

**ERICK SOBREIRO GONÇALVES**

**Análise de padrões operacionais da frota de ônibus de transporte público no município de São Paulo e a influência de fatores climáticos em sua dinâmica**

**São Paulo**

**2018**

**ERICK SOBREIRO GONÇALVES**

**Análise de padrões operacionais da frota de ônibus de transporte público no município de São Paulo e a influência de fatores climáticos em sua dinâmica**

Dissertação apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do título de Mestre em Ciências.

**São Paulo**  
**2018**

**Análise de padrões operacionais da frota de ônibus de transporte público no município de São Paulo e a influência de fatores climáticos em sua dinâmica**

Dissertação apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do título de Mestre em Ciências.

Área de concentração:  
Engenharia de Transportes

Orientador:  
Prof. Dr. Claudio Luiz Marte

**São Paulo**  
**2018**

Gonçalves, Erick Sobreiro

Análise de padrões operacionais da frota de ônibus de transporte público no município de São Paulo e a influência de fatores climáticos em sua dinâmica / E. S. G. -- versão corr. -- São Paulo, 2018.

107 p.

Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Transportes.

1. Transporte Público 2. Variabilidade 3. Mineração de Dados 4. CHAID 5. Krigagem I. Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Transportes II. t.

Este exemplar foi revisado e alterado em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, 02 de Fevereiro de 2018

Autor:

Orientador:

*A minha família, sempre compreensiva, sempre presente.*

## AGRADECIMENTOS

*Agradeço ao meu pai, que me apresentou a então “informática” através da aquisição de um IBM PS/2 em 1993. A minha mãe, sempre incentivando meu desenvolvimento intelectual através do constantemente oferecimento de livros, responsáveis por aguçar minha curiosidade.*

*A minha família, esposa e filho, que abriram carinhosamente mão do meu convívio em função deste objetivo.*

*Ao meu orientador, Professor Claudio Marte, que além de permitir uma escolha de um tema difícil, muito contribuiu neste projeto.*

*Aos professores e amigos, como Professor Leopoldo, Renato Duarte, entre outros que incentivaram meu ingresso nesta empreitada.*

*Aos professores e colegas da Escola Politécnica, que graças ao conhecimento compartilhado, são em boa medida corresponsáveis por esse trabalho.*

*A São Paulo Transportes e o Sr. Maurício Lima, que gentilmente cederam os dados utilizados nesse trabalho.*

*A tantos que muito me inspiraram....*

*“If at first the idea is not absurd, then there is no hope for it.”*

*“Se, a princípio, a ideia não é absurda, então não há esperança para ela.”*

*Albert Einstein*



## RESUMO

Atualmente existem muitas fontes de dados oriundos de movimentação de pessoas e veículos. Deduzir o comportamento do tráfego a partir desses dados pode ser de grande utilidade para a gestão do tráfego.

Entretanto, os métodos existentes podem não abordar todas as dimensões referentes ao comportamento do tráfego, em especial sob efeito de adversidades, como por exemplo - eventos climáticos.

O objetivo dessa dissertação é apresentar uma combinação de métodos estatísticos e de análise espacial, para uso na compreensão dos padrões temporais e espaciais de variabilidade das velocidades, em condições normais e sob o impacto da precipitação.

O método consiste na mineração de dados de micro trechos de deslocamento (microviagens), através da combinação de estatísticas descritivas, árvores de regressão CHAID (Chi square automatic interaction detection) e análises espaciais (Krigagem).

Para a validação do método foram utilizados dados obtidos, durante março de 2014, de operação da frota de ônibus municipal da cidade de São Paulo, fornecidos pela São Paulo Transportes (SPTrans). Os dados de precipitação de chuva foram fornecidos pelo Instituto Nacional de Meteorologia (INMET).

Os resultados demonstraram o impacto da precipitação sobre a variabilidade das velocidades por dia da semana e hora do dia. Adicionalmente, com base nos conjuntos de dados classificados via CHAID, foram produzidas representações da distribuição das velocidades sobre a superfície através uso da Krigagem.

Observou-se uma consistente queda das velocidades no período vespertino, principalmente entre 16:00 e 19:00 horas, frequentemente agravados pela presença de precipitação. Foram selecionados casos de alta variabilidade, a partir da técnica CHAID, que foram analisados espacialmente. Os efeitos da queda da variabilidade foram observados geograficamente, demonstrando que a infraestrutura de transporte é um importante fator em sua caracterização.

O estudo demonstra a viabilidade desta abordagem e a importância de incorporar a geografia no entendimento da variabilidade das velocidades em geral.

Palavras-Chave: Variabilidade. Velocidade. Ônibus. Precipitação. GPS. Transporte público. São Paulo. CHAID. Krigagem. SIG. Mineração de dados.

## ABSTRACT

Nowadays there is a wide number of movement data sources which can be useful in transportation management. However, the existing methods may not cover all traffic dimensions, especially under inclement weather.

The objective of this work is presenting a combination of statistical and spatial methods which allows to understand space and time speed patterns, considering regular operation conditions in comparison with rainfall events.

The method consists in sequential bus tracking data segments (micro travel) mining by the combination of descriptive statistics, CHAID (Chi square automatic interaction detection) regression trees and spatial analysis (Kriging).

To validate the method, São Paulo bus fleet tracking data provided by SPTrans, and rainfall precipitation, provided by INMET was combined and processed.

The results presented rainfall as a relevant factor of speed variability by day of week and time of the day. A consistent speed decrease in the afternoon period, mainly between 4:00 PM and 7:00 PM, was observed. Such effects are more intense in face of rainfall precipitation.

The highest variability cases obtained by CHAID classification was selected and spatially analyzed showing the transportation infrastructure is an important variability component.

Such results show the importance of a geographic approach over speed variability.

Keywords: Bus operation. Rainfall. GPS. Public transportation. São Paulo. CHAID. Average speed. Kriging. GIS. Data mining.

## SUMÁRIO

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>INTRODUÇÃO .....</b>  | <b>4</b>  |
| 1.1      | ESTRUTURA DESTE TRABALHO.....  | 5         |
| 1.2      | OBJETIVOS E JUSTIFICATIVA .....  | 6         |
| <b>2</b> | <b>REVISÃO BIBLIOGRÁFICA .....</b>   | <b>8</b>  |
| 2.1      | VARIABILIDADE.....   | 8         |
| 2.2      | IMPACTO DO CLIMA NOS SISTEMAS DE TRANSPORTE .....                                    | 9         |
| 2.2.1    | <i>Usuários e comportamento.....</i>   | <i>10</i> |
| 2.2.2    | <i>Diferentes abordagens para a análise da influência climática .....</i>            | <i>10</i> |
| 2.3      | MÉTODOS DE ANÁLISE DE DADOS.....   | 11        |
| 2.3.1    | <i>Classification and Regression Trees - C&amp;RT .....</i>                          | <i>11</i> |
|          | <i>(BREIMAN et al., 1984)(BREIMAN et al., 1984)(BREIMAN et al., 1984) .....</i>      | <i>11</i> |
| 2.3.2    | <i>CHAID – Chi-square Automatic Interaction Detector .....</i>                       | <i>13</i> |
| 2.4      | MÉTODOS DE ANÁLISES ESPACIAIS .....  | 14        |
| 2.4.1    | <i>CHAID e Análise Geográfica .....</i>  | <i>14</i> |
| 2.4.2    | <i>Métodos de regressão geográfica .....</i>   | <i>15</i> |
| 2.4.3    | <i>Uso da interpolação de dados em análises de transporte .....</i>                  | <i>16</i> |
| 2.4.4    | <i>Comparação entre métodos de interpolação.....</i>                                 | <i>16</i> |
| 2.4.5    | <i>Krigagem Baesiana Empírica .....</i>  | <i>18</i> |
| 2.4.6    | <i>Considerações Finais do Capítulo.....</i>   | <i>19</i> |
| <b>3</b> | <b>MATERIAIS E MÉTODOS.....</b>  | <b>21</b> |
| 3.1      | CARACTERIZAÇÃO DOS ARQUIVOS COM INFORMAÇÕES DOS ÔNIBUS FORNECIDOS PELA SPTRANS ..... | 21        |
| 3.2      | MODELAGEM DE DADOS.....  | 22        |
| 3.2.1    | <i>Dados alfanuméricos.....</i>  | <i>22</i> |
| 3.2.2    | <i>Dados pluviométricos.....</i>   | <i>23</i> |
| 3.2.3    | <i>Combinação entre dados de rastreamento e dados pluviométricos.....</i>            | <i>25</i> |
| 3.3      | PROCESSAMENTO GEOGRÁFICO DE DADOS .....  | 26        |
| 3.3.1    | CÁLCULO DE GRANDEZAS A PARTIR DOS DADOS DE TELEMETRIA.....                           | 26        |

|          |   |           |
|----------|---|-----------|
| 3.3.2    | DEFINIÇÃO DO MÉTODO DE ANÁLISE DA VARIABILIDADE.....  | 28        |
| 3.3.3    | SELEÇÃO DA ÁREA DE ESTUDO .....   | 29        |
| 3.4      | FLUXO DE MODELAGEM E PROCESSAMENTO DE DADOS.....  | 30        |
| <b>4</b> | <b>PÓS PROCESSAMENTO E RESULTADOS .....</b>   | <b>32</b> |
| 4.1      | CARACTERIZAÇÃO GERAL DOS DADOS .....  | 33        |
| 4.1.1    | <i>Caracterização das velocidades das microviagens .....</i>  | 33        |
| 4.1.2    | <i>Caracterização do intervalo de aquisição de dados de telemetria .....</i>  | 36        |
| 4.1.3    | <i>Caracterização da distribuição das microviagens em dias da semana.....</i>   | 38        |
| 4.1.4    | <i>Caracterização da distribuição das microviagens por hora do dia .....</i>  | 38        |
| 4.1.5    | <i>Caracterização da distribuição pluviométrica.....</i>  | 40        |
| 4.2      | ESTATÍSTICAS DESCRITIVAS .....  | 42        |
| 4.3      | ANÁLISE DOS PADRÕES DAS MICROVIAGENS ATRAVÉS DO MÉTODO CHAID.....   | 49        |
| 4.3.1    | <i>Caracterização dos padrões de velocidade através da técnica CHAID.....</i>   | 51        |
| 4.3.2    | <i>Caracterização dos padrões de velocidade em situação de precipitação através da técnica CHAID</i>                              | 54        |
| 4.3.3    | <i>CHAID – Comparação de resultados: Precipitação = 0 versus Precipitação &gt;0.....</i>  | 57        |
| 4.3.4    | <i>CHAID – Resultado do processamento conjunto de dados com e sem ocorrência de precipitação.....</i>                             | 59        |
| 4.3.5    | <i>CHAID – Breve visão sobre a variância dos preditores .....</i>   | 63        |
| 4.3.6    | <i>CHAID - Considerações gerais:.....</i>   | 65        |
| 4.4      | ANÁLISES ESPACIAIS .....  | 66        |
| 4.4.1    | <i>Contextualização dos resultados.....</i>   | 66        |
| 4.4.2    | <i>Cálculo do desvio padrão das superfícies.....</i>  | 69        |
|          | <i>As superfícies resultantes do processo de Krigagem foram submetidas a um processamento para cálculo do desvio padrão. ....</i> | 69        |
| 4.4.3    | <i>Distribuição das microviagens sobre a área de estudo – Visão geral .....</i>   | 69        |
| 4.4.4    | <i>Análise de casos selecionados com base na classificação CHAID e estatísticas descritivas:</i>                                  | 72        |
| 4.4.5    | <i>Comparação geral: 16:00hs, 17:00hs e 18:00hs em março de 2014 na ausência e presença de chuva:</i>                             | 73        |

|           |   |           |
|-----------|---|-----------|
| 4.4.6     | <i>Casos extremos CHAID:</i> .....                | 76        |
| 4.4.7     | <i>Visão Geral: média dos desvios CHAID</i> ..... | 78        |
| <b>5</b>  | <b>79</b>   |           |
| <b>5.</b> | <b>DISCUSSÃO</b> .....                            | <b>80</b> |
| 5.1       | INDICADORES DE VARIABILIDADE .....                | 80        |
| 5.1.1     | <i>Médias</i> .....                               | 81        |
| 5.1.2     | <i>Precipitação e médias</i> .....                | 82        |
| 5.1.3     | <i>Desvio Padrão</i> .....                        | 82        |
| 5.1.4     | <i>Desvios e precipitação</i> .....               | 84        |
| 5.1.5     | <i>Coefficiente de variação</i> .....             | 85        |
| 5.2       | CHAID – CLASSIFICAÇÃO DE CASOS AFINS .....        | 86        |
| 5.3       | DETERMINISMO GEOGRÁFICO .....                     | 89        |
| 5.4       | CONSIDERAÇÕES FINAIS .....                        | 91        |
| <b>6</b>  | <b>REFERÊNCIAS</b> .....                          | <b>94</b> |

# 1 INTRODUÇÃO

A eficiência no fornecimento de serviços nos sistemas de transporte público é um tema em constante revisão.

Um amplo debate vem ocorrendo acerca de soluções que tragam benefícios, tanto aos usuários de transporte público, quanto aos usuários da infraestrutura de transporte em geral.

Nesse contexto, inúmeras soluções vêm sendo desenvolvidas ou sendo propostas, como, por exemplo: aplicativos, soluções de inteligência preditiva, veículos autônomos, entre outros. Para que essa diversidade de soluções prospere, é preciso compreender profundamente os fatores envolvidos no contexto urbano, onde se dá a operação de transporte público. Técnicas computacionais, como a inteligência artificial, mostram-se fundamentais no tratamento de quantidades maciças de dados. Essas técnicas consideram modelos cada vez mais complexos, e simulações mais próximas da realidade, que abordam questões típicas de deslocamento, de fatores comportamentais, e de fatores externos. Mais recentemente, sistemas baseados em internet das coisas e gerenciados por artefatos de software, aptos em prover análises decisórias, vêm, mesmo de forma inicial, se mostrando como um possível caminho para um melhor aproveitamento da infraestrutura rodoviária existente.

Estudos evidenciam, conforme será apresentado à frente, a importância de inúmeros fatores normalmente ignorados na construção de soluções de transporte para cidades.

Faz-se necessária uma abordagem interdisciplinar, que contemple psicologia - no âmbito das escolhas dos usuários, história e geografia econômica - na dinâmica das atividades microrregionais, análise do clima e relevo urbano e sua disposição na superfície municipal e ainda fatores computacionais - como a complexidade inerente à manipulação de dados e informações de origem espaço-temporal em tempo real.

A aplicação de técnicas capazes de estabelecer relações entre inúmeras fontes de dados é um importante passo no amadurecimento de sistemas de inteligência e previsão do uso do sistema viário urbano. Tais ferramentas podem vir a se constituir,

juntamente com sistemas inteligentes de tráfego, em um padrão de práticas e soluções que se estabeleçam como um padrão operacional nas “cidades inteligentes”.

O estabelecimento de um método ou um conjunto de técnicas capazes de identificar padrões em dados de transporte, frente a um grande volume de dados em dimensões temporais e espaciais, é o alicerce para importantes iniciativas no âmbito da gestão operacional do tráfego, tais como simulações em tempo real.

O uso de simulações em tempo real na gestão do trânsito permitirá intervenções em campo que certamente contribuirão em uma operação síncrona entre os diferentes modos de transporte, garantindo um patamar superior na qualidade nos serviços junto a seus usuários.

## **1.1 ESTRUTURA DESTE TRABALHO**

Este trabalho está organizado de forma a fornecer a compreensão sobre os inúmeros métodos aqui aplicados sequencialmente.

O capítulo 2 (Revisão bibliográfica) apresenta a base teórica pregressa para o desenvolvimento deste trabalho. São apresentados os conceitos de variabilidade em transporte e suas diferentes abordagens, bem como os métodos estatísticos e de análise geográfica utilizados na sua identificação.

No capítulo 3 (Materiais e métodos) são detalhados os processos de modelagem e validação dos dados, empreendidos a partir do material fornecido pela São Paulo Transportes – SPTrans, referente aos dispositivos de rastreamento embarcados em seus ônibus. São apresentados também os dados de precipitação de chuva fornecidos pelo INMET – Instituto Nacional de Meteorologia.

Os resultados (capítulo 4) estão dispostos dentro uma proposta lógica que busca adensar as respostas para os fenômenos observados, partindo de métodos estatísticos clássicos (descritivos), específicos (CHAID) e tendo como última etapa a análise da disposição destes fenômenos na dimensão geográfica-espacial.

A discussão dos resultados (capítulo 5) procura abordar as observações obtidas através dos diferentes métodos, relacionando-as com os objetivos propostos e a literatura existente e confirmando ou não observações de outros autores.

## 1.2 OBJETIVOS E JUSTIFICATIVA

O presente trabalho tem por objetivo propor um método para a caracterização dos padrões operacionais dos ônibus de transporte da cidade de São Paulo.

Complementarmente, será testada a hipótese de senso comum, a qual credita-se à precipitação de chuva a degradação nas condições de transporte público no âmbito da mobilidade.

Assim, foram definidos como objetivos complementares:

- a) Caracterização dos padrões de operação através das velocidades médias em dias úteis – variações diárias;
- b) Caracterização dos padrões de operação através de suas velocidades médias das 7:00 as 21:00 horas – variações hora a hora;
- c) Variabilidade das velocidades médias frente a eventos de precipitação de chuva – comparação dos padrões na ausência e em diferentes escalas de precipitação, dentro do apresentado no período de estudo;
- d) Aplicação de um método analítico-estatístico que permita observar as variações diante de um grande volume de dados, estabelecendo relações de causa e efeito;
- e) Aplicação de um método analítico que permita identificar geograficamente os fenômenos observados nas análises estatísticas.

Uma vez comprovada a sua eficácia, os métodos propostos podem ser empregados em grande escala, para dados de ônibus e outros veículos, produzindo informações úteis na tomada de decisões operacionais, inerentes à gestão do transporte público em São Paulo. Espera-se contribuir efetivamente na identificação de padrões com impacto direto na queda da qualidade dos serviços de transporte



público, simplificando a execução ações preventivas e corretivas, de forma a beneficiar os usuários do sistema.

## 2 REVISÃO BIBLIOGRÁFICA

A caracterização geral dos padrões de deslocamento de veículos de transporte e possíveis variações, causadas sob a influência de fatores adversos, é um tema explorado por diversos autores. A maioria das abordagens refere-se à investigação das variações nos tempos de viagem individual, bem como medidas de variabilidade em distâncias e velocidades percorridas para intervalos regulares de tempo.

As variações gerais no transporte são, principalmente, resultado de congestionamentos recorrentes, cujas causas podem ser atribuídas a diferentes fatores. Do ponto de vista individual, a percepção de fluidez está fortemente associada à variabilidade no tempo gasto para percorrer uma determinada distância.

### 2.1 VARIABILIDADE

A variabilidade é o resultado da ação de diversos fatores sobre o fluxo de tráfego, como por exemplo: diversidade climática, veículos com diferentes portes e performance, não padronização de redes viárias, atrasos produzidos por intersecções, falhas nos sistemas de trânsito, falhas ou acidentes com veículos e pedestres, entre outros fatores (LIU; RECKER; CHEN, 2004).

Pesquisadores têm investigado a variabilidade dos tempos de viagem para rodovias e vias urbanas usando instrumentos de telemetria, tais como: detectores de massa metálica, identificadores por rádio frequência, dispositivos *Bluetooth*, rastreadores GPS (*Global Positioning System*), aplicativos para telefones, dados de pedágios, reconhecimento de caracteres de placas, entre outras diferentes tecnologias capazes de rastrear veículos nas redes viárias (KAMGA; YAZICI, 2014).

A variabilidade nos tempos de viagem é considerada por diversos autores como uma medida para o nível dos serviços de transporte (CHEN; SKABARDONIS; VARAIYA, 2003) e para análises de custo benefício nos investimentos em transporte regional (TAYLOR, 2009) (LYMAN; BERTINI, 2008)(WAARA et al., 2015).

O artigo *Temporal and Weather Related Variation Patterns of Urban Travel time: Considerations and Caveats for Value of Travel Time, Value of Variability, and Mode Choice Studies*, (KAMGA; YAZICI, 2014) analisa o impacto do clima na variabilidade dos tempos de viagem nos táxis da cidade de Nova Iorque.

Este artigo relata a revisão bibliográfica referente ao comportamento do trânsito em vias urbanas e rodovias, através do uso de ferramentas estatísticas, como o desvio padrão, covariância, probabilidade, entre outras.

Em vias urbanas, identificou-se uma relação direta entre o clima adverso e o aumento da variabilidade nos tempos de viagem (VAN LINT; VAN ZUYLEN; TU, 2008) (TU; VAN LINT; VAN ZUYLEN, 2007) (CHIEN; KOLLURI, 2012), que por sua vez se dá, de forma claramente segmentada, conforme a hora do dia. Constatou-se que as médias mais elevadas se dão no fim dos horários de pico (FRANKLIN; KARLSTROM, 2009). Os tempos de viagem sofrem um aumento no intervalo de confiança durante horários de pico e eventos climáticos extremos. Vias urbanas e rodovias apresentam características estruturais diferentes, no entanto, conclui-se que fatores climáticos como precipitação, neve ou mesmo neblina afetam a variabilidade dos tempos de viagem nos casos estudados (YAZICI; KAMGA; SINGHAL, 2013).

## **2.2 IMPACTO DO CLIMA NOS SISTEMAS DE TRANSPORTE**

Dentre os fatores considerados para a variabilidade nas condições de transporte, o impacto do clima é um tema amplamente investigado por diversos pesquisadores. Trata-se um tema extenso, abrangendo escolhas individuais dos usuários em relação aos modos de transporte, mudanças de condução e segurança dos condutores, eventos extremos como alagamentos e deslizamentos, acidentes, entre outros fatores.

### **2.2.1 Usuários e comportamento**

Pesquisas sobre o efeito clima no trânsito vêm provocando resultados contraditórios. Alguns estudos indicam que o clima age como um impeditivo no uso do transporte público, enquanto outros sugerem o oposto (KHATTAK; DE PALMA, 1997a) (CHANGNON, 1996) (GUO; WILSON; RAHBEE, 2007) (STOVER; MCCORMACK, 2012) (KASHFI; BUNKER; YIGITCANLAR, 2016).

Diversos estudos vêm investigando o comportamento da demanda de usuários em ônibus de transporte público durante eventos de precipitação de chuva. Stover e McCormack (2012) analisaram o tema considerando características climáticas do hemisfério norte, concluindo que a precipitação de chuva é a única variável significativa no efeito da diminuição da demanda de passageiros nas quatro estações do ano. Os autores concluíram que houve uma redução na demanda pelos usuários da ordem de 5,05% no inverno, de 9,73% na primavera, de 7,36% no verão e 5,97% no outono. Esses resultados são semelhantes aos encontrados por outros autores (KASHFI; BUNKER; YIGITCANLAR, 2016).

Não foram encontrados artigos retratando a realidade brasileira.

### **2.2.2 Diferentes abordagens para a análise da influência climática**

O clima adverso pode reduzir significativamente a velocidade de operação e a capacidade de tráfego de veículos sobre a malha viária.

Em relação a medições em tempos totais de viagem, autores reportam aumento de 30% em condições de neve e 15% em condições de chuva (LAMM; CHOUERI; MAILAENDER, 1990).

Um estudo promovido pela Administração Federal de Rodovias dos Estados Unidos (FHWA) confirmou a redução das velocidades durante condições climáticas inclementes (HRANAC et al., 2006) em rodovias. Em condições de chuva leve foram identificados cerca de 1,9 km/h de redução nas velocidades de operação, durante

períodos de trânsito livre, e valores entre 6,4 km/h e 12,9 km/h foram observados quando da ocupação máxima da via.

Apesar da redução das velocidades de operação não ser extrema, a capacidade geral das vias costuma ser comprometida da ordem de 3% a 5%, quando um evento de precipitação de chuva está em andamento (SMITH et al., 2004).

Akin; Sisiopiku e Skabardonis (2011) concluíram, após análise de contadores de tráfego da cidade de Istambul, que o clima inclemente teve impacto nas velocidades e nas taxas de fluxo nas duas seções de vias estudadas. A chuva reduziu a velocidade média entre 8% e 12%.

## 2.3 MÉTODOS DE ANÁLISE DE DADOS

### 2.3.1 Classification and Regression Trees - C&RT

(BREIMAN et al., 1984)(BREIMAN et al., 1984)(BREIMAN et al., 1984)

O artigo *Temporal and Weather Related Variation Patterns of Urban Travel time: Considerations and Caveats for Value of Travel Time, Value of Variability, and Mode Choice Studies* (KAMGA; YAZICI, 2014), apresenta um método para a análise de milhares de rotas origem-destino dos táxis na cidade de Nova Iorque. Os autores depararam-se com os desafios inerentes à análise e validação de milhões de dados de GPS. Foram identificados e removidos registros com velocidades e datas-horas inconsistentes.

Foi criado um banco de dados com informações climáticas obtidas através do sitio *Weather Wunderground*. Foram criadas categorias escalares de eventos climáticos, como neve, precipitação de chuva e tempo bom foi construído através de fontes públicas de dados. As condições climáticas foram classificadas baseadas na intensidade do fenômeno conforme classificação do sitio, de onde os dados foram obtidos.

Para a análise do conjunto de dados, os autores fizeram uso do algoritmo *Classification and Regression Trees - C&RT* (BREIMAN et al., 1984).

Trata-se de um modelo não paramétrico utilizado em uma grande variedade de estudos estatísticos, mineração de dados, inteligência artificial, entre outros, que usa entradas de dados categóricas e contínuas para prever uma variável dominante.

Em termos práticos, o método constrói partições de dados homogêneos, segmentando-os em árvores binárias, cujos “nós” representam um evento específico.

Há importantes características analíticas que o tornam bastante adequado ao processamento de grandes volumes tais como:

- Natureza não paramétrica - não requer nenhuma relação de função ou qualquer propriedade de distribuição das variáveis independentes;
- Possui robustez na manipulação e identificação de dados fora das médias e padrões;
- Capaz de lidar com variáveis categóricas em muitos níveis de análise, como por exemplo identificar e agrupar dados nas 24 horas de um dia, nos 7 dias da semana, nas mais diversas condições climáticas, categorizando os tempos de viagem.

Como resultado, a aplicação do algoritmo foi além da classificação das relações diretas entre tempos de viagem e eventos climáticos. A caracterização temporal dos horários de pico demonstrou que estes não são necessariamente os mais congestionados. Os registros próximos ao meio dia possuíam maior lentidão, quando comparados a faixas de horário de pico.

Condições climáticas tidas como inclementes, em geral aumentam o tempo de viagem conforme esperado, mas surpreendentemente reduzem a variabilidade, dependendo da severidade e horário em questão. Os autores ainda concluíram que o coeficiente de variabilidade é a melhor medida para a análise das condições de tráfego e de eventuais perdas nos tempos de viagem.

O método de árvores de regressão apresentou-se como uma ferramenta importante na análise de relações complexas entre clima e tráfego, permitindo distinguir o impacto de cada variável, independentemente da sazonalidade inerente aos horários de pico e variações nos dias da semana.

### 2.3.2 CHAID – Chi-square Automatic Interaction Detector

A partir dos anos sessenta, diversos métodos vêm sendo propostos para fins de análise frente a grandes volumes de dados.

O artigo *An Exploratory Technique for Investigating Large Quantities of Categorical Data* (KASS, 1980), apresenta uma técnica de regressão semelhante ao C&RT, denominada CHAID – *Chi-square Automatic Interaction Detector*.

Esta técnica é na verdade um aprimoramento de uma técnica de regressão conhecida como *AID Automatic Interaction detection* (MORGAN; SONQUIST, 1963).

O algoritmo CHAID processa dados em múltiplas, porém bem definidas etapas de classificação, até que todos os preditores calculados sejam testados. Tais preditores são valores atribuídos a uma classe específica de dados por eles representada. Seu processamento segue uma série de passos cíclicos que se repetem até que os valores mais adequados sejam encontrados.

O primeiro passo é estabelecer a melhor segmentação de dados para cada preditor. São estabelecidas categorias gerais de dados. Na segunda etapa, diversos preditores são comparados e o melhor é escolhido, subdividindo os dados de acordo com a escolha dos preditores dominantes. Na terceira etapa, cada subgrupo é novamente analisado de forma independente, produzindo mais subdivisões para análises.

O tipo do preditor determina qual será o agrupamento dos dados categóricos que deve representar os níveis mais altos de significância obtidos através do teste qui-quadrado, garantindo dessa forma, que haja um número de observações suficientes para constituir uma categoria (KASS, 1980).

Este método permite a análise da interação de dados categóricos e dados contínuos entre si, levando à construção de uma árvore cujos nodos representam os vetores mais significativos na definição das categorias classificadas.

Em termos práticos, o método apresenta-se como uma ferramenta capaz de identificar se uma variável contínua, como por exemplo, velocidade, é influenciada por fatores com volume de precipitação de chuva ou por variáveis categóricas como dia da semana ou hora do dia.

CHAID e CRT são métodos de regressão com objetivos semelhantes. Em uma comparação simplificada, o método de regressão CHAID determina segmentação de um conjunto de dados através do teste qui-quadrado, estabelecendo de forma hierárquica, grupos quantitativamente equivalentes, enquanto o método C&RT produz preditores que se derivam em árvores binárias conforme sua importância frente ao conjunto de dados estudado.

