfego, ou seja, os métodos tentam reconstruir a matriz que gerou os volumes observados (Hazelton, 2001). Contudo, dependendo das premissas sobre o processo que gera a matriz de viagens – que tanto pode ser

[1] **Francisco Moraes de Oliveira Neto**, Universidade Federal do Ceará, Campus de Russas – UFC/Russas (oliveiranefm@gmail.com).

[2] **Anselmo Ramalho Pitombeira Neto**, Departamento de Engenharia de Produção, Universidade Federal do Ceará – DEPRO/UFC (anselmoufc@gmail.com).

[3] **Carlos Felipe Grangeiro Loureiro** (felipe@det.ufc.br).

[4] **Bruno Vieira Bertoncini**, Departamento de Engenharia de Transportes, Universidade Federal do Ceará – DET/UFC (bruviber@det.ufc.br).

Manuscrito recebido em 30/09/2015 e aprovado para publicação em 10/03/2016.

Este artigo é parte de TRANSPORTES v. 24, n. 1, 2016 ISSN: 2237-1346 (online). DOI:10.14295/transportes.v24i1.101

entendido como determinístico ou probabilístico – considera-se, no primeiro caso, que a matriz de interesse é uma matriz constante ou que representa o conjunto dos fluxos médios; já no caso probabilístico, deseja-se realmente estimar, no sentido propriamente estatístico, a média (ou qualquer outro parâmetro) da distribuição populacional dos fluxos OD.

Com o intuito de contribuir para uma melhor compreensão da modelagem deste problema, o presente trabalho traz uma discussão conceitual sobre os principais métodos de reconstrução de matrizes OD estáticas, baseados nos conceitos de maximização da entropia, mínimos quadrados generalizados e inferência bayesiana. Embora na última década, especialmente após o advento de sistemas de controle e monitoramento do tráfego em tempo real, a atenção da comunidade científica tenha se voltado para os métodos de estimação de matrizes OD dinâmicas (Pitombeira-Neto, 2015), justifica-se ainda discutir a formulação e a aplicação dos métodos de reconstrução estática já que suas premissas e objetivos não estão abordados de forma clara e eficaz na literatura, especialmente no que diz respeito a como tais métodos tratam questões relativas à aleatoriedade do fenômeno, ao papel da matriz prévia e ao efeito da qualidade e quantidade de dados de volumes de tráfego observados na rede de transportes.

Para tanto, este artigo está estruturado da seguinte forma: na Seção 2 propomos uma discussão teórica sobre a formulação do problema de obtenção de matrizes OD estáticas sob as óticas determinística e probabilística; a Seção 3 apresenta uma revisão sobre os métodos de reconstrução, na qual analisamos suas formulações matemáticas e premissas adotadas; já na Seção 4, discutimos questões conceituais que consideramos fundamentais acerca das limitações e aplicabilidade dos métodos; por fim, a Seção 5 conclui o artigo com recomendações e perspectivas de estudos futuros em reconstrução de matrizes OD sintéticas.

2. PROBLEMA DE OBTENÇÃO DE MATRIZES OD SINTÉTICAS ESTÁTICAS

Nesta seção, o problema de obtenção de matrizes OD estáticas a partir de contagens de tráfego é analisado sob duas perspectivas: determinística e probabilística. Na primeira, adota-se a premissa de que os indivíduos, ao realizarem suas viagens, se comportam de maneira aproximadamente constante. Isto significa que não há muitas alterações nas decisões de viagens no seu dia a dia (se deslocar ou não, instante de partida, destino e rota escolhida). Assim, uma pesquisa domiciliar para apenas um dia de observação seria suficiente para caracterizar o padrão típico de viagens de um grupo de indivíduos. Essa era a premissa adotada nos anos 1970 quando se iniciaram as primeiras pesquisas em métodos de obtenção de matrizes OD a partir de volumes de tráfego. Partindo-se dessa premissa determinística, pode-se afirmar que o objetivo dos métodos propostos originalmente era encontrar a matriz OD constante que gerou o conjunto de volumes típicos observados na rede de transportes.

Já na perspectiva probabilística, assume-se que as decisões de viagens variam consideravelmente dia a dia e que, portanto, o padrão de deslocamentos deve ser considerado um fenômeno aleatório (Vardi, 1996; Lo *et al.,* 1996; Hazelton, 2001). Esta premissa é consistente com as condições atuais de viagens no meio urbano, caracterizadas por uma diversificação das atividades socioeconômicas que resultam na realização de inúmeros deslocamentos com propósitos diversos. As próprias viagens com propósitos mais corriqueiros, como por motivo trabalho, vêm sofrendo variações quanto às decisões de deslocamento em alguns segmentos da população. Nesse contexto, os objetivos dos métodos de obtenção de matrizes OD sintéticas tanto podem ser de estimar fluxos médios e sua variação, como também de reconstruir a matriz OD para um período específico. As diferenças nas duas perspectivas transcendem o campo conceitual e tornam-se ainda mais evidentes na tentativa de representar matematicamente o fenômeno, conforme descrito e discutido a seguir.

2.1 Formulação Matemática

O problema matemático para obtenção de matrizes OD sintéticas é expresso pela condição de consistência dos volumes de tráfego observados na rede de transportes. Assim, para uma dada rede composta por I arcos conectando J pares de zonas de tráfego, a relação fundamental entre os fluxos OD e os volumes de tráfego nos arcos é expressa como na Equação (1):

$$\mathrm{v} = \mathrm{Mx} \tag{1}$$

em que $\mathbf{v}$ é um vetor de volumes de tamanho igual ao número de arcos com volumes observados; $\mathbf{M}$ é a matriz de utilização dos arcos, ou de alocação do tráfego na rede, em que um elemento m_{ij} é igual à proporção do fluxo no par OD j que usa o arco i; e $\mathbf{x}$ é um vetor de tamanho J que representa o padrão de deslocamento na rede, em que cada elemento x_j é o fluxo de viagens para o par OD j. A matriz de alocação $\mathbf{M}$ pode ser composta por dois fatores relacionados às possíveis rotas utilizadas para as viagens; isto é, de posse do conjunto de K rotas possíveis na rede, a relação (1) também pode ser expressa conforme a Equação (2):

$$\mathrm{v} = \Delta \mathrm{Px} \tag{2}$$

sendo $\mathbf{P}$ a matriz de escolha de rotas, na qual cada elemento p_{kj} representa a proporção da demanda j que utiliza a rota k; e Δ é a matriz de incidências arco-rota, cujos elementos δ_{ik} assumem valor 1 se o arco i pertence à rota k e 0 caso contrário. Note que o vetor de fluxos nas rotas, denotado por $\mathbf{h}$, é dado por $\mathrm{h} = \mathrm{Px}$.

Na maioria das redes de transportes existe uma indeterminação das Equações e , ou seja, o número de incógnitas (fluxos OD) a serem determinadas é maior do que o número de equações independentes (número de arcos com contagens de tráfego não redundantes). Por este motivo, as primeiras estratégias adotadas consistiam em resolver um problema de otimização a partir da introdução de uma função-objetivo estritamente convexa, para encontrar uma solução única dentro do conjunto viável de soluções definido pelo sistema de equações representado em . É importante ressaltar que o conjunto de soluções viáveis desse sistema pode ser vazio. Uma das possíveis causas para a inexistência de soluções viáveis é a inadequação do modelo de escolha de rotas adotado.

Tabela 1: Modelagem da matriz de alocação de tráfego

	Escolha de rotas			
	Determinística		Probabilística	
Matriz de Viagens	Rede sem restrição de capacidade	Rede com restrição de capacidade	Rede sem restrição de capacidade	Rede com restrição de capacidade
Determinística	*Constante*	*Estacionária*	*Constante*	*Estacionária*
Probabilística	*Constante*	*Variável*	*Constante*	*Variável*
Probabilística e dinâmica	*Constante*	*Variável*	*Constante*	*Variável*

Em relação à matriz **P**, a mesma é modelada de forma exógena, sendo interpretada como uma matriz de probabilidades de escolha de rotas. Neste caso, o vetor de volumes passa a ser um vetor de valores esperados de volumes, E(**v**) ou μ_v, e E(**v**) = Δ**Px**. Assim, o vetor **v** é uma variável aleatória que inclui variação devido à escolha das rotas para um dado vetor **x** de viagens. A premissa básica para estimar **P** é que os usuários buscam escolher rotas de menor caminho ou impedância de viagem, com a modelagem de escolha discreta sendo amplamente aplicada neste contexto. Prato (2009) apresenta uma revisão abrangente da literatura sobre os métodos de modelagem da geração do conjunto de alternativas e de escolha de rotas em uma rede.

Em função do tipo de consideração sobre o <u>fenômeno dos deslocamentos</u> (determinístico ou probabilístico), é possível caracterizar a matriz **P** como constante, estacionária ou variável. A situação em que **P** é considerada constante ocorre quando a rede não tem restrição de capacidade. Destacam-se também aqui dois casos do <u>fenômeno de escolha de rotas</u>. No primeiro, a escolha de rotas é um fenômeno determinístico que independe do nível de carregamento na rede, no qual os usuários têm pleno conhecimento do seu nível de serviço e utilizam rotas de menor caminho. Desta forma, **P** se reduz a uma matriz binária, caso em que apenas uma rota é usada para cada par OD. No segundo, existe uma incerteza por parte dos usuários sobre os níveis de desempenho ou custos dos arcos da rede, caso em que a escolha de rotas é probabilística. Em ambos os casos, o resultado da alocação é proporcional à matriz OD de viagens, conforme expresso em (2).

