



**UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE TRANSPORTES
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TRANSPORTES**

PHELIPE DE FIGUEREDO BEDA RABAY

**USO DE DADOS SECUNDÁRIOS DE RASTREAMENTO DA FROTA NA
ETAPA DE CARACTERIZAÇÃO DA PROBLEMÁTICA DO SISTEMA
DE TRANSPORTE PÚBLICO**

FORTALEZA

2017

PHELIPE DE FIGUEREDO BEDA RABAY

USO DE DADOS SECUNDÁRIOS DE RASTREAMENTO DA FROTA NA
ETAPA DE CARACTERIZAÇÃO DA PROBLEMÁTICA DO SISTEMA
DE TRANSPORTE PÚBLICO

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Transportes da Universidade Federal do Ceará, como parte dos requisitos para a obtenção do título de Mestre em Ciências (M. Sc.) em Engenharia de Transportes.

Orientador: Prof. Mário Angelo Nunes de Azevedo Filho, DSc.

FORTALEZA

2017

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

R111u Rabay, Phelipe de Figueredo Beda.

Uso de dados secundários de rastreamento da frota na etapa de caracterização da problemática do sistema de transporte público / Phelipe de Figueredo Beda Rabay. – 2017.
69 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Transportes, Fortaleza, 2017.

Orientação: Prof. Dr. Mário Ângelo Nunes de Azevedo Filho.

1. Transporte Público. 2. Rastreamento da Frota. 3. Ônibus. 4. GPS. I. Título.

CDD 388

PHELIPE DE FIGUEREDO BEDA RABAY

USO DE DADOS SECUNDÁRIOS DE RASTREAMENTO DA FROTA NA ETAPA DE
CARACTERIZAÇÃO DA PROBLEMÁTICA DO SISTEMA DE TRANSPORTE
PÚBLICO

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Transportes da Universidade Federal do Ceará, como parte dos requisitos para a obtenção do título de Mestre em Ciências (M. Sc.) em Engenharia de Transportes. Área de concentração: Planejamento e operação de sistemas de transportes.

Aprovada em 30/11/2017.

BANCA EXAMINADORA

Prof. Dr. Mário Angelo Nunes de Azevedo (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Bruno Vieira Bertoncini
Universidade Federal do Ceará (UFC)

Prof. Dr. Breno Barros Telles do Carmo
Universidade Federal Rural do Semi-Árido (UFERSA)

AGRADECIMENTOS

Aos meus pais, que deram seu melhor para que eu me tornasse o homem que sou hoje.

Ao professor Mário Angelo, que me acolheu e acreditou na minha capacidade e nas minhas ideias, mesmo nos momentos mais difíceis da construção deste trabalho.

Aos professores e professoras do Petran, profissionais excepcionais, verdadeiros exemplos do que é ser apaixonado por educação.

Aos amigos e amigas que me acompanharam durante o mestrado, dentro e fora do ambiente acadêmico, 2JPX e agregados, que fizeram essa jornada mais leve e divertida.

À Prefeitura de Fortaleza, em especial aos colegas do PAITT, por disponibilizar os dados necessários para a realização do trabalho e por me receber sempre de portas abertas para quaisquer esclarecimentos.

À Hiane Braun, por me amar, me inspirar e me acompanhar até o último minuto, me lembrando que tudo iria dar certo. Muito obrigado.

Por fim, a todas e todos que contribuíram direta e indiretamente para que eu finalizasse esta etapa.

RESUMO

Os sistemas automáticos de coleta de dados produzem grande quantidade diária de informações sobre o sistema de transporte público. A armazenagem e tratamento destes dados possibilita análises que contribuem para a compreensão do sistema, gerando informações valiosas para o seu gerenciamento. Dentre essas ferramentas, está o sistema de localização automática de veículos, que produz dados de qualidade e com grande potencial para auxiliar no entendimento do sistema. Muitas vezes, porém, as agências de transporte não exploram todo o potencial desses dados, os utilizando apenas para controle operacional da frota. Observando a importância da compreensão da problemática que envolve um sistema de transporte público e a dificuldade de se obter dados para esse fim, este trabalho se propõe a utilizar os dados disponíveis, de localização da frota de ônibus urbanos da cidade de Fortaleza, para verificar sua capacidade de contribuir para a caracterização do sistema, etapa fundamental no processo de planejamento. Assim, o objetivo principal deste trabalho é propor uma metodologia de adequação desses dados para produzir informações relevantes para a compreensão do funcionamento do sistema de transporte público. A etapa de adequação dos dados soluciona problemas que podem ser comuns aos dados de várias cidades, como a identificação dos veículos, linhas e sentido das viagens. A metodologia utiliza o cruzamento de dados geográficos e de bilhetagem eletrônica para superar o problema da qualidade dos dados disponíveis, no que se refere à localização da frota. Foi possível, partindo de uma base de dados aparentemente incompleta, e com algumas informações inconsistentes, determinar os tempos de viagem, a velocidade individual e média dos veículos de uma determinada linha, em qualquer trecho do itinerário, o intervalo entre veículos e a formação de comboios. Os resultados possibilitaram importantes análises acerca do desempenho e da qualidade do sistema.

Palavras Chave: transporte público, ônibus, GPS, AVL, rastreamento da frota

ABSTRACT

Automatic data collection systems produce a large amount of information about the public transport system. The storage and treatment of this data enables analysis that contributes to an understanding of the system, generating valuable information for its management. Among these tools, is the Automatic Vehicle Location (AVL) System, that produces data of quality and with great potential to aid in the understanding of the system. However, it is observed that transport agencies do not exploit the full potential of the data. They use it just for vehicles operation control. Considering the importance of understanding the transit problems and the difficulty of obtaining data for this purpose, this work proposes the analysis of the use of the available fleet location data in the city of Fortaleza. The idea is the preparation of a better characterization of the system, a fundamental step in the planning process. Thus, the main objective of this work is to propose a methodology to turn this data in to relevant information for an understanding of the public transportation system. The first stage of data manipulation solves problems that can be the same in other cities, such as the identification of the vehicles, routes and travel direction. The method uses the crossing of geographical and automatic fare collection data to overcome some problems of specific fleet location. Starting from an apparently incomplete database with some inconsistent information, it was possible to determine the travel times, the individual and average vehicles' speed, of a certain line in any part of the itinerary, the interval between vehicles and bunch formation. The results allowed important analysis about the performance and quality of the system.

Keywords: transit, bus, GPS, GTFS, fleet tracking, AVL

LISTA DE FIGURAS

Figura 1:	Fluxograma das etapas do processo de planejamento	20
Figura 2:	Velocidades operacionais da linha 112 em segmentos ao longo do itinerário	23
Figura 3:	Velocidades operacionais de linhas ao longo do dia	24
Figura 4:	Comparação das velocidades operacionais de linhas beneficiadas com faixas exclusivas em 2013 e 2015	24
Figura 5:	Comparação das velocidades operacionais de linhas não beneficiadas com faixas exclusivas em 2013 e 2015	25
Figura 6:	Velocidades operacionais em vias de Fortaleza	25
Figura 7:	Diagrama dos dados do <i>feed</i> GTFS	27
Figura 8:	Etapas do processo de adequação dos dados	31
Figura 9:	Registros de GPS da linha 079, antes do filtro espacial	40
Figura 10:	Registros de GPS da linha 079, após a aplicação do filtro espacial.	41
Figura 11:	Gráfico de marcha individual, com dados indesejados	45
Figura 12:	Gráfico de marcha individual, após a limpeza dos dados.....	45
Figura 13:	Gráfico de marcha total com dados indesejados.....	46
Figura 14:	Gráfico de marcha total após a limpeza dos dados.....	46
Figura 15:	Headways no final do trecho 11 da linha 079.....	50
Figura 16:	Gráfico de marcha da linha 079.....	51

LISTA DE TABELAS

Tabela 1: Amostra dos dados do SBE	34
Tabela 2: Formato inicial dos dados do AVL.....	36
Tabela 3: Dados de AVL após o tratamento inicial.....	37
Tabela 4: Variável dist_traveled no arquivo shapes	38
Tabela 5: Amostra do dicionário de veículos	39
Tabela 6: Dados de bilhetagem da linha 079.....	39
Tabela 7: Dados do AVL da linha 079	39
Tabela 8: Correção dos valores inconsistentes de ida e volta.....	43
Tabela 9: Identificação das viagens	44
Tabela 10: Velocidade de operação individual por trechos.....	47
Tabela 11: Velocidade de operação da linha 079 por trecho e horários.....	48
Tabela 12: Velocidade de operação da linha 026 por trecho e horários.....	49

SUMÁRIO

1	INTRODUÇÃO.....	15
1.1	Problema de Pesquisa.....	16
1.2	Objetivos.....	17
1.3	Estrutura da Dissertação	18
2	PLANEJAMENTO DO SISTEMA DE TRANSPORTE PÚBLICO.....	19
2.1	Processo de Planejamento.....	19
2.2	Sistemas de informações.....	21
2.2.1	<i>Rastreamento da Frota – AVL</i>	<i>21</i>
2.2.2	<i>Feed GTFS.....</i>	<i>26</i>
2.2.3	<i>Bilhetagem Eletrônica</i>	<i>28</i>
3	METODOLOGIA.....	30
3.1	Formato dos dados.....	30
3.2	Indicadores obtidos.....	32
3.2.1	<i>Velocidades.....</i>	<i>32</i>
3.2.2	<i>Tempo de espera.....</i>	<i>33</i>
3.2.3	<i>Formação de comboios.....</i>	<i>33</i>
4	ESTUDO DE CASO	34
4.1	Tratamento inicial do banco de dados	36
4.1.1	<i>Redução do tamanho do banco de dados</i>	<i>36</i>
4.1.2	<i>Inserindo a variável dist_traveled no feed GTFS.....</i>	<i>37</i>
4.2	Identificação inicial da linha.....	38
4.2.1	<i>Busca dos veículos no sistema de bilhetagem.....</i>	<i>38</i>
4.2.2	<i>Eliminação dos pontos fora do itinerário</i>	<i>40</i>
4.3	Determinação da distância percorrida e sentido do deslocamento	41
4.3.1	<i>Atribuição inicial do dist_traveled</i>	<i>41</i>
4.3.2	<i>Atribuição IDA/VOLTA</i>	<i>42</i>
4.3.3	<i>Ajuste de dist_traveled</i>	<i>42</i>
4.4	Identificação das viagens.....	43
4.4.1	<i>Avaliação do dist_traveled.....</i>	<i>43</i>
4.4.2	<i>Identificação das viagens efetivas</i>	<i>44</i>
4.6	Determinação das velocidades	47
4.7	Determinação dos tempos de espera	50
4.8	Identificação da formação de comboios.....	50

5	CONCLUSÕES E RECOMENDAÇÕES.....	52
	REFERÊNCIAS.....	54
	APÊNDICE A: MAPAS DE CALOR DAS VELOCIDADES OPERACIONAIS DAS LINHAS.....	57
	APÊNDICE B: GRÁFICOS DE MARCHA DAS LINHAS.....	59

1 INTRODUÇÃO

O Sistema de Transporte Público de Passageiros (STPP) é importante para o funcionamento de uma cidade, pois garante a uma boa parcela da sociedade o deslocamento necessário para acessar serviços e desenvolver suas atividades. No Brasil, estima-se que o STPP seja utilizado em cerca de 30% das viagens diárias e, em muitas cidades, represente mais da metade das viagens motorizadas (ASSOCIAÇÃO NACIONAL DE TRANSPORTES PÚBLICOS, 2014).

O STPP promove o uso racional do espaço urbano e dos recursos naturais, pois demanda de 6 a 15 vezes menos espaço por passageiro transportado do que o transporte motorizado individual (CONFEDERAÇÃO NACIONAL DOS TRANSPORTES, 2002). Além disso, ele democratiza o acesso à cidade, oferecendo um meio de transporte de longa distância, a um custo acessível à maior parte da população, principalmente aquela que não tem condições de custear transporte individual. A maioria dos usuários do STPP depende exclusivamente desse sistema para realizar as suas viagens, tornando-o um serviço essencial e influenciando a qualidade de vida dessas pessoas.

Neste contexto, a lei da Política Nacional de Mobilidade Urbana (PNMU) (BRASIL, 2012) institui uma série de diretrizes para o desenvolvimento urbano, incluindo a questão do planejamento de transportes, e determina que os transportes não motorizados e os transportes públicos coletivos devem ter prioridade sobre os transportes motorizados e individuais.

Segundo de Arroxelas Galvão (1996), a partir da segunda metade do século XX, o modelo de desenvolvimento adotado priorizou o modo rodoviário de transporte, e, desde então, são aplicadas diversas políticas de incentivo à indústria automobilística, desde oferta de infraestrutura a incentivos econômicos, onde o crescimento desta passou a ser fator de desenvolvimento econômico. Esse histórico, que moldou as cidades às necessidades do veículo particular, torna a priorização dos modos mais eficientes de transporte um desafio ainda maior.

Para propor a solução mais adequada a uma situação específica, é necessário entender o problema e suas causas, assim como os impactos que as propostas podem ter caso sejam aplicadas. Não ter conhecimento de quais são as possíveis causas do problema pode trazer propostas de soluções que produzam resultados negativos ao sistema, ou a má utilização de recursos. Uma análise aprofundada sobre o tipo de problema, assim como suas

causas e efeitos, é a ação fundamental para a resolução dos problemas, pois isto faz parte do processo de decisão (GARCIA, 2016).

Segundo Soares (2014), o processo de planejamento consiste em identificar e diagnosticar problemas e direcionar recursos para solucionar situações indesejadas. Esse processo pode ser resumido em duas grandes fases: compreensão da problemática e proposição de soluções. A proposição de soluções adequadas e justas para a problemática de um sistema só pode ocorrer através do seu profundo entendimento, por meio da identificação, mensuração e validação dos reais problemas a serem solucionados. Assim, considerar o entendimento da problemática como base fundamental para a negociação dos interesses conflitantes das partes, envolvidas e afetadas por um sistema em análise, é crucial à tomada de decisão consciente.

Dentro da fase de compreensão da problemática, que será abordada mais profundamente adiante no trabalho, existe uma etapa de caracterização, que objetiva a proposição de indicadores que descrevam e mensurem a atual situação do sistema. Tais indicadores são compostos por variáveis que, por sua vez, são alimentadas por dados, coletados por diversos métodos.

Por muito tempo, a única fonte de informações para se entender o funcionamento dos sistemas de transportes foi o levantamento de dados pelos métodos tradicionais de pesquisa e observações de campo, tais como pesquisas de embarque e desembarque e anotações em cadernetas de cumprimento de viagens e horários e postos de controle ao longo da rota.

Nas duas últimas décadas, porém, está em crescimento o uso de tecnologias automáticas de coleta de dados, que serão descritas mais detalhadamente no capítulo 2 do trabalho. Dentre estas tecnologias, está o rastreamento automático da frota (do inglês *Automatic Vehicle Location* – AVL), que gera dados de localização dos veículos e diversas informações com bastante potencial para descrever o funcionamento do sistema (CORTÉS *et al.*, 2010). O presente trabalho trata do uso deste tipo de dado na etapa de caracterização de um sistema de transporte público.

1.1 Problema de Pesquisa

O levantamento de dados pelos métodos tradicionais de observações de campo pode se mostrar bastante oneroso e demandar bastante tempo para obtenção dos resultados, limitando a caracterização do sistema.

Por outro lado, o uso dos dados de técnicas automáticas de coleta de dados como o AVL está cada vez mais difundido nas grandes cidades brasileiras. O potencial dos dados, por sua vez, não é totalmente explorado pela maioria das agências de transporte. Os dados são usados, sobretudo, para monitoramento da frota durante a fase de operação, verificação da posição do veículo no momento de algum problema relatado e checagem do cumprimento de viagens. No entanto, os dados coletados pelo AVL podem ser empregados na geração de indicadores e variáveis úteis na etapa de caracterização do STPP, tais como: velocidades de operação agregada por linha ou trechos, tempos de espera entre um veículo e outro da mesma linha em determinado ponto do trajeto, formação de comboios e a variação desses dados ao longo do dia. A partir da caracterização das vias, pode-se verificar a influência do tipo de via e do uso do solo no desempenho do sistema e gerar subsídios importantes para a tomada de decisão.

Em cidades onde o AVL já está difundido, seu uso no esforço de caracterização do STPP pode apresentar um baixo custo marginal de levantamento de dados, pois é um sistema automático de coleta já em funcionamento, embora com outra finalidade. Como, na maioria das vezes, serão utilizados dados secundários, isso demanda um esforço para automatizar o armazenamento e organização dos dados.

1.2 Objetivos

Observado o problema em questão, aponta-se como objetivo principal desta pesquisa propor uma metodologia para utilização de dados secundários de AVL na etapa de caracterização da problemática do Sistema de Transporte Público de Passageiros.

Para atingir este objetivo, são traçados os seguintes objetivos específicos:

- Analisar o papel do AVL no planejamento de um sistema de transportes e suas principais aplicações;
- Determinar indicadores que possam ser coletados a partir de dados de AVL;
- Propor um método para obtenção destes indicadores a partir de uma base de dados existente;
- Verificar se os dados de AVL fornecem informações relevantes para a caracterização de um STPP.

1.3 Estrutura da Dissertação

O presente trabalho está estruturado da seguinte forma:

Capítulo 2: aborda aspectos teóricos sobre o tema, com uma visão geral sobre o processo de planejamento, e dá ênfase ao uso de tecnologias auxiliares neste processo, em especial o rastreamento da frota, apresentando os principais trabalhos que utilizam esses dados no processo de planejamento;

Capítulo 3: detalha a proposta metodológica para o uso dos dados de rastreamento no processo de caracterização do sistema de transportes, especificando o formato dos dados e os indicadores obtidos por meio da sua análise;

Capítulo 4: aplica a metodologia no estudo de caso da cidade de Fortaleza e apresenta os resultados obtidos;

Capítulo 5: apresenta conclusões sobre o trabalho realizado e recomendações para trabalhos futuros.

2 PLANEJAMENTO DO SISTEMA DE TRANSPORTE PÚBLICO

Os sistemas de transportes são compostos não apenas pelos elementos que interagem entre si para produzir oportunidades de deslocamentos, mas também pela demanda que aproveita essas oportunidades para realizar viagens. Isso os torna sistemas complexos, que contém elementos de natureza técnica, como veículos e infraestrutura, que seguem leis físicas, tradicionalmente estudadas pela engenharia, bem como elementos fortemente influenciados pela natureza humana, como a demanda e a operação, o que exige uma abordagem diferente para sua compreensão, mais ligada às ciências sociais e econômicas.

Além disso, os projetos de transportes podem ter implicações para a economia, o meio ambiente, a qualidade de vida e as atividades de uma determinada região. Uma influência sobre interesses muitas vezes conflitantes, o que torna ainda mais importante o esforço de compreensão e análise dos problemas e impactos de projetos e sistemas de transportes.

Segundo Cascetta (2009), um sistema de transporte pode ser definido como um conjunto de elementos e suas interações, que produzem tanto a demanda por viagens, dentro de uma determinada área, quanto a oferta para satisfazer esta demanda. Como a maioria dos componentes de um sistema social e econômico interage em diferentes níveis de intensidade, uma abordagem geral da engenharia de sistemas é isolar os elementos mais relevantes para um problema em questão e agrupar esses elementos e suas relações dentro do sistema de análise.