## **2.4 MÉTODOS DE ANÁLISES ESPACIAIS**

### **2.4.1 CHAID e Análise Geográfica**

O artigo *Evaluating the importance of accessibility to congestion response using a GIS-based travel simulator* (CASAS, 2003) combina o método CHAID a análises geográficas.

Foram utilizados sistemas geográficos para simular rotas, que representassem escolhas individuais de usuários da malha viária, frente a uma situação de congestionamento. O método CHAID foi utilizado para avaliar os dados produzidos pelo simulador, elencando as rotas mais prováveis a serem tomadas pelos usuários mediante uma situação de tráfego, atraso nos tempos de viagem e possíveis opções de rotas.

Os resultados apresentam de forma detalhada como a disponibilidade de alternativas pode influenciar nas opções de rotas, permitindo observar geograficamente os prováveis caminhos de utilização. A associação da análise geográfica e do método CHAID permitiu observar que as escolhas mais populares são as menos divergentes das rotas inicialmente planejadas.



### 2.4.2 Métodos de regressão geográfica

No âmbito dos sistemas de informações geográficas, diversos autores vêm buscando métodos de análise e regressão que sejam capazes de produzir informações úteis na identificação de variações de velocidade, tempos de viagem e demanda por transportes (HAYNES et al., 2006) (QUDDUS, 2013) (LOWRY, 2014).

Destes, a interpolação de dados de tráfego (LOWRY, 2014) é um tema particularmente interessante para a compreensão da variabilidade de velocidades, distâncias e tempos de viagem na dimensão geográfica.

O autor analisou dados anuais da média diária de tráfego – *AADT*. Trata-se do volume total de veículos trafegando em uma determinada via durante um ano, dividido por 365 dias. Sua principal aplicação em engenharia de tráfego é referente a análise de demanda.

Os dados em questão foram comparados com estudos prévios (QUDDUS, 2013), onde foram empregados métodos de regressão linear para estimar médias de tráfego em vias municipais do estado de Indiana, Estados Unidos. Isso foi feito através de quatro variáveis explanatórias: população municipal, total em milhas de vias do município, classificação da via entre rural e urbana e proximidade com rodovias estaduais. Com base nesses estudos, foi proposta uma versão modificada do método de centros de stress (SHIMBEL, 1953), agregando ao mesmo modelos de demanda.

Foram processados dados das contagens anuais feitas pelo departamento de engenharia de tráfego, distribuídos em diversas regiões urbanas, com a identificação de pontos de demandas para origem-destino. O modelo foi enriquecido com dados de cadastro municipal, com informações como a natureza de utilização do imóvel e sua área útil construída.

Os resultados foram satisfatórios no que tange à estimação da distribuição espacial das médias de tráfego de veículos ao longo da malha viária, em municipalidades de pequeno e médio porte.

Os métodos de regressão espacial são úteis na identificação da capacidade operacional e respectivas demandas em vias urbanas. A representação geográfica resultante permite ainda identificar pontos críticos passíveis, quando aplicável, de

intervenção gerencial por parte das autoridades de trânsito, beneficiando usuários do sistema de transporte e da malha viária como um todo.

Neste contexto, os métodos de regressão espacial (interpolação), apresentam-se como uma ferramenta útil na compreensão da distribuição de indicadores sobre uma determinada localidade geográfica e período de tempo.

### **2.4.3 Uso da interpolação de dados em análises de transporte**

Selby e Kockelman (2013) contribuíram de forma significativa na utilização de métodos de interpolação aplicados a engenharia de transportes.

Os autores analisaram métodos distintos de regressão geográfica aplicados à predição de tráfego diário a partir de dados de contadores de tráfego no estado do Texas, Estados Unidos, comparando os métodos de Krigagem (CRESSIE, 1990), distância geográfica ponderada (BRUNSDON; FOTHERINGHAM; CHARLTON, 1998), estatísticas descritivas e regressão não geográfica.

Conclui-se que a Krigagem, e a distância geográfica ponderada são incapazes de substituir inteiramente a contagem de trânsito, entretanto, a aplicação de tais métodos reduz de forma significativa a necessidade dessas contagens quando os erros da interpolação espacial apresentarem-se como baixos. Tais métodos também foram considerados úteis em contagens de tráfego em tempo real, velocidades e outras estimativas (por exemplo antecipar e evitar um congestionamento através de técnicas de simulação de tráfego), como também predições demográficas (populações, densidades, censos, etc) (SELBY; KOCKELMAN, 2013).

### **2.4.4 Comparação entre métodos de interpolação**

O uso da Krigagem universal se mostrou útil na predição de estimativas diárias de tráfego em áreas não rodoviárias, conforme observado em Wake County, Carolina do Norte. Foram testados três modelos de semi variogramas (Gaussiano, exponencial e esférico) e quatro métodos de estimativa (mínimos quadrados, quadrados mínimos

ponderados, probabilidade e máxima verossimilhança). Os resultados sugerem a Krigagem universal como um método adequado à previsão geral de tráfego. As vantagens observadas sobre métodos não espaciais são mais evidentes em áreas urbanas onde há uma maior densidade de pontos de medição, com mais pontos próximos (EOM et al., 2006).

Paez (2006) executou regressões sobre dados de tráfego em áreas específicas de interesse, tratando os erros da matriz de covariância de forma similar à Krigagem. O resultado foi um estimador mínimo dos quadrados equivalente à interpolação por vizinho natural, com dados estimados para cada localidade do estudo.

Morency (2011) aprimorou o método para obter coeficientes de variabilidade espacial, com objetivo de estimar a distância percorrida por uma população vulnerável. Tais métodos não levam em consideração dados viários em rede e sim dados entre dois pontos (distância Euclidiana).

Há ainda autores cujos trabalhos estabeleceram as bases para análise espacial de tráfego, através da avaliação dos resultados da regressão de dados de tráfego em rede (BLACK; THOMAS, 1998).

Os autores supracitados consideram os diversos métodos de regressão e estimação espacial ferramentas úteis de análise de tráfego, independentemente do fato dos dados em questão originarem-se de leituras obtidas através de rastreadores, constituindo-se distâncias, velocidades e tempo obtidos em rede ou dados da mesma natureza obtidos como leituras simples entre dois pontos (Euclidianas).

No processo de análise comparativa entre Krigagem e a distância geográfica ponderada, Selby e Kockelman (SELBY; KOCKELMAN, 2013) submeteram dados de distâncias em rede, Euclidianos e não geográficos, a um minucioso comparativo, chegando à conclusão de que em áreas urbanas os erros apresentados pelos dois métodos espaciais variam entre 0,5% e 1,5% de diferença, conforme a Tabela I. Métodos não espaciais apresentam erros de cerca de 20% superiores aos métodos espaciais de interpolação. Em vias urbanas a Krigagem em redes mostrou os melhores resultados.

Estes resultados demonstram que a interpolação de dados de transporte como tempo, distância e velocidade pouco diferem se obtidos a partir de trajetos em redes ou em leituras individuais, fornecendo as bases para análises espaciais de

velocidades e distâncias Euclidianas, processadas a partir de pontos de GPS, data e hora de aquisição. Como será apresentado neste trabalho, independentemente do modelo de coleta (em rede ou não) os erros estão associados à disponibilidade de pontos de leitura, significativamente maiores em ambiente urbano. Adicionalmente, consolida-se a Krigagem como um método adequado para a compreensão de fenômenos de variabilidade no âmbito espacial.

**Tabela 2.5.4-1 – Resultados obtidos por Selby e Kockelman (2013) na comparação de métodos de interpolação e distâncias Euclidianas e em rede. Erros absolutos % destacados em vermelho.**

| Interpolação          | Tipo                   |                        | Houston (sem vias vicinais) | Houston (geral) | Austin (geral) | Rodovias     | Estradas periféricas | Vias urbanas |
|-----------------------|------------------------|------------------------|-----------------------------|-----------------|----------------|--------------|----------------------|--------------|
| Kriging               | Euclidean              | Erro absoluto %        | 63.1%                       | 62.0%           | 55.8%          | 13.7%        | 59.0%                | 61.0%        |
|                       |                        | Erro Médio             | 2.9%                        | 3.9%            | -6.5%          | -1.1%        | 0.6%                 | -2.8%        |
|                       |                        | Erro quadrático médio  | 1.4E+11                     | 2.0E+11         | 4.0E+10        | 1.7E+10      | 3.6E+10              | 5.1E+11      |
|                       |                        | Tipo de semivariograma | Spherical                   | Exponential     | Exponential    | Exponential* | Exponential          | Exponential  |
|                       | Network                | Erro absoluto %        | 62.4%                       | 61.1%           | 55.9%          | 14.5%        | 59.6%                | 60.5%        |
|                       |                        | Erro Médio             | 2.9%                        | 4.1%            | -0.3%          | 0.5%         | 2.1%                 | -2.1%        |
| Erro quadrático médio |                        | 1.4E+11                | 2.0E+11                     | 4.0E+10         | 1.6E+10        | 3.3E+10      | 4.8E+11              |              |
|                       | Tipo de semivariograma | Exponential            | Exponential                 | Exponential     | Exponential*   | Exponential  | Exponential          |              |
| GWR                   | Euclidean              | Erro absoluto %        | 70.2%                       | 68.6%           | 51.8%          | 23.1%        | 83.5%                | 65.0%        |
|                       |                        | Erro Médio             | 7.8%                        | 7.4%            | 1.1%           | -6.7%        | 5.0%                 | 0.0%         |
|                       |                        | Erro quadrático médio  | 1.0E+11                     | 1.5E+11         | 2.8E+10        | 5.4E+10      | 4.8E+10              | 4.9E+11      |
|                       |                        | Kernel                 | Bi-squared                  | Bi-squared      | Bi-squared     | Bi-squared   | Bi-squared           | Bi-squared   |
|                       | Network                | Erro absoluto %        | 70.1%                       | 68.1%           | 60.1%          | 21.4%        | 75.7%                | 63.5%        |
|                       |                        | Erro Médio             | 8.6%                        | 8.2%            | 2.4%           | -5.2%        | 5.6%                 | 0.8%         |
| Erro quadrático médio |                        | 1.2E+11                | 1.6E+11                     | 2.6E+10         | 4.1E+10        | 4.5E+10      | 4.9E+11              |              |
|                       | Kernel                 | Bi-squared             | Bi-squared                  | Bi-squared      | Bi-squared     | Bi-squared   | Bi-squared           |              |
| FGLS                  | (aspacial)             | Erro absoluto %        | 103.6%                      | 103.0%          | 115.3%         | 38.4%        | 114.0%               | 80.6%        |
|                       |                        | Erro Médio             | 8.5%                        | 9.1%            | -8.6%          | -10.9%       | 6.8%                 | -3.4%        |
|                       |                        | Erro quadrático médio  | 4.91E+11                    | 4.63E+11        | 3.59E+11       | 7.62E+11     | 1.48E+11             | 3.25E+12     |

## 2.4.5 Krigagem Baesiana Empírica

Criada por Georges Matheron e Daine G. Krige, a Krigagem é um método popular de interpolação que assume que a distância ou direção entre pontos de uma amostra constituem-se como uma correlação espacial, que pode ser utilizada para explicar a variação sobre uma superfície (CRESSIE, 1990; ESRI, 2015; SELBY; KOCKELMAN, 2013).

Sua formula é expressa como:

$$\hat{Z}(s_0) = \sum_{i=1}^N \lambda_i Z(s_i)$$

Onde:

$Z(s_i)$ : medida do valor em um local

$\lambda_i$ : peso (importância) desconhecido para a medida de valor de um local

$s_0$ : local de predição

$N$  = número de valores medidos

A krigagem baesiana empírica (ESRI, 2015) é um método de interpolação geoestatística que automatiza aspectos da construção de um modelo de krigagem válido, considerando a ponderação local entre semivariogramas, mimizando erros padrão de predição.

Outros modelos de krigagem assumem que o processo segue uma média geral (ou tendência especificada) com variações individuais em torno dessa média. Grandes desvios acabam sendo aproximados às médias, mantendo valores de desvio restritos a uma certa amplitude em relação à média. No entanto, a Krigagem baesiana empírica não assume uma tendência para uma média geral. Dessa forma, grandes desvios são tão propensos a aumentar quanto reduzir os resultados apresentados na superfície interpolada. Assim, as funções aleatórias intrínsecas ocorrem inerentemente às tendências dos dados de forma automatizada.

Este método é parte da ferramenta Spatial Analyst do Software ArcGIS e seus resultados serão apresentados à frente.

#### **2.4.6 Considerações Finais do Capítulo**

Com vistas ao cumprimento dos objetivos aqui propostos, esse trabalho pretende combinar uma visão geral dos indicadores de variabilidade com o uso de medidas de dispersão, aprofundando a compreensão do fenômeno através do uso de árvores de regressão - para a mineração de dados, comparando, por fim, classes e padrões identificados no âmbito espacial.

Os métodos estatísticos e geográficos aqui apresentados foram aplicados por seus autores em problemas típicos de transporte, como tempos de viagem, velocidades e ainda a influência da sazonalidade temporal e do clima em sua variabilidade. Tais autores são as bases para que o presente trabalho apresente um método na melhoria da oferta de serviços públicos à população de São Paulo e outras grandes cidades brasileiras.

### 3 MATERIAIS E MÉTODOS

#### 3.1 CARACTERIZAÇÃO DOS ARQUIVOS COM INFORMAÇÕES DOS ÔNIBUS FORNECIDOS PELA SPTRANS

Foram obtidos, junto à SPTrans, dados provenientes dos veículos rastreados do transporte público paulistano durante março de 2014.

Esses arquivos possuem os registros de localização geográfica, data e hora de aquisição dessas coordenadas, bem como a linha e veículo de origem. Trata-se de um leiaute típico de sistemas de rastreamento, cuja informação contida está diretamente atrelada ao seu uso fim.

Tabela 3.1-1 – Modelo de arquivo MO.txt fornecido pela São Paulo Transportes

| Campos originais arquivo MO.txt |   |
|---------------------------------|---|
| Campo                           | Conteúdo  |
| Data 1                          | Data/hora do Servidor de recebimento dos dados            |
| Data 2                          | Data/hora do momento de captura do dado                   |
| Código AVL                      | Código único identificador do dispositivo de rastreamento |
| Latitude                        | Latitude  |
| Longitude                       | Longitude   |
| Código da linha                 | Código de identificação da linha                          |
| Código de referências           | Código de referência a pontos de interesse                |
| Sentido                         | Sentido ida ou volta                                      |

## 3.2 MODELAGEM DE DADOS

### 3.2.1 Dados alfanuméricos

Foi constituído um banco de dados Postgree com aproximadamente 35MM de registros gerados a partir dos dispositivos rastreadores dos ônibus de transporte da cidade de São Paulo no mês de março de 2014. A tabela principal deste banco de dados foi produzida através da combinação dos arquivos de movimento diário, fornecido pela SPTrans.

Testes prévios demonstraram que as ferramentas de processamento estatístico e geográfico, utilizadas neste trabalho, possuíam limitações em relação ao volume de registros processados. Essas limitações foram determinantes na definição da seleção do conjunto de dados analisados.

Diversas inconsistências nos dados foram encontradas, representando aproximadamente 3% dos registros gerados no mês de março de 2014. Diante deste fato, um processo de limpeza e validação foi efetuado.

Dentre os problemas encontrados, destacam-se registros com ausência de coordenadas de latitude, longitude ou ambas e datas-hora inconsistentes. Em relação as inconsistências de data-hora, é importante observar que é esperado que os registros encontrados nos campos “data hora AVL” sejam diferentes do campo “data hora servidor”. No entanto, esta observação refere-se a registros com valores visivelmente incorretos, contando, em sua maioria, com registros referentes aos anos de 1900 e 1960, ocorrendo principalmente no campo “data-hora” avl.

Kamga e Yazici (2014) relatam dificuldades similares na modelagem de dados obtidos através dos dispositivos de posicionamento global – GPS da frota de táxi da cidade de Nova Iorque. Dentre as possíveis causas sugeridas pelos autores, destacam-se erros provenientes de hardware, software ou a combinação de ambos na aquisição de coordenadas GPS. Os autores descartaram 1,4% dos registros adquiridos, um percentual inferior aos 3% relatados neste trabalho.



No caso de São Paulo a causa é desconhecida, mas é possível que tais erros sejam advindos, entre outros fatores, da falta de manutenção nos dispositivos, como por exemplo: atualizações de software e substituição de baterias do relógio interno.

Foram descartados todos os registros com intervalos de leitura maior que cinco minutos. Esse descarte tem como objetivo garantir uma frequência temporal uniforme, de aquisição de dados, eliminando equipamentos com envios intermitentes.

Uma vez concluídos os trabalhos de limpeza e validação de dados, foram selecionados registros correspondentes a todas as segundas-feiras, terças-feiras, quartas-feiras, quintas-feiras e sextas-feiras do mês de março, na faixa de horário compreendida entre 7:00 horas e as 20:59 horas.

Adicionalmente, o número de registros foi circunscrito ao tráfego de veículos em uma área de estudo definida, conforme explicado à frente.

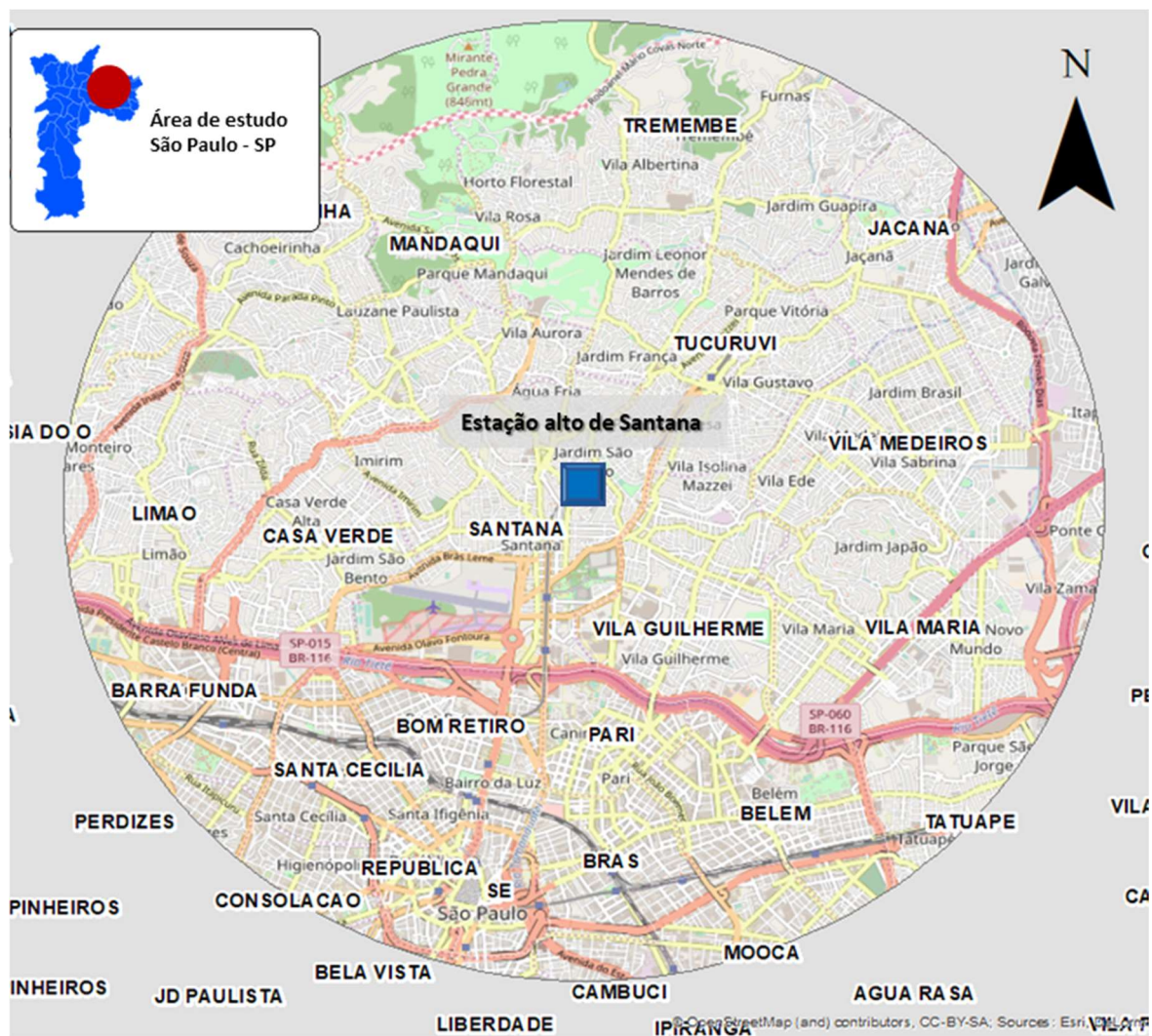
Concluída a limpeza e seleção, foram criados campos adicionais para simplificar futuros processamentos:

- Campo “hora cheia” – derivado do campo data/hora do servidor para fins de agrupamento;
- Campo dia da semana – derivado do campo data/hora do servidor para fins de agrupamento;
- Campo “contagem” – para fins de processamento estatístico.

### **3.2.2 Dados pluviométricos**


Há 2,6 km de distância das Marginais do Rio Tietê, a Estação Mirante de Santana localiza-se na latitude  $-23.496294^{\circ}$  e longitude  $-46.620088^{\circ}$ , a 786 metros de altitude na região do Jardim São Paulo, São Paulo capital (Figura 3.2.2-1). Dentre seus instrumentos, destaca-se o pluviômetro, instrumento responsável pelas leituras de precipitação de chuva em milímetros, fornecidas pelo INMET através da sua divisão de dados meteorológicos - SADMET (INMET, 2017).

Figura 3.2.2-1 – Mapa com a localização da estação meteorológica Alto de Santana



As leituras pluviométricas foram obtidas através de solicitação junto ao INMET. Os dados foram disponibilizados através de uma tabela com estrutura exemplificada pela Tabela 3.2.2-1. Foi necessária a conversão do horário UTC para horário local em São Paulo. As informações foram importadas para o banco de dados principal e processadas conforme explicado a seguir.

Tabela 3.2.2-1 – Modelo de arquivo fornecido contanto dados pluviométricos fornecidos pelo INMET

|  MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO-MAPA<br>INSTITUTO NACIONAL DE METEOROLOGIA - INMET<br>ESTAÇÃO METEOROLÓGICA AUTOMÁTICA DE SÃO PAULO/SP |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |
|--|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| Alt.   | 792,06m           |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |
| Lat.   | 23°30'S           |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |
| Lon.   | 46°37'W           |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |
|  | PRECIPITAÇÃO (mm) | PRECIPITAÇÃO (mm) | PRECIPITAÇÃO (mm) | PRECIPITAÇÃO (mm) | PRECIPITAÇÃO (mm) | PRECIPITAÇÃO (mm) | PRECIPITAÇÃO (mm) | PRECIPITAÇÃO (mm) | PRECIPITAÇÃO (mm) | PRECIPITAÇÃO (mm) | PRECIPITAÇÃO (mm) | PRECIPITAÇÃO (mm) | PRECIPITAÇÃO (mm) | PRECIPITAÇÃO (mm) | PRECIPITAÇÃO (mm) | PRECIPITAÇÃO (mm) |
| HORA UTC   | 0000              | 0100              | 0200              | 0300              | 0400              | 0500              | 0600              | 0700              | 0800              | 0900              | 1000              | 1100              | 1200              | 1300              | 1400              | 1500              |
| 01-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 02-jan-2014  | 0,0               | 1,4               | 1,2               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 03-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 04-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 05-jan-2014  | 0,2               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 06-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 07-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 08-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 09-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 10-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 11-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 12-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,2               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 13-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 14-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 15-jan-2014  | 2,0               | 0,0               | 0,0               | 2,2               | 8,8               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 16-jan-2014  | 0,2               | 0,0               | 1,6               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 2,4               |
| 17-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,2               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 18-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 19-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 20-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 21-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 22-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 23-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 24-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |
| 25-jan-2014  | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               | 0,0               |

### 3.2.3 Combinação entre dados de rastreamento e dados pluviométricos

Os dados pluviométricos estão distribuídos em intervalos que representam o total acumulado em cada hora. Assim, um registro constando 6 mm de precipitação de chuva as 14:00 horas informa a precipitação acumulada entre 14:00 horas e 14:59 horas. Comparativamente, os dados de telemetria provenientes dos ônibus de transporte público de São Paulo estão sendo continuamente produzidos, em intervalos que variam entre 30 segundos a 5 minutos (tipicamente a cada 80 segundos), conforme critérios estabelecidos, que serão posteriormente detalhados. Diante de tais limitações, optou-se por considerar o valor da precipitação registrado para uma hora em todas as leituras de telemetria de ônibus geradas no mesmo período, independentemente da leitura ter sido gerada conforme exemplo, às 14:01 horas ou às 14:56 horas (Figura 3.2.3-1).

Figura 3.2.3-1 – Diagrama ilustrativo: generalização da precipitação horária



### **3.3 PROCESSAMENTO GEOGRÁFICO DE DADOS**

Através do software ArcGis um banco de dados no formato proprietário “Esri Geodatabase” foi criado a partir de uma única tabela constando dados de telemetria e precipitação já integrados.

Uma vez processada a importação de ambos os dados, o passo seguinte foi a criação de pontos a partir das coordenadas geográficas geradas pelos dispositivos rastreadores embarcados nos ônibus, disponíveis na tabela. Esse processamento gerou uma nova entidade no “Geodatabase” que serviu de base para os processamentos subsequentes.

#### **3.3.1 CÁLCULO DE GRANDEZAS A PARTIR DOS DADOS DE TELEMETRIA**

A ferramenta “Track Intervals to Features” é parte integrante do módulo “Tracking Analyst” do software Arcgis. Este recurso calcula grandezas em um intervalo entre dois pontos. Presume-se que a ferramenta explore a equação de Haversine (e suas derivações) (SHUMAKER; SINNOTT, 1984) a qual fornece a distância entre dois pontos na superfície terrestre a partir de suas latitudes e longitudes.

As opções de parametrização da ferramenta permitem que se estabeleçam critérios de agrupamento e indexação dos dados a serem processados. Neste trabalho definiu-se o código identificador do dispositivo embarcado como indexador sequencial.

Este código é único e garante que os dados gerados por cada veículo sejam processados individualmente. O processamento foi estabelecido em sequência cronológica contínua, garantindo que o resultado final representasse valores desempenhados em intervalos recorrentes de leitura, sobre um mesmo veículo ao longo da sua operação diária.

Esse procedimento gerou quatro novos campos no “Geodatabase” (Figura 3.3.3-1):

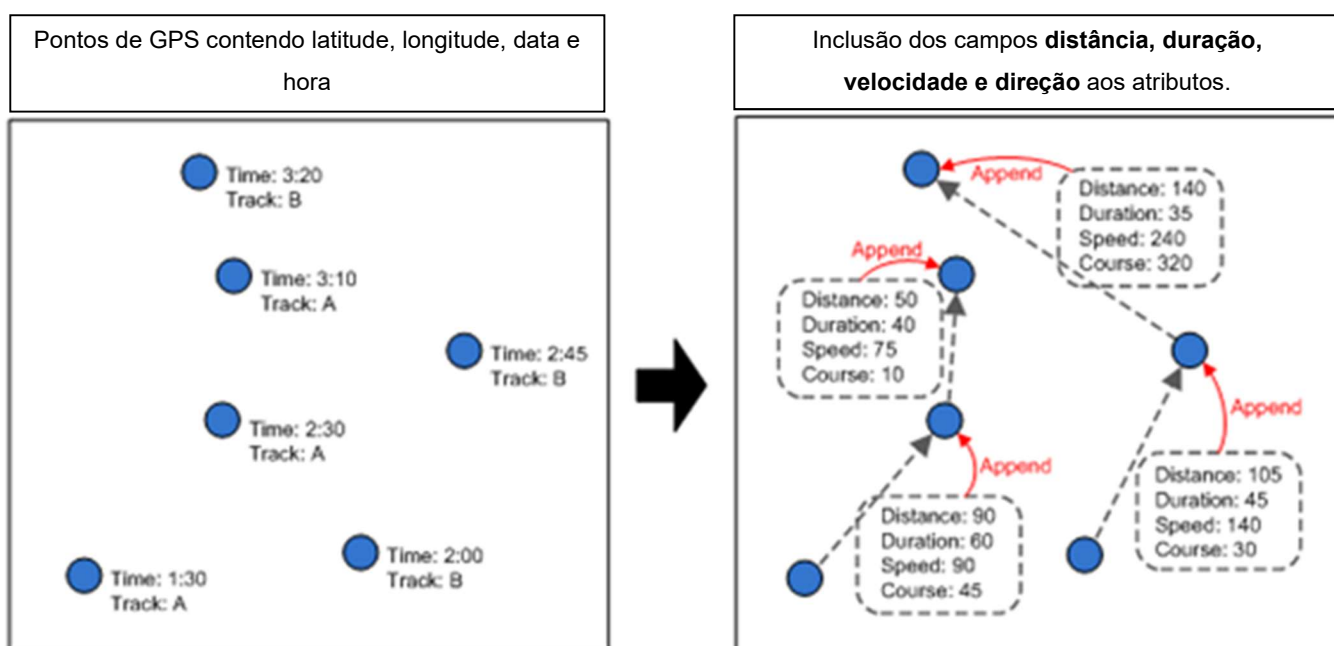
**Velocidade média;**

**Distância euclidiana entre pontos de leitura;**

**Intervalo de tempo entre leituras;**

**Sentido de deslocamento.**

Figura 3.3.3-1 – Diagrama explicativo do cálculo de grandezas a partir de coordenadas geográficas e data/hora: ferramenta “track intervals to features”



### 3.3.2 DEFINIÇÃO DO MÉTODO DE ANÁLISE DA VARIABILIDADE

Kamga e Yazici (2014) analisaram tempos totais de viagem nos táxis de Nova Iorque. Os autores consolidaram diversas pesquisas de menor escala abordando o impacto do clima na variabilidade dos tempos de viagem, através de diversas estratégias de análise baseadas em medidas de dispersão.