A matriz **P** é considerada <u>estacionária</u> quando a rede está congestionada (ou com restrição de capacidade) e a matriz de viagens é representada por um processo estático e determinístico. Em redes congestionadas, existe uma interdependência entre a escolha da rota e o carregamento na rede. Essa dependência ocorre porque o nível de serviço dos arcos da rede depende dos volumes nos próprios arcos. Neste contexto, assume-se que após um processo de aprendizagem sobre os custos das rotas, a rede atingiria um estado estacionário, ou de *equilíbrio*, correspondendo ao que é conhecido na literatura como equilíbrio determinístico do usuário (Wardrop, 1952; Smith, 1979) ou equilíbrio estocástico do usuário (Daganzo e Sheffi, 1977), para os casos em que o fenômeno de escolha de rotas seja considerado como determinístico ou probabilístico, respectivamente. Este estado estacionário corresponde a um contexto em que as probabilidades de escolha de rotas permanecem constantes, existindo então uma consistência entre os volumes médios e os custos médios nos arcos.

Em redes com restrição de capacidade, portanto, não é possível dissociar o problema de reconstrução de uma matriz OD do problema de obtenção de uma matriz de alocação. Vários autores têm proposto determinar a matriz de escolha de rotas com base no *princípio de equilíbrio da rede*. Logo, no caso de redes congestionadas, os modelos para reconstrução da matriz OD são resolvidos por meio de técnicas *biníveis*, nas quais se utiliza um algoritmo iterativo que alternadamente reconstrói a matriz OD e computa a matriz de escolha de rotas correspondentes, por meio de um método de alocação de tráfego. Sobre este tema, recomenda-se a leitura dos trabalhos de Fisk (1988, 1989), Yang *et al.* (1992) e Yang (1995). Recentemente, Bertoncini *et al.* (2013) propuseram um algoritmo com estrutura *binível* que, segundo os autores, não pressupõe equilíbrio na rede.

A matriz **P** é considerada <u>variável</u> quando a matriz OD de viagens é probabilística, do tipo estática (com parâmetros fixos) ou dinâmica (com parâmetros variando ao longo do tempo), e a rede apresenta restrição de capacidade com funções de desempenho não-lineares. Assim, **P** também varia conforme a variação dos fluxos OD. No caso da matriz OD ser resultado de um processo aleatório com parâmetros estáticos, a rede pode atingir um estado estacionário de probabilidade dos volumes e custos nos arcos, mas não um estado estacionário na matriz **P**. Este estado estacionário da variabilidade dos volumes e custos nos arcos corresponde ao conceito de equilíbrio estocástico geral, conforme proposto por Watling (2002a, 2002b). No caso dinâmico, a rede pode não atingir um estado de equilíbrio ou estacionário, uma vez que vários fatores se modificariam durante o tempo necessário para se atingir tal estado, fazendo com que o ponto de equilíbrio estivesse constantemente se deslocando. Vale ressaltar que em análises de padrões estocásticos de viagens, em curtos períodos de tempo, é difícil garantir a hipótese de equilíbrio na rede.

A modelagem dinâmica para obtenção de matrizes é um tema bastante recente. Como exemplo, Pitombeira-Neto (2015) propôs um modelo bayesiano que considera como dinâmico o padrão de deslocamento na rede, incorporando a incerteza na informação prévia de viagens. Conforme destacado anteriormente, o contexto dinâmico do fenômeno dos deslocamentos não é foco deste artigo, e, portanto, não será considerado nas discussões que seguem. A Tabela 1 resume as diferentes formas de modelagem da matriz de probabilidades **P**.

Cabe aqui ressaltar que a discussão proposta a seguir nas Seções 2.2 e 2.3 considera que os volumes observados nos arcos não contêm erros de observação, de forma que toda a aleatoriedade dos volumes de tráfego é proveniente da aleatoriedade dos fluxos OD e da escolha das rotas. Ademais, divide-se a discussão em duas partes: na primeira, assume-se que o fenômeno de viagens é determinístico (a matriz OD populacional é constante). Neste caso, é possível apenas reconstruir a matriz OD com base em volumes observados e em uma aproximação inicial do padrão de deslocamentos na área em estudo. Na segunda parte, considera-se que o padrão de viagens em si é aleatório ou probabilístico (a matriz OD é uma variável aleatória com parâmetros fixos ou estáticos), sendo possível tanto reconstruir uma dada matriz quanto estimar o padrão médio de viagens a partir de uma estimativa prévia da matriz média e de volumes observados nos arcos da rede.

2.2 Padrão de Viagens Determinístico

Seja a premissa de que o fenômeno dos deslocamentos, representado pelo vetor de fluxos OD **x**, é determinístico. A matriz de escolha de rotas **P**, como mostra a Tabela 1, pode ser modelada como constante ou estacionária. No caso de redes congestionadas, assume-se que os usuários estabelecem suas preferências por determinadas rotas como resultado de um processo de aprendizagem na rede, até atingir um estado de equilíbrio. Suponha que se esteja interessado em saber o número de viagens x_j no par OD j de uma região em estudo, na qual se conhece o total de viagens N. Uma alternativa adotada na prática é coletar uma amostra aleatória de viagens de tamanho n e observar quantas viagens no par OD j foram realizadas. Suponha também que a proporção populacional de viagens no par OD j seja p_j. Se as observações na amostra são independentes e n é muito pequeno em relação a N, então o número n_j de viagens do tipo j na amostra segue aproximadamente uma distribuição binomial com parâmetros n e p_j. Com a amostra de tamanho n, uma estimativa pontual do parâmetro populacional p_j é:

$$\hat{p}_j = \frac{n_j}{n} \tag{3}$$

Com média e variância dadas respectivamente por:

$$\mathrm{E}[\hat{p}_j] = p_j \tag{4}$$

$$\mathrm{Var}[\hat{p}_j] = \frac{\mathrm{Var}\ [n_j]}{n^2} = \frac{p_j(1-p_j)}{n} \tag{5}$$

Uma estimativa pontual para o número de viagens no par OD j a partir da amostra de tamanho n é:

$$\hat{x}_j = N\hat{p}_j \tag{6}$$

Para uma amostra suficientemente grande, o estimador dado pela Equação é aproximadamente normal, segundo o Teorema do Limite Central, com média Np_j e variância $N^2p_j(1-p_j)/n$. Este estimador é eficiente, consistente e não tendencioso. Um intervalo de confiança aproximado para x_j com base na amostra pode ser obtido por meio da substituição de p_j no cálculo da variância pela estatística amostral $\hat{p}_j$. O uso de volumes de tráfego observados na rede, neste contexto, tem como objetivo melhorar a estimativa pontual $\hat{x}_j$. Ou seja, incorporar a informação dos volumes, por meio da restrição expressa na Equação , reduzindo a variância da estimativa pontual e sua distância em relação ao valor populacional x_j. Portanto, a nova estimativa pontual ajustada pelos volumes de tráfego, denotado por $\breve{x}_j$, estaria mais próxima do valor real e também seria mais precisa.

Na literatura existem vários métodos de otimização que podem ser aplicados com este propósito (Cascetta, 2009). Uma característica importante de todos os métodos de otimização é que $\hat{x}_j$ é considerada uma matriz alvo para a matriz $\breve{x}_j$. Outra premissa importante é que se assume a suposição de que a matriz **P** é conhecida, ou que pode ser obtida como resultado de um processo binível. Vale ressaltar que os métodos de otimização consideram que os volumes observados em apenas um dia (considerado típico) são volumes médios na rede e que a escolha de rotas é determinística, fazendo com que os volumes alocados não variem ao longo do tempo. Um problema encontrado na prática é que o tamanho da amostra n de viagens pode não ser grande o suficiente para que uma dada interação espacial j seja observada. Desta forma, em geral o vetor amostral de viagens apresenta muitos vazios, ou seja, pares OD para os quais o número de viagens observadas é zero ($n_j = 0$). Assim, modelos de demanda por viagens são empregados para estimar o vetor de viagens com base em relações socioeconômicas e da oferta de transportes. Vale ressaltar neste ponto que os métodos de obtenção de matrizes OD sintéticas foram inicialmente desenvolvidos para "ajustar" uma matriz antiga (coletada em uma época anterior àquela que se desejava obter a matriz de viagens), que pode ser derivada de uma realidade completamente distinta da atual.

2.3 Padrão de Viagens Probabilístico

Considere agora o caso em que o fenômeno de viagens, representado pelo vetor **x**, é probabilístico. Neste caso, o número x_j de viagens do tipo j é uma variável aleatória com parâmetros estáticos, ou seja, o vetor **x** segue algum tipo de distribuição de probabilidade com vetor médio e matriz de variância-covariância constante e independente ao longo do tempo. Como os volumes nos arcos da rede são função de **x** (Equação (1)) e da escolha das rotas, estes também serão variáveis aleatórias. Neste contexto, a matriz **P** deve ser considerada constante ou variável (ver Tabela 1), em redes sem ou com restrição de capacidade, respectivamente. No caso de escolha probabilística de rotas e redes congestionadas, assume-se que os usuários tendem, no longo prazo, a escolher suas rotas de tal forma que ambos os volumes e custos nos arcos variem conforme uma distribuição de probabilidade estacionária, que por sua vez corresponde a um estado de equilíbrio estocástico geral, conforme definido por Watling (2002a, 2002b). Note que a definição tradicional de equilíbrio estocástico (Sheffi, 1985) considera que existe aleatoriedade apenas na percepção dos usuários sobre os custos da rede, enquanto que Watling sugere que pode existir um estado de equilíbrio estocástico geral quando se assume aleatoriedade também nos fluxos OD.

