



## **PADRÕES DE VIAGENS E VARIÁVEIS SOCIOECONÔMICAS, DE USO DO SOLO E PARTICIPAÇÃO EM ATIVIDADES**

**CIRA SOUZA PITOMBO  
EJI KAWAMOTO**

### **INTRODUÇÃO**

Um dos tópicos mais importantes na análise de demanda por transportes é a possível relação entre as necessidades individuais de realização de atividades espacialmente dispersas, a estrutura urbana, as características individuais e domiciliares, o sistema de transporte e as diferenças no comportamento relacionado a viagens, tendo em vista que os indivíduos deslocam-se no meio urbano considerando suas necessidades de realizar atividades fora do domicílio, suas próprias características individuais, atributos domiciliares e características do seu meio.

Tornou-se, então, um dos objetivos da análise de viagens baseada em atividades investigar as variáveis que influenciam a programação da sequência de viagens individuais. Foi sugerido através de diversos trabalhos (BHAT; SINGH, 2000; KWAM, 2000; STRATHMAN; DUEKER, 1995) que indivíduos consideram alguns fatores principais ao realizar os seus respectivos itinerários de viagens: (1) participação em atividades; (2) características socioeconômicas individuais e domiciliares; (3) características das áreas urbanas; (4) atributos do sistema de transporte.

A influência das características socioeconômicas individuais e domiciliares e a participação das pessoas em diversas atividades na formação de padrões de viagens vem sendo estudada nos últimos anos (BHAT;

KOPPELMAN, 1991; LU; PAS, 1999). De fato, na literatura vigente predomina a afirmação de que características dos deslocamentos pessoais podem ser determinadas pelo sexo, posse de automóveis (STRAMBI; VESPUCCI; BILT, 2004), papel do indivíduo no domicílio e alocação de tarefas (SRINIVASAN; ATHURU, 2005) e participação em atividades (classificadas geralmente como subsistência, lazer, saúde, compras etc.).

Contudo, deve-se ressaltar que características socioeconômicas e participação em atividades são apenas partes de um conjunto de variáveis que permite explicar o comportamento concernente à escolha do padrão de viagens encadeadas. Assim como a posse de automóveis influencia a escolha modal, é provável que a disponibilidade de determinados modos de viagem, numa dada zona, exerça influência na formação de padrões de viagem nessa zona. Ainda é possível que o padrão de uso do solo das zonas ajude a explicar as escolhas de destinos.

Aos poucos, novas dimensões vêm sendo incorporadas à análise, em busca da representação mais realística do comportamento referente ao encadeamento de viagens e da construção de estruturas mais adequadas para previsão da demanda por transportes. Isso pode ser exemplificado pela introdução de variáveis que caracterizam a configuração urbana e outras, relacionadas ao sistema de transportes. Dessa forma, surgiram novos trabalhos que consideram, além de características socioeconômicas e participação em atividades, o uso do solo e aspectos ligados ao sistema de transportes.

Kitamura (1985) examinou relações entre a tendência de os indivíduos encadearem as viagens e as características de uma cidade hipotética linearmente disposta. A análise mostrou que a tendência de encadear viagens é função da utilidade de um conjunto de oportunidades e do tipo de atividade a ser realizada. Srinivasan (2000) investigou como as características da vizinhança (uso do solo, rede de transporte e medidas de acessibilidade, quantificadas com auxílio de um Sistema de Informações Geográficas – SIG) afetam o comportamento de viagem em relação à escolha modal e à cadeia de viagens. Dentre os resultados obtidos,

observa-se, por exemplo, que indivíduos residentes em zonas de tráfego mistas (com altas densidades comerciais e residenciais), além de realizarem cadeias de viagens com atividades diferentes de trabalho, fazem esse tipo de viagem a pé.

Assim, com base na literatura, espera-se que a escolha desses padrões faça parte de um processo de tomada de decisão individual, no qual se consideram as atividades a serem realizadas, características individuais e do domicílio, a localização relativa da residência e um conjunto de oportunidades e atributos do sistema de transportes.

Dessa forma, o objetivo deste capítulo é analisar o comportamento referente ao encadeamento de viagens urbanas na Região Metropolitana de São Paulo (RMSP), em consideração a três dos quatro grupos de variáveis mencionados anteriormente: (1) participação em atividades; (2) características socioeconômicas individuais e domiciliares e (3) características do meio urbano (uso do solo). Propõe-se encontrar relações entre padrões de encadeamento de viagens (variável dependente) e variáveis socioeconômicas individuais e domiciliares, participação em atividades, além de uso do solo (variáveis independentes).

Há dois objetivos secundários, que, parte do método adotado, são fundamentais para atingir a finalidade principal: propor um conjunto de variáveis que representem características de uso do solo – vale ressaltar que nesta pesquisa variáveis de uso do solo representam a intensidade e a distribuição geográfica das atividades no meio urbano – e testar a significância do conjunto de variáveis ora proposto.

## MÉTODO

A Figura 1 apresenta uma síntese do método adotado. Cada etapa ilustrada na figura seguinte corresponde às seções subsequentes do presente texto.

FIGURA 1 – SÍNTESE DO MÉTODO ADOTADO



## ESCOLHA DAS TÉCNICAS DE ANÁLISE MULTIVARIADA DE DADOS (AM)

O conjunto de técnicas estatísticas conhecidas como análise multivariada de dados (AM), permite ao pesquisador sintetizar a complexidade dos dados, detectando padrões, relações e outras particularidades e auxiliando o desenvolvimento de hipóteses e modelos adequados.

