

USO DE REDES NEURAIIS PARA PREVISÃO DE PRODUÇÃO DE VIAGENS: UMA ANÁLISE AGREGADA

Samille Santos Rocha
Marcela Navarro Pianucci
Cira Souza Pitombo
André Luiz Cunha
Universidade de São Paulo

Escola de Engenharia de São Carlos/Departamento de Engenharia de Transportes

RESUMO

Redes Neurais Artificiais (RNAs) são modelos matemáticos alimentados por informações contidas em um banco de dados, que simulam o mecanismo de aprendizagem do cérebro humano. Desta forma, uma rede neural pode ser entendida como um cérebro artificial, que “aprende” uma determinada informação a partir de envio repetitivo de exemplos que traçam certas características. Entende-se que a análise de viagens urbanas reflete certos tipos de comportamentos, tanto individual, quanto de uso do solo, e que estas características podem ser introduzidas no ambiente de um cérebro artificial, para, por exemplo, se conhecer o comportamento relativo a viagens. Este trabalho tem o objetivo de utilizar as RNAs para estimar a produção de viagens por zona de tráfego na Região Metropolitana de São Paulo. Os dados utilizados foram da Pesquisa Origem-Destino de 2007. Além de RNAs, se estimou a produção de viagens através da tradicional técnica de Regressão Linear Múltipla, como técnica confirmatória. A análise dos resultados de ambas as técnicas mostrou que os dois modelos tiveram um bom poder preditivo, mostrando finalmente a viabilidade das RNAs para estudos de demanda por transportes.

ABSTRACT

Artificial Neural Networks (ANNs) are mathematical models fed by information contained in a database, which simulate the learning mechanism of the human brain. Thus, the neural network can be understood as an artificial brain that “learns” from a particular information repetitive transmission of examples mapping certain characteristics. It is understood that the analysis of urban trip reflects some types of behavior, both individually and land use, and that these features can be introduced into an artificial brain, for example, to investigate the travel behavior. The aim of this paper is to use the ANN to estimate trip production by traffic analysis zone in the São Paulo Metropolitan Area. The data used were from the Origin-Destination Survey, carried out in 2007. Beyond neural networks analysis, trip production was estimated by a traditional technique of Multiple Linear Regression, as a confirmatory technique. The results of both techniques showed that two models presented a good accuracy, finally showing the viability of ANNs to transportation demand forecasting.

1. INTRODUÇÃO

A geração de viagens, primeira etapa do método sequencial Quatro Etapas, envolve a previsão do total de viagens produzidas ou atraídas à zona de tráfego, levando-se em conta fatores agregados socioeconômicos, de uso do solo ou acessibilidade. Os modelos mais tradicionais e usuais para estimativas de geração de viagens são modelos lineares múltiplos, fator de crescimento e classificação cruzada (Ortúzar e Willumsen, 2011)

No entanto, a Regressão Linear Múltipla (RLM), embora amplamente popular na área de modelagem de demanda por transportes, não seria a técnica mais indicada para previsão de viagens urbanas. Viagens são dados de contagem e não atendem à suposição de normalidade, por exemplo. Além disso, a relação existente entre produção de viagens e variáveis explicativas, como população, por exemplo, não é linear. Adicionalmente, em geral os resíduos provenientes de modelos de atração ou produção de viagens não são distribuídos aleatoriamente com média zero. Além disso, também não atendem à suposição de homoscedasticidade.

Outras considerações são importantes nas análises de tais modelos, como: (1) O papel da intersecção da equação (normalmente espera-se que a reta passe pela origem do sistema); (2) Valores relativos ou absolutos para as variáveis independentes; (3) Zonas nulas; (4) O tamanho

e homogeneidade das zonas de tráfego e (5) O problema da não linearidade entre variável dependente e independentes.

O uso de Redes Neurais Artificiais (RNAs) vem se tornando cada vez mais popular nas últimas décadas com o avanço computacional. RNAs são aplicáveis em diversas áreas da engenharia, incluindo modelagem de problemas de transportes considerando sua grande habilidade de extrair informação a partir de grande base de dados. As RNAs foram inspiradas no funcionamento do cérebro humano. Os modelos matemáticos criados por RNAs são adquiridos através de experiências armazenadas e captadas por um algoritmo de aprendizagem, que modificam os pesos sinápticos ou sinais de comunicação. Segundo Haykin (2001), a RNA é uma máquina projetada para modelar a maneira como o cérebro realiza determinada função.

Inicialmente, as redes neurais passam pelo processo de aprendizagem, ou seja, se introduz um conjunto de exemplos e, a partir disto, se extrai as características necessárias para representação de algum tipo de informação, similares aos exemplos introduzidos nas redes. Segundo Bocanegra (2002), por ter esta característica de aprender e generalizar o conhecimento, as RNAs podem apresentar melhor desempenho do que os métodos tradicionais. Este autor utilizou um modelo de RNA treinada para estimar o potencial de viagens.

Muitos trabalhos são encontrados na área de planejamento de transportes, como o de Raia Jr. (2000) que desenvolveu um modelo destinado a estimar um índice potencial de viagens para o planejamento urbano através de dados da pesquisa OD e o uso das RNAs. No trabalho de Mark et al. (2004), os autores utilizaram as RNAs para prever o tempo de viagem entre dois pontos da rede de transportes de uma autoestrada. Já Lin et al. (2005) concluíram que as RNAs têm um potencial maior que outros métodos estatísticos, para estimar e prever o tempo de viagem em redes viárias urbanas.