O primeiro passo para tratar este fenômeno é adotar alguma premissa sobre o processo de geração do vetor aleatório **x**. Por exemplo, considere a premissa que a variável

x_j resulta de um processo binomial com parâmetros N e p_j, em que N é uma variável aleatória que representa o total de viagens na área de estudo em um determinado dia, e p_j a probabilidade de uma determinada viagem ocorrer no par OD j. O valor esperado de x_j é dado por $\lambda_j = p_j \mathrm{E}[N]$. Uma estimativa da proporção média de viagens $\hat{p}_j$ para o par OD j pode ser obtida a partir de amostras de viagens na área de estudo por vários dias. Por sua vez, a estimativa do total esperado de viagens pode ser obtida a partir de modelos de demanda de viagens. Assim, tem-se um estimador do parâmetro médio $\hat{\lambda}_j$ das viagens no par OD j.

Nesse contexto específico, o uso da informação dos volumes de tráfego observados na rede pode ter dois objetivos: (a) estimar o padrão médio de viagens com informação *a priori* dada por $\hat{\lambda}_j$; e (b) reconstruir uma determinada matriz de um dado dia em que não se tem qualquer observação direta de viagens. Para o objetivo (a), os métodos de inferência estatística, em particular do tipo bayesiana, são os mais adequados. Na inferência bayesiana, o número médio populacional de viagens λ_j é tratado como uma variável aleatória, à qual se atribui uma distribuição de probabilidades *a priori*, representando a incerteza do analista ou crença com respeito a este padrão médio de viagens. O vetor de volumes observados é usado para atualizar a distribuição *a priori* para uma distribuição de probabilidades *a posteriori* por meio da aplicação do teorema de Bayes. Pode-se então definir um estimador bayesiano a partir de medidas de tendência central da distribuição *a posteriori*. Vale ressaltar que múltiplos dias de observação de volumes podem ser usados, gerando uma expectativa de que grandes amostras reduziriam a subespecificação inerente ao problema de obtenção de matrizes OD sintéticas.

Já com relação ao objetivo (b), existem na literatura distintos métodos de reconstrução que visam reconstituir a matriz OD específica que gerou os volumes observados em um dado período de tempo, sendo os três principais discutidos a seguir.

3. MÉTODOS DE RECONSTRUÇÃO DE MATRIZES OD ESTÁTICAS

Conforme destacado anteriormente, os primeiros métodos para obtenção de matrizes OD diretamente dos volumes de tráfego tiveram como base técnicas de otimização que visavam reconstruir uma dada matriz OD assumindo que o fenômeno de viagens era determinístico. Inicialmente, foram propostos expedientes sem considerar informação prévia sobre o padrão de deslocamentos na área em estudo; em seguida, funções de distância em relação a uma matriz OD prévia foram incorporadas tendo por objetivo eliminar a indeterminação do sistema de equações . Dentro desta linha se enquadram os métodos de maximização da entropia (ou minimização da informação) e dos mínimos quadrados generalizado.

3.1 Maximização da Entropia

Uma das primeiras abordagens para a obtenção de matrizes OD sintéticas, sem a necessidade de calibração de modelos de demanda, surgiu depois do trabalho de Wilson (1970), que propôs a entropia para o estudo de movimentos de pessoas ou cargas, no meio urbano ou regional. Este princípio vem sendo usado como uma ferramenta de modelagem no planejamento de transportes há várias décadas. A função de entropia, conforme proposto por Wilson para o fenômeno de viagens, é dada pela Equação . Esta função corresponde ao número de microestados possíveis (viagens individuais) para um dado mesoestado (matriz OD).

$$\mathrm{W}(\mathbf{x}) = \frac{\left(\sum_{j=1}^{J} x_j\right)!}{\prod_{j=1}^{J} x_j!} \tag{7}$$

De acordo com Wilson (1970), o conhecimento sobre o fenômeno, ou fatores macros (exógenos) que explicam o mesmo, é inserido ao problema na forma de restrições, que formam um conjunto viável de matrizes OD. Um modelo matemático descritivo relacionando estes fatores exógenos é então obtido ao resolver um problema de otimização com função-objetivo de maximização da entropia (ME). Por exemplo, o modelo gravitacional de distribuição de viagens pode ser derivado de tal proposta, incorporando as restrições dos totais de viagens originadas e destinadas em cada zona, e a restrição do custo total das viagens (Ortúzar e Willumsen, 2011).

A premissa do método ME é que o mesoestado mais provável será aquele de maior incerteza sobre os seus microestados. Isto significa que quando o comportamento que rege o padrão de viagens é desconhecido pelo analista, é mais razoável ou racional assumir que o estado mais provável é aquele de maior incerteza sobre os microestados, ou o estado com o maior número de microestados possíveis. Esta premissa também pode ser interpretada, segundo Wilson (1970), como a hipótese mais fraca sobre o padrão de viagens entre zonas. Willumsen (1978) propôs inserir como restrição ao ME a consistência nos volumes (Equação (1)), em uma tentativa de melhorar os resultados utilizando dados mais simples de serem obtidos. O resultado é um estimador da matriz de viagens consistente com os volumes de tráfego (Ortúzar e Willumsen, 2011). Paralelamente, Van Zuylen (1979) sugeriu inserir informação adicional na forma de uma matriz de viagens previamente conhecida. A matriz resultante do modelo seria então uma matriz consistente com os volumes observados e próxima à matriz prévia. Detalhes destas primeiras formulações podem ser encontrados em Van Zuylen e Willumsen (1980) e em Willumsen (1981).

O método de Van Zuylen (1979), similar ao ME, é derivado de uma função que mede a quantidade de informação na ocorrência de eventos discretos, neste caso viagens em um dado par OD, passando por um arco específico onde volumes são observados. A matriz prévia, juntamente com a matriz de alocação, é inserida na formulação para estimar a probabilidade de ocorrência destes eventos. O objetivo é obter uma matriz usando o mínimo de informação prévia que seja consistente com os volumes observados nos arcos.

Outra forma de inserir informação prévia, como discutido por Van Zuylen e Willumsen (1980), é assumir que o vetor de fluxos OD $\mathbf{x} = [x_j: j = 1, ..., J]$ é uma variável aleatória multivariada do tipo multinomial com parâmetros N e $\mathbf{p} = [p_j: j = 1, ..., J]$, em que N é o total de viagens na região de estudo e p_j é a probabilidade de ocorrência de uma viagem no par OD j. Usando a matriz prévia para estimar as probabilidades de viagens para cada par OD, encontra-se o que é conhecido como *distância entrópica* de $\mathbf{x}$ em relação à matriz prévia. Note que um modelo similar é obtido com a premissa de que o padrão de viagens é um processo de Poisson (Spiess, 1987). A partir destas premissas, a função-objetivo do problema ME com informação prévia, denotada pelas probabilidades estimadas $\hat{p}_j$, comumente adotada na literatura é dada conforme a Equação (8).

$$f(\boldsymbol{x}) = \frac{\left(\sum_{j=1}^{J} x_j\right)!}{\prod_{j=1}^{J} x_j!} \prod_{j=1}^{J} \hat{p}_j^{x_j} \tag{8}$$

Pode-se mostrar que a solução ótima do modelo de maximização de entropia satisfaz as seguintes equações não-lineares (Van Zuylen e Willumsen, 1980):

$$x_j = \hat{x}_j \prod_{i=1}^{I} a_i^{m_{ij}} \quad j = 1, 2, ..., J \tag{9}$$

Em que $\hat{x}_j$ são os fluxos OD prévios; a_i são fatores de balanceamento; e m_{ij} são os elementos da matriz de alocação correspondentes ao arco i e ao par OD j. A solução de ME com restrição dos volumes, assim como no modelo gravitacional, é um modelo do tipo multiplicativo. Estes modelos são casos particulares da formulação mais geral de Bregman (Lamond e Stewart, 1981). Neste método geral, outras restrições (de igualdade ou desigualdade) podem ser adicionadas, como por exemplo capacidade dos arcos e restrição de estacionamento nas zonas. Os parâmetros de quaisquer destes modelos multiplicativos podem ser obtidos por um método iterativo multi-proporcional como proposto por Murchland (1977).

Segundo Ortúzar e Willumsen (2011), o método ME tem sido adotado provavelmente porque é um método simples e de fácil implementação computacional, embora tenha suas limitações. Uma delas é o caráter multiplicativo da solução do modelo (Equação), o qual implica que células vazias na matriz prévia serão mantidas na solução final. Uma forma de contornar este problema é adotar um valor baixo diferente de zero para as células vazias.

O método ME com informação prévia como proposto originalmente (Equação) não é invariante à aplicação de um fator de escala uniforme à matriz prévia. Isto porque se assume que o total de viagens na área de estudo é constante. Assim, distorções são esperadas para o caso em que o objetivo é atualizar/ajustar uma matriz prévia que representa um contexto no passado. Isto ocorre quando há acréscimo de demanda (crescimento dos fluxos OD e consequentemente nos volumes nos arcos) e uma matriz desatualizada é usada como matriz prévia. Uma solução para este problema, conforme proposto por Bell (1983), é usar a matriz prévia como uma restrição estrutural, na qual se procura manter as proporções da referida matriz em vez de minimizar a distância da matriz modelada em relação a prévia. Esta abordagem é equivalente a inserir no método ME com matriz prévia uma restrição dos totais de viagens, conforme mostrado em Ortúzar e Willumsen (2011).

Outra limitação importante para aplicação do método é a premissa de que o vetor de volumes observados é consistente (i.e. o conjunto de restrições dado pelo sistema possui solução). No método ME, assume-se que os volumes observados são precisos. A matriz prévia é utilizada somente como informação complementar, tendo em vista que os volumes não são suficientes para se obter uma matriz única devido à indeterminação do problema. Contudo, inconsistências nos volumes dos arcos são comuns na prática devido a dois fatores: (a) erros de contagem que violam a condição de continuidade dos volumes; e (b) erros devido às premissas de escolha de rotas adotadas que podem não ser compatíveis com os volumes observados. Por exemplo, a premissa de que os usuários têm pleno conhecimento da rede e minimizam tempos de viagens pode não corresponder à realidade. Assim, é possível que rotas relevantes não sejam consideradas na análise tornando o sistema de equações em inviável.