As técnicas de AM classificam-se em dois grandes grupos: (1) técnicas confirmatórias e (2) técnicas exploratórias. As primeiras podem ser consideradas extensões de técnicas estatísticas tradicionais (regressão linear simples – regressão linear múltipla; Anova – Manova). A rigor, esse conjunto inclui muitas outras técnicas, tais como análise fatorial, regressão logística multinomial etc. São técnicas predominantemente confirmatórias, sendo que a inferência estatística é a sua base.

O segundo grupo citado tem como principal objetivo extrair informação válida, previamente desconhecida, a partir de grandes bases de dados. Técnicas exploratórias permitem que pesquisadores explorem modelos analíticos não necessariamente fundamentados na inferência estatística.

Tais métodos, na maioria das vezes, permitem uma estimação empírica ao invés de suposição estatística. Citam-se, como exemplo de técnicas de AM exploratórias, Redes Neurais Artificiais, Árvores de Decisão e Algoritmos Genéticos.

A aplicação conjunta dos dois tipos de técnicas pode trazer benefícios e resultados mais interessantes: as técnicas exploratórias permitem encontrar relações e modelos inicialmente desconhecidos, ao passo que as confirmatórias podem mensurar o nível de ajuste dos modelos, ou a significância das variáveis envolvidas.

Neste capítulo, foram utilizadas duas técnicas de AM: (1) árvore de decisão (AD) – técnica exploratória: para encontrar relações entre as variáveis dependente e independentes – e (2) regressão linear múltipla (RLM) – técnica confirmatória: para trazer inferência estatística ao estudo.

## ESTUDO DE CASO

O estudo é baseado nos dados da pesquisa origem-destino (O/D) da RMSP, realizada em 1997, por meio de entrevista domiciliar, pela Companhia do Metropolitano de São Paulo (CMSP). Na época, a RMSP contava com uma população de aproximadamente 17 milhões de habitantes, distribuída em 39 municípios.

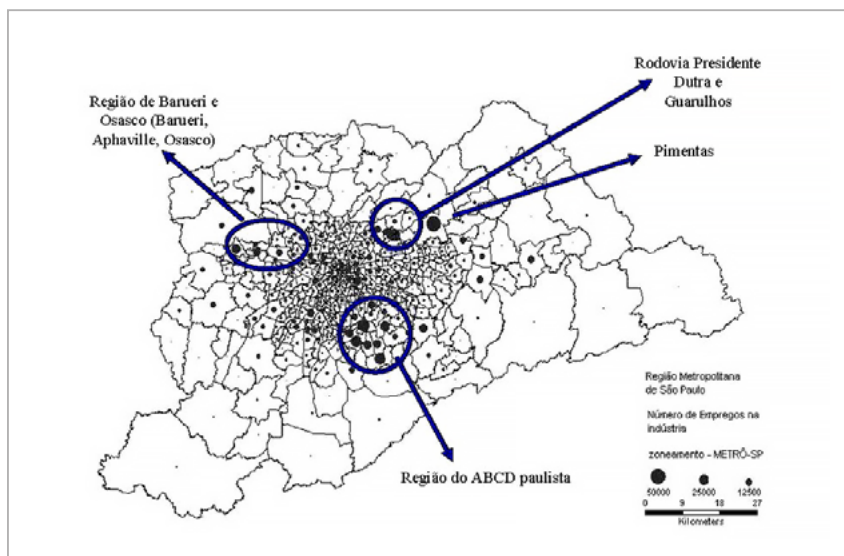
A região foi subdividida em 389 zonas de tráfego e a pesquisa domiciliar obteve dados socioeconômicos e características das viagens realizadas. O banco de dados da RMSP é composto originalmente por 98.780 indivíduos.

A fase de tratamento de dados incluiu as seguintes etapas: (1) eliminação de dados incompletos; (2) eliminação de indivíduos que tenham realizado uma ou mais de quatro viagens, com intuito de limitar a complexidade da análise; (3) eliminação dos indivíduos que não tenham tido como origem e destino final a residência e (4) eliminação dos indivíduos que não viajaram no dia anterior ao da pesquisa.

Partindo da hipótese de que a distribuição de atividades no meio urbano interfere nos deslocamentos individuais, realizou-se uma análise da distribuição espacial de empregos industriais e comerciais na área de estudo.

Observa-se que, no caso de empregos na indústria (Figura 2), há uma concentração maior em algumas regiões: (1) Pimentas, Rodovia Presidente Dutra e Guarulhos; (2) grande ABCD e (3) zonas pertencentes aos municípios de Barueri e Osasco. Apesar de se verificar uma leve concentração nas zonas centrais, pode-se afirmar que há basicamente três polos industriais.

