

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MATHEUS BIGARELLI DANTAS DA COSTA

**UMA ABORDAGEM GIS PARA CÁLCULO DE DESGASTE EM PROGRAMAS
DE MANUTENÇÃO**

CURITIBA

2022

MATHEUS BIGARELLI DANTAS DA COSTA

**UMA ABORDAGEM GIS PARA CÁLCULO DE DESGASTE EM PROGRAMAS
DE MANUTENÇÃO**

A GIS APPROACH TO WEAR CALCULATION IN MAINTENANCE PROGRAMS

Trabalho de Conclusão de Curso de Graduação apresentado como requisito para obtenção do título de Bacharel em Engenharia da Computação do Curso de Bacharelado em Engenharia da Computação da Universidade Tecnológica Federal do Paraná.

Orientador: Prof^ª. Dr^ª. Nádia Puchalski
Kozievitch

Coorientador: Prof^ª. Dr^ª. Keiko Veronica Ono
Fonseca

CURITIBA

2022



[4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Esta licença permite compartilhamento, remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

MATHEUS BIGARELLI DANTAS DA COSTA

**UMA ABORDAGEM GIS PARA CÁLCULO DE DESGASTE EM PROGRAMAS
DE MANUTENÇÃO**

Trabalho de Conclusão de Curso de Graduação
apresentado como requisito para obtenção
do título de Bacharel em Engenharia da
Computação do Curso de Bacharelado em
Engenharia da Computação da Universidade
Tecnológica Federal do Paraná.

Data de aprovação: 19/setembro/2022

Nádia Puchalski Kozievitch
Doutorado
Universidade Tecnológica Federal do Paraná

Keiko Veronica Ono Fonseca
Doutorado
Universidade Tecnológica Federal do Paraná

Juliana de Santi
Doutorado
Universidade Tecnológica Federal do Paraná

Ricardo Dutra da Silva
Doutorado
Universidade Tecnológica Federal do Paraná

CURITIBA
2022

Dedico este trabalho aos meus pais, que tornaram possível tudo que fui, sou e serei e deram a melhor criação que um filho poderia querer.

AGRADECIMENTOS

Na minha visão, o "eu" de hoje é o resultado de duas coisas combinadas que acumulamos em forma de experiências ao longo de nossa vida: o ambiente em que nos situamos e as escolhas que fazemos. Por isso, acredito ser o que sou hoje por causa de todas as pessoas que passaram e tiveram alguma presença na minha vida, mesmo que essa presença seja curta. Então agradeço a todos com quem já convivi, mesmo que brevemente, por ajudarem a me tornar a pessoa que sou hoje.

Com isso em mente, é lógico que algumas pessoas causam maiores impactos nas nossas vidas do que outras, conforme o tempo de convívio aumenta. Por isso, gostaria de agradecer a minha família pelo apoio e paciência nessa jornada, por todas as vezes que recusei algum convite, porque tinha prova ou algum trabalho para entregar ou simplesmente tinha que acompanhar a matéria.

Agradeço também a minha namorada, com quem vivo, pela paciência de me esperar até meia noite para passarmos o pequeno tempo restante dos dias juntos.

Agradeço a cada professor com quem tive aula na UTFPR, porque muito além do conteúdo, cada um deles mostrou também uma visão diferente do meio acadêmico. Alguns mais calmos e flexíveis, outros mais rígidos e metodológicos, cada um mostrou um jeito de pensarmos sobre o modo como aprendemos e trabalhamos.

Agradeço também a minha orientadora por todos os "a minha sugestão..." que me levaram a concluir este trabalho.

Agradeço, finalmente, a todos os outros servidores da UTFPR que ajudam a manter essa instituição de ensino, um lugar transformador de conhecimento e caráter.

All models are wrong, but some are useful.
(BOX, 1976)

RESUMO

O crescimento populacional, principalmente em áreas urbanas, implica no crescimento das necessidades de infraestrutura. O setor de transportes é uma área dessa infraestrutura que pede melhorias. Para empresas que têm veículos pesados, tais como caminhões, tratores, escavadeiras, entre outros, a manutenção dos mesmos é uma parte relevante da administração delas. A manutenção preditiva é um subcampo da manutenção que se preocupa em utilizar dados para prever variáveis de interesse. É composta pelas etapas de processamento de dados, diagnóstico, prognóstico e a etapa da decisão da manutenção. Este trabalho atua nas duas primeiras etapas: através de dados geográficos, propõe um índice de desgaste de veículos que possa ser utilizado em tarefas de diagnóstico. Como caso de exemplo, foram avaliados os dados da linha Inter II da cidade de Curitiba. Como resultado, foi calculado o índice de desgaste para o exemplo, além de ser feita uma análise do mesmo nas vias da linha que também tinham reclamações na Central 156.

Palavras-chave: *smart cities*; gis; altimetria; radares e lombadas; índice de desgaste.

ABSTRACT

The population growth, mainly in urban areas, implies the corresponding growth in the needs of the infrastructure. The transport sector is one of the areas within this infrastructure which need improvements. For companies that have heavy machinery, such as trucks, tractors, excavators, and others, the maintenance of such equipment is a relevant part of the machines' administration. Predictive maintenance is a subfield of maintenance that is concerned about using data to predict variables of interest. It is composed of the following steps: data processing, diagnosis, prognosis and decision making. This work plays a part in the first two steps, namely data processing and diagnosis. Geographical data is used to propose a vehicle wear index that can be used for such diagnosis tasks. As an example scenario, the "Inter II" bus line data of Curitiba was analysed. As result, the wear index for the vehicles that are assigned to that line were computed, and the data in the Central 156 database for that line was also analysed in conjunction with the proposed index.

Keywords: smart cities; gis; altimetry; speed bumps and radars; wear index.

LISTA DE FIGURAS

Figura 1 – Dinâmica das populações urbana e rural de 1950 a 2050. Retirada de: https://population.un.org/wup/Country-Profiles/	17
Figura 2 – Escopo do projeto.	19
Figura 3 – Sistema para predição do RUL de máquinas. Retirado de (ZHANG, 2018).	20
Figura 4 – Sistema de predição com dados GIS e <i>deep learning</i> . Retirado de (CHEN, 2021).	21
Figura 5 – Arquitetura do sistema de roteamento para pessoas com limitação de locomoção. Retirado de (BARCZYSZYN, 2018).	28
Figura 6 – Visão geral da metodologia aplicada. Autoria própria.	31
Figura 7 – Partes do desenvolvimento. Autoria própria	32
Figura 8 – Conjunto de dados utilizados. Autoria própria.	34
Figura 9 – Arquitetura do sistema. Baseado em (BARCZYSZYN, 2018).	35
Figura 10 – Cálculo do grau de inclinação para as interseções consecutivas da altimetria com a linha de ônibus. Retirado de (BARCZYSZYN, 2018).	37
Figura 11 – Componentes do grau de inclinação. Autoria própria.	37
Figura 12 – Graus de inclinação entre uma curva de nível de altimetria e outra. Autoria própria.	40
Figura 13 – Grau de inclinação máximo para cada trecho de 1 km na linha 022. Autoria própria.	40
Figura 14 – Redutores de velocidade e trechos da linha 022. Autoria própria.	41
Figura 15 – Índice de desgaste dado a altimetria ao longo do tempo para vários veículos. Autoria própria	42
Figura 16 – Índice de desgaste dados os redutores ao longo do tempo para vários veículos. Autoria própria	42
Figura 17 – Índice de desgaste. Autoria própria	43
Figura 18 – Ruas da linha 022 alvo de reclamações no 156. Autoria própria.	43
Figura 19 – Trechos com grau de inclinação máximo acima de 30. Autoria própria	44
Figura 20 – Redutores da linha 022.	44
Figura 21 – Análise conjunta dos dados do 156, da altimetria e dos redutores.	45

Figura 22 – Redutor provável de pertencer a outra rua que não a do trecho analisado.	
Autoria própria	45
Figura 23 – Radar mostrado no mapa do QGIS que não existe no Google Street View.	
Retirado do Google Street View.	46
Figura 24 – Radar que foi retirado entre maio de 2021 e janeiro de 2022.	47

LISTA DE TABELAS

Tabela 1 – Abreviaturas encontradas na base e seu correspondente.	40
Tabela 2 – Número de redutores por tipo	48
Tabela 3 – Correspondência encontradas entre a Central 156 e a linha 022, com o número de reclamações.	54
Tabela 4 – Correspondência encontradas entre a Central 156 e a linha 022, com o número de reclamações.	61

LISTAGEM DE CÓDIGOS FONTE

Listagem 1 – Divisão da linha em trechos de 1 km.	56
Listagem 2 – Cálculo dos graus de inclinação para as altimetrias.	57
Listagem 3 – Cálculo do grau de inclinação máximo para cada trecho.	58
Listagem 4 – Filtragem dos redutores da linha 022.	58
Listagem 5 – Filtragem dos redutores da linha 022.	58
Listagem 6 – Filtragem dos redutores da linha 022.	59
Listagem 7 – Tratamento da acumulação de esforços nos índices de desgaste.	60
Listagem 8 – Determinação dos nomes das ruas que compõem a linha 022.	61
Listagem 9 – Correspondência entre as ruas da Central 156 e da linha 022.	62

LISTA DE ABREVIATURAS E SIGLAS

Siglas

BD	Banco de Dados
FMECA	<i>Failure Modes Effects and Criticality Analysis</i> (inglês para análise do modo de falha, efeitos e criticidade)
GIS	<i>Geographic Information Systems</i>
IPPUC	Instituto de Pesquisa e Planejamento Urbano de Curitiba
ONU	Organização das Nações Unidas
RUL	<i>Remaining Useful Life</i>
SGBD	Sistema Gerenciador de Bancos de Dados
SQL	<i>Structured Query Language</i>
TBF	<i>Time Between Failures</i>
URBS	Urbanização de Curitiba

SUMÁRIO

1	INTRODUÇÃO	17
1.1	Objetivos	19
1.1.1	Objetivo geral	19
1.1.2	Objetivos específicos	19
1.2	Justificativa	20
1.3	Estrutura do trabalho	20
2	REFERENCIAL TEÓRICO	22
2.1	<i>Smart Cities</i>	22
2.1.1	Dados abertos	23
2.2	GIS	24
2.2.1	Geoprocessamento	25
2.3	Banco de dados	25
2.4	Manutenção preditiva	26
2.5	Trabalhos relacionados	27
3	MATERIAIS E MÉTODOS	30
3.1	Materiais/Tecnologias	30
3.2	Métodos	30
3.2.1	Revisão bibliográfica	30
3.2.2	Coleta de dados	31
3.2.3	Desenvolvimento	31
3.2.4	Análise de resultados	32
3.2.5	Conclusão	32
4	RESULTADOS	33
4.1	Descrição dos dados	33
4.2	Arquitetura e modelagem	33
4.2.1	<i>Analytics</i>	35
4.2.2	Divisão da rota em trechos	36
4.2.3	Cálculo do grau de inclinação	36
4.2.4	Grau de inclinação máximo do trecho	37
4.2.5	Contagem de redutores para um trecho	38

4.2.6	Determinação do trecho em que um veículo se encontra	38
4.2.7	Índices de desgaste intermediários	38
4.2.8	Índice de desgaste	39
4.2.9	Correspondência com dados da Central 156	39
4.3	Caso de uso	39
4.4	Discussões	44
5	CONCLUSÃO	49
	REFERÊNCIAS	51
	APÊNDICE A TABELA COMPLETA COM AS CORRESPONDÊNCIAS ENCONTRADAS ENTRE A LINHA 022 E OS REGISTROS DE RECLAMAÇÕES DA CENTRAL 156.	55
	APÊNDICE B IMPLEMENTAÇÃO DO SISTEMA	56
	B.1 Implementação do sistema	56
B.1.1	Trechos da linha 022	56
B.1.2	Graus de inclinação da linha 022	56
B.1.3	Graus de inclinação máximo para cada trecho da linha 022	57
B.1.4	Redutores em cada trecho da linha 022	57
B.1.5	Trecho em que os veículos se encontram	59
B.1.6	Índices de desgaste individuais	59
B.1.7	Índice de desgaste composto	59
B.1.8	Correspondência com dados da Central 156	60

1 INTRODUÇÃO

Segundo dados do Departamento de Assuntos Econômicos e Sociais da Organização das Nações Unidas (ONU)¹, mais da metade da população mundial residia em ambientes urbanos em 2018. Além disso, a taxa de crescimento da porcentagem de pessoas nos ambientes urbanos aumentou no decorrer dos anos. Note na Figura 1 (ONU, 2018) que a projeção para 2050 para o Brasil é de que mais que 80% da população resida em ambientes urbanos. Ainda com os dados fornecidos pela ONU, a estimativa é de 68% para a porcentagem da população mundial que morará em áreas urbanas até 2050 (ONU, 2018).

Esse crescimento gera a necessidade de se aumentar a eficiência na administração das cidades, um tópico recorrente no tema *Smart Cities*. A literatura tem diversas definições para o termo, mas a que será utilizada nesse trabalho é a que está disposta na Carta Brasileira para Cidades Inteligentes ²:

CIDADES INTELIGENTES São cidades comprometidas com o desenvolvimento urbano e a transformação digital sustentáveis, em seus aspectos econômico, ambiental e sociocultural, que atuam de forma planejada, inovadora, inclusiva e em rede, promovem o letramento digital, a governança e a gestão colaborativas e utilizam tecnologias para solucionar problemas concretos, criar oportunidades, oferecer serviços com eficiência, reduzir desigualdades, aumentar a resiliência e melhorar a qualidade de vida de todas as pessoas, garantindo o uso seguro e responsável de dados e das tecnologias da informação e comunicação.