Uma extensão do ME foi proposta por Willumsen (1984) para tratar o primeiro problema de imprecisão na medição dos volumes de tráfego. Nesse método, assume-se que os volumes observados contêm algum grau de incerteza. A função objetivo é então formada pela soma de duas parcelas, uma que busca minimizar a distância em relação à matriz prévia e outra que busca minimizar a distância dos volumes *previstos pelo modelo* em relação aos volumes observados. O resultado continua sendo um modelo do tipo multiplicativo que pode ser resolvido por um método iterativo. De forma geral, a função-objetivo dos métodos de otimização pode ser formulada segundo a Equação (Cascetta e Nguyen, 1988):

$$f(\mathbf{x}) = \gamma_1 f_1(\mathbf{x}, \hat{\mathbf{x}}) + \gamma_2 f_2(\mathbf{v}, \hat{\mathbf{v}}) \tag{10}$$

Em que $\mathbf{x}$ é o vetor de fluxos OD; $\mathbf{v} = \mathbf{Mx}$ é o vetor de *volumes modelados*; $\hat{\mathbf{x}}$ é o vetor de fluxos OD correspondentes à matriz prévia; e $\hat{\mathbf{v}}$ é o vetor de volumes observados. Por sua vez, $f(.)$ é uma função-objetivo composta pelas parcelas $f_1(.)$ e $f_2(.)$ referentes aos fluxos OD e aos volumes nos arcos, respectivamente; e γ_1 e γ_2 são pesos dados a cada função.

A seguir, apresenta-se o modelo de mínimos quadrados generalizado − tradicionalmente difundido de forma inadequada como um modelo para *estimação* de matrizes OD sintéticas − cuja função-objetivo pode ser obtida como um caso particular da Equação .

3.2 Método de Mínimos Quadrados Generalizado

Também chamado de GLS (do inglês, *Generalized Least Squares*), este método é uma generalização do método de mínimos quadrados; enquanto neste último a medida de distância adotada é a distância euclidiana, o GLS emprega a *distância de Mahalanobis*. Dada a matriz OD real $\mathbf{x}$, a matriz OD prévia $\hat{\mathbf{x}}$ e os volumes OD observados $\hat{\mathbf{v}}$, tem-se:

$$\hat{\mathbf{x}} = \mathbf{x} + \boldsymbol{\varepsilon}_{\mathbf{x}} \tag{11}$$

$$\hat{\mathbf{v}} = \mathbf{v} + \boldsymbol{\varepsilon}_{\mathbf{v}} \tag{12}$$

Em que $\mathbf{v}$ corresponde ao vetor de volumes reais; $\boldsymbol{\varepsilon}_{\mathbf{x}}$ corresponde a um vetor de erros de amostragem na matriz OD prévia; e $\boldsymbol{\varepsilon}_{\mathbf{v}}$ a um vetor de erros de observação nos volumes. Admitindo que os volumes reais se relacionam à matriz OD real pela expressão $\mathbf{v} = \hat{\mathbf{M}}\mathbf{x}$, em que $\hat{\mathbf{M}}$ é uma matriz de alocação estimada, e supondo adicionalmente que ambos os vetores de erro são independentes com média $\mathbf{0}$ e matrizes de covariância $\mathbf{Z}$ e $\mathbf{W}$, respectivamente, o estimador GLS da matriz real $\mathbf{x}$ que gerou os volumes OD é dado pelo vetor $\mathbf{x}^{\mathrm{GLS}}$ que minimiza a soma das distâncias de Mahalanobis (Cascetta, 1984; McNeil e Hendrickson, 1985; Bell, 1991):

$$\mathbf{x}^{\mathrm{GLS}} = \underset{\mathbf{x}\geq\mathbf{0}}{\operatorname{argmin}}\left\{(\mathbf{x}-\hat{\mathbf{x}})^{\mathrm{T}}\mathbf{Z}^{-1}(\mathbf{x}-\hat{\mathbf{x}}) + (\hat{\mathbf{M}}\mathbf{x}-\hat{\mathbf{v}})^{\mathrm{T}}\mathbf{W}^{-1}(\hat{\mathbf{M}}\mathbf{x}-\hat{\mathbf{v}})\right\} \tag{13}$$

A Equação corresponde a um problema de otimização quadrática, cuja solução ótima é única, e pode ser resolvido por meio do algoritmo do gradiente projetado ou qualquer outro algoritmo para solução de problemas quadráticos. Uma limitação do modelo é que ele precisa de uma matriz prévia para ser aplicado. A formulação dada por pode ser derivada também a partir do princípio de maximização da verossimilhança, conforme visto em Cascetta (2009). A função de verossimilhança expressa a probabilidade condicional dos dados observados em relação aos possíveis valores dos parâmetros populacionais. No presente caso, os dados observados são a matriz prévia $\hat{}$ e o vetor de volumes observados $\hat{\mathbf{v}}$; enquanto o parâmetro populacional é a matriz OD esperada $\mathbf{x}$ que gerou os volumes observados. Logo, admitindo-se que a matriz prévia e o vetor de volumes observados são independentes, a função de log-verossimilhança é dada pela Equação .

$$l(\mathbf{x}) = \log f(\hat{\mathbf{x}},\hat{\mathbf{v}} \mid \mathbf{x}) = \log f(\hat{\mathbf{x}} \mid \mathbf{x}) + \log f(\hat{\mathbf{v}} \mid \mathbf{x}) \tag{14}$$

Em que $l(\mathbf{x})$ é o logaritmo da função de verossimilhança; e $f(.)$ é uma função de probabilidade condicional. Admitindo que os vetores de erros $\boldsymbol{\varepsilon}_{\mathbf{x}}$ e $\boldsymbol{\varepsilon}_{\mathbf{v}}$ seguem distribuições normais multivariadas com média zero, a função-objetivo do problema de maximização da verossimilhança é equivalente à função quadrática de (13). O parâmetro a ser encontrado neste caso é um valor constante, ou seja, uma matriz específica que gerou os volumes observados. A matriz prévia representa uma estimativa direta ou inicial da matriz desejada que é ajustada com base em informação de volumes de tráfego. Nesta abordagem, portanto, assume-se que a matriz prévia e o vetor de volumes observados são oriundos da mesma realização (mesmo dia ou período) do processo de viagens. Isto fica mais claro nos exemplos de aplicação do método apresentados nos artigos de Cascetta (1984) e Bell (1991).

3.3 Reconstrução a partir da Inferência Bayesiana

Como alternativa aos métodos de otimização, alguns autores (Maher, 1983; Tebaldi e West, 1999; Hazelton, 2001) propõem o método estatístico de inferência bayesiana como forma mais adequada de se considerar a informação prévia na reconstrução da matriz OD sintética. Dada a matriz real $\mathbf{x}$, especifica-se uma distribuição de probabilidades *a priori* $f(\mathbf{x})$ que expressa a incerteza ou *grau de crença* (segundo a terminologia da inferência bayesiana) antes da observação dos dados empíricos, que no presente caso correspondem aos volumes de tráfego observados. Uma matriz prévia pode ser utilizada para especificar a distribuição *a priori*; mas caso não esteja disponível, pode-se especificar uma distribuição de probabilidades *não informativa*. Deve-se especificar também uma função de verossimilhança $f(\hat{\mathbf{v}} \mid \mathbf{x})$ dos dados observados. Obtém-se, então, uma distribuição de probabilidades *a posteriori* da matriz real $\mathbf{X}$, denotada por $f(\mathbf{x} \mid \hat{\mathbf{v}})$, por meio da aplicação do teorema de Bayes:

$$f(\mathbf{x} \mid \hat{\mathbf{v}}) = \frac{f(\hat{\mathbf{v}} \mid \mathbf{x}) f(\mathbf{x})}{\int f(\hat{\mathbf{v}} \mid \mathbf{x}) f(\mathbf{x})} \tag{15}$$

A integral no denominador da expressão é chamada de *constante de normalização*, e pode ser de difícil cálculo computacional. No entanto, em muitos métodos é possível obter propriedades da distribuição *a posteriori* sem a necessidade de calcular a constante de normalização (Gelman *et al.*, 2003), sendo necessário apenas o cálculo do numerador de . Na formulação bayesiana do problema de reconstrução proposta por Hazelton (2001), o vetor de fluxos OD $\mathbf{x} = [x_j$: $j = 1, ..., J]$ é considerado uma variável aleatória, em que cada componente x_j segue uma distribuição de Poisson independente com média λ_j. Dessa forma, para um mesmo período de referência, a matriz OD a cada dia é uma realização do vetor aleatório $\mathbf{x}$. Em uma formulação completa, deve-se especificar também uma distribuição *a priori* $f(\ddot{\mathbf{e}})$ para o vetor de fluxos médios $\lambda = \left[\lambda_j : j=1, ..., J\right]$. Logo, a distribuição *a posteriori* de $\mathbf{x}$ é dada por:

$$f(\mathbf{x} \mid \hat{\mathbf{v}}) \propto \int f(\hat{\mathbf{v}} \mid \mathbf{x},\boldsymbol{\lambda}) f(\mathbf{x} \mid \boldsymbol{\lambda}) f(\boldsymbol{\lambda}) \mathrm{d}\boldsymbol{\lambda} \tag{16}$$

Em que na expressão se omitiu a constante de normalização. Para a avaliação da distribuição *a posteriori* dada em (16), deve-se definir cada um dos termos dentro da integral. O termo $f(\hat{\mathbf{v}} \mid \mathbf{x},\boldsymbol{\lambda})$ corresponde à função de verossimilhança do vetor de volumes observados. Caso não haja erros de observação, $f(\hat{\mathbf{v}} \mid \mathbf{x},\boldsymbol{\lambda}) = \mathrm{I}\{\hat{\mathbf{v}} = \mathbf{M}\mathbf{x}\}$, em que $\mathrm{I}\{.\}$ denota a *função indicadora*, cujo valor é 1 se a condição $\hat{\mathbf{v}} = \mathbf{M}\mathbf{x}$ for verdadeira, e 0 caso contrário; e $\mathbf{M}$ é a matriz de alocação (dada como conhecida e constante). O termo $f(\mathbf{x} \mid \boldsymbol{\lambda})$ denota a probabilidade condicional do vetor de fluxos OD $\mathbf{x}$ dado um vetor de fluxos médios $\boldsymbol{\lambda}$.