Colombaroni e Fusco (2011) propuseram RNA para modelar o comportamento dos motoristas de automóveis. Os veículos do experimento foram colocados em fila na via urbana, acoplados com um GPS para a coleta dos dados. Verificaram que as RNAs fornece uma boa aproximação dos padrões de condução e pode ser adequadamente aplicado em modelos de micro simulação.

Alguns autores, como Kikuchi et al. (1993) e Chin et al. (1994) fizeram o uso das RNAs para analisar a geração de viagens com base em características socioeconômicas e determinaram uma matriz OD com base no volume de tráfego para três horários de pico, respectivamente. Observam-se também alguns trabalhos que utilizaram técnicas de RNAs para estimativa de distribuição de viagens (Rasouli e Nikras, 2013, Gonçalves et al., 2015).

O objetivo deste artigo é analisar o potencial das RNAs para estimar viagens produzidas por zona de tráfego. Os resultados obtidos com as RNAs serão comparados aos resultados de um modelo linear tradicional. Neste trabalho serão utilizados dados agregados da Pesquisa Origem-Destino, realizada em 2007, na Região Metropolitana de São Paulo. As próximas seções deste artigo descrevem as etapas metodológicas realizadas. A seção 2 apresenta o referencial teórico relativo à técnica abordada. Em seguida, a seção 3 e suas respectivas subseções apresentam o banco de dados utilizado, a aplicação das RNAs e do modelo linear calibrado para previsão de produção de viagens. Finalmente, na seção 4 são discutidas as principais conclusões deste trabalho.

2. REDES NEURAIS ARTIFICIAIS (RNAs)

As Redes Neurais Artificiais (RNAs) são utilizadas em diversas áreas do conhecimento, porém, somente nesta última década passaram a ser utilizadas com mais frequência na Engenharia de Transportes. As RNAs são sistemas de computação adaptativos, inspirados nas características de processamento de informação encontradas nos neurônios reais e nas características de suas interconexões (Haykin, 2001).

São capazes de processar uma grande quantidade de dados e fazer previsões com acurácia. Tem poder de generalização, no que se refere à capacidade que elas possuem de produzir saídas razoáveis, a entradas que nunca lhe tenham sido apresentadas, solucionando problemas complexos (Silva, 2003).

A RNA é baseada no comportamento das redes neurais biológicas. O cérebro humano apresenta uma enorme capacidade de processar informações, tais como, o reconhecimento da fala e a segmentação de imagens. Devido a isso, pesquisadores tentam reproduzir as reações do cérebro em máquinas e tem-se obtido resultados promissores, apesar do pouco conhecimento do cérebro humano. As redes são formadas por um conjunto de neurônios artificiais que interagem entre si e apresentam comportamentos semelhantes ao funcionamento dos neurônios biológicos (Rodrigues da Silva et al., 2008).

O neurônio forma a base para o projeto de RNAs. O neurônio artificial, que simula o comportamento do neurônio biológico, é uma unidade de processamento matematicamente simples. Este recebe uma ou mais entradas, que correspondem às conexões sinápticas com outras unidades similares a ele, com seus respectivos pesos e transforma em saídas, cujos valores dependem diretamente da somatória ponderada de todas as saídas dos outros neurônios a esse conectado (Haykin, 2001).

A Figura 1 apresenta um modelo não linear de um neurônio, onde o processo sináptico é representado pelos pesos (ω) que amplificam cada um dos sinais recebidos e a função de ativação modela a forma como o neurônio responde ao nível de excitação, limitando e definindo a saída da rede neural.

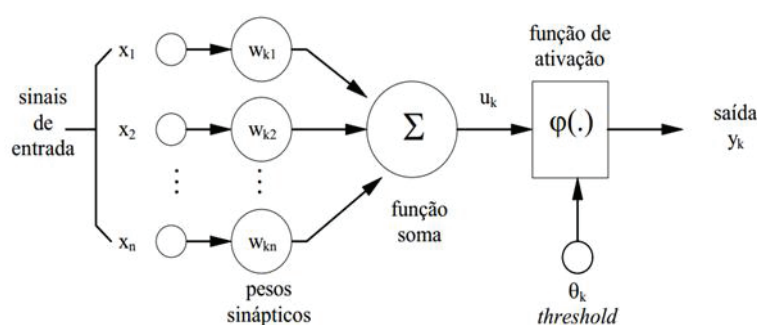


Figura 1 Modelo não linear de um neurônio (Haykin, 2001)

Existem três elementos básicos do modelo neural apresentado acima: (1) Sinapses: são caracterizadas por um peso w , e a função do w é multiplicar o sinal x_j na entrada da sinapse j , conectado ao neurônio k . Se positivo o w é sinapse excitatória, se negativo o w é sinapse inibitória; (2) Somatório: somam os sinais de entradas, ponderados pelas respectivas sinapses do neurônio, realizadas por combinações lineares; (3) Função de ativação: limita a amplitude

da saída de um neurônio. A entrada é normalizada dentro de um intervalo fechado $[0,1]$ ou $[-1,1]$. A função de ativação pode ser do tipo limiar, linear ou sigmoide e esta irá depender do objetivo do estudo.

Os neurônios de uma rede neural podem estar dispostos de diversas formas, ou seja, a estrutura vai depender do algoritmo de aprendizado usado para treinar a rede. De acordo com Haykin, (2001) a arquitetura das redes se apresenta em três diferentes tipos: as redes progressivas de única camada, as redes progressivas de camadas múltiplas e as redes recorrentes.

