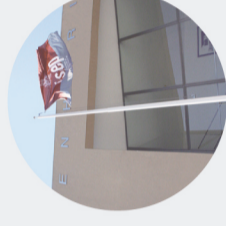


# Identificação de defeitos em veículos ferroviários de mercadorias baseado num sistema wayside e com recurso a um Autoencoder

TOMÁS SIMÕES JORGE

novembro de 2023



# IDENTIFICAÇÃO DE DEFEITOS EM VEÍCULOS FERROVIÁRIOS DE MERCADORIAS BASEADA NUM SISTEMA WAYSIDE E COM RECURSO A UM AUTOENCODER

IDENTIFICAÇÃO DE DEFEITOS EM VEÍCULOS FERROVIÁRIOS DE  
MERCADORIAS BASEADA NUM SISTEMA WAYSIDE E COM RECURSO A UM AUTOENCODER

**TOMÁS SIMÕES JORGE**  
Outubro de 2023



**IDENTIFICAÇÃO DE DEFEITOS EM VEÍCULOS FERROVIÁRIOS DE MERCADORIAS  
BASEADA NUM SISTEMA *WAYSIDE* E COM RECURSO A UM *AUTOENCODER***

TOMÁS SIMÕES JORGE (1180479)

Dissertação submetida para satisfação parcial dos requisitos do grau de

**MESTRE EM ENGENHARIA CIVIL – ESTRUTURAS**

Orientador: Prof. Dr. Diogo Ribeiro (ISEP)

Co-Orientador: Dr. Rúben Silva (FEUP)

Co-Orientadora: Dra. Araliya Mosleh (FEUP)

**OUTUBRO DE 2023**

Eu, Tomás Simões Jorge, estudante nº 1180479, do Mestrado em Engenharia Civil do Instituto Superior de Engenharia do Porto, declaro que não fiz plágio nem auto-plágio, pelo que o trabalho intitulado “Identificação de defeitos em veículos ferroviários de mercadorias baseado num sistema *wayside* e com recurso a um *Autoencoder*” é original e da minha autoria, não tendo sido usado previamente para qualquer outro fim. Mais declaro que todas as fontes usadas estão citadas, no texto e na bibliografia final, segundo as regras de referência adotadas na instituição.

Porto e ISEP, 2023/10/11

A handwritten signature in blue ink that reads "Tomás Simões Jorge". The signature is written in a cursive style with a light blue background behind the text.

# ÍNDICE GERAL

Índice Geral .....	iii
Resumo.....	v
Abstract .....	vii
Agradecimentos .....	ix
Índice de Texto .....	xi
Índice de Figuras.....	xv
Índice de Tabelas.....	xxi
CAPÍTULO 1    Introdução.....	23
CAPÍTULO 2    Estado da Arte.....	29
CAPÍTULO 3    Caso de estudo: Modelação dinâmica do sistema veículo-via .....	77
CAPÍTULO 4    Metodologias para identificação de defeitos isolados baseadas numa rede neuronal do tipo <i>Autoencoder</i> .....	101
CAPÍTULO 5    Algoritmo para deteção e identificação de múltiplos defeitos.....	121
CAPÍTULO 6    Conclusões, desenvolvimentos futuros e publicações .....	141
Referências Bibliográficas .....	145
Anexo 1.....	153
Anexo 2.....	159



## RESUMO

A identificação precoce dos defeitos das rodas ferroviárias pode evitar danos graves quer ao nível do veículo quer da via reduzindo consideravelmente os custos de manutenção tanto para as administrações ferroviárias como para os operadores do material circulante. No presente trabalho é apresentada uma metodologia não supervisionada baseada em técnicas de inteligência artificial, que permite a identificação de multi-danos, por exemplo danos “*Out of Roundness*” e desequilíbrios de carga, com recurso a respostas dinâmicas induzidas na via pela passagem de um veículo ferroviário de mercadorias do tipo *Laagrss*.

Nesta fase inicial, para testar a eficácia da metodologia, foram utilizados registos numéricos das passagens dos comboios que foram simulados com base em modelos numéricos do veículo, de irregularidades da via e das rodas e a interação roda-carril. Estas respostas são posteriormente sujeitas a um conjunto de etapas: i) análise *outlier*, ii) localização do dano, iii) identificação do tipo de dano e iv) classificação da severidade do dano.

Na primeira fase, a deteção dos danos é realizada com base num limite de confiança usando o sinal completo. De seguida, é realizada a localização dos diversos danos presentes no veículo, através de uma segmentação do sinal, permitindo estabelecer janelas de corte de forma a isolar os vários danos, a fim de tornar o problema menos complexo. Numa terceira fase, já com os vários sinais cortados, as principais características são extraídas através da combinação da técnica da análise de componentes principais e de um *autoencoder* esparso, com a “*Bottleneck layer*” a desempenhar um papel crucial na obtenção dos indicadores. Estes indicadores após sofrerem uma fusão, com recurso à distância de *Mahalanobis*, são comparados em termos de extensão, aceleração e frequência, com o objetivo de distinguir qual o tipo de defeito presente no veículo, se poligonização, liso ou desequilíbrio de cargas. Por fim, com base numa metodologia desenvolvida nesta tese para tratar defeitos isolados, é possível distinguir as diferentes severidades do dano.

**Palavras-chave:** Danos nas rodas OOR; desequilíbrio de cargas; deteção e classificação de danos; comboios de mercadorias; sistemas de monitorização *wayside*; *autoencoder* esparso empilhado; inteligência artificial.





## ABSTRACT

Early identification of railway wheel defects can prevent serious damage to both the vehicle and the track, considerably reducing maintenance costs for both railway administrations and rolling stock operators. To this end, this paper presents an unsupervised methodology based on artificial intelligence techniques, which allows the identification of multi-damage, for example “Out of Roundness” damage and unbalance loads, using dynamic responses induced on the track by the passage of a Laagrss-type rail freight vehicle.

In this initial phase, to test the effectiveness of the methodology, numerical records of train passages were used and simulated using numerical models of the vehicle, track and wheel irregularities and wheel-rail interaction. These responses are then subjected to a series of steps: i) outlier analysis, ii) damage localization, iii) damage type identification and iv) damage severity classification.

In the first phase, damage detection is carried out based on a confidence limit using the complete signal. Next, the various damages present on the vehicle are located by segmenting the signal, allowing cutting windows to be established to isolate the various damages in order to make the problem less complex. In a third phase, with the various signals cut out, the main characteristics are extracted by combining the principal component analysis technique and a sparse autoencoder, with the "Bottleneck layer" playing a crucial role in obtaining the indicators. After being merged using the Mahalanobis distance, these indicators are compared in terms of strain, acceleration and frequency, with the aim of distinguishing the type of defect present in the vehicle, whether polygonization, flat or unbalance loads. Finally, based on a methodology developed in this thesis to deal with single damage, it is possible to distinguish the different severities of the damage.

**Keywords:** OOR wheel damage; unbalance loads; damage detection and classification; freight trains; wayside monitoring systems; stacked sparse autoencoder; artificial intelligence.



## **AGRADECIMENTOS**

Ao chegar ao fim destes cinco anos não podia estar mais contente com as coisas que conquistei. Foi uma longa jornada, com altos e baixos, mas repleta de momentos bons e importantes para o meu crescimento pessoal e profissional. Sempre fui fascinado pela Engenharia Civil e desde pequeno me foi inculcido o gosto por esta área. Lembro-me de ir para as obras com o meu pai desde muito novo e ficar pasmado a olhar para a evolução das construções. É incrível pensar que cresci e que continuo a admirar da mesma forma todos os grandes edifícios do Porto nas caminhadas de final de dia. É como se costuma dizer, quando se faz o que se gosta e junto das pessoas certas, tudo se torna mais fácil e enriquecedor.

Esta caminhada não teria sido possível sem a ajuda de algumas pessoas. Portanto, quero expressar os meus sinceros agradecimentos a todos os que me acompanharam durante este percurso, através da sua amizade e compreensão, e que de direta ou indiretamente contribuíram para a realização desta tese.

Em especial, gostaria de agradecer: Aos meus pais e irmã que são a minha razão de viver. Não há palavras para agradecer o que fazem por mim e o quanto se esforçam e trabalham para me dar as melhores condições para ser melhor. A pessoa resiliente e esforçada que sou hoje deve-se principalmente a vocês! Aos avós incansáveis, carinhosos e babados que tenho. A felicidade de poder chegar de fim de semana e estar com vocês nem que seja 10 minutos é inexplicável. É uma lufada de ar fresco para o resto da semana. Que bom que é ter-vos como segundos pais. Aos tios, padrinhos, primos que mesmo não estando sempre comigo se lembram de me dar uma palavra amiga. Esta vitória também é vossa! A uma das pessoas mais importantes da minha vida, a Filipa. Entrámos juntos nesta aventura e saímos dela juntos. Pessoa que levou com as minhas choradeiras, más disposições e risos de felicidade e que sempre me apoiou a ser melhor a cada dia. Esta vitória deve-se grande parte a ti, sem dúvida. Que contemos mais vitórias juntos! Ao prof. Diogo Ribeiro que me abriu as portas para áreas desafiantes e que sem a sua ajuda nada disto seria possível. Agradecer pela pessoa fantástica e trabalhadora que é, e por todos os conhecimentos que me passou. É sem dúvida uma pessoa com um conhecimento invejável e no qual me inspiro diariamente. Aos meus colegas de trabalho, Jorge, António e Rúben. Pessoas trabalhadoras e sempre prontas a ajudar. Sem dúvida que foram fundamentais no sucesso deste trabalho! Á equipa do projeto WAY4SAFERAIL, em principal aos meus orientadores, prof. Diogo Ribeiro, Rúben e Araliya, pelo carinho, atenção e transmissão de conhecimentos. Foi um orgulho trabalhar com pessoas de renome da área da ferrovia!



## ÍNDICE DE TEXTO

1.1	Enquadramento .....	23
1.2	Motivação .....	26
1.3	Objetivos .....	26
1.4	Estrutura da tese.....	27
2.1	Considerações iniciais .....	29
2.2	Sistemas de monitorização para deteção de danos .....	30
2.3	Tipos de veículos ferroviários de mercadorias .....	32
2.4	Classificação das rodas OOR .....	34
2.4.1	Defeitos discretos .....	35
2.4.2	Irregularidade das rodas.....	42
2.5	Desequilíbrio de cargas.....	47
2.6	Metodologias de simulação numérica do sistema dinâmico veículo-estrutura.....	49
2.6.1	Metodologia numérica baseada em cargas móveis .....	49
2.6.2	Metodologia numérica baseada em caminhos virtuais.....	50
2.6.3	Metodologia numérica que considera a interação veículo-estrutura.....	51
2.7	Técnicas baseadas em Inteligência Artificial de identificação de danos .....	54
2.7.1	Levantamento de metodologias desenvolvidas na literatura .....	54
2.7.2	Extração de indicadores .....	57
2.7.3	Fusão dos dados .....	71
2.7.4	Discriminação de indicadores.....	72
3.1	Descrição do modelo proposto.....	77
3.2	Modelação do veículo.....	77

3.3	Modelação da via .....	80
3.4	Modelação do perfil de irregularidades da via .....	82
3.5	Modelação do perfil de irregularidade das rodas.....	83
3.5.1	Modelação do perfil de irregularidade da roda poligonizada baseado no número de harmónicos.....	83
3.5.2	Modelação do perfil de irregularidade de liso na roda .....	85
3.6	Interação dinâmica veículo-via .....	86
3.7	Simulações .....	87
3.7.1	Sistema de monitorização <i>Wayside</i> virtual.....	88
3.7.2	Cenários de base.....	88
3.7.3	Cenários de dano .....	89
3.7.4	Respostas dinâmicas da via .....	91
4.1	Introdução.....	101
4.2	Metodologia 1.....	102
4.3	Metodologia 2.....	108
4.4	Análise de sensibilidade .....	113
4.4.1	Velocidade .....	113
4.5	Discussão dos resultados .....	118
5.1	Introdução.....	121
5.2	Algoritmo para análise multi dano.....	122
5.2.1	Deteção do dano.....	123
5.2.2	Localização do dano.....	123
5.2.3	Identificação do tipo de dano .....	125
5.2.4	Classificação da severidade do dano .....	130
5.3	Validação do algoritmo .....	132
5.3.1	Deteção do dano.....	135
5.3.2	Identificação do tipo de dano .....	135

5.3.3	Classificação da severidade do dano .....	138
6.1	Conclusões .....	141
6.2	Desenvolvimentos futuros.....	143
6.3	Publicações .....	144
6.3.1	Artigos em conferências nacionais.....	144
6.3.2	Artigos em revistas internacionais .....	144





## ÍNDICE DE FIGURAS

Figura 1 – Monitorização da integridade estrutural e as suas vantagens. ....	25
Figura 2 – Sistemas de monitorização: a) <i>onboard</i> e b) <i>wayside</i> (adaptado de [46]).....	30
Figura 3 – Sensores de medição: a) acelerómetro b) extensómetro c) sensores de fibra de <i>Bragg Grating</i> . .....	31
Figura 4 – Veículo de mercadorias com <i>bogie</i> : a) visão geral do vagão [49] b) suspensões helicoidais [50] c) esquema simplificado do veículo d) registo de aceleração característico.....	32
Figura 5 – Veículo de mercadorias sem <i>bogie</i> : a) visão geral do vagão [50] b) suspensões de lâminas [50] c) esquema simplificado do veículo d) registo de aceleração característico.....	33
Figura 6 – Causas da roda OOR (adaptado de [51])......	34
Figura 7 – Tipos de rodas OOR.....	35
Figura 8 – <i>Spalling</i> [51].....	36
Figura 9 – <i>Shelling</i> [51].....	36
Figura 10 – Lisos nas rodas [48]. ....	37
Figura 11 – Características geométricas de um liso na roda (adaptado de [57]).....	38
Figura 12 – Liso arredondado [56]. ....	39
Figura 13 – Irregularidade no carril equivalente ao liso da roda. ....	40
Figura 14 – Características geométricas e respetivos registos tipo em termos de aceleração e deslocamento.....	40
Figura 15 – Corrugação [51]. ....	42
Figura 16 – Esquema ilustrativo com parâmetros que permitem caracterizar a poligonização.....	43
Figura 17 – Imperfeições circulares: excentricidade radial ( $\Delta r$ ) (adaptado da EN 15313-2016). ....	44
Figura 18 – Exemplos de diferentes ordens possíveis: a) $\theta = 1$ , excentricidade, b) $\theta = 2$ , ovalização, c) $\theta = 3$ , triangular, d) $\theta = n$ , poligonal. ....	45

Figura 19 – Esboço da roda e das variáveis de medição. Perfil radial $R(x)$ , raio médio da roda $R_0$ , distância circunferencial $x$ e o ângulo correspondente $\varphi$ [48].....	46
Figura 20 – Esquemas corretos e incorretos de distribuição de cargas. ....	48
Figura 21 – Desequilíbrio de cargas: a) longitudinal b) transversal (adaptado de [79]). ....	49
Figura 22 – Metodologia numérica baseada em cargas móveis [81].....	50
Figura 23 – Metodologia numérica baseada em caminhos virtuais: a) Fase 1 b) Fase 2 [81].....	51
Figura 24 – Possíveis cenários na procura do ponto de contacto [85]. ....	52
Figura 25 – Diferenças entre Inteligência Artificial, <i>Machine Learning</i> e <i>Deep Learning</i> (adaptado de [98]). ....	61
Figura 26 – Exemplos de alguns tipos de redes neuronais [100]. ....	62
Figura 27 – Esquema ilustrativo da arquitetura base de um <i>autoencoder</i> (adaptado de [99]). ....	63
Figura 28 – Funcionamento de um neurónio artificial (adaptado de [102]). ....	64
Figura 29 – Identificação das camadas e ligações da rede. ....	66
Figura 30 – Gradiente da função de custo, adaptado de [104]. ....	67
Figura 31 – Arquitetura adotada no SSAE utilizado no caso de estudo. ....	70
Figura 32 – <i>Outlier</i> . ....	73
Figura 33 – Distâncias maximizadas e minimizadas na análise de <i>clusters</i> [110]. ....	74
Figura 34 – Alocação dos indicadores em diferentes clusters (adaptado de [25]). ....	76
Figura 35 – Vagão de carga: a) vista geral do vagão b) itinerário de viagem (adaptado de [115]). ....	78
Figura 36 – Modelo dinâmico do vagão de carga <i>Laagrss</i> (adaptado de [21]). ....	79
Figura 37 – Modelo numérico do vagão de carga <i>Laagrss</i> [115]. ....	79
Figura 38 – Modelo numérico da via: a) representação multicamadas b) modelo de elementos finitos (adaptado de [17]).....	81
Figura 39 – Perfil de irregularidade da via (direção vertical, V e transversal, L) (adaptado de [27]). ....	83
Figura 40 – Espetros da irregularidade da roda ( $Lw$ ) e ordem harmónica ( $\theta$ ) de rodas poligonais com ordens harmónicas dominantes: a) 6-8 harmónicos b) 12-14 harmónicos c) 17-18 harmónicos. ....	85
Figura 41 – Exemplos de perfis de danos poligonais para cada severidade de dano. ....	85
Figura 42 – Exemplos de perfis de lisos nas rodas de cada severidade. ....	86

Figura 43 – Esquematização do modelo de interação veículo-via. ....	87
Figura 44 – Sistema de monitorização <i>wayside</i> virtual (em perspetiva 3D). ....	88
Figura 45 – Cenários de <i>baseline</i> (sem dano). ....	89
Figura 46 – Cenários de dano considerados: a) identificação dos vagões b) poligonização c) lisos d) desequilíbrios de carga. ....	90
Figura 47 – Diferenças entre registos de acelerações no sensor 1 considerando cenários de base: a) influência da velocidade b) influência da carga c) influência da irregularidade da via. ....	92
Figura 48 – Comparação de registos de aceleração com dano e sem dano e de diferentes severidades com presença de poligonização. ....	93
Figura 49 – Comparação de registos de aceleração de diferentes severidades e diferentes velocidades. ....	94
Figura 50 – Respostas no domínio da frequência na posição 1 referente à poligonização para 60 km/h. ....	95
Figura 51 – Comparação de registos de aceleração com dano e sem dano e de diferentes severidades com presença de lisos. ....	96
Figura 52 – Comparação de registos de aceleração das diferentes severidades de lisos com a alteração da velocidade. ....	97
Figura 53 – Respostas no domínio da frequência na posição 1 referente aos lisos para 60 km/h. ....	98
Figura 54 – Comparação de registos de extensões e acelerações dos diferentes cenários de desequilíbrios de carga, para a velocidade de 60 km/h. ....	99
Figura 55 – Fluxograma referente à metodologia 1 com os principais passos da metodologia de identificação de danos isolados. ....	103
Figura 56 – Extração de indicadores para uma velocidade a partir do acelerómetro 1 referente aos cenários de dano de lisos. ....	104
Figura 57 – Extração de indicadores para 1 velocidade a partir do acelerómetro 1 referente aos cenários de dano de lisos. ....	106
Figura 58 – Fusão de dados para 1 velocidade referente aos lisos: a) 1º nível - fusão de características para o sensor 2 b) 2º nível - fusão de todos os sensores. ....	106
Figura 59 – Classificação automática da severidade da poligonização com base na análise de clusters para 1 velocidade, utilizando 6 acelerómetros (3 de cada lado). ....	107
Figura 60 – Classificação automática da severidade dos lisos com base na análise de clusters para 1 velocidade, utilizando 6 acelerómetros (3 de cada lado). ....	107

