



SIMULAÇÃO PARA AVALIAÇÃO DO EFEITO DE DADOS PERDIDOS NO VALOR DO VOLUME MÉDIO DIÁRIO ANUAL EM RODOVIAS ESTADUAIS DO CEARÁ

Marcos José Timbó Lima Gomes

Maria Elisabeth Pinheiro Moreira

Programa de Mestrado em Engenharia de Transportes - PETRAN

Universidade Federal do Ceará – UFC

RESUMO

Muitos órgãos rodoviários realizam contagens contínuas, através de equipamentos eletrônicos instalados em alguns trechos rodoviários, objetivando obter dados para subsidiar os vários estudos de planejamento e operação de tráfego. Entretanto, tais equipamentos estão altamente propícios a falhas, o que dificulta a obtenção de dados confiáveis pelos programas de monitoramento de tráfego. Mas, para ter informações mais precisas, é comum que alguns órgãos rodoviários estimem os dados faltosos, com a utilização de técnicas e modelos complexos, sem no entanto, investigar a dispensabilidade de tais métodos. Assim, esta comunicação técnica tem como objetivo analisar as perdas de dados dos contadores permanentes do DERT-CE e avaliar a necessidade da estimação de dados faltosos para o cálculo do Volume Médio Diário Anual – VMDA.

ABSTRACT

Several road departments carry out continuous traffic counts in order to obtain data to subsidize planning and traffic operation studies. These data are collected using electronic equipments installed in some sectors of highways. However, such equipments may fail with consequent loss of data, making it difficult to obtain reliable data for the traffic monitoring program. With the purpose of overcome this problem some highway agencies estimate the missing values, using complex techniques and models, without an investigation of the possibility of making the studies not including such missing data. This paper presents an analysis of the traffic data collected by the permanent counters of DERT-CE and an evaluation about the necessity or not of the estimation of missing values for the calculation of the Annual Average Daily Traffic - AADT.

1. INTRODUÇÃO

Em 1993, o Departamento de Edificações, Rodovias e Transportes – DERT-CE, implantou um Plano de Contagem de Tráfego nas rodovias estaduais, com o objetivo de subsidiar informações precisas aos estudos rodoviários de forma a facilitar o processo de tomada de decisão nas áreas de planejamento, segurança e manutenção rodoviária. Cada decisão tomada freqüentemente envolve a alocação de recursos financeiros para execução de melhorias na infra-estrutura das rodovias. O plano de contagem, atualmente é composto de 12 postos de contagem permanentes (contagem contínua), 93 postos sazonais e 169 postos de cobertura (contagem de curta duração).

Os postos permanentes realizam contagens contínuas, efetuadas por equipamentos eletrônicos que através de um sensor indutivo, instalado na pista da rodovia, determina a presença ou não de um veículo. Estes equipamentos são conhecidos como ATR (*Automatic Traffic Recorders*) e os dados de volume armazenados são normalmente registrados em períodos de uma hora e por faixa de tráfego. Segundo o *Federal Highway Administration* – FHWA (2001), tais postos ajudam os órgãos rodoviários a entender as variações diárias, semanais, mensais e anuais, e os padrões de viagem sazonais, além de permitir o desenvolvimento de mecanismo necessário para converter as contagens de curta duração em estimativas precisas das condições anuais do tráfego (fatores de expansão).

Entretanto, tais equipamentos estão altamente propícios a falhas, o que dificulta a obtenção de



informações precisas pelos programas de monitoramento de tráfego. Segundo Zhong (2003), no Estado de Alberta, Canadá, em um período de 7 anos, mais da metade dos dias contados foram perdidos e em alguns anos os percentuais de dados válidos variaram entre 10% a 30%.

No Estado do Ceará, os postos de contagem permanentes também possuem um elevado percentual de dados faltosos. Para um período de contagem entre 1997 a 2001 o percentual médio de dados válidos para todos os postos permanentes é 56%.

Segundo a *American Association of State Highway and Transportation Officials – AASHTO* (1992), os custos, para o órgão rodoviário, provocados pelas perdas de dados e conseqüentemente o fornecimento de informações imprecisas aos estudos rodoviários, são bastante elevados, sendo computados todas as vezes que estatísticas errôneas forem usadas em projetos de planejamento, pavimentação, estudos de acidentes, avaliações na qualidade do ar, e outros.