**FIGURA 2 - DISTRIBUIÇÃO DE EMPREGOS NA INDÚSTRIA**



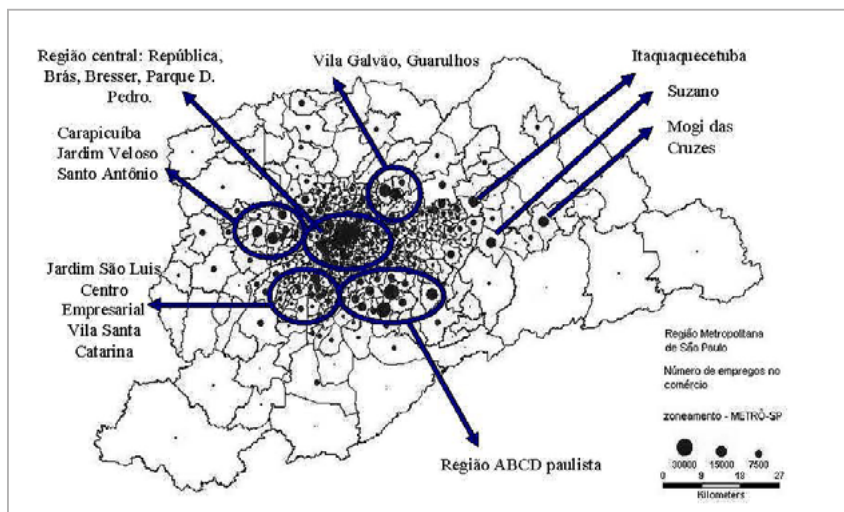
Fonte: CMSP (1988).

No caso de empregos no setor comercial (Figura 3), há forte concentração de tais atividades na região central. Também se observa um grande número de empregos, que se concentra basicamente em seis polos: (1) Itaquaquecetuba, Suzano e Mogi das Cruzes; (2) Guarulhos e Vila Galvão; (3) Região central, República, Brás; (4) Carapicuíba, Jardim Veloso e Santo Antônio e (5) região um pouco a sudoeste do centro, abrangendo principalmente Jardim São Luís, Centro Empresarial e Vila Santa Catarina; (6) ABC paulista. Nota-se que a atividade comercial é relativamente dispersa pela RMSp.

Observando a distribuição geográfica do número de empregos nos dois setores econômicos, verificou-se que empregos no setor industrial concentram-se em alguns polos principais da região de estudo, enquanto

há oferta de empregos no setor comercial distribuída por toda a região. Assim, optou-se por fazer uma análise separada por setor econômico para melhor investigação dos resultados.

**FIGURA 3 – DISTRIBUIÇÃO DE EMPREGOS NO COMÉRCIO**



Fonte: CMSP (1988).

Foram utilizadas duas subamostras, provenientes da amostra original, caracterizadas pelo setor de emprego dos indivíduos: (1) trabalhadores no setor industrial e (2) trabalhadores no setor comercial.

O intuito é evidenciar a possível influência do uso do solo comparando os dois casos específicos. Espera-se que o uso do solo exerça maior influência no caso dos industriários, pois se constata que há uma relativa concentração do número de empregos na indústria em determinadas zonas de tráfego.

## REPRESENTAÇÃO DAS VARIÁVEIS INDEPENDENTES

Conforme o objetivo principal da pesquisa, as variáveis independentes (categóricas ou numéricas) utilizadas no trabalho pertencem a três grupos principais de variáveis: (1) características socioeconômicas; (2) participação em atividades e (3) características de uso do solo.

Nesta etapa, são selecionadas as variáveis que possuem relações com os padrões de viagem. As variáveis pertencentes aos dois primeiros grupos foram escolhidas com base na literatura, bem como na disponibilidade de dados:

(A) Socioeconômicas: renda familiar (R\$); número de automóveis no domicílio; posição do indivíduo no domicílio (1 – chefe; 2 – cônjuge; 3 – filho; 4 – parente; 5 – empregada e 6 – visitante); total de pessoas no domicílio; grau de instrução (1 – não alfabetizado; 2 – pré-escola; 3 – 1º grau incompleto; 4 – 1º grau completo; 5 – 2º grau incompleto; 6 – 2º grau completo; 7 – superior incompleto; 8 – superior completo); sexo; idade; uso de vale-transporte.

(B) Participação em atividades de subsistência: trabalha (1 – assalariados; 2 – autônomos; 3 – não trabalha); estuda (1 – não estuda; 2 – pré-escola; 3 – 1º / 2º / 3º graus; 4 – outros).

As variáveis do terceiro grupo, descritas detalhadamente na próxima seção, foram propostas baseadas nos princípios do modelo de oportunidades intervenientes (SCHNEIDER, 1959).

### **Variáveis de uso do solo**

As variáveis de uso do solo fazem parte do conjunto de fatores que afetam decisões individuais de realização de viagens. A incorporação de tais variáveis ao estudo de viagens encadeadas envolve uma série de dificuldades, tais como representar ou mensurar tais atributos.

Para a elaboração das variáveis de uso do solo, partiu-se do princípio adotado no modelo de oportunidades intervenientes, que admite que, numa área urbana, todas as viagens são tão curtas quanto possível, sendo apenas longas o necessário para atingirem o destino aceitável mais próximo, de forma que o propósito do viajante seja satisfeito. Neste capítulo, as variáveis de uso do solo foram representadas em termos de grau de “oportunidade acumulada” por faixa de distância.