Figura 61 – Classificação automática da severidade dos desequilíbrios de carga com base na análise de clusters para 1 velocidade, utilizando 5 extensómetros do lado do desequilíbrio.....	108
Figura 62 – Fluxograma referente à metodologia 2 com os principais passos da metodologia de identificação de danos isolados. ....	109
Figura 63 – Extração de indicadores para 1 velocidade a partir do acelerómetro 6 referente aos cenários de dano de lisos.....	109
Figura 64 – Fusão de dados para uma velocidade referente aos lisos: a) 1º nível - fusão de características para o sensor 2 b) 2º nível - fusão de todos os sensores.....	111
Figura 65 – Classificação automática da severidade da poligonização com base na análise de clusters para uma velocidade, utilizando 6 acelerómetros (3 de cada lado): a) Amplitude 1 b) Amplitude 2. ....	112
Figura 66 – Classificação automática da severidade dos lisos com base na análise de clusters para uma velocidade, utilizando 6 acelerómetros (3 de cada lado). ....	112
Figura 67 – Classificação automática da severidade dos desequilíbrios de carga com base na análise de clusters para uma velocidade, utilizando 10 extensómetros. ....	113
Figura 68 – Classificação automática da severidade da poligonização com base na análise de clusters para 3 velocidades, utilizando 6 acelerómetros (3 de cada lado).....	114
Figura 69 – Classificação automática da severidade dos lisos com base na análise de clusters para 3 velocidades, utilizando 6 acelerómetros (3 de cada lado).....	114
Figura 70 – Classificação automática da severidade dos desequilíbrios de carga com base na análise de clusters para 3 velocidades, utilizando 5 extensómetros do lado do desequilíbrio. ....	115
Figura 71 – Classificação automática da severidade da poligonização com base na análise de clusters para 3 velocidades, utilizando 6 acelerómetros (3 de cada lado).....	116
Figura 72 – Classificação automática da severidade dos lisos com base na análise de clusters para 3 velocidades, utilizando 6 acelerómetros (3 de cada lado).....	116
Figura 73 – Classificação automática da severidade dos desequilíbrios de carga com base na análise de clusters para 3 velocidades, utilizando 10 extensómetros. ....	117
Figura 74 – Principais etapas do algoritmo proposto para a deteção e classificação de multi-dano. ....	122
Figura 75 – Deteção automática de danos OOR e desequilíbrios de carga à velocidade de 80 km/h.....	123
Figura 76 – Segmentação de sinais em extensão e aceleração (adaptado de [13]). ....	123

Figura 77 – Fluxograma proposto para a etapa da identificação do tipo de dano. ....	125
Figura 78 – Aplicação dos testes para identificação do tipo do dano: a) 1º teste: Extensões b) 2º teste: Acelerações c) 3º teste: FFT. ....	129
Figura 79 – Classificação automática da severidade dos 3 defeitos com base na análise de clusters para a velocidade de 80 km/h e com corte de sinal: a) poligonização b) lisos c) desequilíbrio de cargas. .	131
Figura 80 – Cenários multi dano considerados: a) configuração tipo 1 (CT1) b) configuração tipo 2 (CT2) c) configuração tipo 3 (CT3). ....	133
Figura 81 – Sinal completo com multi dano em acelerações e em extensões no sensor 1 (lado direito): a) CT1) b) CT2) c) CT3. ....	134
Figura 82 – Detecção automática de multi danos à velocidade de 80 km/h. ....	135
Figura 83 – Etapas para identificação do tipo do dano para o caso de multi dano: a) 1º teste: Extensões b) 2º teste: Acelerações c) 3º teste: FFT. ....	138
Figura 84 – Classificação automática da severidade dos 3 defeitos com multi dano para uma velocidade de 80 km/h: a) poligonização b) lisos c) desequilíbrio de cargas. ....	139



## ÍNDICE DE TABELAS

Tabela 1 – Valores limites dos comprimentos dos defeitos nas rodas. ....	37
Tabela 2 – Principais estudos realizados para lisos nas rodas e respectivas conclusões. ....	41
Tabela 3 – Características da corrugação.....	42
Tabela 4 – Características da poligonização.....	43
Tabela 5 – Valores admissíveis para o efeito de falta de arredondamento $\Delta r$ (mm).....	44
Tabela 6 – Principais estudos realizados para poligonização nas rodas e respectivas conclusões. ....	47
Tabela 7 – Principais metodologias para detecção de lisos nas rodas e respectivas conclusões. ....	55
Tabela 8 – Principais metodologias para detecção de poligonização nas rodas e respectivas conclusões. .	56
Tabela 9 – Principais metodologias para detecção de desequilíbrios de carga e respectivas conclusões. ..	57
Tabela 10 – Funções de ativação e respectivas características. ....	65
Tabela 11 – Regiões de decisão num problema de classificação (adaptado de [102])......	68
Tabela 12 – Hiperparâmetros utilizados para o treino no SSAE referentes ao caso de estudo.....	71
Tabela 13 – Estudos relacionados com a análise cluster e respectivas conclusões. ....	76
Tabela 14 – Propriedades geométricas e mecânicas do modelo numérico do vagão <i>Laagrss</i> [115]. ....	80
Tabela 15 – Propriedades mecânicas da via (adaptado de [48]). ....	81
Tabela 16 – Parâmetro de rugosidade para as classes da via [119]......	82
Tabela 17 – Coeficiente para a densidade espectral de potência [119]. ....	83
Tabela 18 – Intervalos das ordens dos harmónicos estudadas.....	84
Tabela 19 – Intervalos para comprimento dos lisos. ....	86
Tabela 20 – Hiperparâmetros utilizados no SSAE para 1 velocidade (80 km/h) e para os 3 tipos de dano. .....	105

## ÍNDICE DE TABELAS

Tabela 21 – Hiperparâmetros utilizados no SSAE para 1 velocidade (80 km/h) e para os 3 tipos de dano. .....	110
Tabela 22 – Hiperparâmetros utilizados no SSAE da metodologia 2 para 3 velocidades. ....	115
Tabela 23 – Falsas classificações relacionadas com a análise de sensibilidade da velocidade. ....	117
Tabela 24 - Classificação dos métodos de identificação de danos.....	121
Tabela 25 – Exemplificação do corte dos sinais completos realizados através da segmentação. ....	124
Tabela 26 – Hiperparâmetros utilizados no SSAE para a identificação do tipo de dano. ....	127
Tabela 27 – Intervalos de valores referentes a cada tipo de defeito para identificação do tipo de dano. .....	129
Tabela 28 – Hiperparâmetros utilizados no SSAE para 1 velocidade (80 km/h) e para os 3 tipos de defeitos.....	130
Tabela 29 – Intervalos de valores referentes a cada tipo de defeito para classificação. ....	132
Tabela 30 – Identificação do tipo de dano com base nos limites de DI definidos na fase de teste (CT1). .....	136



# CAPÍTULO 1

## INTRODUÇÃO

### 1.1 ENQUADRAMENTO

As rodas de veículos ferroviários estão sujeitas a um desgaste constante, que com o passar do tempo, e associado muitas vezes a inadequados planos de manutenção, esse desgaste pode dar origem a defeitos. Nesse sentido, o uso de estratégias de monitorização com base em sensores são uma alternativa muito vantajosa em termos económicos e tecnológicos para identificar a necessidade de manutenção do material rodante. Nas últimas décadas, os investigadores propuseram dois tipos de sistemas de monitorização comumente referidos como sistemas *onboard* e sistemas *wayside*, sendo a principal diferença a localização dos dispositivos de medição. Nos sistemas *onboard*, os sensores são instalados no veículo para a monitorização da condição tanto do material circulante quanto da via férrea, sendo a maioria das técnicas baseada em sensores acústicos, vibração, ultrassom e imagem [1]. No entanto, esses tipos de sistemas envolvem altos custos para a monitorização da condição das rodas do veículo, uma vez que os sensores precisam de ser instalados em cada veículo para obter uma boa representatividade. Portanto, técnicas de deteção *onboard* são mais adequadas para monitorizar a condição da via [2]. Por outro lado, os sistemas *wayside* são uma solução mais viável para identificar defeitos nas rodas, uma vez que a condição de todas as rodas é estimada indiretamente pela medição das respostas em sensores instalados na via-férrea.

Os defeitos mais comuns nas rodas são os defeitos de circularidade da roda, mais conhecidos como defeitos “*out of roundness*” (OOR) na terminologia inglesa. Como o nome sugere, estes defeitos referem-se a deformações na superfície da circunferência da roda que a tornam não circular. Essa deformação pode ter muitas formas e causas diferentes e provocam forças dinâmicas de alta amplitude, que consequentemente podem conduzir a falhas tanto no veículo como na via-férrea. Bian et al. [3] concluíram que o impacto dinâmico na via e nas travessas aumentava sob estas irregularidades geométricas. Os danos OOR mais frequentemente estudados são os lisos nas rodas (defeitos discretos) e as rodas poligonais (irregularidade periódica). Muitos autores relataram o impacto dos lisos nas respostas de medição, e concluíram que estas eram bastante relacionáveis com a forma e o tamanho do liso, a composição e a velocidade de circulação do veículo [4-6]. No caso das rodas poligonais, este fenómeno

depende muito do tipo de veículo e da sua operação, e o impacto na via está relacionado com as ordens dos harmônicos dominantes [7-9]. Para além dos danos OOR, é ainda comum o aparecimento de fenómenos de instabilidade de carga fundamentalmente em comboios de mercadorias [10, 11].

Para identificar a ocorrência destas anomalias de forma atempada, foram desenvolvidas ideias inovadoras que envolvem a criação de metodologias automáticas usando tecnologia avançada, que permitem a deteção e a classificação destes defeitos. Estas metodologias combinam técnicas de processamento de sinal e aprendizagem automática com base na extração de características, tornando possível distinguir entre situações sem dano e com dano. Até ao momento, muitos estudos têm mostrado bons resultados na deteção de defeitos ferroviários usando essas abordagens, quer no caso dos danos OOR [12-16] quer no desequilíbrio de cargas [17, 18].

A extração de características permite transformar os dados medidos numa forma alternativa em que a correlação com o dano é mais facilmente observada [19]. Existem várias técnicas para extrair características, sendo as mais comuns a transformada contínua *wavelet* (CWT) [20-22], a análise de componentes principais (PCA) [20, 23, 24], os modelos autorregressivos (AR) [20, 25, 26] e os modelos autorregressivos exógenos (ARX) [15, 20, 27].

Embora esses estudos tenham mostrado resultados robustos na deteção e classificação, estes apresentam algumas limitações para lidar com os efeitos não lineares dos efeitos operacionais, de que é exemplo a velocidade. Na maioria destes estudos, é usada uma etapa de normalização de características quando é necessário remover os efeitos operacionais. Geralmente, são usados métodos de variáveis latentes, como PCA [28] ou regressão linear múltipla (MLR) [29].

Assim, para lidar com os fenómenos não lineares e reduzir grandes volumes de dados, foram criadas as primeiras redes neuronais artificiais [30]. Ao longo dos anos, novas e mais avançadas redes foram desenvolvidas, sendo o *autoencoder* (AE) um dos principais tipos. Posteriormente, Olshausen et al. [31] propuseram a criação de uma codificação esparsa e Ng et al. [32] introduziram um termo de penalização esparsa na camada oculta para obter características mais compactas e eficazes, surgindo o *autoencoder* esparsa (SAE). Começam a surgir também combinações de vários SAE's, conhecidos como *autoencoders* esparsos empilhados (SSAE) [33].

Atualmente, estes tipos de AE's são amplamente usados em várias áreas científicas, especificamente na engenharia estrutural, para a deteção de danos em estruturas, especialmente em pontes [34-36]. Em Finotti et al. [36], uma metodologia baseada na combinação do SAE, para extração de indicadores, e do *Support Vector Machine* (SVM), para classificação dos dados, foi aplicada em respostas de aceleração de um viaduto rodoviário no Brasil, obtendo uma elevada eficiência na classificação de danos estruturais.

Muitos destes algoritmos baseados em *autoencoder* incluem já uma classificação de indicadores para discriminar os diferentes tipos de danos e severidades. No entanto, podem ser combinados com outras técnicas frequentemente utilizadas para classificar os dados. Existem os métodos não supervisionados, nos quais o modelo é treinado sem rótulos fornecidos ao conjunto de treino, como *k-means* [25], mapas auto-organizáveis (SOM) [37], isolamento de florestas [13], modelo de mistura gaussiano [38], *k-medoids* [39] e análise de cluster [40], e os métodos supervisionados, que assumem que a hierarquia das categorias no banco de dados é conhecida, como classificadores de *Naive Bayes* [41], classificadores *k-Vizinhos Mais Próximos* [42], *Support Vector Machine* [24] e árvores de decisão [43].

Importa referir que com a aplicação de todas estas metodologias e técnicas não se pretende a total substituição das inspeções visuais já utilizadas há largas décadas, mas sim a criação de um procedimento compatibilizado que permita uma monitorização da integridade estrutural (MIE) adequada. Na Figura 1 é apresentado um esquema que retrata o ciclo de MIE e as suas vantagens.

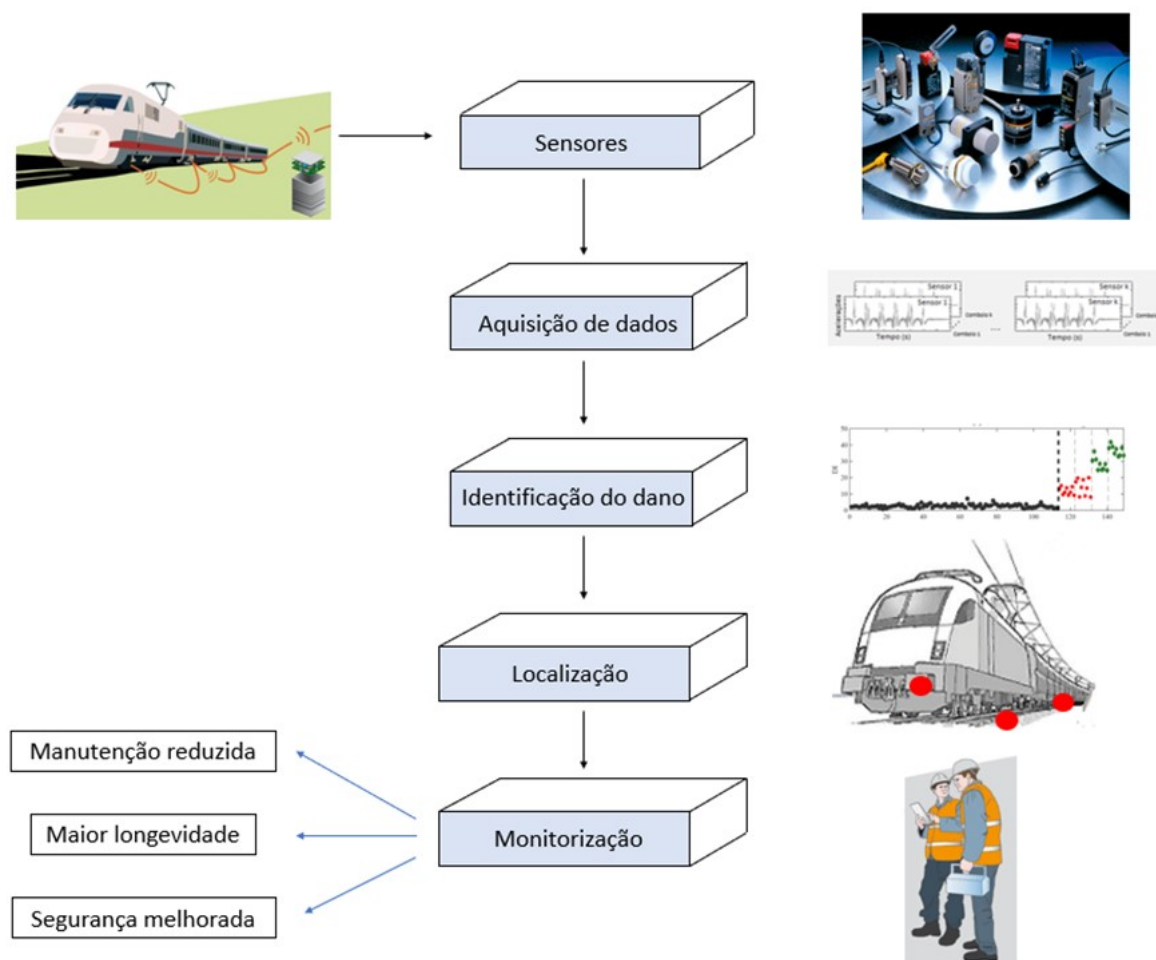


Figura 1 – Monitorização da integridade estrutural e as suas vantagens.

## 1.2 MOTIVAÇÃO

Estudos anteriores demonstram que o estudo de metodologias automáticas baseadas em inteligência artificial para a identificação de danos e instabilidades na circulação, tem sido cada vez mais explorado e apresentado bons resultados [17, 20, 21, 27]. Contudo, todos estes estudos referidos anteriormente foram realizados com base em metodologias que recorrem à extração de indicadores com cariz linear. Como na área da engenharia ferroviária, os efeitos ambientais e operacionais têm uma grande influência, podendo influenciar negativamente as metodologias, é de extrema importância adotar outro tipo de técnicas capazes de anular mais eficientemente estes efeitos. Para isso, com o intuito de dar um passo em frente e tornar a metodologia mais abrangente e independente de variáveis dinâmicas, principalmente a velocidade, recorre-se a um tipo de rede neuronal artificial, o *autoencoder* esparso, que permite classificar os dados de forma não linear. Como estas técnicas baseadas em *autoencoder* ainda têm uma aplicação muito limitada no contexto da engenharia ferroviária, é possível dizer que este trabalho é um estudo inovador e cujas conclusões poderão revelar-se de elevada importância para uma atualização e melhoria das atuais metodologias de IA aplicadas neste contexto.

Dado que estas metodologias possibilitam a identificação e classificação de diversos defeitos nos veículos ferroviários, elas desempenham um papel fundamental no planeamento da manutenção do material circulante e na garantia de segurança durante a circulação ferroviária. Portanto, justifica-se a aposta no estudo e aplicação de ferramentas inovadoras baseadas em inteligência artificial para a monitorização automática e expedita destes problemas.

## 1.3 OBJETIVOS

O trabalho a desenvolver enquadra-se no âmbito do projeto WAY4SAFERAIL – WAYside monitoring system FOR SAFE RAIL transportation, onde o principal objetivo é o desenvolvimento de ferramentas de monitorização das condições dos rodados e de fenómenos de instabilidade de carga de um veículo ferroviário, contribuindo assim para a melhoria da segurança do transporte ferroviário.

Alinhado com os objetivos do projeto, com este trabalho pretende-se o desenvolvimento de uma metodologia de deteção e identificação de múltiplos defeitos num comboio de mercadorias, como danos OOR nos rodados (poligonização e lisos) e desequilíbrios de carga, com base nos registos de acelerações e extensões medidos na via férrea pela passagem de um comboio.

## 1.4 ESTRUTURA DA TESE

A presente dissertação encontra-se dividida em cinco capítulos, sendo estruturado da seguinte forma:

No presente capítulo, Capítulo 1, é realizado um enquadramento do tema e uma pequena apresentação das motivações e objetivos principais que se pretende atingir com o desenvolvimento desta tese.

No Capítulo 2 é apresentado o estado da arte, sendo a secção 2.1 destinada a um enquadramento geral do que será tratado nas secções que compõem o capítulo; na secção 2.2 são apresentados os dois tipos de sistemas de monitorização, os *onboard* e os *wayside*, com a apresentação das respetivas vantagens e desvantagens; a secção 2.3 explora as principais características dos veículos ferroviários de mercadorias que mais frequentemente circulam sobre as vias ferroviárias; na secção 2.4 são explicados sucintamente os danos OOR e os desequilíbrios de carga. Resumidamente, no caso dos danos OOR, são apresentadas as principais causas para o aparecimento de rodas OOR bem como os vários tipos de roda OOR. É realizada ainda uma caracterização geométrica e um levantamento de estudos realizados para os danos mais importantes e que serão alvo de estudo nos capítulos 4 e 5. No caso dos desequilíbrios de carga, são apresentados os limites que serão usados para definir situações de instabilidade de carga; a secção 2.6 apresenta as várias metodologias que permitem reproduzir o sistema dinâmica veículo-via, para ser possível simular registos numéricos próximos da realidade; a secção 2.7 apresenta as variadas técnicas avançadas que permitem a identificação de danos, utilizadas em contextos de aprendizagem de máquina. Serão abordados os fundamentos teóricos de cada uma das técnicas e as suas funcionalidades e vantagens.

