

CONSOLIDAÇÃO DA BASE DE DADOS DE UM PROGRAMA DE MONITORAMENTO CONTÍNUO DO VOLUME DE TRÁFEGO DAS 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

Este trabalho discute e analisa o programa de monitoramento contínuo do volume de tráfego em rodovias Estaduais do Ceará, os quais geram informações para o planejamento e a operação do tráfego. Propõe um sistema de banco de dado que assegure aos usuários a utilização de somente dados de contagens válidos, filtrando e eliminando as inconsistências do atual sistema de contagem de tráfego do Departamento de Edificações, Rodovias e Transportes do Ceará - DERT. Tal preocupação com a qualidade dos dados disponíveis deve-se ao fato que estes subsidiam a tomada de decisão, que envolve frequentemente a alocação de recursos financeiros para construções/melhorias de infra-estrutura de rodovias.

ABSTRACT

This paper discusses and analyses the program for continuous monitoring of traffic flow volumes in state roads of Ceará, which produces information for traffic planning and operation. A database system is proposed in order to assure the users that only valid data are used, filtering and removing the inconsistencies of the traffic counting system that is currently used by the Department of Building, Roads and Transport of the State of Ceará (DERT). Such care with the data quality is due to the fact that this information supports decision making, which involves the allocation of large amounts of money, and other resources, in the construction or maintenance of roads infrastructure.

1. INTRODUÇÃO

Para obter dados que representem a realidade das condições do tráfego e ter maiores êxitos nos projetos de rodovias, muitos órgãos rodoviários criaram programas de monitoramento de tráfego. Estes programas variam entre os órgãos rodoviários, conforme a aplicação de recursos financeiros e a automatização das contagens de tráfego (FHWA, 2001), e envolve a coleta de muitos tipos de dados como volume, composição, velocidade e pesagem de veículos (MOHAMAD *et al.*, 1998).

Apesar da diversidade, ALBRIGHT (1993) lembra que nenhum programa de monitoramento de tráfego deve ser desenvolvido separadamente. Isto porque várias divisões dentro de um órgão rodoviário coletam dados de tráfego para finalidades específicas e a integração destes em um banco de dados comum reduz a possibilidade de duplicação de coleta de dados de um mesmo local.

Além da integração dos dados coletados em campo, o programa de monitoramento de tráfego deve garantir a qualidade dos dados, assegurando que são válidos, permitindo a edição para a remoção de dados inválidos, armazenando-os e gerando informações estatisticamente confiáveis aos estudos rodoviários.

Assim, o presente trabalho objetiva analisar o programa de monitoramento contínuo do volume de tráfego, enfatizando a qualidade dos dados, através da identificação de dados errôneos e perdidos, sendo proposto um aplicativo que identifique, controle e filtre as possíveis inconsistências do atual sistema de contagem de tráfego das rodovias estaduais do Ceará.

2. PROGRAMA DE MONITORAMENTO DE VOLUME DE TRÁFEGO

Um programa de monitoramento de volume de tráfego concentra-se principalmente na estimação do Volume Médio Diário Anual (VMDA) para todos os segmentos de uma malha rodoviária. O método mais seguro para obter os dados do VMDA é a instalação de contadores que realizam contagens automáticas e contínuas ao longo do ano, denominadas de *Automatic Traffic Recorders* – ATR, nas seções rodoviárias da malha. Porém, isto não é prático por causa do alto custo envolvido na compra e instalação de tais equipamentos (MOHAMAD *et al.*, 1998). Normalmente são adotados procedimentos para estimar este parâmetro, baseado em amostras de contagem, consistindo da instalação de postos que realizam contagens contínuas e de curta duração em toda a malha rodoviária.

Segundo o FHWA (2001), o principal objetivo dos contadores contínuos é o desenvolvimento de mecanismos para converter contagens de curta duração em estimativas médias anuais (fatores de ajustes). Visto que, tais equipamentos normalmente registram (no mínimo) volumes horários por direção (e faixa), mas pode-se gerar informações adicionais, tais como: (i) conhecimento das variações e tendências sazonais, semanais e diárias do volume de tráfego; (ii) desenvolvimento de fatores de distribuição direcional do tráfego; e (iii) determinação do volume horário de projeto.