O termo “oportunidade” refere-se, neste estudo, à oferta de empregos por setor econômico. Considerando as duas subamostras adotadas, esse termo representa “parcela (em %) do total de empregos na indústria” (para o

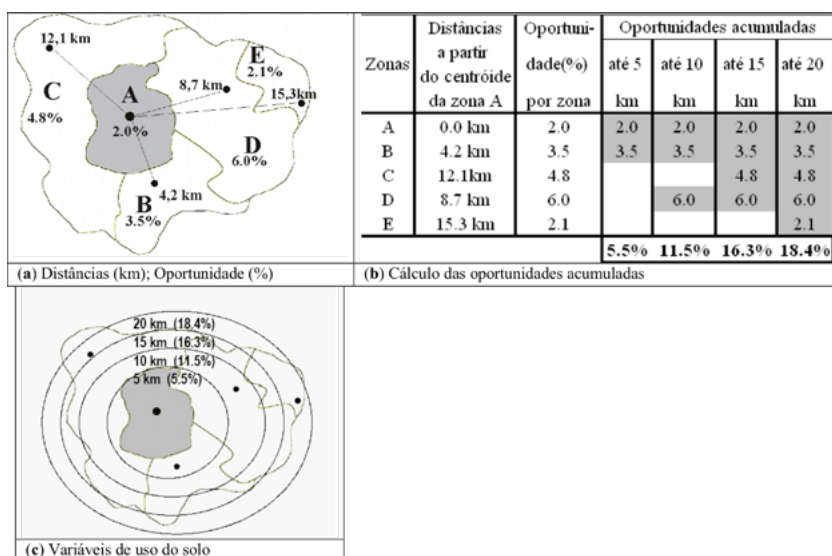


caso de industriários) e “parcela (em %) do total de empregos no comércio” (para o caso de comerciários). Assim, acumularam-se tais parcelas de empregos em faixas de distância, consideradas entre o centroide da zona de residência até centroides das zonas localizadas a 5 km, 10 km, 15 km e 20 km, gerando, dessa forma, o termo “oportunidade acumulada”.

A Figura 4 (a, b e c) exemplifica a proposta das variáveis de uso do solo. A zona de origem é a central “A” (sombreada). Na etapa (a), estão representados os valores das “oportunidades” (em %) em cada zona, assim como distâncias em linha reta a partir do centroide de “A”. Na etapa (b), são mostrados os cálculos das “oportunidades acumuladas” para cada uma das quatro faixas de distância. Finalmente, na etapa (c), são ilustradas as variáveis de uso do solo propostas a partir da zona de tráfego A.

Esse conjunto de variáveis de uso do solo representam basicamente a distribuição e a intensidade de atividades no meio urbano (oportunidade acumulada na indústria até 5 km da zona da residência, oportunidade acumulada na indústria até 10 km da zona da residência e assim sucessivamente).

**FIGURA 4 – PROPOSTA DO CONJUNTO DE VARIÁVEIS DE USO DO SOLO**



## REPRESENTAÇÃO DA VARIÁVEL INDEPENDENTE

Este estudo tem como finalidade representar as sequências de atividades realizadas, segundo os motivos de viagem, modos de transporte e destinos escolhidos pelos indivíduos no período de um dia. Para que os padrões de viagens representem os atributos propostos, foram realizadas as seguintes etapas: (1) representação da sequência de atividades ou motivo de viagem, (2) representação da sequência de modos principais utilizados, (3) representação da sequência de destinos e (4) obtenção dos padrões finais – combinação das etapas anteriores. Os atributos de viagem foram agrupados e representados através de letras e números conforme a Quadro 1. Dessa forma, as categorias da variável dependente foram representadas por códigos alfanuméricos, correspondentes à sequência dos atributos de viagens mencionados.

QUADRO 1

### AGRUPAMENTO DOS ATRIBUTOS DE VIAGEM PARA CODIFICAÇÃO DOS PADRÕES<sup>1</sup>

<b>Motivo de viagem</b>	(H) Residência; (W) Trabalho; (S) Escola; (A) Outras Atividades
<b>Modo de Transporte</b>	(P) Particular; (T) Coletivo; (N) Não motorizado
<b>Destinos</b>	(1) até 5 km; (2) 5 a 10 km; (3) 10 a 15 km; (4) mais de 15 km

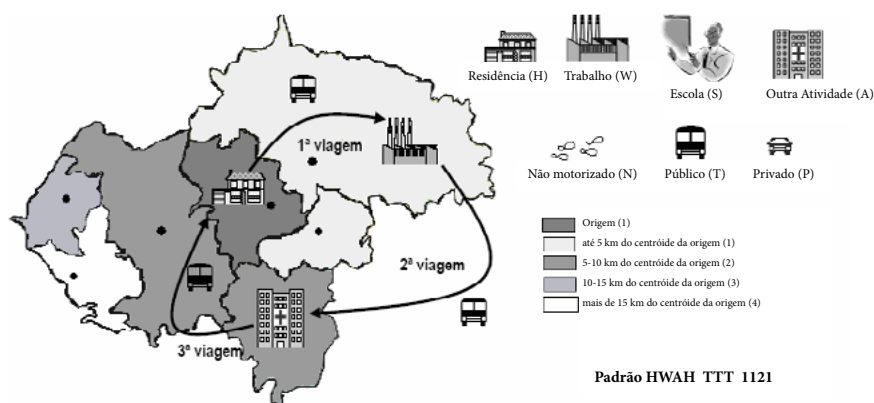
Assim, na primeira etapa, os padrões de viagens, representados por uma sequência de letras (H, W, S, A), indicam as atividades realizadas pelos indivíduos ao longo do dia e a ordem cronológica em que as atividades são desenvolvidas (HSH: casa – escola – casa). Analogamente, na etapa subsequente, as letras P, T e N indicam o modo de transporte principal utilizado nas viagens: NP – não motorizado (1ª viagem) – particular (2ª viagem). Na terceira etapa, os padrões de viagens referentes aos destinos foram representados por números (1, 2, 3 e 4), que expressam as distâncias do centroide da zona de domicílio até o centroide da zona de destino. O primeiro e o último número caracterizam a origem e o destino final (no caso, a residência), sempre representados pelo número 1. Conforme essa

<sup>1</sup> Distâncias dos destinos consideradas até o centroide da zona de tráfego da residência do indivíduo.

especificação, a sequência 1421 representa indivíduos que saíram inicialmente do domicílio (1) e realizaram três viagens: 1º destino – localiza-se a uma distância acima de 15 km do centroide da zona de residência (4); 2º destino – situa-se entre 5 e 10 km do centroide da zona de residência (2) e 3º destino – residência (1).