No Capítulo 3 são apresentadas as ferramentas numéricas utilizadas para a aquisição de dados indicativos do comportamento do sistema veículo-via, tratando-se concretamente de medições de acelerações e extensões na via. A secção 3.1, apresenta e justifica a metodologia numérica utilizada para a aquisição de dados; a secção 3.2 aborda o veículo em estudo. Numa primeira fase é realizada uma descrição do veículo utilizado no estudo - vagão do tipo *Laagrss*, de seguida é apresentado o modelo numérico do vagão realizado no programa ANSYS®[44], destacando os elementos e métodos utilizados de forma a simular o real comportamento de cada componente principal do veículo, nomeadamente suspensões, massa e elementos de ligação dos componentes; a secção 3.3 retrata a modelação numérica da via, também com recurso ao programa de elementos finitos ANSYS®, com especial foco nos elementos de representação do solo de fundação, balastro, travessas e carris; a secção 3.4 e 3.5 tratam da modelação com recurso ao programa MATLAB®[45] dos perfis de irregularidades da via e das rodas, respetivamente, definindo os tipos de irregularidades consideradas e os seus pressupostos, nomeadamente a qualidade da via, a ordem dos harmónicos e o comprimento do liso, considerados na caracterização geométrica do dano nas rodas dos veículos; a secção 3.6 demonstra a ferramenta utilizada na formulação do problema de interação dinâmica do sistema veículo-via (VSI) aplicado em MATLAB®, tendo como principal função acoplar todos

## CAPÍTULO 1

os modelos numéricos do sistema veículo-via resolvendo o problema dinâmico em ordem a adquirir respostas nos pontos de controlo; na secção 3.7 são apresentadas as diversas simulações.

No Capítulo 4 são propostas duas metodologias de classificação automática de danos isolados (danos OOR e desequilíbrios de carga) com extração de indicadores baseados na técnica *autoencoder*. Para cada uma das metodologias é apresentada uma descrição dos passos adotados, bem como os resultados obtidos, desde a aquisição de dados até à classificação final dos dados. É ainda realizada uma análise de sensibilidade à velocidade e ao número de sensores, bem como uma análise comparativa à eficiência demonstrada na aplicação de cada uma das metodologias, a fim de definir a metodologia a adotar no estudo do Capítulo 5.

No Capítulo 5, é apresentado o algoritmo desenvolvido para a identificação automática de múltiplos danos. Este algoritmo foi desenvolvido considerando combinações dos três tipos de defeitos considerados neste trabalho (poligonização e lisos nos rodados e desequilíbrios de carga) e envolve vários níveis, desde a deteção à discriminação do tipo de dano, à classificação individual das severidades. Para cada um dos níveis, são apresentadas as técnicas usadas, que são calibradas com os resultados obtidos com a simulação de danos isolados (secção 5.2). Após esta calibração, é realizada uma validação numérica com simulações de passagens com danos múltiplos em diferentes combinações e localizações, a fim de testar a eficiência do algoritmo (secção 5.3).

Por último, no Capítulo 6, são apresentadas as principais conclusões, os desenvolvimentos futuros e as publicações realizadas com base no trabalho realizado nesta tese.

## CAPÍTULO 2

### ESTADO DA ARTE

#### 2.1 CONSIDERAÇÕES INICIAIS

A monitorização da integridade estrutural (MIE) é um processo que envolve um acompanhamento, um controlo e uma avaliação periódica e organizada de dados fornecidos por equipamentos colocados em zonas específicas para o devido efeito. Para a implementação de uma monitorização eficaz numa infraestrutura ferroviária é necessário adotar um correto procedimento para a deteção de danos, com base na resposta dinâmica do sistema veículo-via. Um dos grandes desafios passa pela criação de técnicas capazes de diferenciar alterações motivadas por diversos fatores, quer ambientais ou operacionais, daquelas provocadas por dano.

Torna-se bastante difícil parametrizar comportamentos de algumas estruturas de Engenharia Civil, nomeadamente as infraestruturas ferroviárias, devido à sua complexidade, aos diferentes materiais constituintes e ao seu comportamento não linear quando sujeito a ações exteriores, como cargas de veículos, agentes atmosféricos etc. Estas não permitem a instrumentação de um grande número de sensores devido ao seu tamanho, sendo mais frequente os ensaios de vibração ambiental, pois não requerem a interrupção do trânsito/serviço da estrutura comparativamente com os de vibração forçada, porém, transmitem uma excitação de baixas frequências o que resulta na aquisição de uma variedade de informação mais limitada [44].

No sentido de solucionar este problema é apresentada uma estratégia de aprendizagem de máquina com recurso a técnicas multivariadas que será alvo de estudo na dissertação. Para isso, neste capítulo, é feita uma breve revisão bibliográfica e teórica. Inicialmente aborda-se os dois principais tipos de sistemas de monitorização (*onboard* e *wayside*) e algumas vantagens, desvantagens e campos de aplicação. De seguida, é realizada uma análise dos possíveis danos dos rodados ferroviários, dos quais se destacam a poligonização, os lisos e a corrugação e uma caracterização geométrica dos danos passíveis de serem caracterizados. Ainda neste campo, é realizado um breve estudo dos desequilíbrios de carga.

Com o intuito de facilitar na perceção dos fundamentos da metodologia de deteção e classificação que será aplicada, é feito um estudo das possíveis metodologias que permitem reproduzir o sistema dinâmico

veículo-estrutura, bem como uma breve descrição das várias etapas da metodologia que será aplicada ao caso de estudo. Temas como a extração de indicadores, fusão de dados, análise *outlier* e análise *cluster* serão abordados.

## 2.2 SISTEMAS DE MONITORIZAÇÃO PARA DETEÇÃO DE DANOS

Devido à interação veículo-via são desenvolvidas forças dinâmicas que provocam danos nos rodados dos veículos ferroviários [45]. Por esta razão, é de extrema importância recorrer-se a técnicas inovadoras baseadas em metodologias de deteção de danos robustas, com o intuito de se identificar/detetar esses mesmos danos numa fase precoce e de os classificar segundo a sua gravidade, evitando interrupções no normal funcionamento dos sistemas de transporte ferroviário.

Sendo assim, são utilizados os chamados sistemas de monitorização, que não são mais que um conjunto de ferramentas e processos projetados com a finalidade de rastrear e analisar a integridade de determinado sistema (neste caso o veículo) em tempo real. Estes têm sido explorados e desenvolvidos com uma variedade de sensores e mecanismos de deteção. Para além da seleção de sensores apropriados, a conceção de um sistema de monitorização que incorpora estes sensores, continua a ser uma questão bastante importante. Durante a última década, os investigadores propuseram dois tipos de sistemas de monitorização que são normalmente designados por sistemas do lado da via (*wayside*) e do lado do veículo (*onboard*), sendo a sua principal diferença a localização dos aparelhos de medição. Como podemos ver na Figura 2, nos sistemas *onboard* os aparelhos de medição são acoplados no veículo, enquanto que nos sistemas *wayside* as medições são efetuadas na via.

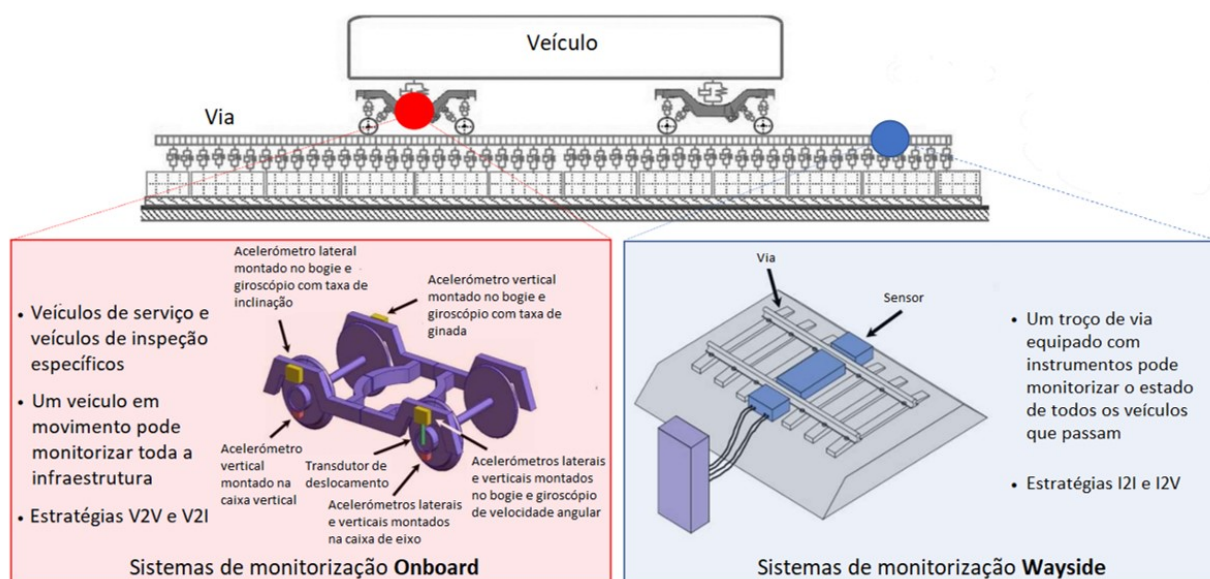


Figura 2 – Sistemas de monitorização: a) *onboard* e b) *wayside* (adaptado de [46]).

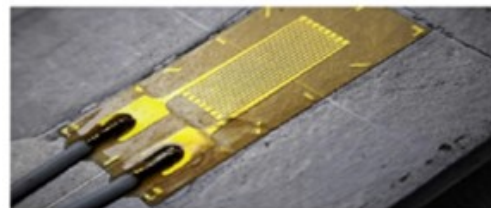


Na sua maioria, as técnicas *onboard* baseiam-se em tecnologias acústicas, de vibração, ultrassónicas e de detecção de imagem [47], em que devem ser instalados sensores nos rodados para um diagnóstico abrangente dos danos e uma gestão eficaz do estado dos rodados. De acordo com Lagneböck [2], estes tipos de sistemas acarretam elevados custos, devido ao facto de ser necessário instalar-se sensores em todos os veículos para se conseguir efetuar uma boa monitorização. Para além de originar um elevado custo de instalação, o grande número de veículos existentes representa um grande desafio de organização e manutenção da tecnologia para detecção de danos em cada um deles.

Segundo alguns autores [48], as técnicas de detecção *onboard* são normalmente utilizadas para monitorizar o estado da via e não para monitorizar o estado dos rodados. Em contrapartida, os sistemas de medição *wayside* são a solução mais viável para identificar defeitos nos rodados, uma vez que o estado de todos os rodados é avaliado durante a passagem dos comboios em serviço. Tipicamente, nestes sistemas, o estado dos rodados é indiretamente estimado através da medição das respostas do carril. Para a medição destes dados são utilizados, por exemplo, acelerómetros (a), medidores de deformação/extensómetros (b) e sensores de fibra de *Bragg Grating* (c).



a)



b)



c)

Figura 3 – Sensores de medição: a) acelerómetro b) extensómetro c) sensores de fibra de *Bragg Grating*.

Segundo Guedes et al. [48], a aplicação de sensores na via permite a extração de um vasto leque de dados relativamente a todos os veículos que sobre esta circulam, fazendo com que seja possível monitorizar-se várias gamas de veículos com um número reduzido de sensores, o que não acontece com o sistema *onboard*. Sendo assim, conclui-se que dependendo do subsistema a monitorizar (via ou veículo), deverá escolher-se o sistema de monitorização mais adequado.

### 2.3 TIPOS DE VEÍCULOS FERROVIÁRIOS DE MERCADORIAS

Existem vários tipos de comboios ferroviários de mercadorias, cada um projetado para atender às necessidades específicas do transporte de cargas. As características específicas destes veículos podem variar de acordo com a região, a infraestrutura ferroviária disponível e as regulamentações locais, podendo ainda ser compostos por uma combinação de diferentes tipos de vagões para atender às necessidades de transporte de cargas variadas. Além disso, existem ainda diferenças na constituição destes veículos que são importantes conhecer, uma vez que na fase da detecção e classificação dos danos nos rodados e dos desequilíbrios de carga, farão com que os registos das passagens dos mesmos variem, influenciando depois a aplicação da metodologia. As frotas de veículos de mercadorias são normalmente compostas por dois tipos de veículos, com *bogie* e sem *bogie*. Nas Figura 4 e 5, é possível observar as suas principais diferenças.

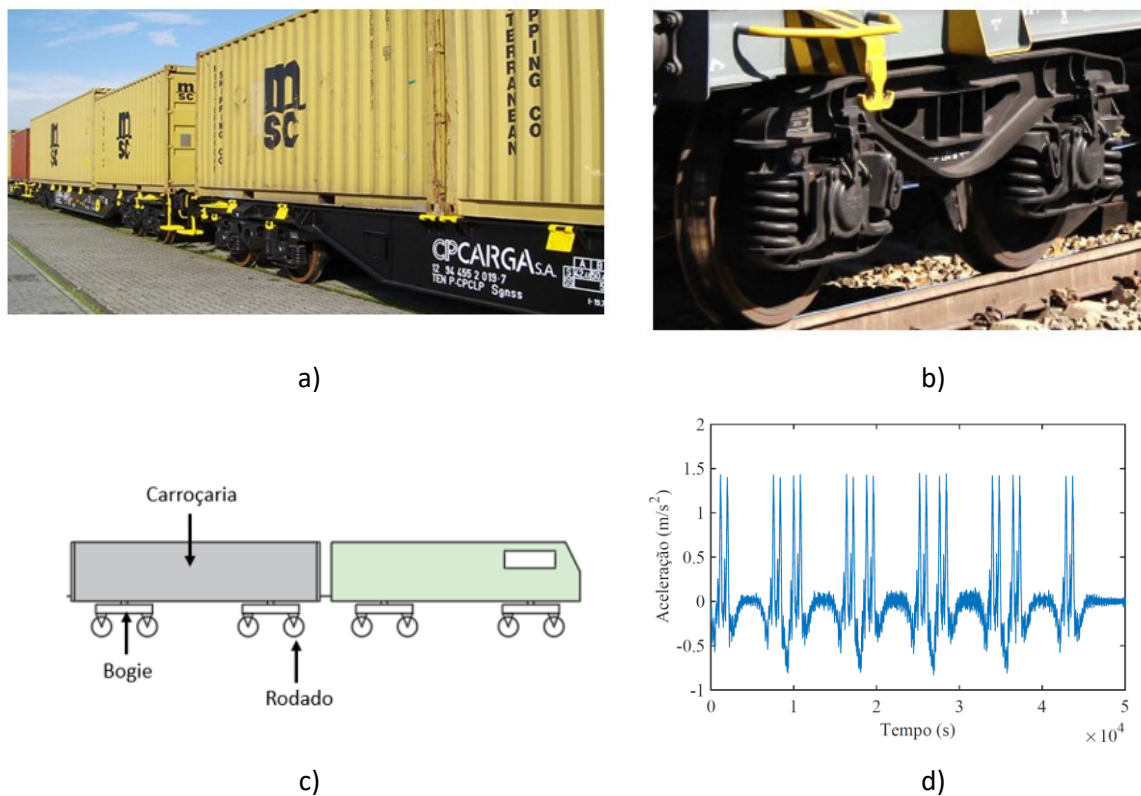


Figura 4 – Veículo de mercadorias com *bogie*: a) visão geral do vagão [49] b) suspensões helicoidais [50] c) esquema simplificado do veículo d) registo de aceleração característico.

Na Figura 4 a) apresenta-se um exemplo de um veículo do tipo com *bogie*, o vagão *Sgnss*, formado por um conjunto de dois eixos suspensos em molas helicoidais (Figura 4 b)) que são agrupados e montados numa estrutura articulada. Essa estrutura permite que o *bogie* se movimente independentemente do resto do veículo, proporcionando uma melhor capacidade de adaptação às curvas e imperfeições da via, logo promovendo uma maior estabilidade (Figura 4 c)). Os veículos ferroviários com *bogies* geralmente

apresentam um número maior de eixos e, portanto, uma capacidade de carga mais elevada. Para além disso, como a carga é distribuída por mais eixos, a pressão exercida sobre a via é menor, logo permite diminuir o desgaste na via. Por fim, na Figura 4 d), é apresentado um registo numérico em acelerações medidas na via de 5 veículos do tipo *Sgnss*, sendo bastante visível os 2 picos consecutivos que representam a passagem dos eixos de cada *bogie*, com acelerações máximas de  $1.5 \text{ m/s}^2$ .

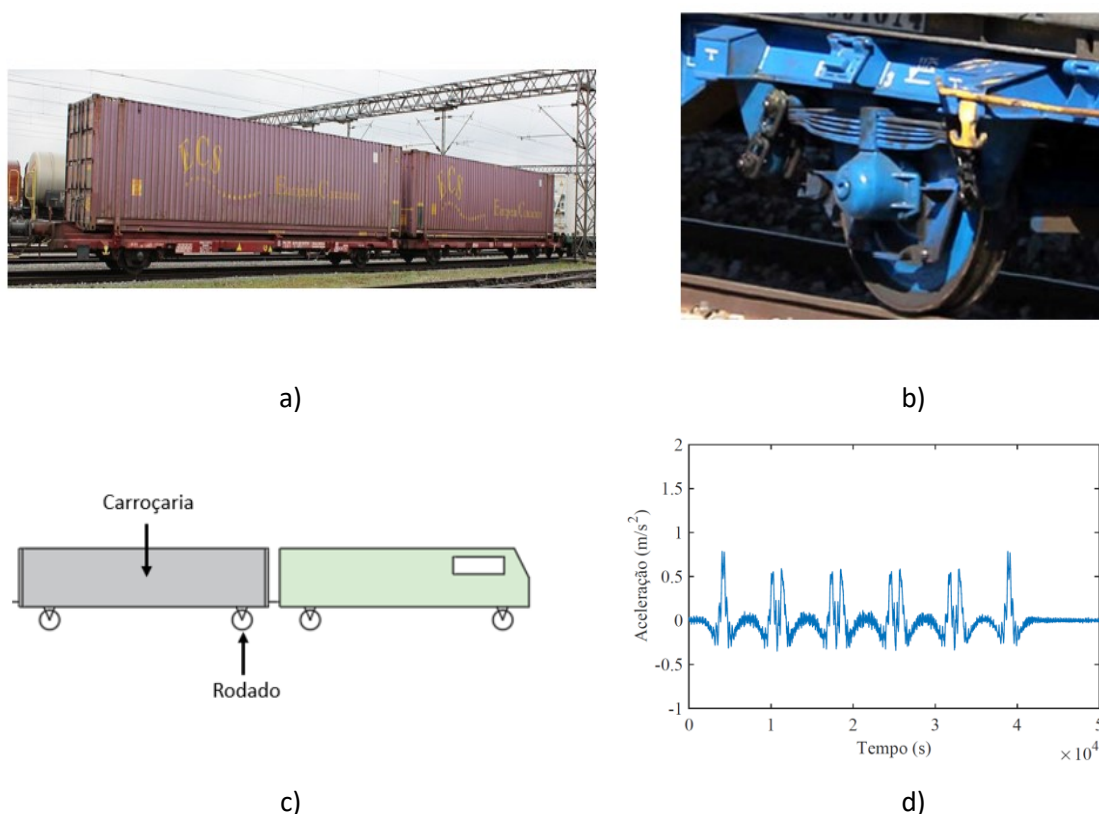


Figura 5 – Veículo de mercadorias sem *bogie*: a) visão geral do vagão [50] b) suspensões de lâminas [50] c) esquema simplificado do veículo d) registo de aceleração característico.