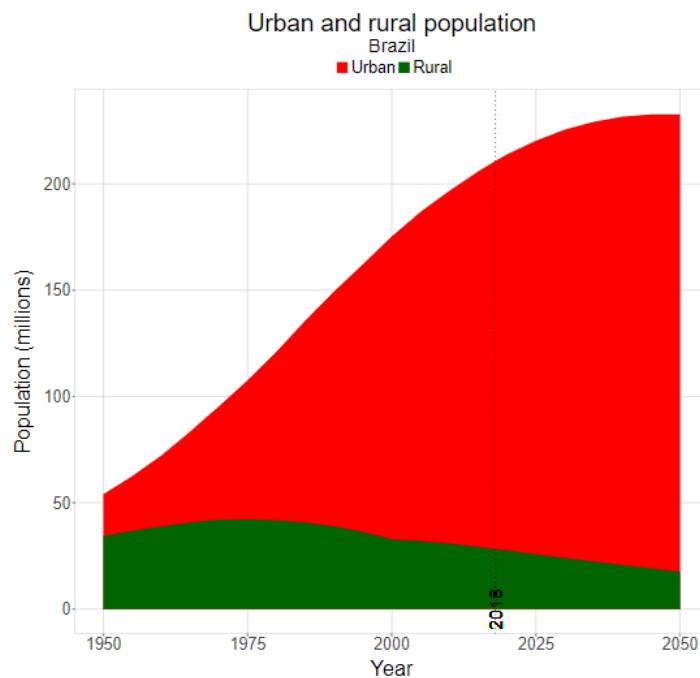


Figura 1 – Dinâmica das populações urbana e rural de 1950 a 2050. Retirada de: <https://population.un.org/wup/Country-Profiles/>

¹ <https://population.un.org/wup/Download/>

² <https://www.cartacidadesinteligentes.org.br/>

Para exemplificar um pouco melhor os problemas abordados no contexto de *Smart Cities*, tem-se o setor de transportes. O uso do transporte aumentará em 50% nos próximos 20 anos, de acordo com a Comissão Europeia (PRYTZ, 2014). Conforme (PRYTZ, 2014), com o aumento da necessidade de transporte, surge também o aumento correspondente da necessidade de infraestrutura. Para melhorar a eficiência e eficácia do setor, foram criados *Softwares* para gestão de frotas, políticas de manutenção preventiva, entre outros. No entanto, ainda há espaço para melhorias em problemas relacionados à administração e operação dos veículos.

A manutenção preditiva é um subcampo da manutenção que se preocupa em utilizar dados para prever variáveis de interesse, como por exemplo tempo restante de vida útil do veículo. Uma tarefa de manutenção preditiva é composta pelas etapas de processamento de dados, diagnóstico, prognóstico e a etapa da decisão da manutenção (BOUSDEKIS; MENTZAS, 2017). Este trabalho atua nas duas primeiras etapas. A partir de dados sobre rotas, foi proposto um índice de desgaste de veículos que pode ser utilizado em tarefas de diagnóstico.

Neste trabalho serão utilizadas somente as variáveis de altimetria e redutores de velocidade (lombadas e radares) de rotas escolhidas. As definições dessas variáveis do ponto de vista quantitativo, assim como este trabalho em si, foram inspirados em (BARCZYSZYN, 2018). Em (BARCZYSZYN, 2018), por exemplo, a métrica de esforço de um trecho é o esforço inerente a um trecho de uma rota utilizada. O índice de desgaste proposto refere-se aos esforços já exercidos por um determinado veículo que se acumularam neste mesmo veículo em forma de desgaste, à medida que este percorreu os trechos na sua rota, um após o outro.

Para enriquecer a análise, as métricas de esforço extraídas e os respectivos trechos associados foram correlacionados em uma base de dados de reclamações da Central 156 de Curitiba (referente a dados de buracos em vias), com o objetivo de buscar padrões e *insights*.

Como caso de estudo, um sistema foi desenvolvido com os dados de ônibus da Cidade de Curitiba e o estudo abrangeu as áreas de conhecimento dispostas na Figura 2:

1. *Geographic Information Systems* (GIS): Um sistema para lidar com dados espaciais (RINOS; KALEEL, 2012);
2. *Smart Cities*: Cidade que oferece serviços avançados e inovativos através do uso de Tecnologias de Informação e Comunicação. (HÖJER; WANGEL, 2015);
3. Banco de dados: Coleções de dados que descrevem tipicamente uma organização (RAMAKRISHNAN; GEHRKE, 2007).
4. Manutenção preditiva: Filosofia ou atitude que, posto de forma simples, usa as condições de operação de máquinas para otimizar a operação da mesma (MOBLEY, 2002).

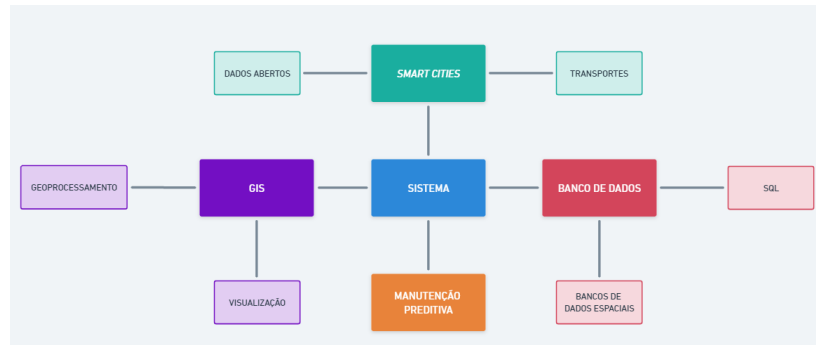


Figura 2 – Escopo do projeto.

1.1 Objetivos

1.1.1 Objetivo geral

O objetivo deste trabalho é desenvolver um subsistema baseado em dados geográficos para as etapas de processamento de dados e diagnóstico da manutenção preditiva, com o objetivo de criar um índice de desgaste de veículos.

1.1.2 Objetivos específicos

1. propor e computar índices de desgaste baseado em métricas de esforço despendido pelo veículo alvo.
2. propor um índice que agregue ou combine os índices de desgaste computados
3. conectar dados de outras bases para enriquecimento de análises de desgastes.
4. analisar e discutir possíveis métodos ou aplicações para o índice composto proposto.

A abordagem e metodologia adotada consistem na revisão da literatura das áreas mencionadas na Figura 2. A seguir, métricas de esforço são estabelecidas para dimensionar o esforço despendido pelo veículo em função de características da rota percorrida ou escolhida a partir da localização com variáveis altimetria e obstáculos de redução de velocidade.

Um índice de desgaste é estabelecido com o uso das métricas de esforço para caracterizar o desgaste total acumulado em função de cada uma das variáveis escolhidas.

Em um caso de uso, são estudadas rotas para identificar trechos de maior desgaste, ou seja, com maiores diferenças de altimetria e maior contagem de redutores. O índice é construído a partir da composição de índice de desgaste de altimetria e do índice de desgaste de redutores de velocidade. Dados de outras bases são conectadas para melhor entendimento do problema de caracterização da rota percorrida. Neste trabalho, optou-se por correlacionar dados da Central de Atendimento 156 sobre condições das vias.

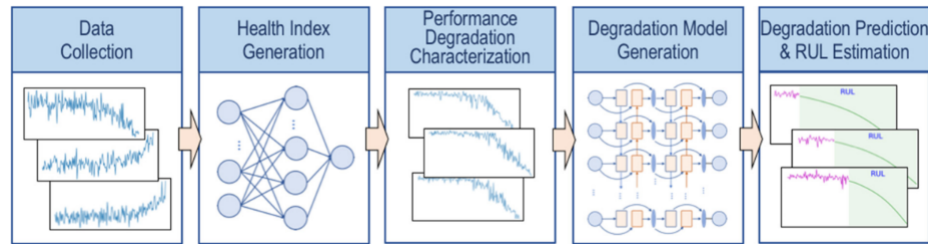


Figura 3 – Sistema para predição do RUL de máquinas. Retirado de (ZHANG, 2018).

Computados os índices a partir de dados reais, os resultados e conclusões são analisadas para refletir sobre a validação da abordagem e metodologia adotadas.

1.2 Justificativa

A literatura é vasta quando se procura informações sobre predição aplicada ao cenário de manutenção e falhas de veículos e outros tipos de máquinas. Por predição, no presente contexto, entende-se qualquer método ou modelo que tenha como entrada dados relevantes para o problema tratado e tenha como saída um valor estimado de uma variável de interesse. Diversos autores trazem soluções envolvendo predição de *Remaining Useful Life* (RUL) (tempo de vida útil restante até a falha) utilizando dados de sensores. Outros trazem predição do *Time Between Failures* (TBF) utilizando os registros de manutenção da máquina, os sistemas que apresentaram falha, o tempo decorrido e outras informações similares. A Figura 3 ilustra um exemplo desses sistemas de predição. Nela, está caracterizado um sistema baseado em dados, em que após a coleta, é gerado um *health index*, o qual é utilizado para análise e geração de um modelo de degradação, a partir do qual o RUL é estimado.

No entanto, alguns poucos trabalhos abordam o problema da perspectiva de dados geográficos, ou dados GIS. Alguns trazem essa abordagem de diagnóstico que incluem os dados GIS, mas apenas feita usando uma abordagem *black box* (CHEN, 2019). Nesse tipo de modelagem, não há como ter uma análise e interpretação do modelo.

Outros trabalhos já avançaram um pouco mais na análise dos dados GIS (CHEN, 2021). A Figura 4 mostra o sistema desenvolvido e as variáveis analisadas, incluindo informações de clima, tráfego, terreno e manutenção, além de um *health index* construído a partir dos dados de manutenção utilizando o modelo de risco proporcional de Cox como entrada para um modelo de *deep learning* cujo objetivo era estimar o RUL.

1.3 Estrutura do trabalho

O Capítulo 2 apresenta a revisão do conteúdo teórico necessário para o projeto, além dos trabalhos relacionados. O Capítulo 3 traz os materiais, as tecnologias e os métodos utilizados no decorrer do trabalho. O Capítulo 4 apresenta os resultados e discussões, trazendo

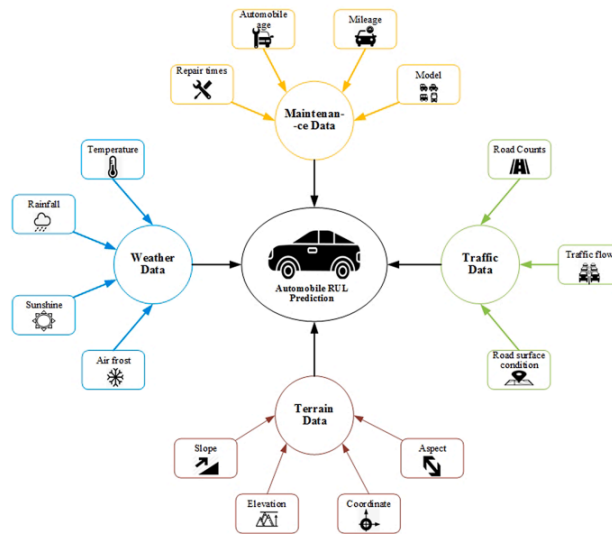


Figura 4 – Sistema de previsão com dados GIS e *deep learning*. Retirado de (CHEN, 2021).

figuras também para ilustrar ambos, sobre a solução proposta. Por fim, o Capítulo 5 apresenta as conclusões do projeto.

2 REFERENCIAL TEÓRICO

Este capítulo trata dos conceitos abordados no trabalho: GIS, *Smart Cities*, Manutenção Preditiva e Banco de Dados.

2.1 *Smart Cities*

Cidades inteligentes, ou *smart cities*, são um fenômeno relativamente recente se comparado com a idade das civilizações organizadas, o calendário gregoriano ou qualquer outra medida de tempo que se tenha para remeter às primeiras sociedades. Por volta de 1972, começaram a surgir problemas relacionados à globalização e, mais especificamente, a globalização dos problemas ambientais (HÖJER; WANGEL, 2015). As consequências, no entanto, eram locais e um exemplo é a produção de iPhones: apesar da Apple ser uma empresa americana, o produto é feito na Índia (ROSA, 2022). Ainda, conforme (MODERNO, 2015):

Segundo a Companhia de Saneamento Básico do Estado de São Paulo (Sabesp), em um banho econômico de chuveiro elétrico, utilizamos aproximadamente 15 litros de água. Portanto, poderíamos tomar 850 banhos com a mesma quantidade de água utilizada na produção de um único smartphone.

O importante aqui não são os números em si, mas a ideia de que existem problemas ambientais locais (na Índia, neste caso) que surgiram em decorrência da globalização a partir da década de 70. No entanto, à medida que o tempo passou, as consequências deixaram de ser locais e começaram a se tornar globais e concentrando-se no ambiente urbano (HÖJER; WANGEL, 2015).

A necessidade de se criar meios, métodos, grupos, entre outros, para buscar soluções no âmbito de desenvolvimento urbano sustentável e cidades sustentáveis é atualmente apoiada em outro fenômeno em paralelo: as tecnologias de informação e comunicação (*Information and Communication Technologies - ICT*). As ICTs ampliam a possibilidade de se reduzir custos ligados à comunicação e informação, além de possibilitar novos produtos, serviços e modelos de negócio.

Finalmente, no final do século XX, começam a surgir os primeiros movimentos que dariam origem ao fenômeno das cidades inteligentes, tais como o Movimento *Smart Growth* (1990) e o *Cybernetically Planned Cities* (1960). O termo Cidade inteligente é associado desde como uma cidade que oferece serviços avançados e inovativos através do uso de ICTs (HÖJER; WANGEL, 2015) ou como uma cidade que integra as infraestruturas física, de TI e social para se aproveitar da inteligência coletiva da mesma (CHOURABI, 2012).

Para este trabalho, será adotada a definição disposta na Carta Brasileira para Cidades Inteligentes ¹ já mencionada no Capítulo 1.

¹ <https://www.cartacidadesinteligentes.org.br/>

2.1.1 Dados abertos

Segundo definição da *Open Knowledge International* disposta no Portal Brasileiro de Dados Abertos ²:

... dados são abertos quando qualquer pessoa pode livremente acessá-los, utilizá-los, modificá-los e compartilhá-los para qualquer finalidade, estando sujeito a, no máximo, a exigências que visem preservar sua proveniência e sua abertura.

Os dados abertos são pautados por três leis, propostas por David Eaves, não sendo estas propriamente as leis do direito, mas sim testes que podem ser feitos para validar se um dado é de fato aberto ou não. As três leis são:

1. Se o dado não pode ser encontrado e indexado na Web, ele não existe;
2. Se não estiver aberto e disponível em formato compreensível por máquina, ele não pode ser reaproveitado;
3. Se algum dispositivo legal não permitir sua replicação, ele não é útil.