Hazelton (2000, 2001) propõe o uso de uma distribuição normal multivariada para $f(\mathbf{x} \mid \boldsymbol{\lambda})$ com média dada por $\boldsymbol{\lambda}$ e matriz de covariância dada por $\boldsymbol{\Lambda} = \operatorname{diag}(\boldsymbol{\lambda})$, em que $\operatorname{diag}(\boldsymbol{\lambda})$ corresponde a uma matriz cuja diagonal principal é preenchida pelos elementos do vetor $\boldsymbol{\lambda}$ e todos os outros elementos são nulos. Adotando também uma distribuição *a priori* normal multivariada para $\boldsymbol{\lambda}$, com média $\boldsymbol{\lambda}_0$ e matriz de covariância $\boldsymbol{\Sigma}_0$, a Equação (16) toma a seguinte forma:

$$f(\mathbf{x} \mid \hat{\mathbf{v}}) \propto \mathrm{I}\{\hat{\mathbf{v}} = \mathbf{M}\mathbf{x}\} \int |\boldsymbol{\Lambda}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\lambda})^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}(\mathbf{x}-\boldsymbol{\lambda}) - \frac{1}{2}(\boldsymbol{\lambda}-\boldsymbol{\lambda}_0)^{\mathrm{T}}\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\lambda}-\boldsymbol{\lambda}_0)\right\} \mathrm{d}\boldsymbol{\lambda} \tag{17}$$

Em que |.| denota o operador determinante de uma matriz e os termos $\mathbf{\Lambda}^{-1}$ e $\mathbf{\Sigma}_0^{-1}$ referem-se às correspondentes matrizes inversas. Como a expressão ainda permanece computacionalmente intratável, Hazelton (2001) propõe uma aproximação na presença de uma distribuição *a priori* muito informativa (ou seja, com pequena dispersão em torno de $\boldsymbol{\lambda}_0$), dada pela Equação (18):

$$f(\mathbf{x} \mid \hat{\mathbf{v}}) \propto \mathrm{I}\{\hat{\mathbf{v}} = \mathbf{M}\mathbf{x}\} \exp\left\{-\tfrac{1}{2}\left(\mathbf{x}-\boldsymbol{\lambda}_0\right)^{\mathrm{T}} (\mathbf{\Sigma}_0 + \mathbf{\Lambda}_0)^{-1} \left(\mathbf{x}-\boldsymbol{\lambda}_0\right)\right\} \tag{18}$$

Logo, o máximo *a posteriori* (MAP), correspondente à moda da distribuição *a posteriori* $f(\mathbf{x} \mid \hat{\mathbf{v}})$, pode ser aproximado pelo máximo do logaritmo da expressão sujeito às restrições de consistência dos volumes e não negatividade do vetor de fluxos OD, como em (19):

$$\mathbf{x}^{\mathrm{MAP}} = \underset{\hat{\mathbf{v}}=\mathbf{M}\mathbf{x},\mathbf{x}\geq\mathbf{0}}{\arg\max} \left\{-\left(\mathbf{x}-\boldsymbol{\lambda}_0\right)^{\mathrm{T}} (\mathbf{\Sigma}_0 + \mathbf{\Lambda}_0)^{-1} \left(\mathbf{x}-\boldsymbol{\lambda}_0\right)\right\} \tag{19}$$

Em que $\mathbf{\Lambda}_0 = \mathrm{diag}(\boldsymbol{\lambda}_0)$. Comparando as Equações e, observa-se que o modelo de otimização para a obtenção de $\mathbf{x}^{\mathrm{MAP}}$ corresponde ao modelo GLS com matriz de covariâncias dos fluxos $\mathbf{Z} = \mathbf{\Sigma}_0 + \mathbf{\Lambda}_0$, matriz prévia $\hat{\mathbf{x}} = \boldsymbol{\lambda}_0$ e matriz de covariâncias dos volumes observados $\mathbf{W} = \mathbf{0}$, uma vez que nessa formulação assume-se que não há erros nos volumes observados. Note que $\mathbf{Z}$ neste caso resulta da soma da variação correspondente à incerteza prévia com relação aos fluxos OD médios reais, na forma da matriz de covariâncias $\mathbf{\Sigma}_0$ da distribuição *a priori* dos fluxos OD médios, e da variação dos fluxos ODs realizados de acordo com distribuições de probabilidades de Poisson independentes, cuja variabilidade se expressa na matriz diagonal de covariâncias $\mathbf{\Lambda}_0$. Ademais, o modelo pode ser facilmente estendido para levar em consideração os erros de observação de volumes, assumindo que o vetor de erros também segue uma distribuição normal multivariada com média $\mathbf{0}$ e matriz de covariâncias $\mathbf{W}$. Após a exposição dos principais métodos de reconstrução de matrizes OD, na Seção 4 a seguir são discutidas questões, que consideramos fundamentais, relativas às premissas básicas dos métodos apresentados, tendo como referência a abordagem apresentada na Seção 2.

4. QUESTÕES CONCEITUAIS DOS MÉTODOS DE RECONSTRUÇÃO

4.1 Aleatoriedade e Informação Prévia

Com base no exposto na seção 2, a primeira premissa a ser questionada com respeito aos métodos de reconstrução ME e GLS é que se assume que o fenômeno de viagens para um dado período, de médio ou curto prazo, é determinístico. Assim, o vetor de viagens $\mathbf{x}$ é considerado aproximadamente fixo. A matriz de probabilidade de escolha de rotas $\mathbf{P}$ é também considerada como resultado de um processo aproximadamente determinístico, pois se considera que possíveis variações nos volumes ocorrem apenas devido a erros de medição, não assumindo, portanto, qualquer variação devido ao processo de escolha de rotas.

Em contrapartida, nos métodos com base na inferência bayesiana, geralmente assume-se que o vetor de deslocamentos $\mathbf{x}$ segue um processo probabilístico, em geral poissoniano. A matriz $\mathbf{P}$ pode ser resultado de um processo determinístico ou probabilístico. No entanto, no método de reconstrução proposto por Hazelton (2000, 2001), assume-se simplificadamente que a matriz de alocação $\mathbf{M}$ é conhecida e constante, e que os volumes observados num único dia são volumes típicos representativos dos volumes médios da rede; ou seja, na prática o método de Hazelton é aplicável apenas quando o vetor $\mathbf{x}$ de viagens varia pouco, em redes não congestionadas. Entretanto, em muitas das aplicações, os métodos são testados em redes de pequeno porte e sem restrição de capacidade, caso em que $\mathbf{P}$ é uma matriz binária e de fácil determinação.

Outra questão fundamental na formulação do problema, não discutida satisfatoriamente na literatura internacional, é a utilização de informação prévia, ou *a priori*. Devido à ineficácia dos primeiros métodos que utilizavam somente a informação proveniente dos volumes de tráfego, passou-se a introduzir informação adicional com relação ao padrão de viagens, na forma de uma matriz OD prévia que na prática entendemos ser de quatro tipos: (a) <u>matriz OD amostrada</u> (obtida de métodos de levantamento diretos); (b) <u>matriz OD expandida</u> com base em modelos de demanda; (c) <u>matriz OD desatualizada</u> (obtida a partir de modelos calibrados com dados amostrais do passado); ou ainda (d) <u>matriz OD subjetiva</u>, construída por analistas com base no seu conhecimento prévio sobre o fenômeno. Ademais, dependendo do tipo de formulação do modelo e do método de solução adotado, esta matriz prévia pode ser interpretada de três formas distintas: (i) como uma *aproximação inicial* da matriz que se deseja obter, a qual deve ser ajustada ou expandida; (ii) como uma *matriz alvo*, da qual se deseja se distanciar o mínimo possível; ou ainda (iii) como uma *matriz média* ou parâmetro de uma distribuição de probabilidades.

No método ME com informação adicional proposto por Van Zuylen e Willumsen (1980), a premissa é que a matriz OD mais provável é aquela que se aproxima o máximo possível de uma matriz OD prévia, desde que seja consistente com os volumes observados. A solução deste problema representa equivalentemente uma situação de mínimo de informação; isto significa que se considera a informação de volumes como precisa, e que a matriz prévia apresenta uma incerteza elevada quando comparada com as observações de volumes. Ressaltamos, contudo, que nas aplicações iniciais do método ME, a matriz prévia era considerada uma informação *a priori* defasada sobre o fenômeno das viagens. Esta ideia traduzia bem o interesse da época, quando se iniciaram os estudos em matrizes OD sintéticas, em que se buscava reduzir custos por meio da reutilização de matrizes obtidas em estudos anteriores. Entretanto, quanto maior a indeterminação do problema, maior a importância da matriz prévia. Neste caso, a solução do método ME também tende a ser cada vez mais próxima da estrutura desta informação *a priori*, se configurando de fato como um alvo da otimização, contrariando o propósito das aplicações do método ME no planejamento de transportes.