Na Figura 5 a) apresenta-se um exemplo de um veículo do tipo sem *bogie*, o vagão *Laagrss*. Este tipo de veículo é constituído por vagões com dois eixos cada, sem *bogie*. Sendo assim, todos os eixos estão conectados ao veículo por uma suspensão em lâmina (Figura 5 b)) sem suspensão secundária e sem movimento independente (Figura 5 c)). O tipo de suspensão usada tem um comportamento altamente não-linear e apresenta algumas vantagens como o facto de serem mais duráveis e capazes de suportar cargas pesadas, terem uma manutenção relativamente simples quando comparado às molas helicoidais, e fornecerem uma rigidez lateral ao veículo, o que pode ser benéfico em alguma situação de carga pesada. Na Figura 5 d), é apresentado um registo numérico de uma passagem com 5 veículos *Laagrss* e onde se observam diferenças relativamente à anteriormente apresentada para o veículo com *bogie*. Neste caso, apresenta apenas dois picos sucessivos devido à passagem dos dois eixos de cada vagão, com acelerações máximas de  $1 \text{ m/s}^2$ , sendo por isso inferiores às obtidas para o veículo com *bogie*.

## 2.4 CLASSIFICAÇÃO DAS RODAS OOR

As rodas dos veículos ferroviários estão sujeitas a um desgaste constante devido a vários fatores, como por exemplo: à operação intensiva, às quantidades significativas de calor geradas devido ao atrito entre a roda e a via, e à manutenção inadequada. Com o passar do tempo, estes desgastes podem evoluir e tornar-se em danos.

Alguns dos danos que surgem nos rodados podem ser enquadrados num grande grupo denominado de danos de não circularidade da roda, do inglês “*out-of-roundness*” (OOR). Uma roda OOR é, como o nome o indica, uma roda com algum tipo de deformação na superfície da circunferência que a torna não arredondada. Esta deformação pode provocar danos por aumento da força dinâmica no material circulante e na via, pelo que devem ser o mais rapidamente identificados.

As causas para o aparecimento deste tipo de problemas são diversificadas, dependendo de fatores de operação, da condição da via e características do comboio. Estas são demonstradas de forma resumida no esquema da Figura 6.

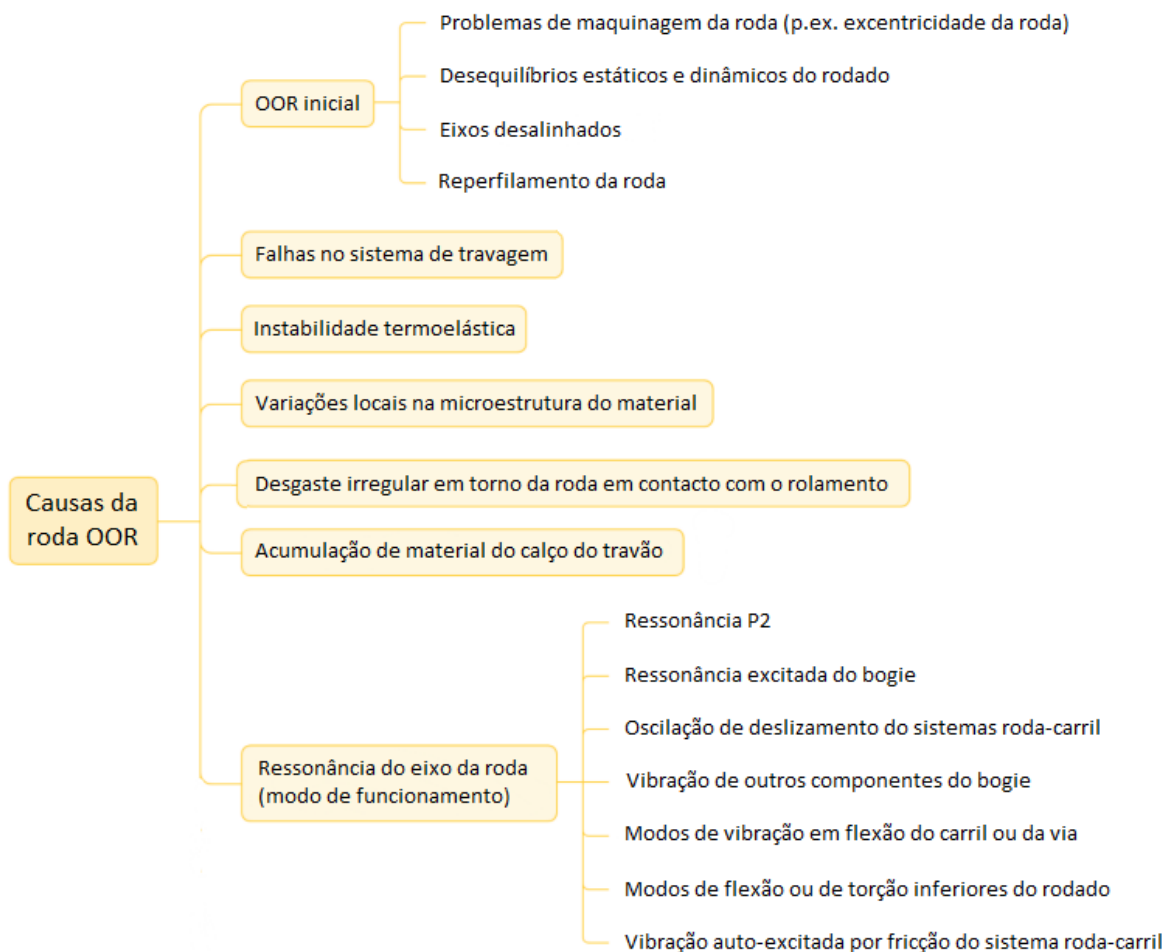


Figura 6 – Causas da roda OOR (adaptado de [51]).

Em termos de classificação das rodas OOR, estas podem ser organizadas conforme a Figura 7, dividindo-se em três principais grupos: defeitos discretos, irregularidades periódicas, irregularidades não periódicas/aleatórias.

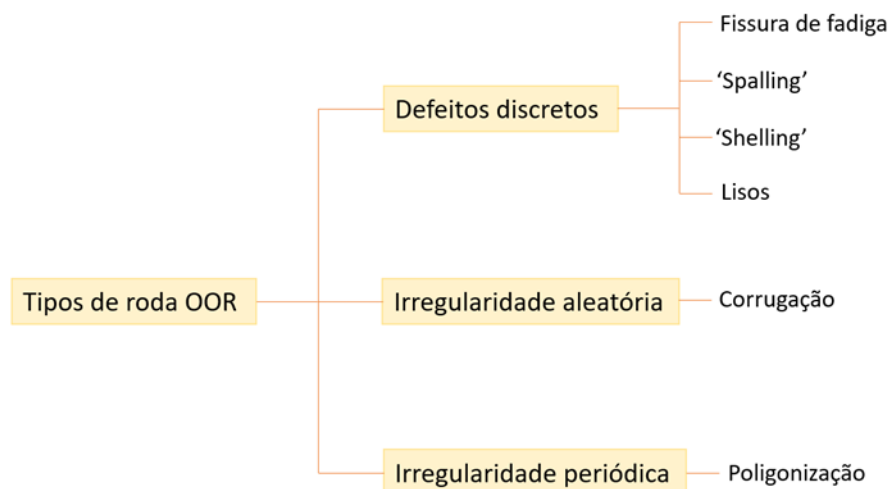


Figura 7 – Tipos de rodas OOR.

Estes defeitos OOR, bem como os restantes defeitos que não estão relacionados com a não circularidade da roda, como a rutura da flange, são todos detalhadamente descritos na norma EN 15313-2016 [52]. Nas figuras seguintes são demonstrados alguns dos exemplos presentes na norma apenas para casos OOR, no entanto será dado maior destaque aos lisos e à poligonização, por se tratarem de danos que são alvo de estudo deste trabalho.

## 2.4.1 Defeitos discretos

Um defeito discreto no perfil da roda é uma variação do raio nominal da roda numa secção pontual da banda de rodagem, o que pode resultar em impacto e carga no contato entre a roda e a via a cada rotação da roda [53]. Esse tipo de defeito representa um risco significativo para a segurança, pois pode desencadear a fratura da via e até mesmo causar um descarrilamento.

### 2.4.1.1 “Spalling”

*Spalling* é o termo utilizado para o fenómeno de fadiga por contacto com o rolamento que ocorre quando as fissuras superficiais de origem térmica se encontram, resultando no afastamento de parte da banda de rodagem [54]. As fissuras térmicas podem surgir devido ao aquecimento e ao arrefecimento rápido da roda durante e após a travagem em bloco, podendo formar-se um padrão multidirecional ou um padrão mais regular.



Figura 8 – *Spalling* [51].

#### 2.4.1.2 “*Shelling*”

Em relação ao *Shelling*, é o termo utilizado para todo o tipo de fissuras induzidas no subsolo. À semelhança do *Spalling* manifesta-se pela perda de material na banda de rodagem da roda, sendo que a principal causa deve-se às forças verticais excessivas de contacto roda-carril.



Figura 9 – *Shelling* [51].

Normalmente, este conceito é frequentemente difícil de distinguir em relação ao *Spalling*, devido à semelhança entre os danos no que toca aos problemas que induz na banda de rodagem da roda, como é possível ver na Figura 8 e Figura 9.

#### 2.4.1.3 Lisos nas rodas

Um plano na superfície de um rodado, em que o aparecimento do mesmo aparece de forma pontual, é designado por liso nas rodas, Figura 10. Este é causado pelo deslizamento da roda sobre o carril que pode ocorrer devido a vários fatores, tais como, sistemas de travagem com anomalias ou mal ajustados, elevadas forças de travagem e defeitos no carril ou ações climáticas que promovam o aparecimento de gelo, neve ou geada sobre o carril [48].



Figura 10 – Lisos nas rodas [48].

#### 2.4.1.3.1 Valores limites normativos

De acordo com a GCU – General Contract of Use for Wagons (2021), é normalmente considerado um comprimento plano máximo do rodado de 60 mm para rodas com diâmetro aproximadamente igual a 840 mm (diâmetro aproximadamente igual ao do veículo de mercadorias em análise).

Segundo estudos feitos anteriormente por Brizuela et al. [55], os comprimentos planos considerados nas simulações numéricas dos sistemas de interação veículo-via variam entre 20-60 mm. No entanto, com o objetivo de detetar precocemente os lisos nas rodas, Mosleh et al. [26] sugerem que esse limite seja reduzido, uma vez que desgastes da ordem de 10-20 mm podem já resultar em cargas de impacto significativas, podendo causar danos no sistema, mesmo quando estão nos estágios iniciais.

Segundo a norma EN 15313 (2016), são definidos um conjunto de valores limite relativamente aos comprimentos dos defeitos nas superfícies das rodas, dependendo das velocidades de circulação, da carga por eixo e do diâmetro da roda. Conforme demonstra a Tabela 1, a partir dos 80 mm estamos perante uma situação extrema de dano, sendo por isso necessário uma intervenção na roda.

Tabela 1 – Valores limites dos comprimentos dos defeitos nas rodas.

M		M ≤ 18			18 < M ≤ 22,5				22,5 < M		
V (km/h)		V ≤ 160	160 < V ≤ 200	200 < V	V ≤ 120	120 < V ≤ 160	160 < V ≤ 200	200 < V	V ≤ 100	120 < V ≤ 160	120 < V
d	1 000 < d	80	60	40	80	60	50	35	X	X	X
	840 < d ≤ 1 000	60	50	30	60	50	35	25	60	50	30
	630 < d ≤ 840	40	30	25	40	30	25	20	40	X	X
	550 < d ≤ 630	35	25	X	X	X	X	X	X	X	X
	d < 550	30	X	X	X	X	X	X	X	X	X

M: carga por eixo em toneladas (t).  
X reservado (não se conhece aplicação).  
d diâmetro da roda (mm).

### 2.4.1.3.2 Caracterização geométrica

No caso dos lisos nas rodas, a sua formação resulta de um processo evolutivo, uma vez que até à formação de um liso acima dos valores limite normativos este passa por algumas fases. Na Figura 11 é demonstrado um esquema representativo, que permite comparar as várias fases de desgaste de um rodado até à criação de um liso e ainda as variáveis que o permitem caracterizar.

As principais fases de desgaste são três. Primeiramente, temos um desgaste inicial de pequena dimensão, em que se dá início à criação do liso. De seguida, numa fase intermédia, há a formação dos cantos do liso, inicialmente não arredondados, mas com o contínuo desgaste da roda conduz à formação de cantos arredondados. Por fim, dá-se a formação de um liso completamente desgastado.

A sua caracterização geométrica depende das seguintes variáveis:

$d$  – Espessura de desgaste;

$L_0$  – Comprimento inicial do desgaste;

$R_w$  – Raio;

$L$  – Comprimento final do desgaste.

Normalmente, a caracterização da forma dos cantos arredondados torna-se difícil, pelo que é usual adotar-se uma aproximação considerando os cantos não arredondados, tal como apresentado na fase intermédia de desgaste (linha tracejada a vermelho, [56]).

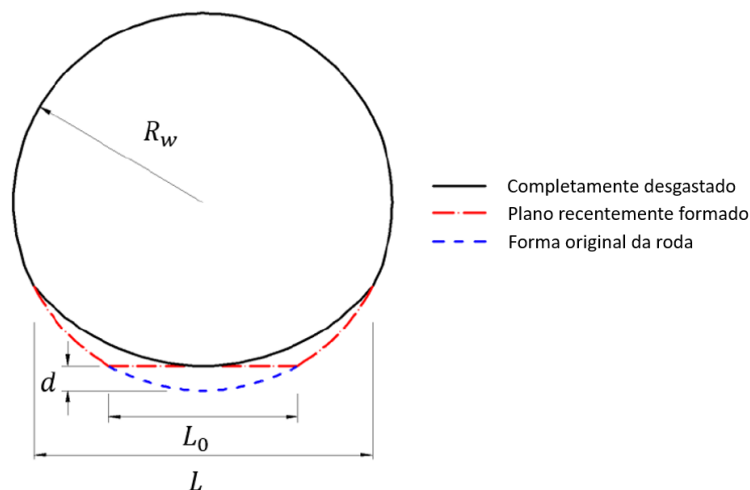


Figura 11 – Características geométricas de um liso na roda (adaptado de [57]).

A caracterização dos lisos pode ser efetuada com base nas equações que se descrevem de seguida, sendo que estas dependem da fase de desgaste em que a roda se encontra.



Na Equação 2.1, em situações de não arredondamento nos cantos (plano recentemente formado), a espessura de desgaste do defeito ( $d$ ) é diretamente relacionada com o raio da roda ( $R_w$ ) e o comprimento do liso ( $L_0$ ).

$$R_w^2 = (R_w - d)^2 + \left(\frac{L_0}{2}\right)^2 \quad (2.1)$$

$$d = \frac{d^2}{2r} + \frac{L_0^2}{8r} \quad (2.2)$$

Uma vez que o valor de  $d$  é um valor insignificante em comparação ao valor de  $L_0$ , é comum desprezar-se a parcela  $\frac{d^2}{2r}$ , surgindo então:

$$d \approx \frac{L_0^2}{8r} \quad (2.3)$$

Para os casos em que não há desgaste e estamos perante uma roda original, Wu et al. [58] propõe a seguinte expressão:

$$x \approx d - \frac{1}{2r} \left(z - \frac{L_0}{2}\right)^2, \text{ se } 0 \leq z \leq L_0 \quad (2.4)$$

Em que:

$z$  – Coordenada do liso.

Em condições reais, como as rodas do veículo são sujeitas a um contínuo desgaste, os cantos dos lisos arredondam-se e o comprimento do liso tende a aumentar. Na ausência do conhecimento real dos perfis de desgaste das rodas, Wu et al. [58] propõe a função quadrática para representar geometricamente os cantos arredondados do liso, como se mostra na Figura 12.

$$x = \begin{cases} \bar{a}z^2 & \Leftarrow 0 \leq z \leq z_1 \\ d - \frac{1}{2r}(z - z_2)^2 & \Leftarrow z_1 \leq z \leq z_2 \end{cases} \quad (2.5)$$

Sendo:

$\bar{a}$  – Parâmetro.

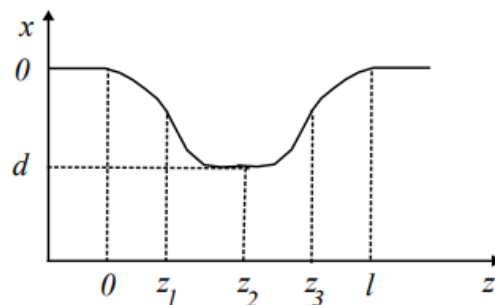


Figura 12 – Liso arredondado [56].

Segundo vários autores, os modelos numéricos de interação veículo-via definem a roda como sendo um ponto de contacto. Para isso, é possível transformar o defeito da roda num defeito equivalente de carril igualmente espaçado, sobre o qual circula uma roda sem defeito. A configuração do defeito da roda incluído na via é dada por:

$$z = \frac{d}{2} \left( 1 - \cos 2\pi \frac{x}{L} \right) H(x - (2\pi R_w - L)), \quad 0 \leq x \leq 2\pi R_w \quad (2.6)$$

Onde:

$z, x$  – Direção vertical e longitudinal da via;

$H$  – Função de Heaviside.

Na Figura 13, mostra-se o perfil de irregularidade do carril equivalente a um liso numa roda do veículo. A cada perimento da roda que passe apresenta-se uma deformação, que representa a pancada proveniente do liso.

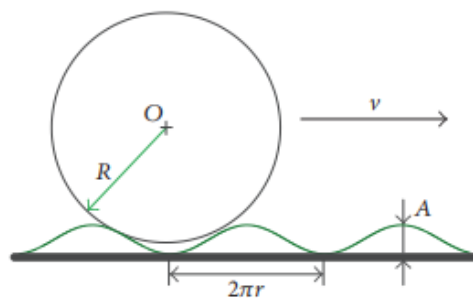


Figura 13 – Irregularidade no carril equivalente ao liso da roda.

A deteção de um liso na roda pode ser observada principalmente em termos de acelerações e deslocamentos. Embora haja uma pequena variação no deslocamento na área onde ocorre o liso, é na aceleração que se obtêm variações mais evidentes. Isso ocorre fundamentalmente devido ao impacto repentino que induz valores de aceleração muito elevados, facilitando na distinção entre um caso sem dano e um caso com dano. Na Figura 14, é possível observar-se um esquema ilustrativo que representa as leituras teóricas de aceleração e deslocamento ao longo do tempo durante a passagem de um liso.

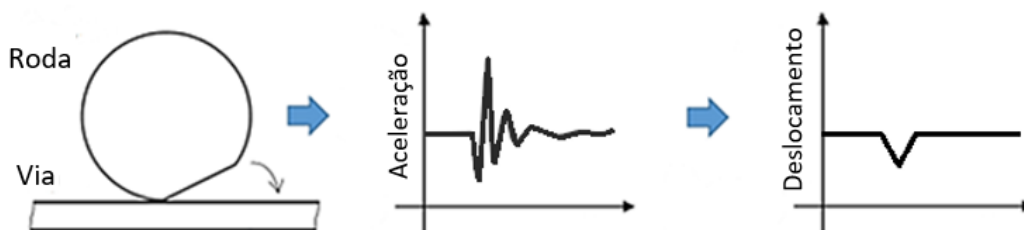


Figura 14 – Características geométricas e respetivos registos tipo em termos de aceleração e deslocamento.