A partir da codificação atribuída nas etapas iniciais, os padrões finais foram representados por três conjuntos de caracteres: o primeiro, referente à sequência de atividades (HWAH, por exemplo); o segundo, correspondente à sequência dos modos de viagens (TTT, por exemplo), e o terceiro, pertinente à sequência de destinos (1121, por exemplo). Assim, obtém-se o padrão final HWAHTTT1121, ilustrado na Figura 5.

**FIGURA 5 - REPRESENTAÇÃO DO PADRÃO HWAH TTT 1211**



Após a representação das categorias da variável dependente, foram excluídos da contagem aqueles indivíduos que realizavam padrões de viagem menos frequentes em cada uma das amostras devido à limitação de 128 categorias de variáveis dependentes do software utilizado. Por fim, foram obtidas duas subamostras: (A) industriários – 4.102 indivíduos e (B) comerciários – 6.043 indivíduos.

## 7. APLICAÇÃO DA AD E ANÁLISE DE RESULTADOS

Nesta etapa da análise, foram processadas duas árvores de decisão (AD) através do software S-Plus 6.1. Utilizou-se uma variante do algoritmo

da *classification and regression tree* (CART), que estabelece uma relação entre variáveis independentes e variável dependente. O algoritmo foi ajustado mediante sucessivas divisões binárias no conjunto de dados, de modo a tornar os subconjuntos resultantes cada vez mais homogêneos em relação à variável dependente. Essas divisões são representadas por uma estrutura de árvore binária, sendo que cada nó corresponde a uma divisão (BREIMAN *et al.*, 1984).

Assim, partiu-se das amostras de industriários e comerciários, de 128 categorias de variável dependente (padrões de viagens) e das variáveis independentes numéricas e categóricas que fazem parte dos três grupos de variáveis mencionados no estudo. As árvores, ilustradas na Figura 6, foram geradas com desvio mínimo de 0,15, apresentando como resultado final 10 folhas (industriários) e 8 folhas (comerciários).

Para o caso dos industriários, observaram-se as seguintes relações entre as variáveis envolvidas: (1) as características socioeconômicas (número de automóveis no domicílio – AUT; renda familiar – RF; uso de vale-transporte – usa VTRA) influenciam predominantemente a sequência de modos de transportes utilizados; (2) a participação em atividades (estuda) influencia a sequência de atividades; (3) o uso do solo (oportunidade acumulada na indústria até um raio de 5 km da residência – IND 5 km; oportunidade acumulada na indústria até um raio de 10 km da residência – IND 10 km; oportunidade acumulada na indústria até um raio de 15 km da residência – IND 15 km) exercem influência na sequência de destinos.

Para o caso de comerciários, observaram-se as mesmas relações entre características socioeconômicas e escolha da sequência de modos de transporte, no entanto verificou-se a seleção de mais uma variável referente à participação em atividades (trabalha). Essa variável exerce influência na escolha da sequência de motivos de viagem, tendo poder discriminante devido a características da amostra, por haver maior número de trabalhadores autônomos no setor comercial. Além disso, verifica-se a seleção de apenas uma das variáveis de uso do solo na amostra de comerciários (COM 5 km). No caso de trabalhadores no setor comercial, as variáveis de uso do solo não exercem interferência significativa na escolha da sequência de destinos. Esse resultado já era esperado, uma vez que

empregos no setor comercial estão bastante dispersos e não se observam polos de concentração tão demarcados como no caso do setor industrial.

Pode-se citar alguns exemplos de como cada variável se relaciona ao comportamento de viagens em ambas as amostras ou em alguns dos dois casos especificamente.

(1) Indivíduos que residem em domicílios com automóveis possuem maior tendência a utilizar automóveis (PP) para realizar suas viagens.

(2) Indivíduos que possuem vale-transporte (variável que também pode ser caracterizada indiretamente como de custo de transporte) utilizam com maior frequência transporte público (TT). Verificam-se casos específicos, em ambas as amostras, de indivíduos que optam pelo uso de transporte público para realização de viagens longas (141), mesmo quando possuem automóveis no domicílio. Nesses casos (destacados na Figura 6), considerando a distância a ser percorrida (mais que 15 km), o custo da viagem provavelmente será menor com a utilização de transporte público (tarifa de transporte público fixa – especificamente para a RMSP).

(3) Indivíduos com maiores rendas utilizam mais frequentemente automóveis para realização de viagens. Para o caso de industriários, por exemplo, indivíduos com renda familiar superior ou igual a R\$ 2.970 utilizam automóvel (PP) com maior predominância.

(4) Especialmente no caso de industriários, indivíduos que residem em zonas de tráfego com poucas oportunidades de empregos nas suas proximidades ( $IND\ 5\ km < 3.22\%$ , por exemplo) tendem a realizar viagens mais longas ao trabalho (HWH TT 141).

(5) Os comerciários realizam, com maior frequência, viagens curtas (111). Pode-se explicar tal resultado com base na maior dispersão geográfica de atividades comerciais na RMSP. Provavelmente, indivíduos que trabalham no setor do comércio residem próximos aos seus respectivos locais de trabalho.