Selby e Kockelman (2013) analisaram dados provenientes de contadores de tráfego de todo o estado do Texas. Os autores abordaram a variabilidade das velocidades medidas sem a distinção individual de veículos.

Lowry (2014) analisou dados anuais da média diária de tráfego – *AADT* abordando volume de tráfego. Não menos interessante, seu trabalho buscou avaliar a demanda e a capacidade das vias através de análises espaciais.

As abordagens supracitadas compartilham o objetivo de caracterizar a variabilidade, seja temporal, geográfica ou estatística a fatores como clima, capacidade da via, sazonalidade e demanda, entre outros.

#### 3.2.3.1 Microviagens: Conceito

Com base nos recursos disponíveis, este trabalho estabeleceu a comparação da variabilidade apresentada nos intervalos de leitura individuais como método de análise. Os pontos foram tratados de forma análoga a abordagem de Selby e Kockelman (2013) junto a medidores de velocidade por indução (loop).

Empreender neste trabalho, técnica semelhante ao cálculo dos tempos de viagem (KAMGA; YAZICI, 2014) constitui-se num desafio técnico significativo. Foram observadas lacunas na transmissão de dados, bem como variação nas rotas. Apesar do grande volume, a construção de sequências completas de viagens a partir de dados brutos de rastreamento constitui um desafio técnico cujo tempo envidado inviabilizaria a apresentação dos resultados no prazo disponível.

Este contexto definiu a estratégia de processamentos de dados conforme detalhado à frente.

No âmbito espacial, as análises volumétricas (LOWRY, 2014) mostraram-se eficientes, produzindo resultados aderentes ao contexto urbano, no qual o presente trabalho está inserido.

Apesar das limitações técnicas, os dados produzidos pelos dispositivos de telemetria dos ônibus urbanos da cidade de São Paulo, quando processados e analisados podem constituir-se como de grande valor na gestão de trânsito e transporte na capital paulista.

Neste contexto, estabeleceu-se como objeto a análise de variabilidade das velocidades médias individuais e análise da variabilidade das distâncias euclidianas percorridas aqui denominadas microviagens. O termo microviagem será aqui utilizado para referir-se ao resultado do processamento das grandezas entre dois pontos de GPS sequencialmente estabelecidos no tempo e espaço tais como velocidade, distância e tempo.

Embora já descrito, é importante ressaltar que estes pontos são gerados a partir de dados de um mesmo rastreador único e estão ordenados de forma cronológica.

O intervalo de tempo de duração dessas microviagens foi utilizado para estabelecer um conjunto de dados com limites regulares de aquisição.

Os dados referentes à direção do deslocamento não foram utilizados devido à parametrização estabelecida ser individual-cronológica, como já descrita, eliminando a possibilidade de por exemplo, processar um intervalo entre dois veículos passando próximos em sentido contrário ou registros de veículos produzidos em uma mesma localidade em dias ou horas diferentes.

### **3.3.3 SELEÇÃO DA ÁREA DE ESTUDO**

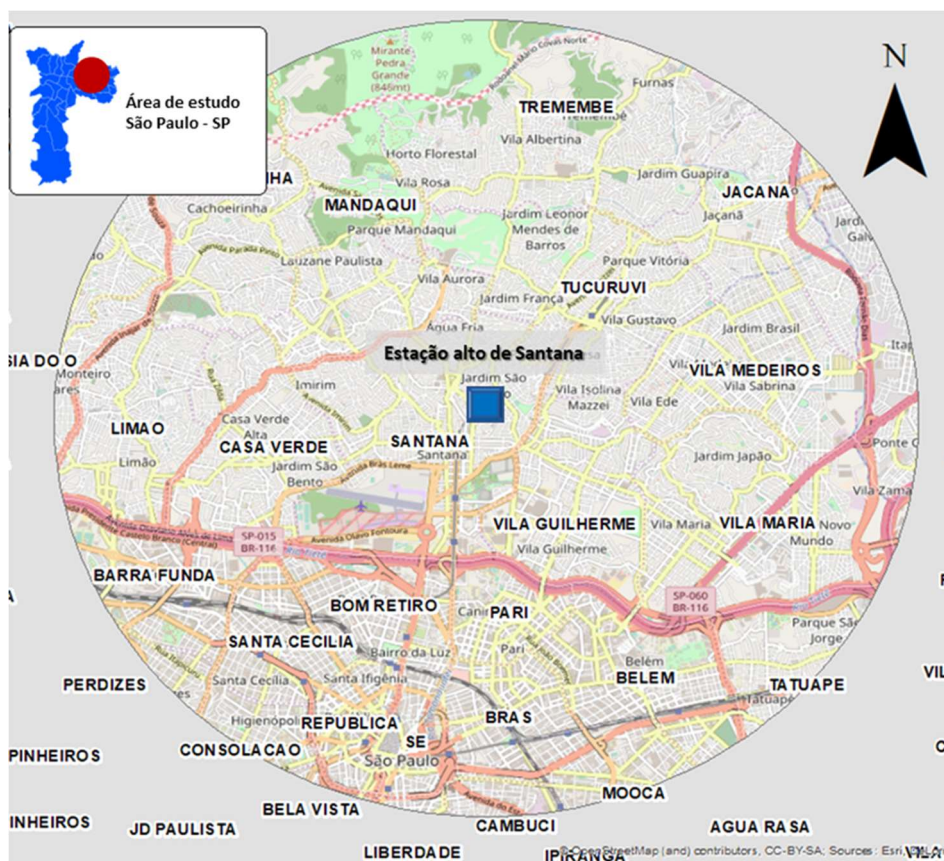
Os autores investigados neste trabalho não estabelecem critérios de definição de abrangência de fenômenos meteorológicos em seus estudos. Kamga e Yazici (2014) utilizam dados meteorológicos históricos da região da ilha de Manhattan, Nova Iorque, Estados Unidos, obtidos via portais de internet especializados.

Hranac et al (2006) e Akin et al. (2011) utilizaram dados obtidos através de estações meteorológicas ao longo de rodovias e vias urbanas com área de abrangência variada. Observa-se que não há um padrão bem estabelecido para a amplitude geográfica sobre os eventos meteorológicos.

Neste contexto, definiu-se uma área de 7 quilômetros no entorno na estação meteorológica do Alto de Santana como objeto de estudo (Figura 3.3.3-1).

Como justificativa, o raio de abrangência cobre diferentes tipos de vias urbanas, incluindo vias centrais, vias de alta velocidade - como a marginal Tietê - e ainda parte do centro da cidade, bairros como Limão, Barra Funda, Santa Cecília, Santana e regiões menos densamente povoadas, próximas ao Horto Florestal e de onde atualmente encontra-se, em construção, o Rodoanel Norte.

Figura 3.3.3-1 – área de estudo (figura 2)



### 3.4 FLUXO DE MODELAGEM E PROCESSAMENTO DE DADOS

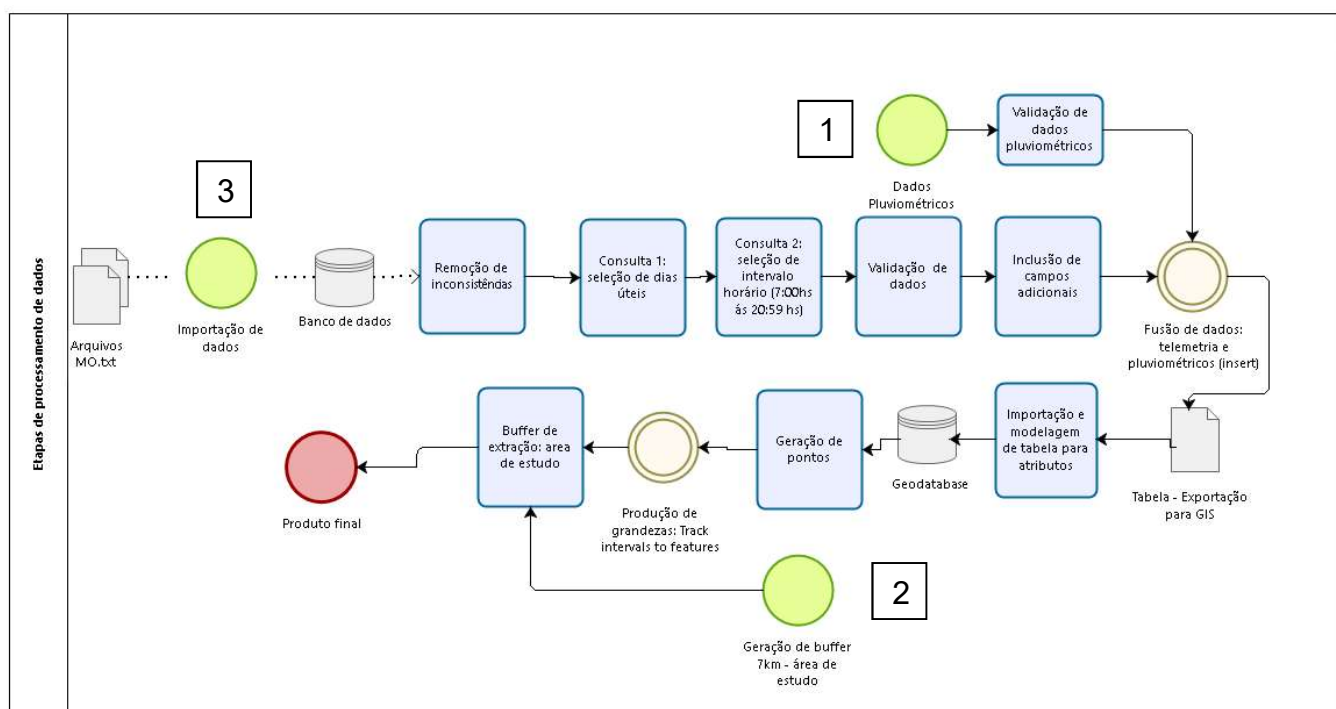
O presente item visa recapitular de forma estruturada as etapas já descritas que compõem o processo de modelagem de dados.

Tratam-se de três processos interdependentes (Figura 3.4-1):



- O primeiro compreende a validação dos dados pluviométricos cedidos pelo INMET – Instituto Nacional de Meteorologia. Como tais dados são oficiais, a validação se deu no âmbito da integridade de dados;
- O segundo compreende a definição da área de estudo e procedimentos de geoprocessamento para a produção de uma feição correspondente;
- O terceiro abrange todo o tratamento dos dados de telemetria e incorporação dos processos já descritos, culminando no banco de dados geográfico tratado como “produto final”.

Figura 3.4-1 – Etapas da modelagem de dados

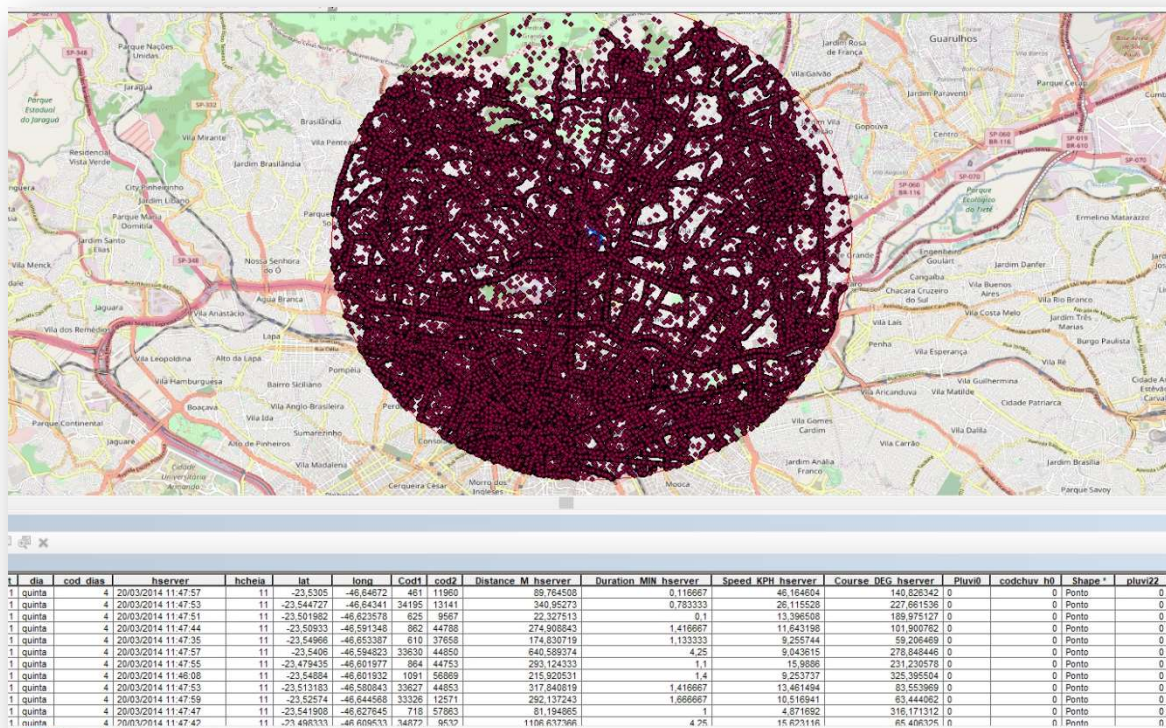


## 4 PÓS PROCESSAMENTO E RESULTADOS

O produto final oriundo do processo de modelagem e seleção resultou em uma entidade contida em um Esri Geodatabase cujos atributos compreendem dados de telemetria, dados de microviagens e dados pluviométricos (Figura 4.1).

Embora a descrição do processo de modelagem de bancos seja necessária, não é objeto de estudo e sim ponto de partida para o cumprimento da proposta deste trabalho.

Figura 4.1 – Produto final pós modelagem de dados



## **4.1 CARACTERIZAÇÃO GERAL DOS DADOS**

Autores (LIU; RECKER; CHEN, 2004) abordam a variabilidade como resultado da influência de um conjunto de fatores de diferentes características sobre o trânsito urbano, rodoviário e até mesmo o rural.

Como já relatado, a análise dos dados de contadores de tráfego da cidade de Istambul demonstrou que o clima inclemente produziu impacto nas velocidades e nas taxas de fluxo nas duas seções de vias estudadas. Diante da precipitação, as reduções se deram em média entre 8% e 12% (AKIN; SISIOPIKU; SKABARDONIS, 2011).

Os autores aqui mencionados serviram como exemplos. De fato, em todos os métodos investigados, o estudo da variabilidade se deu através da identificação prévia dos padrões operacionais em condições consideradas normais.

Para esse trabalho a caracterização dos dados contidos no produto do processamento de dados brutos de rastreamento é a primeira de três etapas analíticas que abordarão respostas aos objetivos propostos: padrões gerais da operação de ônibus e a variabilidade em si, incluindo a influência climática e a distribuição geográfica destes padrões.

### **4.1.1 Caracterização das velocidades das microviagens**

Estudos vêm investigando o comportamento da demanda de usuários em ônibus de transporte durante horários de pico, eventos climáticos, entre outras situações (STOVER; MCCORMACK, 2012), sem chegar a um consenso definitivo se existe um padrão de comportamento da demanda (CHANGNON, 1996; GUO; WILSON; RAHBEE, 2007; KASHFI; BUNKER; YIGITCANLAR, 2016; KHATTAK; DE PALMA, 1997b)

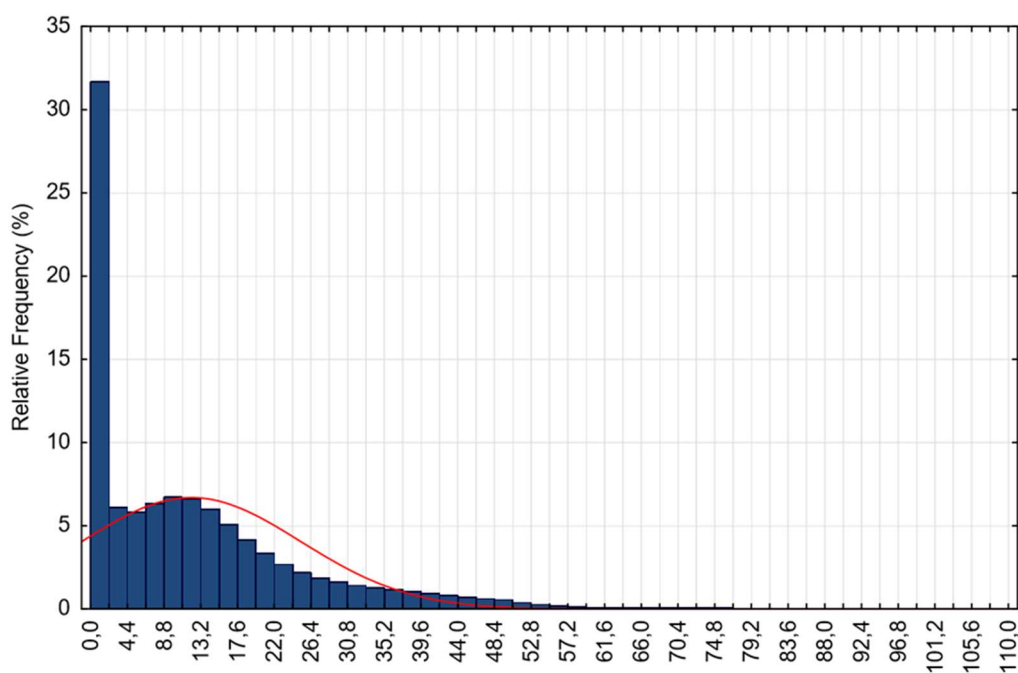
No mês de março de 2014, 243 milhões de passagens foram processadas ao longo de cerca de 18 mil pontos de parada (SPTRANS, 2017).

A análise da distribuição normal das velocidades das microviagens permitiu constatar a extensão do fenômeno de embarque e desembarque dos ônibus de transporte público de São Paulo. Cerca de 32% dos registros apresentaram velocidade média igual a zero (Gráfico 4.1.1-1).

É possível que além do embarque e desembarque de passageiros, outros fenômenos, como por exemplo semáforos ou tráfego intenso, contribuam neste contexto. Segundo a Prefeitura de São Paulo em cerca de 50% do tempo os ônibus encontram-se em movimento. O restante é dividido entre semáforos e pontos de parada de forma equivalente (PERON, 2015). Outros trabalhos constataram, em vias expressas, tempos de parada em semáforo na ordem de 6% (HOSHINA, L. N. N; CHIOVETTI, P. B.; DELUCA, 2015). Embora haja certa disparidade entre estas duas constatações, é preciso levar em consideração que, de uma forma ou de outra, os veículos, em parte de seu tempo, não se movimentam.

Assim, este trabalho toma com premissa que os registros com velocidades próximas a zero enquadram-se, conforme supracitado, em eventos de embarque, desembarque, semáforos e ocasionalmente, congestionamentos.

Gráfico 4.1.1-1 – Histograma (normal): Velocidades médias (km/h) - microviagens

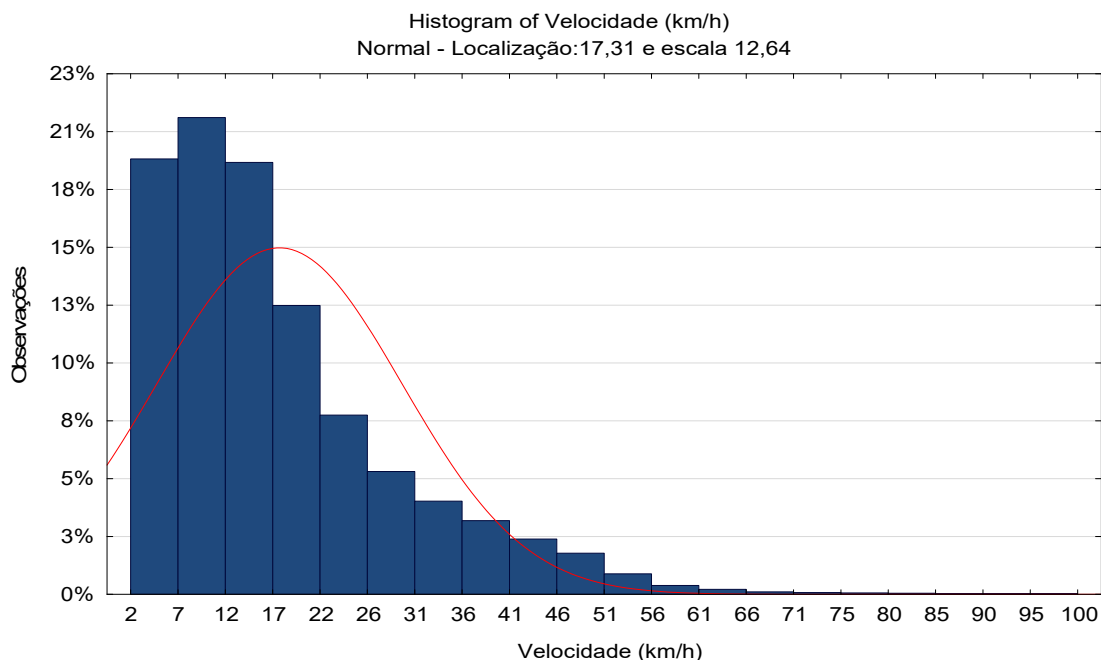


Os estudos sobre variabilidade dos tempos de viagens (CHIEN; KOLLURI, 2012; KAMGA; YAZICI, 2014; VAN LINT; VAN ZUYLEN; TU, 2008), em geral não tratam de ônibus urbanos, bem como as análises baseadas em observações pontuais de variabilidade em velocidades ou fluxo de tráfego geral (AKIN; SISIOPIKU; SKABARDONIS, 2011; LOWRY, 2014; SELBY; KOCKELMAN, 2013) não distinguem o tipo de transporte sobre os eventos medidos.

É possível afirmar que as características operacionais dos ônibus de transporte urbano, como as encontradas em São Paulo e provavelmente análogas às grandes metrópoles ao redor do mundo, são um obstáculo à análise da variabilidade, seja das velocidades, dos tempos de viagem, dos fluxos de veículos entre outros. Considerando tais fatos, as classes de registros categorizadas na distribuição normal como microviagens em velocidades entre zero e dois quilômetros foram excluídas da análise deste trabalho. O objetivo dessa exclusão foi garantir que a análise da variabilidade das velocidades se atenha aos eventos relacionados aos veículos em movimento. Com a aplicação desse filtro o conjunto final de dados analisados contabilizou 19 milhões de registros.

Uma nova análise da distribuição normal das microviagens com velocidades a partir de 3 quilômetros por hora demonstrou uniformidade na representatividade das microviagens até 20 quilômetros por hora de velocidade média. Conforme o esperado, microviagens com médias a partir de 30 quilômetros por hora não são mais significativas do que 1%, confirmando que a representatividade das categorias de velocidades médias se reduz, na medida que os valores observados aumentam (Gráfico 4.1.1-2).

**Gráfico 4.1.1-2 – Histograma (normal): Microviagens com velocidades médias a partir de 3 km/h**



**Gráfico 4.1.2-1 – Histograma (normal): Distâncias Euclidianas - microviagens**

#### 4.1.2 Caracterização do intervalo de aquisição de dados de telemetria

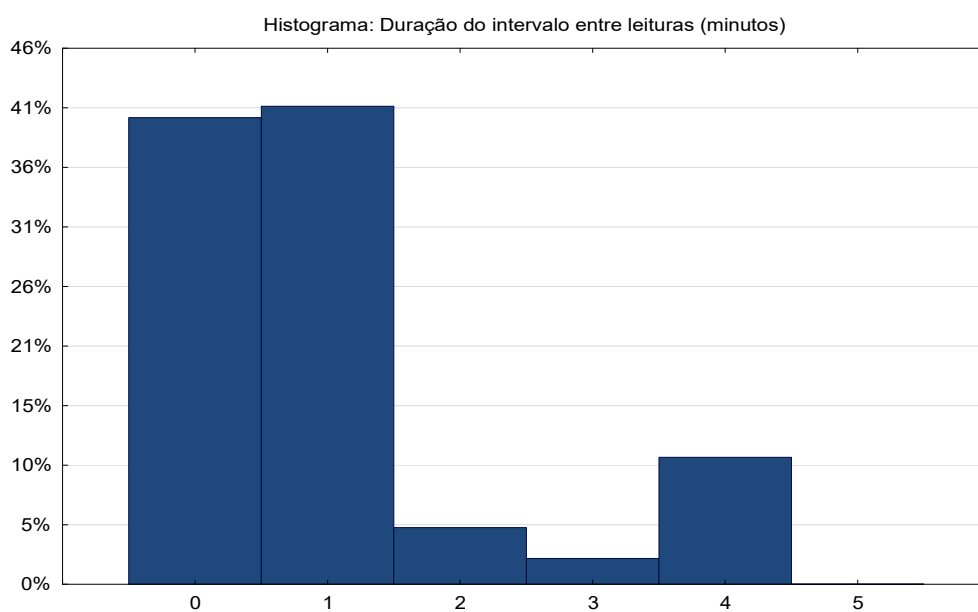
As causas dos intervalos alongados de envio de dados permanecem incertas. Conforme observado, a dispersão geográfica dos pontos pode elevar os erros apresentados frente a processos de interpolação de superfície (SELBY; KOCKELMAN, 2013).

Conforme mencionado, foram desconsiderados os registros cuja sequência possuem intervalos de leitura maior que cinco minutos. Esse fato é considerado irrelevante frente a um volume de exclusão com representatividade abaixo de 1% dos registros. Presume-se que não um, mas uma combinação de fatores possa vir a causar essa situação como por exemplo a permanência por longos períodos em áreas cuja cobertura de dados móveis possa ser deficitária, devido a características físicas ambientais ou técnicas, falhas gerais de software e falhas gerais de hardware.

A análise da distribuição normal da frequência das leituras demonstra que, conforme esperado, a maioria das microviagens encontra-se com intervalos de leitura

de até 2 minutos, ou seja, em intervalos regulares significativamente curtos (Gráfico 4.1.2-1)

Gráfico 4-1.2-1 – Histograma (normal): Duração do intervalo de leitura (em minutos) - microviagens



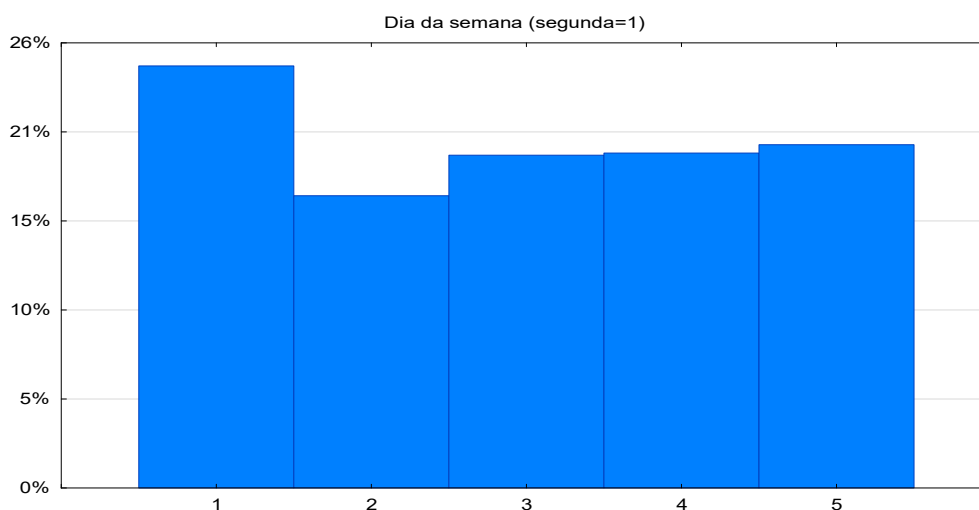
As microviagens em análise são materializadas em distâncias Euclidianas das leituras dos dispositivos de rastreamento dos ônibus. Selby e Kockelman (2013) demonstraram, através da comparação com métodos de interpolação dos dados de tráfego, que o erro médio observado comparando distâncias euclidianas e não euclidianas é significativamente baixo, de cerca de 1,5%. Os autores utilizaram dados dos detectores de massa metálica dispostos em distâncias conhecidas para analisar velocidades.

A restrição de intervalos de leitura para até 5 minutos, empreendida no presente trabalho, bem como a supressão das velocidades médias menores que 3 km/h, devem produzir, embora não seja objeto desse estudo, erros semelhantes aos observados por Selby e Kockelman na caracterização da variabilidade das velocidades entre redes Euclidianas e não Euclidianas, permitindo aplicações na gestão do transporte público municipal em São Paulo.

### 4.1.3 Caracterização da distribuição das microviagens em dias da semana.

A análise da distribuição normal das microviagens durante os dias úteis semanais demonstra um padrão uniforme. Um número maior de veículos foi registrado às segundas feiras (1) em março de 2014. Nas terças (2) um número ligeiramente menor de veículos foi registrado. As quartas (3), quintas (4) e sextas (5) apresentaram uma quantidade semelhante de veículos circulantes (Gráfico 4.1.4-1).