### 2.4.1.3.3 Levantamento de estudos realizados

Os lisos nas rodas são defeitos que estão diretamente relacionados com o atrito desenvolvido entre a roda e o carril aquando da travagem. Diversos autores tem estudado este tipo de defeito, por muitos considerado o dano mais gravoso ao nível dos rodados, pelo que na Tabela 2, se resumem os principais trabalhos desenvolvidos até à data e principais conclusões obtidas em cada estudo.

Tabela 2 – Principais estudos realizados para lisos nas rodas e respetivas conclusões.

<i>Estudos dos lisos nas rodas</i>	<i>Trabalho e conclusões</i>
Contacto multi-ponto de um veículo de alta velocidade - dinâmica do sistema por Ren et al. [59].	Estudou-se o efeito de um modelo de contacto roda-carril numa simulação em que o impacto foi causado por lisos nas rodas. Os resultados mostraram que a correta modelação das características geométricas longitudinais do liso tiveram maior efeito sobre o impacto máximo roda-carril do que no modelo de contacto roda-carril.
Investigação sobre a dinâmica do impacto roda/carril com um modelo tridimensional para lisos por Zunsong et al. [60].	Construiu-se um modelo tridimensional (3D) para lisos e um modelo dinâmico de alta velocidade de acoplamento veículo-via, que teve em conta o comprimento, a profundidade e a largura do liso. Os efeitos de 3 tipos de largura e 5 comprimentos foram estudados para uma faixa segura de velocidade.
Influência de um rodado flexível nas respostas dinâmicas de um vagão de alta velocidade devido a lisos nas rodas por Wu et al. [5].	Analisou-se os efeitos da flexibilidade da roda na resposta dinâmica roda-carril de lisos e estudou-se o impacto dos lisos na força e na aceleração vertical da caixa do eixo.
Cargas de impacto devida a roda planas e "shelling" por Dukkipati et al. [61].	Estudou-se as características da carga de impacto das rodas planas. Descobriu-se que a forma e o tamanho dos lisos, a carga por eixo e a velocidade do veículo foram os principais fatores da carga de impacto.
Análise dinâmica do impacto roda-carril da via férrea sob o liso por análise de elementos finitos por Bian et al. [62].	Construiu-se um modelo 3D de elementos finitos para estudar o impacto dos lisos e concluiu-se que o impacto dinâmico na via e nas travessas aumentavam sob os lisos e conforme o seu tamanho.
Lisos nas rodas podem causar ou agravar a poligonização da roda por Ye et al. [63].	Realizou-se simulações e testes de campo e provou-se que os lisos podem agravar problemas poligonais nas rodas.
O papel da geometria de contacto no impacto roda-carril devida a lisos nas rodas por Steenbergen et al. [64].	Estudou-se o impacto roda-carril causado pelos lisos e caracterizou-se os principais mecanismos e parâmetros, estabelecendo um padrão para a taxa de desgaste do contacto roda-carril.
Modelação, validação e análise de modelo veículo-via 3D com propriedades de via lineares e não lineares na presença de lisos por Uzzal et al. [65].	Construiu-se um modelo veículo-via 3D para estudar a carga dinâmica dos pontos de contacto roda-carril e a resposta dinâmica para lisos nas rodas.

## 2.4.2 Irregularidade das rodas

### 2.4.2.1 Corrugação

A corrugação é uma irregularidade aleatória da superfície do piso da roda com uma amplitude e um comprimento de onda pequenos, que costumam surgir em rodas que são travadas por blocos. Segundo alguns autores, [51, 53], o comprimento de onda dominante deste tipo de OOR é de 3-8 cm, enquanto que a amplitude é inferior a 10 µm, conforme é observável na Tabela 3.

Tabela 3 – Características da corrugação.

<i>Designação</i>	<i>Comprimento de onda</i>	<i>Amplitude</i>	<i>Possível causa</i>
Corrugação	30 - 80 mm	10 µm	Travagem no piso da roda

Vernersson [66] realizou estudos experimentais e teóricos (por meio de simulações numéricas) sobre o desenvolvimento desse tipo de defeito. A hipótese proposta e confirmada durante estes estudos é que, durante a travagem por bloco, certas regiões da superfície do piso da roda tornam-se mais quentes devido a uma instabilidade termoelástica, resultando em pontos quentes. Essas regiões aquecidas elevam-se acima da superfície da roda devido à expansão térmica, o que as torna mais suscetíveis ao desgaste em comparação com as outras partes da superfície do piso da roda. À medida que a roda vai reduzindo a sua temperatura, o volume de material nessas áreas quentes diminui, formando vales, o que resulta num padrão de ondulação, Figura 15. É importante destacar que a corrugação é uma das principais causas de ruído durante o rolamento das rodas [51].



Figura 15 – Corrugação [51].

### 2.4.2.2 Poligonização

Segundo Iwnicki et al. [67], a poligonização é o processo que resulta em rodas poligonais, em que a roda é caracterizada por uma variação regular e periódica do raio. O efeito poligonal ocorre quando a superfície da roda, que originalmente é aproximadamente circular e suave, sofre desgaste e começa a desenvolver uma forma ondulada. Essa deformação pode ser comparada à vibração de uma corda, exibindo características específicas, como comprimentos de onda diferentes dependendo do número de harmônicos gerados e amplitudes. À semelhança dos lisos nas rodas, este tipo de defeito tem sido amplamente observado em todos os tipos de material circulante, incluindo comboios de alta velocidade [68-70], veículos de metro [71, 72] e locomotivas [73].

Quando se estuda uma roda com defeitos ondulatórios, são estudados dois parâmetros principais, o comprimento de onda ( $\lambda$ ), que ditará o número de harmônicos, e a amplitude do harmônico ( $A$ ). Através da Figura 16 é possível observar-se um esquema ilustrativo que apresenta estas características.

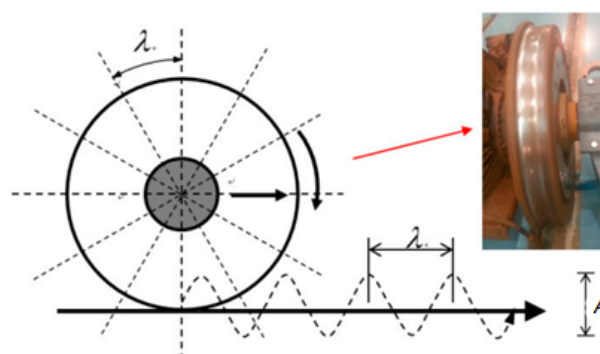


Figura 16 – Esquema ilustrativo com parâmetros que permitem caracterizar a poligonização.

Na Tabela 4 são apresentados os valores correntes quer para o comprimento de onda quer para a amplitude, de uma roda poligonizada.

Tabela 4 – Características da poligonização.

<i>Designação</i>	<i>Comprimento de onda</i>	<i>Amplitude</i>	<i>Possível causa</i>
Poligonização	140 mm - perímetro	> 0,2 mm	Mecanismos de frequência fixa

Caso as frequências de vibração causadas pelos defeitos coincidam com as frequências naturais da via, poderão ocorrer fenómenos de ressonância, amplificando a resposta dinâmica que poderá causar problemas nos componentes do veículo e também na via. Dado que estes fenómenos são bastante comuns, é de extrema importância conhecer melhor como funciona este dano, daí ser dada especial atenção à poligonização na realização deste trabalho.

**2.4.2.2.1 Valores limites normativos**

A norma EN 15313-2016 prevê uma série de valores limite de desgaste ( $\Delta r$ ) para as irregularidades na roda. Os valores de  $\Delta r$  variam consoante o diâmetro das rodas e velocidades de circulação e podem ser encontrados no Anexo G. É de salientar que estes valores são meramente indicativos não representando um requisito legal.

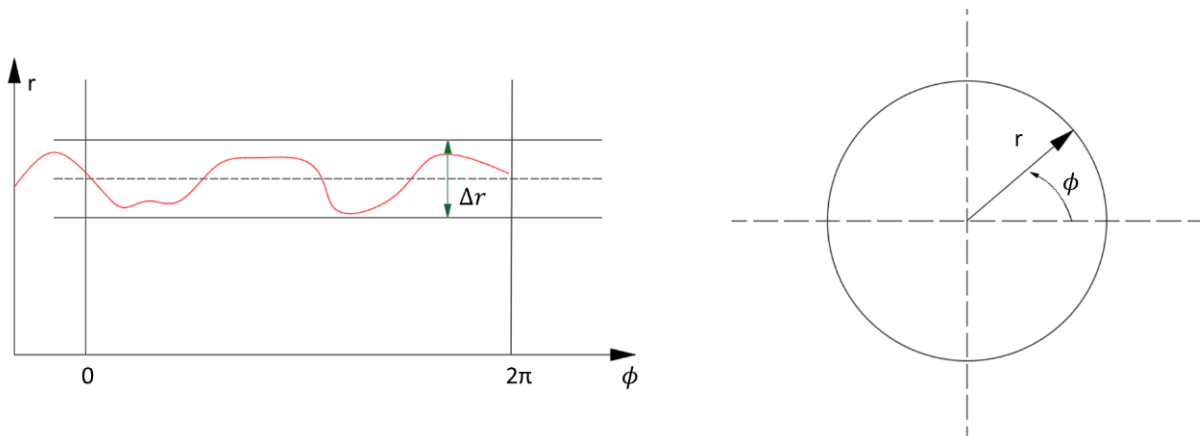


Figura 17 – Imperfeições circulares: excentricidade radial ( $\Delta r$ ) (adaptado da EN 15313-2016).

Segundo a norma, os valores dos defeitos circulatorios admissíveis, ( $\Delta r$ ), devem ser definidos no plano de manutenção.

Tabela 5 – Valores admissíveis para o efeito de falta de arredondamento  $\Delta r$  (mm).

<i>Intervalos de velocidades em função do diâmetro da roda</i>	<i>Defeitos circulatorios admissíveis</i>
<b><i>D &gt; 840 mm</i></b>	
• $v_{m\acute{a}x} \leq 60 \text{ km/h}$	1,5
• $60\text{km/h} < v_{m\acute{a}x} \leq 160\text{km/h}$	1,0
• $160\text{km/h} < v_{m\acute{a}x} \leq 200\text{km/h}$	0,7
• $v_{m\acute{a}x} > 200 \text{ km/h}$	0,5
<b><i>380 mm &lt; D ≤ 840 mm</i></b>	
• $v_{m\acute{a}x} \leq 200 \text{ km/h}$	0,7
• $v_{m\acute{a}x} > 200 \text{ km/h}$	0,5
<b><i>D ≤ 380 mm</i></b>	0,3

**2.4.2.2.2 Caracterização geométrica**

Como já referido anteriormente, a caracterização de uma roda poligonizada pode ser efetuada com base na ordem e na amplitude do harmónico.

A ordem não é mais que o número de lados do polígono da roda, ou seja, o número de picos de onda ao longo da circunferência da roda. Na Figura 18 é possível observar diferentes números de harmónicos.

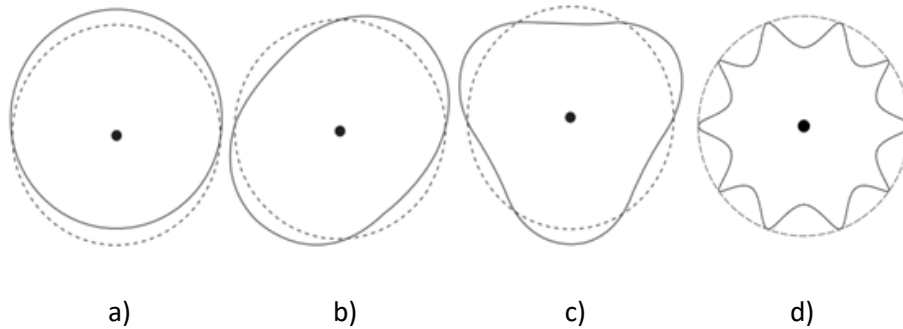


Figura 18 – Exemplos de diferentes ordens possíveis: a)  $\theta = 1$ , excentricidade, b)  $\theta = 2$ , ovalização, c)  $\theta = 3$ , triangular, d)  $\theta = n$ , poligonal.

A variação da ordem está relacionada com os comprimentos de onda, pelo que quanto maior for o comprimento de onda menor será a ordem dos harmónicos. O comprimento de onda pode ser definido de acordo com a Equação 2.7:

$$\lambda_{\theta} = \frac{2\pi R}{\theta}, (\theta = 1, 2, 3, \dots, n) \quad (2.7)$$

Onde:

$\lambda_{\theta}$  – Comprimento de onda;

$R$  – Raio da roda;

$\theta$  – Ordem dos harmónicos.

Em estudos conduzidos por Johansson et al. [72], o perfil de irregularidade do defeito poligonal é caracterizado em função do espaço, podendo ser definido de acordo com a seguinte expressão:

$$w(x_w) = \sum_{\theta=1}^M a_{\theta} \sin\left(\frac{2\pi}{\lambda_{\theta}} x_w + \psi_{\theta}\right) \quad (2.8)$$

Onde:

$w(x_w)$  – Perfil de irregularidade;

$a_{\theta}$  – Amplitude;

$x_w$  – Distância percorrida pela roda;

$\psi_{\theta}$  – Ângulo de fase.

Ainda segundo os mesmos autores, as rodas novas apresentam, apesar de baixas, irregularidades iniciais. O espectro do nível de irregularidade inicial é obtido de uma forma empírica através de dados medidos em rodas novas, conforme a seguinte equação:

$$L_{w\theta} = 24,7 \log(\lambda_\theta) + 8,47 \quad (2.9)$$

Para cada comprimento de onda, a amplitude pode ser obtida de acordo com a seguinte equação:

$$a_\theta = \sqrt{2} \cdot 10^{L_{w\theta}/20} \cdot w_{ref} \quad (2.10)$$

em que  $w_{ref} = 1\mu m$ .

Em termos práticos, para se poder avaliar a profundidade de desgaste originado pelo efeito poligonal, é necessário recorrer-se à seguinte expressão:

$$r(x) = R(x) - R_0 \quad (2.11)$$

Estes princípios são ilustrados na seguinte Figura 19:

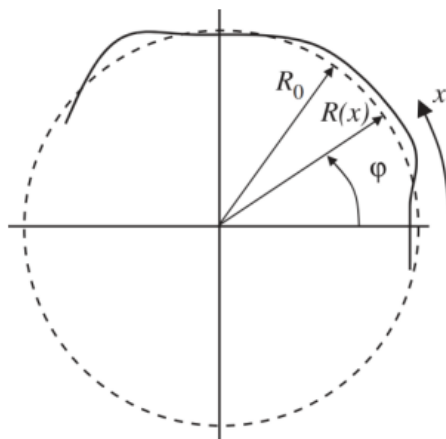


Figura 19 – Esboço da roda e das variáveis de medição. Perfil radial  $R(x)$ , raio médio da roda  $R_0$ , distância circunferencial  $x$  e o ângulo correspondente  $\phi$  [48].

#### 2.4.2.2.3 Levantamento de estudos realizados

Devido à ocorrência frequente deste tipo de dano nos veículos ferroviários, diversos autores têm apresentado estudos sobre o fenômeno da poligonização. Devido à dependência deste fenômeno em relação ao tipo de veículo e à sua velocidade, estes estudos têm sido bastante abrangentes, quanto à sua formação e evolução bem como os impactos na circulação. À semelhança da tabela apresentada para os lisos nas rodas, na Tabela 6, são apresentados alguns dos principais estudos e respectivas conclusões relativos à poligonização.



Tabela 6 – Principais estudos realizados para poligonização nas rodas e respectivas conclusões.

<i>Estudos da poligonização nas rodas</i>	<i>Trabalho e conclusões</i>
Medição e avaliação de defeitos OOR em rodas de locomotivas elétricas/ Investigação experimental da principal característica do desgaste poligonal nas rodas de locomotiva por Tao et al. [73].	Estudos realizados em locomotivas elétricas de alta potência na China, incluindo testes modais das rodas, testes de resposta de frequência da via, testes de vibração das locomotivas na via e simulações da evolução do desgaste das rodas. Os resultados indicaram que os modos de torção do rodado foram responsáveis pelo aparecimento da poligonização. Estas medições extensas indicaram rodas poligonizadas com ordens 17-19 e 24 em três tipos de locomotivas.
Análise da causa raiz para o aparecimento de rodas poligonizadas em locomotivas por Fröhling et al. [74].	Locomotiva sul africana com velocidade máxima de 120 km/h apresentou uma ordem harmónica dominante de 20. Foi proposta uma medida para evitar a formação de rodas poligonais que envolve o uso de um sistema de controlo para detetar e interromper a vibração torsional e aumentando o amortecimento do contato roda-carril.
Medição e avaliação de defeitos OOR em rodas de locomotivas elétricas por Tao et al. [75].	Estudou-se rodas de comboios de mercadorias chinesas em que a sua velocidade máxima é 120 km/h. Após um estudo de mais de 2000 rodas de 7 tipos de locomotivas, concluiu-se que a ordem mais comum é 18-19.
Interação veículo-via e mecanismos de desgaste irregular nas superfícies das rodas e via por Nielsen et al. [76].	Aproximadamente 60% das rodas de metro em Estocolmo apresentam poligonização com uma ordem dominante de 3. Descobriu-se que o desenvolvimento das ordens de roda poligonal 14-20 foi causado por antirressonância enquanto o aumento das ordens 5-7 foi devido à ressonância P2 e à excitação da travessa a uma velocidade de 50 km/h.
Problemas de desgaste e RCF de sistemas roda-carril de metros da China: fenómenos, causas e medidas a tomar por Xuesong et al. [77].	Estudo detalhado em rodas de metro na China e concluiu-se que este tipo de defeito é altamente dependente do tipo de metro e da operação do mesmo. Observou-se que o mesmo tipo de metro num tipo de via desenvolveu desgaste poligonal nas rodas enquanto que noutra tipo de via já não. Isso pode estar relacionado com a forma da via e/ou distribuição do raio da curva.
Investigação sobre o mecanismo de desgaste poligonal de alta ordem das rodas do metro e as suas medidas de mitigação por Tao et al. [8].	Estudou-se o desgaste poligonal assimétrico em rodas de metro. Obteve-se ordens de 12 a 14, sendo este fenómeno mais grave nas rodas do lado direito do que a das rodas do lado esquerdo. Concluiu-se que a primeira ressonância de flexão do conjunto das rodas, causada por vias com curvas acentuadas, foi a causa da poligonização das rodas. O comprimento total das curvas à direita é muito maior do que o das curvas à esquerda, o que desempenha um papel vital na formação do desgaste poligonal assimétrico.
Investigação experimental e numérica sobre a formação de poligonização em rodas de metro por Cai et al. [78].	Através de um dispositivo de medição de desgaste, conseguiu-se verificar que cerca de 70% das rodas medidas, todas com a mesma gama de quilometragem operacional (70.000 a 150.000 km), apresentavam uma ordem de harmónicos dominante entre 6-8.

## 2.5 DESEQUILÍBRIO DE CARGAS

O desequilíbrio de cargas é um problema de instabilidade que afeta especialmente veículos de mercadorias, podendo causar graves problemas tais como: desgaste excessivo na via e no veículo, devido

à pressão desigual entre os eixos, aumento do consumo de energia, devido à resistência adicional exercida no carril, e em casos mais graves o descarrilamento dos veículos, uma vez que o excesso de peso pode levar a um maior risco de instabilidade nas curvas e ao desengate dos vagões. Para mitigar este tipo de problemas, é essencial haver um controlo rigoroso da carga quer por parte do operador ferroviário quer por parte da gestora das infraestruturas, de forma a garantir que a carga seja distribuída uniformemente nos vagões para que nem os eixos nem as rodas sejam sobrecarregados, conforme ilustra a Figura 20.