Já no método GLS, Cascetta (1984) assumiu que a matriz mais provável era encontrada a partir de uma matriz amostrada recentemente, operando explicitamente como uma matriz alvo, desde que consistente com os volumes de tráfego. Ademais, erros de observação no vetor de volumes (através da matriz de covariâncias **W**) e erros de amostragem da matriz OD prévia (através da matriz de covariâncias **Z**) são considerados de forma explícita no GLS. Portanto, Cascetta parecia não acreditar na ideia de que uma matriz antiga, defasada, pudesse fornecer informação substancial para uma situação presente, propondo o uso de uma matriz prévia proveniente de uma amostra recentemente coletada, não necessariamente representativa, mas compatível com a realidade atual. Entendemos que esta abordagem é mais coerente com os objetivos do planejamento estratégico, assim como do operacional, de sistemas de transportes.

Em seu estudo, Cascetta (1984) também comparou os métodos GLS e ME com matriz prévia. A partir de uma análise empírica em uma rede hipotética não congestionada, e na qual os volumes observados eram consistentes, Cascetta concluiu que mesmo com uma estimativa arbitrária para a matriz **Z**, que mede a precisão do vetor $\hat{\mathbf{x}}$, o método GLS era ainda superior ao método ME com matriz prévia. Destacamos, entretanto, que para o caso analisado a principal diferença entre os dois métodos é a consideração ou não de erros nas medidas de volumes. Quando erros nos volumes não são considerados, e a indeterminação do problema não é tão elevada (como é o caso da rede analisada por Cascetta), os dois métodos resultam aproximadamente na mesma solução ótima.

Já na abordagem bayesiana, a matriz prévia é vista como uma estimativa *a priori* sobre o padrão médio da matriz de viagens. Os volumes são então usados como evidência da situação corrente que serve para atualizar a percepção do analista acerca do fenômeno. Portanto, à primeira vista, a matriz prévia é realmente um conhecimento *a priori* sobre o fenômeno, e não um alvo como nos casos dos métodos GLS e ME com informação prévia. O desafio aqui é definir a crença do analista sobre a sua estimativa *a priori*. Assim, o quanto a informação prévia afeta a solução do problema depende do grau de entendimento do analista sobre o fenômeno.

Cabe aqui ressaltar que o método bayesiano também pode ser aplicado na modelagem determinística do fenômeno de viagens, em que o vetor **x** e a matriz de escolha de rotas **P** não variam, ou não se assume qualquer modelo probabilístico para representar o processo de deslocamento (Maher, 1983). Para tanto, assume-se que existe uma matriz constante que ao ser alocada na rede gera volumes com dado desvio em relação aos volumes observados, ou erros de medida independentes do processo que gera as viagens.

O método pode ser derivado da Equação (18), fazendo $\Lambda_0 = \mathbf{0}$ e assumindo que os volumes observados contém erros de medidas. Assim, a crença prévia do analista é voltada para a definição da variação da matriz OD subjetivamente construída em relação a uma matriz real de viagens que é fixa ou determinística. Esta percepção poderia ser fruto de um processo de amostragem das viagens; neste caso, o método de inferência bayesiana de Maher se assemelharia ao GLS de Cascetta.

Na prática qualquer um dos métodos apresentados na Seção 3 pode ser aplicado para reconstruir uma matriz de um dia ou período específico, porém uma dificuldade adicional ao se aplicar o GLS ou a inferência bayesiana é que se os fluxos OD forem na realidade resultados de um processo aleatório deve-se definir qual a estrutura da variância dos erros dos volumes observados nos arcos (Hazelton, 2000). Os erros de medição são, em geral, proporcionais aos volumes nos arcos e podem estar correlacionados de acordo com a topologia da rede. Assim, costuma-se adotar simplificações, como o uso de matrizes identidade, ou a simples não consideração dos erros nos dados observados.

É interessante notar que o ME com matriz prévia pode ser equiparado ao método bayesiano para encontrar uma matriz OD advinda de um processo determinístico. Para tanto, se a informação de volumes é precisa, e adota-se como crença sobre a matriz de viagens uma distribuição de probabilidades multinomial, com parâmetros dados pela matriz OD prévia, o problema de otimização para encontrar o máximo *a posteriori* é equivalente ao método da maximização da entropia. Por outro lado, pode-se também contra-argumentar que o método ME com matriz prévia é adequado para se reconstituir um vetor **x**, o qual é resultado de um processo probabilístico. Neste caso, a matriz prévia é muito informativa, e assume-se que a função de probabilidade condicional $f(\boldsymbol{x}|\lambda)$ é multinomial. Contudo, o método ME com matriz prévia parece ter sido desenvolvido originalmente para o caso determinístico, já que não é possível encontrar uma matriz prévia verdadeiramente representativa da realidade.

Com respeito ainda à consideração de que o método ME pode ser análogo a um método bayesiano, a matriz prévia não necessariamente precisa ser coletada para um período recente, ou no mesmo dia em que os volumes são observados, como é o caso do GLS. Como esta matriz representa no método ME uma aproximação para o valor esperado da matriz de viagens, ela pode ser obtida em qualquer período considerado típico para o tráfego analisado. Inclusive matrizes com base em estudos no passado, ou complementadas com informação subjetiva, podem ser usadas. A principal deficiência do método ME é que a incerteza, ou matriz de variâncias, sobre esta informação *a priori* não é considerada (Maher, 1983).

Acreditamos, contudo, que o método de reconstrução por inferência bayesiana se torna ainda mais apropriado quando o padrão dos deslocamentos é assumido como probabilístico. A grande vantagem desta abordagem é sua fundamentação em modelos de probabilidade que possam representar adequadamente os fluxos OD. Além disso, aspectos subjetivos (conhecimento do analista) ou ainda outras fontes de informação podem ser incorporadas na forma de uma distribuição de probabilidade *a priori*, como uma função da crença do analista sobre o padrão médio de viagens e da própria variabilidade do fenômeno, ou do processo que gera as viagens. Com respeito aos modelos de probabilidade, a maioria dos estudos assume por conveniência que a geração de viagens ocorre segundo um processo de Poisson e, portanto, a probabilidade condicional $f(\boldsymbol{x}|\lambda)$ é uma função de probabilidade de Poisson. Assumir isto é equivalente a dizer que as viagens são eventos raros e independentes entre si no espaço e no tempo. Entretanto, não temos conhecimento de qualquer estudo empírico abrangente de verificação desta premissa, nem de avaliação das implicações desta simplificação na matriz reconstruída.

Gostaríamos de ressaltar, em suma, que na prática uma das principais dificuldades para a viabilização dos métodos de reconstrução diz respeito à qualidade da matriz prévia. Muitos trabalhos empíricos têm como base redes hipotéticas de pequeno porte, restringindo a aplicação dos métodos de otimização somente a análises operacionais. Contudo, em redes de grande porte é difícil não só obter uma estimativa prévia do padrão de viagens, mas também que seja representativa do fenômeno de viagens atual. Em geral, dispõem-se de uma matriz OD de um estudo realizado no passado, e nesse caso o método baseado em inferência bayesiana nos parece ser o mais adequado.

4.2 Qualidade e Quantidade das Observações de Volumes

A inconsistência na Equação entre o vetor de volumes e a matriz OD é um problema comum na prática. Como visto, essa inconsistência tem normalmente duas origens: a primeira é o erro relacionado à medição dos volumes; e a segunda é a inadequação dos modelos de alocação, que podem produzir estimativas da matriz de alocação fora da realidade. Outra fonte de inconsistência é a própria variação temporal do tráfego: volumes observados em pequenos intervalos de tempo produzem fluxos inconsistentes nos nós da rede. Este último aspecto é ignorado na modelagem estática, pois nesse caso assume-se que o intervalo de tempo de observação é longo o suficiente para eliminar esse tipo de inconsistência.

Da mesma forma, a variabilidade dos volumes devido a uma possível natureza probabilística das viagens também não é considerada nos métodos de reconstrução aqui discutidos, pois todos partem da observação de apenas um único vetor de volumes (amostra $n = 1$). Qualquer incerteza na informação do conjunto de volumes na rede é devido aos erros de medição ou ao modelo de alocação de tráfego. Vale reforçar que os métodos consideram o conjunto de volumes observados como sendo típicos da rede em análise, portanto representando volumes médios, com variância insignificante. A imprecisão na observação de volumes, ou a qualidade do vetor de volumes, é considerada de forma pragmática nos métodos de otimização (como no GLS) ou através de considerações estatísticas sobre a aleatoriedade dos erros (como no caso da inferência bayesiana). Em geral, assume-se que os erros nas medições de volumes são independentes e seguem distribuições normais com valores esperados nulos e variâncias proporcionais aos próprios valores médios de volumes. Para o caso do fenômeno de viagens determinístico, a variância dos erros nos volumes poderia ser estimada a partir de uma amostra de vários dias de observações de volume nos arcos da rede. Entretanto, no caso do fenômeno probabilístico, não é possível isolar, a partir de uma amostra, a variação devido somente ao erro de medida nos volumes.

Ressaltamos que o método ME com matriz prévia necessita de um conjunto de volumes com elevada precisão para que o método encontre uma solução com mínimo de informação da matriz prévia. Argumenta-se que esta restrição de consistência, ou alta precisão da observação de volumes, é crítica nos métodos entrópicos. Há na literatura algumas propostas para corrigir vetores de volumes observados de forma a garantir sua consistência (Van Zuylen e Branston, 1982). Entretanto, a principal barreira ao bom desempenho dos métodos de otimização refere-se à quantidade de informação de volumes, ou o número de arcos com volumes observados. Em geral, quanto maior a quantidade de arcos com informação não redundante, ou mesmo volumes de conversão observados nos nós (Bertoncini e Kawamoto, 2012), melhor é o desempenho dos métodos (Marzano *et al.*, 2009). Em contrapartida, a solução dos problemas de otimização tende a ser mais próxima da matriz prévia quanto maior for a indeterminação do problema.