Embora tenham sido propostas inicialmente para dados governamentais, elas podem ser aplicadas no caso mais geral (DADOS.GOV.BR, 2022). Os dados são considerados abertos se:

1. São completos - Todos os dados públicos são disponibilizados. Dados são informações eletronicamente gravadas, incluindo, mas não se limitando a, documentos, bancos de dados, transcrições e gravações audiovisuais. Dados públicos são dados que não estão sujeitos a limitações válidas de privacidade, segurança ou controle de acesso, reguladas por estatutos;
2. São primários - Os dados são publicados na forma coletada na fonte, com a mais fina granularidade possível, e não de forma agregada ou transformada;
3. São atuais - Os dados são disponibilizados o quão rapidamente seja necessário para preservar o seu valor;
4. São acessíveis - Os dados são disponibilizados para o público mais amplo possível e para os propósitos mais variados possíveis;
5. São processáveis por máquina - Os dados são razoavelmente estruturados para possibilitar o seu processamento automatizado;
6. Possuem acesso não discriminatório - Os dados estão disponíveis a todos, sem que seja necessária identificação ou registro;

² <https://dados.gov.br/pagina/dados-abertos>

7. Estão em formato não proprietário - Os dados estão disponíveis em um formato sobre o qual nenhum ente tenha controle exclusivo;
8. Possuem licenças livres - Os dados não estão sujeitos a restrições por regulações de direitos autorais, marcas, patentes ou segredo industrial. Restrições razoáveis de privacidade, segurança e controle de acesso podem ser permitidas na forma regulada por estatutos.

Um dos conjuntos de dados abertos que estão disponíveis na web são os dados referentes ao transporte público de Curitiba³. Neste conjunto de dados há informações geocodificadas sobre posição, rota, itinerários, etc. dos ônibus que circulam pela cidade. Os dados de posição dos veículos e de rotas foram indispensáveis para o desenvolvimento deste trabalho.

2.2 GIS

Outro conceito relevante para este trabalho é o dos sistemas de informação geográfica (*Geographic Information Systems* - GIS). Algumas das definições utilizadas por diferentes autores (RINOS; KALEEL, 2012) estão listadas abaixo:

1. Um sistema para lidar com dados espaciais;
2. Um sistema computadorizado para capturar, armazenar, recuperar, analisar e visualizar dados espaciais;
3. Um sistema para coleta, armazenamento, recuperação por demanda, transformação e visualização de dados espaciais do mundo real.

Os objetivos de um GIS, também segundo (RINOS; KALEEL, 2012), são:

1. Maximizar a eficiência do planejamento e tomada de decisão;
2. Prover meios eficientes de se distribuir dados e trabalhar com eles;
3. Eliminar dados redundantes de uma base de dados. Minimizar duplicação;
4. Capacidade de integrar dados de diferentes fontes;
5. Capacidade de fazer consultas e análises complexas com dados georreferenciados para geração de novas informações.

Os componentes de um GIS, ainda segundo o mesmo autor, são cinco: hardware, software, dados, pessoas e métodos. O componente de dados é especialmente importante para este trabalho, pois aqui encontra-se o banco de dados.

³ <https://www.curitiba.pr.gov.br/dadosabertos/busca/?termo=transporte>

2.2.1 Geoprocessamento

A coleta, processamento e visualização de informações geográficas é uma das partes principais do projeto e aplica-se aqui a seguinte definição de geoprocessamento:

...é uma poderosa ferramenta computacional, que processa dados geograficamente referenciados e pode ser bastante útil na abordagem integrada, essencial ao gerenciamento dos recursos naturais (HAMADA; GONÇALVES, 2007).

Com geoprocessamento é possível fazer inúmeras aplicações diferentes (HAMADA; GONÇALVES, 2007), sendo algumas delas: manejo e conservação de recursos naturais, gestão das explorações agrícolas, planejamento de área urbana, gestão das instalações, entre outros.

2.3 Banco de dados

Um Banco de Dados (BD) é uma coleção de dados. Tipicamente, descrevem as atividades de uma ou mais organizações (RAMAKRISHNAN; GEHRKE, 2003). É importante notar que, apesar da definição de 'coleção de dados' ser um pouco genérica, não são quaisquer dados aleatoriamente armazenados que podem ser chamados de BD (ELMASRI; NAVATHE, 2017).

A tecnologia mais comumente utilizada para se armazenar e manipular dados é o Sistema Gerenciador de Bancos de Dados (SGBD). A definição de SGBD em (RAMAKRISHNAN; GEHRKE, 2003) é de um software projetado para auxiliar na manutenção e utilização de grandes coleções de dados. Já para Elmasri e Navathe (2017), um SGBD é um sistema de software de propósito geral que facilita os processos de definição, construção, manipulação e compartilhamento de BD entre vários usuários e aplicações.

Há alguns tipos de SGBDs disponíveis para escolha, mas o relacional é o mais usado (RAMAKRISHNAN; GEHRKE, 2003). Este é o tipo de SGBD utilizado também neste trabalho.

Os SGBDs, tais como o PostgreSQL, podem ser estendidos para agregar mais funcionalidades além das fornecidas e especificadas pelo padrão *Structured Query Language* (SQL). Um exemplo disso é justamente o alvo deste trabalho: as extensões para banco de dados espaciais.

Por objetos espaciais, ou dados espaciais, entende-se pontos, linhas, retângulos, polígonos, cubos e outras entidades geométricas multidimensionais. Estes tipos de dados ocupam uma certa região do espaço, chamada extensão espacial, que é caracterizada por sua localização e por seu limite (ELMASRI; NAVATHE, 2017).

Para suportar bancos de dados espaciais, existe um padrão SQL, o SQL/MM: Spatial. Este padrão suporta pontos, linhas e dados bidimensionais. Ele foi projetado visando a compatibilidade com padrões relacionados como aqueles propostos pelo *Open GIS Consortium*, uma

organização internacional que publica padrões aceitos pela comunidade (RAMAKRISHNAN; GEHRKE, 2007).

Dentre as funções implementadas em uma extensão de banco de dados espaciais, por exemplo o PostGIS, tem-se:

1. Distância entre pontos, ponto e linha, linha e linha. Distância entre quaisquer duas entidades geométricas (ou geometria), para ser mais genérico;
2. Criação de uma região em volta de uma geometria (*buffer*) englobando uma determinada distância da geometria;
3. Cálculo de interseções booleano (há interseção/não há interseção) e das interseções em si – o ponto como interseção entre retas, por exemplo;
4. Operadores de relação de continência, como um ponto contido em uma reta, por exemplo.

Estes são apenas alguns exemplos, mas podem haver muito mais. Apenas com a leitura da documentação completa pode ser ter uma real noção de todas as ferramentas que uma extensão pode fornecer.

2.4 Manutenção preditiva

A manutenção preditiva é uma filosofia ou atitude que, posta de forma simples, usa as condições de operação de máquinas para otimizar a operação da mesma (MOBLEY, 2002). Por condições de operação, entende-se que, de alguma forma, dados de sensores são lidos para monitorar a operação da máquina continuamente. Esse monitoramento, seja das condições mecânicas, da eficiência do sistema ou outros indicadores, é o que possibilita a determinação do tempo médio entre as falhas da máquina, ou perda de eficiência dela (MOBLEY, 2002).

A manutenção preditiva é uma estratégia de manutenção que está ganhando interesse das empresas (BOUSDEKIS; MENTZAS, 2017). Esse aumento de interesse está ancorado na ideia de que o monitoramento contínuo da operação da máquina em questão proverá todos os dados necessários para assegurar a maximização do tempo entre as manutenções executadas (reduzindo tempo ocioso e custos) e também a minimização do número e do custo de manutenções não planejadas causadas por falhas críticas (MOBLEY, 2002).

A administração de máquinas e a manutenção pode ser realizada seguindo a metodologia proposta em (BOUSDEKIS; MENTZAS, 2017). No trabalho, os autores estabelecem as seguintes etapas: processamento de dados; diagnóstico; prognóstico; tomada de decisão (e posterior tomada de ação, *Failure Modes Effects and Criticality Analysis* (inglês para análise do modo de falha, efeitos e criticidade) (FMECA) e administração de operações industriais).

A etapa de processamento consiste na aquisição dos dados e na manipulação e processamento necessários para armazenar de forma integrada esses dados, possivelmente de diferentes fontes.

A etapa de diagnóstico consiste em "detecção de estado" e "avaliação de saúde". Esses passos se preocupam em aplicar técnicas para detectar um estado fora do que seria considerado como normalidade da operação, ou seja, condições não usuais de operação.

A etapa de prognóstico inclui predição de estados futuros de um sistema inteiro ou de componentes de um sistema. Esta predição pode ser, por exemplo, do tempo até a falha ou o tempo de vida útil restante, etc.

A etapa de tomada de decisão é composta por algoritmos, para que essa tomada de decisão possa ser feita de forma automática, ou para que possam ser providas recomendações de ações futuras baseando-se em informações do prognóstico e conhecimento de especialistas na área.

A etapa de tomada de ação, ou de implementação, como também é chamada, inclui um mecanismo para coleta, armazenamento, agregação, análise e visualização dos dados e das informações geradas para tomada de decisão.

A etapa FMECA é a etapa em que a operação da máquina é analisada para se determinar os possíveis modos de falha.

Finalmente, a etapa de administração de operações industriais é a parte gerencial da metodologia, abrangendo áreas como planejamento de produção, logística e controle de qualidade.

É importante ressaltar que este trabalho se concentrou apenas nas duas primeiras etapas, processamento de dados e diagnóstico. Mais especificamente, a coleta e as transformações necessárias para o início dos cálculos do índice de desgaste compõem a etapa de processamento de dados e o cálculo do índice e avaliação dos resultados compõem a etapa de diagnóstico.

2.5 Trabalhos relacionados

Aplicações semelhantes, que utilizaram abordagens baseadas no uso de dados para análise de operação de máquinas, propuseram conceitos que serão reutilizados ou adaptados neste trabalho.

Na Figura 5 é apresentado um sistema que inclui em sua análise a altimetria, as rampas e a dificuldade que uma pessoa tem ao percorrer esse trecho (BARCZYSZYN, 2018). A ideia de dificuldade em subidas e descidas (pelo dado da altimetria) acentuadas foi traduzida no trabalho como um grau de inclinação (BARCZYSZYN, 2018). Ele será reutilizado na íntegra no presente trabalho. O grau de inclinação, tal como definido no trabalho de referência e reutilizado aqui, é cem vezes a distância vertical entre dois pontos dividida pela distância horizontal entre eles, tal como pode ser observado na Equação 1. O número é expresso em porcentagem.

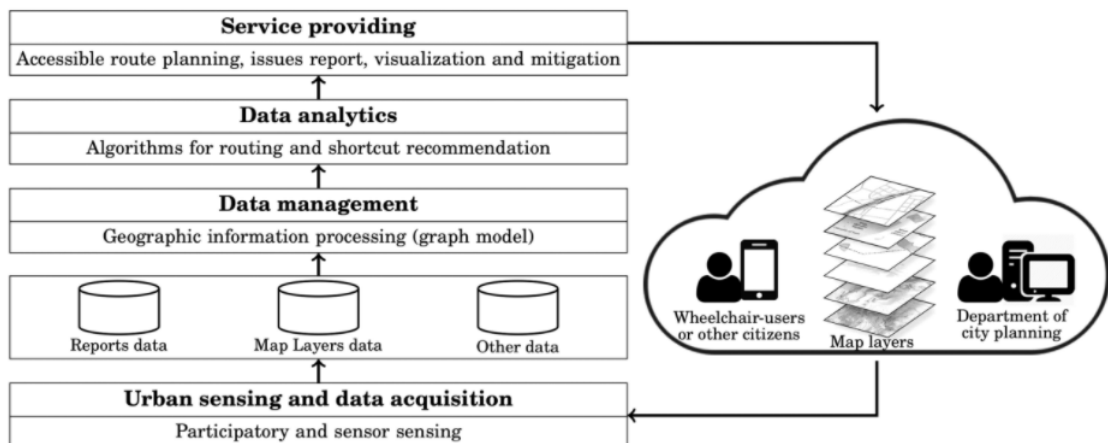


Figura 5 – Arquitetura do sistema de roteamento para pessoas com limitação de locomoção. Retirado de (BARCZYSZYN, 2018).

$$g(p_i, p_j) = 100 \frac{d_v(p_i, p_j)}{d_h(p_i, p_j)} \quad (1)$$

Na Equação 1, $g(p_i, p_j)$ é o grau de inclinação, $d_v(p_i, p_j)$ é a distância vertical entre dois pontos p_i e p_j e $d_h(p_i, p_j)$ é a distância horizontal entre esses pontos.

Um outro conceito que foi utilizado é o de redutores de velocidade. Para o presente trabalho, um redutor de velocidade é um objeto que, por alguma razão, força a redução de velocidade dos ônibus que passam por ele e a subsequente aceleração para manter os padrões de velocidade da via. Pode ser um dos seguintes:

1. Faixa elevada;
2. Linha de trem;
3. Lombada;
4. Ponto de ônibus;
5. Radar;
6. Terminal.

Embora existisse a possibilidade de alguns redutores não estarem em conformidade com a lei, este fato não foi considerado nem analisado para o desenvolvimento deste trabalho. Como exemplo da não conformidade com a lei desses redutores, é possível citar a resolução 600/2016 do DENATRAN, que especifica que a distância mínima do redutor à interseção é de 15 metros, no caso de uma ondulação transversal (lombada) próxima a um cruzamento (SIMETTE, 2018). Para este trabalho, foram considerados todos os redutores presentes na base de dados.

Embora este trabalho traga os conceitos de grau de inclinação de (BARCZYSZYN, 2018) e de redutores de velocidade de (SIMETTE, 2018), ele se diferencia de ambos os outros na

combinação de ambos os conceitos em métricas de esforço que, mais tarde, seriam utilizadas na construção de um índice de desgaste de veículos.

Uma outra aplicação importante, embora não fosse focada em análise de operação de máquinas, é (BARCLAY, 2018). No artigo, foi proposto um índice de desempenho para hospitais agregando algumas variáveis diferentes. No trabalho de Barclay, a solução do problema resultaria nos pesos que seriam dados a cada variável para compor o índice final. Nessa abordagem é possível determinar propriedades que se espera que o índice tenha (por exemplo monotonicidade), e, então calcular os pesos automaticamente baseado nessa expectativa de que o índice tenha essas propriedades.