Gráfico 4.1.4-1 – Histograma (normal): Microviagens por dia da semana



### 4.1.4 Caracterização da distribuição das microviagens por hora do dia

Franklin e Karlstrom (2009) constataram que a variabilidade nos tempos de viagem pode ser explicada, em parte, pela demanda ao longo das horas do dia em ambiente urbano. Segundo Martchouk; Chien e Kolluri (2010) desvios padrão elevados são observados em horários de pico. Ainda sobre a variabilidade horária, o mesmo autor observou que os tempos de viagem podem chegar a mais de 100% dos valores observados sobre a média em horários de pico.

Kwon (2011) observou em rodovias que os horários de picos da tarde possuem maior demanda que os picos da manhã. No entanto, a variabilidade nos picos da tarde



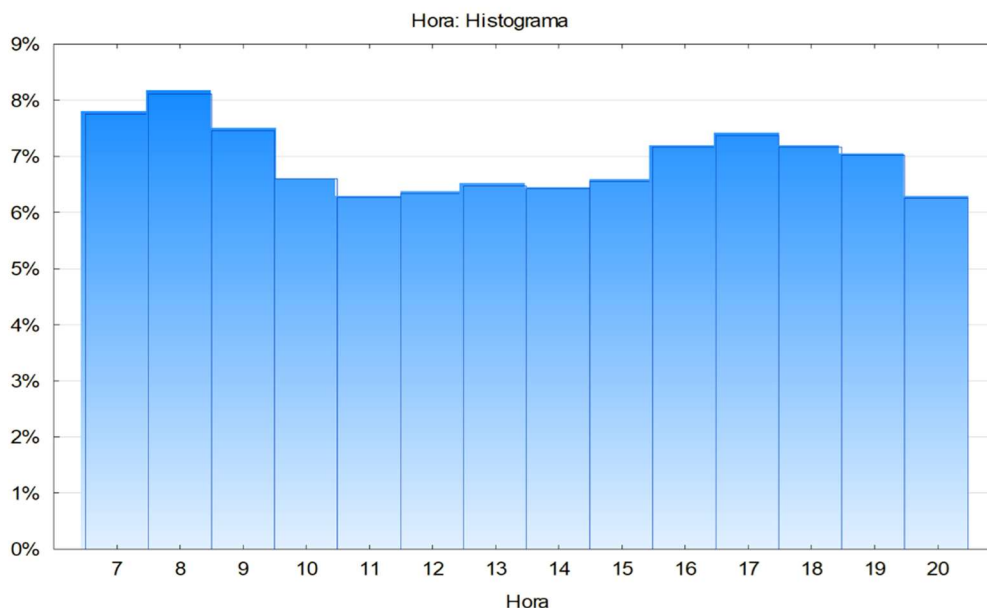
é menor, possivelmente, segundo o autor, devido a possíveis gargalos no fluxo de tráfego.

Diferentemente das velocidades médias e distâncias euclidianas observadas nas microviagens, a distribuição normal dos dados em dias de semana ou horas do dia não caracteriza diretamente fenômenos relativos à variabilidade. Considera-se que a observação da distribuição normal revele intervenções operacionais cotidianas, como por exemplo a inclusão de um número maior de veículos circulantes, visando atender as demandas dos picos da manhã e da tarde.

Apesar destas observações, é esperado que as intervenções operacionais reflitam picos de demanda, cuja ocorrência está diretamente associada à variabilidade.

A análise da distribuição normal demonstra um número maior de microviagens nos horários matinais, entre 6:00hs e 8:00hs da manhã. Já no período da tarde observa-se um aumento das microviagens no horário das 15:00hs às 19:00 horas (Gráfico 4.1.5-1). No caso das intervenções manuais refletirem a variabilidade observada em campo, as observações de Kwon (2011) para rodovias podem ser aplicáveis também ao transporte público. Parte dessa afirmação é comprovada no corredor Rebouças em São Paulo, cujo pico da tarde é 42% mais lento que o da manhã (PERON, 2015).

Gráfico 4.1.5-1 – Histograma (normal): Variação horária diária - microviagens



#### 4.1.5 Caracterização da distribuição pluviométrica

Foi observado o comprometimento de cerca de 3% a 5% dos tempos de viagem em rodovias interurbanas quando um evento de precipitação de chuva está em andamento (SMITH et al., 2004).

Akin; Sisiopiku e Skabardonis (2011) concluíram que o clima inclemente teve impacto nas velocidades médias, com variação entre 8% e 12%.

Muitos dos estudos reconhecem o impacto da precipitação de chuva sobre a variabilidade geral do tráfego, seja em rodovias ou em vias urbanas, no entanto, a quantificação dos fenômenos meteorológicos não é precisa. Diversos trabalhos classificam os eventos climáticos por tipo e intensidade (AKIN; SISIPIKU; SKABARDONIS, 2011; KAMGA; YAZICI, 2014; LAMM; CHOUEIRI; MAILAENDER, 1990), sem detalhar o volume escalar desses eventos.

Assis et. al (1996) propõem na Tabela 4.1.6-1 um dimensionamento da intensidade da precipitação horária, de acordo com o seguinte intervalo de classes:

Tabela 4.1.6-1 – Classificação da intensidade da precipitação segundo Assis et. al (1996)

| Intervalo de precipitação (mm) | Classificação da intensidade de precipitação |
|--------------------------------|--|
| (0,1 – 5)                      | Muito fraca                                  |
| (5 – 10)                       | Fraca  |
| (10 – 15)                      | Moderamente fraca                            |
| (15 - 20)                      | Moderada                                     |
| (20 - 25)                      | Moderamente forte                            |
| (25; 30]                       | Forte  |
| > 30                           | Muito forte                                  |

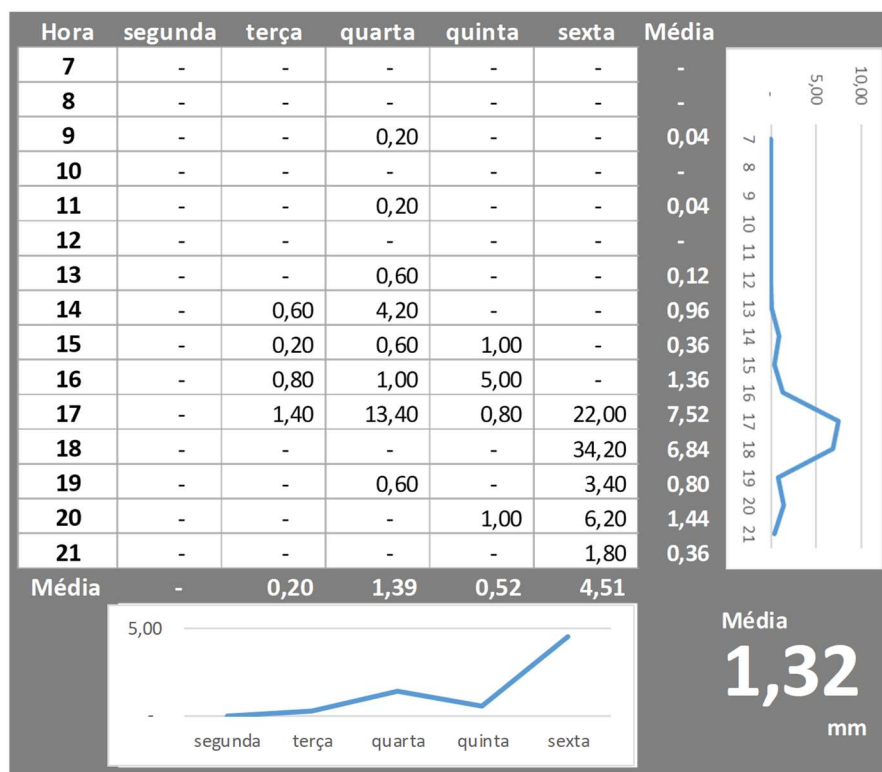
Um dos objetivos desse trabalho é quantificar a variação frente a eventos de precipitação de chuva – comparação dos padrões na ausência e em diferentes escalas de precipitação, dentro do apresentado no período de estudo.

A compreensão de possíveis relações estatísticas entre a variabilidade de velocidades e distâncias, em comparação com a variabilidade da intensidade de precipitação, requer detalhada caracterização de todas as variáveis aqui apresentadas.

A observação da distribuição da intensidade horaria de precipitação no mês de março de 2014, agrupada por dias úteis, revela precipitação fraca (até 10mm) nas segundas e terças (1 e 2) e eventos de precipitação muito fortes (acima de 30mm) nas sextas (5) (Gráfico 4.1.6-2).

O número de registros com presença de precipitação maior que zero corresponde a aproximadamente 9% das microviagens.

Gráfico 4.1.6-2 – Eventos de precipitação (mm) distribuídos ao longo dos dias de semana de março de 2014



A observação dos mesmos eventos, agrupados pelo acumulado nos horários diários, mostra que as chuvas ocorreram com maior frequência nos períodos vespertinos. Destaque para os eventos classificados como fortes, ocorridos próximo ao pico da tarde, a partir das 16:00 horas.

## 4.2 ESTATÍSTICAS DESCRITIVAS

A variabilidade dos tempos de viagem, velocidades e distâncias são quantificadas por diversos autores através de estatísticas descritivas.

Tabela 4.2-1 – Estatísticas descritivas - geral

|                   | Estatísticas gerais - ausência de precipitação |              |       |         |                  |                  |           |               |                          |
|-------------------|--|--------------|-------|---------|------------------|------------------|-----------|---------------|--------------------------|
|                   | N  | % Valid obs. | Média | Mediana | Quartil inferior | Quartil superior | Variância | Desvio Padrão | Coefficiente de Variação |
| Velocidade (km/h) | 17712505                                       | 100          | 17,4  | 14,1    | 8,3              | 22,8             | 161,3     | 12,7          | 73,0                     |

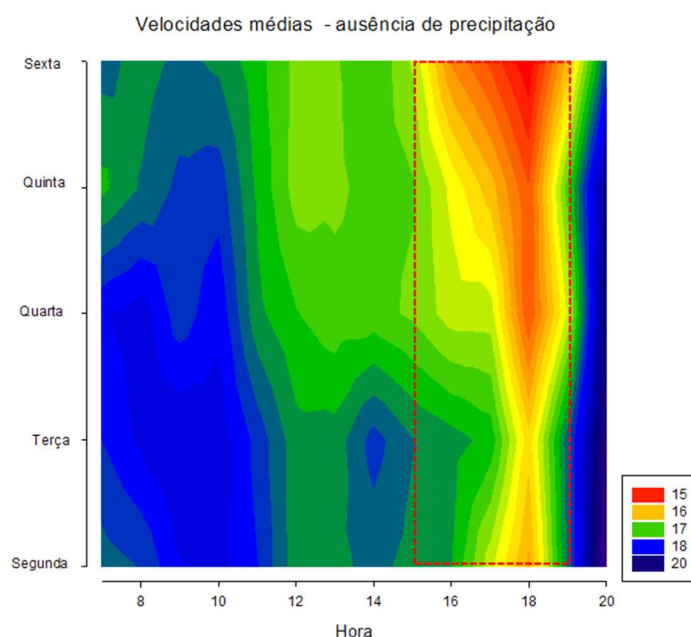
|                   | Estatísticas gerais - presença de precipitação |              |       |         |                  |                  |           |               |                          |
|-------------------|--|--------------|-------|---------|------------------|------------------|-----------|---------------|--------------------------|
|                   | N  | % Valid obs. | Média | Mediana | Quartil inferior | Quartil superior | Variância | Desvio Padrão | Coefficiente de Variação |
| Velocidade (km/h) | 1668325  | 100          | 16,3  | 13,2    | 7,7              | 21,5             | 144,5     | 12,0          | 73,7                     |

As estatísticas gerais mostram, ver Tabela 4.2-1, ligeiras diferenças na presença e ausência de precipitação.

Há pequenas reduções nas médias gerais e desvio padrão, uma queda perceptível, embora a baixa das variâncias e uma pequena alta do coeficiente de variação em eventos com precipitação.

A análise das velocidades médias por dia e hora (gráfico 4.2-2), na ausência de eventos de precipitação, demonstra um padrão geral de quedas das velocidades no período de pico da tarde. Este padrão é observado por volta das 18:00 horas nas segundas e terças, aumentando progressivamente até a sexta-feira, quando atinge amplitude de aproximadamente cinco horas. As segundas e terças têm as melhores médias de velocidade das microviagens ao longo de todo o período.

Gráfico 4.2-2 – Médias das velocidades (km/h) das microviagens por dia de semana e hora do dia



As médias gerais das velocidades apresentadas em dias chuvosos e não chuvosos não parecem significativas. Na ausência de chuva a média geral da velocidade das microviagens é de 17,4 km/h enquanto na presença de chuva, a média geral é de 16,3 km/h (Tabela 4.2-1). Os desvios padrão tem variação mínima, bem como o coeficiente de variação. Para as distâncias euclidianas, eventos sem precipitação apresentaram 308,5 metros de média frente a 307,2 metros em eventos de precipitação.

Para os táxis da cidade de Nova Iorque as médias dos tempos de viagem seguem padrões gerais similares aos desvios padrão. Há um pico das médias dos tempos de viagem a partir das 7:00 horas da manhã. Para o desvio padrão esse pico é observado a partir das 8:00 horas da manhã. Durante os picos, após o almoço, a situação oposta é observada. Os desvios padrão, em geral, caem a partir das 18:00 horas, porém as médias sustentam valores de tempo de viagem altos até as 19:00 horas. Esses períodos são vistos como formação e dispersão de congestionamentos. Desta forma, os desvios padrão atingem seus valores mais altos depois que o congestionamento é acumulado e não começa a diminuir até que esse se dissipe.

Os valores do coeficiente de variação exibem um padrão oposto ao das médias e dos desvios padrão. Entre 6:00 horas e 10:00 horas da manhã (períodos de grande atividade), os valores de coeficiente de variação são baixos, comparados às horas de atividade menos significativa.

A previsibilidade nos tempos de viagem sugere que o coeficiente de variação seja um bom indicador de confiabilidade. Considerando que um baixo coeficiente de variação implica em alta confiabilidade, tempos de viagem altos com altos desvios padrão possuem, contra intuitivamente, níveis mais altos de confiabilidade do tempo de viagem (YAZICI; KAMGA; SINGHAL, 2013). Os autores utilizam o termo “contra intuitivamente” referindo-se a algo oposto ao senso comum.

Neste contexto, maiores tempos de viagem correspondem, de forma análoga, a menores velocidades ou menores distâncias euclidianas percorridas em intervalos de tempo regulares.

Franklin e Karlstrom (2009) observam, através das variações dos desvios padrão, que o aumento unitário nas médias de tempos de viagem está associado a um maior aumento na variabilidade, durante condições de trânsito livre.

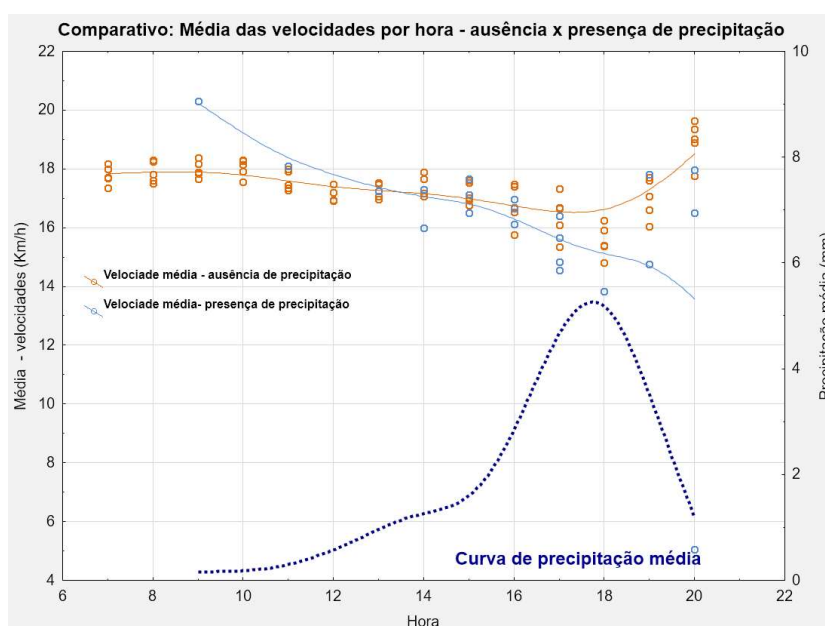
Picos da tarde possuem maior demanda que os da manhã, com menor variabilidade nos períodos de trânsito mais intenso (KWON et al., 2011).

O clima inclemente impacta diretamente velocidades e taxas de fluxo, com reduções das velocidades em torno de 12% e do fluxo das vias em torno de 7%. Tais observações e conclusões referem-se a dados colhidos em ambiente urbano para táxis e veículos em geral (AKIN; SISIPIKU; SKABARDONIS, 2011).

A análise dos dados obtidos, através dos ônibus de transporte da cidade de São Paulo, mostra um comportamento distinto nas médias de velocidades em situações de precipitação e ausência de precipitação hora a hora. As médias das velocidades em situações de ausência de chuva se mantêm uniformes ao longo do dia, com queda observada nos horários de pico de trânsito vespertino-noturno, entre 16:00 e 19:00 horas.

Na presença de precipitação há uma tendência de queda das velocidades, aparentemente relacionada à intensidade da precipitação. No entanto, frente à precipitação média fraca no período da manhã observam-se velocidades significativamente elevadas em relação a ausência de precipitação, mantendo-se até o início do período da tarde. Aparentemente, há uma forte interação entre a precipitação e os horários de pico da tarde, resultando em padrões de microviagens significativamente distintos (Gráfico 4.2-3).

**Gráfico 4.2-3: Médias das velocidades das microviagens: chuva x seco x curva de precipitação**



Neste contexto, confirma-se uma queda nas velocidades em situação de tempo inclemente, conforme observado por Akin, Sisiopiku & Skabardonis (2011) para veículos em geral. É possível, ainda que indiretamente, confirmar as tendências de aumento de demanda no período vespertino observadas por Kwon (KWON et al., 2011) e localmente confirmadas por Peron (2015)

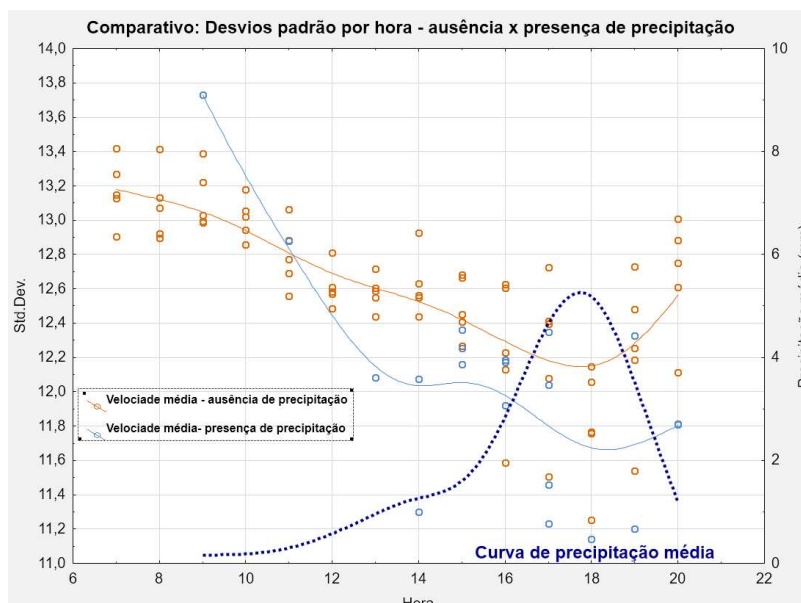
Os valores de desvio padrão das velocidades das microviagens (Gráfico 4.2-4) mostram uma progressiva queda da variabilidade ao longo do dia, com intensificação do fenômeno no pico do fim da tarde.

Chuva muito fraca ou os momentos que antecedem seu início, combinados a horários de menor demanda de transportes (manhã), parecem propiciar um ambiente favorável a uma alta variabilidade das velocidades das microviagens.

No entanto, no início da tarde a chuva passa a ter um efeito oposto, reduzindo progressivamente a variabilidade, de forma significativamente acentuada em relação a ausência de precipitação. A curva de queda dos desvios padrão das velocidades das microviagens apresenta padrão similar, porém inverso ao da precipitação média. No período de pico da tarde, a partir das 16:00 horas, há uma assincronia entre desvios e a curva de precipitação, sugerindo a formação de congestionamentos mais intensos, a partir de eventos de precipitação.



Gráfico 4.2-4: Desvios padrão das velocidades das microviagens: chuva x seco x curva de precipitação



A deterioração das condições de tráfego, seja nos horários de pico, seja frente a condições climáticas adversas ou a combinação de ambos, produz um efeito de aumento do coeficiente de variação das velocidades dos ônibus de transporte da cidade de São Paulo.

Em contrapartida, os táxis de Nova Iorque, em condições semelhantes as supracitadas de degradação do tráfego, apresentam uma queda significativa no coeficiente de variação. Este fato, segundo os autores, permite uma maior previsibilidade dos tempos de viagem nestas condições. Em termos práticos, os autores afirmam que diante de uma condição degradada de tráfego, todos os veículos andaram mais lentamente, em um ritmo mais uniforme, tornando portanto, seu tempo de deslocamento mais previsível, sempre abaixo da média (KAMGA; YAZICI, 2014).

A fórmula clássica de velocidade média é expressa como:

$$Vm = \frac{\Delta s}{\Delta t}$$

Onde:

$V_m$  : velocidade média;

$\Delta s$ : distância;

$\Delta t$  : tempo.

Assim, tempos de viagem e velocidades médias são grandezas que representam de modo diferente o mesmo fenômeno.

O coeficiente de variação é expresso pela razão entre o desvio padrão e a média:

$$Cv = \frac{\sigma}{\mu}$$

Onde:

Cv: Coeficiente de variação

$\sigma$ : Desvio padrão

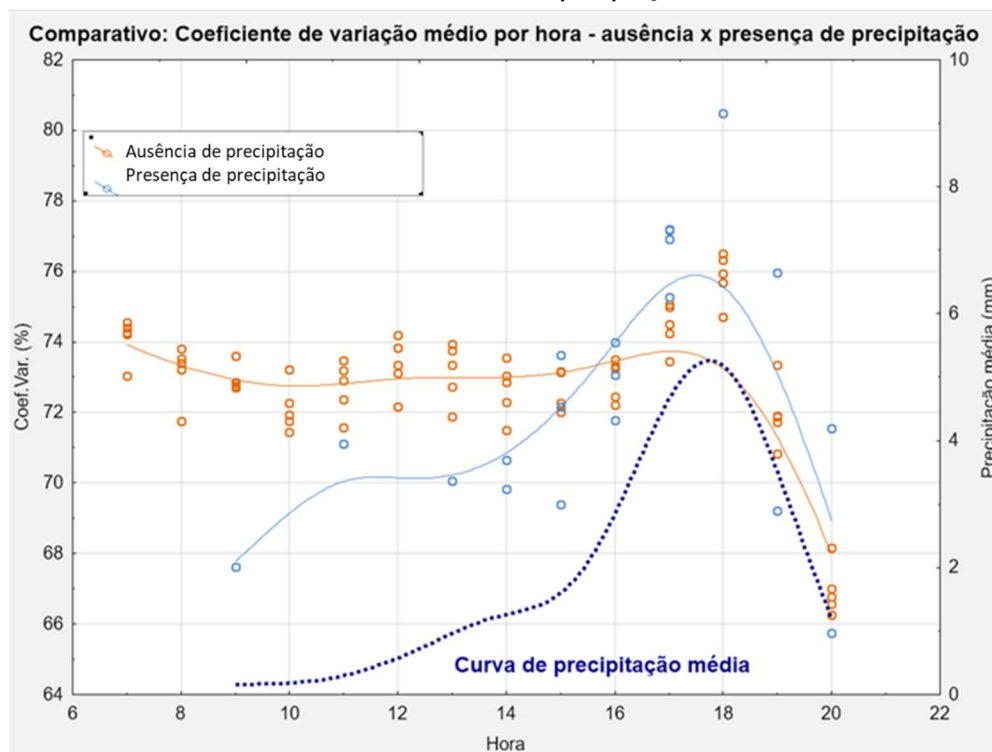
$\mu$ : Média

O desvio padrão é uma medida de dispersão de dados em torno da média amostral.

Assim, o desvio padrão das velocidades, aqui observado, representa uma queda nas mesmas frente a horários de pico, precipitação ou a combinação de ambos. As médias das velocidades mostram a mesma tendência de queda em qualquer situação até as 18:00 horas.

Enquanto no caso de São Paulo as velocidades nos eventos supracitados causadores de variabilidade caem abaixo média, para os táxis de Nova Iorque os tempos de viagem estabelecem-se acima da média. Assim o aumento do coeficiente de variabilidade (gráfico 4.2-5) das velocidades nos ônibus de São Paulo é um evento semelhante a queda dos coeficientes de variabilidade dos táxis de Nova Iorque, sendo portanto uma importante medida de aumento da previsibilidade das velocidades para os ônibus de São Paulo.

Gráfico 4.2-5: Coeficiente de variação das velocidades das microviagens:  
chuva x seco x curva de precipitação



### 4.3 ANÁLISE DOS PADRÕES DAS MICROVIAGENS ATRAVÉS DO MÉTODO CHAID

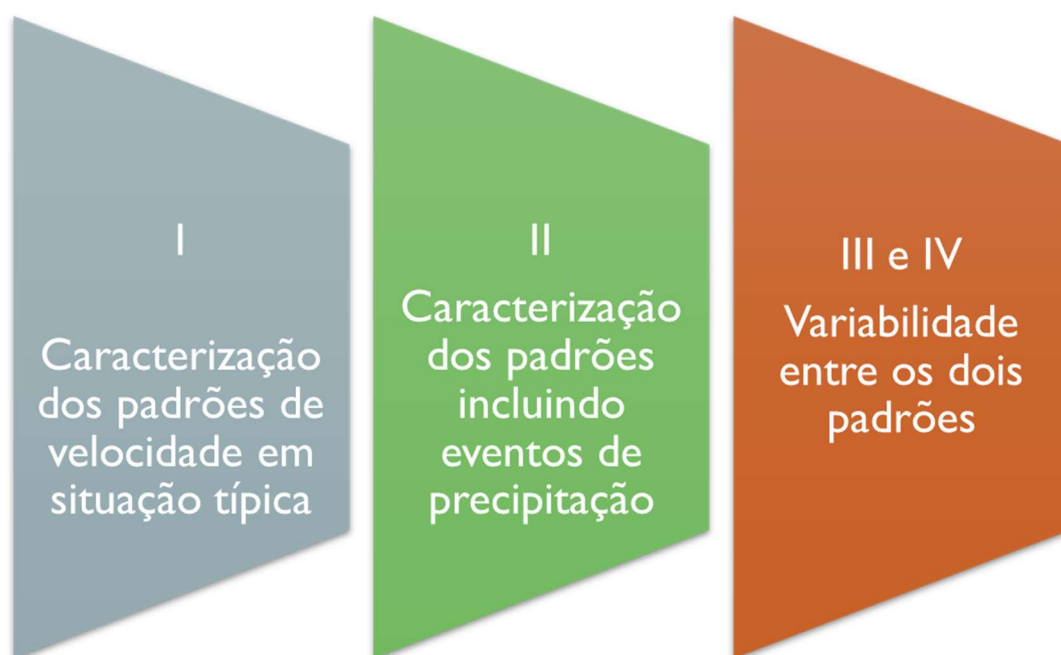
A técnica CHAID (Chi-square Automatic Interaction Detector) (KASS, 1980) foi desenvolvida a partir de métodos pré-existentes de interação e classificação de dados (MORGAN; SONQUIST, 1963). A técnica permite a identificação e quantificação de relações entre variáveis contínuas e categóricas. Assim, a partir de um modelo preditivo, é possível determinar qual combinação de variáveis pode melhor explicar um fenômeno. Na prática, um preditor é estabelecido para representar uma classe real de registros. Além da classificação, o algoritmo disponível no software Dell Statistica fornece a representatividade das classes e sua variância.

Esse capítulo divide-se em quatro partes:

- (I) Na primeira a técnica é empregada na análise dos padrões de velocidade frente à ausência de precipitação, cujo objetivo é a caracterização geral da variabilidade diária e horária.

- (II) Na segunda parte, a técnica é empregada na análise dos padrões de velocidade frente à presença de precipitação.
- (III) Na terceira parte os resultados individuais nas situações de ausência e presença de precipitação são comparados conforme classes horárias produzidas.
- (IV) Na quarta parte a técnica é novamente empregada, desta vez, na totalidade dos registros. O objetivo é estabelecer quais são os vetores predominantes da variabilidade, considerando dias úteis semanais, progressão horária, na presença e ausência de precipitação.

Gráfico 4.3-1: Método de aplicação CHAID



Estas análises resultam em diagramas dispostos em árvore, hierarquicamente distribuído em classes, onde é possível analisar de forma simplificada um número significativo de vetores associados a uma determinada variável.

Na medida que outros fatores são incluídos, as classes passam a representar interações significativas entre as variáveis, aglutinando classes semelhantes.

### **4.3.1 Caracterização dos padrões de velocidade através da técnica CHAID**

A construção das árvores de regressão (gráfico 4.3.1-1) e a tabela analítica processada, a partir de seus resultados (gráfico 4.3.1-2), retrata as características gerais de operação dos ônibus de transporte da cidade de São Paulo, com base nas leituras de velocidades durante o mês de março de 2014.