As contagens de curta duração asseguram a cobertura e diversidade geográfica. No presente estudo, serão abordados somente os dados coletados exclusivamente por postos permanentes.

2.1. Programa de monitoramento contínuo de volume de tráfego nas rodovias estaduais do Ceará

A implantação do Sistema de Contagem de Tráfego nas Rodovias Estaduais do Departamento de Edificações, Rodovias e Transportes do Ceará - DERT, teve início com a execução de um Plano de Contagem do Tráfego no ano de 1994.

Os atuais postos permanentes do DERT são dotados com equipamento eletrônico (modelo AVC100 do fabricante *Pat Traffic Control Corp.*), instalados em locais representativos dos trechos rodoviários, em abrigos de concreto, protegidos de intempéries e vandalismo. O processo de contagem ocorre através de um sensor tipo laço indutivo instalado no pavimento, em ambas direções. Os valores registrados são armazenados em memória por intervalos de uma hora, com capacidade de armazenamento dos dados de 20 dias.

Por o DERT não possuir linhas de transmissão nos postos, as coletas dos dados nos postos permanentes são realizadas por um técnico, utilizando-se de um software do próprio equipamento. Os dados são transferidos para um *notebook*, o qual processa estes dados e os gravam em memória no formato hexadecimal, e posteriormente convertidos para um formato texto ASCII, os quais eram transferidos ao banco de dados computacional do *Sistema de Controle de Tráfego* (CTF), desenvolvido originalmente em linguagem CLIPPER 5.01.

O sistema original possuía somente rotinas para a geração dos fatores de expansão de cada posto e emissão de relatórios sobre o volume de tráfego, sem qualquer ferramenta ou procedimento que pudesse verificar e analisar a qualidade dos dados coletados em campo. Com o passar do tempo, o sistema tornou-se ineficiente e foi abandonado, pois não era robusto para trabalhar com uma grande quantidade de dados armazenados.

2.2. Qualidade de dados de um programa de monitoramento de tráfego

Segundo TURNER (2004) a definição de qualidade de dados é um conceito relativo que pode ter significado diferente para diversos usuários. Por exemplo, dado considerado aceitável por um usuário pode ser inaceitável a outro mais exigente. Então é importante considerar e entender todos os usos dos dados antes de tentar medir ou determinar os níveis de qualidade. Deste modo, os usuários de dados aqui tratados são tipicamente comprometidos em análises e planejamento a médio e longo prazo, e os usos destes dados são de natureza histórica, diferentes dos sistemas de monitoramento e controle de tráfego em tempo real.

Depois dos dados de campo serem coletados e armazenados, é fundamental o planejamento e a implementação de procedimentos ao programa de monitoramento de tráfego, que assegure aos usuários a utilização de somente dados válidos, filtrando e eliminando inconsistências na base de dados. A preocupação na qualidade dos dados disponíveis deve-se ao fato que estes subsidiam a tomada de decisão, que envolve freqüentemente a alocação de recursos financeiros para execução de melhorias na infra-estrutura de rodovias, e a utilização de informações imprecisas ou errôneas pelos estudos de planejamento, pavimentação, estudos de acidentes, etc., podem gerar elevados custos ao órgão rodoviário.

O modo como os órgãos rodoviários identificam e controlam erros dos dados de tráfego é um componente fundamental no programa de monitoramento de tráfego. Assim estes devem manter registros de como são manipulados os dados, e que cada manipulação tenha sido baseada em uma análise estatística. Não se deve excluir dados ou simplesmente substituí-los porque "eles parecem ser errados". Ao invés disso, deve-se estabelecer procedimentos sistemáticos que permitam checar e que possam identificar dados inválidos, com estes controles, e identificar quando esses passos de controle de qualidade tiverem sido executados. Uma medida de qualidade freqüentemente utilizada é a verificação da validade dos dados coletados, baseado em princípios de capacidade de tráfego, tendências típicas ou padrões de tráfego, ou simplesmente percepção e experiências locais. Logo, a validação de dados é uma medida de qualidade que usa dados históricos e regras para a aceitação das atuais medidas de tráfego e uma vez que, algum registro viole tais regras, uma advertência assinala como "registro suspeito", que deve ser verificado por analista antes do armazenamento final no banco de dados.