3 MATERIAIS E MÉTODOS

Este capítulo apresenta a metodologia utilizada para solucionar o problema, assim como os materiais utilizados.

3.1 Materiais/Tecnologias

1. **PostgreSQL (9.6.13)** - SGBD usado para armazenar as tabelas utilizadas;
2. **PostGIS (2.3.1)** - Extensão do PostgreSQL para trabalhar com bancos de dados espaciais;
3. **QGIS (3.26)** - Ferramenta de visualização de dados espaciais;
4. **DBeaver (21.3.5)** - Ferramenta utilizada para interação básica com o banco de dados e a realização de consultas no mesmo;
5. **Whimsical** - Utilizado para diagramação e confecção de ilustrações dos processos utilizados;
6. **Git (2.25.1) e GitHub** - Usados para versionar o código desenvolvido;
7. **DBT Data Build Tool (1.1.0)** - Utilitário de linha de comando usado para criação das tabelas, execução de consultas e auxílio no desenvolvimento.
8. **Python (3.8.10)** - Linguagem de programação usada para escrever o *crawler* que buscou os dados faltantes e os algoritmos para o cálculo dos índices para os veículos.
9. **Scrapy¹ (2.6.1)** - Framework da linguagem Python para *web-crawling*.

3.2 Métodos

A Figura 6 apresenta a visão geral da metodologia aplicada neste trabalho. Embora este trabalho tenha sido realizado em cinco etapas, note que os 4 passos após a revisão bibliográfica poderiam ser aprimorados após revisões.

3.2.1 Revisão bibliográfica

A primeira etapa foi fazer a revisão bibliográfica dos principais conceitos englobados pelo escopo do trabalho, ou seja, *smart cities*, GIS, bancos de dados e manutenção preditiva, apresentados no Capítulo 2.

¹ <https://scrapy.org/>



Figura 6 – Visão geral da metodologia aplicada. Autoria própria.

3.2.2 Coleta de dados

Após os estudos terem sido concluídos, foi feita a etapa de coleta de dados. Nesta etapa, foram feitas verificações dos dados previamente disponíveis na base de dados e a coleta dos faltantes (como a geolocalização dos ônibus da URBS) para iniciar o desenvolvimento do trabalho. Para buscar os dados necessários, foi feito um *crawler* para buscar arquivos entre determinadas datas no servidor do C3SL². Detalhes da base de dados serão apresentados na próxima seção.

3.2.3 Desenvolvimento

A etapa de desenvolvimento, mostrada na Figura 7, pode ser dividida nas seguintes partes:

1. Análise das tabelas individuais;
2. Escolha e divisão em trechos iguais da linha de ônibus de interesse;
3. Projeto e implementação dos algoritmos para extração das métricas de esforço;
4. Análise das correspondências com reclamações da Central 156;
5. Projeto e correlação dos veículos com o trecho percorrido em suas linhas de ônibus;
6. Projeto e implementação do algoritmo de acumulação de esforço no formato de índice de desgaste.

² <http://dadosabertos.c3sl.ufpr.br/curitibaurbbs/>

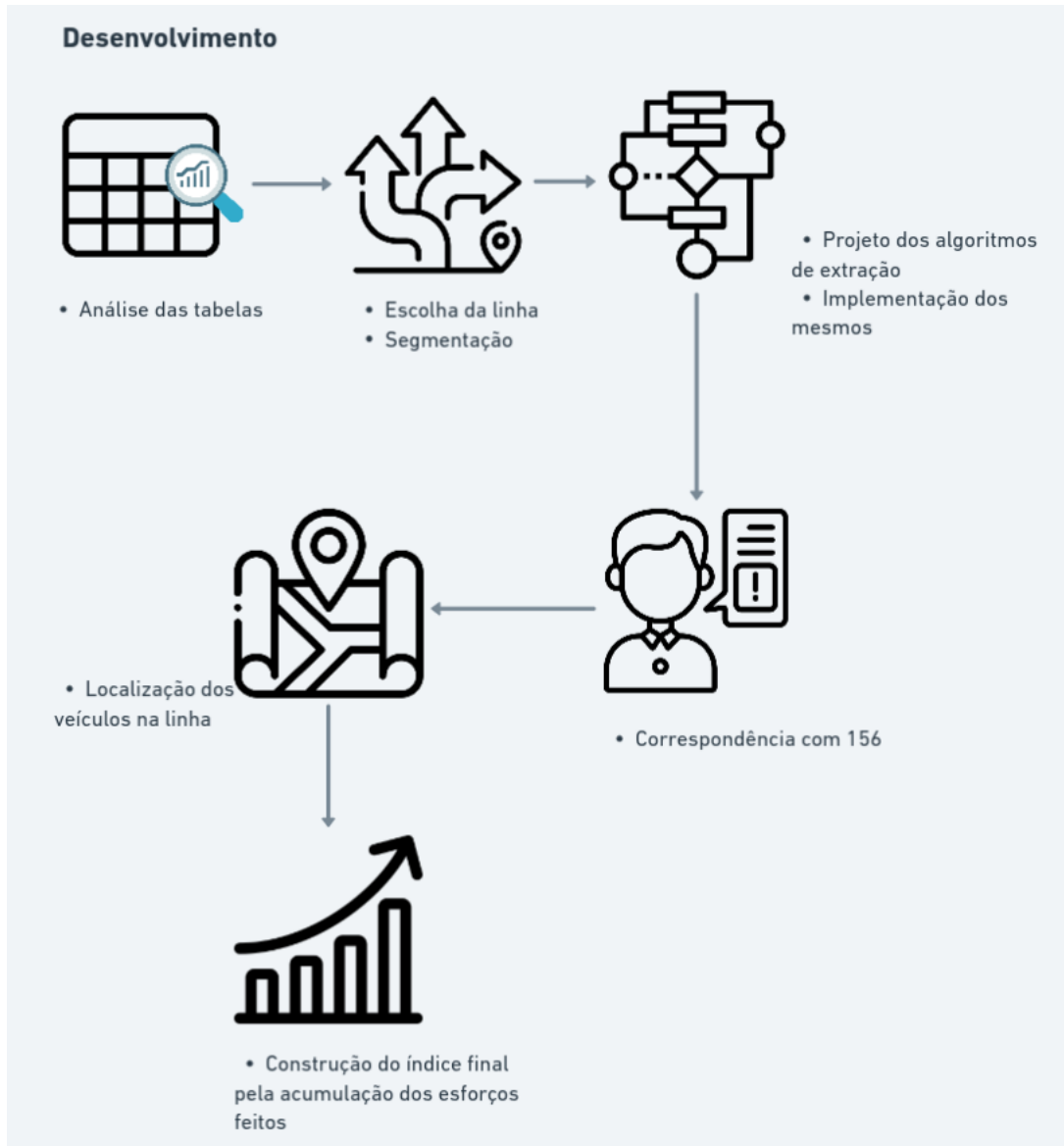


Figura 7 – Partes do desenvolvimento. Autoria própria

3.2.4 Análise de resultados

Uma vez que foi finalizado o desenvolvimento, passou-se para a etapa de análise de resultados. Esta consistiu em extrair informações de valor do modelo criado, assim como determinar suas limitações.

3.2.5 Conclusão

A etapa de conclusão apresenta os resultados obtidos, as limitações e as sugestões de trabalhos futuros.

4 RESULTADOS

Este capítulo apresenta a descrição dos dados, a arquitetura e modelagem, o caso de uso e discussões.

4.1 Descrição dos dados

A Figura 8 apresenta o conjunto de dados utilizados, originários do Instituto de Pesquisa e Planejamento Urbano de Curitiba (IPPUC)¹, Urbanização de Curitiba (URBS)² e Prefeitura de Curitiba. A tabela `eixo_rua` traz uma relação entre os nomes das vias da cidade de Curitiba e suas localizações em formato de geometria. Ela possui diferentes informações acerca da via, tais como o sistema viário, o bairro, a regional, o CEP, etc.

A tabela `lombada_radar` traz informações sobre redutores de velocidade. Não apenas lombadas e radares, como também travessias elevadas, linhas de trem, pontos de ônibus e terminais.

A tabela `linha_de_onibus` traz as linhas de ônibus de Curitiba. Além da geometria da linha, ainda existem informações sobre as categorias das linhas, sentido, entre outras.

A tabela `central156` possui dados para o período entre 01/12/2016 e 01/10/2017. Alguns desses dados são o tipo de solicitação, a data em que foi feita, horário, o assunto, o logradouro ao qual corresponde a reclamação, etc. Dessa tabela, foram utilizados somente os dados que possuem as reclamações sobre buracos nas vias (categoria "TAPA BURACO").

A tabela `veiculos` traz os dados dos veículos para o período entre 28/02/2022 e 06/03/2022. Dentre os atributos, podemos citar os identificadores dos veículos, a latitude e a longitude, a data e hora da amostragem e o código da linha em que o veículo está operando. Para os testes foram utilizados somente os dados da linha Inter II.

A tabela `alt_curva_de_nivel` traz as curvas de nível indicando a elevação da região em relação ao nível do mar.

4.2 Arquitetura e modelagem

A arquitetura do sistema é apresentada na Figura 9. A camada de coleta foi onde foram concentrados os esforços de coleta dos dados e armazenamento no banco de dados. A camada de gerenciamento se preocupou em realizar os processamentos necessários prévios ao início do desenvolvimento dos algoritmos e da computação dos resultados. Os dados descritos na seção anterior (ilustrados na etapa de coleta de dados) foram armazenados em um banco de dados PostgreSQL (etapa gerenciamento de dados). A camada de *Analytics* concentrou a maior parte do tempo gasto, pois nela estão os algoritmos e consultas necessários para realização das

¹ <https://www.ippuc.org.br/>

² <https://www.urbs.curitiba.pr.gov.br/>

eixo_rua	lombada_radar	linha_de_onibus	central156
gid (pk)	gid (pk)	gid (pk)	solicitacao (pk)
objectid	num equip	objectid	tipo
chave	tipo	layer	orgao
codvia	bairro	cd_categor	data
noinicio	endereco	categoria	horario
nofim	cruzamento	cd_linha	assunto
status	latitude	linha	subdivisao
nmvia	longitude	data	descricao
nmvia_ante	comprimento	fonte	logradouro_ass
sist_viari	largura	seta_senti	bairro_ass
sist_via_1	altura	shape_len	regional_ass
hierarquia	tiporadar	geom	meio_resposta
status_den	velocidade_max		observacao
obs	geom (pk)	<i>alt_curva_de_nivel</i>	sexo
obs2		gid (pk)	bairro_cidadao
cod_bairro	veiculos	layer	regional_cidadao
bairro_e	veic (pk)	elevation	orgao_resp
bairro_d	dthr (pk)	data_voo	resposta_final
cod_reg	lat	data_rest	
regional_e	lon	resp_tec	
regional_d	cod_linha	escala	
cep_e	geom	curvas_m	
cep_d		shape_len	
cep_info		geom	
cep_fonte			
shape_len			
geom			

Figura 8 – Conjunto de dados utilizados. Autoria própria.

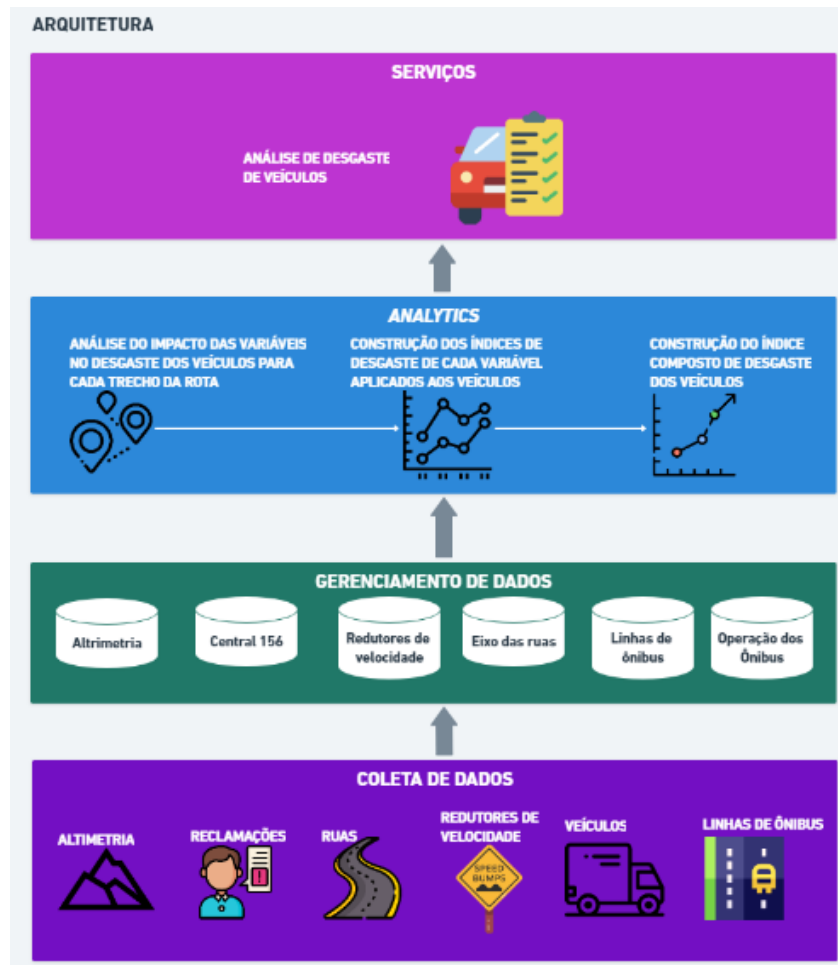


Figura 9 – Arquitetura do sistema. Baseado em (BARCZYSZYN, 2018).

análises e computação dos índices de desgaste. Na etapa de *Analytics*, é realizado o cálculo do índice de desgaste. Para esse índice, foram utilizadas as duas variáveis disponíveis: altimetria e redutores de velocidade. Por fim, na camada de serviços, os resultados foram analisados.

4.2.1 *Analytics*

Primeiramente, a rota escolhida (Inter II) foi dividida em trechos de 1km cada. Para cada trecho foram filtrados e relacionados os dados de altimetria e redutores correspondentes através das informações de localização codificadas nas colunas geométricas de cada tabela. A próxima etapa é responsável por calcular a quantidade de esforço (a métrica de esforço) para cada variável (altimetria e redutores), para cada trecho. A métrica de altimetria é a mesma utilizada em (BARCZYSZYN, 2018) e a métrica de redutores foi inspirada nela. As Equações 1 e 4 mostram como é feito o cálculo da métrica de esforço de altimetria.

$$i_d^{(A)}(n) = i_d^{(A)}(n - 1) + m_e^{(A)} \quad (2)$$

onde $m_e^{(A)}$ é essa métrica de altimetria do trecho atual, e $i_d^{(A)}$ é o valor acumulado total dos esforços realizados anteriormente.