Conforme observado através das médias das velocidades, confirma-se a redução sistemática nos índices de velocidade no período vespertino. A variabilidade entre os períodos é baixa às segundas (2%) e com maior valor observado às sextas (5%). Estes percentuais, no entanto, não refletem a rápida degradação das velocidades, especialmente entre 16:00 horas e 18:00 horas. A observação da variabilidade hora a hora, no período supracitado, apresenta uma redução acumulada entre 10% nas segundas a 12% nas sextas. Van Lint; Van Zuylen e Tu (2008) afirmam que, em rodovias, há variação diária, ainda que discreta, dos tempos de viagem.

Gráfico 4.3.1-1 – Caracterização geral das velocidades no mês de março de 2014 através do método CHAID  
(vermelho: variabilidade hora a hora)

|       | Dia da semana | Hora        | N       | Velocidade CHAID (km/h) | Média (Km/h) por período | Média (Km/h) diária | Variabilidade hora a hora |
|-------|---------------|-------------|---------|-------------------------|--------------------------|---------------------|---------------------------|
| Manha | Segunda       | 7           | 685.655 | 17,69                   |                          |                     |                           |
|       |               | 8           | 391.997 | 17,81                   |                          |                     | 1%                        |
|       |               | 9           | 369.559 | 18,19                   |                          |                     | 2%                        |
|       |               | 10          | 324.275 | 18,27                   | ▲                        | 17,91               | 0%                        |
|       |               | 11          | 307.400 | 18,00                   |                          |                     | -2%                       |
|       |               | 12          | 310.392 | 17,47                   |                          |                     | -3%                       |
|       |               | 13 e 15     | 643.799 | 17,53                   | ▲                        | 17,68               | 0%                        |
|       |               | 16          | 351.082 | 17,40                   |                          |                     | -1%                       |
|       |               | 17          | 358.821 | 16,66                   | ▼                        | 17,46               | -4%                       |
|       |               | 18          | 351.916 | 15,93                   |                          |                     | -5%                       |
| Tarde | Terça         | 19          | 343.625 | 17,62                   |                          |                     | 10%                       |
|       |               | 20          | 302.488 | 19,63                   | ▲                        |                     | 10%                       |
|       |               | 7           | 260.868 | 18,00                   |                          |                     | 1%                        |
|       |               | 8           | 279.570 | 18,24                   | ▲                        | 18,12               | 1%                        |
|       |               | 9 e 10      | 481.432 | 18,35                   | ▲                        |                     |                           |
|       |               | 11          | 379.868 | 17,90                   |                          |                     | -2%                       |
|       |               | 12, 13 e 14 | 614.969 | 17,48                   | ▲                        | 17,82               | -2%                       |
|       |               | 15 e 16     | 162.893 | 17,61                   |                          |                     | 1%                        |
|       |               | 17          | 178.779 | 17,33                   | ▼                        | 17,62               | -2%                       |
|       |               | 18          | 244.094 | 16,26                   |                          |                     | -7%                       |
| Manha | Quarta        | 19          | 240.482 | 17,70                   |                          |                     | 8%                        |
|       |               | 20          | 211.721 | 19,35                   | ▲                        |                     | 9%                        |
|       |               | 7 e 10      | 553.419 | 18,16                   | ▲                        |                     |                           |
|       |               | 8           | 313.159 | 18,30                   | ▲                        | 17,80               | 1%                        |
|       |               | 9           | 221.240 | 17,88                   | ▲                        |                     | -2%                       |
|       |               | 11          | 185.539 | 17,45                   |                          |                     | -2%                       |
|       |               | 12          | 247.327 | 17,21                   |                          |                     | -1%                       |
|       |               | 13, 14 e 15 | 384.748 | 17,08                   | ▲                        | 17,34               | -1%                       |
|       |               | 16          | 194.029 | 16,96                   | ▼                        | 16,96               | -1%                       |
|       |               | 17          | 347.506 | 16,70                   |                          |                     | -2%                       |
| Tarde | Quinta        | 18          | 274.665 | 15,38                   |                          |                     | -9%                       |
|       |               | 19          | 204.914 | 16,62                   |                          |                     | 7%                        |
|       |               | 20          | 238.684 | 19,03                   | ▲                        |                     | 13%                       |
|       |               | 7 e 11      | 544.368 | 17,35                   | ▲                        | 17,44               | 2%                        |
|       |               | 8           | 318.948 | 17,61                   | ▲                        |                     | 1%                        |
|       |               | 9 e 10      | 547.533 | 17,88                   | ▲                        |                     |                           |
|       |               | 12 e 13     | 490.734 | 16,93                   |                          |                     | -6%                       |
|       |               | 14, 15 e 19 | 714.733 | 17,06                   | ▲                        | 17,08               | 1%                        |
|       |               | 16          | 138.628 | 16,54                   | ▼                        | 16,80               | -3%                       |
|       |               | 17          | 220.414 | 16,09                   |                          |                     | -3%                       |
| Manha | Sexta         | 18          | 281.131 | 15,40                   |                          |                     | -4%                       |
|       |               | 20          | 176.558 | 18,89                   | ▲                        |                     | 18%                       |
|       |               | 7 e 9       | 596.199 | 17,67                   | ▲                        | 17,4                | -1%                       |
|       |               | 8 e 10      | 580.493 | 17,53                   | ▲                        |                     | -1%                       |
|       |               | 11          | 246.463 | 17,27                   |                          |                     | -2%                       |
|       |               | 12 e 13     | 505.370 | 16,96                   |                          |                     | 1%                        |
|       |               | 14          | 252.216 | 17,18                   |                          |                     | -2%                       |
|       |               | 15          | 267.330 | 16,77                   | ▼                        | 16,6                | -2%                       |
|       |               | 16          | 293.236 | 15,77                   |                          |                     | -6%                       |
|       |               | 17          | 101.690 | 15,34                   |                          |                     | -3%                       |
| Tarde | Sexta         | 18          | 168.763 | 14,81                   |                          |                     | -4%                       |
|       |               | 19          | 162.251 | 16,06                   |                          |                     | 8%                        |
|       |               | 20          | 144.432 | 17,77                   | ▲                        |                     | 10%                       |

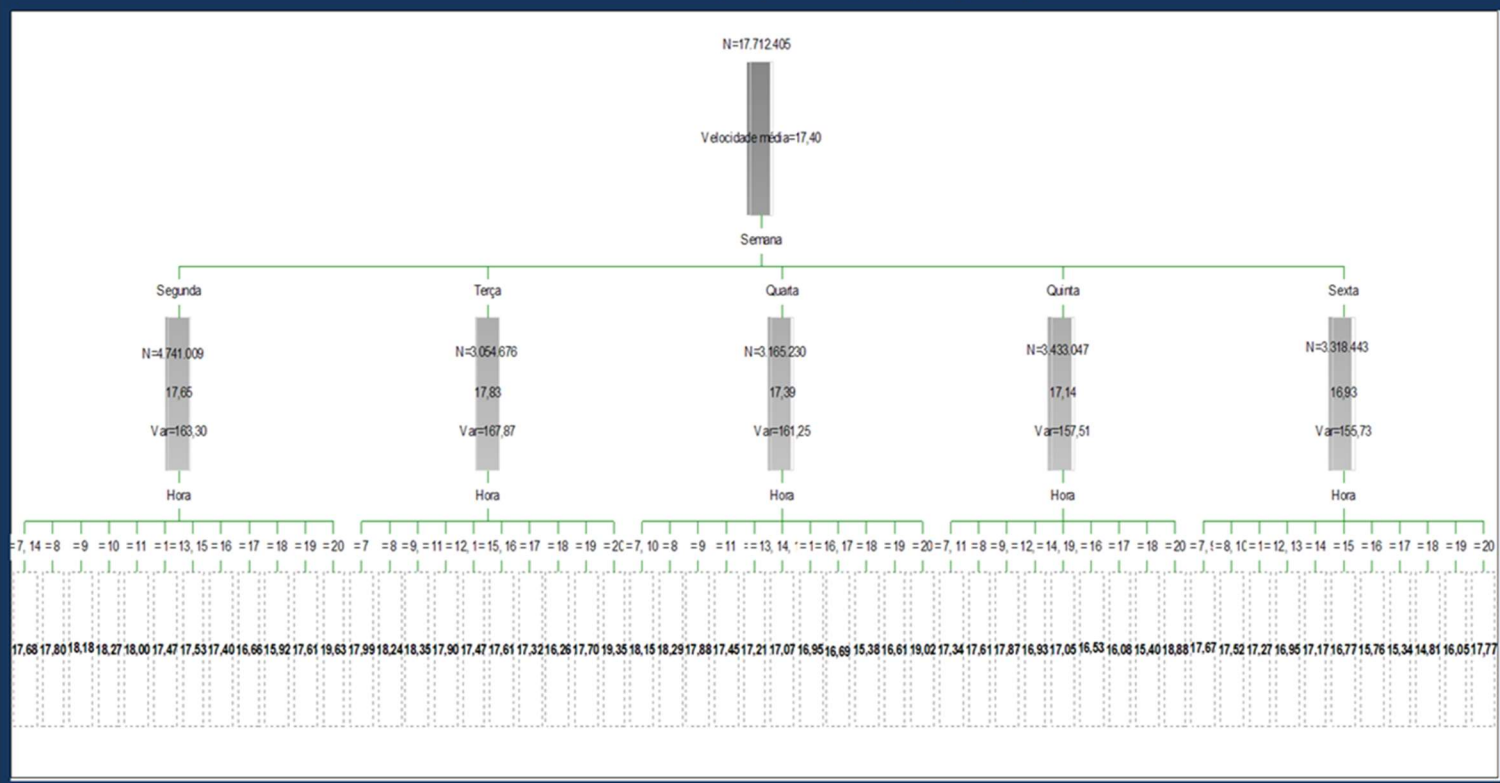
Enquanto a aparente degradação das velocidades se dá ao longo de 3 horas no período da tarde, a uma taxa acumulada de cerca de 11%, a partir das 19:00 horas é observada uma abrupta inversão da variabilidade. A elevação de média é de 8% nas velocidades a partir das 19:00 horas, seguida por uma elevação média de 12% as 20:00 horas.

Neste contexto, a observação de Van Lint e outros não parece aplicável aos ônibus de transporte da cidade de São Paulo. Aparentemente, veículos trafegando entre 16:00 horas e 18:00 horas apresentam velocidades semelhantes (distribuição simétrica), provavelmente devido ao tráfego intenso. É possível ainda especular que as velocidades nos períodos pós 18:00 são predominantemente variadas (distribuição assimétrica), devido à queda na demanda de uso das vias do transporte coletivo.

Para que se estabeleçam conclusões definitivas é preciso um aprofundamento sobre as análises conforme se apresentará adiante.

Gráfico 4.3.1-2 – Caracterização geral das velocidades no mês de março (precipitação=0) de 2014 através do método CHAID:

### Diagrama em árvore: Velocidade (Km/h) Velocidades em situação padrão (precipitação=0)



### **4.3.2 Caracterização dos padrões de velocidade em situação de precipitação através da técnica CHAID**

Cerca de 1,7 milhões de registros, referentes a eventos com presença de precipitação foram processados. Como resultado foram criadas classes de velocidade, representando horas do dia e dias da semana (gráficos 4.3.2-1 e 4.3.2-2).

Do ponto de vista geral, as classes são um reflexo da precipitação e portanto são limitadas a estes eventos. Coincidentemente houve a combinação de feriado e precipitação em 4 e 5 de março de 2014.

A classificação das velocidades nas terças, diante de precipitação muito leve (em torno de 0,40 mm), apresenta uma melhora das velocidades da ordem de 3% entre o meio do dia e o período da tarde, até as 16:00 horas. As 17:00 horas com precipitação ainda leve (em torno de 1,40 mm) há uma redução das velocidades da ordem de 5% em relação a leitura anterior.

Os resultados referentes a quarta-feira (dia 5), parecem refletir os efeitos do feriado de cinzas. O período próximo ao meio dia apresenta uma redução de 11% da velocidade em relação ao início da manhã, possivelmente em função das características do feriado, onde várias organizações retomam a rotina operacional a partir do período da tarde.

O mesmo ocorre durante o avanço da tarde, quando por volta das 16:00 horas há uma redução de 2% em relação ao período anterior, se consolidando subsequentemente, às 17:00 horas, em uma redução de 13% em relação ao período supracitado. Essa redução nas velocidades tem sob contexto uma precipitação mais acentuada, embora moderada, no período (13,40 mm) associada ao possível aumento da demanda sobre as vias e sobre os serviços no fim do feriado.

Nas quintas-feiras, sob precipitação de baixa escala, observa-se uma redução moderada das velocidades, entre 16:00 horas e 17:00 horas, da ordem de 3% e 2% - respectivamente. De forma similar ao observado em períodos secos, uma melhora acentuada, da ordem de 9%, é registrada às 20:00 horas.

Durante as sextas-feiras, há uma queda de 5% entre 17:00 horas e 18:00 horas, acompanhada por melhora abrupta registrada as 19:00 horas (7%) e 20:00 horas (12%). Embora já observado, o cálculo da variabilidade aqui descrito é relativo ao



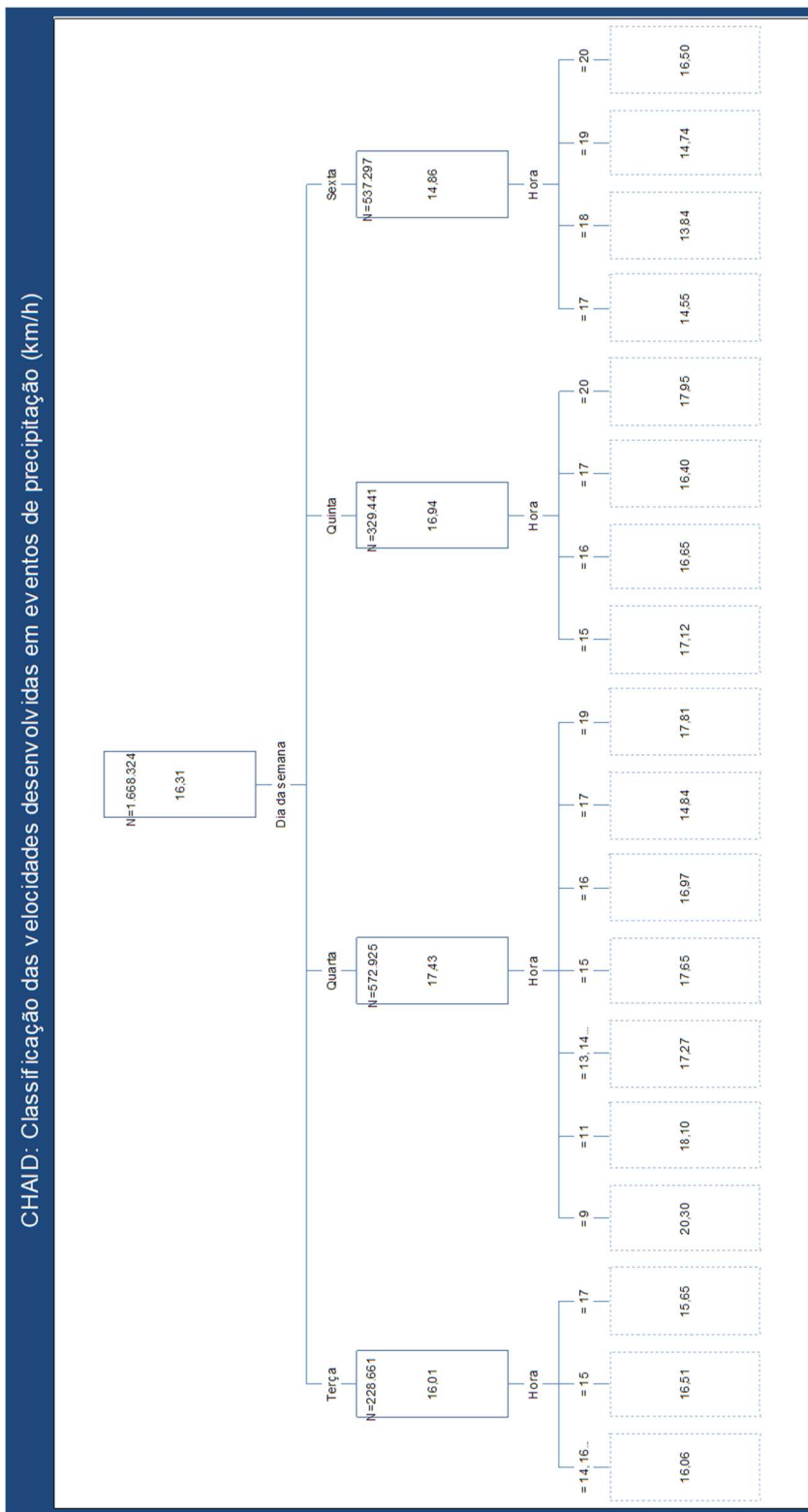
horário anterior, sendo, portanto, uma variação relativamente acentuada frente ao curto período.

Embora haja forte correlação entre as curvas de precipitação e variabilidade das velocidades, não foi possível estabelecer uma relação direta entre volume de precipitação e queda ou aumento nas velocidades. É possível que uma análise com um histórico mais significativo possa fornecer insumos para a identificação de padrões de variabilidade em velocidade, provocados por um volume específico de precipitação.

Gráfico 4.3.2-1 – CHAID: Variabilidade em eventos de precipitação

|       | Dia da semana                                  | Hora        | Velocidades (Km/h): P>0 | Variabilidade | Média diária (Km/h) | Precipitação média (mm) |
|-------|--|-------------|-------------------------|---------------|---------------------|-------------------------|
| Tarde | <b>Terça</b> (ocorrência de feriado: Carnaval) | 12, 13 e 14 | 16,05                   |               | 16,07               | 0,60                    |
|       |  | 15 e 16     | 16,51                   | 3%            |                     | 0,40                    |
|       |  | 17          | 15,65                   | -5%           |                     | 1,40                    |
| Manhã | <b>Quarta</b> (ocorrência de feriado: Cinzas)  | 9           | 20,3                    |               | 17,55               | 0,20                    |
| 11    |  | 18,1        | -11%                    | 0,20          |                     |                         |
| Tarde |  | 13, 14 e 15 | 17,27                   |               |                     | 1,80                    |
|       |  | 16          | 16,97                   | -2%           |                     | 0,50                    |
|       | 17   | 14,84       | -13%                    | 13,40         |                     |                         |
|       |  | 19          | 17,81                   | 20%           | 0,60                |                         |
| Tarde | <b>Quinta</b>                                  | 14, 15 e 19 | 17,13                   |               | 17,03               | 1,00                    |
|       |  | 16          | 16,65                   | -3%           |                     | 2,40                    |
|       |  | 17          | 16,4                    | -2%           |                     | 0,80                    |
|       |  | 20          | 17,95                   | 9%            |                     | 1,00                    |
| Tarde | <b>Sexta</b>                                   | 17          | 14,55                   |               | 14,91               | 6,60                    |
|       |  | 18          | 13,84                   | -5%           |                     | 13,00                   |
|       |  | 19          | 14,74                   | 7%            |                     | 1,50                    |
|       |  | 20          | 16,5                    | 12%           |                     | 2,70                    |

Gráfico 4.3.2-2 – CHAID: Diagrama de árvore - velocidade sob precipitação



### 4.3.3 CHAID – Comparação de resultados: Precipitação = 0 versus Precipitação >0

Este item faz uma comparação direta das velocidades observadas em eventos sem precipitação (= 0) e com precipitação (>0). Neste contexto, os casos foram processados e classificados de forma completamente independente, sem interação entre si ou demais variáveis, como por exemplo, volume de precipitação. Os casos referentes às segundas-feiras, onde não há ocorrência de precipitação, não foram mencionados (gráfico 4.3.3-1).

A variabilidade horária das velocidades as terças feiras é, em média, 8% (1,40 km/h) superior na ausência de precipitação, com maior índice observado as 17:00 horas (10%). Não há precipitação e, portanto, base de comparação no horário das 18:00 horas.

Na quarta-feira, os resultados demonstram velocidades significativamente superiores, da ordem de 14% (2,42%) na presença de precipitação. Essa tendência permanece até as 16:00 horas, com abrupta inversão no horário das 17:00 horas, onde dias secos apresentaram velocidades 11% superiores. Conforme observado anteriormente, os eventos referentes a esse dia estão sob a influência do feriado de carnaval.

Nas quintas-feiras, com exceção das 20:00 horas, há uma variabilidade recorrentemente superior, porém de baixa escala, em eventos com precipitação. A inversão dessa tendência as 20:00 horas (5%) resultou em uma média geral das velocidades de apenas 1% superior na ausência de precipitação.

Nas sextas-feiras, as velocidades na ausência de precipitação são recorrentemente superiores, com variabilidade média de 7% (1,09 km/h).

Embora pontualmente haja variações, com exceção das quartas-feiras, as médias demonstram velocidades superiores na ausência de precipitação. Esse resultado comparativo produzido pelo CHAID é semelhante ao observado nas médias gerais observadas.

Assim como o observado nas médias, não houve interação e comparação direta entre as classes e volumes de chuva, o que provavelmente reduziu a

importância dos eventos extremos e sua representatividade, conforme será descrito no item a seguir.

Gráfico 4.3.3-1 – CHAID: Variabilidade comparativa independente

| Dia da semana                              | Hora        | Velocidades: P=0 | Velocidades: P>0 | Variabilidade (base P=0)% | Variabilidade (Km/h) |
|--|-------------|------------------|------------------|---------------------------|----------------------|
| Terça<br>(ocorrência de feriado: Carnaval) | 12, 13 e 14 | ▲ 17,48          | ▼ 16,05          | 8%                        | 1,43                 |
|  | 15 e 16     | ▲ 17,61          | ▬ 16,51          | 6%                        | 1,10                 |
|  | 17          | ▲ 17,33          | ▼ 15,65          | 10%                       | 1,68                 |
|  |             | ▲ <b>17,47</b>   | ▼ <b>16,07</b>   | 8%                        | 1,40                 |
| Quarta<br>(ocorrência de feriado: Cinzas)  | 9           | ▬ 17,88          | ▲ 20,3           | -14%                      | -2,42                |
|  | 11          | ▬ 17,45          | ▬ 18,1           | -4%                       | -0,65                |
|  | 13, 14 e 15 | ▬ 17,08          | ▬ 17,27          | -1%                       | -0,19                |
|  | 16          | ▬ 16,96          | ▬ 16,97          | 0%                        | -0,01                |
|  | 17          | ▬ 16,70          | ▼ 14,84          | 11%                       | 1,86                 |
|  | 19          | ▼ 16,62          | ▬ 17,81          | -7%                       | -1,19                |
|  |             | ▼ <b>17,11</b>   | ▲ <b>17,55</b>   | -3%                       | -0,44                |
| Quinta                                     | 14, 15 e 19 | ▬ 17,06          | ▬ 17,13          | 0%                        | -0,07                |
|  | 16          | ▼ 16,54          | ▼ 16,65          | -1%                       | -0,11                |
|  | 17          | ▼ 16,09          | ▼ 16,4           | -2%                       | -0,31                |
|  | 20          | ▲ 18,89          | ▬ 17,95          | 5%                        | 0,94                 |
|  |             | ▲ <b>17,14</b>   | ▼ <b>17,03</b>   | 1%                        | 0,11                 |
| Sexta                                      | 17          | ▼ 15,34          | ▼ 14,55          | 5%                        | 0,79                 |
|  | 18          | ▼ 14,81          | ▼ 13,84          | 7%                        | 0,97                 |
|  | 19          | ▬ 16,06          | ▼ 14,74          | 8%                        | 1,32                 |
|  | 20          | ▲ 17,77          | ▲ 16,5           | 7%                        | 1,27                 |
|  |             | ▲ <b>16,00</b>   | ▼ <b>14,91</b>   | ▲ 7%                      | 1,09                 |

#### 4.3.4 CHAID – Resultado do processamento conjunto de dados com e sem ocorrência de precipitação

Neste item os dados foram processados de forma integrada, independentemente da ausência ou presença de precipitação. Todos os casos foram comparados entre si e o volume de precipitação, dia de semana e hora do dia, foram considerados como possíveis vetores da variabilidade das velocidades.

A classificação obtida através do método CHAID apresenta uma predominância da precipitação como vetor de redução das velocidades, quando comparada à variação horária diária (Gráficos 4.3.4-1 e 4.3.4-2).

Nas segundas-feiras, quando não há precipitação, o resultado da classificação é semelhante, porém não idêntico ao processamento isolado do conjunto de dados referente à ausência de precipitação, com variabilidade diferindo entre si em aproximadamente 1% nos horários das 17:00 horas, 18:00 horas e 19:00 horas. Em termos de representatividade, os eventos registrados neste dia representam 24% do total dos registros. Ainda na segunda-feira encontra-se a classe de velocidades mais altas, com 19,63 km/h às 20:00 horas.

Nas terças-feiras o vetor predominante é a precipitação, não havendo distinção horária.

Há uma redução de 7% das velocidades entre ausência de precipitação e a ocorrência de precipitação muito fraca, que persiste em menor escala de variabilidade (3%), com intensidade ligeiramente superior de precipitação.

Nas quartas-feiras, a inversão observada na análise individual dos dados de precipitação se repete diante de chuva muito leve, com um aumento da velocidade de 11%. Enquanto volumes baixos de precipitação (entre 2,6 mm e 4,2 mm) apresentam um pequeno aumento da velocidade (1%) em relação ao índice anterior (0,4 mm), a classe seguinte (precipitação > 4,2 mm) apresenta uma queda acentuada das velocidades na casa de 14%. Quanto à representatividade, quarta-feira representa aproximadamente 14% do total de registros.

Com representatividade de aproximadamente 20% dos registros, a quinta-feira apresenta eventos de aumento e diminuição da velocidade, em diferentes medidas de

precipitação (entre 0,8 mm e 1,1mm). Ainda assim a variabilidade acumulada entre as medidas é de redução de 1% das velocidades na presença de precipitação.

Às sextas-feiras apresentam classes específicas, tanto para o vetor hora quanto para o vetor precipitação. Essa caracterização demonstra que os padrões observados são significativamente distintos.

Na ausência de precipitação há pequena variação das classes (em torno de 1%), em relação à análise anterior dedicada exclusivamente a essa categoria de registros. Ainda, conforme observado anteriormente há queda nas velocidades hora a hora, com exceção de uma ligeira melhora (1%) as 14:00 horas.

Em relação à escala de precipitação, com exceção da variação positiva apresentada entre as classes 4,2 mm e 6,4 mm (8%), o padrão geral é de queda nas velocidades diante de eventos de precipitação.