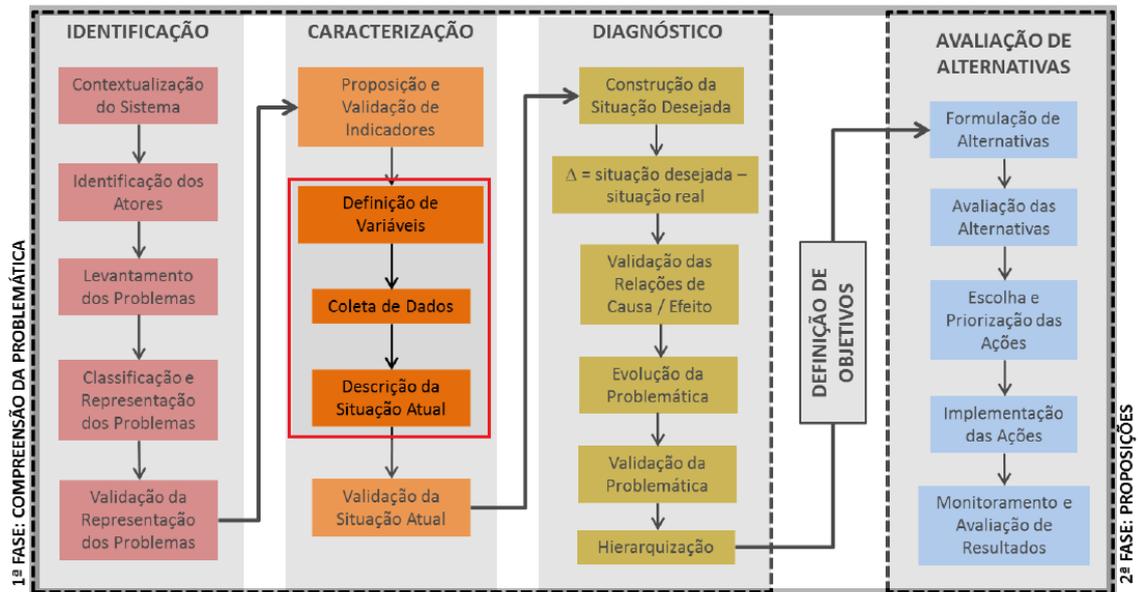
2.1 Processo de Planejamento

A Figura 1 apresenta o fluxograma que sistematiza as etapas do método proposto por Soares (2014) para o processo de planejamento. Como se pode observar, a fase de compreensão da problemática, mencionada no capítulo introdutório deste trabalho, é dividida em três outras etapas: identificação, caracterização e diagnóstico.

A fase de identificação procura estabelecer uma contextualização do sistema, através da identificação dos atores envolvidos e da classificação e representação dos problemas por eles percebidos. A fase de caracterização, por sua vez, busca expressar analiticamente os problemas identificados através de suas características, propondo e validando indicadores que os representem, resultando em uma descrição da situação atual do fenômeno analisado. É nesta fase que está o foco do presente trabalho, onde ocorre a definição de variáveis, a coleta e manipulação de dados para compor os indicadores desejados.

Por fim, a fase de diagnóstico tem como objetivo levantar as informações necessárias para construir a situação desejada, possibilitando o estabelecimento de uma comparação entre a situação desejada e a atual.

Figura 1: Fluxograma das etapas do processo de planejamento



Fonte: Adaptado de Soares (2014)

A proposição e formação de indicadores são fundamentais no processo de tomada de decisão, pois, por meio deles, os tomadores de decisão são capazes de analisar o desempenho, a eficácia e sintetizar o funcionamento do sistema. Suas principais funções são fornecer informações sobre os problemas levantados e contribuir para a previsão e o acompanhamento de ações. Assim, um indicador deve ser capaz de descrever um fenômeno ou parte dele baseando-se em dados confiáveis. Chapleau *et al.* (2011) apontam quatro fontes importantes de informação passiva:

- AFC (*Automated Fare Collection*): Sistema de Bilhetagem Eletrônica;
- APC (*Automatic Passenger Counting*): Contagem Automática de Passageiros;
- AVL (*Automatic Vehicle Location*): Localização Automática de Veículos;
- GIS (*Geographic Information System*): Sistema de Informações Geográficas.

Este trabalho propõe uma metodologia com o uso de três destas fontes, o rastreamento da frota (AVL), a bilhetagem eletrônica (AFC) e, também, os *feeds* GTFS – *General Transit Feed Specification*, o que pode ser traduzido para Especificação Geral de

Alimentação de Dados de Transporte Público. Estes *feeds* são constituídos de um conjunto de arquivos contendo informações sobre os itinerários, horários e viagens do transporte público.

2.2 Sistemas de informações

A evolução tecnológica e a facilidade e barateamento do armazenamento de dados têm proporcionado meios extremamente eficientes de utilização de bases de dados, gerando informações confiáveis com custos cada vez mais baixos. Esta seção apresenta os principais sistemas de informação e coleta de dados utilizados neste trabalho.

2.2.1 Rastreamento da Frota – AVL

As tecnologias de localização de frota estão sendo cada vez mais utilizados para a avaliação de serviços de ônibus urbanos. Okunieff (1997) desenvolveu um trabalho de revisão sobre vários sistemas automáticos de localização de veículos (AVL – *Automatic Vehicle Location*) que incluíam postos de controle, comunicação por rádio, leitura de odômetro e, também, localização utilizando GPS. Dessouky *et al.* (1999) avaliou o uso de AVL para determinar os tempos de chegadas em estações de transbordo de transporte público em Los Angeles. Os autores identificaram a localização por GPS como a tecnologia mais vantajosa, por não requerer estrutura física ao longo das rotas e funcionar, praticamente, em qualquer lugar. Dentre as desvantagens, foi citado que o sinal de GPS não alcança lugares cobertos, como túneis, onde deve ser utilizada uma tecnologia complementar.

A agência de transporte público de Portland (TriMet) começou a usar um sistema automatizado de localização de ônibus para gerenciar e coletar dados sobre o desempenho de sua frota de ônibus no final da década de 1990. Esses dados fornecem ao TriMet uma abundância de informações úteis que vêm sendo usadas para melhorar o desempenho e a eficiência de seu sistema de transporte público (BERKOW *et al.*, 2007).

O estudo do uso de dados de GPS nos transportes se divide em três linhas de pesquisa. Primeiro, tem-se o estudo do uso de dados *online* para controle operacional e de informação ao usuário. Segundo, tem-se o uso das velocidades dos ônibus para estimar o tempo de viagem do tráfego misto. Em terceiro, o uso de dados consolidados de GPS podem ser utilizados para planejamento de horários e melhoria na qualidade do serviço (Cortés *et al.*, 2010).

Sobre a primeira linha, os dados em tempo real são usados em estratégias de controle, sendo as mais utilizadas as do tipo *holding* e *stop-skipping*, onde o operador procura

padronizar os *headways* (intervalos entre veículos), fazendo com que determinados veículos atrasem ou ignorem pontos de parada respectivamente. Cortés *et al.* (2010) desenvolveram métodos utilizando as duas estratégias.

Sobre a segunda linha de estudo, Berkow *et al.* (2007) discutiram o uso de dados de transporte público para avaliar o desempenho de vias arteriais. Pu e Lin (2008) avaliaram vários estudos sobre o uso desses dados na estimação de velocidades do tráfego misto e chegaram à conclusão que, apesar das dificuldades no processamento e nas incertezas do transporte público, os dados são confiáveis para estimar o desempenho de vias arteriais.

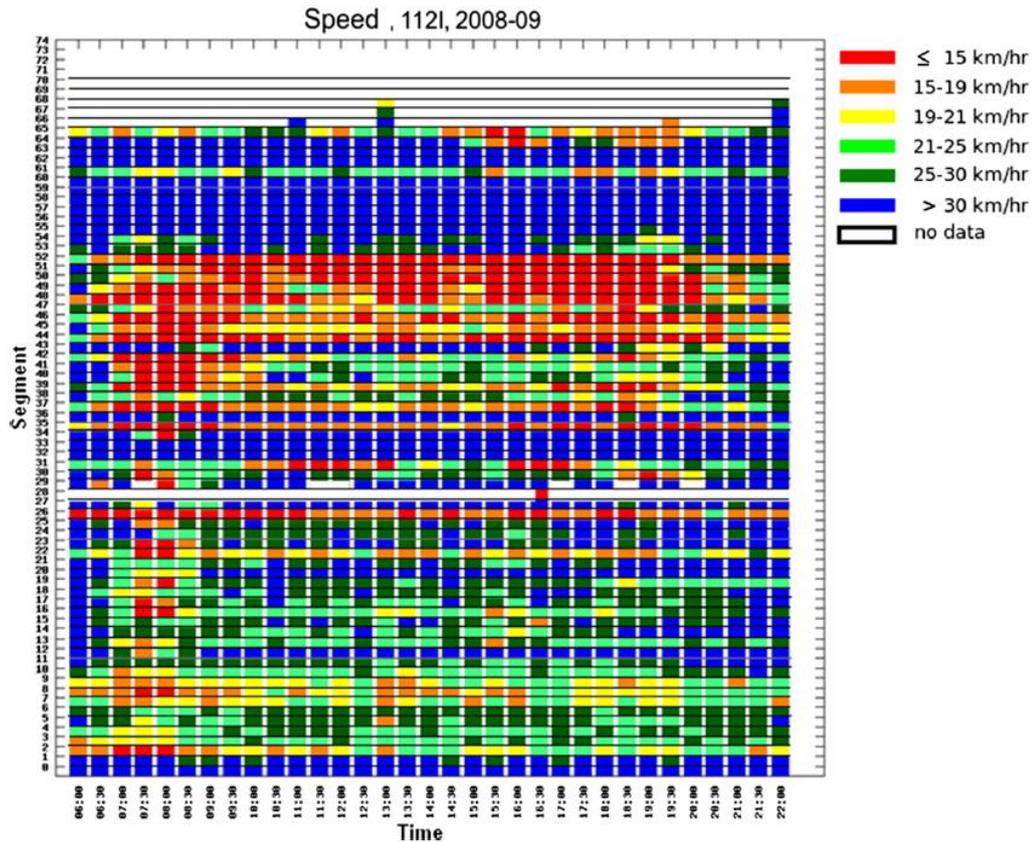
A terceira linha de pesquisa, na qual se encontra o presente trabalho, tem crescido nos últimos anos. Chakroborty e Kikuchi (2004) investigaram a previsão dos tempos de viagem em corredores urbanos a partir da utilização de dados de aparelhos de GPS instalados em ônibus. Foi relatado que a diferença nos tempos de viagem foi estável, o que possibilitou a elaboração de modelos de previsão dos tempos de viagem.

Outra característica importante a ser observada no sistema é a variabilidade dos tempos de viagem dentro de uma mesma linha ou trecho. Nesse sentido, Ye *et al.* (2016) observaram a variabilidade dos tempos de viagem dos ônibus em um corredor e elaboraram modelos de regressão para identificar os principais fatores que causam esta variabilidade. Fatores como uso do solo, extensão da rota e quantidade de semáforos podem afetar bastante essa variabilidade. Também foi analisada a influência da existência de faixas exclusivas, que são eficientes em reduzir a variabilidade dos tempos de viagem.

Cortés *et al.* (2010) avaliaram o sistema de transporte por ônibus de Santiago, no Chile, por meio da utilização de dados de GPS gerados pelos mais de 6000 veículos da frota, distribuídos em 700 diferentes linhas. Os autores relataram o desafio de se trabalhar com um volume muito grande de dados de uma forma sistemática. Assim, a metodologia desenvolvida pode ser aplicada em cada linha ou segmentos determinados pelo analista. O estudo aplicou também uma metodologia para avaliar a consistência dos dados, pois os aparelhos de GPS emitem uma localização instantânea que não necessariamente se encontra em uma rota pré-estabelecida, sendo necessário corrigir os dados pontuais para os itinerários oficiais. Os autores relataram que as informações fornecidas pelos registros GPS podem ser traduzidas em indicadores confiáveis que permitem, aos tomadores de decisão, monitorar as velocidades comerciais dos ônibus nas redes de transportes, independentemente de seu tamanho. A Figura 2 apresenta as velocidades operacionais da linha 112 (ida), coloridas de acordo com a eficiência na operação. O eixo x representa os horários, a partir das 6h da manhã, com intervalos de 30 minutos, enquanto o eixo y representa o número do segmento dentro do

itinerário. É possível observar uma baixa velocidade nos trechos 45 a 52, seguida de altas velocidades nos trechos 54 a 60. A Figura 3 apresenta um comparativo entre várias linhas, atribuindo um parâmetro de desempenho relativo a uma velocidade de referência para cada horário. Dessa maneira, consegue identificar as linhas que operam abaixo do esperado (valores maiores que 1,0).

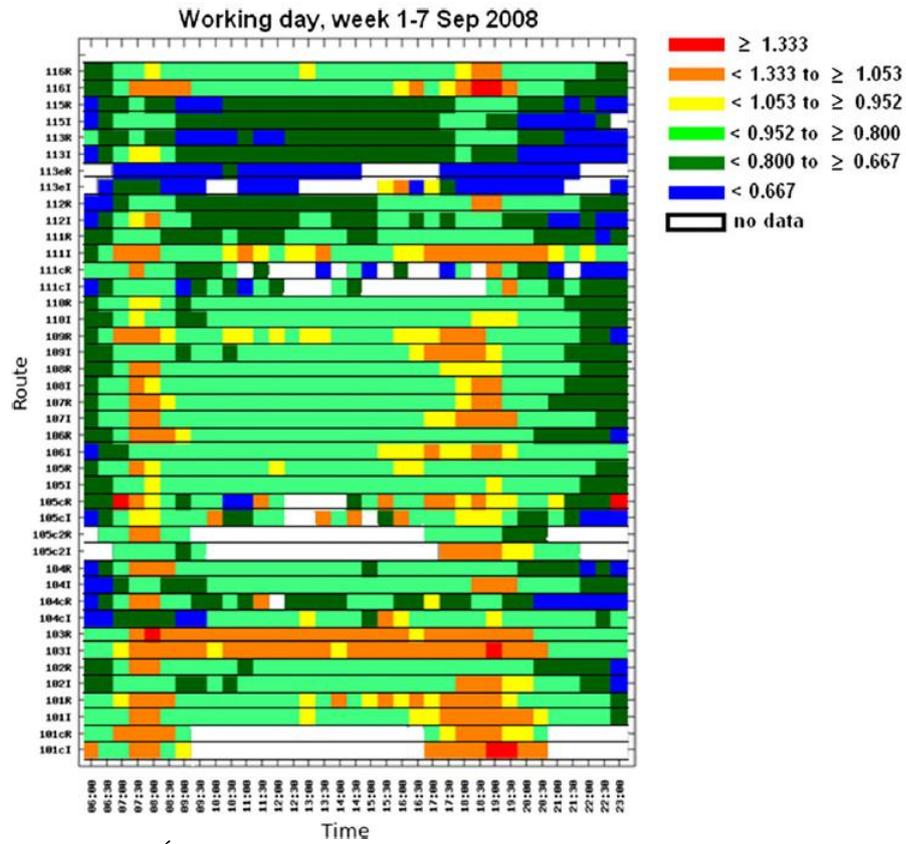
Figura 2: Velocidades operacionais da linha 112 em segmentos ao longo do itinerário



Fonte: (CORTÉZ *et al.*, 2011)

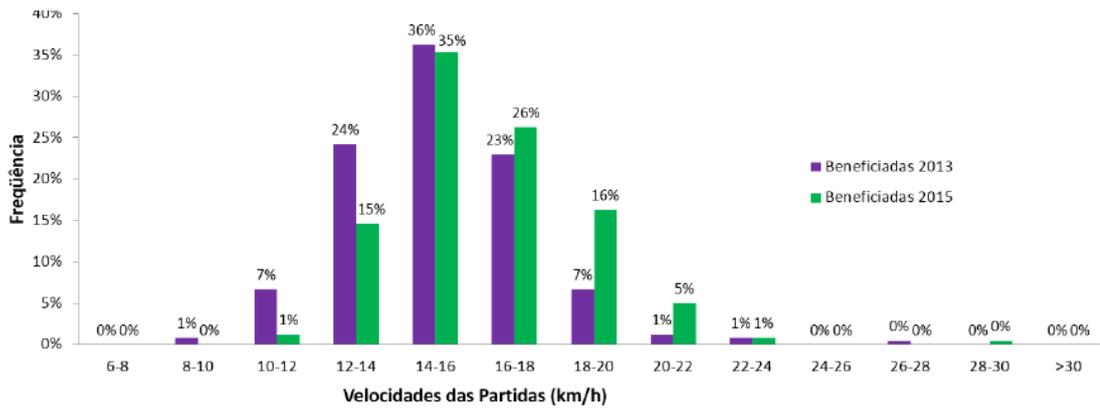
Em São Paulo, Arbex e Cunha (2016) utilizaram dados de GPS em conjunto com dados GTFS para avaliar as velocidades de linhas de transporte público no período de 2013 a 2015, quando foram implantadas diversas faixas exclusivas para ônibus na cidade, para avaliar o impacto das mudanças. É interessante ressaltar que foram utilizados no processo os dados no formato GTFS fornecidos pela Prefeitura de São Paulo. O trabalho concluiu que a implantação das faixas exclusivas e preferenciais teve impacto positivo sobre a velocidade de operação das linhas beneficiadas. A Figura 4 demonstra o comparativo das velocidades entre os anos de 2014 e 2015, para linhas beneficiadas e não beneficiadas pela implantação de faixas exclusivas de transporte público.