A rede progressiva de múltiplas camadas (*multilayer perceptron*) apresenta uma ou mais camadas escondidas, cuja função é intervir entre a entrada externa e a saída da rede, tornando a rede capaz de extrair estatísticas de ordem mais alta, que as redes de única camada. A Figura 2 mostra uma rede de múltiplas camadas (10-4-2), com 10 nós de entrada, quatro neurônios escondidos e dois neurônios de saída, totalmente conectada. Neste trabalho, foi utilizado o modelo *multilayer perceptron*. Este modelo apresenta maior poder computacional, em termos de precisão e rápida convergência, que aquele apresentado pelas redes sem camadas intermediárias. Além disso, possui habilidade de aprender da experiência através de treinamento e permite grande quantidade de dados de entrada.

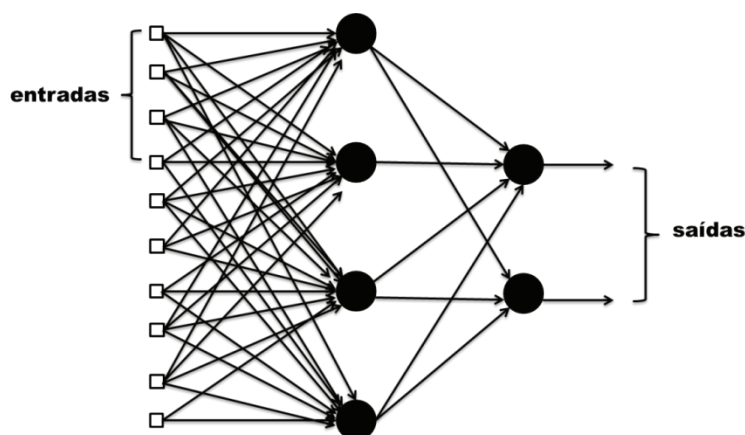


Figura 2 - Rede progressiva de múltiplas camadas

3. PROCEDIMENTOS METODOLÓGICOS

Esta seção apresenta as principais etapas metodológicas utilizadas neste trabalho. Nela, primeiramente, apresenta-se a área de estudo e são descritos os dados utilizados. Além, disso, são apresentadas as duas técnicas aplicadas e os resultados obtidos.

3.1 DADOS

Os dados utilizados neste trabalho são referentes à Pesquisa O/D realizada na Região Metropolitana de São Paulo (RMSP) em 2007, pela Companhia do Metropolitano de São Paulo. A pesquisa levantou informações relativas a 30 mil domicílios, escolhidos aleatoriamente e distribuídos nas 460 zonas de tráfego e 39 municípios, conforme Figura 3.

A amostra final utilizada é composta por 460 observações (zonas de tráfego) e 80 variáveis agregadas numéricas relativas à população, empregos, matrículas escolares, automóveis, etc. Em geral, tais variáveis são altamente correlacionadas. Medidas para evitar multicolinearidade

foram realizadas na etapa de calibração do modelo linear tradicional. As variáveis independentes (por zona de tráfego) utilizadas foram:

1. Produção de viagens; Número de domicílios; 3. Número de famílias; 4. População; 5. Matrículas Escolares; 6. Empregos; 7. Número de automóveis particulares; 8. Área total;
9. População por faixa etária (11 variáveis);
10. População por grau de instrução (5 variáveis);
11. População por gênero (2 variáveis);
12. População por Renda Familiar (em Reais de Outubro de 2007) – (5 variáveis);
13. Renda Média Familiar; 14. Renda Média per capita;
14. Matrículas por Tipo de Escola (2 variáveis);
15. População por Condição de Atividade (8 variáveis);
16. População que Trabalha por Vínculo Empregatício (9 variáveis);
17. Famílias por Número de Automóveis Particulares (5 variáveis);
18. Empregos por Setor (7 variáveis);
19. Serviços (9 variáveis);
20. Empregos por Vínculo Empregatício (11 variáveis).

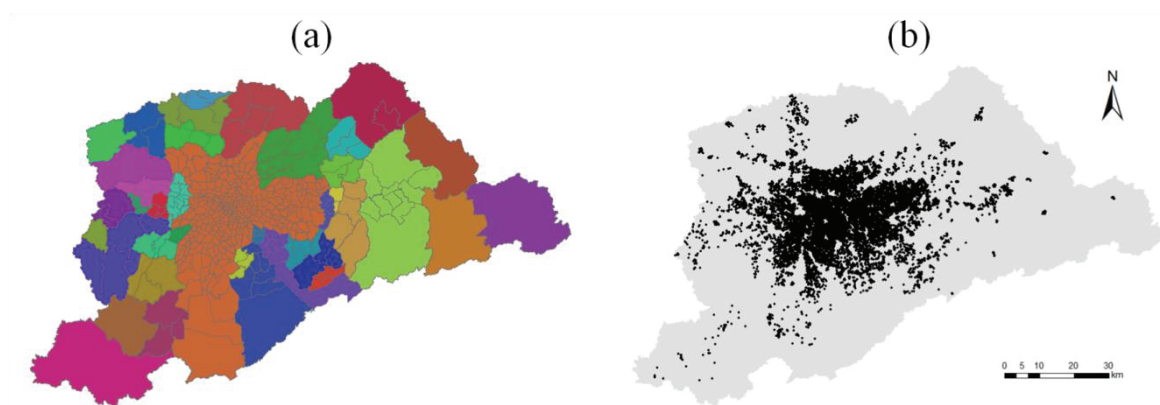


Figura 3 Representação da RMSP (a) e dos domicílios entrevistados pela Pesquisa O/D (b).
Fonte: Lindner (2015)

As variáveis utilizadas foram normalizadas com o objetivo de manter um formato uniforme e diminuir a grande variabilidade dos dados, mantendo os dados em uma escala de 0 a 1. A fórmula para normalização dos dados é mostrada na Equação 1.

$$N = \frac{x - x_{i \min}}{x_{i \max} - x_{i \min}} \quad (1)$$

N : valor normalizado da variável;

x : valor absoluto da variável (produção de viagens);

$x_{i \min}$: menor valor de uma variável i ;

$x_{i \max}$: maior valor de uma variável i .