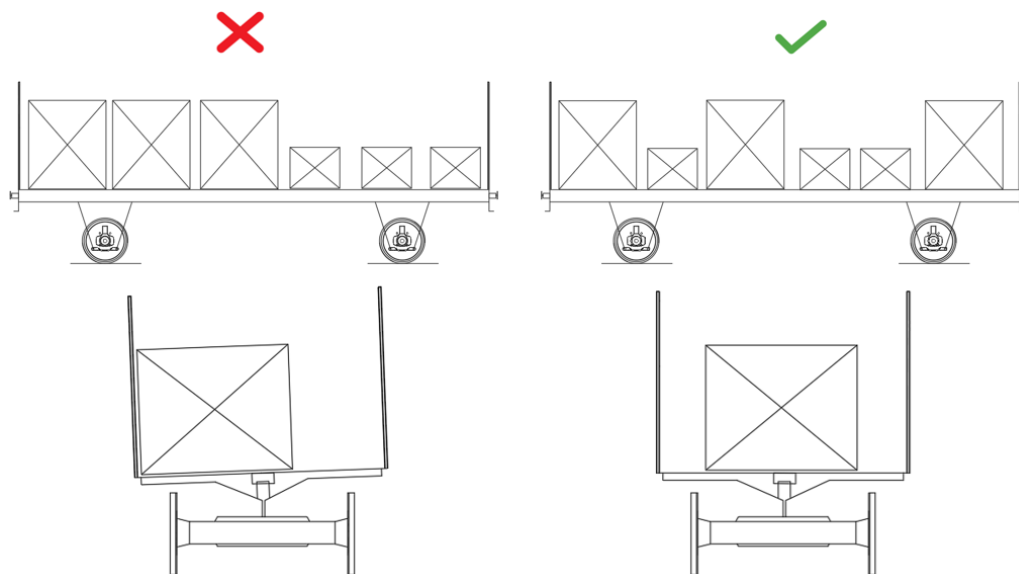


Figura 20 – Esquemas corretos e incorretos de distribuição de cargas.

Geralmente, é consensual que a localização ideal do centro de gravidade da carga deve estar situada lateralmente e longitudinalmente no centro do veículo. Existem diversas normas e diretrizes internacionais que definem as melhores práticas para a distribuição de carga em vagões ferroviários, contudo as mais usuais são as Normas da Associação Internacional de Transporte Ferroviário (UIC).

Apesar de não existirem requisitos universais sobre a distribuição de massa para um vagão carregado, o código UIC [79] apresenta limites para o deslocamento longitudinal e lateral do centro de massa da carga útil.

Na Figura 21 é apresentado um esquema ilustrativo referente aos limites definidos no código UIC para o desequilíbrio de cargas num vagão sem *bogies* com 2 eixos. Na direção longitudinal, a relação entre as cargas por eixo é 2:1, enquanto na direção transversal o rácio a considerar entre carga por roda é 1,25:1.

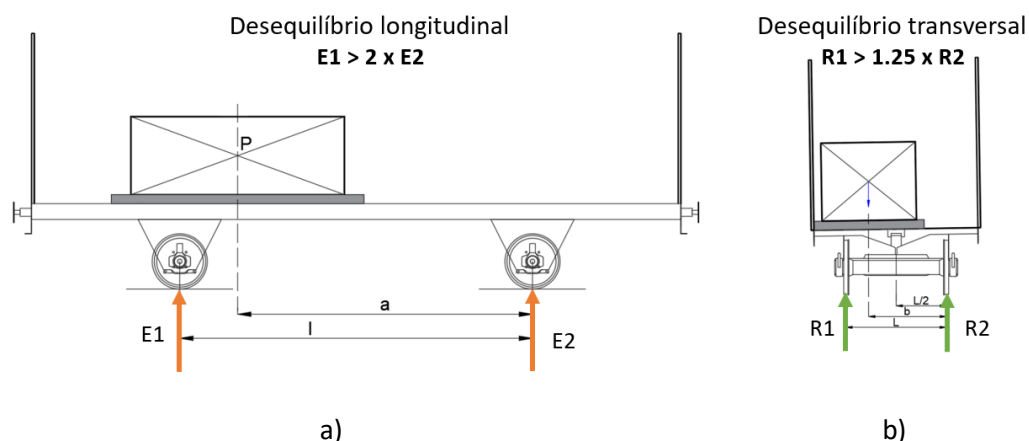


Figura 21 – Desequilíbrio de cargas: a) longitudinal b) transversal (adaptado de [79]).

## 2.6 METODOLOGIAS DE SIMULAÇÃO NUMÉRICA DO SISTEMA DINÂMICO VEÍCULO-ESTRUTURA

“A análise da resposta dinâmica da via devida à passagem de um comboio deve considerar os efeitos de uma estrutura móvel, o comboio, sobre outra estrutura imóvel, a via” [80]. Consoante o objetivo e a complexidade do problema, a escolha da metodologia varia. É de salientar que deverá ter-se em consideração que quanto mais complexa a abordagem, maior será o custo computacional, contudo maior será a precisão.

Das várias metodologias possíveis, destacam-se as seguintes: 1) o método de cargas móveis, 2) o método do caminho virtual e 3) os métodos de interação veículo-estrutura. São de seguida descritas sucintamente cada uma das três metodologias, sendo que estas se encontram descritas de forma pormenorizada em [81].

### 2.6.1 Metodologia numérica baseada em cargas móveis

O método de cargas móveis é baseado na premissa de que a ação do veículo sobre a estrutura pode ser representada por meio de cargas móveis de magnitude fixa, que correspondem à carga estática de cada conjunto de rodas [82, 83]. Dessa forma, o comboio é caracterizado por um conjunto de cargas constantes, representadas por  $F_i$ , separadas entre si de acordo com a geometria do comboio, conforme ilustrado na Figura 22. Nesse método, a interação entre o veículo e a estrutura é desprezada, restringindo-se ao estudo da resposta dinâmica da estrutura.

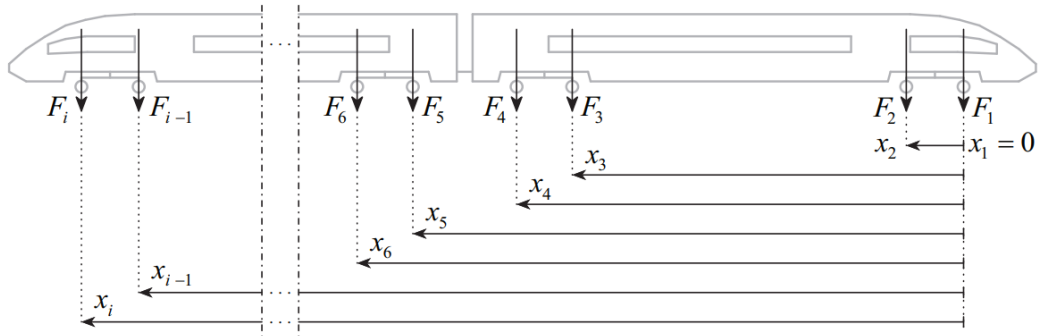


Figura 22 – Metodologia numérica baseada em cargas móveis [81].

A resposta dinâmica da estrutura pode ser calculada segundo a seguinte equação:

$$M \cdot \ddot{a} + C \cdot \dot{a} + K \cdot a = F_{ext} + F_i \tag{2.12}$$

onde:

$M$  – Matriz de massa;

$C$  – Matriz de amortecimento;

$K$  – Matriz de rigidez da estrutura;

$F_{ext}$  – Cargas nodais exteriores;

$F_i$  – Cargas móveis que representam a ação do veículo sobre a estrutura.

### 2.6.2 Metodologia numérica baseada em caminhos virtuais

O método do caminho virtual é uma abordagem desvinculada da interação dinâmica entre o veículo e a estrutura. Permite calcular uma resposta aproximada tanto para o veículo quanto para a estrutura. O método consiste em duas etapas. Na primeira etapa, é calculada a trajetória virtual para as rodas do veículo, sendo calculada a resposta dinâmica da estrutura sob a ação do veículo, usando um modelo de carga móvel. Na segunda etapa, é prescrita a história temporal dos deslocamentos da estrutura nas rodas do veículo, a fim de obter uma resposta desvinculada. As irregularidades da via podem ser adicionadas à trajetória virtual para considerar o seu efeito na resposta do veículo. A Figura 23 ilustra as duas etapas mencionadas anteriormente.

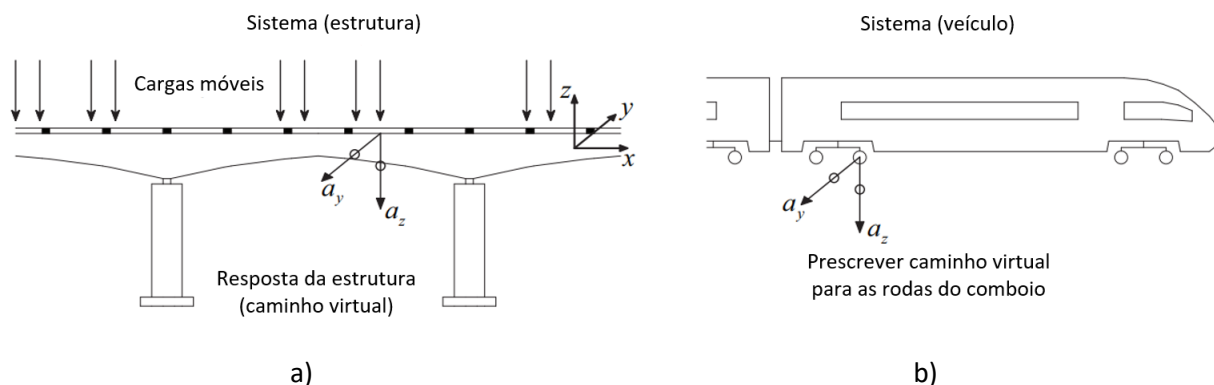


Figura 23 – Metodologia numérica baseada em caminhos virtuais: a) Fase 1 b) Fase 2 [81].

Ao contrário do modelo de carga móvel, este método permite o cálculo, embora simplificado, da resposta do veículo.

### 2.6.3 Metodologia numérica que considera a interação veículo-estrutura

Existem quatro tipos de metodologias que consideram a interação veículo-estrutura: método iterativo, método de condensação, método direto e um método de contacto roda-carril. As metodologias que consideram a interação dinâmica entre estes dois subsistemas são as mais realistas, uma vez que conseguem captar as respostas acopladas de ambos.

Como estas metodologias são as mais próximas de casos reais, Neves et al. [84] optou por estudar de forma pormenorizada a metodologia que adota o método direto e criar uma ferramenta denominada VSI - *Vehicle-Structure Interaction Analysis*, que permite efetuar simulações de interação dinâmica veículo-estrutura.

Como no Capítulo 3 será utilizada a ferramenta VSI, que permite a obtenção das simulações numéricas das passagens de tráfego com dano e sem dano, e esta é baseada no método direto, será dada ênfase apenas a este método.

O método direto é um método que envolve a interação acoplada entre um veículo e uma estrutura, no qual as equações de equilíbrio do veículo e da estrutura são complementadas por equações adicionais que relacionam os deslocamentos dos nós de contacto do veículo com os deslocamentos nodais correspondentes da estrutura [84]. Esse método utiliza o método dos multiplicadores de *Lagrange* na análise do método de elementos finitos (MEF).

A formulação original desse método era restrita à interação vertical, contudo, através do trabalho desenvolvido por Montenegro et al. [81], foi posteriormente expandida para lidar com o contacto tridimensional, levando em consideração a geometria da roda e do carril, bem como as leis não lineares para o contacto normal e tangencial (modelo de contacto roda-carril).

A formulação do contacto roda-carril divide-se em três problemas principais, nomeadamente i) o geométrico, ii) o normal e iii) o tangencial, em que para a simulação do contacto é utilizado um elemento finito de contacto. Através deste elemento, implementado em MATLAB®, é possível avaliar-se a localização do ponto de contacto entre a roda e o carril com base em movimentos relativos entre os dois subsistemas. Este primeiro passo é denominado de problema de contacto geométrico, e é possível de ser resolvido através da parametrização das superfícies dos corpos de contacto, nomeadamente a roda e o carril. A potencial posição do ponto de contacto pode ser avaliada pelas seguintes equações não lineares:

$$\begin{cases} t_r \cdot d_{wr} = 0 \\ t_w \cdot n_r = 0 \end{cases} \quad (2.13)$$

Onde:

$t_r, t_w$  – Vetores tangentes às superfícies de contacto do carril e da roda, respetivamente;

$n_r$  – Vetor normal à superfície de contacto na via;

$d_{wr}$  – Vetor que define a posição relativa entre os pontos de contacto na superfície da roda e no carril.

Podem ocorrer dois tipos de cenários (ocorrência de ponto de contacto e não ocorrência) que resultam da solução válida das equações não lineares acima referidas para a procura de contacto entre a roda e o carril. Na Figura 24, é possível observar esses dois cenários.

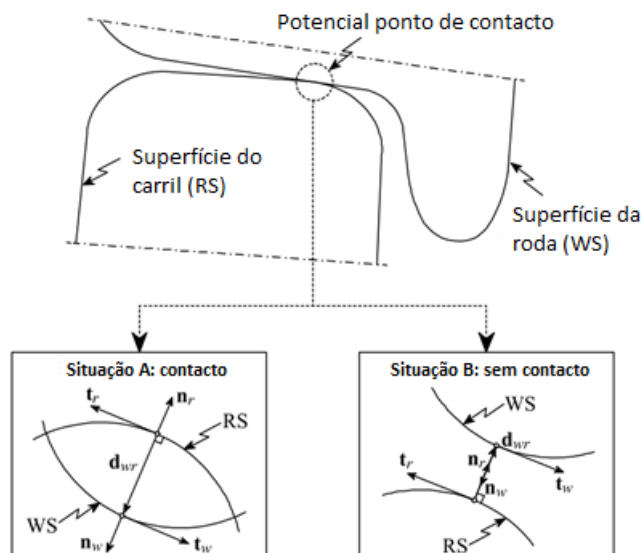


Figura 24 – Possíveis cenários na procura do ponto de contacto [85].

A condição definida pela Equação 2.14 é necessária, ainda assim não suficiente para garantir um contacto efetivo. Para isso é definida uma condição adicional à formulação para garantir que as duas superfícies se intersectam uma à outra:

$$d_{wr}^t \cdot n_r^t \leq 0 \quad (2.14)$$

Para a obtenção do ponto de contacto é feito um processo iterativo. De forma a otimizar o processo, evitando divergências ou convergências muito morosas, são utilizados alguns passos que permitem otimizar a formulação, que podem ser encontrados de forma detalhada em [81].

Em relação ao problema de contacto normal, este é analisado com base na teoria não linear de Hertz. Esta teoria é apenas aplicável a contactos pontuais e lineares, sendo por isso utilizada muito no contacto roda-carril. Resumidamente, quando dois corpos são comprimidos um contra o outro, estes tendem a deformar-se na região em torno do primeiro ponto de contacto, formando uma mancha de contacto com um formato elíptico.

A força normal de contacto  $F_n$  entre a roda e o carril pode ser calculada segundo a seguinte equação:

$$F_n = K_h \cdot d^{3/2} \quad (2.15)$$

Em que:

$K_h$  – Coeficiente generalizado de Hertz;

$d$  – Deformação do elemento de contacto ( $d = \|d_{wr}\|$ ).

Após o cálculo das forças normais, é possível avaliar as forças tangenciais que surgem na interface de contacto devido ao contacto de entre a roda e o carril. Ao contrário do atrito de *Coulomb*, em que o comportamento na zona de contacto é homogéneo (todos os pontos estão a aderir ou a escorregar), quando dois corpos são comprimidos um sobre o outro, a zona de contacto pode partilhar pontos de adesão e de deslizamento em simultâneo. Com base neste facto, é possível obter as forças, forças essas que são relatadas e calculadas em [81].

O acoplamento entre o veículo e o carril é efetuado através do método dos multiplicadores de *Lagrange*, em que as equações de equilíbrio do movimento são complementadas com equações que ligam os dois subsistemas. Estas equações formam um sistema único, cujas incógnitas são deslocamentos e forças de contacto. O sistema matricial pode ser descrito matematicamente como:

$$\begin{bmatrix} \bar{K} & \bar{D} \\ \bar{H} & 0 \end{bmatrix} \begin{bmatrix} \Delta a_F^{i+1} \\ \Delta X^{i+1} \end{bmatrix} = \begin{bmatrix} \psi(a^{t+\Delta t,1}, X^{t+\Delta t,1}) \\ \bar{r} \end{bmatrix} \quad (2.16)$$

Onde:

$\bar{K}$  – Matriz de rigidez efetiva do sistema acoplado;

$\bar{r}$  – Vetor de irregularidade da via;

$\Delta a, \Delta X$  – Deslocamentos incrementais e forças de contacto para mitigar o problema da não linearidade;

$\psi$  – Vetor da força residual;

$\bar{D}, \bar{H}$  – Matrizes de transformação que relacionam as forças de contacto e as deslocações no sistema de coordenadas locais com o sistema de coordenadas globais.

Os sobrescritos  $t + \Delta t$  indicam o instante de tempo atual, enquanto  $i$  e  $i + 1$  referem-se à anterior e atual iteração, respetivamente.

## **2.7 TÉCNICAS BASEADAS EM INTELIGÊNCIA ARTIFICIAL DE IDENTIFICAÇÃO DE DANOS**

As técnicas baseadas em inteligência artificial para a identificação de danos têm ganho destaque na engenharia de estruturas. Estas abordagens inovadoras permitem a deteção precoce de potenciais problemas em estruturas e equipamentos, utilizando algoritmos avançados que permitem evitar a análise de dados por parte de um utilizador, fazendo com que seja uma abordagem muito mais atrativa que as mais tradicionais. Ao analisar padrões e anomalias, as técnicas de IA proporcionam uma resposta rápida e precisa, contribuindo significativamente para a manutenção preditiva e a segurança das infraestruturas. Este campo em constante evolução promete melhorias substanciais na eficiência operacional e na prevenção de danos em diversas aplicações industriais.

### **2.7.1 Levantamento de metodologias desenvolvidas na literatura**

Há uns anos, a inspeção humana era o meio mais comum de deteção de danos. No entanto, esse método é demorado e sujeito a erros humanos, pelo que recentemente têm sido estudadas metodologias que procuram facilitar essa identificação de danos. Como acabam por ser estudos complexos e que requerem muitas análises antes se serem postas em prática, são realizados vários estudos com o intuito de evitar ao máximo eventuais falhas que possam acontecer na identificação destes danos.

Quase todos os estudos que se encontram na área da ferrovia baseiam-se em técnicas lineares (por exemplo o PCA), contudo tornam o estudo mais limitado quando estamos perante fenómenos não lineares. Para isso, procurou-se recorrer ao *autoencoder* que é independente a este tipo de fenómenos. Esta técnica é correntemente utilizada na área da saúde ou até mesmo na área da Engenharia Civil para a deteção de danos em estruturas, contudo na área da ferrovia ainda continua a ser um tema pouco explorado.

Neste contexto, serão apresentadas algumas das metodologias que têm sido estudadas e que serviram de base para o desenvolvimento deste trabalho, juntamente com as conclusões resultantes de cada estudo. Na Tabela 7, Tabela 8 e Tabela 9, serão abordados os lisos nas rodas, a poligonização e o desequilíbrio de cargas, respetivamente.



Tabela 7 – Principais metodologias para detecção de lisos nas rodas e respetivas conclusões.