Para ter informações mais precisas, é comum que alguns órgãos rodoviários estimem os dados faltosos, com a utilização de técnicas e modelos complexos, sem no entanto, investigar a dispensabilidade de tais métodos. Assim, esta comunicação técnica tem como objetivo analisar as perdas de dados dos contadores permanentes do DERT-CE, no cálculo do Volume Médio Diário Anual – VMDA, e avaliar a necessidade para a estimação de dados faltosos.

2. DADOS PERDIDOS

Desde os anos 30, há uma preocupação no tratamento e análise de dados faltosos dos programas de monitoramento de tráfego (Albright, 1991), visto que quando há altos percentuais de dias perdidos nos postos permanentes, normalmente os dados obtidos de tais postos não podem ser usados para o cálculo de estatísticas anuais de tráfego, como por exemplo, do Volume Médio Diário Anual (Zhong *et al.*, 2003). Entretanto alguns órgãos rodoviários comumente estimam os valores perdidos ou inválidos, substituindo-os muitas vezes por valores históricos, sendo esta técnica denominada *imputation*.

De acordo com a AASHTO (1992), a estimação de dados faltosos, não são recomendadas, pois introduzem erros que não podem ser quantificados, e ao mesmo tempo estimula aos órgãos rodoviários que se utilizam destas técnicas um abandono de tais práticas. Mas, a preocupação principal para a não utilização da *imputation* é com o princípio da integridade dos dados, em que as medidas de tráfego devem ser armazenadas sem qualquer modificação ou ajuste. Entretanto, Zhong *et al.*, (2003) afirma que a *imputation* de dados não é proibitiva e que, para algumas análises avançadas pode-se utilizar a técnica de estimação de dados faltosos.

As dificuldades para a utilização da *imputation* é que os métodos mais precisos se utilizam de modelos mais complexos como redes neurais, algoritmos genéticos, e os métodos mais simples, como os de fatores de ajuste ou séries temporais, obtêm resultados não tão precisos. Além disso, Wrigth *et al.*, (1997) afirmam que a perda aleatória dos dados tem influência desprezíveis no cálculo de estimativas anuais do tráfego.



3. ESTUDO DE CASO - METODOLOGIA

3.1 Dados analisados

Para avaliar o efeito das perdas dos dados dos contadores permanentes, no cálculo do VMDA, provocadas por falhas no funcionamento dos equipamentos, desvio nas rodovias, ou dados inválidos, foram escolhidos 3 postos considerados representativos, de acordo com o tipo funcional da rodovia a qual pertence, padrão de variação mensal do tráfego, e magnitude do VMDA. A Tabela 1 mostra as características dos postos selecionados.

Tabela 1: Características dos postos permanentes selecionados para a realização do estudo

Número do Posto	Rodovia	Tipo Funcional	VMDA (veíc/dia)
P01	CE-040	Rodovia Turística	4607
P02	CE-187	Rodovia de Tráfego Geral	2103
P12	CE-292	Rodovia de Tráfego Geral	373

Como os postos selecionados possuem magnitudes do VMDA muito diferentes, para a identificação dos padrões de variação mensal do tráfego, usou-se uma medida relativa, denominada Fator Mensal e obtida pela Equação 1.

$$FatorMensal = \frac{VolumeMédioMensal}{VolumeMédioDiárioAnual} \quad (1)$$

A Figura 1 mostra um gráfico desta variável plotada ao longo do ano identificando os diferentes padrões de variação mensal do tráfego dos postos selecionados.