Formado o banco de dados normalizados, as análises feitas através das ferramentas anunciados, foram realizadas no *software* IBM SPSS, versão 22. A Tabela 1 traz as principais medidas descritivas de algumas variáveis agregadas que fizeram parte da análise. Vale ainda ressaltar

que 70% da amostra foi utilizada para treinamento da RNA e calibração do modelo linear. Os dados restantes (30% da amostra final) foram utilizados para teste da rede e validação do modelo paramétrico.

Tabela 1 Medidas descritivas de algumas variáveis agregadas

MEDIDAS DESCRITIVAS								
Variável Dependente	Média	Mediana	Desvio Padrão	Mínimo	Máximo	Quartil25	Mediana	Quartil75
PRODUÇÃO	0,171	0,1405	0,149	0	1	0,074	0,1405	0,2255
Variável Independente								
Matrículas Escolares	0,144	0,094	0,155	0	1	0,038	0,094	0,1965
População Assalariada com Carteira	0,108	0,058	0,131	0	1	0,023	0,058	0,148
Emprego-Setor Comércio	0,112	0,085	0,114	0	1	0,039	0,085	0,154
População com nível Superior Completo	0,132	0,104	0,133	0	1	0,032	0,104	0,189
Número de Estudantes	0,110	0,061	0,138	0	1	0,018	0,061	0,1405
Número de Empregos sem Endereço Fixo	0,134	0,101	0,131	0	1	0,05275	0,1005	0,171
Empregos Assalariados com Carteira	0,127	0,084	0,140	0	1	0,04	0,084	0,16725
População de 23 a 29 anos	0,084	0,045	0,104	0	1	0,01775	0,045	0,1205
População Em Licença	0,052	0,013	0,099	0	1	0	0,0125	0,063
População com Ginásio Completo Colegial Incompleto	0,121	0,065	0,151	0	1	0,019	0,065	0,172
População Renda Familiar R\$ 3040 a 5700	0,123	0,084	0,138	0	1	0,029	0,0835	0,16925
População Renda Familiar R\$ 5700 ou mais	0,102	0,056	0,134	0	1	0	0,056	0,14525
População com Colegial Completo Superior Incompleto	0,116	0,070	0,137	0	1	0,025	0,07	0,16825

3.2 APLICAÇÃO DAS RNAs

As redes neurais foram utilizadas para prever produção de viagens por zona de tráfego. A rede definiu como principal variável preditora da produção de viagens a variável *População com Primário completo e Ginásio Incompleto* (Figura 4). Em seguida a variável *Área* e a variável *População entre 30 e 39 anos* foram classificadas como segunda e terceira variáveis explicativas mais importantes, respectivamente. Observa-se que possivelmente tais variáveis, relacionadas à população em geral, são altamente correlacionadas. No entanto, a técnica de RNAs não teria a limitação de multicolinearidade. Desta forma, foram definidas como *input* 80 covariáveis.

Para prever a produção de viagens foi definida uma arquitetura personalizada de 1 camada oculta. A amostra de treinamento da rede é equivalente a 326 zonas de tráfego, enquanto que amostra de teste é composta por 134 zonas. A função de erro utilizada para ser minimizada é a soma dos erros quadráticos. A Tabela 2 resume os principais dados para processamento da rede.

Tanto os resultados para treinamento (70% da amostra) quanto para teste (30% da amostra) ficaram bons, com erros baixos. Foram consideradas algumas medidas de desempenho tanto para o modelo de RNAs *multilayer perceptron* como para o modelo linear tradicional. Tais medidas são descritas a seguir. Os resultados das medidas de desempenho são apresentados na Tabela 3.

$$EM = \frac{1}{N} \cdot \sum_{i=1}^N (x_i - y_i) \quad (2)$$

EM = erro médio; x_i = valor observado; y_i = valor previsto.

$$RME = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^N (x_i - y_i)^2} \quad (3)$$

REM = Raiz do erro médio

$$ER = \frac{\sum_{i=1}^N (x_i - y_i)^2}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad (4)$$

ER = Erro relativo

$$r = \frac{1}{N-1} \cdot \sum_{i=1}^N \frac{(x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sigma_x \cdot \sigma_y} \quad (5)$$

r = coeficiente de correlação; $\bar{x}$ = valor médio observado; $\bar{y}$ = valor médio estimado; σ_x = desvio padrão dos valores observados; σ_y = desvio padrão dos valores estimados.

$$desvpad(x - y) = \sqrt{E((x - y) - \mu)^2} \quad (6)$$

Desvpad (x-y) – desvio padrão dos erros; μ = média dos erros; x-y = erros.