Figura 3: Velocidades operacionais de linhas ao longo do dia



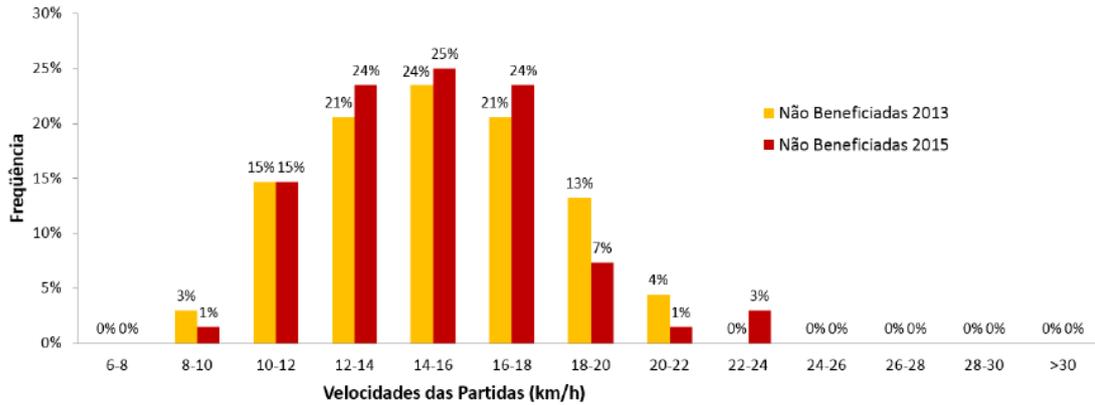
Fonte: (CORTÉZ *et al.*, 2011)

Figura 4: Comparação das velocidades operacionais de linhas beneficiadas com faixas exclusivas em 2013 e 2015



Fonte: (ARBEX; CUNHA, 2016)

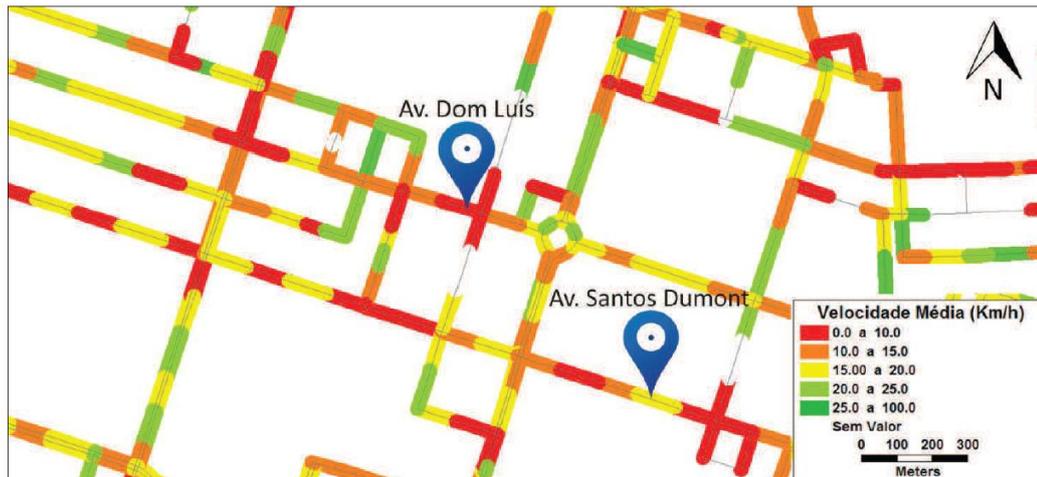
Figura 5: Comparação das velocidades operacionais de linhas não beneficiadas com faixas exclusivas em 2013 e 2015



Fonte: (ARBEX; CUNHA, 2016)

No que se refere a Fortaleza, os dados de GPS da frota de ônibus já são utilizados em pesquisas envolvendo a qualidade do serviço de ônibus. Maia (2015) aplicou uma metodologia para analisar as velocidades operacionais do transporte público em uma região de Fortaleza antes e após a implantação de intervenções viárias com faixas exclusivas. Diferente dos trabalhos apresentados anteriormente, os dados de localização foram agregados e não foram utilizados para avaliar o desempenho de uma linha ou sentido específico, mas de toda a operação nas vias selecionadas para o estudo. O autor sugeriu melhorias futuras no processo utilizado para a obtenção dos resultados. A Figura 6 apresenta as velocidades operacionais em vias onde foi implantado um binário.

Figura 6: Velocidades operacionais em vias de Fortaleza



Fonte: (Maia, 2015)

No caso da cidade de Fortaleza, Maia (2015) obteve o desempenho geral do transporte público nas vias escolhidas, mas não pode agregar dados para determinar o desempenho de linhas e sentidos específicos, pois os dados utilizados não continham as

informações necessárias, limitando o tipo de resultado. Já no caso das cidades de Santiago e São Paulo, a base de dados utilizada continha dados completos sobre o rastreamento da frota, possibilitando a análise de toda a frota com agregação de dados no nível de linhas e sentidos. Foi relatada a dificuldade de se adequar as coordenadas aos itinerários oficiais, o que aumentou o trabalho de processamento.

No presente trabalho, a utilização de *feeds* GTFS e de dados de bilhetagem eletrônica, abordadas mais adiante, possibilitou a adequação dos dados incompletos da cidade de Fortaleza para que estes pudessem fornecer informações mais específicas sobre a operação do sistema de transporte público da cidade.

Visto que muitas outras grandes cidades, assim como Fortaleza, também possuem dados de rastreamento armazenados, mas para fins alheios às análises apresentadas (dados secundários), espera-se que os métodos apresentados sirvam para resolução de problemas de outras cidades.

2.2.2 *Feed GTFS*

A Especificação Geral sobre Feeds de Transporte Público (GTFS – *General Transit Feed Specification*) foi desenvolvida em 2005, em uma parceria entre o Google e a TriMet, em Portland (EUA), para criar um recurso de planejamento de viagens de transporte público dentro do Google Maps. Esta especificação foi desenvolvida para ser simples de serem criados os arquivos pelas agências de transporte público, para o sistema ser alimentado de forma colaborativa. A ideia é de que o sistema seja facilmente acessado por desenvolvedores e robusto o suficiente para descrever um sistema de transportes. O GTFS estabelece uma série de arquivos em formato CSV (*comma separated values* ou valores separados por vírgula) que, juntos, descrevem as paradas, viagens, rotas e informações de tarifas sobre o serviço de uma agência. O *feed* foi aberto, para uso geral, pelo Google em 2007 e seu uso se propagou amplamente, à medida que as agências publicavam seus planos operacionais no formato padronizado. O formato permite análises de frequências, rotas e paradas que antes eram dificultadas pelo formato não digitalizado e não padronizado das informações de transporte público. De acordo com Wong (2013) o GTFS é o padrão mais usado para a troca estática de dados de transporte público nos Estados Unidos. A Figura 7 identifica os arquivos e campos obrigatórios e opcionais que compõem o *feed* GTFS.

Os *feeds* publicados pelas agências possibilitam que desenvolvedores criem aplicativos que usam esses dados de maneira interativa. A maioria dos aplicativos tem o objetivo de fornecer informações personalizadas aos usuários sobre os sistemas de transporte, mas existem usos alternativos para o GTFS. Um relatório discutiu um estudo de caso no Departamento de Transporte da Flórida, onde havia uma necessidade de localizar paradas de ônibus de alta atividade em toda a região, a fim de identificar áreas de interesse para a segurança de pedestres. Agregar e manter a base de dados atualizada seria um trabalho complicado. Em vez disso, seus parceiros de pesquisa usavam *feeds* GTFS e *scripts* simples para a manutenção do banco de dados (CATALÁ *et al.*, 2011).

Os *feeds* GTFS evitam erros de codificação manual e facilitam a integração de dados entre vários provedores, melhorando a qualidade geral dos dados. Além disso, tornam mais simples a tarefa de atualizar as informações de operação quando a programação muda por algum motivo, o que antes era feito manualmente (PUCHALSKY; JOSCHI, SCHERR, 2012).

Porém, dados estáticos de programação têm capacidade limitada para ajudar em processos de planejamento, pois muitas medidas de desempenho utilizam dados sobre o que de fato é executado, ao invés do que foi planejado. Uma das motivações do presente trabalho foi a de verificar a viabilidade de se utilizar dados de rastreamento da frota em conjunto com dados GTFS para medir o desempenho na operação de um sistema de transporte público.

2.2.3 Bilhetagem Eletrônica

Os dados do GPS e de bilhetagem eletrônica forneceram subsídio para vários estudos na última década. Grande esforço tem sido feito para usar os dados gerados pelas transações automatizadas de cobrança de tarifas em sistemas de transporte público. Os objetivos principais, da maioria dos estudos sobre este assunto, se concentraram na estimativa de padrões de viagem, variabilidade de uso e geração de matrizes origem-destino (PELLETIER; TRÉPANIER; MORENCY, 2011).

Munizaga e Palma (2012) propuseram uma metodologia para estimar uma matriz OD de transporte público a partir de dados de *smartcard* e GPS para Santiago, Chile. Dentre as contribuições do trabalho, está o uso de tempo generalizado ao invés de distância física para estimar ponto de descida para transferências e destinos. As taxas de sucesso alcançadas sugerem que esta fonte de informação tem um grande potencial e pode realmente substituir uma boa parte dos levantamentos de dados para matrizes Origem-Destino.

Fora do uso comum dos dados de bilhetagem eletrônica, Devillaine, Munizaga e Trépanier (2012) apresentaram um método para a estimativa da localização, horário, duração e finalidade das atividades realizadas pelos usuários de transporte público com o uso de bancos de dados de bilhetagem eletrônica e outras informações disponíveis sobre o uso do solo e comportamento do usuário. O método é aplicado a casos em Santiago, Chile e Gatineau, Quebec, Canadá, para identificar o propósito da atividade e o prazo para caracterizar o comportamento do usuário.

Freitas (2015) propôs uma metodologia para caracterização da problemática do sistema de transporte público de Fortaleza a partir dos dados de bilhetagem eletrônica. De maneira complementar, ele buscou reconstituir a matriz origem-destino que congrega os fluxos observados, obtendo resultados específicos sobre o transporte público da cidade. Dentre os resultados, destacam-se o comportamento da demanda ao longo do tempo e do espaço, o número de viagens realizadas por linha e por tipo usuário, o comparativo da origem da viagem com domicílio do usuário e a recuperação da matriz origem-destino do transporte público.

3 METODOLOGIA

Este capítulo descreve a proposta Metodológica para o uso de dados secundários de AVL na caracterização de um sistema de transporte público, abordando aspectos sobre os dados utilizados e as variáveis e indicadores extraídos.

3.1 Formato dos dados

O formato dos dados deve ser tal que, para cada coordenada emitida e armazenada, existam os seguintes dados associados:

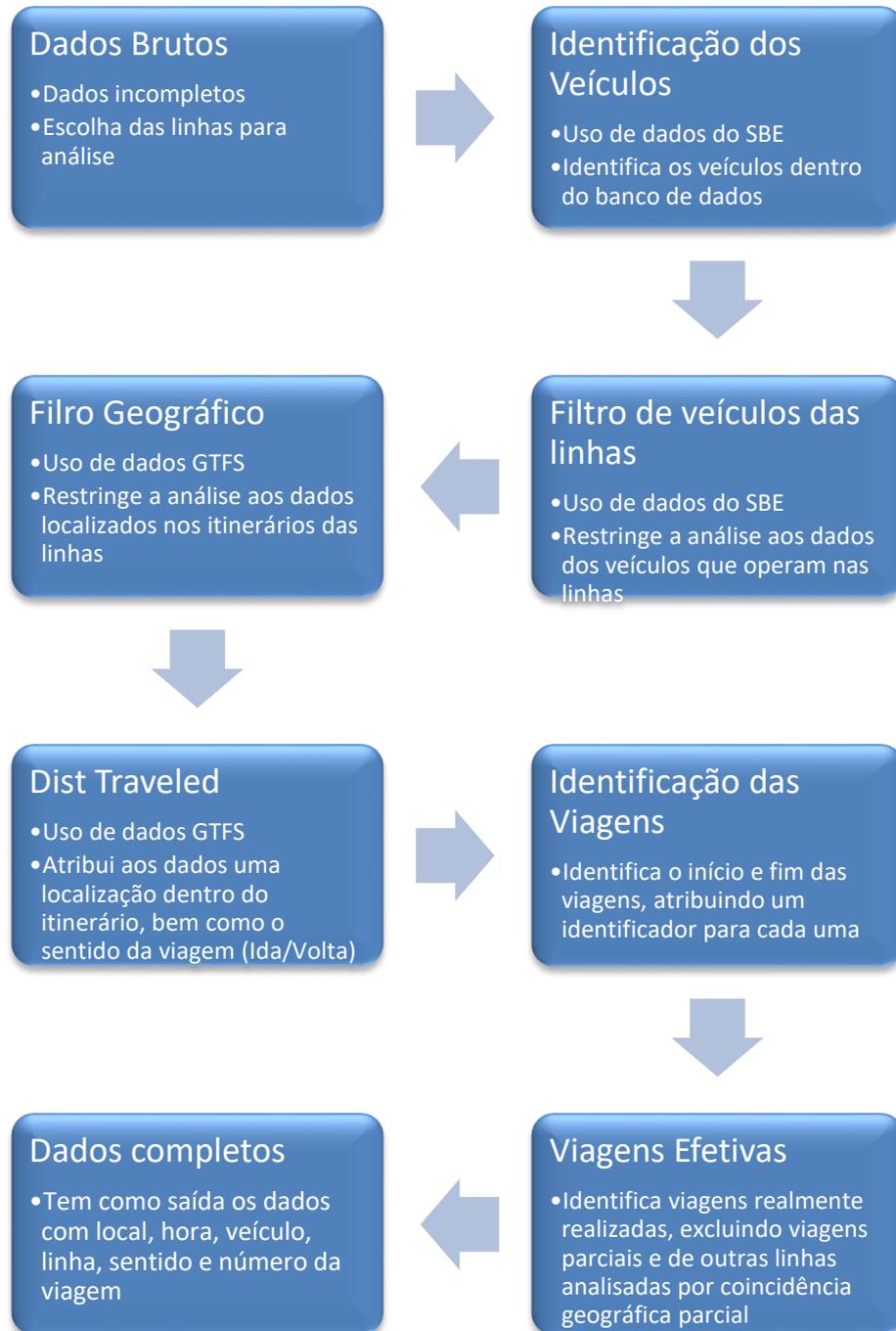
- Horário;
- Identificação do veículo;
- Identificação da linha;
- Identificação do sentido (ida/volta);
- Identificação da viagem.

Estas informações são necessárias para que o dado geográfico possa ser usado nas análises de linhas e viagens específicas. Se a coordenada armazenada não estiver identificada por essas informações, não se pode compará-la às outras para calcular as variáveis e indicadores. Dados não identificados, além de não poderem ser usados nas análises, prejudicam os resultados, pois deixam vazios no conjunto.

Parte deste trabalho se dedica a demonstrar técnicas para se adequar o banco de dados disponível ao formato necessário, seja combinando ou complementando dados, solucionando diversos problemas que possam estar associados à qualidade dos dados disponíveis. Neste trabalho, foram utilizadas duas ferramentas principais, programação em Python e edição de planilhas em Microsoft Excel.

Para isso, foram também utilizados dados geográficos sobre os itinerários e horários planejados das linhas de transporte público. Estes dados podem ser encontrados no *feed* GTFS, fornecido pela agência de transportes a entidades que os utilizam para integrar informações sobre transporte público a mapas fornecidos aos seus usuários. Também foram utilizados dados do SBE para corrigir inconsistências e vazios nos dados do AVL. A Figura 8 apresenta o fluxograma das etapas do processo de adequação dos dados.

Figura 8: Etapas do processo de adequação dos dados



Fonte: Elaborado pelo autor.

3.2 Indicadores obtidos

O objetivo da etapa de caracterização dentro do processo de planejamento é propor indicadores que descrevam e mensurem os problemas levantados na etapa de identificação. Para isso, é necessário utilizar dados para extrair variáveis que irão compor os indicadores. Os dados do AVL permitem o cálculo de variáveis importantes na caracterização do sistema de transporte público, detalhadas mais a seguir.

A maioria dos estudos utilizando dados de AVL focam na determinação da velocidade operacional da frota. Este trabalho demonstra, porém, que os dados são capazes de fornecer outras informações acerca da qualidade do serviço prestado ao usuário do transporte público, como o tempo de espera e a formação de comboios de veículos.

3.2.1 Velocidades

Segundo Cortés *et al.* (2011), a velocidade comercial é um fator chave para a avaliação do sistema de transporte público, pois está relacionada tanto com o nível de serviço quanto com os custos do sistema. Historicamente, as velocidades de operação são determinadas através de postos de controle no início e fim das viagens, resultando em seus tempos totais e oferecendo informações úteis para o dimensionamento da frota. Este método, porém, impossibilita uma análise desagregada do itinerário, necessária para uma melhor avaliação de intervenções. Nesse sentido, o AVL fornece dados que permitem calcular a velocidade dos veículos em qualquer ponto ou trecho do itinerário, resultando em uma compreensão mais aprofundada da problemática.

As velocidades dos veículos são calculadas por meio de uma expressão clássica, uma vez que os dados estejam no formato desejado. Basicamente, para cada ponto fornecido por um veículo, são conhecidas as distâncias acumuladas desde o início da viagem e os seus tempos associados. Para um determinado par de pontos desejados, é utilizada a definição clássica de velocidade média:

$$v_m = \frac{\Delta S}{\Delta t}$$

Onde:

v_m : *velocidade no trecho desejado*

ΔS : *comprimento do trecho (diferença das distâncias acumuladas)*

Δt : *tempo decorrido entre as duas leituras*

A definição dos trechos pode se dar por diversos critérios, como o tipo de via (via segregada, faixa preferencial, tráfego misto), o uso do solo, entre pontos de parada, trechos de

tamanho padronizado ou percentis do itinerário total. Para este trabalho, a definição dos trechos seguiu um padrão de 5% do total do itinerário. A vantagem desta divisão é que a quantidade total de trechos por sentido é igual em qualquer linha, independente do seu comprimento, facilitando a comparação entre diferentes linhas.

Outra análise interessante a ser feita com relação às velocidades é a sua variabilidade ao longo do dia em um determinado trecho. Para cada trecho, é feita uma média das velocidades de todos os veículos que passaram por ele durante um determinado período de tempo, no caso, períodos de uma hora. Uma alta variabilidade indica um trecho mais sensível às mudanças características de uma zona urbana: horários de pico, influência de polos geradores e eventos específicos.

3.2.2 *Tempo de espera*

Conhecendo-se o horário em que cada veículo de uma determinada linha passou por determinado ponto do itinerário, é possível calcular o intervalo de tempo decorrido entre um veículo e outro ao longo do dia (*headway*), nesse ponto. Esta variável pode ser comparada ao intervalo determinado pela tabela de horários planejada, disponível no *feed* GTFS, fornecendo uma avaliação sobre pontualidade do serviço. O tempo médio de espera, para o usuário, pode ser representado, aproximadamente, pela metade do intervalo decorrido entre a passagem dos veículos.

3.2.3 *Formação de comboios*

Os *headways* também fornecem, de maneira indireta, informações sobre a formação de comboios entre os veículos de uma linha. *Headways* muito abaixo da média prevista indicam veículos muito próximos, em comboio, resultando em problemas como, por exemplo, longos tempos de espera. Os comboios também podem ser identificados visualmente no gráfico de marchas, obtido através da agregação dos dados totais de uma linha.

4 ESTUDO DE CASO

O sistema de transporte público por ônibus da cidade de Fortaleza é um sistema de porte médio, com frota de aproximadamente 2 mil veículos, operando em cerca de 320 linhas. A infraestrutura conta com 7 terminais físicos de integração e sistema de bilhetagem eletrônica (SBE) que também permite integração (FORTALEZA, 2017).

O SBE armazena os dados de validações dos usuários, associando a validação à linha, veículo, horário da validação, tipo de usuário e sentido da viagem. Os dados armazenados pelo SBE não informam, no entanto, a localização geográfica do veículo no momento da validação. A Tabela1 apresenta uma amostra dos dados de bilhetagem.