Embora estas questões com relação à qualidade e quantidade das medidas de volumes nos arcos tenham sido objeto de pesquisas no passado, elas têm perdido destaque nos dias atuais. Atualmente, as tecnologias de detecção veicular permitem obter observações de volumes em todos os arcos da rede de modo contínuo e com boa precisão. Talvez a questão ainda remanescente com relação aos volumes veiculares seja a viabilidade dos métodos de reconstrução para redes urbanas reais de grande porte que apresentam elevada indeterminação na Equação (1). Além disso, a variabilidade ou aleatoriedade inerente ao fenômeno de viagens deve ser considerada de forma mais robusta a partir de análises empíricas, já que se dispõe de grandes quantidades (vários dias de informação de volumes) de dados de volumes nos dias atuais.

4.3 Premissa de Equilíbrio na Rede

A premissa de equilíbrio na alocação do tráfego é bastante aceita e difundida. Conforme discutido, por trás desta premissa está a consideração de que os usuários utilizam uma rede de transportes fixa, que não sofre alteração na sua oferta, e que após um longo período de aprendizado, ou interação entre os usuários e o sistema, a rede atingiria um estado em que qualquer mudança de comportamento não acarretaria em qualquer benefício marginal para os seus usuários. Considerando um processo aproximadamente determinístico das viagens, esta premissa parece ser adequada. Ou seja, se as viagens realizadas têm as mesmas características a cada dia (propósito, destino e horários de saída), sem que haja alteração na oferta de transportes no longo prazo, é possível assumir que, após um certo período de tempo, os indivíduos irão apreender sobre o desempenho da rede e tender a usar um certo número restrito de rotas, ou mesmo uma única rota de preferência. O sistema então convergiria para um estado estacionário nos deslocamentos, com poucas variações em torno do seu comportamento médio. Entendemos que esta visão corresponde ao conceito tradicional de equilíbrio na rede, em que tanto os fluxos OD como os fluxos nas rotas são determinísticos.

Por outro lado, considerando o processo de viagens como probabilístico e dinâmico, parece difícil assumir tal equilíbrio. Se os usuários tendem a mudar as decisões de viagens com certa frequência, influenciados por resultados de decisões anteriores, possivelmente o sistema não atingiria um estado de equilíbrio nos custos. Assim, para se reconstruir uma matriz específica (ou uma realização do processo estocástico), assumir que esta realização do processo aleatório de viagens estaria em equilíbrio parece não ser adequado. Seria necessário considerar a dinâmica do fenômeno de viagens tanto no processo de alocação de tráfego, ou de escolha de rotas, quanto na geração da matriz OD de viagens. Ou seja, observações de volumes não apenas do período em que se deseja reconstruir a matriz, mas de períodos anteriores forneceriam informação relevante para se obter a matriz corrente.

5. CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES

Como discutido ao longo deste artigo, os métodos de reconstrução podem ser aplicados para obter matrizes OD sintéticas tanto numa perspectiva determinística, quanto em uma abordagem probabilística do fenômeno de viagens. Os métodos de otimização, ME com matriz prévia e GLS, buscam encontrar a matriz mais provável que gerou um dado conjunto de observações de volumes. Neste caso, a matriz prévia informada pode ser considerada um alvo, cujo efeito na solução depende da indeterminação do problema e da qualidade do conjunto de volumes observados. Já o método de inferência bayesiana considera a matriz prévia como uma informação *a priori* sobre o padrão médio de viagens.

Na perspectiva determinística do padrão de deslocamentos, concluímos pela adequabilidade dos métodos de otimização. Destacamos, entretanto, que o método ME pode ser derivado a partir da inferência bayesiana, assumindo que a distribuição de probabilidade *a priori* do vetor de viagens é multinomial. Argumentamos também que o método ME é mais indicado do que o GLS para situações em redes de maior escala, onde há um alto grau de incerteza na matriz prévia e uma elevada indeterminação no sistema de equações de fluxos OD da rede, tendo em vista que o primeiro pressupõe o uso do mínimo de informação prévia que seja consistente com os volumes de tráfego. Ao minimizar a informação adicional, o método ME pressupõe que os volumes de tráfego podem ser observados de maneira precisa, tendo um peso elevado em comparação à matriz prévia. Por outro lado, quanto maior a indeterminação do problema, maior será a importância da matriz prévia. Na prática, espera-se uma elevada indeterminação na Equação (1) para redes de cidades de médio e grande portes, mesmo com sistemas amplos de monitoramento do tráfego urbano.

Já o método GLS parece ser mais adequado para estudos operacionais em redes de pequeno porte, nas quais observações de viagens e volumes do tráfego atual possam ser obtidas, pois neste caso considera-se a matriz prévia como uma estimativa inicial recente da matriz real que se deseja obter. Além disso, diferente do ME, o método GLS considera de forma explícita incertezas tanto nos fluxos OD prévios como nos dados de volumes. Contudo, para que o método possa ser aplicado, torna-se necessária a adoção de simplificações pouco realistas devido à dificuldade prática de estimar a imprecisão na matriz prévia e nos volumes observados.

Por sua vez, ao modelarmos o padrão de deslocamentos como um fenômeno probabilístico, destacamos as vantagens de aplicar um método de inferência bayesiana. Neste tipo de abordagem a matriz prévia é vista como uma estimativa *a priori* do padrão médio de viagens, sendo considerada muito informativa. A principal premissa para derivação do método é assumir que os deslocamentos são bem modelados por um processo de Poisson multivariado, que pode ser aproximado por uma distribuição de probabilidades normal multivariada, desde que os fluxos OD sejam elevados. É interessante notar também que a formulação da Equação (19) é uma aproximação matemática para o problema do método ME com matriz prévia. A consistência e implicações da premissa de Poisson, assim como da sua aproximação pela Normal, ainda precisam ser melhor verificadas e avaliadas empiricamente.

Ressaltamos, entretanto, que qualquer que seja a abordagem adotada, o desempenho satisfatório dos métodos requer informação prévia e observações de volumes de boa qualidade. A matriz prévia deve ser não viesada e com baixa dispersão. Além disso, a situação ideal é que todos os volumes não redundantes (ou seja, que não são combinações lineares de outros volumes nos arcos) sejam observados e apresentem boa precisão. Contudo, tanto a qualidade quanto a quantidade (cobertura espacial) das medidas de volumes não parece ser uma limitação nos dias atuais. Mesmo assim, em redes de grande porte ainda existem dificuldades práticas de se obter uma matriz representativa do padrão atual de viagens. Destacamos ainda que a grande maioria dos estudos apresentados da literatura internacional foram baseados em simulação computacional de redes hipotéticas de pequeno porte, sendo ainda incipientes os esforços empíricos de verificação e validação dos métodos de reconstrução em redes de larga escala, com dados reais.

Outra questão relevante que buscamos analisar diz respeito à aplicabilidade prática de se reconstruir uma matriz OD estática. Os primeiros métodos de obtenção de matrizes OD sintéticas surgiram como uma alternativa de menor custo e tempo às pesquisas domiciliares, com o objetivo de estimar uma matriz média para fins de planejamento estratégico. Se os métodos de otimização são na verdade métodos de reconstrução de uma única realização do padrão de viagens, questionamos a coerência em se modelar este complexo fenômeno por meio de uma única observação. Além disso, deve-se avaliar em qual escala da rede de transportes a reconstrução dos fluxos OD pode ser mais eficaz. Talvez estes métodos sejam mais adequados em análises operacionais das condições do tráfego em tempo real em sub-redes de pequeno porte, monitoradas continuamente por equipamentos de detecção veicular.

A complexidade da dinâmica do tráfego é outra questão que merece destaque para a reconstrução de matrizes OD sintéticas. Até aqui, as comunidades técnica e científica vem analisando redes de tráfego sob a premissa de que os fluxos OD são independentes no espaço e no tempo. Esta hipótese de independência nas decisões de viagens é válida provavelmente em redes não congestionadas, em que os destinos são poucos atrativos e o tempo de viagem é pouco afetado pela intensidade do tráfego. Já em redes congestionadas, existe maior influência da dinâmica do tráfego nas decisões de viagem. Ou seja, o padrão de viagens deveria ser modelado como um fenômeno estocástico no qual a rede poderia não atingir um estado de equilíbrio (premissa adotada para obtenção da matriz de alocação de tráfego). Assim, os fluxos OD seriam uma realização de uma variável aleatória cujos parâmetros variam de forma dinâmica resultando em carregamentos, ou volumes na rede, não equilibrados. Neste contexto, é mais plausível assumir que existe algum tipo de dependência entre viagens, especialmente quando existem destinos com alta atratividade. Portanto, a premissa de equilíbrio estacionário não seria mais adequada.

Deve-se ainda ter em mente que as viagens em uma rede de transportes resultam de vários processos de decisão nos subsistemas de atividades e do uso do solo, que sofrem forte influência espacial e temporal, tanto no caso urbano como regional. Como discutido neste artigo, as relações de dependência espacial e temporal nas viagens (sendo mais evidente em redes congestionadas) não foram consideradas no desenvolvimento de métodos de reconstrução de matrizes OD sintéticas. Uma questão digna de investigação diz respeito aos ganhos de se inserir tal complexidade em comparação aos métodos tradicionais. Portanto, acreditamos que a combinação do poder de informação dos dados desagregados do uso do solo com aqueles oriundos de um sistema de controle de tráfego em tempo real possibilite a análise empírica e a modelagem da variabilidade do fenômeno do padrão de deslocamentos, por meio da investigação da dependência espacial e temporal das decisões de viagens (tanto na sua geração, como nas escolhas de destinos e rotas).