Gráfico 4.3.4-1 – CHAID: Diagrama em árvore – análise integrada de registros

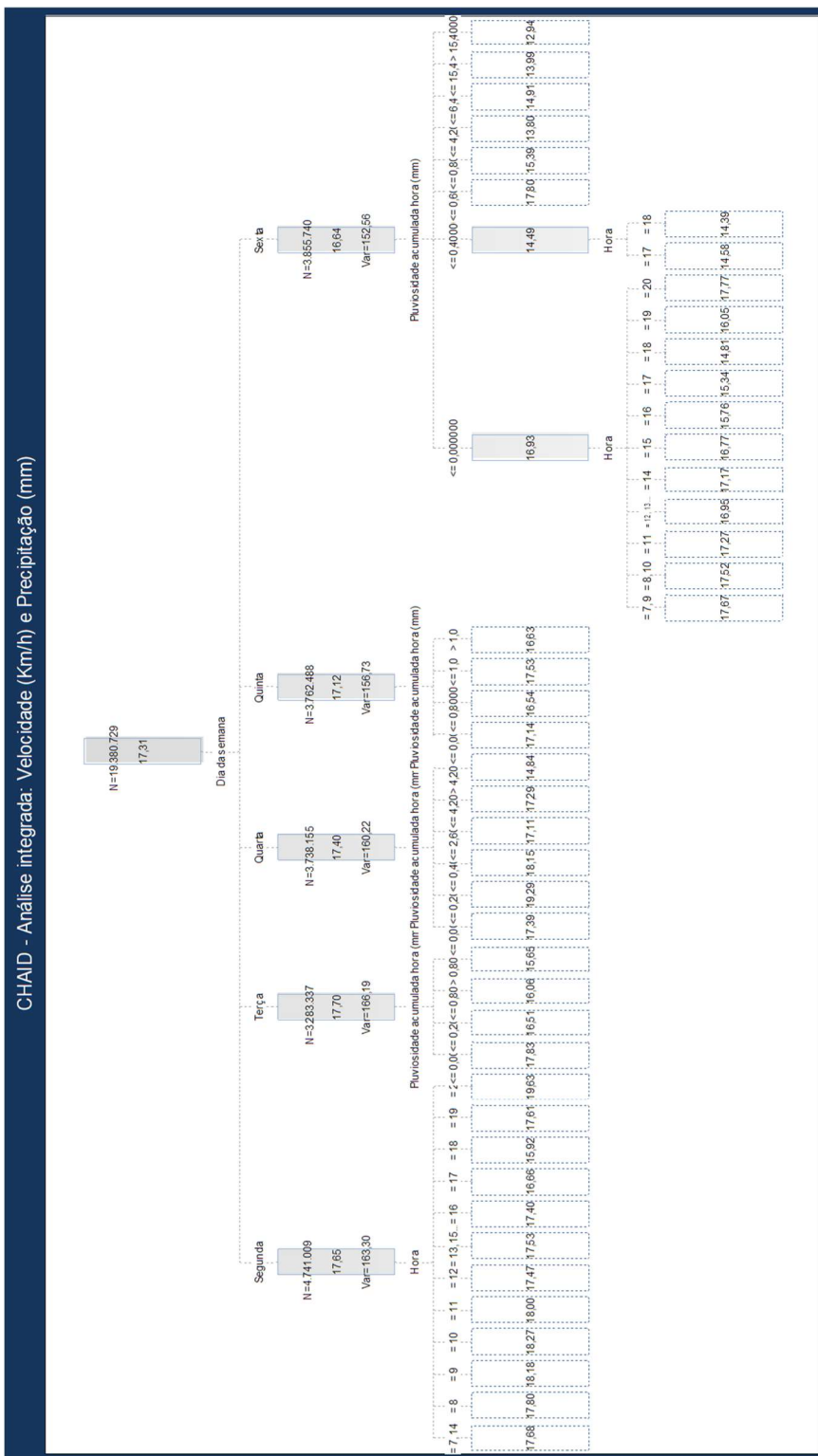


Gráfico 4.3.4-2 – CHAID: Tabela comparativa de variabilidade direta entre classes (vetores)

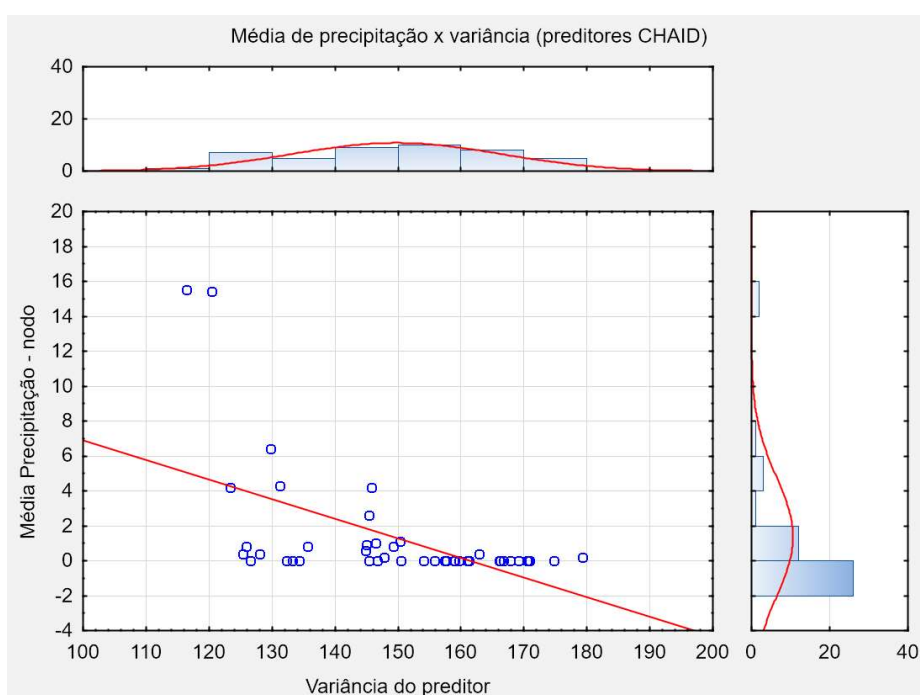
| Semana  | Vetor (hora ou precipitação) | Vetor predominante | Velocidade (Km/h) | Representatividade da classe (n=19MM) | Varição individual | Variabilidade entre casos extremos | Médias       |
|---------|------------------------------|--------------------|-------------------|---------------------------------------|--------------------|------------------------------------|--------------|
| Segunda | 7 e 14                       | Hora               | 17,69             | 3,6%                                  |                    | 34%                                | 17,68        |
|         | 8                            |                    | 17,81             | 2,0%                                  | 1%                 |                                    |              |
|         | 9                            |                    | 18,19             | 1,9%                                  | 2%                 |                                    |              |
|         | 10                           |                    | 18,27             | 1,7%                                  | 0%                 |                                    |              |
|         | 11                           |                    | 18,00             | 1,6%                                  | -1%                |                                    |              |
|         | 12                           |                    | 17,47             | 1,6%                                  | -3%                |                                    |              |
|         | 13 e 15                      |                    | 17,53             | 3,3%                                  | 0%                 |                                    |              |
|         | 16                           |                    | 17,40             | 1,8%                                  | -1%                |                                    |              |
|         | 17                           |                    | 16,66             | 1,9%                                  | -4%                |                                    |              |
|         | 18                           |                    | 15,93             | 1,8%                                  | -4%                |                                    |              |
|         | 19                           |                    | 17,62             | 1,8%                                  | 11%                |                                    |              |
|         | 20                           |                    | <b>19,63</b>      | 1,6%                                  | 11%                |                                    |              |
| Terça   | 0                            | Precipitação       | 17,83             | 15,9%                                 |                    | 34%                                | 16,51        |
|         | 0,2                          |                    | 16,51             | 0,2%                                  | -7%                |                                    |              |
|         | 0,8                          |                    | 16,06             | 0,6%                                  | -3%                |                                    |              |
|         | 0,9                          |                    | 15,65             | 0,4%                                  | -3%                |                                    |              |
| Quarta  | 0                            | Precipitação       | 17,40             | 16,5%                                 |                    | 34%                                | 17,35        |
|         | 0,2                          |                    | 19,29             | 0,6%                                  | 11%                |                                    |              |
|         | 0,4                          |                    | 18,15             | 0,3%                                  | -6%                |                                    |              |
|         | 2,6                          |                    | 17,11             | 1,3%                                  | -6%                |                                    |              |
|         | 4,2                          |                    | 17,29             | 0,3%                                  | 1%                 |                                    |              |
| >4,2    | 14,84                        | 0,4%               | -14%              |                                       |                    |                                    |              |
| Quinta  | 0                            | Precipitação       | 17,14             | 17,8%                                 |                    | 34%                                | 16,96        |
|         | 0,8                          |                    | 16,55             | 0,7%                                  | -3%                |                                    |              |
|         | 1                            |                    | 17,53             | 0,7%                                  | 6%                 |                                    |              |
|         | 1,1                          |                    | 16,64             | 0,3%                                  | -5%                |                                    |              |
| Sexta   | 7 e 9                        | Hora               | 17,67             | 3,1%                                  |                    | 34%                                | 16,65        |
|         | 8 e 10                       |                    | 17,53             | 3,0%                                  | -1%                |                                    |              |
|         | 11                           |                    | 17,27             | 1,3%                                  | -1%                |                                    |              |
|         | 12 e 13                      |                    | 16,96             | 2,6%                                  | -2%                |                                    |              |
|         | 14                           |                    | 17,18             | 1,3%                                  | 1%                 |                                    |              |
|         | 15                           |                    | 16,77             | 1,4%                                  | -2%                |                                    |              |
|         | 16                           |                    | 15,77             | 1,5%                                  | -6%                |                                    |              |
|         | 17                           |                    | 15,34             | 0,5%                                  | -3%                |                                    |              |
|         | 18                           |                    | 14,81             | 0,9%                                  | -3%                |                                    |              |
|         | 19                           |                    | 16,06             | 0,8%                                  | 8%                 |                                    |              |
|         | 20                           |                    | 17,77             | 0,8%                                  | 11%                |                                    |              |
|         | 0,6                          |                    | 17,80             | 0,3%                                  |                    |                                    |              |
|         | 0,8                          |                    | 15,39             | 0,4%                                  | -14%               |                                    |              |
|         | 4,2                          |                    | 13,80             | 0,2%                                  | -10%               |                                    |              |
| 6,4     | 14,91                        | 0,6%               | 8%                |                                       |                    |                                    |              |
| 15,4    | 13,99                        | 0,3%               | -6%               |                                       |                    |                                    |              |
| > 15,4  | <b>12,94</b>                 | 0,2%               | -8%               |                                       |                    |                                    |              |
|         |                              |                    | <b>12,94</b>      | 0,2%                                  |                    |                                    | <b>15,18</b> |



### 4.3.5 CHAID – Breve visão sobre a variância dos preditores

A análise dos índices de variância apresentados nos preditores das classes, produzidos pelo método CHAID, demonstram uma tendência geral de queda da variância frente ao aumento da precipitação. Esses resultados são coerentes com as observações descritivas das médias, desvio padrão e coeficiente de variação.

Gráfico 4.3.5-1- Variância CHAID



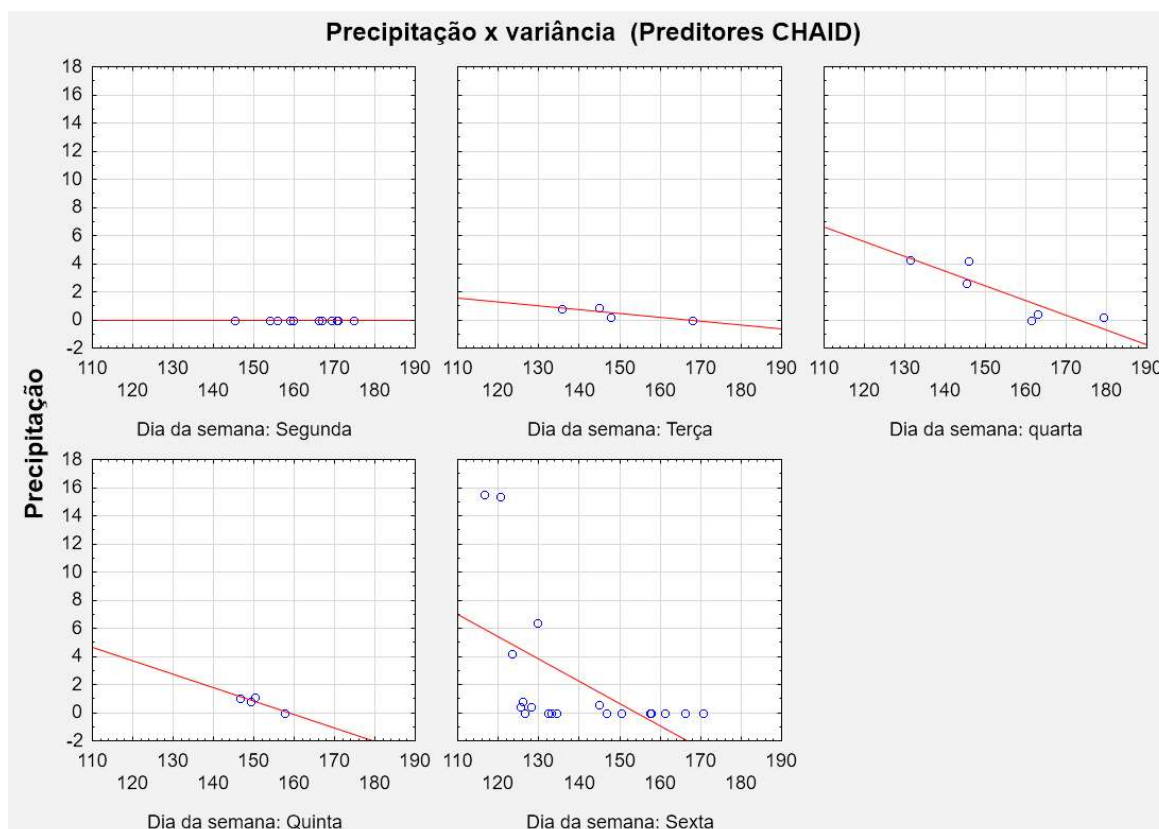
No entanto, as classes com baixa variância não se resumem exclusivamente a eventos de precipitação. Quedas de velocidade em geral produzem efeito semelhante na variância (descritiva), que se manifesta de forma mais evidente através da análise CHAID.

Durante as segundas feiras, na ausência de precipitação, é notada variabilidade inerente à variação diária das condições de trânsito diante do horário.

Frente à chuva muito fraca (terças e quintas) e médias mais significativas de precipitação, que abrangem de chuva fraca à forte (quartas e sextas), ocorre um padrão recorrente de redução da variância, inclusive em casos extremos, conforme

observado no gráfico 4.3.5-1. Em termos gerais, a redução da variância diante da precipitação foi observada em todas as classes segmentadas pelo CHAID (Gráfico 4.3.5-2).

Gráfico 4.3.5-2 – Variância CHAID por dia da semana



Destacam-se, na presença de precipitação, leituras de variância mais evidentes. Uma possível interpretação é o fato de que a precipitação move o patamar de variabilidade esperado a uma escala inferior, sendo a precipitação, portanto, não a principal causa da variabilidade, mais sim um fator de intensificação da degradação, tanto de velocidades quanto de distâncias percorridas.

#### 4.3.6 CHAID - Considerações gerais:

Diversos autores relatam variabilidade de velocidade e tempos de viagem em relação à presença de chuva.

Lamm et al. (1990) reportam aumento nos tempos de viagem de até 30% frente a precipitação. Hranac et al. (2006) descrevem redução de velocidade entre 1,9 km/h e 12,9 km/h em rodovias diante de diferentes níveis de precipitação.

Segundo Smith et al., (2004), apesar da redução das velocidades de operação não ser extrema, a capacidade geral das vias costuma ser comprometida da ordem de 3% a 5%, quando um evento de precipitação de chuva está em andamento. Akin; Sisopiku e Skabardonis (2011) concluíram, através de dados de contadores de tráfego, que precipitação inclemente causou redução nas velocidades e taxas de fluxo da via entre 8% e 12%.

Enquanto as variações médias diárias e sob precipitação são baixas (entre 1% a 3%), a classificação CHAID aqui apresentada demonstrou variações significativas hora a hora.

Tais dados convergem com os índices observados por diversos autores nos diferentes cenários considerados, sendo particularmente representativos nos horários de pico da tarde e nos horários que se seguem no início da noite. De forma semelhante, eventos de precipitação também causam significativa variabilidade pontual, embora a escala de precipitação não tenha sido associada a um padrão quanto às velocidades.

A variabilidade de 34% observada entre a classe de velocidades mais altas (segunda-feira, 20:00 horas) e a de velocidades mais baixas (sexta-feira, precipitação maior que 15,4 mm) demonstra o potencial de amplitude da variabilidade em si, além de indicar possíveis pontos críticos ou gargalos na fluidez do tráfego nos ônibus de transporte de São Paulo.

As árvores de regressão permitiram observar de forma detalhada como o fenômeno ocorre, seu impacto pontual e ainda, o possível efeito das mudanças de característica de demanda, tráfego e operação em situações específicas como, por exemplo, feriados.

Adicionalmente, pode se afirmar, ainda que especulativamente, que as classes representadas explicam a percepção individual da queda da qualidade no serviço de transporte público.

#### **4.4 ANÁLISES ESPACIAIS**

O uso de métodos estatísticos diversos, incluindo o CHAID, combinados a análises geográficas não é novo. Casas (CASAS, 2003) fez uso desta técnica na identificação de rotas pouco congestionadas em situações de trânsito adverso.

A análise de dados de transporte e tráfego, bem como sua variabilidade no âmbito geográfico, vem sendo consolidada através de diversos estudos. Neste contexto, métodos de regressão geográfica aplicados às velocidades registradas, seja por GPS, seja por contadores veiculares ou outros métodos, vêm produzindo importantes resultados (HAYNES et al., 2006; LOWRY, 2014; QUDDUS, 2013).

Selby e Kockelman (2013) analisaram as características de diversos métodos de regressão e interpolação sobre dados de velocidade em distâncias Euclidianas e não Euclidianas. Os autores concluíram que interpolações para distâncias euclidianas e não euclidianas possuem características semelhantes, assim como métodos de interpolação de superfície, como a distância geográfica ponderada (BRUNSDON; FOTHERINGHAM; CHARLTON, 1998) e a Krigagem (CRESSIE, 1990), sendo esta última ligeiramente superior, inclusive em comparação a métodos não geográficos.

##### **4.4.1 Contextualização dos resultados**

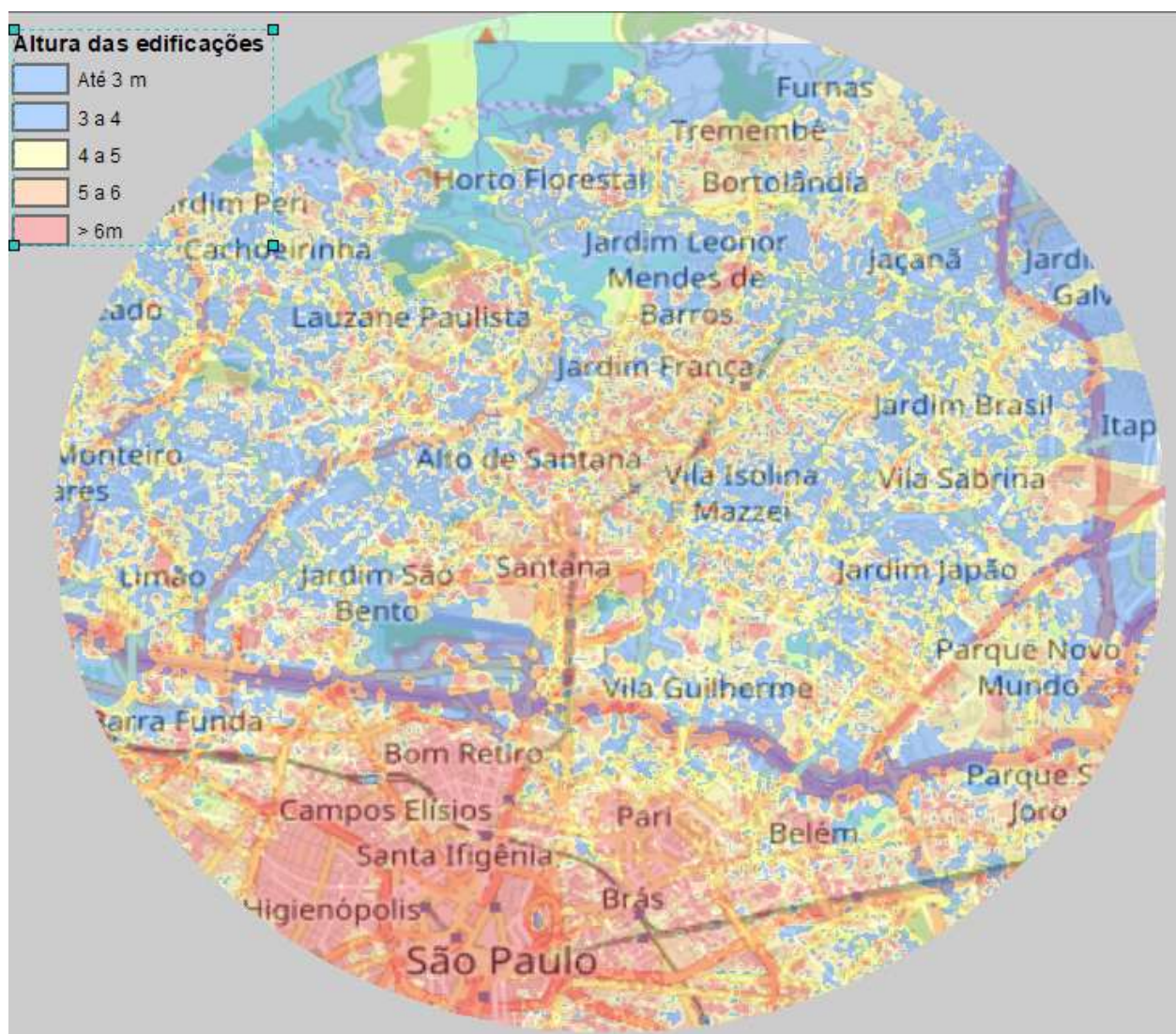
A área de estudo em questão constitui-se de forma heterogênea. No limite norte, encontram-se áreas de periferia próximas ao horto florestal que, embora densamente urbanizadas, são predominantemente de baixa verticalização. Em direção ao sul começa a se observar verticalização e polos geradores de demanda por viagens como: universidades, templos religiosos, centros comerciais, o Aeroporto Campo de Marte, o Terminal Rodoviário do Tietê, o Pavilhão de Eventos do Anhembi

e o Shopping Center Norte. A Marginal Tietê tem um papel relevante na mudança das características urbanas, especialmente ao sul da área de estudo, onde encontra-se o centro de São Paulo e bairros conhecidos do seu entorno, como: República, Santa Efigênia, Brás, Campos Elísios, Higienópolis, entre outros.

Com objetivo de estabelecer uma base de referência em relação a possíveis polos geradores de demanda por serviços de transporte, um mapa (Figura 4.4.1-1) da altura das edificações foi produzido, a partir de dados da prefeitura de São Paulo (SMDU, 2015).

Os resultados confirmam a heterogeneidade da área de estudo e fornecem uma base importante para a compreensão dos resultados no âmbito geográfico.

Figura 4.4.1-1 – Área de estudo: Altura das edificações

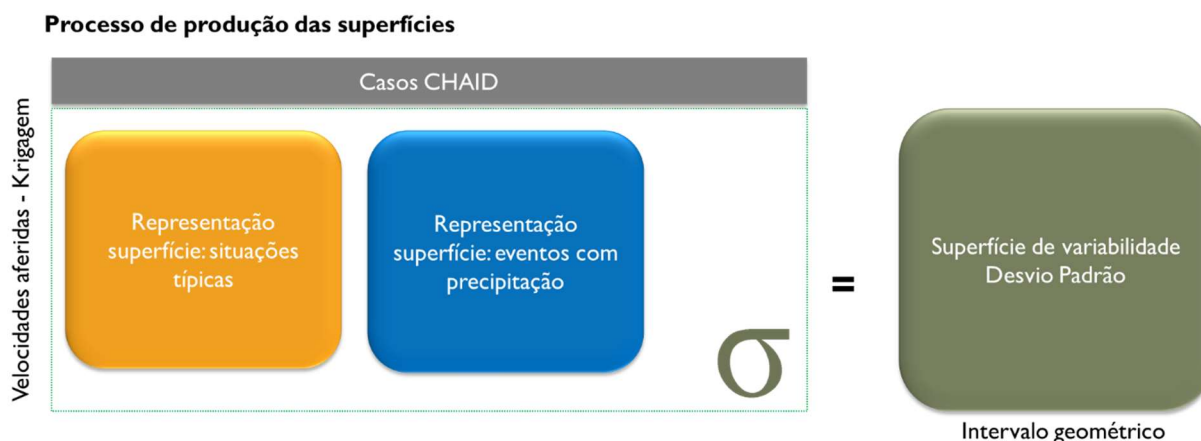


#### 4.4.2 Cálculo do desvio padrão das superfícies

As superfícies resultantes do processo de Krigagem foram submetidas a um processamento para cálculo do desvio padrão.

Os desvios-padrão foram calculados a partir de critérios gerais estabelecidos, a partir das observações dos resultados CHAID.

Figura 4.4.2 -1 – Fluxo de produção para superfícies de desvio padrão



#### 4.4.3 Distribuição das microviagens sobre a área de estudo – Visão geral

A análise geográfica das velocidades das microviagens, no mês de março de 2014, mostra um padrão semelhante de distribuição, tanto na ausência quanto na presença de precipitação (Figura 4.4.3-1).

As áreas próximas aos eixos viários, tanto no extremo norte da área de estudo, quanto próximo à Via Dutra (à leste) e Marginal Tietê (ao sul), apresentam um padrão de velocidades superiores, independentemente da condição climática.

Enquanto ao norte as áreas mais críticas quanto às velocidades (em vermelho) estão dispersas em uma grande área, há um claro padrão de lentidão na região correspondente ao centro de São Paulo, incluindo seu entorno, com Santa Cecília, República, Consolação, Sé, entre outros.

O fenômeno da redução da variabilidade observado na ocasião de eventos de precipitação, conforme análise das medidas de dispersão e classificação CHAID, manifesta-se geograficamente, com uma predominância da classe das velocidades entre 10 km/h e 20 km/h no mapa.

Os resultados geográficos também convergem no que tange à redução geral das velocidades, observadas com a substituição de boa parte das velocidades acima de 30 Km/h (azul escuro) em precipitação igual a zero, por velocidades entre 10km/h (amarelo) e 30 km/h (azul claro) nas suas respectivas regiões. As regiões correspondentes a acessos e viadutos aparentam ser significativamente afetadas, enquanto regiões centrais, embora prejudicadas, mantêm certa homogeneidade.

Neste contexto, a variabilidade parece ser mais intensa nas regiões de entorno das vias expressas, hipótese esta confirmada pelo cálculo do desvio das duas superfícies (Figura 4.4.3-2).

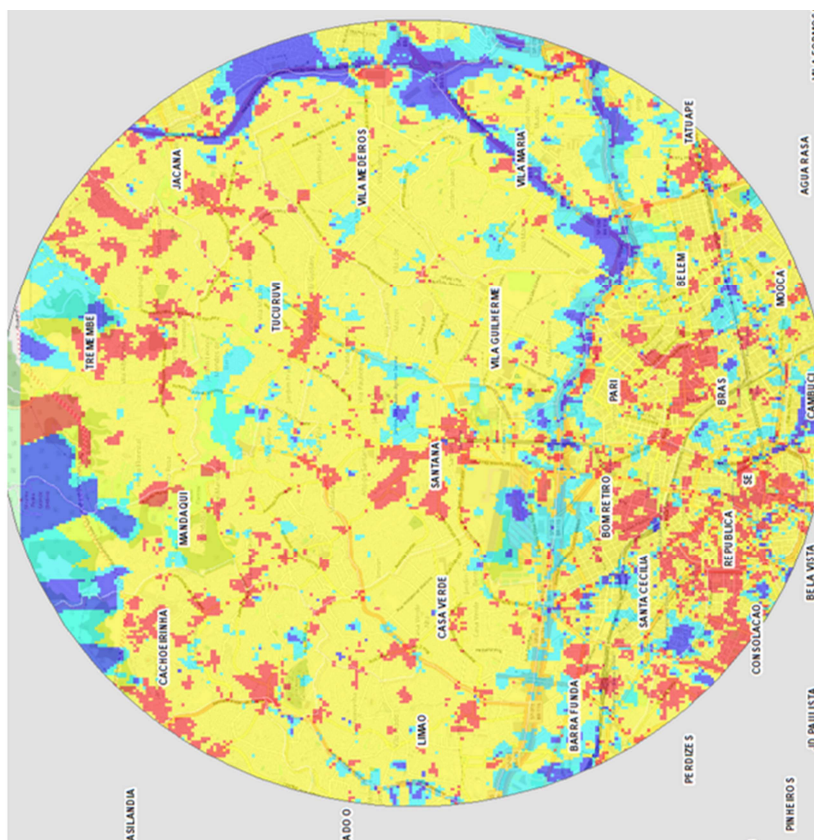
No entanto, desvios padrão acima de 15 km/h não são exclusividade das regiões de entorno da via Dutra e Marginal Tietê. Esses pontos estão presentes ao norte da área de estudo e ao sul, um pouco abaixo entre a região do Mandaqui e Tucuruvi.

No entorno da região central de São Paulo se observa sua ocorrência na região da Consolação, Cambuci, Belém Santa Cecília, entre outros. Este fenômeno também é observado na região ao leste da via Dutra.