Tabela 1: Amostra dos dados do SBE

Id	Linha	Linha2	Carro	Data_hora	Cod	Cod2	Ida_Volta
4858535	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:22	2	02-ESTUDANTE ETUFOR	Ida
4891654	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:22	12	12-VALE-TRANPORTE AVULSO	Ida
4423098	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:22	2	02-ESTUDANTE ETUFOR	Ida
5165546	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:22	4	04-VALE TRANSPORTE	Ida
6140564	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:28	4	04-VALE TRANSPORTE	Ida
6024962	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:28	2	02-ESTUDANTE ETUFOR	Ida
5137057	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:29	4	04-VALE TRANSPORTE	Ida
4353256	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:30	4	04-VALE TRANSPORTE	Ida
2071962	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:33	4	04-VALE TRANSPORTE	Ida
2071962	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:33	4	04-VALE TRANSPORTE	Ida
5286056	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:33	4	04-VALE TRANSPORTE	Ida
6453544	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:35	4	04-VALE TRANSPORTE	Ida
6391113	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:35	2	02-ESTUDANTE ETUFOR	Ida
2816868	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:35	4	04-VALE TRANSPORTE	Ida
6346808	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:35	4	04-VALE TRANSPORTE	Ida
0	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:36	27	PAGANTE EM HORARIO DIFERENCIADO	Ida
6130293	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:39	4	04-VALE TRANSPORTE	Ida
6099320	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:39	12	12-VALE-TRANPORTE AVULSO	Ida
0	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:41	27	PAGANTE EM HORARIO DIFERENCIADO	Ida
4910716	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:41	12	12-VALE-TRANPORTE AVULSO	Ida
4239287	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:41	4	04-VALE TRANSPORTE	Ida
5917573	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:41	4	04-VALE TRANSPORTE	Ida
6395457	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:41	12	12-VALE-TRANPORTE AVULSO	Ida
5278042	79	Antônio Bezerra/Náutico	35006	13/04/2016 06:41	4	04-VALE TRANSPORTE	Ida

Fonte: Elaborado pelo autor.

Onde os campos indicam:

- Id: identificador;
- Linha: Número da linha;
- Linha2: Nome da linha;
- Carro: Número do veículo;
- Data_hora: Data e horário da validação;
- Cod: Código do tipo de validação;
- Cod2: Tipo de validação;

- Ida_Volta: Sentido da viagem.

Toda a frota de veículos é monitorada em tempo real por um sistema de localização automática (AVL), que opera separadamente do SBE. Os dados armazenados relacionam a coordenada geográfica (latitude e longitude) emitida por um veículo, totalizando 43 campos. A estrutura destes dados de rastreamento pode ser observada na Tabela 2.

É importante observar que os dados armazenados pelo sistema de AVL não informam o sentido da viagem realizada no momento da emissão das coordenadas de GPS, além de conter muitos campos em branco ou com informações inconsistentes com relação à linha (*routecode*), o que impossibilita a maioria das análises. Os trabalhos de Cortéz *et al.* (2011) e Arbex e Cunha (2016), por exemplo, não seriam possíveis utilizando dados desta qualidade.

A agência de transportes de Fortaleza, ETUFOR, também fornece dados no formato GTFS, como os descritos previamente neste trabalho. Para a aplicação do método, foi utilizado um conjunto de dados referente ao dia 13/04/2016, uma quarta feira, dia útil. Foi escolhida para a análise a linha 079 – Antônio Bezerra / Náutico, que apresenta as seguintes características:

- Ponto inicial em um terminal físico de integração;
- Ponto final fora dos terminais de integração;
- Percorre trechos de características diversas, como:
 - Trecho de rodovia;
 - Corredor exclusivo;
 - Corredor preferencial;
 - Trecho de tráfego misto;

De maneira complementar, são apresentados resultados para as linhas 026 – Antônio Bezerra / Messejana, 200 – Antônio Bezerra / Centro / BRT e 222 – Antônio Bezerra / Papicu / Antônio Sales.

Tabela 2: Formato inicial dos dados do AVL

Ordem	Dados	
	Informações	Descrição
1	metricid	Identificador único
2	busservicecode	NA
3	calculatedspeed	Velocidade instantânea
4	counter1	NA
5	counter2	NA
6	counter3	NA
7	direction	Sentido do movimento
8	gprsonline	NA
9	gpsonline	NA
10	ignition	NA
11	ignitiontimehour	NA
12	infrabuscompany	NA
13	internalalert	NA
14	latitude	Coordenada geográfica
15	lifetimer	NA
16	longitude	Coordenada geográfica
17	metrictimestamp	Data e hora
18	mtcatstop	NA
19	mtcgpsdisconnected	NA
20	mtcinput2	NA
21	mtcinput3	NA
22	mtcinput4	NA
23	mtcinput5	NA
24	mtcoutput1	NA
25	mtcoutput2	NA
26	mtcoutput3	NA
27	mtcoutput4	NA
28	mtcspeedexceeded	NA
29	mtclamper	NA
30	odometer	Marcação do odômetro (em observação)
31	panealert	NA
32	panicalert	NA
33	patternhint	NA
34	the_geom	Etiqueta geométrica (em observação)
35	processed	NA
36	roletcounter	NA
37	routecode	NA
38	speed	NA - mesmo valor do campo 3
39	temperature	NA
40	transmissioncount	NA
41	voltage	NA
42	device_deviceid	Número do aparelho de GPS (M2M)
43	vehicle_vehicleid	Número do veículo (M2M)

NA – não aproveitado

Fonte: Elaborado pelo autor.

4.1 Tratamento inicial do banco de dados

4.1.1 Redução do tamanho do banco de dados

Como mencionado anteriormente, o AVL, no caso de Fortaleza, armazena 43 campos de dados, muitos dos quais não serão necessários para o método, pois não se relacionam com as variáveis procuradas. Portanto, é oportuno diminuir a quantidade de dados

para tornar o processo de cálculo mais leve. Uma rotina de eliminação de campos foi aplicada com essa finalidade e, ao final, o banco de dados, como mostra a Tabela 3, permanece com 13 campos apenas. Estes serão complementados, posteriormente, durante a aplicação do método.

Tabela 3: Dados de AVL após o tratamento inicial

Ordem	Dados	
	Informações	Descrição
1	metricid	identificador único
2	calculatedspeed	velocidade instantânea
3	direction	sentido do movimento
4	latitude	coordenada geográfica
5	longitude	coordenada geográfica
6	metrictimestamp	data e hora
7	odometer	marcação do odômetro (em observação)
8	the_geom	etiqueta geométrica (em observação)
9	device_deviceid	número do aparelho de GPS (M2M ¹)
10	vehicle_vehicleid	número do veículo (M2M)
11	carro	identificação da ETUFOR
12	segundos	tempo
13	bigtag	Identificador único

Fonte: Elaborado pelo autor.

Os campos 11 a 13 foram adicionados ao arquivo, identificando o veículo de acordo com o código da agência de transportes, o tempo em segundos decorrido desde o início do dia, gerando um novo identificador. A identificação do veículo pelo número da agência de transportes auxilia na determinação da linha em que está operando, o tempo em segundos permite o cálculo do tempo de deslocamento do veículo entre os pontos de interesse, enquanto o campo *bigtag* constitui um campo único (sem duplicatas) que possibilita a ligação com a base de dados de bilhetagem.

4.1.2 Inserindo a variável *dist_traveled* no *feed GTFS*

Apesar da agência de transportes de Fortaleza fornecer os dados no formato GTFS, alguns campos e arquivos não obrigatórios do *feed* são ignorados. É o caso do campo *dist_traveled*, do arquivo *shapes.txt*.

O arquivo *shapes.txt* contém as coordenadas geográficas de todas as linhas de ônibus, ponto a ponto, formando seus itinerários. O campo *dist_traveled* determina a distância acumulada desde o início do trajeto até o ponto em análise, como um odômetro, e é de fundamental importância para a aplicação do método. Assim, foi necessário o cálculo desta

variável, através da soma das distâncias entre os pontos consecutivos do itinerário, como demonstrado na Tabela 4.

Tabela 4: Variável `dist_traveled` no arquivo `shapes`

<code>shape_id</code>	<code>shape_pt_lat</code>	<code>shape_pt_lon</code>	<code>shape_pt_sequence</code>	<code>shape_dist_traveled</code>
shape079-I	-3,73752	-38,584548	1	0
shape079-I	-3,737483	-38,584552	2	4,14
shape079-I	-3,737105	-38,584587	3	46,32
shape079-I	-3,736826	-38,584613	4	77,46
shape079-I	-3,736792	-38,584329	5	109,24
shape079-I	-3,737457	-38,584263	6	183,50
shape079-I	-3,737515	-38,584257	7	189,98
shape079-I	-3,738309	-38,584184	8	278,58
shape079-I	-3,738487	-38,584266	9	300,35
shape079-I	-3,738526	-38,584546	10	331,77
shape079-I	-3,738642	-38,585382	11	425,55
shape079-I	-3,738771	-38,58629	12	527,46
shape079-I	-3,737821	-38,586329	13	633,12
shape079-I	-3,737649	-38,586326	14	652,23
shape079-I	-3,737602	-38,585853	15	705,05
shape079-I	-3,737578	-38,585437	16	751,35
shape079-I	-3,738166	-38,585414	17	816,74
shape079-I	-3,738642	-38,585382	18	869,76
shape079-I	-3,738758	-38,585362	19	882,84
shape079-I	-3,738862	-38,585352	20	894,45
shape079-I	-3,739237	-38,585369	21	936,16
shape079-I	-3,739421	-38,585589	22	968,03

Fonte: Elaborado pelo autor.

4.2 Identificação inicial da linha

4.2.1 Busca dos veículos no sistema de bilhetagem

Para contornar o problema da falta de informação sobre a linha, uma das abordagens possíveis é buscar o veículo no banco de dados do SBE, onde essa informação possui maior consistência. No caso de Fortaleza, o veículo é identificado de maneira diferente no AVL e no SBE. No primeiro, a identificação se dá pelo campo `vehicle_id`, uma numeração de 4 dígitos originada na empresa que opera o rastreamento. Já no segundo, a identificação se dá pelo número de 5 dígitos do carro, que é especificado pela ETUFOR e identifica a empresa que opera aquele veículo, o ano de fabricação e sua numeração dentro da empresa.

Foi utilizado um dicionário fornecido pela ETUFOR que relaciona estas duas numerações. Assim, foi possível encontrar os veículos que operam em determinada linha através dos dados do SBE e identificá-los dentro dos dados do AVL, eliminando os demais. O processo pode ser observado nas Tabelas 5 a 7.

Tabela 5: Amostra do dicionário de veículos

vehicleid	numbus
1187	35258
1188	35116
1189	35205
1190	35313
1192	35139
1193	35237
1194	35610

Fonte: (M2M, 2017)

Tabela 6: Dados de bilhetagem da linha 079

Id	Linha	Linha2	Carro	Data_hora	Cod	Cod2	Ida_Volta
4476679	79	Antônio Bezerra/Náutico	35139	13/04/2016 19:19	4	04-VALE TRANSPORTE	Volta
3286445	79	Antônio Bezerra/Náutico	35139	13/04/2016 19:21	4	04-VALE TRANSPORTE	Volta
1218258	79	Antônio Bezerra/Náutico	35139	13/04/2016 19:25	2	02-ESTUDANTE ETUFOR	Volta
0	79	Antônio Bezerra/Náutico	35139	13/04/2016 19:39	27	PAGANTE EM HORARIO DIFERENCIADO	Volta
6238972	79	Antônio Bezerra/Náutico	35139	13/04/2016 19:39	4	04-VALE TRANSPORTE	Volta
3002630	79	Antônio Bezerra/Náutico	35139	13/04/2016 19:47	4	04-VALE TRANSPORTE	Volta
6308748	79	Antônio Bezerra/Náutico	35139	13/04/2016 19:47	4	04-VALE TRANSPORTE	Volta
6308748	79	Antônio Bezerra/Náutico	35139	13/04/2016 19:48	4	04-VALE TRANSPORTE	Volta
566183	79	Antônio Bezerra/Náutico	35139	13/04/2016 19:48	2	02-ESTUDANTE ETUFOR	Volta
6222608	79	Antônio Bezerra/Náutico	35139	13/04/2016 19:49	12	12-VALE-TRANPORTE AVULSO	Volta
541776	79	Antônio Bezerra/Náutico	35139	13/04/2016 19:52	2	02-ESTUDANTE ETUFOR	Volta
599143	79	Antônio Bezerra/Náutico	35139	13/04/2016 19:52	2	02-ESTUDANTE ETUFOR	Volta
0	79	Antônio Bezerra/Náutico	35139	13/04/2016 20:12	27	PAGANTE EM HORARIO DIFERENCIADO	Ida
2548894	79	Antônio Bezerra/Náutico	35139	13/04/2016 20:16	2	02-ESTUDANTE ETUFOR	Ida
5463463	79	Antônio Bezerra/Náutico	35139	13/04/2016 20:27	12	12-VALE-TRANPORTE AVULSO	Ida
6311998	79	Antônio Bezerra/Náutico	35139	13/04/2016 20:27	4	04-VALE TRANSPORTE	Ida
0	79	Antônio Bezerra/Náutico	35139	13/04/2016 20:29	27	PAGANTE EM HORARIO DIFERENCIADO	Ida
0	79	Antônio Bezerra/Náutico	35139	13/04/2016 20:33	27	PAGANTE EM HORARIO DIFERENCIADO	Volta
0	79	Antônio Bezerra/Náutico	35139	13/04/2016 20:37	27	PAGANTE EM HORARIO DIFERENCIADO	Volta
0	79	Antônio Bezerra/Náutico	35139	13/04/2016 20:37	27	PAGANTE EM HORARIO DIFERENCIADO	Volta
5747890	79	Antônio Bezerra/Náutico	35139	13/04/2016 20:37	4	04-VALE TRANSPORTE	Volta
4909556	79	Antônio Bezerra/Náutico	35139	13/04/2016 20:38	12	12-VALE-TRANPORTE AVULSO	Volta
6287017	79	Antônio Bezerra/Náutico	35139	13/04/2016 20:38	4	04-VALE TRANSPORTE	Volta

Fonte: Elaborado pelo autor.

Tabela 7: Dados do AVL da linha 079

latitude	longitude	seg	hora	routecode	device_deviceid	vehicle_vehicleid	linha	dist_traveled	longe	menor_dst	IDA	viagem	
-3,733500004	-38,55369949	72273	20:04:33		0	2564	1192	79	4648,13	0	0,620383257	1	8
-3,7332201	-38,55239868	72303	20:05:03		0	2564	1192	79	4808,39	0	9,206940263	1	8
-3,733200073	-38,55239868	72333	20:05:33		0	2564	1192	79	4808,66	0	9,828257964	1	8
-3,733109951	-38,55189896	72363	20:06:03		0	2564	1192	79	4851,80	0	27,31229612	1	8
-3,732389927	-38,54859924	72393	20:06:33		0	2564	1192	79	5188,36	0	35,11648732	1	8
-3,731980085	-38,54629898	72453	20:07:33		0	2564	1192	79	5512,55	0	18,87322305	1	8
-3,731659889	-38,5428009	72488	20:08:08		0	2564	1192	79	5934,02	0	43,96968566	1	8
-3,732579947	-38,54259872	72518	20:08:38		0	2564	1192	79	6038,63	0	44,35814747	1	8
-3,733360052	-38,54119873	72548	20:09:08		0	2564	1192	79	6234,80	0	21,12822869	1	8
-3,73416996	-38,53910065	72578	20:09:38		0	2564	1192	79	6485,40	0	19,36330438	1	8
-3,734210014	-38,53900146	72608	20:10:08		0	2564	1192	79	6496,60	0	8,128821538	1	8
-3,734910011	-38,53689957	72638	20:10:38		0	2564	1192	79	6771,20	0	17,08198796	1	8
-3,735919952	-38,53350067	72668	20:11:08		0	2564	1192	79	7136,43	0	13,08514852	1	8
-3,736190081	-38,53279877	72698	20:11:38		0	2564	1192	79	7221,68	0	38,12581785	1	8
-3,736850023	-38,53099823	72758	20:12:38		0	2564	1192	79	7437,29	0	4,351094196	1	8
-3,737040043	-38,53039932	72773	20:12:53		79	2564	1192	79	7503,87	0	7,76838003	1	8
-3,737620115	-38,52809906	72803	20:13:23		79	2564	1192	79	7768,09	0	14,83993398	1	8
-3,737639904	-38,52809906	72863	20:14:23		79	2564	1192	79	7768,33	0	14,67451758	1	8
-3,738500118	-38,52539825	72893	20:14:53		79	2564	1192	79	8079,76	0	5,625321147	1	8
-3,739259958	-38,5223999	72923	20:15:23		79	2564	1192	79	8446,86	0	7,164455921	1	8
-3,739989996	-38,52009964	72953	20:15:53		79	2564	1192	79	8709,44	0	27,71028324	1	8
-3,740859985	-38,51739883	72983	20:16:23		79	2564	1192	79	9032,43	0	7,386305732	1	8

Fonte: Elaborado pelo autor.

Como pode ser visto nas Tabelas 5 e 6, o carro 35139, de *vehicle_id* 1192, estava operando na linha 079 no horário 20:10. Porém, neste horário, o dado registrado pelo AVL não identifica corretamente a linha (campo *routecode*), como demonstrado na Tabela 7. É então criado o campo “linha”, contendo a informação resultante deste cruzamento de dados.

4.2.2 *Eliminação dos pontos fora do itinerário*

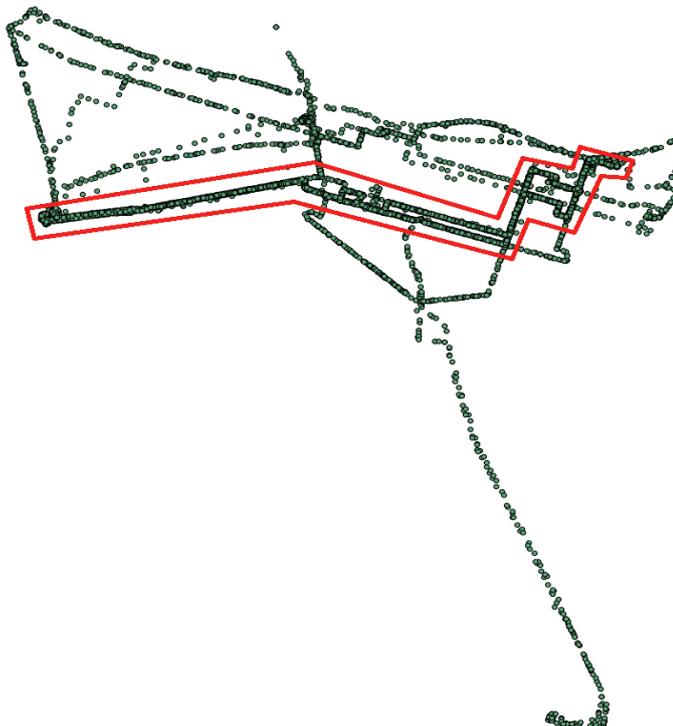
O processo anterior elimina a maioria dos pontos não interessantes para a análise. Porém, ainda constam muitos dados fora do itinerário analisado. Isto se dá por motivos como:

- Um mesmo veículo pode operar em mais de uma linha ao longo do dia;
- O GPS permanece ligado antes e depois das viagens, em caminhos e horários que não interessam à análise.

Para minimizar a ocorrência desses dados, foi criado um filtro que identifica e elimina pontos localizados fora do itinerário da linha em análise. Cada coordenada armazenada no AVL é comparada com as coordenadas do *shape* referente à linha analisada, sendo calculada sua distância até o ponto mais próximo do itinerário. Em seguida, são eliminados os dados que estão a mais de 100 metros do itinerário.