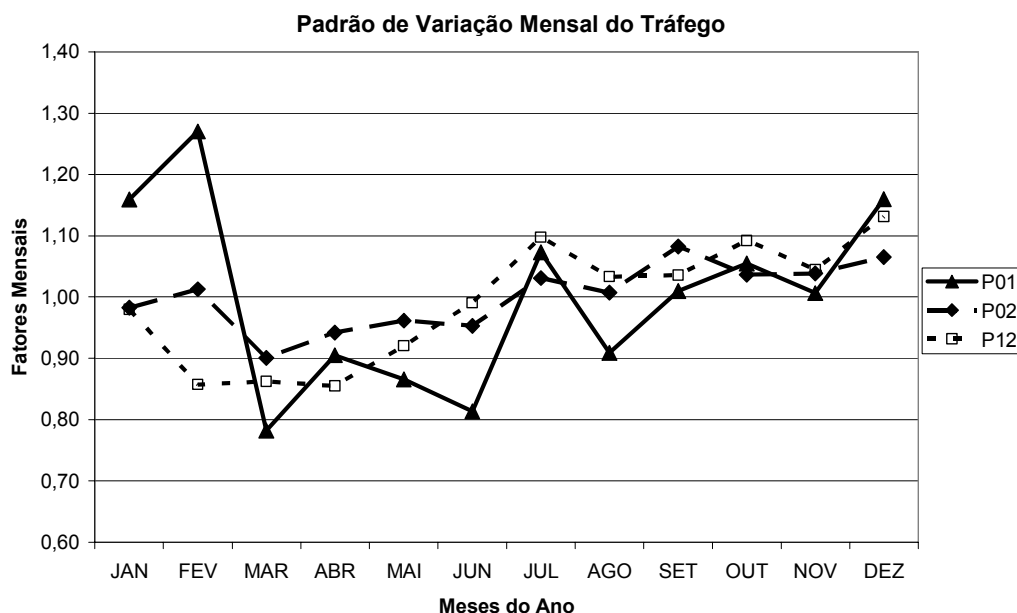


Figura 1: Padrão de variação mensal do tráfego dos postos permanentes

Além destas características, os postos necessitavam ter um ano completamente sem falhas. Mas, para o período analisado de 1997 a 2001 nenhum dos 12 postos atendia a tal exigência, sendo utilizado para o estudo, o ano de 1998, visto que apresentava a menor quantidade de falhas entre todos os anos da série histórica. A Figura 2 mostra um gráfico que exhibe os dias de dados perdidos para o ano de 1998 nos 12 postos permanentes.

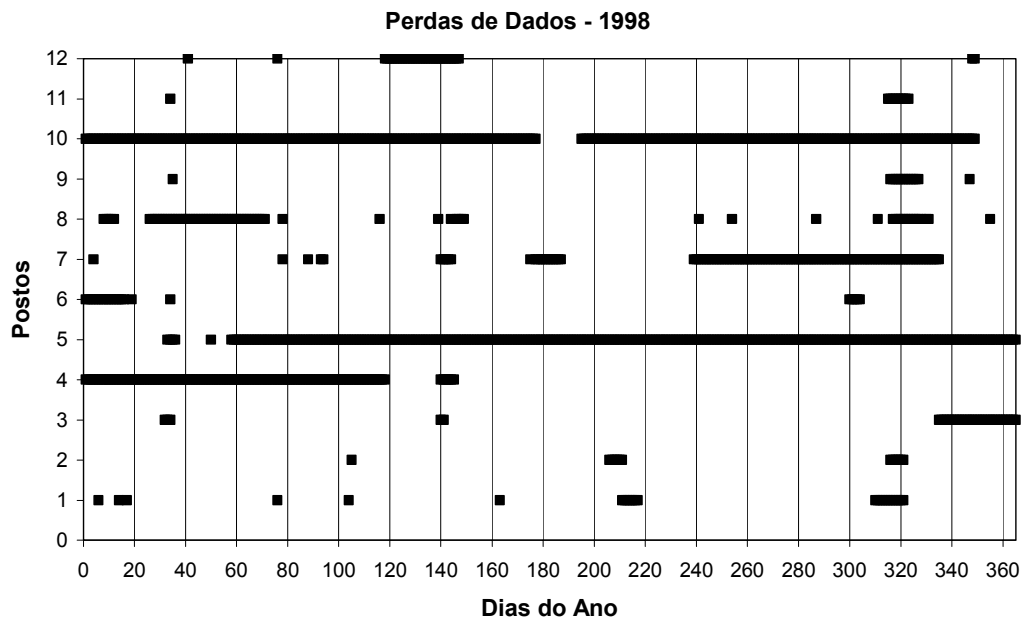


Figura 2: Gráfico de dias perdidos para os 12 postos permanentes no ano de 1998

3.2 Simulação

O estudo consistiu da comparação do $VMDA_{REAL}$, calculado com base nos dados existentes e válidos dos postos analisados, e o $VMDA_{SIMULADO}$ calculado após a retirada de uma certa quantidade de dados, de acordo com os seguintes cenários: i) perda aleatória de um percentual de 5% dos dados; ii) perda aleatória de um percentual de 20% dos dados; iii) perda aleatória de um percentual de 50% dos dados.