Finalmente, com os dois índices de desgaste anteriores, um para altimetria e outro para redutores de velocidade (para cada trecho), foi possível computar o índice de desgaste através de uma função que combinasse ambos. Nesse caso, foi calculada a média aritmética.

As etapas estão descritas no Algoritmo 1:

Algoritmo 1 – Algoritmo de cálculo de índice de desgaste

- 1: $trecho_atual \leftarrow determinar_trecho_atual(veiculo.geom)$
 - 2: $m_e^{(A)} \leftarrow determinar_metrica_esforco_altimetria(trecho_atual)$
 - 3: $m_e^{(R)} \leftarrow determinar_metrica_esforco_redutores(trecho_atual)$
 - 4: $i_d^{(A)} \leftarrow i_d^{(A)} + m_e^{(A)}$
 - 5: $i_d^{(R)} \leftarrow i_d^{(R)} + m_e^{(R)}$
 - 6: $i_d \leftarrow \frac{i_d^{(A)} + i_d^{(R)}}{2}$
-

Fonte: Aatoria própria (2022).

A última etapa é responsável por correlacionar essas informações com os dados da Central 156. O objetivo era compreender se havia alguma correlação entre a presença/ausência de reclamações e os valores das métricas de esforço.

4.2.2 Divisão da rota em trechos

O trabalho utilizou dados da linha 022 do sistema de transporte público de Curitiba, a linha Inter II. A escolha baseou-se no fato de que esta linha de ônibus percorre um grande perímetro da cidade, passando por vários bairros diferentes e, portanto, poderia estar sujeita a diversas condições, tais como trechos com grande variação de altimetria ou trechos com grande quantidade de redutores, enquanto outros trechos teriam pouca variação de altimetria ou poucos redutores.

A rota foi dividida em trechos de 1 km cada. Como a rota tem, no total, aproximadamente 38,2 km, o trecho final foi deixado com tamanho menor que os demais.

4.2.3 Cálculo do grau de inclinação

Quanto mais íngreme a subida ou descida, mais força (da física) é demandada dos sub-sistemas do veículo para contrabalancear a força gravitacional que atua contra o movimento do mesmo, ou seja, com maior demanda de força desses sub-sistemas, maior é o desgaste dos mesmos.

A Figura 10 mostra visualmente o método, como feito em (BARCZYSZYN, 2018):

$$g(p_i, p_j) = \frac{100d_v(p_i, p_j)}{d_h(p_i, p_j)} \quad (3)$$

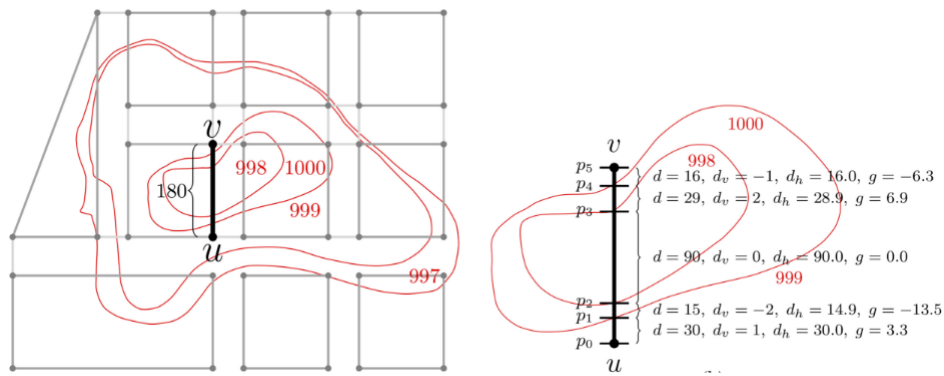


Figura 10 – Cálculo do grau de inclinação para as interseções consecutivas da altimetria com a linha de ônibus. Retirado de (BARCZYSZYN, 2018).

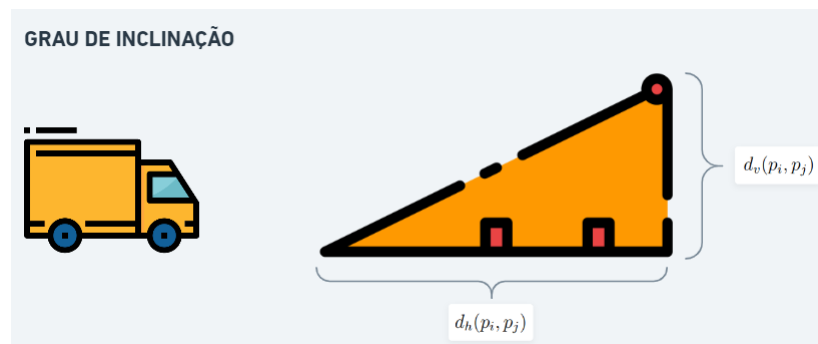


Figura 11 – Componentes do grau de inclinação. Autoria própria.

Onde $d_h(p_i, p_j) = \sqrt{d(p_i, p_j)^2 - d_v(p_i, p_j)^2}$ é o deslocamento horizontal entre os pontos p_i e p_j , $d_v(p_i, p_j)$ é o deslocamento vertical entre os pontos e $d(p_i, p_j)$ é a distância entre eles. O número é expresso em porcentagem, tal como em (BARCZYSZYN, 2018). Note que o grau de inclinação pode ser negativo, no caso de uma descida. A Figura 11 mostra os componentes de distância do cálculo do grau de inclinação.

4.2.4 Grau de inclinação máximo do trecho

Quando um trecho possuir diferentes níveis de altimetria, o grau de inclinação máximo foi calculado utilizando o cálculo de esforço máximo de (BARCZYSZYN, 2018):

$$E_{max} = \max_{p_i \in (u,v), p_i \neq v} \{|g(p_i, p_{i+1})|\} \quad (4)$$

onde é E_{max} é o maior grau de inclinação dos pontos de interseção da rota com as curvas de nível de altimetria, p_i e p_{i+1} são pontos consecutivos de interseção entre o trecho e as curvas de nível.

O grau de inclinação máximo do trecho servirá posteriormente como a métrica de esforço necessário para o veículo percorrer o trecho.

4.2.5 Contagem de redutores para um trecho

A métrica de esforço para a segunda variável foi determinada como a contagem dos redutores ao longo do trecho de um quilômetro. Então, se dentro do mesmo quilômetro há 5 redutores no total, não importando o tipo, o valor da métrica de esforço é 5.

4.2.6 Determinação do trecho em que um veículo se encontra

A correlação do dado geolocalizado do veículo (tabela veículos) com os trechos da rota do mesmo (tabela linha_de_onibus) foi feita de acordo com o Algoritmo 2. Na prática, a etapa foi necessária para não utilizar dados como a saída de garagem, por exemplo.

Algoritmo 2 – Algoritmo de determinação de trecho de veículo

```

1: trechos_possiveis ← []
2: para todos trecho em linha faça
3:   dist_veiculo_trecho ← calcular_distancia_veiculo_trecho(veiculo, trecho)
4:   se dist_veiculo_trecho < 25m então
5:     trechos_possiveis.adicionar_elemento(trecho)
6:   finaliza se
7: finaliza para
8: trecho_atual ← min(trechos_possiveis, trechos_possiveis.num_seq)

```

Fonte: Autoria própria (2022).

4.2.7 Índices de desgaste intermediários

Os índices de desgaste intermediários (variável altimetria e variável redutores) são calculados acumulando os esforços nos trechos, como indicado abaixo:

$$i_d^{(A)}(n) = i_d^{(A)}(n - 1) + m_e^{(A)}(n) \quad (5)$$

$$i_d^{(A)}(0) = 0 \quad (6)$$

Da Equação 5, $i_d^{(A)}(n)$ é o índice de desgaste relativo apenas à altimetria acumulado até a n-ésima travessagem de um trecho cujo esforço é $m_e^{(A)}(n)$. A primeira travessagem, por exemplo, pode ter sido no trecho 0, a segunda no trecho 1 e a terceira no trecho 0 novamente, caso esta seja a rota do respectivo ônibus. Os desgastes são somados somente ao término da travessagem de um trecho.

Para a variável redutores o índice intermediário é:

$$i_d^{(R)}(n) = i_d^{(R)}(n-1) + m_e^{(R)}(n) \quad (7)$$

$$i_d^{(R)}(0) = 0 \quad (8)$$

A Equação 7 é análoga à Equação 5. Sendo assim, $i_d^{(R)}(n)$ é o desgaste devido aos redutores de um dado veículo na n-ésima travessagem e assim por diante.

4.2.8 Índice de desgaste

Os índices de desgaste intermediários são então agregados em um índice de desgaste final, utilizando a média aritmética (para fins de simplificação), como indicado abaixo:

$$i_d(n) = \frac{i_d^{(A)}(n) + i_d^{(R)}(n)}{2} \quad (9)$$

Onde $i_d(n)$ é o desgaste total de um dado veículo na n-ésima travessagem.

Em poucas palavras, o que esse índice expressa, de forma numérica, é que quanto mais subidas e descidas (e quanto mais íngremes, pior!) e quanto mais redutores de velocidade houver na rota que o veículo está percorrendo, mais desgaste esse veículo terá.

4.2.9 Correspondência com dados da Central 156

Como a correspondência das rotas percorridas pelos ônibus com os endereços presentes nos registros de reclamação do 156 não é direta através da junção por uma coluna geométrica, foi utilizada comparação textual, como indicado no Algoritmo 3. A primeira etapa encontra os nomes das ruas da linha e a segunda encontra as correspondências. Como ao trabalhar com interseções entre colunas geométricas, existe a possibilidade de capturar ruas transversais a rua que efetivamente pertence à rota do ônibus, foi feito um filtro baseado no tamanho dessa interseção, em que foram removidas as ruas cuja interseção não tinha comprimento maior que 50 metros (valor verificado visualmente no QGIS). A Tabela 1 mostra as abreviações de nomes de ruas e para qual valor as mesmas foram mapeadas.

4.3 Caso de uso

Para o caso de uso, inicialmente a linha 022 (Inter 2) foi dividida em trechos de 1 km cada. Para os mesmos trechos foram identificadas as curvas de nível referentes às altimetrias na linha 022 (Figura 12). O grau de inclinação máximo foi então calculado para cada trecho (Figura 13).

Algoritmo 3 – Algoritmo de correspondência linha-156

```

1: ruas_linha022 ← []
2: para todos rua em tabela de eixo das ruas faça
3:   buffer ← gerar_buffer(linha022.geom)
4:   tamanho_intersecao ← calcular_tamanho_intersecao(rua.geom, linha022.geom)
5:   se tamanho_intersecao > 50 então
6:     ruas_linha022.adicionar_elemento(rua)
7:   finaliza se
8: finaliza para
9: correspondencias ← []
10: para todos reclamacao em tabela do 156 faça
11:   para todos rua em ruas_linha022 faça
12:     nome_via_linha022 ← preprocessar_texto(rua.nmvia)
13:     nome_via_central ← preprocessar_texto(reclamacao.logradouro_ass)
14:     se nome_via_linha022 = nome_via_central então
15:       correspondencias.adicionar_elemento(nome_via_linha022)
16:     finaliza se
17:   finaliza para
18: finaliza para

```

Fonte: Autoria própria (2022).

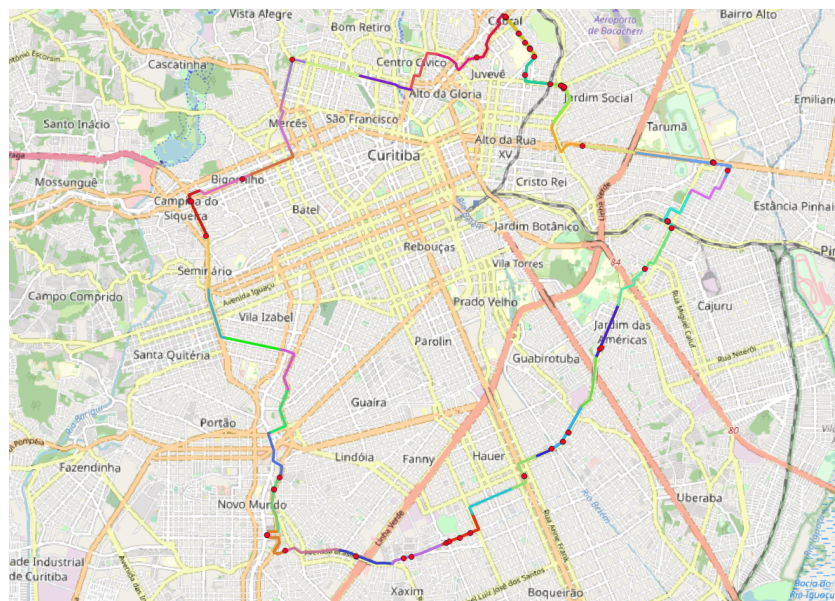


Figura 14 – Redutores de velocidade e trechos da linha 022. Autoria própria.

Os redutores de velocidade foram então selecionados, para a rota da linha 022, como ilustra a Figura 14. Note que há trechos que possuem mais redutores que outros, como o trecho da figura na parte de cima, à direita.

Sete veículos foram então selecionados (BL302, BL315, ML308, EL309, DL040, DL302 e DR803) e o índice de altimetria foi calculado (Figura 15). Note que a propriedade de monotonicidade crescente está presente, já que estamos sempre somando quantidades positivas ao índice acumulado, pois nenhum dos esforços necessários para percorrer os trechos é negativo. A Figura 16 ilustra o índice de desgaste dados os redutores de velocidade da linha 022 para os mesmos 7 veículos.