Figura 4.4.3-1 – Comparativo: Krigagem das velocidades gerais (microviagens)

Precipitação > 0



Precipitação = 0

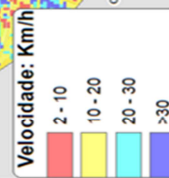
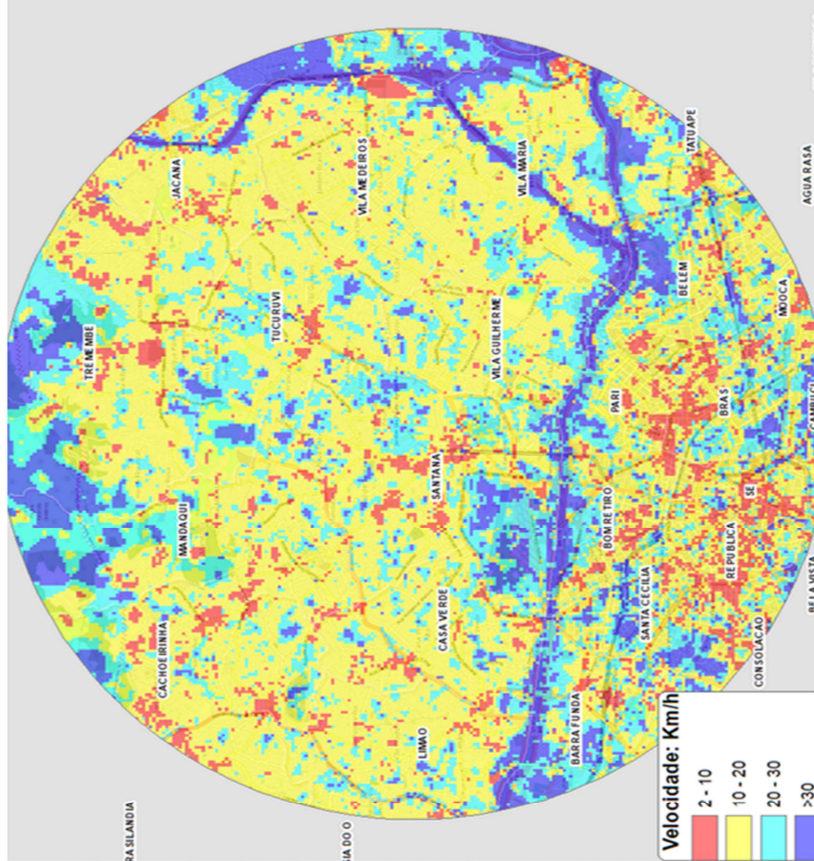
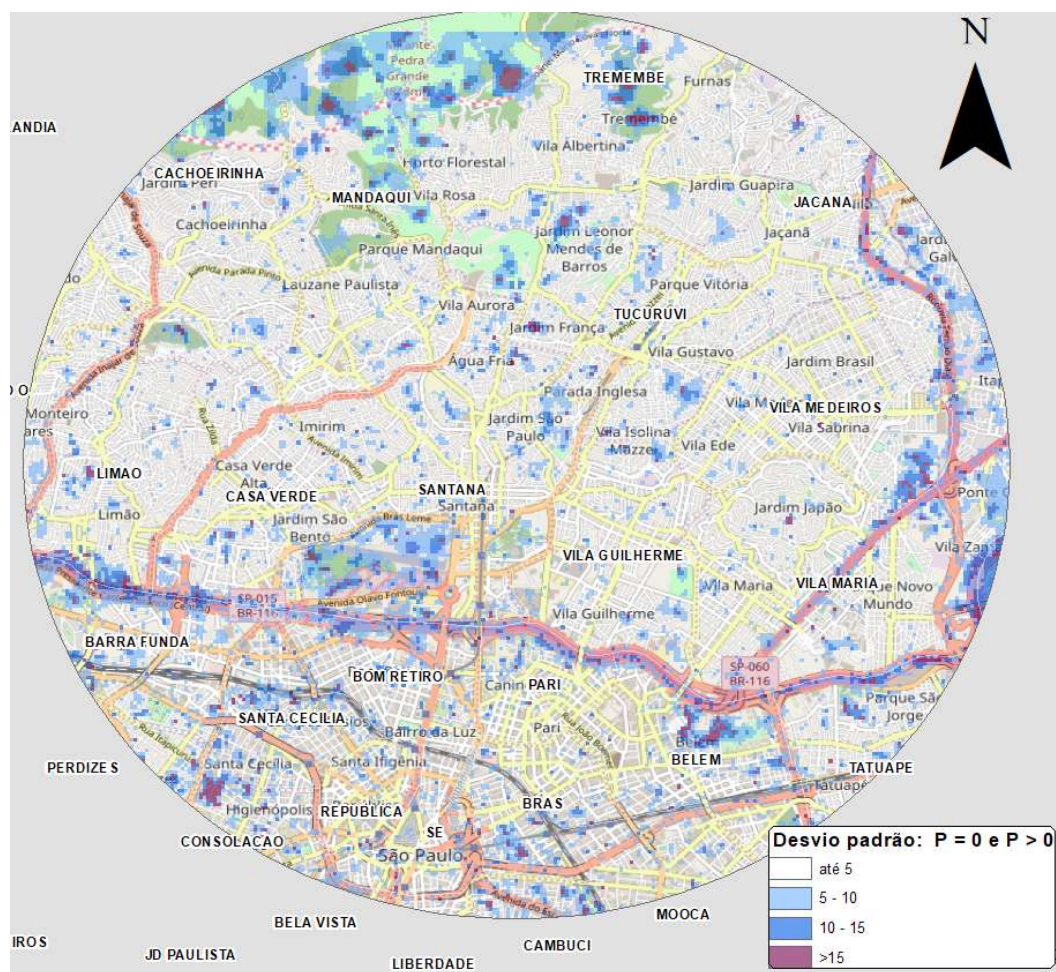


Figura 4.4.3-2 – Desvios padrão das velocidades (microviagens)



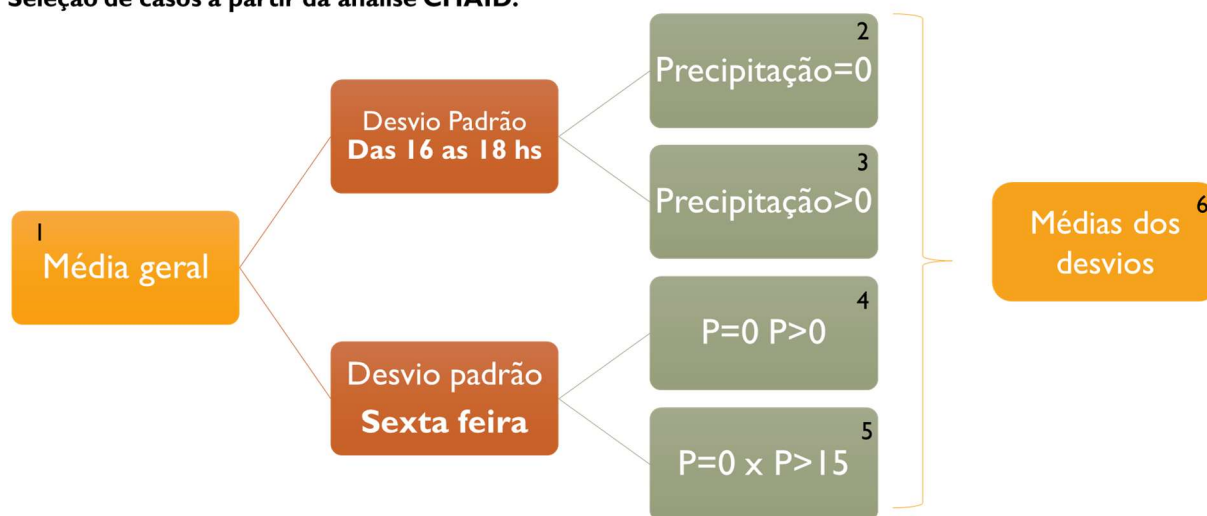
#### 4.4.4 Análise de casos selecionados com base na classificação CHAID e estatísticas descritivas:

Para compreender melhor a dinâmica das microviagens foi produzida uma análise dos desvios padrão, a partir das superfícies calculadas sobre as velocidades das mesmas (ver Figura 4.4.4-1). Foram utilizados critérios de seleção baseados nos resultados da classificação, através do método CHAID, analisando os horários das 16:00 horas, 17:00 horas e 18:00 horas, comparativamente na presença e ausência de chuva.

Também foi produzida uma análise comparando as microviagens em eventos de chuva muito fraca (até 1mm/hora) e chuva de moderada à forte (registros acima de 15mm/hora).

Figura 4.4.4-1 – Esquema descritivo das representações geradas a partir da classificação CHAID

Seleção de casos a partir da análise CHAID:



#### 4.4.5 Comparação geral: 16:00hs, 17:00hs e 18:00hs em março de 2014 na ausência e presença de chuva:

Foram calculados os desvios padrão das superfícies geradas a partir da análise das microviagens entre 16:00hs e 18:00hs. Para a classificação da superfície optou-se pela classificação por intervalo geométrico.

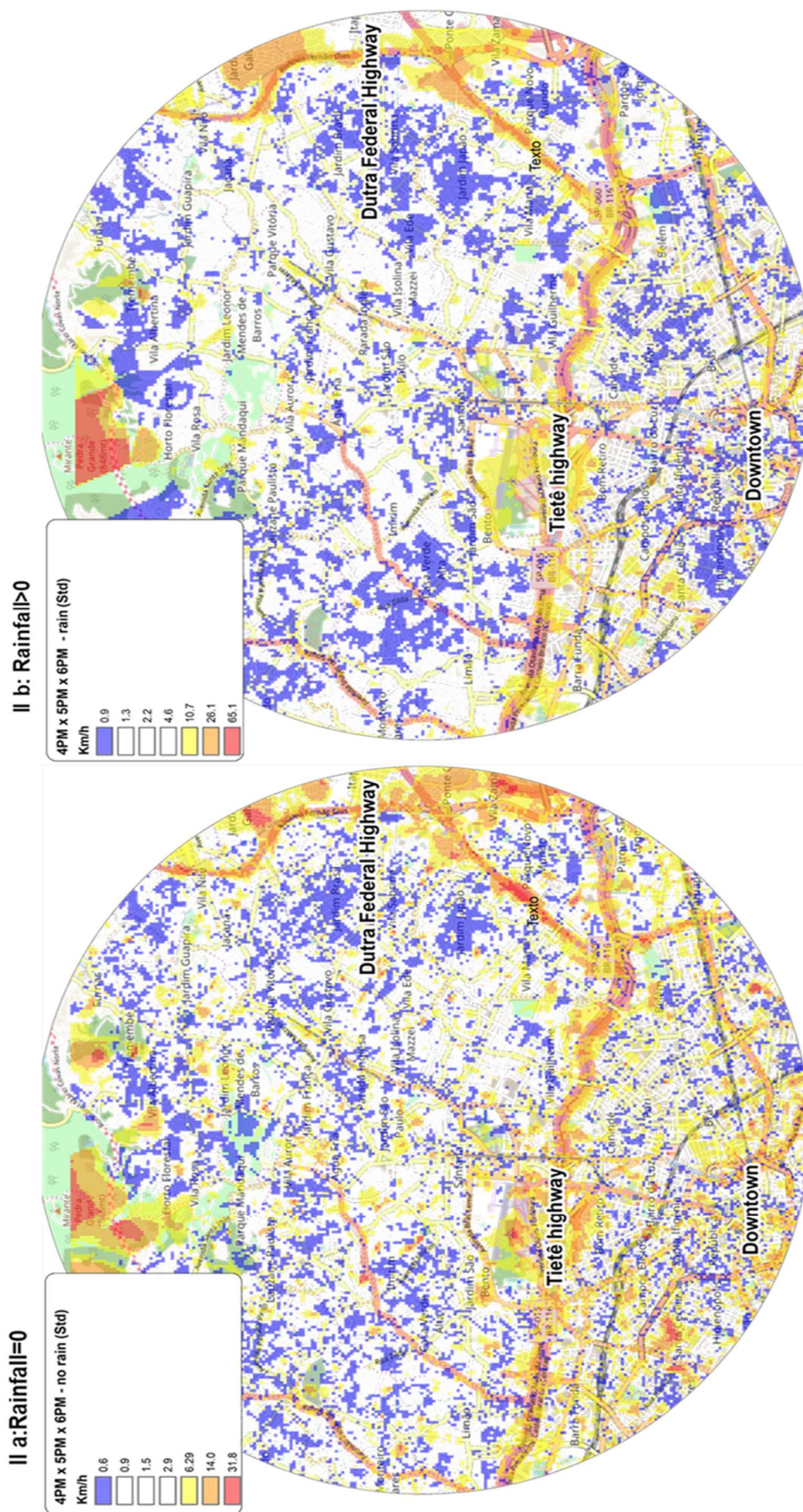
O esquema de classificação de intervalo geométrico é baseado em intervalos de classe que possuem uma série geométrica. O algoritmo cria intervalos geométricos minimizando a soma dos quadrados do número de elementos em cada classe. Isso garante que cada classe tenha aproximadamente o mesmo número de valores e que a mudança entre os intervalos seja bastante consistente (ESRI, 2017).

Em ambos os casos as classes até cerca de 2,6 Km/h de desvio foram suprimidas (em transparente) para uma melhor interpretação dos resultados. Isso significa que em aproximadamente 60% da superfície calculada os desvios de velocidade são relativamente baixos, conforme o período analisado.

Tanto na presença ( $p > 0\text{mm/h}$ ), quanto na ausência de precipitação ( $p = 0\text{mm/h}$ ) as classes remanescentes mostram que a variabilidade associada à progressão horária do pico da tarde segue um padrão já observado até aqui.

Diferentemente das análises estatísticas realizadas anteriormente, quando a combinação entre dia da semana, hora do dia e ocorrência de precipitação causam fenômenos distintos de variabilidade, a análise das superfícies mostra um padrão (Figura 4.4.5-1) que parece ser fortemente definido por características espaciais, independentemente da questão climática.

Figura 4.4.5-1 e 4.5.5-2 – Desvios padrão das velocidades (microviagens): 16:00s – 18:00hs. Respectivamente P=0 mm e P>0 mm



#### 4.4.6 Casos extremos CHAID:

A análise CHAID estabeleceu classes com base em nas velocidades das microviagens e fatores como dia da semana, hora do dia e presença de precipitação.

Dentre os casos classificados, destacam-se dois, resultado do processamento das microviagens, são eles: Sexta feira geral, com ausência de chuva, com média de velocidade de 16,93 km/h e Sexta feira com precipitação maior ou igual a 15,40 mm/h e velocidade de 12,94 km/h.

Dois mapas foram produzidos considerando os casos e o desvio padrão entre os dois. A classificação utilizou o algoritmo de intervalo geométrico (equivalência entre classes) presente no software ArcGIS.

O resultado sugere (Figura 4.4.6-1) que quando a precipitação (P) é igual 0, há grande variação na distribuição geográfica das classes sobre a superfície da área estudada, com exceção do entorno de rodovias e áreas extremamente periféricas, pouco representativas. A superfície produzida pelas velocidades das microviagens indica uma dinâmica de fluidez das vias urbanas no sentido das vias expressas. A variabilidade geográfica é mais heterogênea em áreas menos povoadas. O oposto ocorre nas áreas mais povoadas, circundadas por polos geradores de viagens como escolas, hospitais, centros comerciais entre outros.

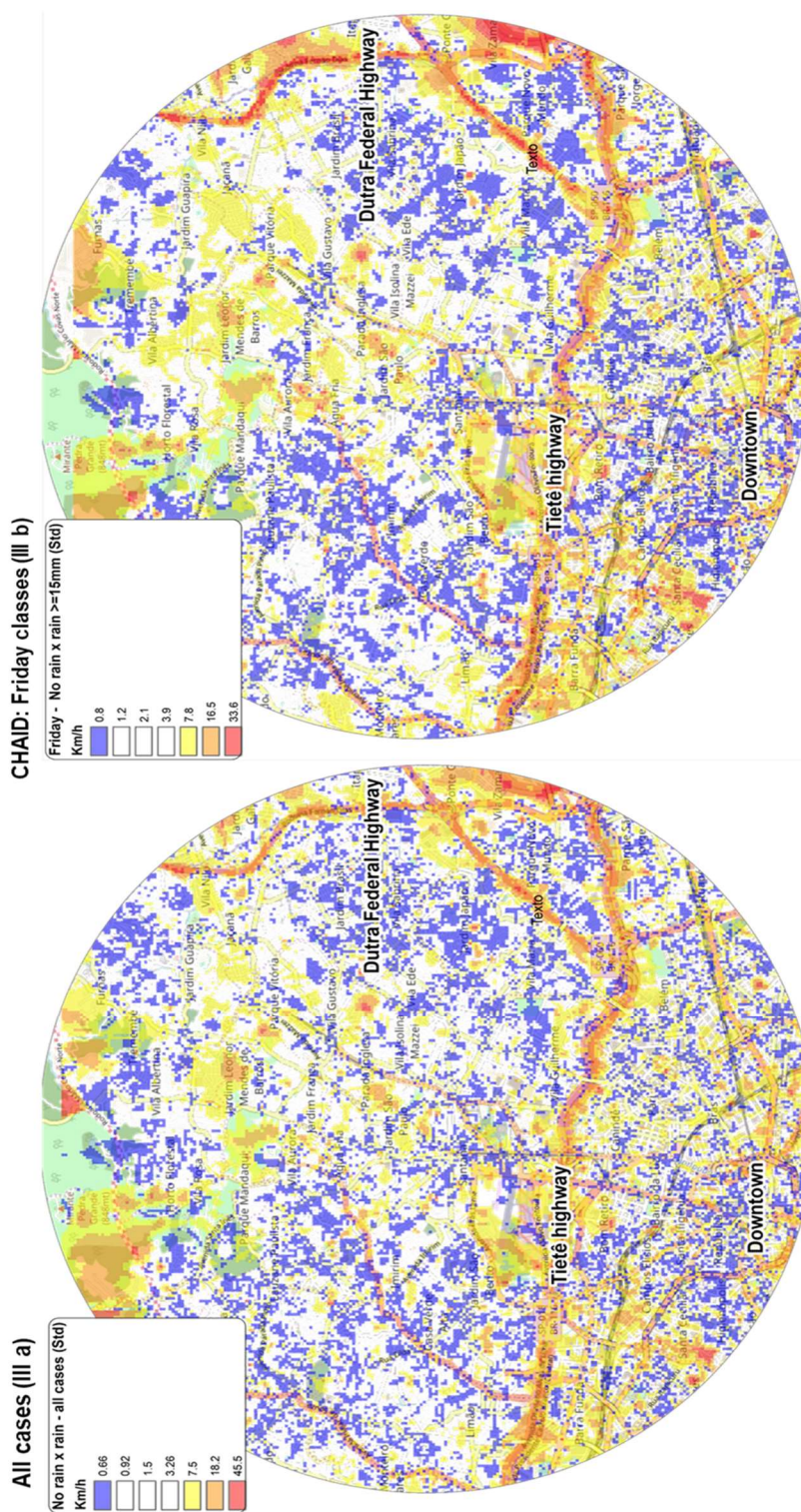
Em eventos de extrema precipitação as classes tornam-se uniformes, porém, observa-se uma generalização de velocidades baixas em regiões onde o fenômeno não era esperado, como em regiões ao centro, norte e nordeste da área de estudo. O pentágono formado pelas regiões de Tremembé, Mandaqui, Jaçanã, Santana e Vila Medeiros, com Tucuruvi ao centro tem sua maioria coberta pela cor vermelha, classe representada por velocidades de até 9 km/h.

Há uma variabilidade aproximada de 8 km/h, com picos que podem ir até 33 km/h na região do Tucuruvi. O mapa de desvios padrão é, em geral, um reflexo dos eventos de precipitação intensa, apresentados em seu respectivo mapa.

O cálculo dos desvios das velocidades, nas superfícies calculadas entre os casos extremos, fornece uma visão clara da dinâmica da variabilidade das microviagens e sugere indícios de que certos padrões de superfície são determinantes

na variabilidade das velocidades, distâncias euclidianas e mesmo tempos de viagem dentro das paisagens urbanas.

Figura 4.4.6-1 e 4.4.6-2 – Desvios Padrão - Casos CHAID respectivamente  $p=0 \text{ mm} \times p>0 \text{ mm}$  e  $p=0 \text{ mm} \times p>15 \text{ mm}$

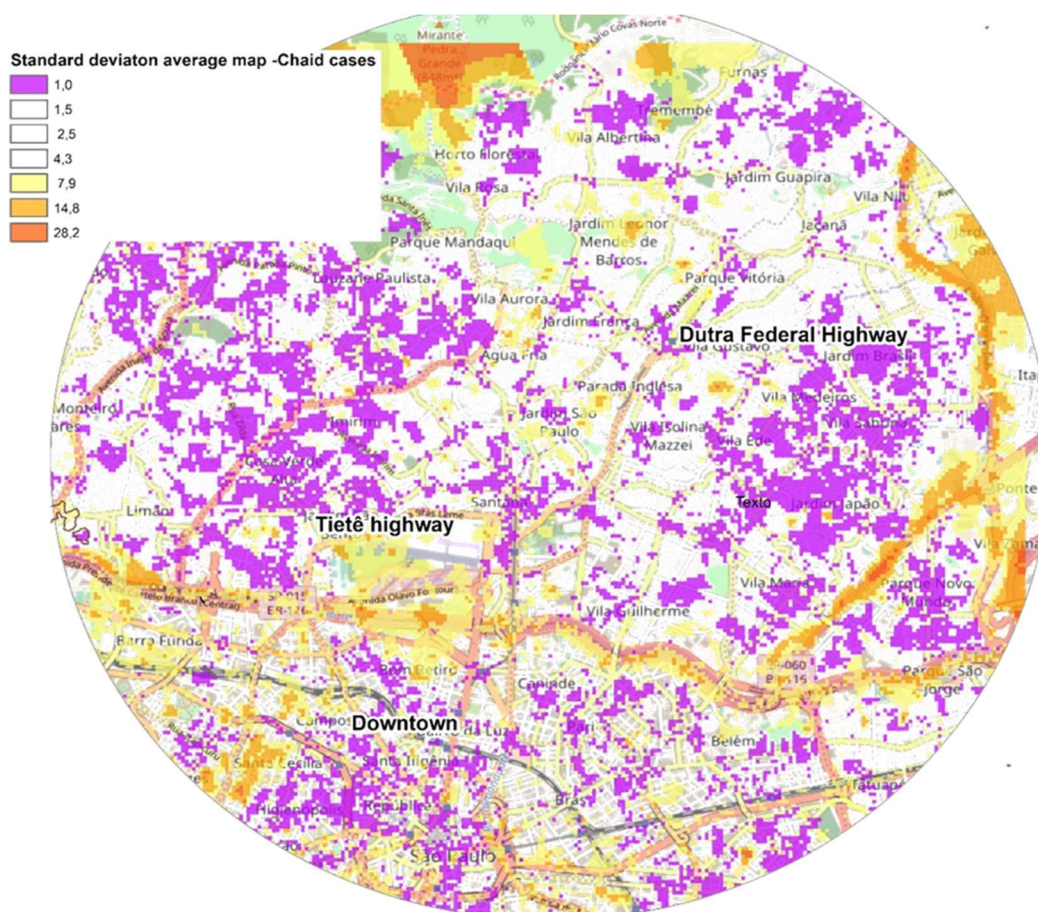


#### 4.4.7 Visão Geral: média dos desvios CHAID

A partir da premissa de que os cenários analisados possuem um padrão recorrente de variabilidade espacial, foi produzida uma superfície a partir das médias dos desvios. Os resultados confirmam a recorrência de padrões espaciais observados na Figura 4.4.7-1 e Tabela 4.4.7-1. A análise da distribuição das classes apresenta desvios padrão semelhantes, reforçando as observações anteriores.

Quando comparadas as classes entre si, situações de horário de pico com presença de precipitação possuem desvios ligeiramente maiores que em outros casos.

Figura 4.4.7-1 Consolidado: Média dos desvios padrão dos casos CHAID





5

Tabela 4.4.7-1 Comparativo entre classes - desvios CHAID

**Class surface representativeness (%)**

| Map class - average (Km/h) | No rain x rain (all classes) | Friday - No rain x rain >=15mm (Std) | 4PM x 5PM x 6PM - rain (Std) | 4PM x 5PM x 6PM - no rain (Std) | Standard deviation |
|----------------------------|------------------------------|--------------------------------------|------------------------------|---------------------------------|--------------------|
| 0,7                        | 23%                          | 21%                                  | 22%                          | 18%                             | 2%                 |
| 1,1                        | 8%                           | 9%                                   | 12%                          | 11%                             | 2%                 |
| 1,8                        | 19%                          | 18%                                  | 25%                          | 21%                             | 3%                 |
| 3,7                        | 25%                          | 24%                                  | 27%                          | 24%                             | 1%                 |
| 8,1                        | 19%                          | 19%                                  | 10%                          | 17%                             | 4%                 |
| 18,7                       | 7%                           | 8%                                   | 2%                           | 6%                              | 3%                 |
| 44,0                       | 1%                           | 1%                                   | 1%                           | 1%                              | 0%                 |
| <b>Standard deviation</b>  | 9%                           | 8%                                   | 11%                          | 8%                              |                    |

## 5. DISCUSSÃO

O estabelecimento de um método para a caracterização dos padrões de operação dos ônibus de transporte da cidade de São Paulo para a quantificação do impacto da precipitação sobre sua operação, aqui proposto como principal objetivo, constitui-se como um desafio técnico desde a modelagem e processamento dos dados, até a compreensão geral dos resultados estatísticos e geoprocessados.

A construção da compreensão do problema veio através da combinação de técnicas que abordam diferentes aspectos e dimensões do conjunto de dados trabalhados, culminando em uma detalhada caracterização da operação dos ônibus de transporte de passageiros urbanos, na área de estudo aqui definida no período de março de 2014.

Ao longo das investigações e análises, os objetivos complementares tornaram-se componentes do objetivo principal, fundamentais para o estabelecimento do problema.

### 5.1 INDICADORES DE VARIABILIDADE

Muitos dos trabalhos aqui referenciados vêm usando medidas de dispersão como base para caracterizar a variabilidade, seja em tempos de viagem, seja em leituras de velocidade ou distâncias (AKIN; SISIOPIKU; SKABARDONIS, 2011; CARRION; LEVINSON, 2012; CHANGNON, 1996; CHEN; SKABARDONIS; VARAIYA, 2003; EOM et al., 2006; FRANKLIN; KARLSTROM, 2009; GUO; WILSON; RAHBEE, 2007; KAMGA; YAZICI, 2014; KASHFI; BUNKER; YIGITCANLAR, 2016; KHATTAK; DE PALMA, 1997a; KWON et al., 2011; LIU; RECKER; CHEN, 2004; LOWRY, 2014; LYMAN; BERTINI, 2008; MARTCHOUK; MANNERING; SINGH, 2010; SELBY; KOCKELMAN, 2013; SKABARDONIS; VARAIYA; PETTY, 2003; TAO et al., 2016; TAYLOR, 2009; TU; VAN LINT; VAN ZUYLEN, 2007; VAN LINT; VAN ZUYLEN; TU, 2008).

Embora haja peculiaridades em cada estudo e os autores nem sempre concordem em relação a interpretação de um fenômeno, muitas das conclusões

observadas são semelhantes aos resultados obtidos junto aos ônibus de transporte da cidade de São Paulo em março de 2014.

### 5.1.1 Médias

Franklin e Karlstorm (2009) observaram, nas vias urbanas da Suécia que as variações nos tempos de viagem podem ser explicadas por segmentação horária diária e respectivas demandas.

Os autores ainda verificaram que o atraso médio dos períodos de pico provavelmente está associado com as características da estrada, como o fluxo, velocidade, localização em relação ao centro da cidade e níveis de serviços (congestionamento).

Chien e Kolluri (2012) observaram em rodovias americanas que no ciclo semanal, quintas e sextas apresentam tempos de viagem mais elevados.

Kwon et al. (2011) constataram em rodovias americanas uma maior demanda no pico da tarde com queda da variabilidade no período da tarde, possivelmente devido a gargalos como acessos a rodovias secundárias e vias urbanas.

O aumento dos tempos de viagem é um fenômeno está relacionado à queda das velocidades desenvolvidas em uma determinada distância.

Nesse trabalho foram observados fenômenos semelhantes em relação as quedas nas médias das velocidades aferidas pelas microviagens, em situações semelhantes.

O gráfico 4.2-2 caracteriza as médias de velocidades por dia da semana e hora do dia, sem a interferência da precipitação, permitindo observar a variação das médias das velocidades no decorrer das dimensões temporais, conforme observado por Franklin e Karlstorm. Este gráfico (4.2-2) demonstra a queda gradativa das velocidades diárias desenvolvidas pelos ônibus como observado em rodovias por Kwon et al. E ainda, a intensificação desse fenômeno as quintas e sextas conforme afirmado por Chien e Kolluri.

Considerando o fato de que nenhum dos autores trabalhou especificamente com ônibus de transporte público, considera-se que a opção metodológica de exclusão das microviagens com velocidades abaixo de 2,9 km/h, feita durante a fase de modelagem de dados, permitiu aferir situações de movimento dos ônibus. Adicionalmente, essa sinergia entre resultados remete a considerar a possibilidade da utilização dos ônibus como instrumentos de sensoriamento das condições de trânsito na cidade de São Paulo.

### **5.1.2 Precipitação e médias**

Hranac et al. (2006) identificou em rodovias americanas uma queda nas velocidades de cerca de 1,9 km/h para chuva leve durante períodos de trânsito livre. Em condições de pico, as vias apresentaram lentidão, com redução de até 12,9 km/h.

A observação dos resultados gerais contidos na tabela 4.2-1 sugere um impacto discreto da precipitação sobre as velocidades. No entanto, o gráfico 4.2-3 apresenta em sua curva de precipitação uma queda nas velocidades semelhante ao observado por Hranac, com cerca de 6 km/h de amplitude entre 9:00 horas e 20:00 horas.

Há uma relação inversa entre a curva de velocidade na presença de precipitação e a precipitação em si, semelhantes, aparentemente definidas pelo horário.

### **5.1.3 Desvio Padrão**

Martchouk et al. (2010) constataram que os tempos de viagem podem variar mais de 100% em um único dia. Foram observados desvios padrão elevados durante períodos de pico. De forma semelhante, Franklin e Karstorm observaram uma distinta variação dos tempos de viagem através de seus desvios padrão ao longo do dia. Ainda sobre esse tema, Akin et. al (2003) observaram variações significativas das

velocidades em horários diários e na presença de eventos climáticos, pelas análises das médias e desvios padrão.

Kamga e Yazici (2014) afirmam que tempos de viagem mais altos também possuem maior previsibilidade, especialmente nos horários de pico.

Tu et. al (2007), observaram em rodovias holandesas um aumento na variação dos tempos de viagem frente a clima adverso.

A quantificação das variações das velocidades das microviagens, utilizando a relação de desvio padrão sobre o tempo, se mostrou efetiva na compreensão de certos padrões não evidentes na análise das médias das velocidades.

Na observação dos desvios hora a hora em situações de ausência de precipitação, consideradas como “situação normal de operação” (laranja), o gráfico (4.2-4) apresenta uma queda gradativa da variabilidade com seu ponto mais baixo as 18:00 horas. Há uma maior previsibilidade, neste caso das velocidades, no pico das 18:00 horas, conforme observado por Kamga e Yazici em seus estudos sobre tempos de viagem em ambiente urbano em Nova Iorque. Segundo os autores, essa previsibilidade se dá em função da normalização dos tempos de viagem causada pelo excesso de veículos em circulação.