Para a realização da simulação, adotou-se que um posto específico, com 'N' dados válidos, a remoção de 5% destes dados ($d_1 = 0,05.N$). Com os dados restantes ($N - d_1$) calculou-se o $VMDA_{SIMULADO}$. Em seguida foi realizada a reposição dos dados retirados, e repetiu-se o passo anterior em 999 vezes. Portanto, para cada cenário, de cada posto, totalizaram 1000 valores de $VMDA_{SIMULADOS}$. Em seguida, foi calculado, o erro percentual entre o valor real e cada valor simulado, através da Equação 2.

$$\text{Erro Percentual} = \frac{|VMDA_{REAL} - VMDA_{SIMULADO}|}{VMDA_{REAL}} \times 100\% \quad (2)$$

Finalizando foi realizada a distribuição dos erros percentuais, calculando-se os erros médios, bem como os 50°, 75° e 95° percentis, para avaliação do cenário que mais se aproximou do valor real.

Os cenários avaliados consideram a perda de um dia de forma isolada, mas como pode ser visto na Figura 2, as falhas ao acontecerem tende a prolongar-se por 2 ou mais dias consecutivos, pelo fato destas perdas de dados resultarem de falhas no equipamento, e o intervalo de tempo de reparo ser superior a 1 dia.

Sendo assim, simulou-se mais 6 cenários, agora considerando a perda consecutivas dos dados,



os quais tentam emular os padrões de perda de dados mostrados na Figura 2. Para os novos cenários, os períodos de perda foram os seguintes: i) 1 semana; ii) 2 semanas não consecutivas; iii) 2 semanas consecutivas; iv) 4 semanas, sendo 2 consecutivas; v) um mês; vi) 2 meses.

4. ANÁLISE DOS RESULTADOS

Os erros médios encontrados para os cenários que consideram as perdas aleatórias e isoladas variaram de 0,10% a 1,63%, tendo os melhores resultados o posto P02 (Tabela 2). Analisando os 95º percentis dos erros, verifica-se que mesmo com um alto percentual de dados perdidos (50%), o maior erro obtido foi de 3,97%, indicando um valor muito próximo do VMDA real. Nos demais cenários considerando as perdas consecutivas, o maior erro encontrado para o 95º percentil foi de 2,97%, obtido na simulação com perda de 2 meses.

Tabela 2: Resultados da simulação considerando perdas isoladas nos postos pesquisados

	P01			P02			P12		
	Perda 5%	Perda 20%	Perda 50%	Perda 5%	Perda 20%	Perda 50%	Perda 5%	Perda 20%	Perda 50%
ERRO MÉDIO	0,37%	0,79%	1,63%	0,10%	0,22%	0,42%	0,22%	0,48%	0,97%
50º PERCENTIL	0,31%	0,69%	1,40%	0,09%	0,19%	0,34%	0,19%	0,42%	0,81%
75º PERCENTIL	0,52%	1,13%	2,33%	0,14%	0,31%	0,59%	0,32%	0,69%	1,38%
95º PERCENTIL	0,87%	1,91%	3,97%	0,24%	0,52%	1,08%	0,52%	1,18%	2,35%

Apesar do efeito dos dados perdidos no valor do VMDA poder ser desprezível, observa-se que quanto maior o percentual de falhas mais impreciso é a estimativa. Isto pode ser visto através da Figura 3, que mostra que quanto maior a quantidade de dados perdidos, maior é a dispersão dos dados e conseqüentemente menos confiável é a estimativa.