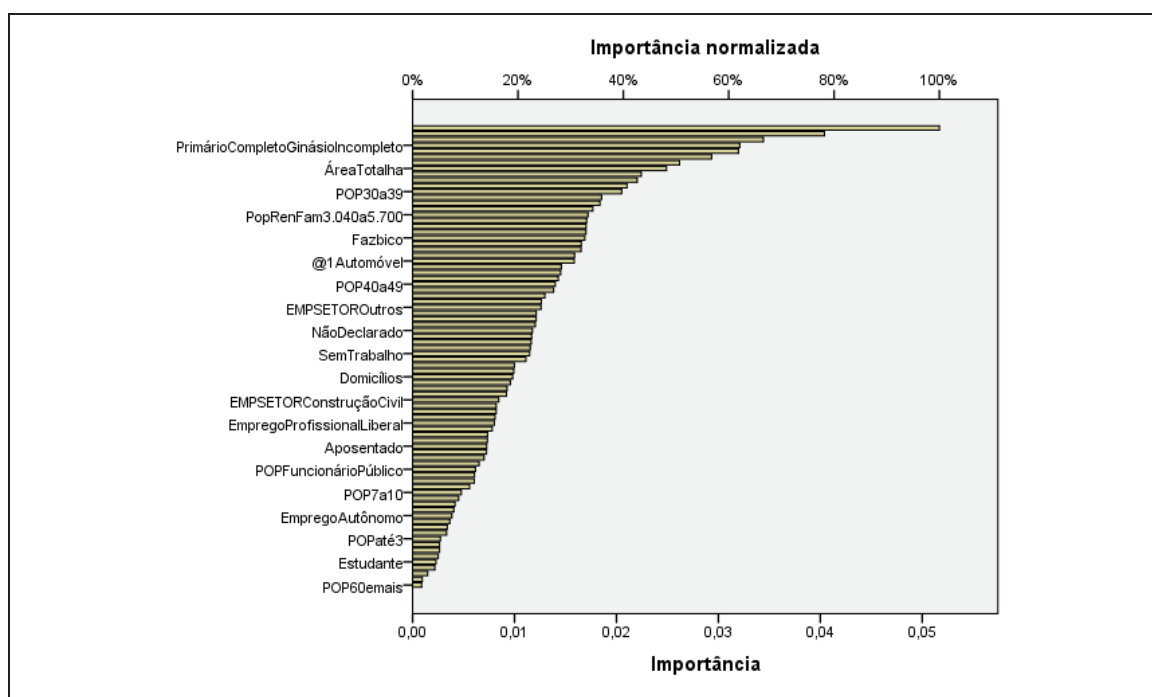


Figura 4 Importância das variáveis produção de viagens

Tabela 2 Dados para processamento da RNA

Processamento RNA-PRODUÇÃO		
Amotra	Treinamento	326
	Teste	134
Camada de entrada	Covariáveis	80
Camadas ocultas	Número de camadas	1
	Número de unidades	2
	Função de ativação	Tangente Hiperbólica
Camada de saída	Variáveis	1
	Número de unidades	1
	Função de ativação	Identidade
	Função de erro	Soma dos Quadrados

Tabela 3 Medidas de desempenho calculadas para RNA

PRODUÇÃO		
ERROS	TREINO	TESTE
	70%	30%
EM	0,003	-0,002
RME	0,029	0,048
ER	0,034	0,170
CORREL	0,983	0,911
DESVPAD	0,029	0,048

Estabelecido o modelo, foi possível verificar os resultados da previsão da produção de viagens nas zonas de tráfego da RMSP. Além da saída com os valores previstos, o IBM SPSS 22.0 fornece gráfico de dispersão entre valores observados e previstos e gráficos de resíduos. Desta maneira, há uma forma eficiente e visual de verificar a qualidade dos resultados obtidos a partir das RNAs. A Figura 5 ilustra estes resultados. Verifica-se que há uma relação linear entre valor estimado e observado. Além disso, o gráfico de resíduos indica que a maior parte dos resíduos tem distribuição aleatória com média muito próxima à zero.