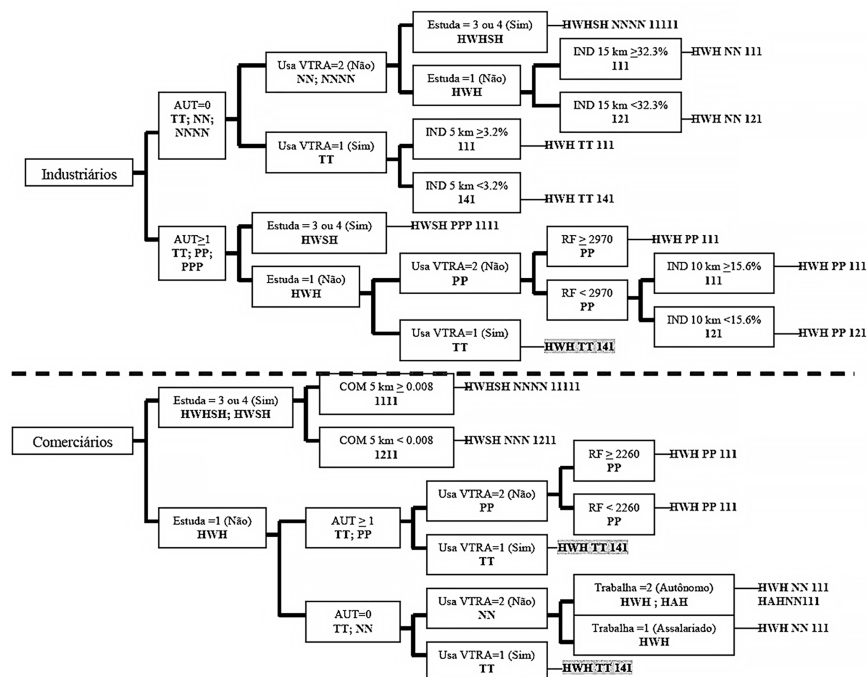
(6) Em ambas as amostras, a variável “estuda” influencia a sequência de atividades. Observa-se que aqueles indivíduos que estudam, além de trabalhar nos setores industrial ou comercial, realizam, predominantemente, sequências de atividades relacionadas ao trabalho e ao estudo.

Nesta etapa do trabalho, foram analisados os resultados obtidos com o processamento das árvores para as duas amostras. Verificou-se a influência

de algumas variáveis independentes nas sequências de viagens realizadas pelos indivíduos (seja na sequência de atividades, na sequência de modos de transporte ou na sequência de destinos). Quanto ao conjunto de variáveis de uso do solo ora proposto, constatou-se a maior influência de tais variáveis na amostra composta por industriários. Esse fato justifica-se pela concentração geográfica de oferta de empregos na indústria na RMSF.

Visando corroborar tais afirmações e buscando associar inferência estatística às variáveis independentes, a próxima etapa do método consiste na aplicação da regressão linear múltipla (RLM) para estimação de parâmetros que comprovem a contribuição das variáveis independentes na realização de determinados padrões de viagens (*teste de significância das variáveis independentes*). Observa-se na Figura 6 as árvores obtidas para cada amostra e apenas o padrão de viagem mais frequente em cada folha. Dentro dos retângulos estão representados os valores das variáveis selecionadas para partição dos dados, bem como a sequência de atributos de viagem influenciada pelas respectivas variáveis. Por exemplo, “AUTO = 0” implica o uso de transporte público ou modo não motorizado – TT ou NN ou NNNN).

FIGURA 6 – AD OBTIDAS PARA INDUSTRIÁRIOS E COMERCIÁRIOS



## TESTE DE SIGNIFICÂNCIA DAS VARIÁVEIS INDEPENDENTES

A aplicação da regressão linear múltipla para mensurar a significância das variáveis independentes, deu-se em conjunto com a AD. Para obtenção dos resultados nesta fase confirmatória, foram realizadas as seguintes etapas: (1) discretização das variáveis independentes; (2) representação das variáveis dependentes; (3) análise de correlação entre variáveis e (4) obtenção de modelos lineares e análise dos resultados.

Na primeira etapa, foi realizada a discretização das variáveis independentes. As variáveis contínuas foram divididas em classes com o intuito de reduzir o efeito da eventual não-linearidade na relação entre variáveis independentes e variável dependente. Cada classe foi associada a uma variável dummy e os valores para escolha das classes das variáveis dummy foram obtidos com a aplicação da AD. Dessa forma, a variável IND 5 km, por exemplo, que antes era uma variável contínua, torna-se uma variável binária, assumindo os seguintes valores: (A) IND 5 km < 3.2% (0); (B) IND 5 km > 3.2% (1). Esse processo foi repetido considerando todas as variáveis independentes selecionadas pela AD, bem como os valores resultantes para partição dos dados.

Para representação das variáveis dependentes, considera-se o seguinte exemplo: (1) observa-se o indivíduo 1, pertencente à amostra de industriários; (2) no processamento da AD, o indivíduo 1 é classificado como pertencente ao nó terminal 8; (3) as pessoas que compõem a folha 8 possuem a probabilidade de 0.22 de realizar o padrão HWHTT121, por exemplo; (4) essa probabilidade de realizar o padrão HWHTT121 (0.22) seria o valor da variável dependente (y) correspondente ao indivíduo 1. No estudo, foi considerada a probabilidade de ocorrência (variando de 0 a 1), para cada indivíduo das amostras, de três padrões de viagem: HWH TT 141, HWH NN 111 e HWH PP 111. Assim, foram obtidos três modelos lineares por amostra (seis no total), nos quais as variáveis dependentes corresponderam às probabilidades de ocorrência por indivíduo de cada um dos três padrões de viagem adotados.