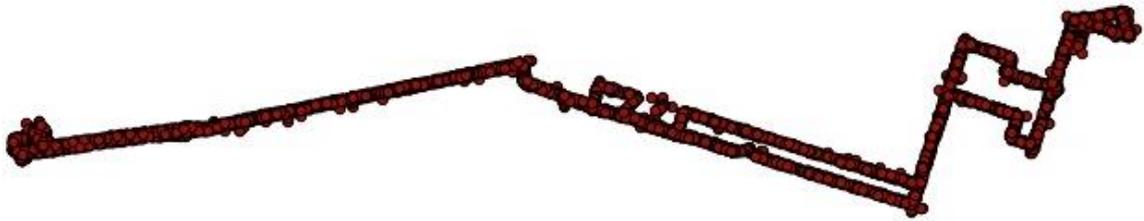
A Figura 9 apresenta os dados dos veículos que operaram na linha 079 durante todo o dia. Após a aplicação do filtro, tem-se os pontos apresentados na Figura 10.

Figura 9: Registros de GPS da linha 079, antes do filtro espacial



Fonte: Elaborado pelo autor.

Figura 10: Registros de GPS da linha 079, após a aplicação do filtro espacial.



Fonte: Elaborado pelo autor.

Apesar de limpar bastante os dados indesejados, esta rotina ainda falha em eliminar dados de veículos cujo trajeto coincide, em parte, com o itinerário da linha analisada. Estes dados podem ser eliminados em uma etapa posterior, ao analisar a consistência das viagens realizadas.

4.3 Determinação da distância percorrida e sentido do deslocamento

4.3.1 Atribuição inicial do *dist_traveled*

Os dados agora estão todos dentro do itinerário. Pode-se, então, atribuir a cada um dos dados um valor de *dist_traveled*, afim de se saber em que ponto do trajeto está. Como não se sabe o sentido da viagem, é feito um teste, atribuindo ao ponto o *dist_traveled* correspondente ao trajeto de ida. A necessidade deste teste se dá pelo fato de que, no caso do trajeto de ida coincidir com o de volta (utilizando a mesma via), os pontos do *shape* de ida e volta se tornam muito próximos. Assim, os dados podem ser associados ao ponto errado do *shape*.

Podem ser utilizados dois métodos para isto:

- a) Interpolando o *dist_traveled* dos dois pontos mais próximos do *shape*

Este método compara a coordenada registrada com as duas coordenadas mais próximas dentre os pontos que compõem o arquivo *shape* da linha analisada, interpolando o valor do *dist_traveled*.

- b) Criando uma linha correspondente ao *shape* e associando o dado ao ponto mais próximo desta linha.

A linguagem Python possui um módulo chamado *shapely*, que permite criar uma linha com base em coordenadas em sequência, formando um trajeto contínuo e atribuindo uma distância acumulada a cada ponto do

trajeto. Pode-se, então, calcular o ponto mais próximo desta linha para cada dado armazenado, associando a eles uma distância acumulada.

Após este processamento, todos os dados possuem um *dist_traveled* associado, referente ao trajeto de ida.

4.3.2 Atribuição IDA/VOLTA

No passo seguinte, é feita uma análise do *dist_traveled* atribuído na etapa anterior. Os dados estão organizados por ordem de veículo, seguida de horário. Ou seja, os dados consecutivos pertencem ao mesmo veículo, em instantes imediatamente consecutivos. Assim, dada a evolução do *dist_traveled*, pode-se identificar se o veículo está no trajeto de ida ou volta: Caso o valor do *dist_traveled* aumente, significa que o veículo está no trajeto de ida, caso contrário, está no trajeto de volta.

Foi observado que existem valores nulos ou inconsistentes, ocorrendo por dois motivos principais:

- a) O veículo está parado ou em velocidade muito baixa, podendo o GPS registrar um deslocamento nulo ou negativo por conta da imprecisão;
- b) O veículo passa pelo mesmo ponto mais de uma vez, como em um laço de quadra, alça de viaduto ou retorno no ponto final ou dentro do terminal, registrando a posição incorreta dentro do *shape*.

Em ambos os casos, a rotina irá identificar o sentido incorreto, sendo necessário um ajuste. Para isso, é feita uma média móvel dos valores de IDA/VOLTA, o que elimina valores de IDA em meio a valores de VOLTA, e vice-versa.

4.3.3 Ajuste de *dist_traveled*

Sabendo o sentido correto do deslocamento, é possível atribuir novamente o *dist_traveled* ao dado, desta vez utilizando o *shape_ida* e *shape_volta* para cada situação distinta. A Tabela 8 exemplifica o processo de ajustes no sentido da viagem e nas distâncias acumuladas, destacando dois valores inconsistentes corrigidos pelo método da média móvel (campo “mm”)

Tabela 8: Correção dos valores inconsistentes de ida e volta

vehicle_vcarro	linha	sentido	lat	lon	ponto	dist_ac1	sentnovo	dist_act	sentNum	mm	sentnovo	dist_act2	horaDec
1446	35009	79	555876,4	9587514	POINT (55	13009,93	Ida	13009,93	1	7	Ida	13009,93	20,66833
1446	35009		556032,1	9587984	POINT (55	13504,71	Ida	13504,71	1	5	Ida	13504,71	20,69639
1446	35009		556121,1	9587997	POINT (55	13606,4	Ida	13606,4	1	5	Ida	13606,4	20,70111
1446	35009		556343	9587900	POINT (55	14434,78	Ida	14434,78	1	3	Ida	14434,78	20,73583
1446	35009		555921	9587678	POINT (55	13179,29	Volta	14931,7	-1	-1	Volta	14931,7	20,8675
1446	35009		555820,8	9587389	POINT (55	12882,04	Volta	15236,03	-1	-3	Volta	15236,03	20,88806
1446	35009	79	555642,9	9587384	POINT (55	12731,84	Ida	12731,84	1	-5	Volta	15421,1	20,89583
1446	35009	79	555532	9587437	POINT (55	12883,16	Volta	15542,93	-1	-7	Volta	15542,93	20,91333
1446	35009		555432	9587464	POINT (55	11593,22	Volta	15646,22	-1	-9	Volta	15646,22	20,91722
1446	35009	79	555099,2	9587728	POINT (55	11187,31	Volta	16198,28	-1	-9	Volta	16198,28	20,93694
1446	35009		554877,1	9587475	POINT (55	11063,96	Volta	16645,03	-1	-9	Volta	16645,03	20,96667
1446	35009	79	554632,3	9586720	POINT (55	10455,79	Volta	17438,57	-1	-9	Volta	17438,57	21,00667
1446	35009	79	554432,2	9586454	POINT (55	10140,85	Volta	17865,59	-1	-11	Volta	17865,59	21,04417
1446	35009	79	554243,3	9586506	POINT (55	9636,045	Volta	18061,45	-1	-11	Volta	18061,45	21,05222
1446	35009	79	553876,9	9586626	POINT (55	9251,741	Volta	18446,83	-1	-11	Volta	18446,83	21,06389
1446	35009		552977,8	9586902	POINT (55	8302,685	Volta	19387,6	-1	-9	Volta	19387,6	21,10833
1446	35009	79	552644,5	9587007	POINT (55	7948,732	Volta	19736,82	-1	-9	Volta	19736,82	21,1225
1446	35009		552611,4	9587016	POINT (55	7914,23	Volta	19771,31	-1	-9	Volta	19771,31	21,125
1446	35009		552100,6	9586960	POINT (55	7440,714	Volta	20239,5	-1	-9	Volta	20239,5	21,16306
1446	35009		551778,7	9587078	POINT (55	7098,101	Volta	20581,86	-1	-9	Volta	20581,86	21,17611
1446	35009		551745,2	9587111	POINT (55	7056,523	Ida	7056,523	1	-9	Volta	20625,04	21,19028
1446	35009		551456,8	9587201	POINT (55	6754,261	Volta	21111,3	-1	-9	Volta	21111,3	21,24028
1446	35009		551145,9	9587279	POINT (55	6431,226	Volta	21450,47	-1	-9	Volta	21450,47	21,25806
1446	35009		550790,7	9587509	POINT (55	5941,061	Volta	21910,45	-1	-9	Volta	21910,45	21,29611

Fonte: Elaborado pelo autor.

Pode-se observar que o campo *dist_ac1* refere-se à distância acumulada relativa à viagem de ida, com valores decrescentes para os pontos identificados como volta. O campo *dist_act2* modifica os valores para os pontos identificados como volta, sendo usado como valor final.

4.4 Identificação das viagens

4.4.1 Avaliação do *dist_traveled*

Com os dados organizados no formato correto, é necessário identificar cada viagem distinta. Para isso, é feita uma análise do valor do *dist_traveled* ao longo do arquivo. Sempre que um veículo finaliza uma viagem de VOLTA e inicia uma de IDA, o valor do *dist_traveled* sofre uma queda, pois o contador de distância percorrida volta ao início do trajeto.

Foi criado um contador de viagens, associado a cada dado para identificar a viagem a que está associado. Sempre que ocorre esta reinicialização do *dist_traveled*, o contador registra o início de uma nova viagem. Assim, as análises podem ser feitas por linha, por veículo e por viagem. A Tabela 9 exemplifica a mudança no número da viagem.

Tabela 9: Identificação das viagens

routecode	device_de	vehicle_v	carro	sentnovo	dist_act2	horaDec	dist_na	aux	viagem
32847	1447	1446	35009	Volta	23210,7	21,34167	1300,254	0	6
32847	1447	1446	35009	Volta	23587,47	21,35361	376,7691	0	6
32847	1447	1446	35009	Volta	23803	21,36806	215,5287	0	6
32847	1447	1446	35009	Volta	24508,02	21,41139	705,0134	0	6
32847	1447	1446	35009	Volta	24803,86	21,42528	295,8447	0	6
32847	1447	1446	35009	Volta	26393,76	21,48	1589,899	0	6
32847	1447	1446	35009	Volta	26606,99	21,48528	213,2282	0	6
32847	1447	1446	35009	Volta	26710,42	21,4925	103,4316	0	6
0	1447	1446	35009	Volta	26747	21,50583	36,58387	0	6
0	1447	1446	35009	Ida	62	21,66583	-26684,8	1	7
0	1447	1446	35009	Ida	1088	21,77639	1026,068	0	7
79	1447	1446	35009	Ida	2538	21,81667	1449,333	0	7
79	1447	1446	35009	Ida	2818	21,82639	280,3585	0	7
79	1447	1446	35009	Ida	3181	21,85417	363,2494	0	7
79	1447	1446	35009	Ida	4151	21,88306	969,8228	0	7
79	1447	1446	35009	Ida	4208	21,88417	56,84563	0	7
79	1447	1446	35009	Ida	4607	21,89278	398,9581	0	7

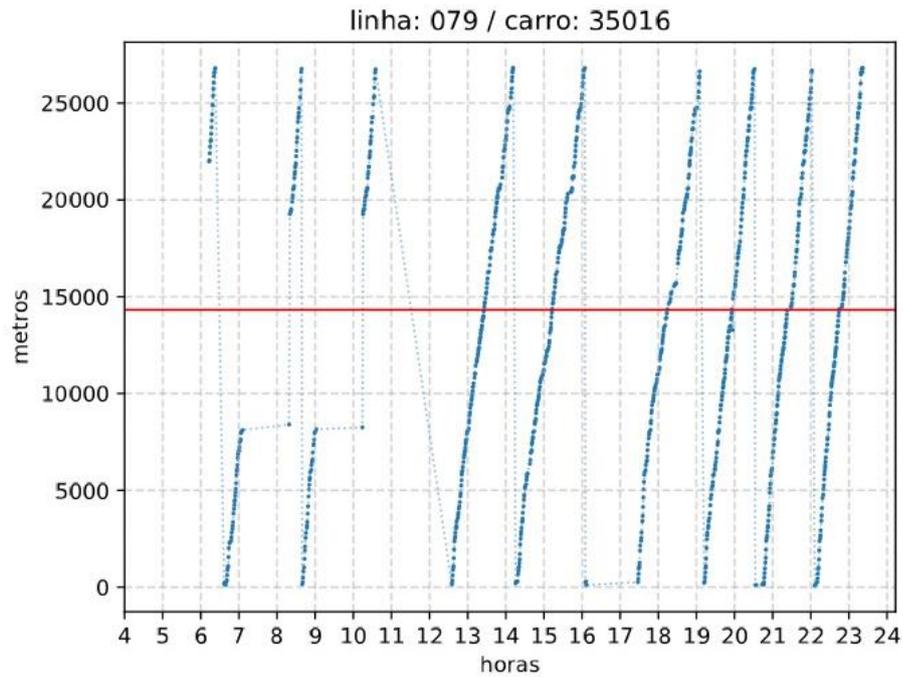
Fonte: Elaborado pelo autor.

4.4.2 Identificação das viagens efetivas

A etapa anterior identifica um número de viagens maior do que foram realizadas efetivamente naquela linha ou por aquele veículo. Isto se dá porque o veículo pode estar operando em outra linha, ou em trajetos diferentes do itinerário analisado, mas com trechos ou pontos coincidentes. Na Figura 11, observa-se a localização do veículo 35016 dentro do itinerário ao longo do dia, aonde o eixo x representa o horário e o eixo y representa a posição relativa ao início do itinerário. O veículo opera na linha 079 apenas a partir de 12h30min, estando em outra linha durante a manhã. Como esta linha tem parte do trajeto coincidente com a linha 079, estes pontos acabam fazendo parte da análise.

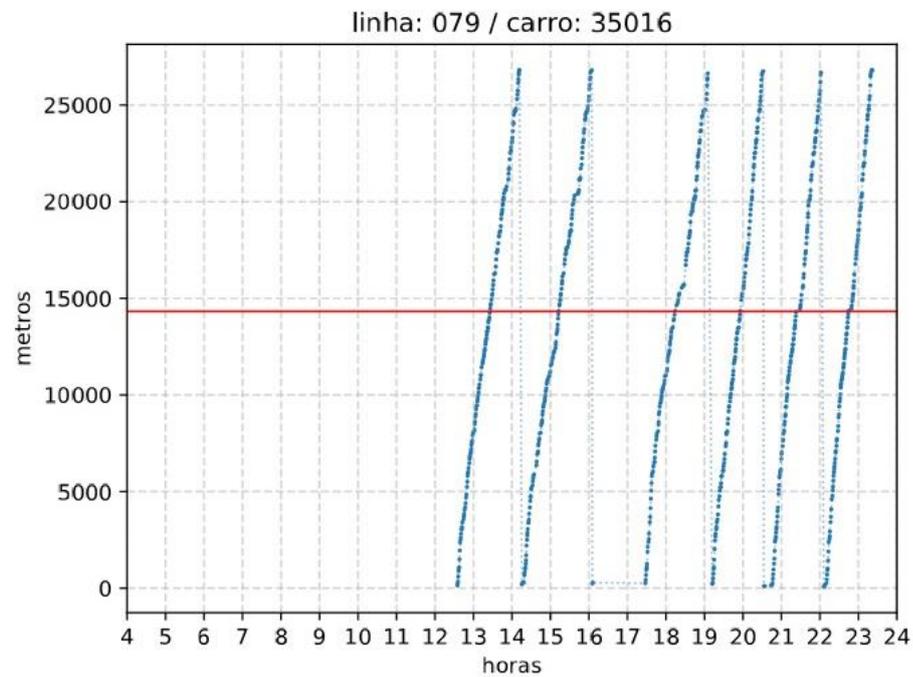
Nesta etapa, o programa considera apenas viagens que iniciem e finalizem em um dos terminais, sendo filtradas as viagens incompletas. Como resultado, são eliminados os últimos dados indesejados. A Figura 12 apresenta o resultado para as viagens do veículo 35016.

Figura 11: Gráfico de marcha individual, com dados indesejados



Fonte: Elaborado pelo autor.

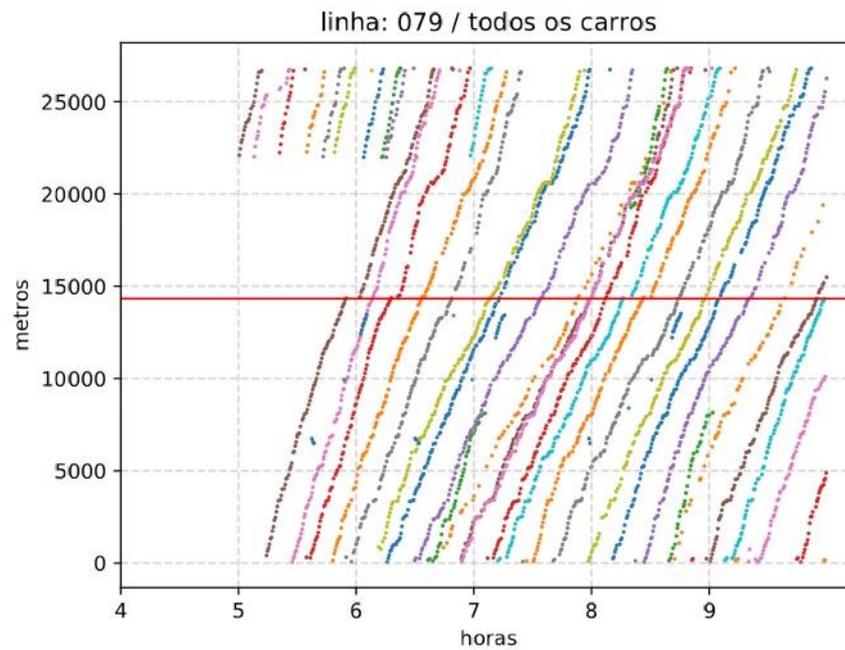
Figura 12: Gráfico de marcha individual, após a limpeza dos dados



Fonte: Elaborado pelo autor.

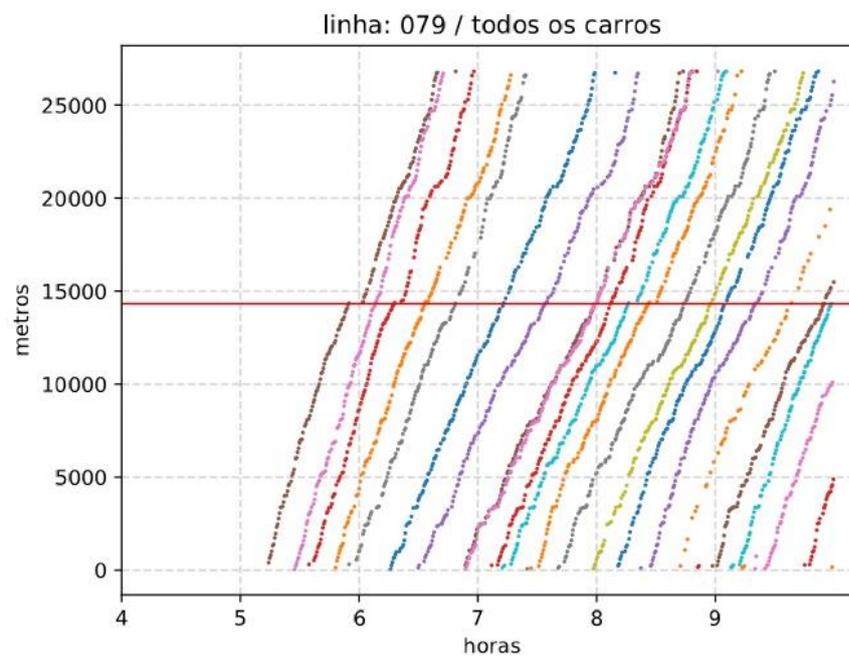
As Figuras 13 e 14 mostram o resultado, na forma de gráfico de marcha, para todos os veículos da linha 079, no período de 5h às 10h, onde cada veículo é representado por uma cor distinta.