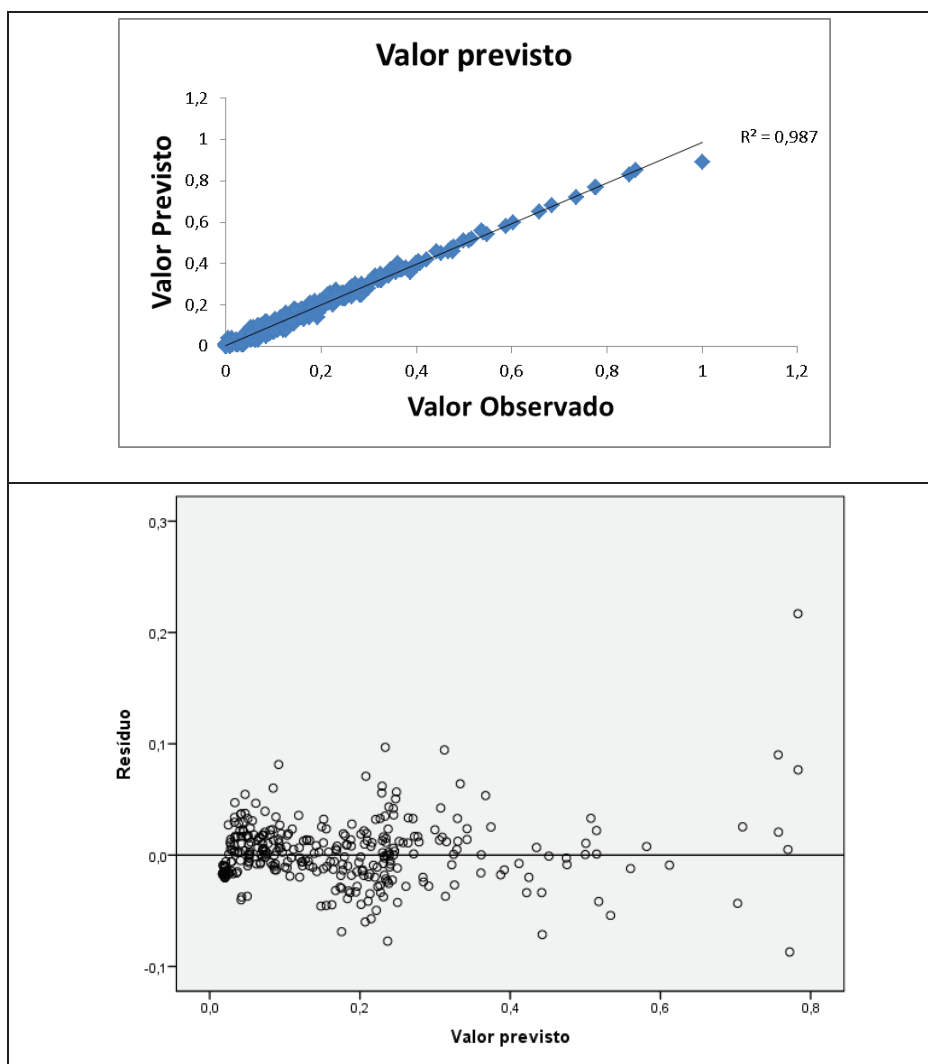


Figura 5: Gráfico de dispersão da variável produção de viagens. (a) Valores estimados x observados. (b) valores previstos x resíduos

3.3 APLICAÇÃO DA RLM

A Regressão Linear Múltipla (RLM) é aplicada em uma infinidade de casos, onde se deseja encontrar uma relação entre uma única variável dependente (numérica) e diversas variáveis independentes (numéricas ou *dummy*). Relação esta supostamente linear. A finalidade é encontrar a combinação linear das variáveis independentes que forneça a máxima correlação com a variável dependente.

Optou-se, neste trabalho, pelo modelo de RLM *Stepwise*. O procedimento constrói, iterativamente, uma sequência de modelos de regressão pela adição ou remoção de variáveis em cada etapa, sendo o critério o teste parcial F. A Tabela ilustra uma aplicação de modelo de regressão *Stepwise* para previsão de viagens por domicílio (y) a partir do número de membros trabalhadores (x_1) e automóveis (x_2) (Ortúzar e Willumsen, 2011). Para aplicação da RLM neste trabalho, seguiu-se o método apresentado na Figura 6.

Tabela 4 Exemplo de regressão *Stepwise* (Ortúzar e Willumsen, 2011).

Etapa	Equação	R ²
1	$y = 2.33 x_1$	0.203
2	$y = 1.80 x_1 + 1.31 x_2$	0.325
3	$Y = 0.91 + 1.44 x_1 + 1.07 x_2$	0.384

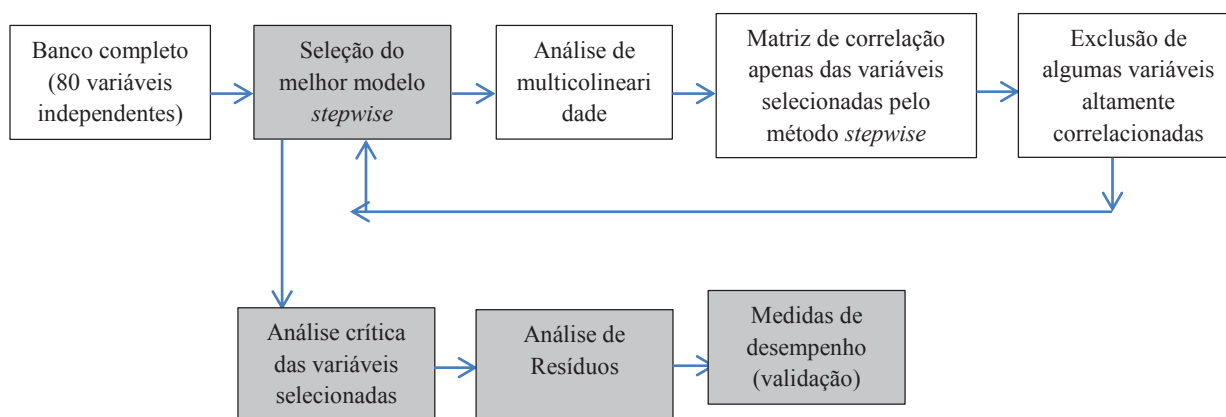


Figura 6 Etapas para aplicação da RLM

Assim, inicialmente foi utilizado o banco de dados original com 80 variáveis independentes normalizadas. Através do método *stepwise*, foi selecionado o melhor modelo (maior coeficiente de determinação), com 20 variáveis independentes selecionadas e 21 parâmetros estimados (incluindo a constante). Em seguida, foi feita uma análise de multicolinearidade deste modelo, detectando-se multicolinearidade.

Numa etapa subsequente, foi calculada uma matriz de correlação das 20 variáveis selecionadas e a variável dependente *Produção de viagens*. Assim, para exclusão do “excesso” de variáveis, o critério de escolha entre duas variáveis independentes altamente correlacionadas foi a maior correlação com a variável dependente.

Finalmente, com um novo conjunto de 14 variáveis, foram calibrados novos modelos pelo método *stepwise*. O modelo escolhido para este trabalho foi aquele descrito na Tabela 5.

Tabela 5 Principais resultados do modelo linear escolhido

MODELO LINEAR: PRODUÇÃO DAS VIAGENS			
Variáveis significativas	R2	Coefficiente	t
Constante	0,855	0,042	9,002
Matrículas Escolares		0,536	18,45
população		0,497	13,84

Em seguida, foi realizada a análise crítica das variáveis independentes significativas selecionadas pelo método estatístico. Observa-se que, tanto as variáveis selecionadas, quanto os sinais e ordem de grandeza dos parâmetros estimados fazem sentido para explicar o fenômeno de produção de viagens. É notória a influência da população na produção de viagens. A variável *Matrículas Escolares* tem um significado similar à variável *População*. Geralmente as matrículas escolares são realizadas em escolas próximas às residências, provavelmente na mesma zona de tráfego da residência. Após análise crítica do modelo obtido, são observados os resíduos do modelo linear. Na Figura 7 pode-se verificar que os resíduos atendem à suposição de normalidade, contudo não atendem à suposição de homocedasticidade. Fazendo testes de normalidade (Kolmogorov-Smirnov e Shapiro-Wilk), observa-se que, nem a variável dependente e nem os resíduos passam em tais testes.