Representadas as variáveis independentes (*dummy*) e as variáveis dependentes (numérica – 0 a 1), fez-se uma análise de correlação entre todas as variáveis envolvidas. Variáveis independentes altamente correlacionadas foram descartadas dos modelos. Assim, se duas variáveis independentes tinham alta correlação, era descartada da análise, entre as duas altamente correlacionadas, aquela que possuía menor correlação com a variável dependente. A Tabela 1 traz cada um dos seis modelos lineares obtidos, valores de R e R<sup>2</sup>, coeficientes associados a cada uma das variáveis, bem como a estatística.

TABELA 1

**RESULTADOS DA APLICAÇÃO DA RLM**

AMOSTRA INDÚSTRIA (4.102 INDIVÍDUOS)									
MODELOS/VARIÁVEIS	CONSTANTE	AUT	VTRA	RENDA	ESTUDA	IND 5KM	IND 10KM	IND 15KM	R <sup>2</sup>
<b>HWHHT141</b>	Coeficiente	0,261	-0,009	-0,109	-0,025	-0,048	-0,101	–	0,725
	stat t	145,850	-9,980	-56,560	-9,090	-16,920	-54,710	–	–
<b>HWHNN111</b>	Coeficiente	0,077	-0,133	0,175	-0,770	-0,360	–	0,026	0,784
	stat t	39,640	-61,780	81,370	-25,330	-11,260	–	12,690	–
<b>HWHPP111</b>	Coeficiente	-0,021	0,102	0,047	0,031	-0,045	–	–	0,716
	stat t	-22,680	101,940	47,300	22,000	-30,627	–	–	20,672
AMOSTRA COMÉRCIO (6.043 INDIVÍDUOS)									
MODELOS/VARIÁVEIS	CONSTANTE	AUT	VTRA	RENDA	ESTUDA	TRABALHA	COM 5KM		R <sup>2</sup>
<b>HWHHT141</b>	Coeficiente	0,199	-0,002	-0,132	-0,019	-0,107	-0,009	0,044	0,783
	stat t	100,240	-28,740	-103,830	-15,290	-74,070	-8,130	23,330	–
<b>HWHNN111</b>	Coeficiente	0,114	-0,087	0,149	-0,026	-0,043	-0,012	-0,026	0,713
	stat t	44,170	-59,560	89,380	-16,520	-22,910	-8,160	-10,740	–
<b>HWHPP111</b>	Coeficiente	-0,033	0,106	0,065	0,034	-0,063	-0,005	0,008	0,746
	stat t	-14,370	74,680	44,330	24,400	-38,080	-3,550	3,740	–

AUT = 0 - (O) e AUT ≥ 1 - (1); VTRA = 1 (usa) - (O) e VTRA = 2 (não usa) - (1); RENDA(IND) < 2970 - (O) e RENDA(IND) ≥ 2970 - (1); RENDA (COM) < 2260 - (O) e RENDA (COM) ≥ 2260 - (1); ESTUDA = 1 (Não) - (O); ESTUDA > 1 (Sim) - (1); TRABALHA = 1 (Assalariado) - (O); TRABALHA = 2 (Autônomo) - (1); IND 5km < 3,2 - (O); IND 5km > 3,2 - (1); IND 10km < 15,6 - (O); IND 10km ≥ 15,6 - (1)

Pode-se afirmar que todos os seis modelos obtidos foram considerados satisfatórios, com valores de R<sup>2</sup> variando de 0,713 (modelo HWHNN111 da amostra de comerciários) a 0,784 (modelo HWHNN111



da amostra de industriários). Por exemplo, 78,4% da variância da variável dependente (probabilidade de cada um dos indivíduos pertencente à amostra de industriários escolher o padrão de viagem HWHNN111) é explicada pelas variáveis independentes incluídas no modelo.

Analisando cada uma das variáveis independentes individualmente, verifica-se que elas foram consideradas significativas, com valores relativamente altos da estatística *t* associados a cada uma delas. Além disso, todos os valores dos coeficientes estimados para cada variável foram coerentes.

## CONCLUSÕES

Conforme os resultados, foi atingido o objetivo principal deste capítulo, que consistia em analisar o comportamento individual subjacente ao encadeamento de viagens sob a perspectiva dos três grupos de variáveis principais. Com utilização conjunta de técnicas de AM, exploratórias e confirmatórias (árvore de decisão + regressão linear múltipla), foi possível encontrar relações entre variáveis envolvidas (AD) e, finalmente, ratificar tais relações e mensurar a significância estatística das variáveis independentes (RLM).

Através dos resultados obtidos, pôde-se analisar a influência dos três grupos de variáveis na sequência de viagens: (1) variáveis socioeconômicas (Renda familiar, Usa Vale Transporte, Número de automóveis) afetam principalmente a sequência de modos de transporte utilizados; (2) participação em atividades (Estuda, Trabalha) interfere na sequência de motivos de viagem, e, finalmente, (3) variáveis de uso do solo (Oportunidade acumulada por faixas de distância a partir do centroide da zona de residência) influenciam a sequência de destinos escolhidos. A influência de tais grupos de variáveis no comportamento relativo a viagens urbanas foi analisada nas últimas décadas na literatura, corroborando os resultados encontrados neste estudo. Relações aqui encontradas pela AD e RLM entre renda, posse de automóveis, uso do solo, atividades e padrões de viagens foram corroboradas pela literatura anterior (AXHAUSEN, 2000; BHAT; SRINIVASAN, 2005; FRANK; PIVO, 2004; PITOMBO *et al.*, 2011).