Recomendamos, por fim, que as seguintes questões teóricas e práticas, inerentes aos métodos de reconstrução de matrizes OD sintéticas, ainda sejam melhor investigadas (embora reconhecendo a prioridade da pesquisa voltada para os métodos de estimação dinâmica): a) eficácia dos métodos em redes reais de médio e grande porte; b) impacto da estimativa da matriz de alocação sobre a matriz reconstruída e verificação da premissa de equilíbrio na rede; c) consideração do caráter probabilístico no padrão de deslocamentos e nas escolha de rotas, assim como verificação das premissas adotadas para os processos aleatórios que geram os fluxos OD e os consequentes fluxos nas rotas; d) análise de sensibilidade para avaliar em que situações (de oferta, de demanda e de informação prévia) os métodos são mais robustos, podendo ser usados como ferramentas que auxiliem efetivamente a tomada de decisão.

AGRADECIMENTOS

Os autores deste trabalho agradecem à CAPES e ao CNPq pelo financiamento da pesquisa.

REFERÊNCIAS

Bell, M. G. H. (1983) The Estimation of an Origin-Destination Matrix from Traffic Counts. *Transportation Science,* v. 17, p. 198–217. DOI: 10.1287/trsc.17.2.198.

Bell, M. G. H. (1991) The Estimation of Origin-Destination Matrices by Constrained Generalized Least Squares. *Transportation Research,* v. 25B, p. 13-22. DOI: 10.1016/0191-2615(91)90010-G.

Bertoncini, B. V. e E. Kawamoto (2012) Modelagem da matriz OD sintética a partir dos volumes observados nas interseções da rede de transportes. *Transportes,* v. 20, n. 2, p. 75–83. DOI: 10.4237/transportes.v20i2.562.

Bertoncini, B. V.; C. F. G. Loureiro e E. Kawamoto (2013) Proposta e Avaliação de Algoritmo de Médias Sucessivas para Reconstrução de Matriz Origem-destino Sintética. *Transportes,* v. 21, n. 2, p. 21–29. DOI: 10.4237/transportes.v21i2.697.

Cascetta, E. (1984) Estimation of Trip Matrices from Traffic Counts and Survey Data: a Generalized Least Squares Estimator. *Transportation Research,* v. 16B, n. 4-5, p.89-99. DOI: 10.1016/0191-2615(84)90012-2.

Cascetta, E. (2009) *Transportation Systems Analysis*: *Models and Applications* (2ª ed.). Springer, New York, USA. DOI: 10.1007/978-0-387-75857-2.

Cascetta, E. e S. Nguyen (1988) A Unified Framework for Estimating or Updating Origin/Destination Matrices from Traffic Counts. *Transportation Research,* v. 22B, p. 437–455. DOI: 10.1016/0191-2615(88)90024-0.

Daganzo, C. F. e Y. Sheffi (1977) On Stochastic Models of Traffic Assignment. *Transportation Science*, v. 11, n. 3, p. 253-274. DOI: 10.1287/trsc.11.3.253.

Fisk, C. (1988) On Combining Maximum Entropy Trip Matrix Estimation with User Optimal Assignment. *Transportation Research B*, v. 22**,** p. 69-79. DOI: 10.1016/0191-2615(88)90035-5.

Fisk, C. (1989) Trip Matrix Estimation from Link Counts: the Congested Network Case. *Transportation Research B*, v. 23, p. 331-336. DOI: 10.1016/0191-2615(89)90009-X.

Gelman, A.; J. B. Carlin; H. S. Stern e D. B. Rubin (2003) *Bayesian Data Analysis* (2ª ed.). Chapman-Hall/CRC, Boca Raton, USA.

Hazelton, M. L. (2000) Estimation of Origin-Destination Matrices from Link Flows on Uncongested Networks. *Transportation Research B,* v. 34, p. 549-566. DOI: 10.1016/S0191-2615(99)00037-5.

Hazelton, M. L. (2001) Inference for Origin-Destination Matrices: Estimation, Prediction and Reconstruction. *Transportation Research B,* v. 35, p.667-676. DOI: 10.1016/S0191-2615(00)00009-6.

Lamond, B. e N. F. Stewart (1981) Bregman's Balancing Method. *Transportation Research B,* v. 15 p. 239–248. DOI: 10.1016/0191-2615(81)90010-2.

Lo, H. P.; N. Zhang e W. H. K. Lam (1996) Estimation of an Origin-Destination Matrix with Random Link Choice Proportions: a Statistical Approach. *Transportation Research B*, v. 30, n. 4, p. 309-324. DOI: 10.1016/0191-2615(95)00036-4.

Maher, M. J. (1983) Inferences on Trip Matrices from Observations on Link Volumes: a Bayesian Statistical Approach. *Transportation Research B,* v. 17, p. 435–447. DOI: 10.1016/0191-2615(83)90030-9.

Marzano, V.; A. Papola e F. Simonelli (2009) Limits and Perspectives of Effective O-D Matrix Correction Using Traffic Counts. *Transportation Research C*, v. 17, p. 120-132. DOI: 10.1016/j.trc.2008.09.001.

McNeil, S. e C. Hendrickson (1985) A Regression Formulation of the Matrix Estimation Problem. *Transportation Science,* v. 19, p. 278–292. DOI: 10.1287/trsc.19.3.278.

Murchland, J. D. (1977) *The Multiproportional Problem*. TSG Note JDM-263, Transport Studies Group, University College London.

Ortúzar, J. de D. e L. G. Willumsen (2011) *Modelling Transport* (4ª ed.). Wiley, Chichester. DOI: 10.1002/9781119993308.

Pitombeira-Neto, A. R. (2015) *Dynamic Bayesian Statistical Models for the Estimation of the Origin-Destination Matrix*. Tese de doutorado apresentada ao Programa de Pós-Graduação em Engenharia de Transportes do Departamento de Engenharia de Transportes da Universidade Federal do Ceará.

Prato, C. G. (2009) Route Choice Modeling: Past, Present and Future Research Directions*.* *Journal of Choice Modelling*, v. 2, n. 1, p. 65-100. DOI: 10.1016/S1755-5345(13)70005-8.

Robillard, P. (1975) Estimating the OD Matrix from Observed Link Volumes. *Transportation Research B*, v. 9, p. 123-128. DOI: 10.1016/0041-1647(75)90049-0.

Sheffi, Y. (1985) *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming methods*, Prentice-Hall, Englewood Cliffs, USA.

Smith, M. J. (1979) The Existence, Uniqueness and Stability of Traffic Equilibria. *Transportation Research B*, v. 13, p. 295-304. DOI: 10.1016/0191-2615(79)90022-5.

Spiess, H. (1987) A Maximum Likelihood Model for Estimating Origin-Destination Matrices. *Transportation Research B,* v. 21, p. 395–412. DOI: 10.1016/0191-2615(87)90037-3.

Tebaldi, C. e M. West (1998) Bayesian Inference on Network Traffic Using Link Count Data. *Journal of the American Statistical Association*, v. 93, n. 442, p. 557-573. DOI: 10.1080/01621459.1998.10473707.

Van Zuylen, H. (1978) The Information Minimizing Method: Validity And Application to Transport Planning. *In*: Jasen, G. R. H. *et al*. (*eds.*) *New Developments in Modelling Travel Demand and Urban Systems*, Saxon, Farnborough.

Van Zuylen, H. e D. M. Branston (1982) Consistent Link Flow Estimation from Counts. *Transportation Research B*, v. 16, p. 473-476. DOI: 10.1016/0191-2615(82)90006-6.

Van Zuylen, H. e L. G. Willumsen (1980) The Most Likely Trip Matrix Estimated from Traffic Counts. *Transportation Research B,* v. 14, p. 281-293. DOI: 10.1016/0191-2615(80)90008-9.

Vardi, Y. (1996) Network Tomography: Estimating Source-Destination Traffic Intensities from Link Data. *Journal of the American Statistical Association,* v. 91, p. 365-377. DOI: 10.2307/2291416.

Wardrop, J. G. (1952) Some Theoretical Aspects of Traffic Research. *Proceedings of the Institution of Civil Engineers part II*, Institution of Civil Engineers, p. 325-378. DOI: 10.1680/ipeds.1952.11259.

Watling, D. (2002a) A second order stochastic network equilibrium model, i: Theoretical foundation. *Transportation Science*, v. 36, n. 2, p. 149-166. DOI: 10.1287/trsc.36.2.149.560.

Watling, D. (2002b) A second order stochastic network equilibrium model, ii: Solution method and numerical experiments. *Transportation Science*, v. 36, n. 2, p. 167-183. DOI: 10.1287/trsc.36.2.167.564.

Willumsen, L. G. (1978) Estimation of an O-D Matrix from Traffic Counts: a Review. Institute for Transport Studies, *Working Paper,* n. 99, Leeds University.

Willumsen, L. G. (1981) Simplified Transport Demand Models Based On Traffic Counts. *Transportation,* v. 10, p. 257–278. DOI: 10.1007/BF00148462.

Willumsen, L. G. (1984) Estimating Time-Dependent Trip Matrices from Traffic Counts. *Proceedings of the 9th International Symposium on Transportation and Traffic Theory*, Delft University, Delft, Netherlands, p. 397-411.

Wilson, A. G. (1970) *Entropy in Urban and Regional Modeling*. Pion, London, England.

Yang, H. (1995) Heuristic Algorithms for the Bilevel Origin-Destination Matrix Estimation Problem. *Transportation Research B*, v. 29, n.4, p. 231-242. DOI: 10.1016/0191-2615(95)00003-V.

Yang, H.; T. Sasaki; Y. Iida e Y. Asakura (1992) Estimation of Origin-Destination Matrices from Link Counts on Congested Networks. *Transportation Research B*, v. 26, p. 417-434. DOI: 10.1016/0191-2615(92)90008-K.