3. EDIÇÃO DE DADOS DE VOLUME TRÁFEGO

A etapa após a coleta e o armazenamento dos dados é denominada pela AASHTO (1992) de etapa de edição e tem como propósito assegurar que as medidas de campo sejam válidas antes dos cálculos de estatísticas de tráfego. A edição para a validação dos dados é uma das etapas críticas em um programa de monitoramento de tráfego, e os procedimentos podem variar de acordo com o tipo e o uso do dado coletado, sendo importante que cada órgão rodoviário documente tais procedimentos. Os programas de monitoramento preocupam-se basicamente com a identificação de três categorias de erros: valores errôneos para a contagem de tráfego; sucessivos valores idênticos e dados de tráfego perdidos (HU, 2002).

3.1. Valores errôneos para a contagem de tráfego

Compõem esta categoria de erro os valores de máximos e mínimos extremos. Um critério usado para determinar a validade de um dado é compará-lo em relação a medidas históricas de um mesmo local, isto porque alguns padrões de características de tráfego repetem-se regularmente. Os padrões dos dados dos contadores permanentes ajudam a identificar os

potenciais erros das medidas de tráfego. Se o padrão de uma característica específica, tal como o volume diário, difere substancialmente de padrões anteriores durante um determinado mês, ou mesmo dia da semana, pode existir erro, devendo ser considerado um dado suspeito.

É de responsabilidade do analista o questionamento da natureza do dado, sendo responsável por aceitá-lo ou rejeitá-lo. Um novo padrão pode ser identificado e aceito. Se os dados forem rejeitados, o analista deve documentar o motivo para a rejeição.

3.2. Sucessivos valores idênticos

Um exemplo desse tipo de erro acontece quando o contador de tráfego registra consecutivamente o mesmo valor para um período longo de contagem. Como as principais causas de inconsistências em uma base de dados de tráfego são ocasionadas por falhas, ou o mau funcionamento dos equipamentos eletrônicos, tais erros podem estar provocando a repetição de valores sucessivamente.

SCHMOYER, *et al.*, (2001) afirma que valores sucessivos iguais a zero são freqüentemente usados para denotar o mau funcionamento dos equipamentos, mas valores diferentes de zero podem ocorrer, principalmente durante o período da madrugada. O autor acrescenta ainda que, a possibilidade de valores repetidos diminuírem com o aumento do tráfego e com o tempo considerado (por exemplo, 4 horas consecutivas com o mesmo valor). Assim, a possibilidade de reais valores consecutivos ocorrerem depende da localização e da hora do dia, enfatizando a importância do conhecimento do padrão de variação de tráfego dos locais monitorados por postos permanentes para ser usados em análises e edições preliminares.

Por fim, o mais comum tipo de erro encontrado em quaisquer programas de monitoramento de tráfego é a falta de dados existentes, ocasionados entre outros motivos por falhas dos equipamentos eletrônicos.

3.3. Dados de tráfego 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 dados perdidos nos postos permanentes, estes não podem ser usados para os cálculos de estatísticas anuais de tráfego, como por exemplo, do Volume Médio Diário Anual. 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 é recomendada, pois pode introduzir erros que não são quantificáveis, mas os órgãos rodoviários que utilizam tais técnicas devem: (i) identificar os valores perdidos na base de dados, que foram estimados, para assegurar o Princípio Básico da Integridade dos Dados (as medidas de tráfego devem ser armazenadas sem quaisquer modificações ou ajustes); (ii) documentar os procedimentos usados para a estimação de dados perdidos e informar qual a magnitude da *imputation* no cálculo das estatísticas de tráfego; e (iii) considerar que os valores perdidos estimados são medidas temporárias que devem ser abandonadas, visto que tais técnicas não são tão importantes quanto à utilização e tratamento dos reais dados coletados.

Entretanto, SHARMA et al., (2003) afirma que a *imputation* de dados não é proibitiva e que, para algumas análises avançadas, pode-se realizar a estimação de dados faltosos. Uma das técnicas de *imputation* mais populares são os fatores de ajustes, desenvolvidos a partir de dados históricos coletados em postos de contagem contínua. A virtude desta técnica é a simplicidade, porém, ZHONG et al., (2002) acrescenta que os resultados obtidos normalmente são menos precisos que os modelos mais sofisticados que se utilizam, por exemplo, de inteligência artificial (redes neurais, algoritmos genéticos, etc).