Figura 13: Gráfico de marcha total com dados indesejados



Fonte: Elaborado pelo autor.

Figura 14: Gráfico de marcha total após a limpeza dos dados



Fonte: Elaborado pelo autor.

4.6 Determinação das velocidades

As velocidades são calculadas por meio da fórmula clássica da velocidade média:

$$v_m = \frac{\Delta S}{\Delta t}$$

Onde ΔS representa a diferença entre a distância acumulada nos dois instantes medidos, e Δt a diferença de tempo.

Como a divisão dos trechos foi feita a cada 5% do comprimento total do itinerário, para cada viagem foram calculadas as velocidades para 20 trechos. A Tabela 10 mostra as velocidades de operação para o veículo 35149, durante a viagem de número 11, sentido de IDA da linha 079.

Tabela 10: Velocidade de operação individual por trechos

Viagem 11		
Trecho	Hora	Velocidade (km/h)
1	06:55:50	2,9
2	06:58:50	13,4
3	06:59:50	44,8
4	07:05:17	8,6
5	07:10:47	8,4
6	07:12:47	16,8
7	07:15:47	19,4
8	07:18:47	13,1
9	07:21:47	12,1
10	07:24:47	16,1
11	07:28:47	11,0
12	07:32:47	11,1
13	07:35:25	16,0
14	07:38:55	11,8
15	07:41:25	17,7
16	07:46:55	8,5
17	07:51:25	9,1
18	07:53:25	20,0
19	07:56:55	13,6
20	07:59:25	17,3

Fonte: Elaborado pelo autor.

Para cada trecho, foi calculada a média das velocidades dos veículos a cada intervalo de uma hora, obtendo como resultado a variação da velocidade no trecho ao longo do dia. A Tabela 11 apresenta as velocidades de operação para as viagens da linha 079 durante todo o dia.

Tabela 11: Velocidade de operação da linha 079 por trecho e horários

		Horário																					
		5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23			
Trechos	1	18	16	5	17	5	8	10	14	13	11	6	15	13	8	19	11	12	10	26			
	2	19	14	18	18	19	28	25	27	21	23	21	20	25	20	24	25	26	30	34			
	3	38	34	36	33	42	46	43	43	39	45	48	48	39	42	52	45	47	53	54			
	4	26	14	17	21	25	23	18	21	21	25	38	20	22	23	31	28	28	30	28			
	5	15	10	9	13	15	12	18	13	19	17	22	13	18	15	13	21	21	31	18			
	6	25	21	24	22	21	16	20	18	26	29	25	24	33	30	26	28	29	37	39			
	7	21	28	20	26	17	20	23	23	29	29	29	27	22	18	21	29	32	34	25			
	8	21	21	16	17	21	15	23	22	21	15	22	19	27	19	20	31	31	36	31			
	9	22	15	12	15	18	19	14	15	20	14	17	15	16	12	15	18	24	26	24			
	10	18	16	16	16	15	14	19	23	18	19	24	18	20	20	23	28	18	21	21			
	11	25	22	13	17	19	21	16	16	18	19	14	18	15	19	19	20	21	21	38			
	12	19	14	11	13	17	15	15	15	21	12	16	14	14	16	23	23	18	17	33			
	13	34	21	18	18	20	20	23	23	21	22	18	20	19	18	23	33	36	35	33			
	14	18	19	14	14	14	17	14	15	18	12	16	23	13	13	20	22	22	25	39			
	15	25	23	14	13	16	16	15	15	15	14	15	18	12	16	17	25	33	31	26			
	16	21	21	10	10	11	12	5	9	13	9	9	18	8	12	20	15	19	14	21			
	17	12	18	10	11	15	16	12	9	13	13	14	10	13	8	13	16	24	18	29			
	18	23	19	15	15	13	13	12	12	13	11	11	11	14	11	16	16	17	20	23			
	19	14	13	10	13	12	15	17	14	12	17	15	12	10	10	11	14	14	17	18			
	20	18	17	17	19	18	17	19	21	17	20	19	20	12	14	17	20	17	15	21			
	21		10	11	10	13	7	14	9	11	13	9	9	5	7	11	14	8	16	6			
	22		24	17	14	18	14	18	17	15	16	16	11	9	8	12	12	15	20	25			
	23		23	22	24	19	20	20	26	19	23	20	17	15	14	15	23	22	28	24			
	24		22	18	14	16	19	13	18	21	16	15	24	10	15	13	19	30	17	20			
	25		33	17	20	14	16	19	31	20	16	14	15	10	13	17	20	23	21	21			
	26		20	22	13	16	14	15	20	20	11	11	11	13	12	19	16	32	20	28			
	27		21	20	21	16	11	17	17	14	14	15	20	12	16	21	32	26	24	35			
	28		30	32	25	27	22	39	27	32	37	26	26	21	32	30	27	34	21	34			
	29		17	17	11	17	15	14	10	14	20	13	20	9	10	19	16	23	22	18			
	30		11	11	12	11	9	11	15	10	12	7	12	10	8	12	17	12	17	27			
	31		10	12	11	15	15	16	12	10	11	15	13	11	22	14	15	24	13	23			
	32		28	21	22	23	26	17	22	30	20	21	23	17	23	21	23	21	28	28			
	33		36	25	29	22	26	29	26	28	23	24	19	14	20	24	21	27	28	26			
	34		28	21	18	25	26	22	28	17	19	19	19	15	20	23	23	31	22	31			
	35		17	27	19	22	19	23	22	22	28	25	21	26	18	17	22	18	26	28			
	36		30	25	28	21	36	28	22	25	30	18	15	21	21	28	25	22	26	33			
	37		21	25	18	10	10	13	16	15	11	11	16	9	6	10	15	18	22	16			
	38		26	18	46	29	18	27	32	20	23	30	24	15	31	28	26	31	32	31			
	39		35	42	47	38	45	47	45	43	41	41	37	30	40	43	39	41	43	51			
	40		17	26	15	13	8	17	18	19	23	6	10	4	23	22	15	9	12	7			

Fonte: Elaborado pelo autor.

Pode ser observada uma baixa velocidade de operação nos trechos 16 e 17, especialmente nos horários de pico, chegando a um valor mínimo de 5 km/h. Estes são trechos de tráfego misto em vias coletoras, o que pode explicar a dificuldade na operação. Já o trecho 3 apresenta uma alta velocidade de operação, sofrendo alguma variação nos horários de pico. Este é um trecho urbano de uma rodovia federal, com 3 faixas por sentido, mais duas faixas na pista marginal. A Tabela 12 apresenta os resultados para a linha 026:

Tabela 12: Velocidade de operação da linha 026 por trecho e horários

		Horário																					
		4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23		
Trechos	1	25	14	14	21	18	13	18	21	9	10	13	13	18	12	14	14	15	6	22	19		
	2	37	29	23	26	25	30	29	36	31	30	35	33	31	34	34	36	33	36	27	38		
	3	36	27	14	18	25	19	30	23	20	20	20	24	18	24	22	24	31	26	38	32		
	4	30	19	12	12	19	13	15	17	14	19	17	19	19	15	14	15	21	27	23	28		
	5	28	28	19	19	21	20	20	17	24	24	23	22	26	19	19	23	22	24	22	26		
	6	23	23	18	14	19	14	18	16	20	14	19	17	14	18	17	17	16	19	25	22		
	7	22	21	15	3	8	12	13	13	13	12	12	13	13	11	11	15	16	16	17	21		
	8	24	20	13	6	10	10	12	10	10	10	11	12	11	11	10	15	17	17	19	18		
	9	25	29	19	12	12	12	10	10	8	11	10	10	7	6	10	14	24	18	31	28		
	10	33	29	24	23	19	23	20	22	20	26	22	16	18	16	10	18	21	24	34	26		
	11	30	26	22	16	17	21	16	19	20	19	17	19	21	13	11	17	22	22	24	27		
	12	46	41	36	38	37	37	39	35	33	37	38	41	35	29	32	38	35	42		42		
	13	60	37	36	36	39	38	40	37	39	40	37	37	40	30	33	40	36	40		51		
	14	57	36	35	36	38	39	40	37	37	38	39	38	42	24	24	36	40	37		48		
	15	58	30	34	38	47	47	43	50	41	46	40	41	39	41	33	41	40	49		41		
	16	51	40	34	37	39	41	40	49	40	48	40	39	35	34	35	40	42	44		47		
	17	28	33	36	34	34	39	36	31	37	36	31	35	34	31	30	34	40	35	36	46		
	18	27	31	25	21	23	26	21	24	23	25	19	26	24	20	19	20	25	30	25	33		
	19	28	31	32	25	32	37	36	23	23	26	22	25	26	20	20	25	26	30	26	39		
	20	3	27	18	16	13	10	14	11	12	12	14	11	9	8	5	8	25	24	20	32		
	21	16	14	7	4	3	1	2	3	5	8	3	2	1	1	4	5	2	1	1	3		
	22	31	20	20	15	14	13	13	13	17	17	13	10	16	17	14	16	17	18	21	35		
	23	33	27	21	15	23	25	26	22	22	29	25	21	23	24	19	27	25	25	34	32		
	24	43	26	19	17	18	23	19	22	18	23	21	22	17	15	15	19	25	24	27	35		
	25	41	35	33	35	37	38	42	41	36	39	38	36	40	36	32	41	39	40	39	42		
	26	47	37	33	33	38	45	49	49	36	39	40	33	36	36	33	50	45	42	40	49		
	27	43	36	31	31	36	43	37	38	38	41	35	32	36	32	30	36	39	34	36	43		
	28	42	40	29	26	30	33	36	32	31	38	33	31	31	33	26	41	39	36	37	46		
	29	47	32	24	23	32	33	32	32	31	37	28	24	29	31	25	32	35	33	35	33		
	30	51	46	33	27	32	41	41	56	47	44	45	43	38	39	40	36	33	31	40	44		
	31	34	28	18	14	22	17	22	19	20	23	20	21	19	17	17	21	23	24	23	33		
	32	33	30	22	19	13	19	24	21	25	20	18	19	21	20	18	24	31	23	31	38		
	33	30	21	16	13	13	11	7	9	14	15	11	11	13	12	10	16	20	17	22	26		
	34	29	22	17	10	13	10	8	13	10	10	12	11	12	12	11	15	18	14	21	26		
	35	20	20	12	8	11	14	11	15	13	15	13	13	12	11	13	17	19	17	17	22		
	36	39	34	27	24	19	25	24	19	23	26	24	25	22	19	16	27	26	19	31	39		
	37	33	27	20	16	16	18	20	18	18	20	21	16	17	20	13	19	21	25	25	25		
	38	23	21	19	23	18	16	13	18	17	17	17	14	12	14	9	13	12	15	26	27		
	39	33	31	32	30	36	27	30	40	27	26	32	26	33	25	24	30	35	34	34	36		
	40		23	21	28	17	19	19	35	25	13	17	27	16	13	8	10	18	18	34	29		

Fonte: Elaborado pelo autor.

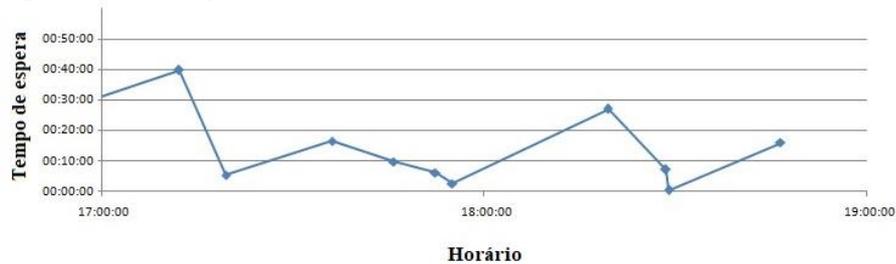
É possível observar a diferença na velocidade de operação entre os trechos 7-9, correspondentes a uma área de tráfego misto e os trechos 12-19, correspondentes a uma rodovia federal. A característica da via não é, contudo, o único fator que influencia a velocidade de operação de um sistema de transportes. Os resultados das velocidades servem de base para uma investigação mais detalhada sobre as causas da baixa eficiência na operação da linha. Adicionalmente, são apresentados, no Apêndice A, os mapas de calor (*heatmaps*), semelhantes às Tabelas 11 e 12, para as linhas 026, 079, 200 e 222.

4.7 Determinação dos tempos de espera

Sabendo que horário cada veículo passou por determinado ponto do trajeto, é possível calcular o tempo médio de espera neste ponto do itinerário. Este tempo pode ser representado pela metade do tempo médio entre veículos (*headway*).

O gráfico da Figura 15, referente ao ponto final do trecho 11 da viagem de ida da linha 079, apresenta os *headways* entre os horários de 17h e 19h. O *headway* médio observado foi de 14 minutos, resultando em um tempo médio de espera de 7 minutos. Os valores, porém, variaram de 1 a 40 minutos. Isto significa que, apesar de longos tempos de espera, ocorre também a passagem de vários veículos em um curto espaço de tempo.

Figura 15: Headways no final do trecho 11 da linha 079



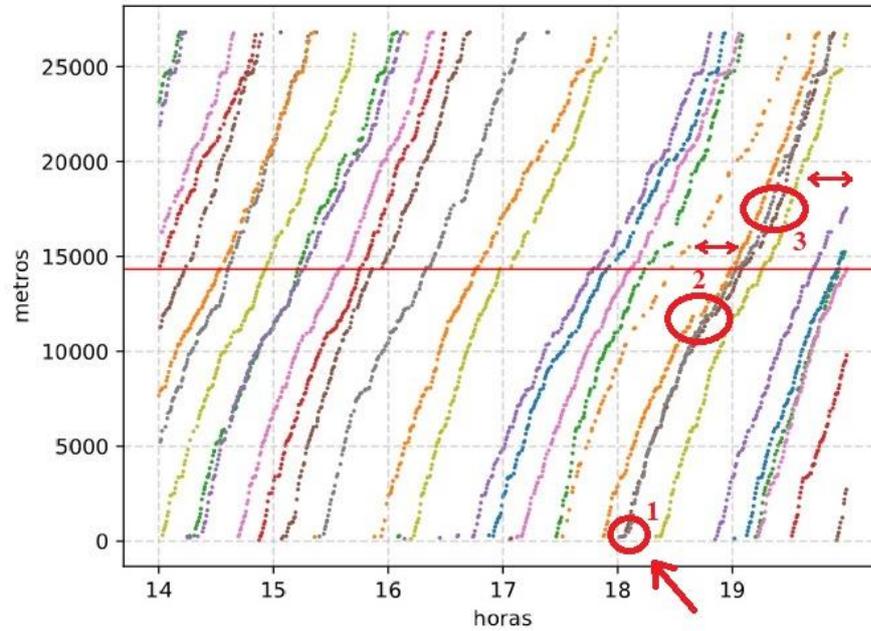
Fonte: Elaborado pelo autor.

Isto ajuda a caracterizar a confiabilidade do sistema, uma vez que, no *feed* GTFS, existe um valor determinado para o *headway* médio, que pode ser encontrado no arquivo *stop_times.txt*. Para este caso, os veículos deveriam passar a cada 13 minutos, com valores variando entre 12 e 16 minutos. Assim, conclui-se que, para este horário, não estão sendo cumpridos os intervalos planejados.

4.8 Identificação da formação de comboios

Analisando os *headways* e os gráficos de marcha, é possível identificar a formação de comboios por horário e trecho, como pode ser visto na Figura 16. Observando o gráfico de marcha, é possível identificar dois veículos com horário de partida por volta das 18h (ponto 1). Às 18h40min, um terceiro veículo é alcançado pelos dois (ponto 2), próximo ao ponto final da viagem de ida, representado pela linha vermelha horizontal. Eles se mantêm próximos durante toda a viagem de volta. O comboio ainda é alcançado por um quarto veículo (ponto 3), fazendo com que aumentem os tempos de espera antes e depois da sua passagem (setas), como pode ser observado pelas lacunas formadas. Adicionalmente, são apresentados, no apêndice B, os gráficos de marchas para as linhas 026, 079, 200 e 222.

Figura 16: Gráfico de marcha da linha 079



Fonte: Elaborado pelo autor.

Pelo exposto neste capítulo, a metodologia aplicada foi capaz de utilizar dados secundários incompletos de AVL para fornecer informações relevantes para a etapa de caracterização da problemática de um sistema de transportes. Por meio do cruzamento de bancos de dados, foi possível complementar as informações armazenadas pelo AVL e tornar possível a extração de resultados para além da velocidade comercial, como o gráfico de marchas, os tempos de espera para o usuário e os pontos de formação de comboio.

5 CONCLUSÕES E RECOMENDAÇÕES

O presente trabalho teve como objetivo geral propor uma metodologia para utilizar dados secundários de AVL na etapa de caracterização da problemática de um sistema de transporte público. A construção do problema de pesquisa se apoiou na necessidade de dados para a etapa de caracterização de um STPP, na dificuldade de se obter dados através de levantamento por pesquisas de campo e na disponibilidade de dados secundários com potencial para auxiliar nesta fase do planejamento.

Para alcançar o objetivo geral, foram traçados objetivos específicos que passavam por analisar o uso deste tipo de dado no processo de planejamento, determinar os indicadores que este sistema poderia fornecer, analisar os dados disponíveis e propor e aplicar um método para adequar este banco de dados para extração e representação destes indicadores.

Foi feita uma revisão acerca do uso de sistemas automáticos de coleta de dados e suas aplicações, sendo constatada a relevância desses dados para as etapas de planejamento. Para o caso do AVL, as principais aplicações destacadas foram a determinação das velocidades de operação em vários níveis de desagregação, a análise do desempenho de vias de tráfego misto através de dados de GPS de ônibus e fornecimento de informações em tempo real para usuários e operadores. As pesquisas limitam-se ao uso de dados completos de rastreamento, não focando na avaliação da qualidade destes dados. Além disso, a presente pesquisa mostrou ser possível a determinação dos tempos de espera para o usuário do sistema, bem como a identificação de comboios que diminuem a confiabilidade do serviço.

Por conta das barreiras devido ao formato e qualidade dos dados disponíveis, foi demonstrada uma série de técnicas para contornar possíveis problemas com a base de dados. Foi possível corrigir as informações inconsistentes ou incompletas do banco de dados do AVL, como a identificação da linha e do sentido da viagem através do cruzamento com bancos de dados do SBE e dos *feeds* GTFS. Ao final do tratamento dos dados, a base corrigida foi capaz de fornecer informações específicas para avaliação do desempenho e caracterização do sistema.

Foram determinadas as velocidades dos veículos e da frota, desagregada por trecho e por horários, ao longo do itinerário. Com isso, foi possível identificar trechos com dificuldade de operação, além da variação das velocidades ao longo do dia.