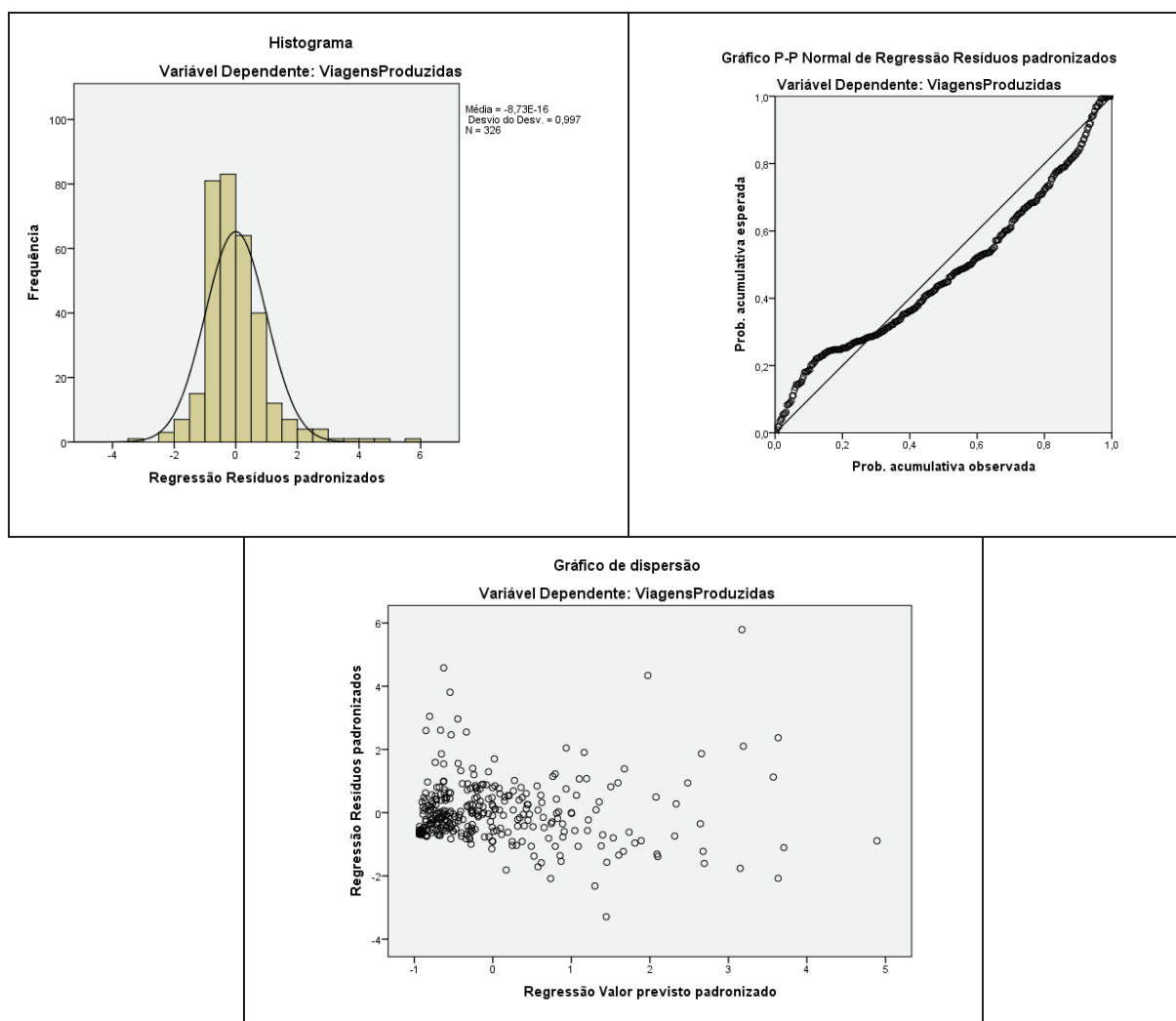


Figura 7 Análise de resíduos (RLM)

Observando as medidas de desempenho, apresentadas na seção anterior, para a RLM, tais medidas são apresentadas na Tabela 6. Verificam-se algumas diferenças sutis entre a análise dos erros de ambas as técnicas. O poder preditivo é similar para ambas as técnicas, alternando maiores ou menores erros a depender da medida de desempenho adotada (Erro Médio – EM; RME - Raiz do Erro Médio; ER- Erro Relativo; r- coeficiente de correlação e DESVPAD – Desvio Padrão do erro). Desta forma, o desempenho das duas técnicas pode ser classificado como muito bom. Além disso, a Figura 8, em seguida, apresenta um diagrama de dispersão entre valores estimados e observados, corroborando a hipótese de alta relação entre os valores. Observa-se também uma relação mais forte para o caso dos resultados da RNA.