<i>Metodologias de detecção de lisos nas rodas</i>	<i>Trabalho e conclusões</i>
Método do mecanismo do paralelogramo por Gao et al. [86].	Detecção de lisos nas rodas com base no mecanismo do paralelogramo. Lisos podem ser detetados medindo a mudança no deslocamento vertical do medidor. Experiências de laboratório mostram a eficácia do método.
Detecção baseada em vibração usando análises de espectros por Wang et al. [87].	Detecção de lisos baseada na vibração em veículos de alta velocidade. A análise de espectros de envolvente foi realizada em sinais de vibração da caixa do eixo causados por lisos. Resultados demonstram que a análise de envolventes permite a detecção e a identificação de lisos com alta eficiência e confiabilidade.
Sistema de monitorização <i>wayside</i> usando análise de espectros de envolvente por Mosleh et al [88].	Criação de uma metodologia baseada em sistemas de monitorização <i>wayside</i> que permite a detecção de lisos nas rodas com recurso a espectros de envolvente.
Detecção de lisos com múltiplos registos usando análises espectrais de <i>Kurtosis</i> por Mosleh et al. [89].	Implementação de um layout multissensorial para detecção de lisos em comboios de mercadorias e de passageiros, com o auxílio do espectro de envolvente e o espectro de <i>Kurtosis</i> . Os resultados mostram que o sistema é eficaz na detecção de rodas planas, independentemente da posição dos sensores ou a gravidade do liso.
Detecção e avaliação através de imagens da banda de rodagem das rodas com recurso a uma rede neuronal convolucional por Trilla et al. [90].	Detecção de defeitos na banda de rodagem através de imagens com recurso a redes neuronais convolucionais. Estas localizam as áreas danificadas nas imagens e estimam o tamanho dos defeitos e permitem avaliar o tipo e a gravidade.
Análise avançada de sinais acústicos para detecção de lisos por Komorski et al. [91].	Detecção de lisos usando sinais acústicos com recurso à transformada de Hilbert e à análise de espectros de envolvente.
Detecção precoce de lisos: abordagem automática baseada em <i>wavelet's</i> para a ferrovia por Mosleh et al. [21].	Método para detecção precoce de danos de forma não supervisionada e capaz de distinguir uma roda sem caso de uma com dano. Os resultados mostram que apenas um sensor é capaz de o fazer, permitindo reduzir os custos de instalação e manutenção.
Abordagem não supervisionada para detecção de lisos através de sistemas de monitorização <i>wayside</i> por Mohammadi et al. [20].	Abordagem que permite a detecção de lisos com recurso a técnicas como o AR, ARX, CWT e PCA. Apresentou-se uma metodologia bastante eficaz para a detecção deste tipo de dano.

Tabela 8 – Principais metodologias para detecção de poligonização nas rodas e respetivas conclusões.

<i>Metodologias de detecção de poligonização nas rodas</i>	<i>Trabalho e conclusões</i>
Monitorização de OOR usando PVDF como técnica de detecção por Song et al. [92].	Método de detecção de OOR baseado na deformação de sensor PVDF. Os sensores mediram a resposta da deformação da via causada pela interação veículo-via para deter defeitos na roda. A metodologia foi verificada com simulações em elementos finitos.
Método de detecção de OOR com base na otimização de parâmetros POVMD e LS-SVM por Fang et al. [93].	Deteção de rodas OOR baseada em POVMD e LS-SVM. Os resultados experimentais mostram que a metodologia é eficaz na decomposição do sinal de vibração e analisa a condição da roda de acordo com suas características.
Método de detecção de poligonização na banda de rodagem baseada em respostas dinâmicas por Xu et al. [94].	Método de detecção de poligonização de rodas baseado na transformada de Fourier de tempo curto (SSTFT). Os resultados numéricos demonstram que o desgaste da roda num veículo de alta velocidade pode ser determinado a tempo e garantir a validade dos dados de inspeção.
Método de diagnóstico de OOR em rodas de metro com ruído forte por Huang et al. [95].	Método de detecção de defeitos combinando a técnica KPCA e a DBN. Esta metodologia uma boa precisão na detecção.
Deteção de rodas poligonizadas baseada em sistemas de monitorização <i>wayside</i> e inteligência artificial por Guedes et al. [27].	Metodologia automática que permite a detecção de rodas poligonizadas usando técnicas lineares (CWT e ARX). É feita uma análise de sensibilidade do número de sensores e concluiu-se que 6 sensores é o ideal.

Em relação aos desequilíbrios de carga, a maior parte dos sistemas de detecção propostos na literatura são *onboard*, em que os dispositivos de detecção são montados ou na carroçaria do veículo, ou nos *bogies*, ou nas suspensões. Ainda assim, existem também alguns estudos com base em sistemas *wayside* que se mostram também muito eficazes, contudo importa salientar que nenhum deles utiliza metodologias automáticas não supervisionadas. Basicamente, estes estudos consistem na criação de métodos para obtenção de valores de certas variáveis, como por exemplo extensões, que depois são comparados a valores limite já previamente definidos e estudados.

Pelo contrário, um estudo recentemente realizado por Silva et al. [17] representou um grande avanço neste tema, visto que foi criada uma metodologia de detecção automática de desequilíbrios de carga apenas com recurso a registos da passagem de um comboio e à inteligência artificial. Uma vez que apenas este estudo possibilitou a capacidade de detecção automática de desequilíbrios de carga com base em metodologias automáticas, procurou-se aprofundar a investigação desta questão específica no caso de estudo, de modo a impulsionar o progresso nesta temática.

De forma a sintetizar estes estudos, é apresentada a Tabela 9.

Tabela 9 – Principais metodologias para detecção de desequilíbrios de carga e respectivas conclusões.

<i>Metodologias de detecção de desequilíbrios de carga</i>	<i>Trabalho e conclusões</i>
Detecção de sobrecarga do comboio e de carga desequilibrada com base no medidor FBG por Jianjun et al. [18].	Aplicação de extensómetros de redes de <i>Bragg</i> em fibra ótica (FBG) para a monitorização da sobrecarga e da carga desequilibrada de comboios, comparando os valores dinâmicos da carga por eixo de roda nos dois carris.
Sistema de sobrecarga e desequilíbrio de carga de alta velocidade na China por Ding et al. [11].	Propôs-se um sistema de detecção de sobrecarga e desequilíbrio de carga composto por células de carga que deteta automaticamente o peso da carga, o peso total, a carga desequilibrada à frente e atrás e a carga desequilibrada à esquerda e à direita.
Estudo sobre um método combinado de filtro passa-alto <i>Butterworth</i> e detetor de bordos <i>Canny</i> na aplicação da detecção da posição de carga em veículos ferroviários por Chang et al. [10].	Foi criado um sistema inovador de detecção da posição de carga em veículos baseado num equipamento de amostragem de imagens <i>CCD (Charge-Coupled Device)</i> .
Sistema inovador de pesagem em movimento de alta velocidade para veículos ferroviários por Allotta et al. [96].	Algoritmo baseado em WIM para verificar todos os tipos de condições de carga de uma grande variedade de veículos utilizando extensómetros.
Identificação precoce de cargas desequilibradas de veículos de mercadorias com base na monitorização <i>wayside</i> e na inteligência artificial por Silva et al. [17].	Metodologia para identificação de desequilíbrios longitudinais e transversais de forma automática, com base em modelos ARX e PCA e com auxílio de registos em aceleração e extensão. A técnica PCA com a utilização de 8 acelerómetros + 8 extensómetros mostrou-se a estratégia mais eficaz.

### 2.7.2 Extração de indicadores

A extração de indicadores refere-se ao processo de transformação dos dados medidos numa forma alternativa na qual a correlação com o dano se torna mais facilmente observável [19]. O principal objetivo ao realizar a extração de características é reduzir o número de recursos de um conjunto de dados, criando novos recursos a partir dos existentes, sendo que esse novo conjunto reduzido deverá ser capaz de resumir a maior parte das informações contidas no conjunto original. A condensação de dados é necessária e benéfica, especialmente na monitorização de longo prazo, onde grandes quantidades de dados têm que ser comparados e analisados. Esta redução permitirá acelerar o processamento dos dados, melhorar a precisão dos resultados e reduzir o risco de *Overfitting* (modelo estatístico ajusta-se demasiado ao conjunto de dados).

O processo de extração de indicadores sensíveis a dano pode ser definido como a identificação de determinados parâmetros da resposta estrutural que possibilitam a distinção entre a condição estrutural danificada e não danificada. Uma característica sensível ao dano é alguma grandeza extraída dos dados de resposta medidos do sistema, que possui correlação com a presença de dano numa estrutura.

Idealmente, uma característica sensível ao dano deve apresentar variações consistentes à medida que o nível de dano aumenta.

Com o intuito de tornar esta detecção de dano mais confiável e eficaz, alguns autores referem que os resultados da extração de indicadores podem ser aprimorados se forem usadas técnicas adequadas, das quais se destaca a análise de componentes principais (PCA), os modelos autorregressivos (AR), os modelos autorregressivos exógenos (ARX), as transformadas contínuas *wavelet* (CWT) ou técnicas mais recentes, e que tem dado de falar, as redes neuronais artificiais (RNA). Estas são exemplos de técnicas eficazes de extração de indicadores sensíveis a dano tanto para uma monitorização estática ou dinâmica, contudo existem muitas mais que têm sido estudadas e melhoradas nestas últimas décadas.

Nos subcapítulos seguintes serão abordadas algumas destas técnicas. Técnicas com o PCA, o AR, o ARX e o CWT já foram previamente utilizadas em trabalhos realizados no projeto WAY4SAFERAIL, contudo o *autoencoder* não. Como é uma novidade no projeto e foi a técnica mais utilizada ao longo do caso de estudo, será dada uma especial atenção no esclarecimento desta técnica.

### 2.7.2.1 Análise das componentes principais (PCA)

A ciência dos dados tem utilizado frequentemente a análise de componentes principais (PCA) para extrair características baseadas na projeção de dados para um novo subespaço sem dimensão [20], através de uma matriz de covariância e uma transformação linear em componentes principais usando vetores próprios (*Eigenvectors*). Considerando uma matriz de dados de entrada  $X$  ( $n \times m$ ), os componentes principais ou scores podem ser calculados com base na seguinte equação:

$$PC = X \cdot T \quad (2.17)$$

Onde:

$PC$  – Matriz ( $n \times m$ ) dos componentes principais;

$X$  – Matriz ( $n \times m$ ) dos indicadores extraídos;

$T$  – Matriz ( $m \times m$ ) de transformação orto-normal linear, sendo cada coluna um *Eigenvector* de  $X$ .

Antes da aplicação do método PCA, deverá ser realizada uma padronização da matriz dos dados de entrada, subtraindo a média e dividindo pelo desvio padrão para cada valor de cada coluna. É importante realizar a padronização antes do PCA, pois esta técnica é bastante sensível à variação das variáveis iniciais.

Após o cálculo das componentes principais, quatro variáveis estatísticas são extraídas das pontuações do PCA, nomeadamente, a média quadrada (RMS), o desvio padrão (SD), a assimetria e o *kurtosis*. Assim, com este método, o número de características extraídas é reduzido apenas a 4.

### 2.7.2.2 Modelos autorregressivos (AR e ARX)

Os modelos AR e ARX para extração de características são aplicados, pois conseguem realizar uma fusão significativa, generalizando precisamente as informações contidas nos dados. Estes métodos de análise de séries temporais consideram o comportamento preditivo do sistema numa localização específica, levando em conta o histórico nesse mesmo ponto de medição, e no caso específico do ARX, ainda a resposta preditiva em outras posições de medição perante o mesmo evento - entradas exógenas.

Os modelos AR (equação 2.18) e ARX (equação 2.19) podem ser definidos pelas equações abaixo:

$$x_j = \sum_{i=1}^{n_a} a_i x_{j-1} + \varepsilon_j \quad (2.18)$$

$$x_j = \sum_{i=1}^{n_a} a_i x_{j-1} + \sum_{k=1}^{n_b} b_k x_{j-k} + \varepsilon_j \quad (2.19)$$

Onde:

$x_j$  – Valor atualizado da resposta;

$n_a$  – Ordem das entradas;

$a_i$  – Parâmetro constante da entrada;

$x_{j-1}$  – Resposta anterior;

$\varepsilon_j$  – Erro aleatório não observável na resposta  $j$ ;

$n_b$  – Ordem das saídas;

e  $b_k$  – Parâmetro constante da saída.

De acordo com Meixedo et al. [25], as ordens do modelo AR/ARX devem ser determinadas em função das características dos dados em estudo, sendo uma questão chave no processo. Para isso, com o intuito de otimizar o valor da ordem do modelo, é frequente utilizar-se o Critério de Informação *Akaike* (*AIC*), que permite avaliar a qualidade de adequação de um modelo estatístico estimado baseando-se no compromisso entre a precisão de adaptação e o número de parâmetros estimados. No contexto dos modelos AR/ARX, o *AIC* é dado por:

$$AIC = N_t \ln(\varepsilon) + 2N_p \quad (2.20)$$

Onde:

$N_p$  – Número de parâmetros estimados;

$N_t$  – Número de pontos de dados previstos;

$\varepsilon$  – Erros da soma média dos resíduos quadráticos (SSR), em que  $\varepsilon = SSR/N_t$ .

Sendo assim, para se obter o valor da ordem ideal deverá ensaiar-se uma vasta gama de ordens e escolher um número em que o valor de *AIC* seja mínimo.

### 2.7.2.3 Transformadas Contínuas *Wavelet* (CWT)

As transformadas contínuas *wavelet* são uma técnica de processamento de sinais que permitem analisar e representar sinais de forma eficiente em diferentes escalas de tempo e frequência, pelo que têm sido amplamente utilizadas em variadas áreas, na qual se enquadra a engenharia.

Embora as transformadas de Fourier geralmente usem ondas sinusoidais como base para a decomposição, outras funções podem ser selecionadas de forma adaptar a forma da *wavelet* às propriedades do sinal. Dois parâmetros definem a análise de *wavelet*: escala e translação. Em termos matemáticos, o CWT ( $T$ ) correspondente ao sinal  $x(t)$  é normalmente definido da seguinte forma [97]:

$$T(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi^* \left( \frac{t-b}{a} \right) dt \quad (2.21)$$

Onde:

$a, b$  – Parâmetros de escala e translação, respetivamente;

$\psi^*$  – Complexo conjugado da onda mãe  $\psi$ .

A escolha da onda mãe depende das propriedades do sinal em análise e das características que se deseja extrair, pelo que é crucial esta escolha para se obter uma representação precisa e eficiente do sinal de interesse. Como são calculados coeficientes *wavelet* para cada ponto, através da implementação de um CWT, são obtidos inúmeros dados. Com o intuito de realizar uma compressão e tornar o processo mais rápido, é correntemente utilizado o PCA, já abordado anteriormente, extraindo-se quatro parâmetros estatísticos. Considerando uma matriz  $X$  de  $n \times m$  elementos, em que são considerados os indicadores CWT extraídos do sinal original relativamente a cada sensor, onde  $n$  corresponde ao número de pontos de medição em cada sensor e  $m$  representa o número de coeficientes *wavelet*, os componentes principais podem ser calculados com base na seguinte equação:

$$S = X \cdot T \quad (2.22)$$

Onde:

$S$  – Matriz,  $n \times m$ , das componentes principais;

$X$  – Matriz,  $n \times m$ , dos indicadores CWT;

$T$  – Matriz de transformação linear ortonormal,  $m \times m$ .

### 2.7.2.4 Redes neurais artificiais (RNA)

O estudo das redes neurais ganhou destaque a partir do momento em que se começou a falar de inteligência artificial (IA). A IA é uma subárea da ciência da computação que se refere à capacidade de uma máquina imitar e atuar como um ser humano em tarefas que são geralmente realizadas por seres humanos.

Dentro do amplo domínio da IA, surge um subgrupo específico, o *machine learning*, que envolve a execução de algoritmos, que de forma automática, criam modelos de representação de conhecimento com base num conjunto de dados. A ideia fundamental por detrás deste processo é capacitar as máquinas através do acesso a dados históricos, permitindo que o algoritmo aprenda e se consiga posteriormente ajustar de maneira iterativa, permitindo efetuar previsões de qualidade em situações futuras e que estejam relacionadas com padrões históricos. Por fim, ainda dentro deste último grupo, surge o *deep learning*, que inclui as redes neurais artificiais, as quais serão estudadas mais detalhadamente de seguida.

De forma resumir as principais diferenças entre os vários domínios, é apresentada a Figura 25.



Figura 25 – Diferenças entre Inteligência Artificial, *Machine Learning* e *Deep Learning* (adaptado de [98]).

No grupo do *deep learning* surgem as redes neurais artificiais. Estas são sistemas de computação com nós interconectados que funcionam como os neurónios do cérebro humano [99]. Usando algoritmos, estas permitem reconhecer padrões escondidos e correlações em dados brutos, agrupá-los e classificá-los, e com o tempo aprender e melhorar continuamente. Estas redes podem aprender e modelar relações entre entradas e saídas de dados que são não lineares e complexas, realizar generalizações e interferências, revelar padrões e predições ocultas, e ainda modelar dados altamente voláteis para prever eventos raros.

Existem vários tipos de redes neurais que foram desenvolvidas ao longo do tempo com o objetivo de responder a problemas específicos. Portanto, a arquitetura da rede e o funcionamento variam de acordo

com o tipo de rede em questão. Na Figura 26, são apresentadas algumas das inúmeras redes neuronais que foram desenvolvidas e estudadas, desde as mais simples até às mais complexas.

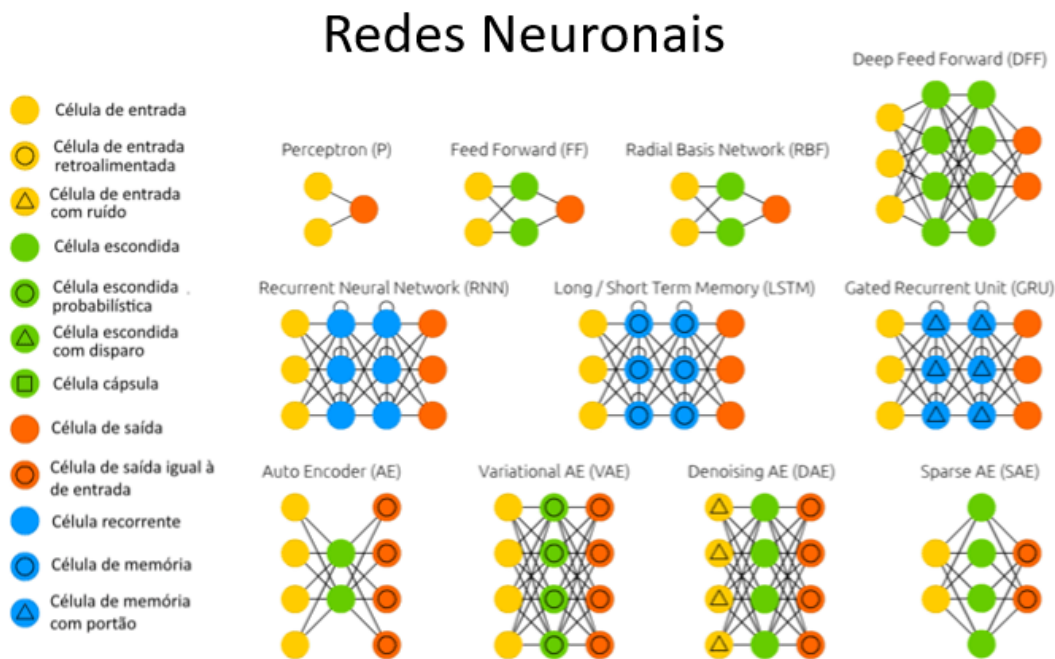


Figura 26 – Exemplos de alguns tipos de redes neuronais [100].

Para dados que apresentam uma complexidade mais elevada, é mais frequente a utilização dos *autoencoders*. Conforme é visível na Figura 26, existem vários tipos de *autoencoders*, dos quais se destacam o *autoencoder* tradicional (AE), o *autoencoder* variacional (VAE), o *denoising autoencoder* (DAE) e o *autoencoder* esparso (SAE).