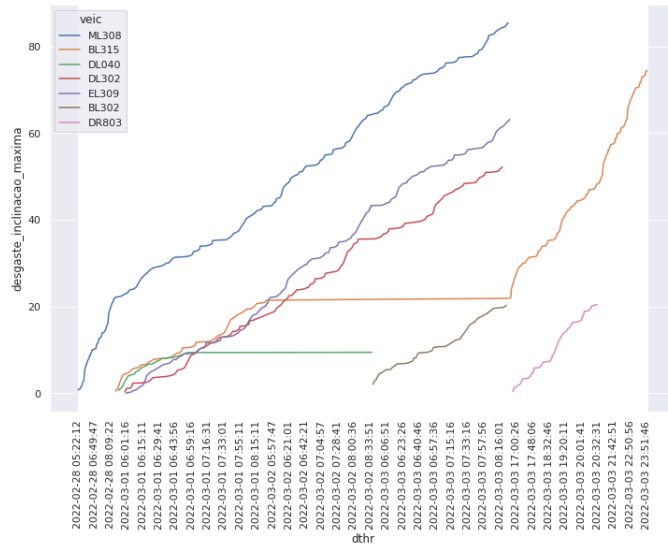


Figura 15 – Índice de desgaste dado a altimetria ao longo do tempo para vários veículos. Autoria própria

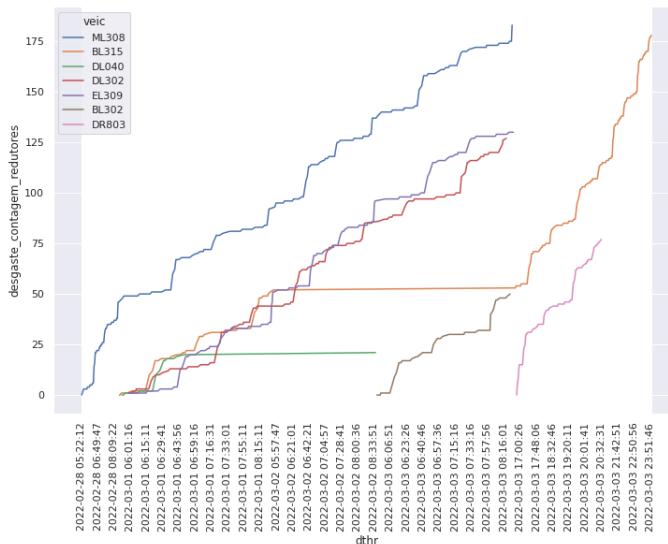


Figura 16 – Índice de desgaste dados os redutores ao longo do tempo para vários veículos. Autoria própria

Finalmente, o índice de desgaste, agregando as duas variáveis principais – altimetria e redutores de velocidade – pode ser verificado na Figura 17. Note novamente a propriedade de monotonicidade crescente, esperada, visto que o desgaste é apenas somado. Outra é o desligamento dos veículos e o tempo fora de operação representados por longos patamares sem incremento de valor no eixo y. E a terceira, por fim, é o fato de os veículos provavelmente estarem atuando em regiões diferentes, visto que as inclinações dos índices de cada veículo variam de um para o outro.

As ruas utilizadas pela linha 022 foram então correlacionadas com dados da Central 156 (Figura 18). A figura apresenta somente as ruas que tiveram alguma reclamação.

Os detalhes da implementação do sistema, assim como alguns dos códigos produzidos, podem ser encontrados no Apêndice B.

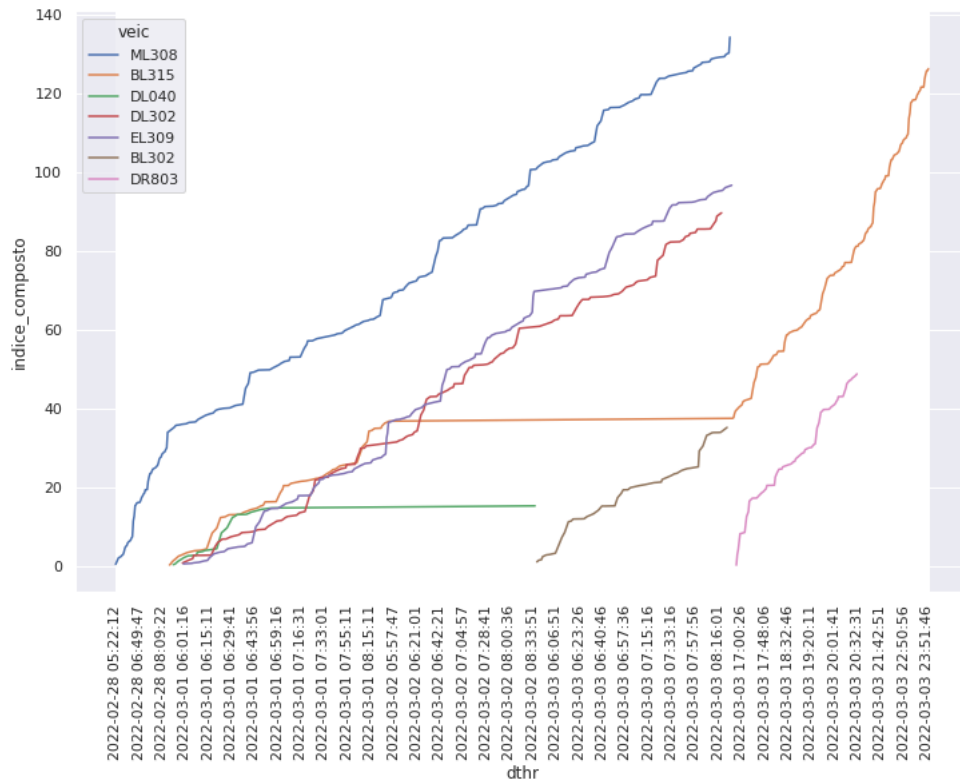


Figura 17 – Índice de desgaste. Autoria própria

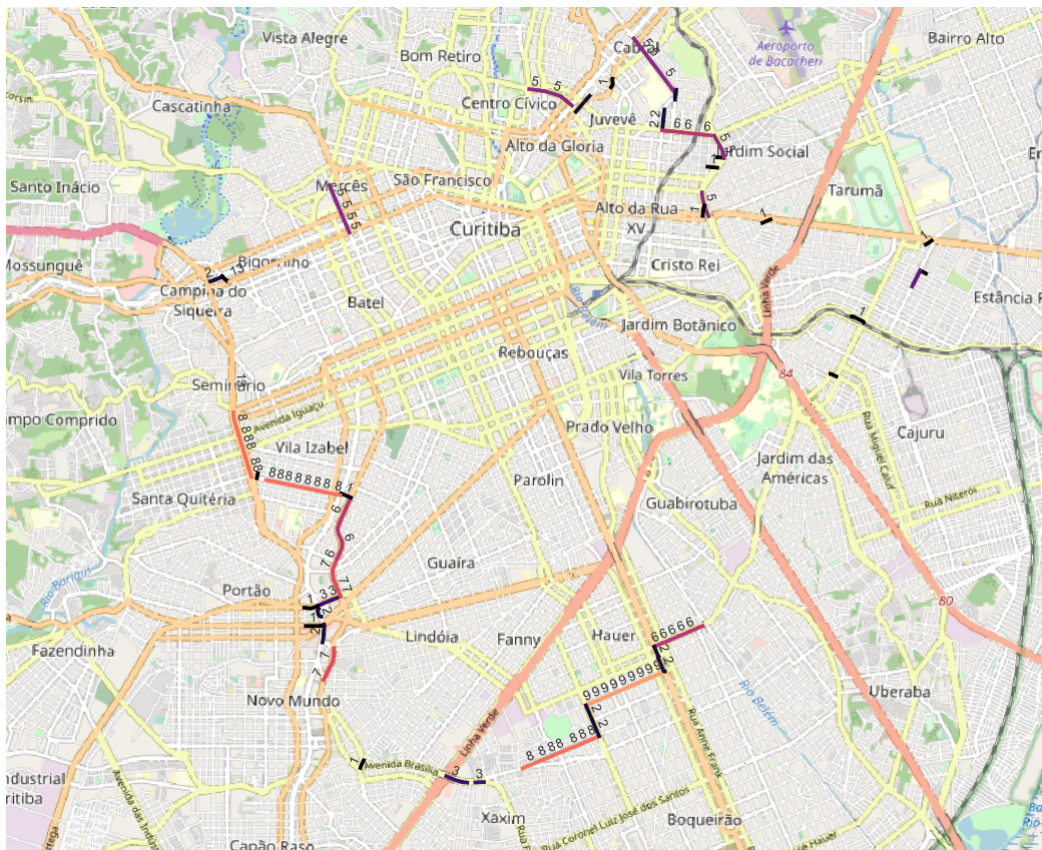


Figura 18 – Ruas da linha 022 alvo de reclamações no 156. Autoria própria.

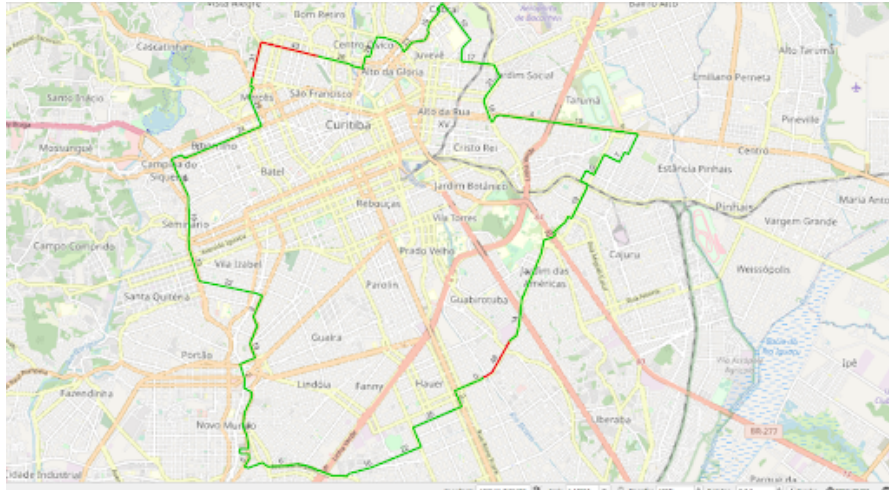


Figura 19 – Trechos com grau de inclinação máximo acima de 30. Autoria própria

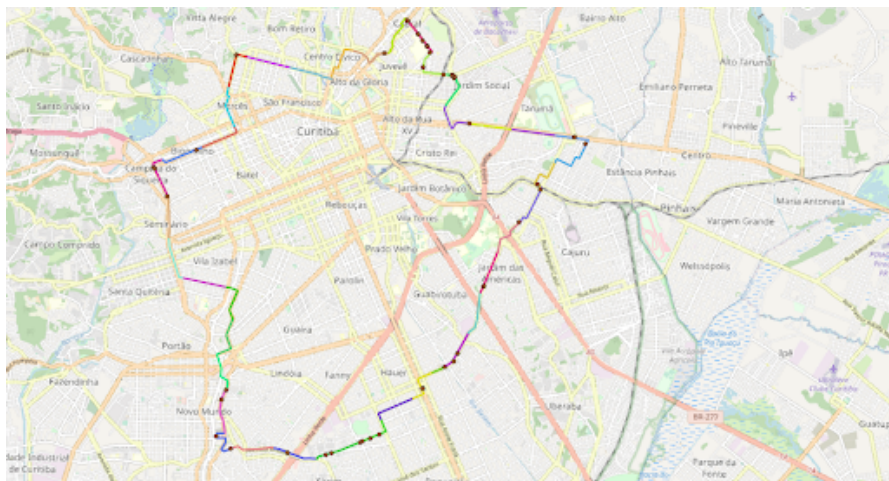


Figura 20 – Redutores da linha 022.

4.4 Discussões

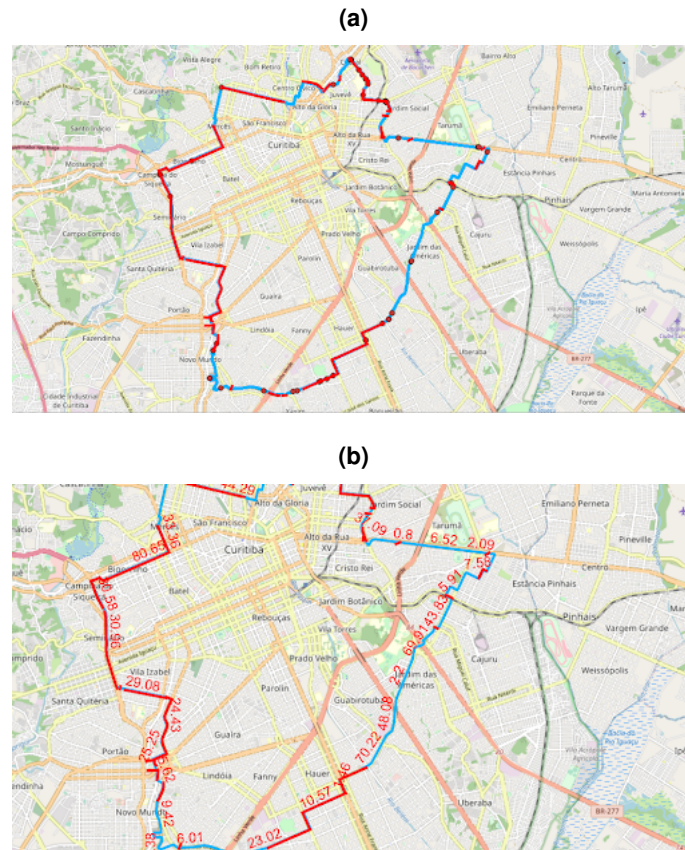
Considerando a linha Inter II, há vários trechos que estão acima do limiar indicados em vermelho (neste trabalho considerado como 30), como ilustra a Figura 19. Isso quer dizer que para cada 1 metro de deslocamento horizontal há uma variação de 30 cm associada no trecho. Há três regiões (3 km de linha) cujo grau de inclinação e, por consequência, a variação de altimetria, é bem elevado, e serão categorizados neste trabalho como "piores trechos da linha". Isso implica, em termos práticos, em mais desgaste para o veículo transitar naqueles trechos que em outros.

A Figura 20 mostra uma informação similar, mas para redutores de velocidade. A região do Cabral, mais acima e à direita na figura, apresenta uma concentração muito maior de redutores do que outras regiões, como por exemplo o Centro Cívico.