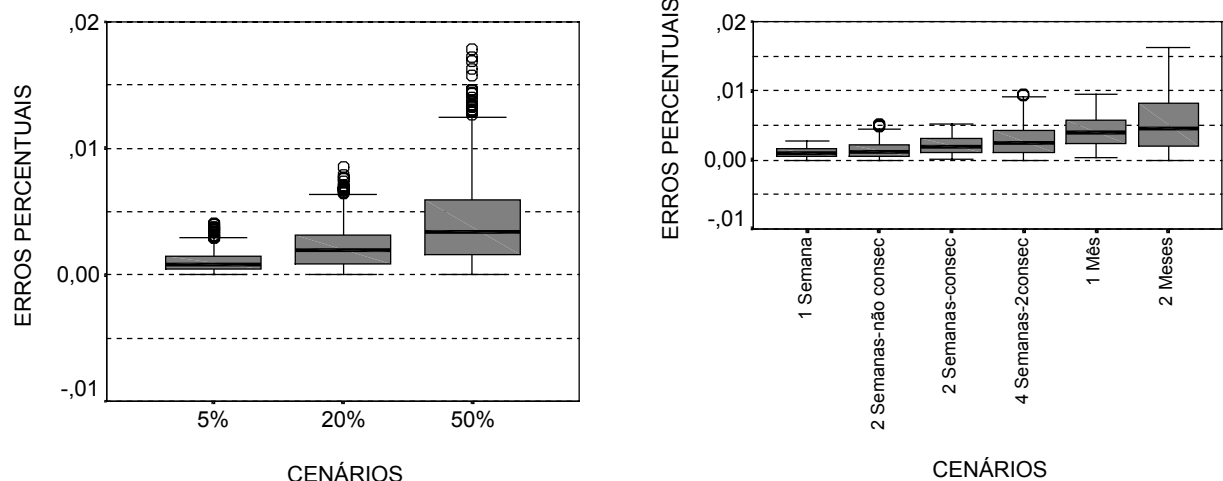


Figura 3: Erros percentuais do posto P02 para os cenários simulados



5. CONSIDERAÇÕES FINAIS

A realização desta pesquisa, com a utilização de dados de contagem dos postos permanentes para as rodovias estaduais do Estado do Ceará, possibilitou uma análise sobre a necessidade do órgão rodoviário de fazer o preenchimento dos dados faltosos, por meio de valores históricos, para o cálculo do VMDA. De forma geral, os resultados encontrados neste estudo vieram confirmar conclusões anteriores (Wright *et al.*, 1997). Além disso, o estudo mostra a dispensabilidade da utilização da prática de *imputation* para o cálculo do VMDA, vindo de encontro com recomendações sugeridas pela AASHTO (1992).

Agradecimentos

Os autores agradecem a cessão dos dados por parte do DERT-CE e aos colegas Francisco Moraes, Hamifrancey Brito Meneses e Paulo Marinho pelas sugestões que muito enriqueceram a pesquisa.

REFERÊNCIAS BIBLIOGRÁFICAS

- Albright, D. (1991) History of Estimating and Evaluating Annual Traffic Volume Statistics. *Transportation Research Record 1305*, Transportation Research Board, Washington, D.C, pp. 103-107.
- AASHTO (1992) *Guidelines for Traffic Data Programs*. America Association of State Highway and Transportation Officials, Washington, D.C.
- FHWA (2001) *Traffic Monitoring Guide*. Federal Highway Administration (2001), U.S. Department of Transportation.
- Zhong, M., Lingras, P.J.; and Sharma, S.C. (2003) Updating Missing Values of Traffic Counts: Factor Approaches, Time Series Analysis versus Genetically Designed Regression and Neural Network Models. *Transportation Research Board, The 82 th Annual Meeting*, Washington, D.C., USA.
- Wright, T., Hu, P.; Young, J.; and Lu, A (1997). Variability in Traffic Monitoring Data, Final Summary Report. *Oak Ridge National Laboratory for the US Department of Energy*.

Marcos José Timbó Lima Gomes (timbo@det.ufc.br)
Maria Elisabeth Pinheiro Moreira (beth@det.ufc.br)
Universidade Federal do Ceará
Departamento de Engenharia de Transportes, Campus do Pici, Bloco 703
CEP 60455-760; Fortaleza - Ceará - Brasil.
Tel: (085) 2874792 - 288.9568; Fax: (085) 288.9572