Tabela 6 Medidas de desempenho calculadas para RLM

R2: 0,855	PRODUÇÃO	
	70%	30%
ME	0,000	-0,005
RME	0,061	0,045
ER	0,146	0,151
CORREL	0,925	0,928
DESVPAD	0,061	0,044

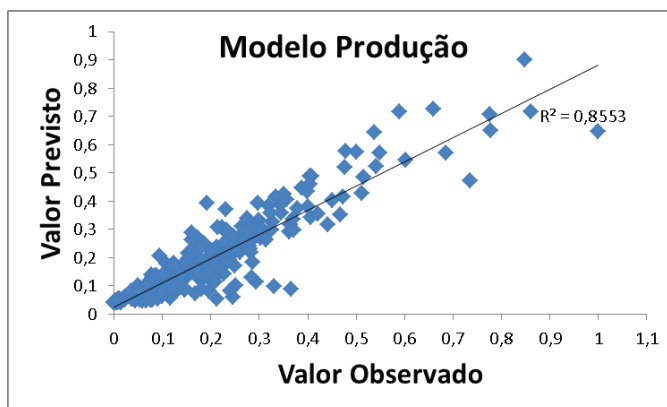


Figura 8 Gráfico de dispersão da variável produção de viagens (RLM).

4. CONCLUSÕES

A proposta deste trabalho apresentou resultados satisfatórios, e evidencia a eficácia da aplicação das RNAs no planejamento de transportes, especificamente na previsão de produção de viagens.

Os resultados encontrados, através da técnica mais usual no planejamento de transportes, a Regressão Linear Múltipla e a RNA foram bem similares e compatíveis. Apesar disso, é necessário considerar que a técnica de RLM possui suposições rígidas, tais como normalidade, homocedasticidade, linearidade, multicolinearidade. Como foi visto neste trabalho, muitas dessas suposições não são atendidas na tentativa de adequar problemas de transportes na aplicação de tal técnica.

Contudo, é importante destacar as vantagens e desvantagens de cada uma das referidas técnicas. A técnica de RNA é um *data mining*, sendo completamente “livre” em termos de suposições e *input*. No entanto, não é um modelo paramétrico como o modelo linear. Através das RNAs não é possível associar a variável dependente a parâmetros estimados das variáveis independentes e nem mensurar se tais variáveis são significativas na explicação do fenômeno. As RNAs, contudo, estabelecem um nível de importância das covariáveis para previsão da variável dependente.

Desta forma, este trabalho corroborou a hipótese de que a RNA é uma técnica adequada para fins de previsão de geração de viagens, minimizando os erros das estimativas, podendo ser utilizada alternativamente às abordagens tradicionais.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio do CNPq.

REFERÊNCIAS BIBLIOGRÁFICAS

- Bocanegra, C. W. (2002) Procedimentos para tornar mais efetivo o uso das redes neurais artificiais em planejamento de transportes. Dissertação (Mestrado). Escola de Engenharia de São Carlos, Universidade de São Paulo. São Carlos.
- Chin, S.; Hwang, H.; Pei, T. (1994) Using neural networks to synthesize origin-destination flows in a traffic circle. *Transportation Research Record*, n. 1457, p. 132-142.
- Colombaroni, C.; Fusco G. (2011) Artificial Neural Network Models for Car Following: Experimental Analysis and Calibration Issues. *Proceedings of 2nd International Conference on Models and Technologies for Intelligent Transportation Systems*, Leuven, Belgium.
- Gonçalves, D.N.S.; M.A. Silva e M. A. d'Agosto (2015) Procedimento para uso de Redes Neurais Artificiais no planejamento estratégico de fluxo de carga no Brasil. *Journal of Transport Literature*, 9(1), 45-49 (2015).
- Haykin, S. (2001) *Redes neurais - princípios e práticas*. 2 ed. . Porto Alegre: Bookman, 2001. 900 p.
- Kikuchi, S.; Nanda, R.; Perinchery, V. (1993) A method to estimate trip o-d patterns using neural networks approach. *Transportation Planning and Technology*, v.17, p. 51-57.
- Lin, H.; Taylor, M.A.P.; Zito R.(2005) A Review of Travel-Time prediction in Transport and Logistics. *Proceedings of the Eastern Asia Society for Transportation Studies*, v. 5, p. 1433 - 1448.
- Lindner, A. (2015) Análise desagregada de dados de demanda por transportes através de modelagem geoestatística e tradicional. 2015. 106p. Dissertação (Mestrado em Ciências) - Departamento de Engenharia de Transportes, Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2015.
- Mark, C.D.; Sadek, A.W.; Rizzo D. (2004) Predicting experienced travel with neural networks: A parametric simulation study. *Intelligent Transportation Systems Conference*, Washington, D.C., p. 906-911, out. 2004.
- Metrô - Companhia do Metropolitano de São Paulo (2007) *Pesquisa Origem Destino 2007 – Região Metropolitana de São Paulo: Síntese das informações*.
- Ortúzar, J.D.; Willumsen, L.G. (2011). *Modelling Transport*. Wiley, 4th Edition.
- Raia Jr, A.A. (2000) *Acessibilidade e Mobilidade na Estimativa de um Índice Potencial de Viagens Utilizando Redes Neurais Artificiais e Sistema de Informações Geográficas*. 2000. 196 p. Tese (Doutorado em Engenharia de Transportes) – Escola de Engenharia de São Carlos, Universidade de São Paulo. São Carlos.
- Rasouli, M. e H. Nikraz (2013) Trip Distribution Modelling Using Neural Network. *Transport Research Forum 2013*, Brisbane, Australia.
- Rodrigues da Silva A.N.; Ramos, R.A.R.; Souza, L.C.L.; Rodrigues, D.S., Mendes, J. F. G.(2008) SIG: Uma plataforma para introdução de técnicas emergentes no planejamento urbano, regional e de transportes: uma ferramenta 3D para análise ambiental urbana, avaliação multicritério, redes neurais artificiais. São Carlos. Ed. EdUFSCar, 2008.
- Silva, C.A.U. (2003) Um método para estimar observáveis GPS usando Redes Neurais Artificiais. 2003. 113p. (Doutorado em Engenharia de Transportes) – Escola de Engenharia de São Carlos, Universidade de São Paulo. São Carlos, 2003.