Em uma última etapa foi analisada a sobreposição das informações do 156 com as informações de altimetria (Figura 21). É possível ver que, apesar de não ser uma regra possível de ser generalizada sem exceções, as regiões com mais redutores, tais como o Cabral e o Xaxim,

também possuem mais registros de reclamações. Ocorreu algo equivalente para a altimetria (Figura 21b): as regiões com graus de inclinação acima de 20 ou 30 também possuem reclamações no 156.

Figura 21 – Análise conjunta dos dados do 156, da altimetria e dos redutores.



Fonte: Autoria própria (2022).

Sobre os desafios, podemos citar o problema da disponibilidade de dados de horímetro (dispositivo que registra dados de tempo de operação), localização e a atualização dos dados.

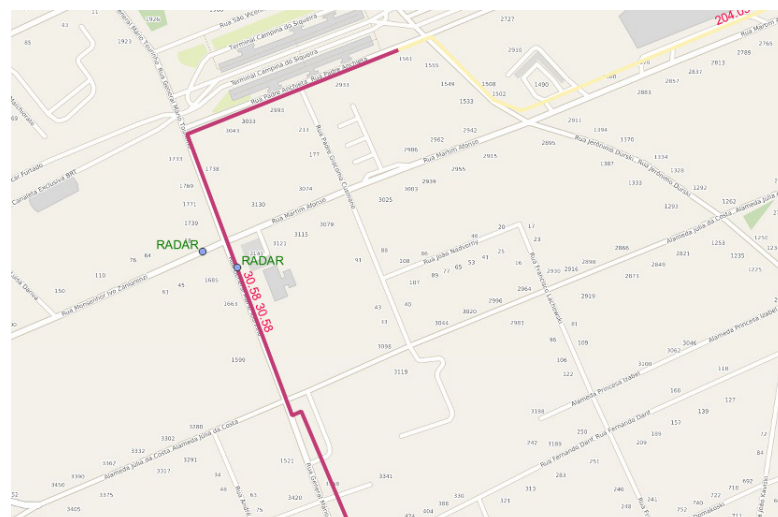


Figura 22 – Redutor provável de pertencer a outra rua que não a do trecho analisado. Autoria própria



Figura 23 – Radar mostrado no mapa do QGIS que não existe no Google Street View. Retirado do Google Street View.

Os dados de altimetria, por exemplo, tem a data de coleta mais recente entre os anos de 2000 e 2001. Um outro exemplo de dado que passou a ser inválido com o tempo é mostrado nas Figuras 22 e 23. As imagens da rua contida especificamente nesta imagem foram capturadas em setembro de 2021, de acordo com o próprio Google. No entanto, não pode ser visto nenhum radar de velocidade disposto nessa rua.

Um outro problema que pode ser citado é orientação do radar: se o campo de visão da câmera atrelada ao redutor estiver perpendicular à via, certamente o redutor não foi posicionado naquele local para operar naquela via. Apesar de existirem aplicativos que mostrem a posição do radar, assim como o limite de velocidade que o mesmo verifica (tais como o Radarbot), não foram coletados dados destes para este trabalho.

Outra questão que deve ser avaliada é o método de agregação dos índices de desgaste intermediários. Apesar da média ser um ponto de partida, não é uma maneira robusta de se calcular o índice composto, visto que as variáveis podem ter influências maiores ou menores sobre o desgaste. Um outro método sugerido em (CHEN, 2021) utiliza três propriedades desejáveis – monotonicidade, tendência e robustez – em um índice e cria um problema de otimização baseado nessas três propriedades com o qual calcula os pesos utilizados para cada variável dentro do índice final.

Por fim, o método de correspondência entre os dados da linha de ônibus e os dados da Central 156 está limitado a nomes de ruas muito parecidos. Casos, por exemplo, de erros de digitação, onde alguém pode escrever "Rua Alferes Polu" em vez de "Rua Alferes Poli" não geraram uma correspondência válida, apesar dos textos diferirem apenas por 1 letra. Para este fim, processamento de linguagem natural poderia ter sido utilizado para aumentar a acurácia dos resultados. No entanto, este trabalho limitou-se ao método mais simples. Uma coisa interessante de se observar é a seguinte: na Figura 18, é possível perceber que as regiões mais afastadas do centro também são as de maior concentração de reclamações.

**Figura 24 – Radar que foi retirado entre maio de 2021 e janeiro de 2022.
(a) Google Street View mostra radar em maio de 2021**



(b) Google Street View não mostra mais o radar em janeiro de 2022.



Fonte: Google Street View.

A partir das análises produzidas, é possível dar algumas sugestões, especificamente para a prefeitura de Curitiba. A primeira delas se refere à quantidade de redutores de velocidade. É possível verificar na Tabela 2 que há 59 redutores na linha 022, no total, sendo 25 deles radares. Uma outra informação útil é que a linha 022 tem aproximadamente 38,3 km. Em outras palavras, isso quer dizer que há mais de 1 redutor por quilômetro e quase um radar por quilômetro. Ao colocar tantos deles, os motoristas são forçados a variar a velocidade do ônibus em grande quantidade várias vezes. A variação é ainda maior nos casos das lombadas e travessias elevadas, que devem ser ultrapassadas em baixa velocidade. Ao aplicar essas grandes variações de velocidade, desgastam-se diversas partes do veículo, tais como freio e motor, para começar. Em redutores como lombadas, ainda há o envolvimento das rodas, pneus, suspensão e toda a estrutura do veículo que sofre o impacto da colisão com o redutor. Sendo assim, a primeira sugestão é tentar limitar um pouco mais a disposição dos redutores na linha ou até alterar a rota para não incluir tantos.

A segunda sugestão é utilizar o próprio algoritmo de cálculo das métricas de esforço da linha para a construção das mesmas, de forma a encontrar uma rota que minimize o desgaste dos veículos, mas ainda assim atenda as regiões necessárias, fazendo alterações precisas para evitar trechos problemáticos identificados usando as métricas propostas.

Tabela 2 – Número de redutores por tipo

Tipo de redutor	Quantidade
Radar	25
Lombada	18
Terminal	6
Ponto	5
Ônibus	3
Linda de Trem	2

A terceira sugestão é utilizar o índice de desgaste proposto em conjunto com os dados da rota do veículo e do itinerário para análise dos tempos entre falhas dos veículos. Ao fazer isso, teríamos uma relação entre as rotas e duas métricas de esforço.

5 CONCLUSÃO

O crescimento populacional, principalmente o urbano, está levando ao aumento consequente dos problemas com infra-estrutura. Um dos setores que sofre com isso é o setor de transportes. Então, é crescente a necessidade de uma metodologia de análise de operação dos veículos, por exemplo de transporte público, e a subsequente tomada de decisão no planejamento de manutenção. A criação de um índice de desgaste para cada veículo pode ser utilizada nesse planejamento de manutenção para entender como a rota que um veículo percorre (ou percorrerá) influencia no desgaste dele, ou até quanto tempo um veículo pode operar antes de começar a apresentar falhas e necessitar de manutenção. Abordando o problema de uma perspectiva GIS (usando dados de altimetria e redutores de velocidade) foi apresentando um índice de desgaste.

A abordagem, ao trazer dados da rota, também tem a vantagem de não necessitar que o veículo envie sinais o tempo todo. Isso por si só já elimina uma série de possíveis fontes de problema, tais como falha no dispositivo de comunicação, falta de rede para enviar os dados, envio de dados errados por falha no sensor, entre outros. Se, por exemplo, o dispositivo de telemática apresentar defeitos que impossibilitem o envio de informações para um servidor em *back-office*, o desgaste ainda pode ser calculado se a rota e o itinerário do veículo forem conhecidos e se os dados GIS da rota já estiverem disponíveis nesse servidor. Essa abordagem tem, então, uma vantagem sobre a abordagem que se utiliza de dados de sensores, que precisam ser lidos e enviados o tempo todo.

O objetivo geral deste trabalho foi atuar nas etapas de processamento de dados e diagnóstico da manutenção preditiva (através de dados geográficos), para propor um subsistema com foco em índice de desgaste de veículos que possa ser utilizado em tarefas de diagnóstico. Para isso, também foram propostas métricas de esforço para as variáveis altimetria (inspirada em (BARCZYSZYN, 2018)) e redutores de velocidade. A partir das variáveis, os índices de desgaste foram criados e, por sua vez, estes foram agregados para construir o índice de desgaste final. Também foi apresentada a correspondência dos dados da linha de ônibus escolhida com os registros de reclamações da Central 156 e foi feita a determinação dos piores trechos.

Foi possível notar, de acordo com os resultados encontrados, que o esforço necessário para os veículos percorrerem cada trecho é diferente entre eles. Também foi possível perceber que alguns veículos sofrem mais desgaste que outros a partir das curvas plotadas nos gráficos. Juntando esses dois fatos, tem-se que alguns veículos podem estar tendo suas vidas úteis subaproveitadas, considerando que existem trechos cuja passagem exige mais esforço e que alguns veículos percorrem apenas trechos específicos. O método proposto, no entanto, tem suas limitações. Ao utilizar apenas distâncias e nenhuma outra informação para determinar se um redutor está ou não contido em uma via, é possível gerar erros nessa determinação e, posteriormente, no índice de desgaste. Outro problema é o da validade dos dados e sua atualização.

Outras questões, tais como o método de agregação dos índices individuais, podem ser consideradas em trabalhos futuros. Pode-se citar o método de (CHEN, 2021) para determinação dos pesos de cada variável no índice, por exemplo. Também podem ser utilizadas técnicas mais robustas de linguagem natural para processamento de textos e *entity matching* para reconhecimento das ruas na Central 156 e outras fontes similares de informação. Além disso, ainda como trabalho futuro, pode-se estender a metodologia para inclusão de mais variáveis, tais como temperatura, pluviometria, cruzamentos de vias preferenciais, semáforos, umidade e pressão do ar, entre outras.

REFERÊNCIAS

- BARCLAY, M.; DIXON-WOODS, M.; LYRATZOPOULOS, G. *The problem with composite indicators*. **BMJ Quality Safety**, v. 28, 08 2018.
- BARCZYSZYN, G.; CAMENAR, L.; NASCIMENTO, D.; KOZIEVITCH, N.; SILVA, R.; ALMEIDA, L.; SANTI, J. de; MINETTO, R. A collaborative system for suitable wheelchair route planning. **ACM Transactions on Accessible Computing**, v. 11, p. 1–26, 08 2018.
- BOUSDEKIS, A.; MENTZAS, G. Condition-based predictive maintenance in the frame of industry 4.0. In: LÖDDING, H.; RIEDEL, R.; THOBEN, K.-D.; CIEMINSKI, G. von; KIRITSIS, D. (Ed.). **Advances in Production Management Systems. The Path to Intelligent, Collaborative and Sustainable Manufacturing**. Cham: Springer International Publishing, 2017. p. 399–406. ISBN 978-3-319-66923-6.
- BOX, G. E. P. Science and statistics. **Journal of the American Statistical Association**, Taylor Francis, v. 71, n. 356, p. 791–799, 1976. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/01621459.1976.10480949>.
- CHEN, C.; LIU, Y.; SUN, X.; CAIRANO-GILFEDDER, C. D.; TITMUS, S. Automobile Maintenance Prediction Using Deep Learning with GIS Data. **Procedia CIRP**, v. 81, p. 447–452, 2019. ISSN 2212-8271. 52nd CIRP Conference on Manufacturing Systems (CMS), Ljubljana, Slovenia, June 12-14, 2019. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2212827119303828>.
- CHEN, C.; LIU, Y.; SUN, X.; CAIRANO-GILFEDDER, C. D.; TITMUS, S. An integrated deep learning-based approach for automobile maintenance prediction with gis data. **Reliability Engineering System Safety**, v. 216, p. 107919, 2021. ISSN 0951-8320. Disponível em: <https://www.sciencedirect.com/science/article/pii/S095183202100435X>.
- CHEN, Z.; XIA, T.; ZHOU, D.; PAN, E. A health index construction framework for prognostics based on feature fusion and constrained optimization. **IEEE Transactions on Instrumentation and Measurement**, v. 70, p. 1–15, 2021.
- CHOURABI, H.; NAM, T.; WALKER, S.; GIL-GARCIA, J. R.; MELLOULI, S.; NAHON, K.; PARDO, T. A.; SCHOLL, H. J. Understanding smart cities: An integrative framework. In: **2012 45th Hawaii International Conference on System Sciences**. 2012. p. 2289–2297.
- DADOS.GOV.BR. **O que são dados abertos?** 2022. <https://dados.gov.br/pagina/dados-abertos>. Acessado em: 2022-05-16.
- ELMASRI, R.; NAVATHE, S. B. **Fundamentals os Database Systems**. 2017. ISBN 978-0-13-397077-7.
- HAMADA, E.; GONÇALVES, R. R. d. V. **Introdução ao Geoprocessamento: princípios básicos e aplicação**. 2007.
- HÖJER, M.; WANGEL, J. Smart sustainable cities: Definition and challenges. **ICT Innovations for Sustainability, Advances in Intelligent Systems and Computing**, p. 333–349, 01 2015.
- MOBLEY, R. **An Introduction to Predictive Maintenance**. Elsevier Science, 2002. (Plant Engineering). ISBN 9780080478692. Disponível em: <https://books.google.com.br/books?id=SjqXzxpAzSQC>.

MODERNO, C. **Quais são os impactos ambientais de um smartphone?** 2015. <https://www.consumidormoderno.com.br/2015/11/01/quais-sao-os-impactos-ambientais-de-um-smartphone/#:~:text=Dentre%20os%20diversos%20recursos%20extra%C3%ADdos,kg%20de%20emiss%C3%B5es%20de%20carbono.> Acessado em: 2022-05-14.

ONU. **68% of the world population projected to live in urban areas by 2050, says UN.** 2018. <https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html#:~:text=News-,68%25%20of%20the%20world%20population%20projected%20to%20live%20in,areas%20by%202050%2C%20says%20UN&text=Today%2C%2055%25%20of%20the%20world's,increase%20to%2068%25%20by%202050.> Acessado em: 22/04/2022.

ONU. **Urban and rural population.** 2018. [https://population.un.org/wup/country-profiles/.](https://population.un.org/wup/country-profiles/) Acessado em 09/04/2022.

PRYTZ, R. **Machine learning methods for vehicle predictive maintenance using off-board and on-board data.** 09 2014. Tese (Doutorado), 09 2014.

RAMAKRISHNAN, R.; GEHRKE, J. **Database Management Systems.** : McGraw-Hill Companies, 2003. ISBN 0-07-123151-X.