Foi demonstrado ser possível a determinação dos tempos de espera para o usuário em um determinado ponto do itinerário. Os resultados mostraram uma alta variação nestes tempos, constatando uma baixa pontualidade no serviço.

Finalmente, foi traçado o gráfico de marchas para a linha em análise, uma representação dos dados que indica a posição dos veículos no itinerário ao longo do tempo. A análise do gráfico permitiu identificar a formação de comboios entre os veículos da frota.

Finalizada a aplicação do método de caracterização para os dados de Fortaleza, conclui-se que os dados armazenados pelo AVL têm uma boa capacidade de fornecer informações relevantes para o planejamento de um sistema de transportes, ainda que a que se tratem de dados secundários, ou seja, levantados e especificados para outros fins. O método de adequação apresentado é indicado para cidades com dados inconsistentes ou incompletos, mas a etapa de determinação dos indicadores é comum a todas as análises de bancos de dados completos.

Os indicadores apresentados identificam e mensuram problemas de operação e qualidade do serviço para o usuário do sistema de transporte público. Estas informações auxiliam a tomada de decisão da gestão, uma vez que os recursos, financeiros e humanos, são limitados e podem ser direcionados aos pontos críticos do sistema.

Foram encontradas algumas limitações na elaboração deste trabalho, que resultaram em sugestões para trabalhos futuros:

- Para as linhas com os dois pontos finais em terminais físicos de integração, ocorre, muitas vezes, um tempo de espera entre as viagens de ida e volta. Isso ocasiona o registro de uma baixa velocidade no primeiro trecho da viagem de volta (21) e/ou no último trecho da viagem de ida (20). Isto pode ser observado nos *heatmaps* de velocidades operacionais das linhas 026 e 222, contidos no apêndice A. Como consequência, pode ser registrada uma diminuição na velocidade global de operação da linha, pois a análise da viagem é precocemente iniciada, uma vez que o veículo está aguardando para deixar o terminal de integração.
- As análises foram feitas para um único dia de observações. Portanto, as variações avaliadas foram ao longo deste dia. Isto pode ocasionar resultados influenciados por características específicas do dia da análise, como acidentes de trânsito, desvios de rota ou congestionamentos devido a eventos diversos. O uso de dados de um período mais longo pode dar mais robustez aos resultados.

REFERÊNCIAS

- Brasil. **Lei nº 12.587, de 03 de janeiro de 2012. Institui as diretrizes da Política Nacional de Mobilidade Urbana e dá outras providências.** Brasília; 2012.
- ARBEX, R. O.; CUNHA, C. B. da. Avaliação das mudanças nas velocidades das linhas de ônibus da cidade de São Paulo após a implantação de faixas exclusivas através da análise de dados de GPS. **TRANSPORTES**, v. 24, n. 4, p. 21, 1 dez. 2016.
- ARROXELAS GALVÃO, O. J. de. Desenvolvimento dos transportes e integração regional no Brasil - uma perspectiva histórica. **Planejamento e Políticas Públicas**, v. 13, p. 183–214, 1996.
- ASSOCIAÇÃO NACIONAL DE TRANSPORTES PÚBLICOS. **Sistema de Informações da Mobilidade Urbana: Relatório Geral 2012.** São Paulo: ANTP, 2014.
- BERKOW, M.; CHEE, J.; BERTINI, R. L.; MONSERE, C. Transit performance measurement and arterial travel time estimation using archived AVL data. **ITE District**, v. 6, 2007.
- CASCETTA, E. **Transportation Systems Analysis: Models and Applications.** 2nd. ed. New York: Springer, 2009.
- CATALA, M.; DOWNING, S.; HAYWARD, D. Expanding the Google transit feed specification to support operations and planning. **World Transit Research**, 1 jan. 2011. Disponível em: <<https://www.worldtransitresearch.info/research/4067>>.
- CHAKROBORTY, P.; KIKUCHI, S. Using Bus Travel Time Data to Estimate Travel Times on Urban Corridors. **Transportation Research Record: Journal of the Transportation Research Board**, v. 1870, p. 18–25, 1 jan. 2004.
- CHAPLEAU, R.; CHU, K. K. A.; ALLARD, B. Synthesizing AFC, APC, GPS and GIS Data to Generate Performance and Travel Demand Indicators for Public Transit. In: **Anais... In: TRANSPORTATION RESEARCH BOARD 90TH ANNUAL MEETING** TRANSPORTATION RESEARCH BOARD. 2011. Disponível em: <<https://trid.trb.org/view/1092733>>. Acesso em: 27 jul. 2018.
- CONFEDERAÇÃO NACIONAL DOS TRANSPORTES. **Transporte de Passageiros.** In: PESQUISA CNT/COPPEAD. Rio de Janeiro, 2002.
- CORTÉS, C. E.; GIBSON, J.; GSCHWENDER, A.; MUNIZAGA, M.; ZÚÑIGA, M. Commercial bus speed diagnosis based on GPS-monitored data. **Transportation Research Part C: Emerging Technologies**, v. 19, p. 695–707, 2011.
- CORTÉS, C. E.; SÁEZ, D.; MILLA, F.; NÚÑEZ, A.; RIQUELME, M. Hybrid predictive control for real-time optimization of public transport systems' operations based on evolutionary multi-objective optimization. **Transportation Research Part C: Emerging Technologies**, Applications of Advanced Technologies in Transportation: Selected papers from the 10th AATT Conference. v. 18, n. 5, p. 757–769, 1 out. 2010.

DESSOUKY, M.; HALL, R.; NOWROOZI, A.; MOURIKAS, K. Bus dispatching at timed transfer transit stations using bus tracking technology. **Transportation Research Part C: Emerging Technologies**, v. 7, n. 4, p. 187–208, 1 ago. 1999.

DEVILLAIN, F.; MUNIZAGA, M.; TRÉPANIÉ, M. Detection of Activities of Public Transport Users by Analyzing Smart Card Data. **Transportation Research Record: Journal of the Transportation Research Board**, v. 2276, p. 48–55, dez. 2012.

FORTALEZA. **Plano de Ação Imediata de Trânsito e Transportes**. 2017. Disponível em: <<http://www.fortaleza.ce.gov.br/scsp/paaitt-plano-de-acoes-imediatas-de-transito-ettransportes>>. Acesso em: 15 jan. 2017.

FREITAS, A. T. de. **Metodologia de caracterização da problemática do sistema de transporte público de passageiros a partir dos dados da bilhetagem eletrônica**. 2015. (Dissertação de Mestrado) Universidade Federal do Ceará, Fortaleza, 2015.

GARCIA, C. **Strategic Assessment of Accessibility on Urban Mobility Networks**. 2016. (Ph. D. thesis), Instituto Superior Técnico - Universidade de Lisboa, 2016.

MAIA, S. S. **Caracterização das velocidades do sistema de transporte público com base em dados do rastreamento dos veículos nas vias de Fortaleza**. 2015. (Trabalho de Conclusão de Curso) Universidade Federal do Ceará. Curso de Engenharia Civil, Fortaleza, 2015.

MUNIZAGA, M. A.; PALMA, C. Estimation of a Disaggregate Multimodal Public Transport Origin-Destination Matrix from Passive Smartcard Data from Santiago, Chile. **Transportation Research Part C: Emerging Technologies**, v. 24, p. 9–18, out. 2012.

OKUNIEFF, P. E. **AVL systems for bus transit: a synthesis of transit practice**: Transit Cooperative Research Program (TCRP). Washington, DC: Transportation Research Board, 1997.

PELLETIER, M.-P.; TRÉPANIÉ, M.; MORENCY, C. Smart Card Data Use in Public Transit: A Literature Review. **Transportation Research Part C: Emerging Technologies**, v. 19, n. 4, p. 557–568, ago. 2011.

PU, W.; LIN, J. Buses as probes: archived versus real-time automatic vehicle location data. **Proc., 2008 Advanced Applications of Transportation Technology**, p. 27–31, 2008.

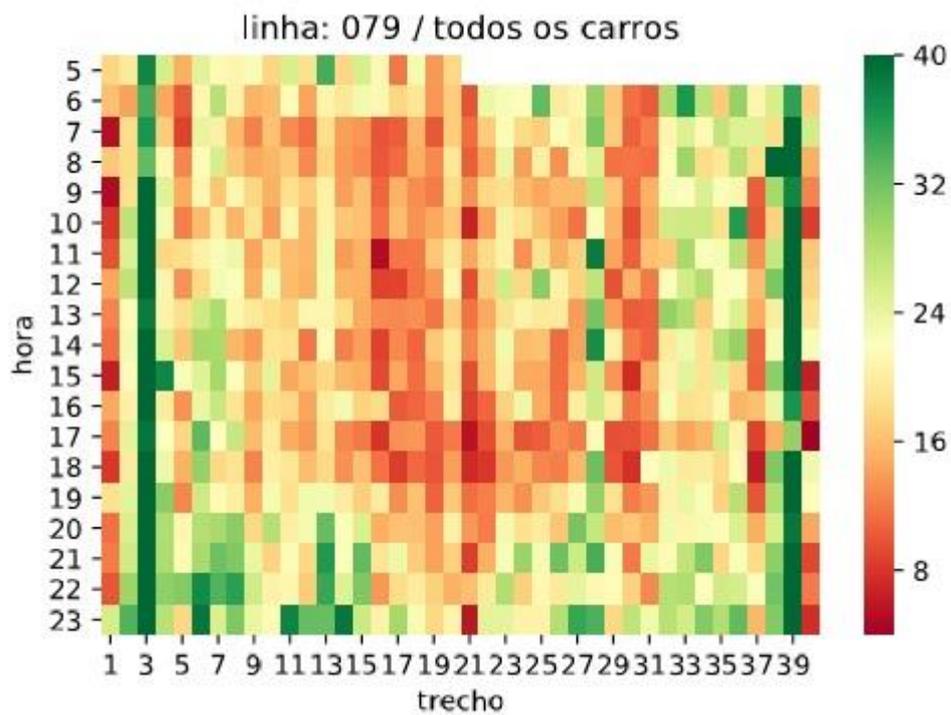
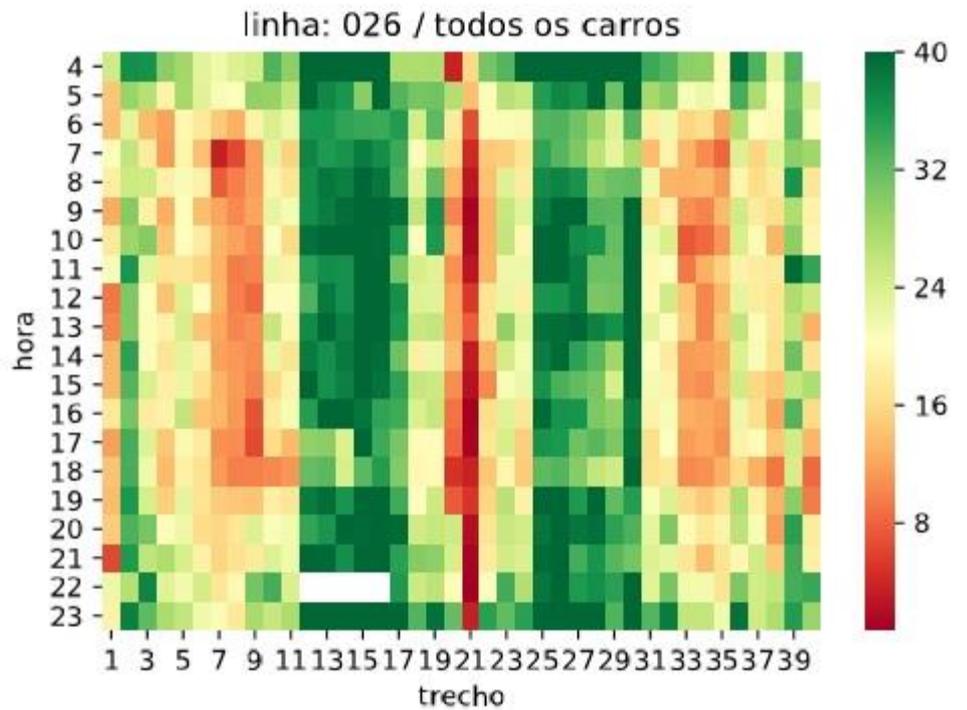
PUCHALSKY, C. M.; JOSHI, D.; SCHERR, W. Development of a regional forecasting model based on Google transit feed. In: 91st annual meeting of the transportation research board, Washington, DC, **Anais...** 2012.

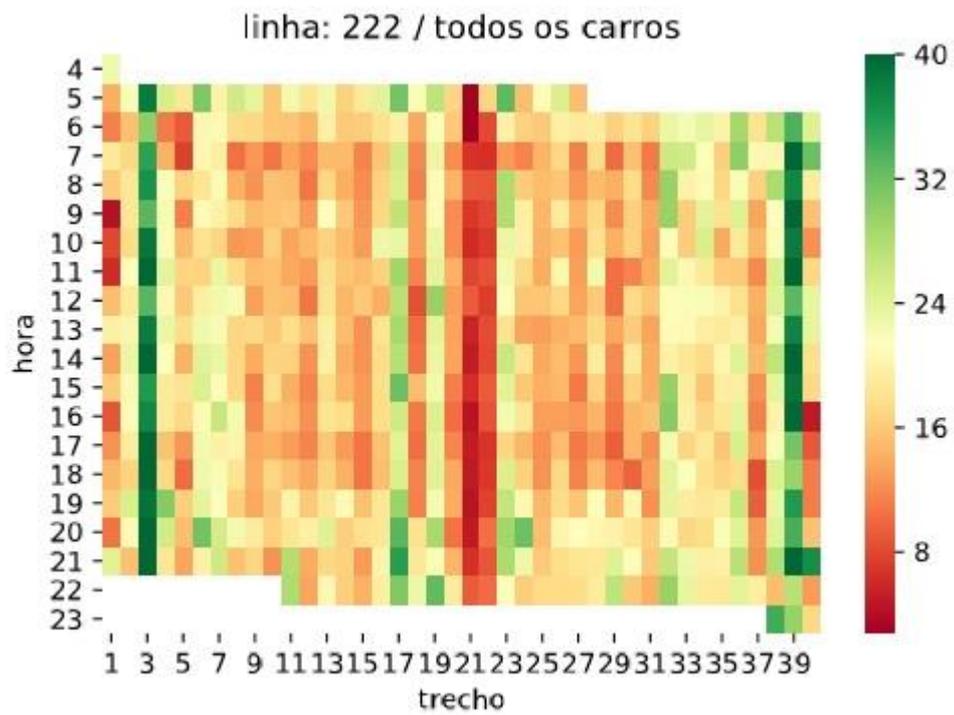
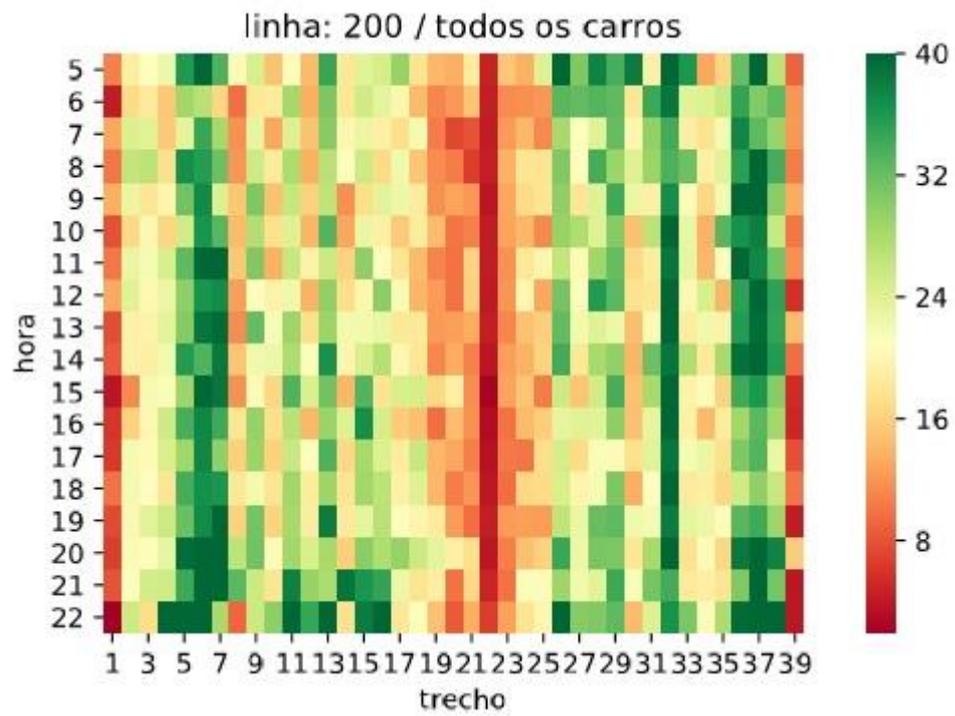
SOARES, F. D. P. **Proposta metodológica de compreensão da problemática das relações entre uso do solo e transportes no planejamento urbano integrado**. 2014. (Dissertação de Mestrado) Universidade Federal do Ceará, Fortaleza, 2014. Disponível em: <<http://www.repositorio.ufc.br/handle/riufc/15178>>. Acesso em: 16 mar. 2017.