Uma das principais limitações desta pesquisa está no limite de categorias para variável dependente do pacote estatístico utilizado, Splus 6.1. O software tem um limite inferior a 128 categorias para processamento dos dados. Além disso, a representação conjunta de uma série de fatores de viagem (atividade, modo de transporte, destinos, período do dia) em cadeias de viagens pressupõe uma grande quantidade de combinações, acarretando abundância de categorias, o que torna complexa a análise.

Espera-se que o presente capítulo constitua uma contribuição ao meio acadêmico, tanto em termos de representação da intensidade e distribuição geográfica das atividades no meio urbano (variáveis de uso do solo), quanto em relação à investigação da influência de um conjunto de variáveis nos deslocamentos dos indivíduos.

Além disso, como na RMSP vem ocorrendo um aumento do emprego no setor terciário em detrimento dos empregos no setor industrial, os resultados indicam que devem ocorrer mudanças no padrão de deslocamentos dos indivíduos, haja vista as diferenças de comportamento relativo à escolha de destinos observadas entre trabalhadores dos dois setores devido à distribuição geográfica de tais atividades.

## REFERÊNCIAS

AXHAUSEN, Kay W. Activity-based modeling: research directions and possibilities. In: SIMMONDS, David; BATES, John J. *New look at multi-modal modeling*: report for the Department of Environment, Londres, Cambridge e Oxford. Zürich: ETH Zurich, 2000. p. 1-14. Disponível em: <https://www.research-collection.ethz.ch/handle/20.500.11850/145480> Acesso em: 5 jan. 2007.

BHAT, Chandra R.; SINGH, Sujit K. A comprehensive daily activity-travel generation model system for workers. *Transportation Research Part A: Policy and Practice*, Amsterdam, v. 34, n. 1, p.1-22, 2000.

BHAT, Chandra R.; KOPPELMAN, Frank S. Activity-based travel demand analysis: history, results and future directions. In: ANNUAL MEETING OF TRANSPORTATION RESEARCH BOARD, 79., 2000, Washington. *Compendium of papers* [...]. Washington: Transportation Research Board, 2000. 1 CD-ROM.

BHAT, Chandra R.; SRINIVASAN, Sivaramakrishnan. A multidimensional mixed ordered-response model for analyzing weekend activity participation. *Transportation Research Part B: Methodological*, v. 39, n. 3, p.255-278, 2005.

BREIMAN, Leo *et al.* *Classification and regression trees*. Belmont, California: Wadsworth International Group, 1984.

FRANK, Lawrence D.; PIVO, Gary. Impacts of mixed use and density on utilization of three modes of travel: single-occupation vehicle, transit, and walking. *Transportation Research Record*, Thousand Oaks, v. 1466, p. 44-52, 1994.

KITAMURA, Ryuichi. Trip chaining in a linear city. *Transportation Research Part A: General*, Amsterdam, v.19, n. 2, p. 115-167, 1985.

KWAN, Mei-Po. Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: a methodological exploration with a large data set. *Transportation Research Part C: Emerging Technology*, Amsterdam, v. 8, n. 1-6, p.185-203, 2000.

LU, Xuedong; PAS, Eric I. Socio-demographics, activity participation and travel behavior. *Transportation Research Part A*, Amsterdam, v. 33, v. 1, p. 1-18, 1999.

CMSP (Companhia do Metropolitano de São Paulo). *Pesquisa origem-destino 1987: Região Metropolitana de São Paulo: síntese das informações*. São Paulo: CMSP, 1988.

SCHNEIDER, Morton. Gravity models and trip distribution theory. *Papers of Regional Science Association*, v.5, p.51-56, 1959.

SRINIVASAN, Karthik K.; ATHURU, Sudhakar R. Analysis of within-household effects and between-household differences in maintenance activity allocation. *Transportation*, New York City, v. 32, p. 495-521, 2005.

SRINIVASAN, Sumeeta. *Linking Land Use and Transportation: Measuring the Impact of Neighborhood-scale Spatial Patterns on Travel Behavior*. 2000. Thesis (PhD in Urban Studies and Planning) – Massachusetts Institute of Technology, Cambridge, 2000.

STRAMBI, Orlando; VESPUCCI, Katia M.; BILT, Karin-Anne van de. Analysis of the evolution of classes of individual patterns and their relation to socio-demographic and economic variables. In: ANNUAL MEETING OF TRANSPORTATION RESEARCH BOARD, 83., 2004, Washington. *Compendium of Papers [...]*. Washington: Transportation Research Board, 2004. 1 CD-ROM.

STRATHMAN, James G.; DUEKER, Kenneth J. Understanding trip chaining. In: FEDERAL HIGH ADMINISTRATION (ed.). *Special reports on trip and vehicle attributes: 1990 NPTS*. Lanham: Federal High Administration, 1995. p. 1-28.

PITOMBO, Cira S.; KAWAMOTO, Eiji; SOUSA, António J. An exploratory analysis of relationships between socioeconomic, land use, activity participation variables and travel patterns. *Transport Policy*, Oxford, v. 18, n. 2, p. 347-357, 2011.