Mas, apesar dos bons resultados obtidos pelas técnicas e modelos mais complexos para estimação de dados faltosos e dependendo da variável que se deseja obter, é importante o órgão rodoviário investigar a dispensabilidade de tais métodos. Em um estudo realizado por WRIGH et al., (1997), que utilizaram dados de postos permanentes das rodovias do estado da Flórida, EUA, simulou falhas em uma base de dados de tráfego e concluíram que a perda aleatória dos dados tem influências desprezíveis no cálculo do Volume Médio Diário Anual (VMDA).

Há também outros procedimentos que podem reduzir bastante o efeito dos dados faltosos nas estatísticas de tráfego, sem necessidade de usar complexas técnicas de substituição de valores. A AASHTO (1992) adotou um método modificado para calcular o VMDA diretamente com dados perdidos. Neste método, primeiramente calcula-se uma média para cada dia da semana, dentro de cada mês do ano. Com estes 84 valores (12 meses x 7 dias da semana) é calculada uma média anual para cada dia da semana. Por fim, com a média destes sete valores é obtido o VMDA. Este processo efetivamente remove a maioria dos erros dos dias perdidos, especialmente quando estes dados são distribuídos de forma desigual por meses ou dias da semana, visto que esta prática assume que os dias do ano têm o mesmo peso no cálculo do VMDA.

Assim, antes do órgão rodoviário adotar alguma técnica de *imputation*, deve ser realizada uma análise sobre a magnitude da perda dos dados e a influência destas falhas nas estatísticas de tráfego desejadas. Vale lembrar também que um bom monitoramento das falhas (falta de energia, problemas de comunicação e outros equipamentos) e a manutenção dos contadores automáticos são mais importantes que a preocupação com tais técnicas.

4. NOVO SISTEMA DE BANCO DE DADOS DE CONTAGEM DE TRÁFEGO

Um novo banco de dados foi desenvolvido e alimentado com os dados brutos (formato hexadecimal convertido no formato ASCII) dos doze postos permanentes instalados nas rodovias estaduais, obtidos entre os anos de 1996 a 2002.

Neste novo sistema, a principal preocupação foi com a qualidade dos dados registrados pelos contadores automáticos, sendo criados procedimentos que visavam identificar, controlar e filtrar possíveis inconsistências. Para isso foi incluído o campo TRF_ERRO, onde os vários tipos de erros eram identificados (ver Tabela 1), e o campo TRF_CONDICAO que dar a condição de validade ou não do valor “suspeito”, sendo ‘0’ caso o erro fosse aceitável, ou ‘1’ para um erro inaceitável. Foi também criada uma rotina que realiza a importação dos arquivos ASCII, comparando com padrões históricos do mesmo posto ou valores limites máximos, e checando eletronicamente a existência de valores extremos, nulos ou repetidos.

Tabela 1: Tipo de erro identificado pelo novo sistema de contagem de tráfego

Tipo de Erro	Descrição dos Erros
0	Sem problema
1	Arquivo com menos de 48 registros
2	Soma total do volume de tráfego diário igual a zero
3	Soma do volume de tráfego diário de um dos sentidos igual a zero
4	Volume de tráfego diário fora do intervalo aceitável
5	Ultrapassa porcentagem máxima de desequilíbrio direcional
6	Ultrapassa o volume horário máximo por faixa
7	Ultrapassa a quantidade de horas com o mesmo volume de tráfego
8	Ultrapassa a quantidade de horas com o volume de tráfego igual a zero

A parte inicial da rotina computacional, denominada ‘Padrão do Tráfego Diário’, foi utilizada para a identificação de valores extremos do volume diário. Assim, ao ser executada a importação dos dados de determinado posto permanente, verifica-se o volume de um determinado dia estava em um intervalo de ± 2 desvios padrões em relação ao volume médio anual do ano anterior (ano base), para o mesmo dia da semana no mesmo local. A utilização de ± 2 desvios padrões deve-se ao fato de que o tamanho da amostra permite afirmar que a variável volume diário aproxima-se de uma distribuição normal e que, neste caso 95,5% dos valores estariam neste intervalo considerado.