No caso dos ônibus paulistanos, presume-se que essa queda de variabilidade seja reflexo da degradação das condições de trânsito ao longo do dia, com auge no fim da tarde, observada por Kwon et. al. Adicionalmente, pressupõe-se que a demanda pelo uso do transporte público deva aumentar nestes períodos, acentuando as curvas de desvios padrão observados (gráfico 4.2-4). É possível inferir, embora não especificamente investigado, que a demanda sobre o serviço de transporte se manifeste através das pequenas variações observadas na curva, representando ausência de chuva ao longo do dia. O aumento da demanda em certos horários pode se constituir como um componente da variabilidade, uma vez que os veículos tendem a permanecer parados por mais tempo e com maior frequência nos pontos de ônibus. A própria taxa de ocupação dos veículos (possivelmente lotados) pode ser responsável pelo desenvolvimento de velocidades moderadas ao longo do trajeto, contribuindo para a queda da variabilidade. Em termos práticos, em situações de alta demanda, a previsibilidade de chegadas e partidas deve melhorar, favorecendo o uso de aplicativos de mobilidade relacionados a chegadas e partidas de ônibus.

#### 5.1.4 Desvios e precipitação

A queda da variabilidade em situações de precipitação ou clima inclemente foi consistentemente citada até aqui, seja por métodos de medição de tempos de viagem (CARRION; LEVINSON, 2012; CHEN; SKABARDONIS; VARAIYA, 2003; FRANKLIN; KARLSTROM, 2009; KAMGA; YAZICI, 2014; KHATTAK; DE PALMA, 1997a; KWON et al., 2011; MARTCHOUK; MANNERING; SINGH, 2010; SKABARDONIS; VARAIYA; PETTY, 2003; VAN LINT; VAN ZUYLEN; TU, 2008), seja pela medição direta das velocidades (AKIN; SISIOPIKU; SKABARDONIS, 2011; LOWRY, 2014; QUDDUS, 2013; SELBY; KOCKELMAN, 2013).

A observação da curva de desvio padrão das velocidades sob precipitação (azul claro) no gráfico 4.2-4 apresenta uma queda na variabilidade diante dos eventos de precipitação. A comparação com a curva de precipitação média acumulada/hora (azul pontilhado), no mesmo gráfico, sugere uma relação de causa e efeito sobre as velocidades e sua variabilidade.

Valores de precipitação abaixo de 1 mm/h parecem criar uma situação favorável de trânsito livre, refletido no gráfico por velocidades médias mais elevadas até mesmo em relação a dias sem precipitação. Uma das possíveis explicações é que a precipitação por si não se dá de forma instantânea. Há visíveis mudanças no ambiente, como temperatura, luminosidade, entre outros.

A alteração no comportamento tanto de motoristas, quanto de usuários de transporte público é um tema controverso. Há observações consistentes de mudanças na demanda por transporte público e privado frente a diferentes situações climáticas (KASHFI; BUNKER; YIGITCANLAR, 2016).

Uma possível explicação para elevados desvios (aumento de velocidade) observados entre 8:00 horas e 11:00 horas no mês de março de 2014 sob chuva, pode estar condicionada à baixa demanda por transporte público e infraestrutura em geral, ocorrida durante o feriado de carnaval favorecendo médias de velocidade excepcionalmente altas neste horário.

Embora esse fenômeno tenha duração de algumas horas, a progressão horária e a degradação das velocidades a estes associados, demonstram um efeito

acentuado no período da tarde, reduzindo os desvios e estabilizando as curvas em situação de presença e ausência de precipitação em feições semelhantes.

No período entre 17:00 horas e 19:00 horas, com precipitação média entre 2mm/h e 5mm/h, observa-se um atraso do impacto nas velocidades, semelhante aos atrasos na formação dos congestionamentos observados por Franklin e Karlstrom.

Segundo os autores, esta assincronia entre as curvas seria consequência do tempo necessário para que o efeito “gargalo” observado nas entradas e saídas das vias expressas se alastre pelo sistema de transporte.

Embora a queda na variabilidade esteja caracterizada, os desvios médios (hora a hora) não apresentam valores extremos. Como exemplo, às 18:00 horas, se observam desvios de velocidades entre 12,1 km/h - sem precipitação - e 11,7 km/h - com precipitação.

#### **5.1.5 Coeficiente de variação**

A comparação do coeficiente de variação (gráfico 4.2-5) com as médias pluviométricas permite traçar uma relação direta entre volume médio de precipitação/hora e variação % das velocidades. Observa-se uma amplitude do coeficiente de 9% (67% a 76%) frente a precipitações médias de 0mm/h a 5mm/h e redução média de velocidades de aproximadamente 1 km/h.

Com base na observação do gráfico 4.2-5, é possível inferir uma relação de correspondência de cerca de 1% de variação para cada milímetro de precipitação acumulado.

Embora o gráfico do coeficiente de variação apresente um padrão inverso ao observado nos táxis de Nova Iorque (YAZICI; KAMGA; SINGHAL, 2013), devido à natureza dos dados estudados (tempos de viagem x médias de velocidade), é possível afirmar que os ônibus de São Paulo apresentam comportamento semelhante. Para os táxis de Nova Iorque, os tempos de viagem aumentam e o coeficiente de variação cai. Para os ônibus de São Paulo a velocidade cai e os coeficientes aumentam.

Confirmam-se os fenômenos observados em relação ao impacto da precipitação nas velocidades e variabilidade, observada através das médias, desvios e coeficiente de variação.

## **5.2 CHAID – CLASSIFICAÇÃO DE CASOS AFINS**

A descrição parcimoniosa de um grande conjunto de dados, composto de diversos vetores categóricos independentes, associados a um conjunto com preditores e suas categorias pode ser ou não indexável. Um típico conjunto de dados pode ter centenas de vetores, cada um composto de diversos elementos (KASS, 1980).

A técnica conhecida como CHAID particiona os dados em subcategorias de dados mutuamente exclusivas, que melhor descrevem as variáveis dependentes (KASS, 1980).

A técnica CHAID, aplicada às velocidades das microviagens, foi utilizada para aprofundar a compreensão da interação entre os vetores objeto deste estudo frente as variações das velocidades.

Um dos problemas da caracterização da variabilidade das velocidades dos ônibus de São Paulo, frente à precipitação, é a definição de padrões de velocidade - referências capazes de estabelecer bases para a identificação dos fenômenos de variabilidade.

A análise isolada da variabilidade das velocidades em dias secos (gráficos 4.3.1-1 e 4.3.1-2) mostra um padrão recorrente de degradação das velocidades, hora após hora, no período da tarde, culminando em uma abrupta melhora a partir das 19:00 horas.

Essa classificação endereça um dos objetivos desta dissertação: a caracterização das velocidades desenvolvidas pelos ônibus de São Paulo. Diferentemente das médias gerais, a classificação via CHAID demonstra padrões diários cíclicos recorrentes e a variabilidade horária, exibindo através dos indicadores, aspectos que podem ser relacionados à percepção individual de melhoria ou queda nas condições de serviço dos ônibus de transporte de São Paulo.



A análise dos eventos exclusivamente sob precipitação, (gráficos 4.3.2-1 e 4.3.2-2) apresentou, com exceção da terça-feira, um padrão recorrente de degradação no período da tarde, conforme tendência observada. Graças a essa análise, o impacto do feriado de carnaval ficou evidente, embora os valores observados nessa modelagem não se apresentem significativamente menores em relação a dias secos.

O gráfico 4.3.3-1 compara os resultados exclusivos (chuva x seco) em uma tabela de horário normalizado, onde em geral repetem-se as observações. Com exceção da quinta, os horários próximos ou o horário de pico (17:00 – 18:00 horas) apresentaram velocidades mais altas em dias secos em relação a dias chuvosos. As análises isoladas e sua comparação direta tiveram por objetivo caracterizar os eventos em cada uma das suas circunstâncias. Até aqui, as classificações dos dados vêm sendo testadas em conjuntos isolados, com o objetivo de caracterização. Neste contexto, os resultados são semelhantes aos indicadores descritivos, com a vantagem de um maior detalhamento desta classificação sobre eventos temporais.

A integração do conjunto total de dados em um mesmo processamento (gráficos 4.3.4-1 e 4.3.4-2) permite que o algoritmo estabeleça agrupamentos de base comparativa em todos os cenários. Como resultado, novas classes são constituídas e organizadas (terças, quartas e quintas), cuja distinção considera a ausência e presença de precipitação, além de vetores temporais como hora do dia e dia da semana. Há uma prevalência do vetor precipitação em comparação com o vetor hora em diversos casos, normalizando a variabilidade horária-diária a uma única classe com ausência de precipitação denominada " $\leq 0$  mm".

As segundas-feiras constituem um bom exemplo da variabilidade em situações consideradas operacionalmente normais, ou seja, isentas de chuva. Não há registros de precipitação neste dia da semana em todo o mês de março. Assim, a classificação horária é apresentada de forma semelhante ao observado nos gráficos 4.3.1-1 e 4.3.1-2, onde há uma degradação das velocidades contínua no período da tarde, seguida por uma melhora significativa a partir das 19:00 horas.

Os resultados distribuídos nos grupos de terça, quarta e quinta demonstram a precipitação como fator determinante na variabilidade, independentemente do volume de chuvas. Mesmo volumes de precipitação considerados baixos (em torno de 0,20

mm/h, por exemplo) apresentam-se como vetores mais significativos do que a faixa horária.

Entretanto, nas sextas feiras as condições de sazonalidade horária demonstraram um efeito marcante na variabilidade. Como resultado, houve uma distinção dos nós em 8 subcategorias classificadas como ausência total de precipitação, precipitação distinta a 0,40 mm/h para as 17:00hs e 18:00hs e classes específicas para precipitação entre 0,60 mm/h até valores maiores que 15,40 mm/h, significando, que os dois cenários são relevantes em termos de amostras.

Neste contexto, as velocidades variaram entre 16,93 km/h, na ausência de chuva, até 12,94 km/h, com precipitação acima de 15,40 mm/h (amplitude=3,99 km/h).

Os resultados mostram que há eventos de variabilidade acentuada, significativos o suficiente para resultarem em uma classe. O número de classes gerados e sua representatividade absoluta sobre o universo estudado levam à conclusão de que grupos menores, porém relevantes do ponto de vista estatístico, apresentaram médias significativamente baixas, sugerindo pontos de lentidão e possível engarrafamento nas sextas (gráficos 4.3.4-1 e 4.3.4-2).

Há uma clara relação da queda da variância dos dados contidos nas classes frente a presença da precipitação (gráficos 4.3.5-1 e 4.3.5-2), semelhante ao observado no desvio padrão, sugerindo que a variância também é uma medida de dispersão adequada, simples e eficaz para avaliação da variabilidade das velocidades.

Através da análise descritiva obtém-se uma visão operacional de perda de aproximada de 1 km/h de velocidade na presença de precipitação. Com a análise CHAID são apresentadas variações mais significativas, de cerca de 4km/h.

Ambas as técnicas contabilizam a variabilidade das velocidades por registro, em intervalo máximo de 5 minutos. Casos individuais podem apresentar perdas de velocidade muito superiores.

O método se mostrou efetivo na caracterização geral e identificação dos eventos causadores da variabilidade e seu impacto nas velocidades cotidianas. Através desse método será possível, em estudos futuros, aprimorar o impacto escalar da

precipitação sobre a velocidade, bem como estabelecer a relevância de variáveis adicionais como acidentes de trânsito, eventos de mobilização de massa, entre outros.

### **5.3 DETERMINISMO GEOGRÁFICO**

Os resultados obtidos através do CHAID estabelecem os padrões variabilidade das velocidades nos ônibus de São Paulo em março de 2014, na dimensão temporal e clima. A partir destes padrões, uma investigação sobre o comportamento espacial da variabilidade foi efetuada.

O uso de métodos de interpolação de superfícies para análise de tráfego vem apresentando resultados consistentes, consolidando-se como um método efetivo em estudos de transporte (BLACK; THOMAS, 1998; EOM et al., 2006; PAEZ, 2006; SELBY; KOCKELMAN, 2013).

A Krigagem (CRESSIE, 1990) é um método eficiente na identificação de padrões de superfície, especialmente na determinação de velocidades através de distâncias euclidianas em ambiente urbano (SELBY; KOCKELMAN, 2013).

A Figura 4.4.2-1 retrata as velocidades desenvolvidas pelos ônibus de transporte de São Paulo, durante todo o mês de março de 2014, sob precipitação e sob ausência da mesma. As superfícies foram produzidas através da ferramenta Krigagem empírica Baesiana do software ArcMap. Em termos gerais, o mapa mostra, de uma forma marcante, o impacto da precipitação, especialmente nas vias expressas. O uso da técnica de intervalo geométrico na definição das classes, permitiu observar a queda da variabilidade das velocidades diante dos eventos de precipitação. Há maior homogeneidade geográfica das classes, especialmente para velocidades entre 10 km/h e 20 km/h, diante de eventos de precipitação, o que confirma observações feitas até aqui por meio de outros métodos.

A comparação da variabilidade por intermédio dos desvios das velocidades entre os horários das 16:00 horas e das 18:00 horas, na presença e ausência de precipitação (Figuras 4.4.5-1 e 4.4.5-2) confirma uma maior homogeneidade das velocidades ante a precipitação.

A Figura 4.4.5-1 caracteriza as velocidades dos ônibus de São Paulo no período do pico da tarde em situação comum. Nessas condições formam-se áreas de grande desvio padrão das velocidades, principalmente em vias expressas e seus acessos, e certas áreas onde há polos geradores de viagens, como centros comerciais ou centros residenciais de alta densidade.

Há uma relação entre regiões densamente povoadas e economicamente ativas, com redução das velocidades. No entanto, é possível identificar, geograficamente, pontos passíveis de intervenção por parte das autoridades de trânsito

A definição de classes, obtida pela análise CHAID, estabelece uma segmentação temporal para a análise espacial, com destaque para as sextas-feiras (Figuras 4.4.6-1 e 4.4.6-2). A análise do desvio padrão entre os cenários extremos desse dia (precipitação = 0 mm x precipitação > 15mm) confirma a ocorrência da variabilidade redutiva, permitindo a identificação e confirmação das observações feitas por meio da análise CHAID em relação a possíveis gargalos de velocidade.

No entanto, a Figura 4.4.6-2 estabelece que nem toda a extensão da área de estudo foi significativamente impactada. Ao contrário, regiões como Limão, Casa Verde e Vila Medeiros e até mesmo República e Consolação não parecem estar com índices de variabilidade baixos em relação a regiões muito próximas aos seus respectivos entornos.

A Figura 4.4.7-1 confirma os padrões de variabilidade observados até aqui e confirmados na tabela 4.4.7-1, onde é quantificada a representatividade de cada classe de desvio.

Certos padrões podem estar relacionados a características como: estrutura viária, economia, demografia, e outras, que, combinadas, resultam em um melhor, ou em um pior desempenho nas velocidades desenvolvidas pelos veículos de transporte em geral em São Paulo.

Além da caracterização temporal, as evidências apontam para a geografia, em um sentido amplo, como fator determinante da variabilidade das velocidades dos ônibus no município de São Paulo.

É possível que características geográficas sejam, ainda, o principal fator na diversidade das conclusões, que - por vezes, apresentam-se contraditórias em relação

ao clima como fator impeditivo ou incentivador do uso do transporte público (KHATTAK; DE PALMA, 1997a) (CHANGNON, 1996) (GUO; WILSON; RAHBEE, 2007) (STOVER; MCCORMACK, 2012) (KASHFI; BUNKER; YIGITCANLAR, 2016).

#### **5.4 CONSIDERAÇÕES FINAIS**

A variabilidade do trânsito não é um fenômeno simplesmente restrito ao ciclo diário e horário. E, embora a precipitação de chuva tenha um papel importante, essa não é, contudo, o único vetor relevante observado nesse fenômeno.

As características físicas da infraestrutura de transportes e da ocupação territorial são igualmente significativas na manifestação da variabilidade das velocidades, e é possível que ainda haja outros fatores relevantes e determinantes para a compreensão do problema.

É possível observar a variabilidade nas velocidades dos ônibus de transporte de São Paulo ante a fenômenos de precipitação. Isso ocorre porque, em termos práticos, a precipitação altera os parâmetros de operação do sistema urbano, influenciando o comportamento de operadores e usuários.

Fatores como a hora e o dia da semana, bem como as feições e a capacidade das vias, as características físicas dos veículos, do relevo, a densidade populacional e a atividade econômica, em geral, seguem padrões cíclicos, sendo, portanto, de difícil percepção, tanto analiticamente, quanto do ponto de vista individual do usuário.

A análise das medidas estatísticas de dispersão, no entendimento do fenômeno da variabilidade, é importante para identificar e caracterizar a variabilidade das velocidades, mas não é suficiente para estabelecer uma base quantitativa do impacto, principalmente da precipitação, em um nível de detalhes que permita intervenções que beneficiem os usuários do sistema de transporte. A incorporação de técnicas de árvores de classificação permite aprofundar significativamente o entendimento da questão, graças à identificação de classes distintas e distribuídas em diferentes graus de influência do fenômeno estudado. Em termos práticos, milhares de microviagens de baixíssima velocidade causam, do ponto de vista do usuário, impacto muito maior do que milhões de viagens feitas regularmente dentro de certa previsibilidade.

A previsibilidade da variabilidade é uma questão controversa. É consenso entre os autores que, por meio da análise de medidas de dispersão – como a variância, o desvio padrão, o coeficiente de variação e a média em eventos de precipitação –, seja possível estabelecer uma relação de previsibilidade em tempos de viagem, e, por conseguinte, em velocidades desenvolvidas. No entanto, do ponto de vista do usuário, há pouco valor agregado em ter a certeza de que sua viagem demorará uma hora a mais que a de praxe, quando uma precipitação moderada ocorre. É preciso estabelecer métodos concretos de intervenção em campo, de forma a, efetivamente, reduzir a variabilidade das velocidades. Nesse ponto, é inserida a dimensão geográfica ao contexto. A variabilidade é cíclica, bastante previsível, e, depois do comportamento humano, seu maior vetor é a geografia. Há grande utilização das vias expressas, como as rodovias Dutra e Bandeirantes, bem como as Marginais, pelos ônibus de transporte público. Neste contexto os ônibus podem ser afetados pelas condições destas vias bem como, dado o volume de veículos circulantes, afetar sua fluidez, principalmente, em situações críticas como os horários de pico.

Embora não tenha sido investigado, é possível que essas vias sejam utilizadas em trechos curtos, pelos ônibus, objetivando, através do desenvolvimento de velocidades mais altas, a redução dos tempos de viagem. O uso dessas vias provavelmente causa um efeito contrário ao desejado, especialmente diante de eventos de precipitação.

A combinação das técnicas de estatística descritiva, CHAID e geoprocessamento estabelece-se como uma abordagem útil na mineração de dados de transporte, permitindo a elucidação de questões quanto à variabilidade das velocidades dos ônibus e do impacto da precipitação na sua operação, através da identificação de horários e locais críticos, passíveis de intervenção operacional na gestão de transporte em São Paulo.

Todos os objetivos estabelecidos, destacando-se a caracterização dos padrões de operação através de suas velocidades, impacto do clima em sua variabilidade e a aplicação de métodos espaciais no aprofundamento das suas características foram alcançados.

A partir destes resultados, espera-se que o esforço empreendido possa contribuir nas ações de planejamento e operação, não apenas dos ônibus de transporte público,

mas da estrutura de transporte como um todo. Para isso, atividades acadêmicas futuras expandindo o escopo aqui apresentado devem ser empreendidas, contemplando séries históricas e toda a extensão do município de São Paulo, objetivando a criação de um grande repositório de dados com aplicações disponíveis na internet para uso do público.

## 6 REFERÊNCIAS

AKIN, D.; SISIOPIKU, V. P.; SKABARDONIS, A. Impacts of weather on traffic flow characteristics of urban freeways in Istanbul. **Procedia - Social and Behavioral Sciences**, v. 16, p. 89–99, 2011.

ASSIS, F. N. DE; ARRUDA, H. V. DE; PEREIRA, A. R. DE. Aplicações de estatística à climatologia: teoria e prática. **Pelotas: UFPel**, 1996.

BLACK, W. R.; THOMAS, I. Accidents on Belgium's motorways: a network autocorrelation analysis. **Journal of Transport Geography**, v. 6, n. 1, p. 23–31, 1998.

BREIMAN, L. et al. **Classification and regression trees**. [s.l.] CRC press, 1984.

BRUNSDON, C.; FOTHERINGHAM, S.; CHARLTON, M. Geographically weighted regression. **Journal of the Royal Statistical Society: Series D (The Statistician)**, v. 47, n. 3, p. 431–443, 1998.

CARRION, C.; LEVINSON, D. Value of travel time reliability: A review of current evidence. **Transportation Research Part A: Policy and Practice**, v. 46, n. 4, p. 720–741, 2012.

CASAS, I. Evaluating the importance of accessibility to congestion response using a GIS-based travel simulator. **Journal of Geographical Systems**, v. 5, n. 1, p. 109–127, 2003.

CHANGNON, S. A. Effects of summer precipitation on urban transportation. **Climatic Change**, v. 32, n. 4, p. 481–494, 1996.

CHEN, C.; SKABARDONIS, A.; VARAIYA, P. Travel-time reliability as a measure of service. **Transportation Research Record: Journal of the Transportation Research Board**, n. 1855, p. 74–79, 2003.

CHIEN, S. I.; KOLLURI, K. K. Evaluation of freeway travel time variability and reliability under adverse weather with TRANSMIT data. **Journal of Civil Engineering and Architecture**, v. 6, n. 1, p. 1, 2012.

CRESSIE, N. The origins of kriging. **Mathematical Geology**, v. 22, n. 3, p. 239–252, 1990.

EOM, J. et al. Improving the prediction of annual average daily traffic for nonfreeway facilities by applying a spatial statistical method. **Transportation Research Record: Journal of the Transportation Research Board**, n. 1968, p. 20–29, 2006.

ESRI. **Empiric Baesian Kriging**. Disponível em: <<http://desktop.arcgis.com/en/arcmap/10.3/guide-books/extensions/geostatistical-analyst/what-is-empirical-bayesian-kriging-.htm>>. Acesso em: 16 jan. 2017.

ESRI. ArcGIS Data Classification Methods. 2017.

FRANKLIN, J.; KARLSTROM, A. Travel time reliability for Stockholm roadways:



modeling mean lateness factor. **Transportation Research Record: Journal of the Transportation Research Board**, n. 2134, p. 106–113, 2009.

GUO, Z.; WILSON, N.; RAHBEE, A. Impact of weather on transit ridership in Chicago, Illinois. **Transportation Research Record: Journal of the Transportation Research Board**, n. 2034, p. 3–10, 2007.

HAYNES, R. et al. Validation of travel times to hospital estimated by GIS. **International Journal of Health Geographics**, v. 5, p. 40, 2006.

HOSHINA, L. N. N.; CHIOVETTI, P. B.; DELUCA, R. S. **Estudo de Viabilidade e Impactos da Aplicação de Ferramentas ITS em Faixas e Corredores de Ônibus**. [s.l.: s.n.].

HRANAC, R. et al. Empirical studies on traffic flow in inclement weather. 2006.

INMET. **SADMET**. Disponível em: <[http://www.inmet.gov.br/portal/index.php?r=home/page&page=central\\_servicos](http://www.inmet.gov.br/portal/index.php?r=home/page&page=central_servicos)>.

KAMGA, C.; YAZICI, M. A. Temporal and weather related variation patterns of urban travel time: Considerations and caveats for value of travel time, value of variability, and mode choice studies. **Transportation Research Part C: Emerging Technologies**, v. 45, p. 4–16, 2014.

KASHFI, S. A.; BUNKER, J. M.; YIGITCANLAR, T. Modelling and analysing effects of complex seasonality and weather on an area's daily transit ridership rate. **Journal of Transport Geography**, v. 54, p. 310–324, 2016.

KASS, G. V. An Exploratory Technique for Investigating Large Quantities of Categorical Data. **Applied Statistics**, v. 29, n. 2, p. 119, 1980.

KHATTAK, A. J.; DE PALMA, A. The impact of adverse weather conditions on the propensity to change travel decisions: A survey of Brussels commuters. **Transportation Research Part A: Policy and Practice**, v. 31, n. 3, p. 181–203, 1997a.

KHATTAK, A. J.; DE PALMA, A. The impact of adverse weather conditions on the propensity to change travel decisions: A survey of Brussels commuters. **Transportation Research Part A: Policy and Practice**, v. 31, n. 3, p. 181–203, 1997b.

KWON, J. et al. Decomposition of travel time reliability into various sources: incidents, weather, work zones, special events, and base capacity. **Transportation Research Record: Journal of the Transportation Research Board**, n. 2229, p. 28–33, 2011.

LAMM, R.; CHOUEIRI, E. M.; MAILAENDER, T. Comparison of operating speeds on dry and wet pavements of two-lane rural highways. **Transportation Research Record**, v. 1280, n. 8, p. 199–207, 1990.

LIU, H. X.; RECKER, W.; CHEN, A. Uncovering the contribution of travel time reliability to dynamic route choice using real-time loop data. **Transportation Research Part A: Policy and Practice**, v. 38, n. 6, p. 435–453, 2004.

LOWRY, M. Spatial interpolation of traffic counts based on origin-destination centrality. **Journal of Transport Geography**, v. 36, p. 98–105, 2014.

LYMAN, K.; BERTINI, R. Using travel time reliability measures to improve

regional transportation planning and operations. **Transportation Research Record: Journal of the Transportation Research Board**, n. 2046, p. 1–10, 2008.

MARTCHOUK, M.; MANNERING, F. L.; SINGH, L. Travel Time Reliability in Indiana Final Report. 2010.

MORENCY, C. et al. Distance traveled in three Canadian cities: Spatial analysis from the perspective of vulnerable population segments. **Journal of Transport Geography**, v. 19, n. 1, p. 39–50, 2011.

MORGAN, @; SONQUIST, @. Problems in the analysis of survey data and a proposal. **Journal of the American Statistical Association**, v. 58, n. 302, p. 415–434, 1963.

PAEZ, A. Exploring contextual variations in land use and transport analysis using a probit model with geographical weights. **Journal of Transport Geography**, v. 14, n. 3, p. 167–176, 2006.

PERON, L. **Contribuição Metodológica para a aplicação de Prioridade Semafórica Condicional em Corredores de Ônibus**. [s.l.] EPUSP, 2015.

QUDDUS, M. Exploring the Relationship Between Average Speed, Speed Variation, and Accident Rates Using Spatial Statistical Models and GIS. **Journal of Transportation Safety & Security**, v. 5, n. 1, p. 27–45, 2013.

SELBY, B.; KOCKELMAN, K. M. Spatial prediction of traffic levels in unmeasured locations: applications of universal kriging and geographically weighted regression. **Journal of Transport Geography**, v. 29, p. 24–32, 2013.

SHIMBEL, A. Structural parameters of communication networks. **The Bulletin of Mathematical Biophysics**, v. 15, n. 4, p. 501–507, 1953.

SHUMAKER, B. P.; SINNOTT, R. W. Astronomical computing: 1. Computing under the open sky. 2. Virtues of the haversine. **Sky and telescope**, v. 68, p. 158–159, 1984.

SKABARDONIS, A.; VARAIYA, P.; PETTY, K. Measuring Recurrent and Nonrecurrent Traffic Congestion. **Transportation Research Record: Journal of the Transportation Research Board**, v. 1856, p. 118–124, 2003.

SMDU, S. M. DE D. U.-. **Mapas da Cidade - MDC**. Disponível em: <[http://www.prefeitura.sp.gov.br/cidade/secretarias/desenvolvimento\\_urbano/dados\\_estatisticos/index.php?p=160798](http://www.prefeitura.sp.gov.br/cidade/secretarias/desenvolvimento_urbano/dados_estatisticos/index.php?p=160798)>. Acesso em: 20 jul. 2017.

SMITH, B. L. et al. **An investigation into the impact of rainfall on freeway traffic flow**. 83rd annual meeting of the Transportation Research Board, Washington DC. **Anais...Citeseer**, 2004.

SPTRANS. **Passageiros Transportados Anualmente**. Disponível em: <[http://www.sptrans.com.br/indicadores/historico\\_passageiros-transportados.aspx](http://www.sptrans.com.br/indicadores/historico_passageiros-transportados.aspx)>. Acesso em: 8 mar. 2017.

STOVER, V. W.; MCCORMACK, E. D. The impact of weather on bus ridership in Pierce County, Washington. **Journal of Public Transportation**, v. 15, n. 1, p. 6, 2012.

TAO, S. et al. The influence of weather on local geographical patterns of bus usage. **Journal of Transport Geography**, v. 54, p. 66–80, 2016.

TAYLOR, A. J. P. Reliability and Cost Benefit Analysis in Australia and New Zealand. **Institute of Sustainable Systems and Technologies. University of South Australia**, 2009.

TU, H.; VAN LINT, H. W. C.; VAN ZUYLEN, H. J. **Impact of adverse weather on travel time variability of freeway corridors**. Transportation Research Board 86th Annual Meeting. **Anais...**2007

VAN LINT, J. W. C.; VAN ZUYLEN, H. J.; TU, H. Travel time unreliability on freeways: Why measures based on variance tell only half the story. **Transportation Research Part A: Policy and Practice**, v. 42, n. 1, p. 258–277, 2008.

WAARA, N. et al. Feasible provision of targeted traveler information in public transportation: Segmentation based on functional limitations. **Transportation Research Part A: Policy and Practice**, v. 74, p. 164–173, 2015.

YAZICI, M. A.; KAMGA, C.; SINGHAL, A. Weather's impact on travel time and travel time variability in New York City. **Transportation Research**, v. 40, p. 41, 2013.