RAMAKRISHNAN, R.; GEHRKE, J. **Sistemas de Gerenciamento de Bancos de Dados.** : AMGH, 2007. ISBN 8577260275.

RINOS, M. M.; KALEEL, m. i. **Geographic Information System: (1st Edition).** : Department of Geography, Southeastern University of Sri Lanka, 2012. ISBN 978-955-627-032-7.

ROSA, M. **iPhone 13 começará a ser fabricado na Índia em fevereiro de 2022.** 2022. <https://www.tudocelular.com/mercado/noticias/n183939/iphone-13-comecara-fabricado-india-fevereiro-2022.html#:~:text=No%20entanto%2C%20o%20procedimento%20n%C3%A3o,utilizada%20para%20os%20novos%20aparelhos.> Acessado em: 2022-05-14.

SIMETTE, G.; PARCIANELLO, Y.; KOZIEVITCH, N. P.; FONSECA, K. V. O. Análise da situação dos redutores de velocidade de curitiba. In: **Anais da XIV Escola Regional de Banco de Dados.** Porto Alegre, RS, Brasil: SBC, 2018. ISSN 2595-413X. Disponível em: <https://sol.sbc.org.br/index.php/erbd/article/view/2836>.

ZHANG, J.; WANG, P.; YAN, R.; GAO, R. X. Long short-term memory for machine remaining life prediction. **Journal of Manufacturing Systems**, v. 48, p. 78–86, 2018. ISSN 0278-6125. Special Issue on Smart Manufacturing. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0278612518300803>.

APÊNDICE A – Tabela completa com as correspondências encontradas entre a linha 022 e os registros de reclamações da Central 156.

Tabela 3 – Correspondência encontradas entre a Central 156 e a linha 022, com o número de reclamações.

Rua	Número de reclamações
martim afonso	13
general mario tourinho	13
doutor roberto barrozo	13
waldemar kost	9
presidente arthur da silva bernardes	8
evaristo da veiga	8
professor ulisses vieira	8
eduardo carlos pereira	7
guilherme pugsley	6
augusto stresser	6
doutor julio cesar ribeiro de souza	6
dos funcionarios	5
nossa senhora da luz	5
deputado mario de barros	5
fernando simas	5
hayton da silva pereira	4
augusto de mari	3
francisco derosso	3
padre germano mayer	2
camoes	2
paulo setubal	2
republica argentina	2
marechal floriano peixoto	2
nicolau maeder	2
padre anchieta	2
munhoz da rocha	1
itatiaia	1
wenceslau glaser	1
victorio vizinoni	1
jeronimo durski	1
rio grande do sul	1
conselheiro carrao	1
das andorinhas	1
presidente beaurepaire rohan	1
luiz parigot de souza	1
simao bolivar	1
professora antonia reginato vianna	1
raul joaquim quadros gomes	1
presidente affonso camargo	1
miguel caluf	1

APÊNDICE B – Implementação do sistema

B.1 Implementação do sistema

B.1.1 Trechos da linha 022

Para realizar a divisão da linha 022 em trechos de 1 km cada, foi implementada a consulta em SQL apresentada na Listagem 1 utilizando a extensão PostGIS para PostgreSQL.

Listagem 1 – Divisão da linha em trechos de 1 km.

```

1 SELECT
2   gid ,
3   i AS num_seq,
4   ST_Length(
5     ST_LineSubstring(
6       ST_LineMerge(geom) ,
7       startfrac ,
8       LEAST(endfrac , 1)
9     )
10  ) tamanho_subcaminho ,
11  ST_LineSubstring(
12    ST_LineMerge(geom) ,
13    startfrac ,
14    LEAST(endfrac , 1)
15  ) AS geom
16 FROM (
17   SELECT gid , geom, ST_Length(geom) len , 1000 sublen
18   FROM transporte.linha_de_onibus ldo
19   WHERE cd_linha = '022' AND gid = 3042
20 ) AS d CROSS JOIN LATERAL (
21   SELECT i , (sublen * i) / len AS startfrac ,
22          (sublen * (i+1)) / len AS endfrac
23   FROM generate_series(0, floor( len / sublen )::INTEGER ) AS t(i)
24   WHERE (sublen * i) / len <> 1.0
25 ) AS d2

```

Fonte: Autoria própria (2022).

B.1.2 Graus de inclinação da linha 022

A consulta para realizar os cálculos dos graus de inclinação para cada par consecutivo de altimetrias, tal como explicado anteriormente e mostrado na Figura 11, está disposta na Listagem 2.

Listagem 2 – Cálculo dos graus de inclinação para as altimetrias.

```

1 SELECT
2     linha_alt_gid ,
3     trecho_num_seq ,
4     CASE
5         WHEN grau_de_inclinacao IS NULL THEN 0
6         ELSE grau_de_inclinacao
7     END grau_de_inclinacao ,
8     km_geom,
9     alt_geom,
10    km_alt_intersecao
11 FROM (
12     SELECT
13         linha_alt_gid ,
14         trecho_num_seq ,
15         ROUND(
16             (
17                 100
18                 * abs( elevacao - LAG(elevacao) OVER (
19                     PARTITION BY trecho_num_seq ORDER BY linha_alt_gid
20                 ) )
21             / ST_Distance(
22                 km_alt_intersecao ,
23                 LAG(km_alt_intersecao) OVER (
24                     PARTITION BY trecho_num_seq ORDER BY linha_alt_gid
25                 )
26             )
27         )::NUMERIC,
28         2) as grau_de_inclinacao ,
29         km_geom,
30         alt_geom ,
31         km_alt_intersecao
32     FROM {{ ref('matheus_linha_alt_intersecoes' ) }}
33 ) a

```

Fonte: Autoria própria (2022).

B.1.3 Graus de inclinação máximo para cada trecho da linha 022

A Listagem 3 traz a consulta para o cálculo do grau de inclinação máximo de cada trecho de 1 km calculado anteriormente.

B.1.4 Redutores em cada trecho da linha 022

A determinação dos redutores de velocidade presentes em cada trecho da linha 022 foi feita em duas partes: primeiro foram filtrados os redutores na linha completa e, após isso, foram determinados quais e quantos desses redutores estavam em quais trechos.

Listagem 3 – Cálculo do grau de inclinação máximo para cada trecho.

```

1  {% set limiar = 30 %}
2
3  SELECT
4      a.*,
5      CASE
6          WHEN inclinacao_maxima > {{ limiar }} THEN TRUE
7          ELSE FALSE
8      END acima_limite
9  FROM (
10     SELECT
11         trecho_num_seq,
12         MAX(ABS(grau_de_inclinacao)) as inclinacao_maxima,
13         km_geom
14     FROM {{ ref('matheus_pct_grade_info') }}
15     GROUP BY trecho_num_seq, km_geom
16 ) a

```

Fonte: Autoria própria (2022).

A Listagem 4 mostra a consulta implementada para executar a primeira parte. Foram filtrados então todos os redutores dentro de uma região de até 25 metros de distância da linha 022.

Listagem 4 – Filtragem dos redutores da linha 022.

```

1  SELECT
2      ldo.gid AS linha_gid ,
3      ldo.geom AS linha_geom ,
4      lr.gid AS redutor_gid ,
5      lr.tipo AS redutor_tipo ,
6      ST_Transform(lr.geom, 29192) AS redutor_geom
7  FROM transporte.linha_de_onibus ldo , arruamento.lombada_radar lr
8  WHERE cd_linha = '022' AND ST_DWithin(
9      ldo.geom,
10     ST_Transform(lr.geom, 29192),
11     25
12 )

```

Fonte: Autoria própria (2022).

Com os redutores da linha 022 em mãos, passou-se para a etapa de determinação dos trechos correspondentes a cada um. A Listagem 5 mostra a consulta implementada para resolver esse problema.

Listagem 5 – Filtragem dos redutores da linha 022.

```

1  SELECT t.num_seq, t.geom AS trecho_geom, r.redutor_tipo, r.redutor_geom
2  FROM {{ ref('matheus_linha022_trechoskm') }} t
3     LEFT JOIN {{ ref('matheus_linha022_redutores') }} r
4     ON ST_DWithin(t.geom, r.redutor_geom, 25)

```

Fonte: Autoria própria (2022).

B.1.5 Trecho em que os veículos se encontram

Finalmente, a determinação dos trechos em que os veículos se encontram dentro de uma linha foi feita de acordo com a consulta disposta na Listagem 6. Foram filtrados todos os trechos que estavam a 25 metros ou menos da localização dos veículos.

Listagem 6 – Filtragem dos redutores da linha 022.

```

1 SELECT vkm.*, i.inclinacao_maxima, r.contagem_redutores FROM (
2   SELECT DISTINCT
3     v.veic,
4     v.dthr,
5     v.geom,
6     {# Fazemos esse janelamento para evitar que pontos
7     que caíram em dois trechos fiquem sobrepostos.
8     Com isso evitamos que o trecho atual do veículo fique
9     pulando 14 -> 15 -> 14 -> ... #}
10    MIN(km.num_seq) OVER (PARTITION BY v.veic, v.dthr) AS num_seq
11 FROM {{ ref('matheus_veiculos_linha022') }} v,
12      {{ ref('matheus_linha022_trechoskm') }} km
13 WHERE
14     {# Procurando pontos de localização dos veículos dentro da linha #}
15     ST_DWithin(km.geom, v.geom, 25)
16 ) vkm,
17   {{ ref('matheus_trechos_inclinacao') }} i,
18   {{ ref('matheus_redutores_por_trecho') }} r
19 WHERE
20     {# Procurando inclinação de cada trecho #}
21     vkm.num_seq = i.num_seq
22     {# Procurando contagem de redutores de cada trecho #}
23     AND vkm.num_seq = r.num_seq
24 ORDER BY veic, dthr

```

Fonte: Autoria própria (2022).

B.1.6 Índices de desgaste individuais

A seguir, é explicada a metodologia de cálculo dos índices individuais.

A Listagem 7 mostra a implementação em Python¹ do tratamento dos dados.

B.1.7 Índice de desgaste composto

Com ambos os índices de desgaste individuais prontos, finalmente, basta agregar os dois números em um único índice de desgaste composto. Isso foi feito apenas calculando a

¹ <https://www.python.org/>

Listagem 7 – Tratamento da acumulação de esforços nos índices de desgaste.

```

1 handler = HandlerVeiculos()
2 veiculo: Veiculo
3
4 dados = []
5
6 veiculo_sendo_tratado = None
7 ultimo_trecho_analisado = None
8 for veiculo in handler.buscar_dados():
9     if veiculo_sendo_tratado != veiculo.veic:
10        veiculo_sendo_tratado = veiculo.veic
11        ultimo_trecho_analisado = veiculo.num_seq
12        dados.append(veiculo.as_dict())
13        continue
14
15     if ultimo_trecho_analisado != veiculo.num_seq:
16        ultimo_trecho_analisado = veiculo.num_seq
17        dados.append(veiculo.as_dict())
18        continue

```

Fonte: Autoria própria (2022).

média aritmética de ambos, apesar de existirem métodos muito mais sofisticados e robustos, como explicados anteriormente.

B.1.8 Correspondência com dados da Central 156

Para determinar então as ruas da linha 022 que apresentavam correspondência na tabela da Central 156, foram dados 2 passos: determinar o nome das ruas da linha 022 e fazer a correspondência desses nomes com os nomes presentes na tabela do 156.

A Listagem 8 traz a consulta para extrair os nomes das ruas da linha 022. A Listagem 9 traz, por sua vez, a correspondência entre o resultado da consulta anterior com as ruas da tabela da Central 156. Note que é tratado todo o texto para caixa baixa e mapeado algumas abreviações encontradas, tais como "cel.", "pres.", "mal.", entre outras. Um exemplo desse tratamento pode ser "R. GEN. MÁRIO TOURINHO", que gera o resultado "general mario tourinho". Apesar de ser um tratamento válido, não é muito robusto. Técnicas de *entity matching* poderiam ser usadas para esse objetivo.

Algumas das correspondências encontradas e a quantidade de reclamações na Central 156 podem ser vistas na Tabela 4. Para que o quadro não ficasse muito longo, foi omitida a maior parte das correspondências. No entanto, a tabela completa pode ser vista no Apêndice .

Listagem 8 – Determinação dos nomes das ruas que compõem a linha 022.

```

1  {% set buffer_size = 25 %}
2  {% set min_intersection_size = 50 %}
3
4  SELECT
5      b.nmvia ,
6      b.nmvia_tratado ,
7      a.linha_geom ,
8      b.eixo_rua_geom ,
9      ST_Buffer(a.linha_geom , {{ buffer_size }}) AS buffered_linha_geom
10 FROM (
11     SELECT ST_Transform(ST_LineMerge(geom) , 29192) AS linha_geom
12     FROM transporte.linha_de_onibus ldo
13     WHERE gid = 3042
14 ) a, (
15     SELECT
16         nmvia ,
17         public.matheus_mapear_abreviaturas(
18             public.matheus_retirar_tipo_logradouro(
19                 public.matheus_normalizar_texto(nmvia)
20             )
21         ) AS nmvia_tratado ,
22         ST_Transform(ST_LineMerge(geom) , 29192) AS eixo_rua_geom
23     FROM public.eixo_rua er
24 ) b
25 WHERE ST_Length(
26     ST_Intersection(
27         b.eixo_rua_geom ,
28         ST_Buffer(a.linha_geom , {{ buffer_size }})
29     )
30 ) > {{ min_intersection_size }}
31     AND b.nmvia_tratado != ' a' AND b.nmvia IS NOT NULL

```

Fonte: Autoria própria (2022).

Tabela 4 – Correspondência encontradas entre a Central 156 e a linha 022, com o número de reclamações.

Rua	Número de reclamações
martim afonso	13
general mario tourinho	13
doutor roberto barrozo	13
waldemar kost	9

Listagem 9 – Correspondência entre as ruas da Central 156 e da linha 022.

```
1 SELECT  
2     c.logradouro ,  
3     l.nmvia_tratado ,  
4     l.nmvia ,  
5     l.eixo_rua_geom ,  
6     l.linha_geom  
7 FROM {{ ref('matheus_buracos_central156') }} c ,  
8     {{ ref('matheus_nome_ruas_linha022') }} l  
9 WHERE c.logradouro = l.nmvia_tratado
```

Fonte: Autoria própria (2022).