Caso o volume do dia importado estivesse fora do intervalo considerado, não necessariamente este era um valor inválido. Na verdade esta é uma condição de advertência que identifica um valor suspeito e neste caso deve-se inspecionar os dados, a magnitude do volume diário, o intervalo e o desvio padrão considerado verificando se não corresponde a um feriado, ou uma data de influência (data próxima a feriados).

A segunda parte do aplicativo, denominada ‘Padrão de Tráfego Direcional Diário’, foi usada para tentar examinar problemas de mau funcionamento dos laços detectores dos contadores que possam provocar um desequilíbrio no tráfego direcional, visto que na maioria das rodovias bidirecionais os picos da manhã e da noite ocorrem em direções opostas. Observa-se ainda que, há limites do uso da distribuição direcional como um indicativo de mau funcionamento dos contadores. Um desequilíbrio direcional pode ser provocado por limitações das condições típicas, de operação da rodovia, como por exemplo, o fechamento de uma pista de tráfego, ou ainda, em áreas turísticas, onde em algumas rodovias, o tráfego da sexta-feira estará desequilibrado em uma direção e o tráfego do domingo na direção oposta. Estes padrões típicos de tráfego devem ser registrados de forma que dados futuros possam ser melhores editados contra informação do tráfego atual.

No aplicativo foi usada a média de oito contagens manuais realizadas nos dois sentidos dos doze postos permanentes para definir um valor percentual de desequilíbrio direcional de 60% com aceitável. Se este valor fosse superado, uma mensagem de advertência era emitida para a aceitação ou não dos dados.

A terceira e última parte do aplicativo busca valores limites máximos como indicativos de erros, como por exemplo, um volume horário máximo por sentido de tráfego, ou uma quantidade máxima de sucessivos valores idênticos. Para os valores sucessivos maiores que ‘zero’ foi considerado que o mesmo volume não deveria ocorrer por mais de 4 horas consecutivas por sentido de tráfego. Este critério foi usado como adequado para a maioria dos

postos, mais para alguns postos de baixo volume de tráfego, utilizou-se 5 horas consecutivas, visto que em alguns períodos a repetição do mesmo volume horário não era rara e a quantidade de dados suspeitos era grande. Já para valores sucessivos iguais a zero foi considerado para todos os postos que não poderia haver um volume igual a zero por 8 horas consecutivas.

5. AVALIAÇÃO DOS DADOS DOS CONTADORES PERMANENTES DO DERT

Após a importação dos dados, realizaram-se consultas no banco de dados e construí-se gráficos para melhor avaliar a qualidade e quantidade dos dados de cada contadores contínuos. A Figura 1 mostra o percentual médio de dados válidos, inválidos e perdidos de todos os postos permanentes durante os anos de 1996 a 2002.

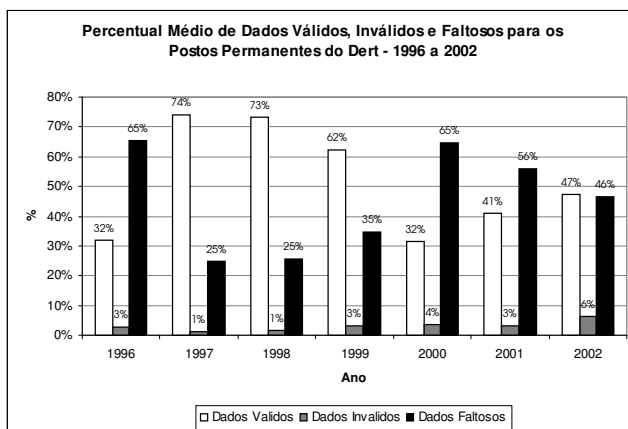


Figura 1: Avaliação do percentual de dados válidos, inválidos e faltosos dos postos permanentes do DERT

Considerando os sete anos da série histórica e todos os postos permanentes, o percentual médio é de 52% de dados válidos, 3% de dados inválidos, e 45% de dados perdidos. Este alto percentual de dados perdidos apresenta-se mais evidente a partir do ano de 2000, e decorre principalmente da falta de planejamento de um cronograma de coleta dos dados armazenados nos contadores, aliado ao elevado custo de deslocamento de um técnico para realização de tal atividade, e ausência de um serviço de manutenção dos equipamentos eletrônicos. Dos doze postos de contagens permanentes, os que apresentaram maiores percentuais de dados perdidos, durante os sete anos da série histórica, são os postos P004 e P005, com valores de 58% e 67% respectivamente, devido principalmente a questões de manutenção dos equipamentos.