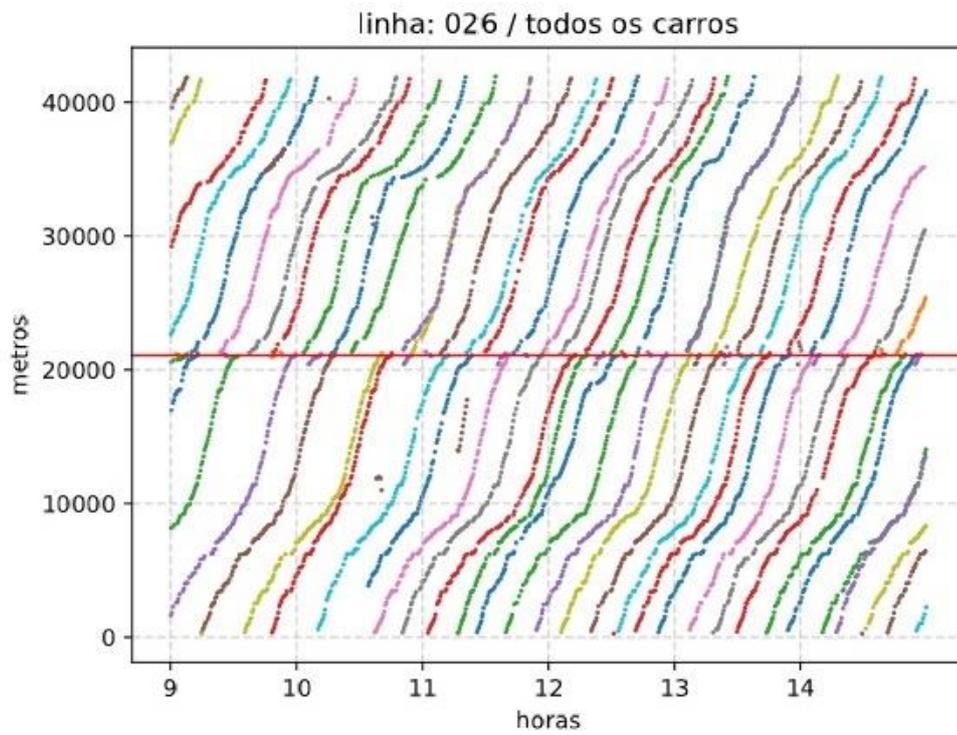
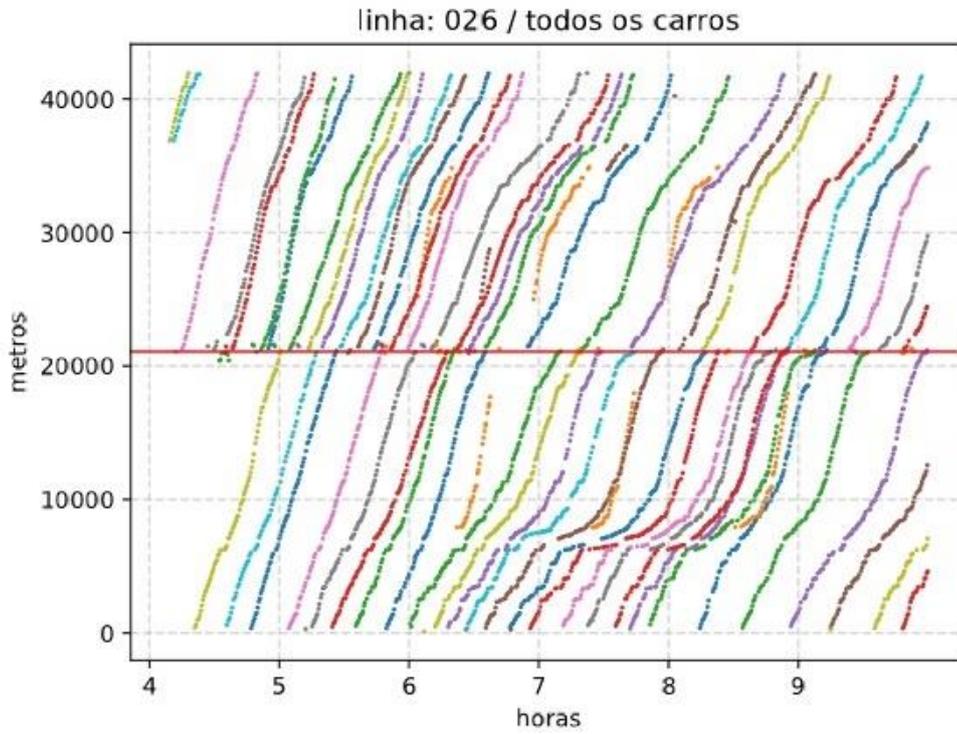
WONG, J. C. **Use of the general transit feed specification (GTFS) in transit performance measurement**. 2013. (Ph. D. thesis) Georgia Institute of Technology, 2013.

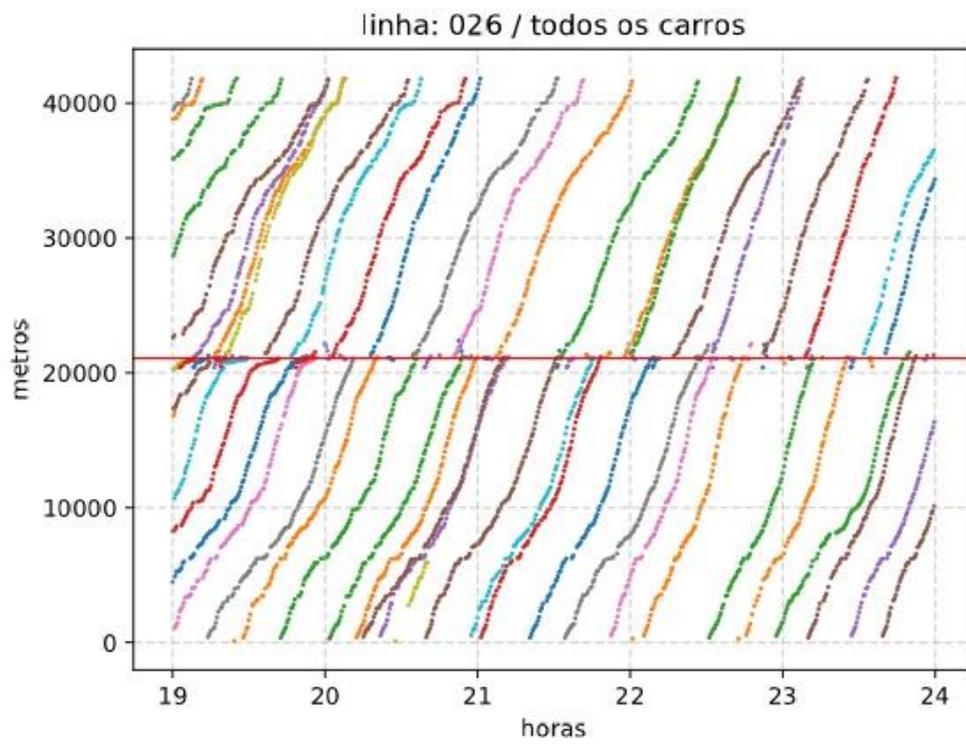
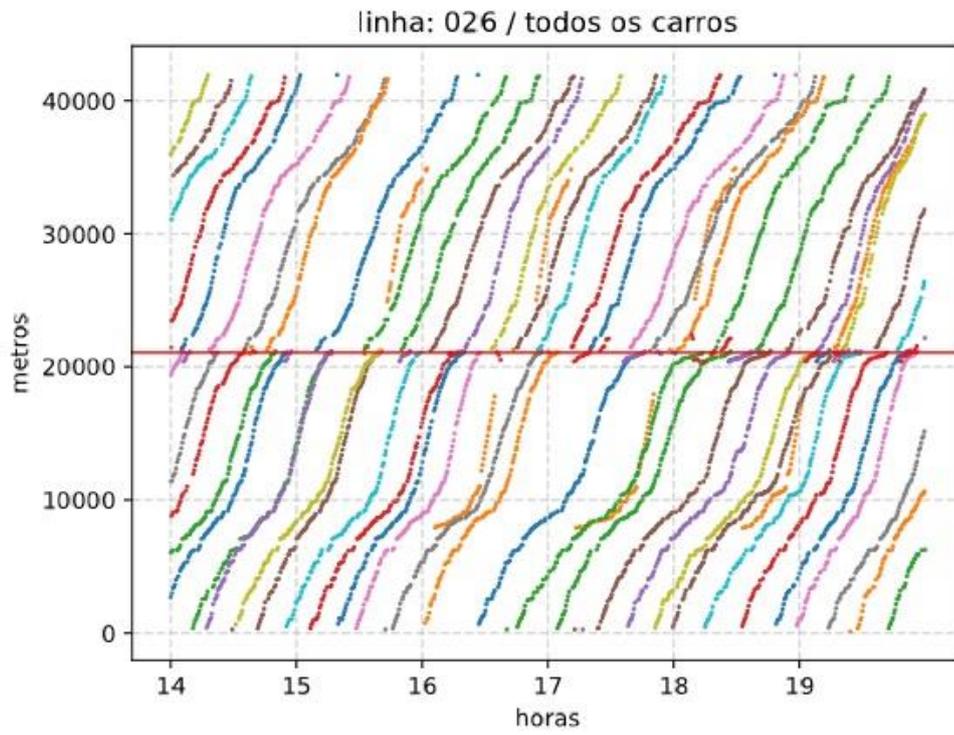
YE, P.; CHEN, Z.; XU, L.; YANG, H.; RAN, B. Use of GPS Data to Analyze Urban Bus Travel Time Variability In China. In: **Anais...** In: Transportation Research Board 95th Annual Meeting Transportation Research Board, Washington, DC, 2016. Disponível em: <<https://trid.trb.org/view/1393896>>. Acesso em: 27 jul. 2018.

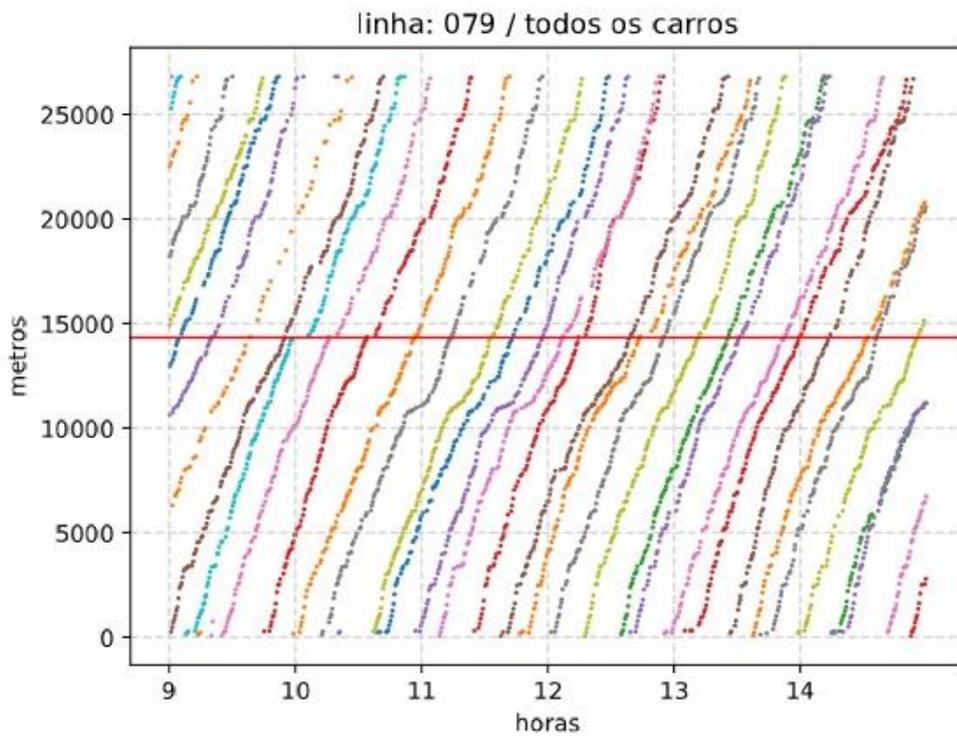
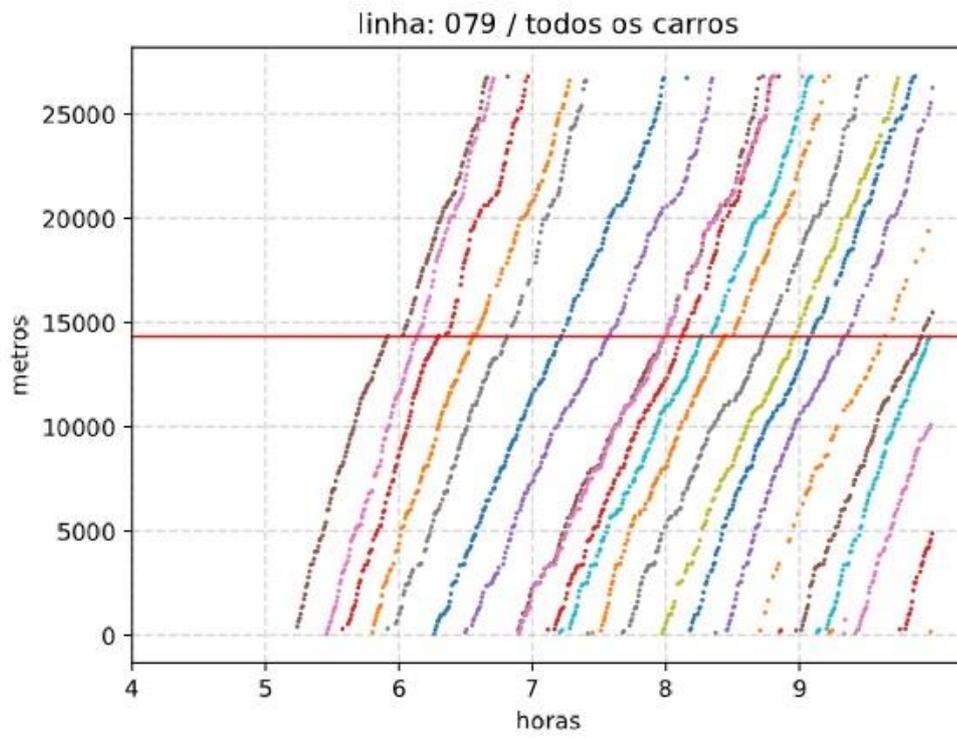
APÊNDICE A: MAPAS DE CALOR DAS VELOCIDADES OPERACIONAIS DAS LINHAS

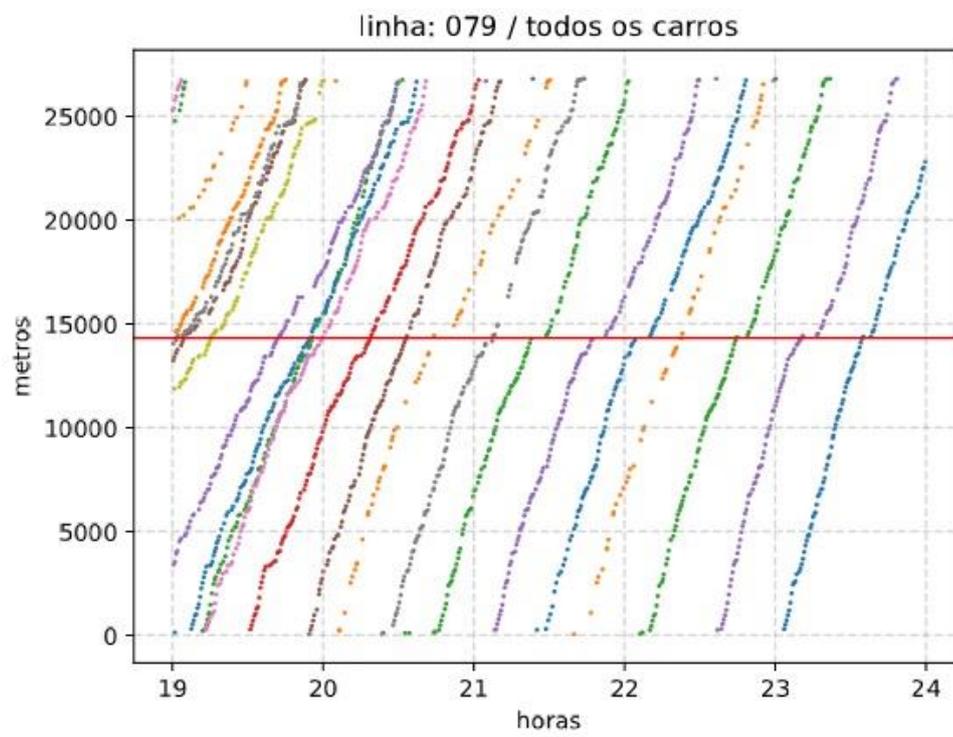
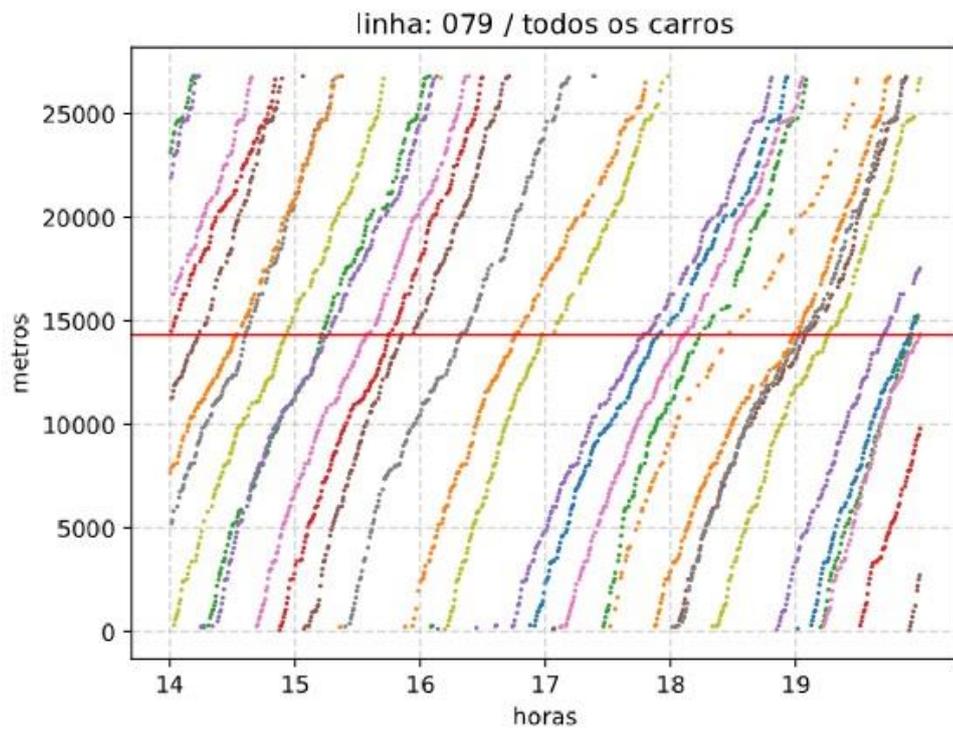




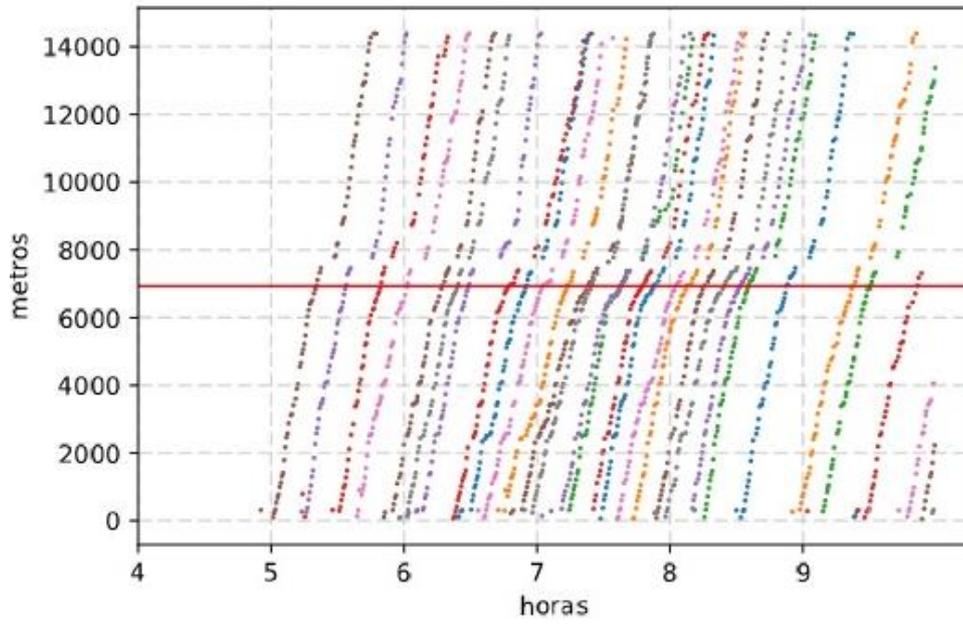
APÊNDICE B: GRÁFICOS DE MARCHA DAS LINHAS







linha: 200 / todos os carros



linha: 200 / todos os carros