Apesar de como sugerido por WRIGH *et al.*, (1997) de que a perda aleatória dos dados tem influências desprezíveis no cálculo do VMMA ou do procedimento de cálculo da AASHTO, que tenta reduzir o efeito dos dados faltosos no valor do VMMA, atualmente o único motivo da existência dos postos permanentes nas rodovias estaduais do Ceará é a obtenção de fatores de expansão para converter contagens de curta duração em estimativas anuais de tráfego. Então, a utilização de tais fatores com esta grande quantidade de dados perdidos, podem ter influência na veracidade das estimativas anuais obtidas a partir de contagens de curta duração.

6. CONCLUSÕES

Segundo a AASHTO (1992), qualquer banco de dados de um programa de monitoramento deve indicar ao usuário a qualidade e quantidade de dados que estão sendo usados nos cálculos das estatísticas de tráfego de um determinado local.

A edição para a validação dos dados é uma etapa crítica de qualquer programa de monitoramento de tráfego e os procedimentos variam de acordo com o tipo e o uso destes dados. Um aplicativo foi proposto, com a finalidade de identificar, controlar e filtrar possíveis inconsistências no programa de monitoramento contínuo do volume de tráfego das rodovias Estaduais do Ceará, para a avaliação da qualidade destes dados. A qualidade dos dados dos contadores permanente do DERT foi avaliada por meio da magnitude dos dados válidos, inválidos e faltosos para os anos da série histórica.

Verificando os dados, conclui-se que não é possível estimar valores perdidos para a variável volume de tráfego, pois o alto percentual de dados faltosos na série histórica não permite avaliar e quantificar os erros que se estariam comentando na estimação de tais valores, nem definir o modelo ou técnica mais adequada para a realização de tal tarefa.

Alguns procedimentos poderiam ajudar na melhora da qualidade dos dados, tais como, um bom monitoramento das falhas (falta de energia ou defeitos no equipamento), redução dos tempos de manutenção preventiva/corretiva dos contadores automáticos e principalmente um estudo para implantação de transferência de dados por meio remoto.

REFERÊNCIAS BIBLIOGRÁFICAS

- AASHTO (1992) *Guidelines for Traffic Data Programs*. America Association of State Highway and Transportation Officials, Washington, D.C.
- 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.
- Albright, D. (1993) Standards, Innovation, and the Future of Traffic Monitoring. *Institute of Transportation Engineers Journal*, USA.
- FHWA (2001) *Traffic Monitoring Guide*. Federal Highway Administration (2001), U.S. Department of Transportation, Washington, D.C., May 2001.
- Mohamad, D.; K. C. Sinha; T. Kuczek e C. F. Scholer (1998) Annual Average Daily Traffic Prediction Model for County Roads. *Transportation Research Board, The 77th Annual Meeting*, Washington, D.C., USA.
- Hu, P. (2002) *Costs and Benefits of Using ITS as An Alternative Data Source: A Case Study*, Transportation Research Board, Washington, D.C., USA.
- Schmoyer, R.; P. Hu e R. Goeltz (2001) Statistical Data Filtering and Aggregation to Hour Totals of ITS Thirty-Second and Five-Minute Vehicle Counts. *Transportation Research Board, The 80th Annual Meeting*, Washington, D.C., USA.
- Sharma, S.; P. Lingras e M. Zhong (2003) Effect of Missing Value Imputations on Traffic Parameters Estimations from Permanent Traffic Counts. *Transportation Research Board, The 82th Annual Meeting*, Washington, D.C., USA.
- Turner, P. E. S. (2004) Defining and Measuring Traffic Data Quality. *Transportation Research Board, The 83th Annual Meeting*, Washington, D.C., USA.
- Wright, T.; P. Hu; J. Young e A. Lu (1997). Variability in Traffic Monitoring Data, Final Summary Report. *Oak Ridge National Laboratory for the US Department of Energy*.
- Zhong, M.; P. Lingras; E. S. C. Sharma (2002) Updating Missing Values of Traffic Counts: Factor Approaches, Time Series Analysis versus Genetically Designed Regression and Neural Network Models. *Transportation Research Part C*, United Kingdom, 2002.