Basicamente, os *autoencoders* (AE) são redes neurais artificiais, treinadas de maneira não supervisionada, que visam primeiro aprender representações codificadas dos dados, compactando a entrada numa representação de espaço latente, mais conhecida por “*bottleneck layer*”, e em seguida, através desta informação comprimida, apenas com a informação essencial, estes reconstróem a informação [99]. Assim, a saída de um *autoencoder* é uma previsão para os dados de entrada.

Esse tipo de rede é composto por três partes [99]:

- Codificador (*Encoder*): é a parte da rede que compacta os dados numa representação de espaço latente (codificando a entrada);
- Espaço latente: Espaço intermédio da rede que contém as representações de conhecimento compactadas e, portanto, a parte mais importante da rede;
- Decodificador (*Decoder*): Esta parte tem como objetivo reconstruir a entrada da representação do espaço latente.



Pela análise da Figura 27, é possível obter uma visão geral do conceito principal do funcionamento de um *autoencoder*.

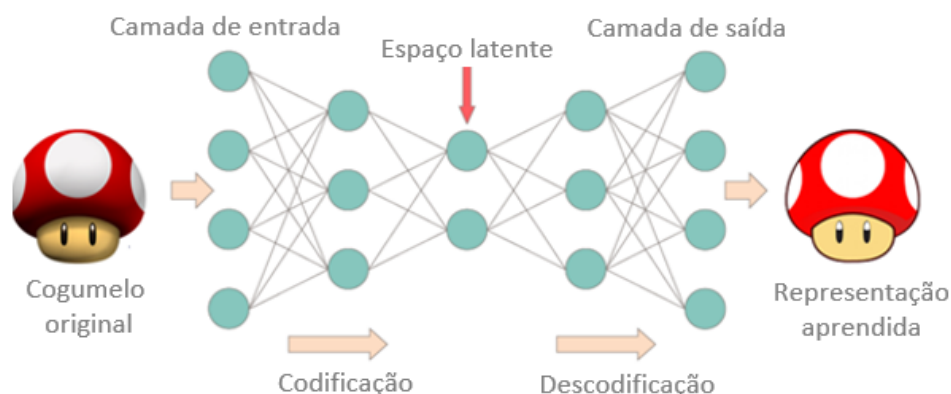


Figura 27 – Esquema ilustrativo da arquitetura base de um *autoencoder* (adaptado de [99]).

À semelhança de todas as redes neurais, também o *autoencoder* é constituído por nós ou também conhecidos por “neurónios”, que formam as camadas constituintes da rede. A condição mínima para se ter uma rede neuronal é que esta seja constituída, no mínimo, por três camadas: uma camada de entrada, uma camada oculta e uma camada de saída. No caso de se ter redes com mais que três camadas (entrada, oculta e saída), estas são chamadas de redes neurais profundas (*deep neural network*).

Inspirados no neurónio biológico, desenvolveu-se um modelo de neurónio matemático/artificial que é a base da inteligência artificial. A imagem do neurónio biológico, também o neurónio artificial (Figura 28) é a unidade básica da rede neuronal, sendo uma célula especializada na transmissão de informações [99]. Cada neurónio artificial presente na rede possui um ou mais sinais de entrada e apenas um sinal de saída que pode ser distribuído como sinal de saída da rede, ou como sinal de entrada para um ou vários outros neurónios da camada posterior (que formam a rede neural artificial). Cada sinal de entrada recebe um valor  $(x_1, \dots, x_n)$ , e estes são ponderados pelos pesos  $(w_1, \dots, w_n)$ , combinados por uma função que equivale ao processamento realizado pela soma ( $\Sigma$ ), acrescido do valor do bias, ou viés,  $(b)$ . Tanto os valores dos pesos como dos bias são inicializados de forma aleatória e vão sendo atualizados conforme o treino e os dados do problema.

Resumidamente, o papel dos pesos é determinar o impacto que o valor de cada variável de entrada tem no valor das variáveis de saída, ou seja, quanto maior for o valor do peso desse neurónio maior será a influência deste neurónio nos nós seguintes. Assim como no modelo biológico, o estímulo pode ser excitativo ou inibidor, representado pelo peso positivo ou negativo, respetivamente.

Já em relação ao bias, é um parâmetro que representa uma conexão extra no neurónio e estabelece uma certa tendência ou inclinação no processo computacional, ajudando o modelo a adaptar-se aos dados fornecidos [101].

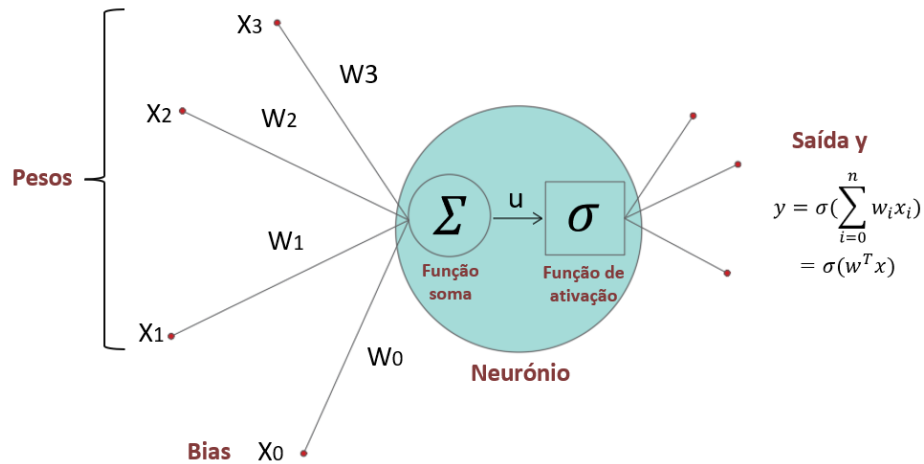


Figura 28 – Funcionamento de um neurónio artificial (adaptado de [102]).

Neste seguimento, imagine-se que é fornecida à rede um conjunto de dados que representam, por exemplo, ou um objeto, ou um número ou um sinal dinâmico obtido através da leitura de um sensor. Quando as entradas são apresentadas ao neurónio, estas são multiplicadas pelos pesos sinápticos correspondentes, gerando as entradas ponderadas, ou seja,  $x_1$  que multiplica  $w_1$ , e assim sucessivamente para cada dado de entrada  $x_n$ , através da continua multiplicação de matrizes. Basicamente, a entrada total recebida por um neurónio, chamada de potencial de ativação ( $u$ ) para uma entrada  $X$ , pode ser definida pela seguinte equação:

$$u = \sum_{i=1}^n x_i w_{ij} + b_j \tag{2.23}$$

Onde:

$u$  – Potencial de ativação;

$x_i$  – Entrada;

$w_{ij}$  – Peso referente à entrada  $i$  e ao neurónio  $j$ ;

$b_j$  – Bias.

A saída  $y$  de um neurónio é definida por meio da aplicação de uma função de ativação  $\sigma$  à entrada total [103]. Existem várias funções de ativação disponíveis, sendo que são estas as responsáveis por decidir se um neurónio deve ser ativado ou não, ou seja, se a informação que o neurónio está a receber é relevante para a informação fornecida ou se deve ser ignorada. Sendo assim, a saída por ser definida por:

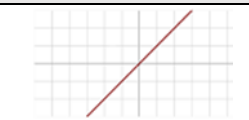
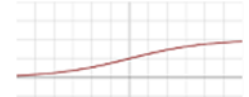
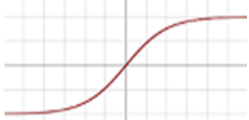
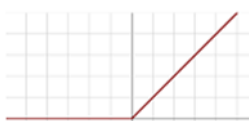
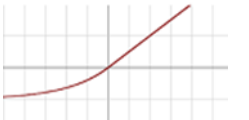
$$y = \sigma\left(\sum_{i=1}^n x_i w_{ij} + b_j\right) \tag{2.24}$$

$\sigma$  – Função de ativação.

A principal razão pelas quais as funções de ativação ou transferência são utilizadas, é que estas conferem a capacidade de não linearidade ao processamento realizado pelas redes neuronais, sendo bastante importante especialmente nas camadas escondidas da rede, fazendo com que possam aprender relações complexas quer entre variáveis dependentes quer independentes. Caso não seja utilizada a função de ativação, os pesos e bias fazem uma transformação linear.

Na Tabela 10, são apresentados alguns exemplos de funções de ativação, bem como algumas das suas características que são importantes para a escolha do tipo de função para cada problema.

Tabela 10 – Funções de ativação e respetivas características.

Tipo de função de ativação	Esquema	Equação	Derivada
Linear		$f(x) = ax + c$	$f'(x) = a$
Sigmóide		$f(x) = \frac{1}{1 + e^{-x}}$	$f'(x) = f(x)[1 - f(x)]$
Tangente hiperbólica (TanH)		$f(x) = \tanh(x)$	$f'(x) = 1 - f(x)^2$
Unidade linear retificada (ReLU)		$f(x) = \begin{cases} 0 & \text{para } x < 0 \\ x & \text{para } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{para } x < 0 \\ 1 & \text{para } x \geq 0 \end{cases}$
Unidade linear exponencial (ELU)		$f(x) = \begin{cases} \alpha(e^x - 1) & \text{para } x < 0 \\ x & \text{para } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} f(x) + \alpha & \text{para } x < 0 \\ 1 & \text{para } x \geq 0 \end{cases}$

Dependendo dos resultados que se pretende obter deverá adequar-se a função de ativação. Por exemplo, para problemas de regressão linear é comum utilizar-se a função linear, enquanto para problemas de classificação binária é usada a função sigmóide e problemas de classificação não binária o softmax.

O processo descrito na Equação 2.23 e 2.24 é aplicado a cada neurónio presente na rede neuronal. Uma vez calculados todos os valores de saída dos neurónios, o neurónio que tiver o maior valor será o mais ativado, indicando que a resposta para o problema é equivalente ao resultado que este neurónio de saída representa. A precisão desta rede neuronal será definida pelo ajuste dos pesos e bias de cada neurónio, de acordo com as equações que se mostram mais abaixo. Com base na arquitetura da rede representada

na Figura 29, são agora apresentados os fundamentos matemáticos de uma rede neuronal, os quais são aplicáveis a todas as redes neurais, independentemente do seu tipo.

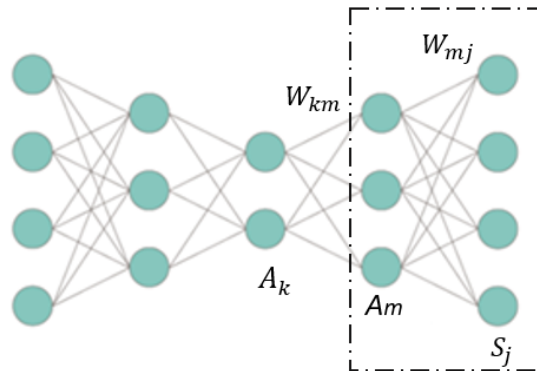


Figura 29 – Identificação das camadas e ligações da rede.

Com vista na aplicação da equação 2.25 e considerando duas camadas consecutivas, nomeadamente a camada \$A\_m\$ e \$S\_j\$, podemos estabelecer-las na forma matricial:

$$S_j = \sigma\left(\sum_{i=1}^n W_{mj}A_m + b_j\right) \quad (2.25)$$

$$S = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_n \end{bmatrix} \quad W_{mj} = \begin{bmatrix} W_{11} & \cdots & W_{1m} \\ \vdots & \ddots & \vdots \\ W_{j1} & \cdots & W_{jm} \end{bmatrix} \quad A = \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_j \end{bmatrix} \quad (2.26)$$

Onde:

\$S\_j\$ – Camada de saída;

\$W\_{mj}\$ – Pesos;

\$A\_m\$ – Camada intermédia;

\$b\_j\$ – Bias da camada de saída.

A atualização dos pesos e dos bias da rede faz-se através da minimização do erro dado por uma função de custo (\$C\$). É possível adotar-se várias funções de custo, como por exemplo, o erro da média absoluta, o erro da soma absoluta, o erro quadrático médio, entre outros, dependendo do tipo de problema. A título de exemplo é adotado o erro quadrático médio que pode ser definido da seguinte forma:

$$C = \frac{1}{N} \sum_{j=1}^N (S_j - y_j)^2 \quad (2.27)$$

Onde:

\$N\$ – Quantidade de inputs;

$C$  – Função de custo;

$S_j$  – Valor de saída da rede (previsão);

$y_j$  – Valor de referência/real.

Como o objetivo é diminuir o valor de custo, ou seja, obter o mínimo da função, é comum recorrer-se ao gradiente descendente, que é um método que permite encontrar, de forma iterativa, os valores dos parâmetros que minimizam determinada função de interesse. Por exemplo, caso se pretenda obter o mínimo da função representada na Figura 30, é necessário recorrer à derivada parcial. Assim, será possível avaliar a taxa de variação da função de custo em relação ao peso  $W_{mj}$ , através dos ângulos das retas tangentes.

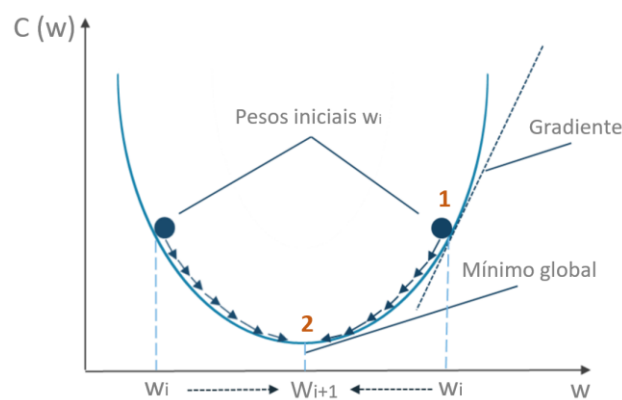


Figura 30 – Gradiente da função de custo, adaptado de [104].

Sendo assim, a atualização dos pesos é realizada em todos os neurónios da seguinte forma:

$$W_{mj(i)} = W_{mj(i-1)} - \mathcal{L} \cdot \frac{\partial C}{\partial W_{mj}} \quad (2.28)$$

Onde:

$W_{mj(i)}$  – Valor do peso atualizado, iteração  $i$ ;

$W_{mj(i-1)}$  – Valor do peso antes de ser atualizado;

$\mathcal{L}$  – Constante da taxa de aprendizagem.

Inicialmente, na primeira iteração, temos um peso inicial aleatório,  $W_{mj(i-1)}$  (1). Na iteração seguinte, de modo a obter-se um peso atualizado,  $W_{mj(i)}$  (2), este valor de peso vai ser reduzido por um valor proporcional à derivada parcial da função de custo em relação a este peso  $\frac{\partial C}{\partial W_{mj}}$ , ou seja, quanto maior for esta derivada, maior será esta redução do peso. Pelo contrário, se o valor do custo estiver já próximo do mínimo, esta atualização já não será tão significativa.

A parcela da derivada parcial  $\frac{\partial C}{\partial W_{mj}}$  pode ser obtida da seguinte forma:

$$\frac{\partial C}{\partial W_{mj}} = 2 \cdot \sum_j (S_j - y_j) \cdot \sigma' \left( \sum_m W_{mj} \cdot A_m + b_j \right) \cdot \sum_m A_m \tag{2.29}$$

É possível observar que na Equação 2.29, esta parcela depende do erro entre a previsão e o real ( $S_j - y_j$ ) e dos pesos e dos neurónios da camada anterior.

À semelhança dos pesos, a atualização dos bias é realizada da mesma forma.


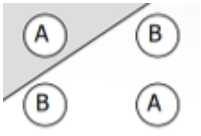
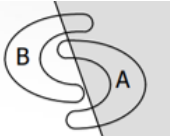


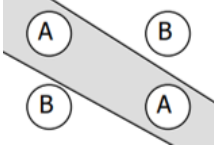
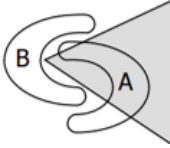
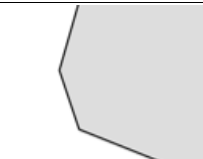

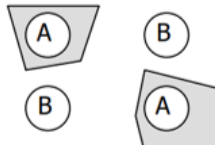
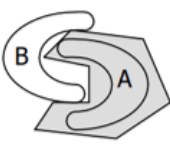
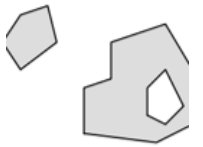
$$b_{j(i)} = b_{j(i-1)} - \mathcal{L} \cdot \frac{\partial C}{\partial b_j} \tag{2.30}$$

$$\frac{\partial C}{\partial b_j} = 2 \cdot \sum_j (S_j - y_j) \cdot \sigma' \left( \sum_m W_{mj} \cdot A_m + b_j \right) \tag{2.31}$$

Como o objetivo desta explicação era apenas dar uma breve introdução sobre o funcionamento da rede neuronal e não fazer uma explicação matemática exaustiva, importa referir que apenas foi apresentada a dedução das fórmulas relativas à camada mais próxima da camada de saída, contudo este processo teria de ser realizado para todas as camadas da rede. O resto das deduções podem ser encontradas em [105].

Como é possível compreender, as redes neuronais são bastante potentes, contudo a sua complexidade pode ser melhorada conforme o número de camadas que esta apresenta. Em problemas de classificação, de que é exemplo o caso de estudo, o número de camadas pode ser preponderante na melhor ou pior classificação. Através da Tabela 11, é evidente que uma rede mais completa resulta numa classificação mais eficaz e próxima da perfeição, ao passo que uma única camada não consegue realizar a classificação de forma satisfatória.

Tabela 11 – Regiões de decisão num problema de classificação (adaptado de [102]).

Estrutura	Classes com regiões separadas	Classes com regiões misturadas	Formas de região mais gerais
<p>Uma camada</p> 			
<p>Duas camadas</p> 			
<p>Três camadas</p> 			

### 2.7.2.4.1 Autoencoder esparso (SAE)

No contexto desta tese, foi utilizado um *autoencoder* do tipo esparso (SAE), uma vez que é uma das redes neuronais mais abordadas em trabalhos de identificação automática de danos. A principal diferença deste tipo de *autoencoder* em relação ao *autoencoder* tradicional atrás apresentado, é que o SAE introduz restrições através da utilização de parâmetros de esparsidade, permitindo que apenas um subconjunto de neurónios seja ativado. Ao aplicar esta restrição, é possível tornar a codificação mais eficiente e melhorar a capacidade de generalização, uma vez que apenas as características mais importantes são mantidas [106].

Para que isso aconteça, são adicionados dois termos à função de custo ( $C$ ), um termo de penalização esparso ( $\Omega_{\text{esparsidade}}$ ), que permite limitar o valor médio de ativação do neurónio da camada oculta [31], e um termo de regularização de pesos ( $\Omega_{\text{pesos}}$ ), denominado regularização L2, com o propósito de evitar o sobreajuste da rede [31].

Sendo assim, a função de custo aplicada ao SAE, pode ser obtida pela seguinte equação:

$$C = \frac{1}{N} \sum_{n=1}^N \sum_{k=1}^K (S_j - y_j)^2 + \lambda \cdot \frac{1}{2} \sum_{l=1}^L \sum_{j=1}^{n_l} \sum_{i=1}^{k_l} (w_{ji}^l)^2 + \beta \cdot \sum_{k=1}^K KL(\rho \parallel \hat{\rho}_i) \quad (2.32)$$

Onde:

$\lambda$  – Coeficiente para o termo de regularização L2;

$L$  – Número de camadas ocultas;

$k_l, n_l$  – Tamanho da camada  $l$  de entrada e de saída, respetivamente;

$\beta$  – Coeficiente para o termo de regularização de esparsidade;

$KL$  – Convergência de *Kullback-Leibler*;

$\rho, \hat{\rho}_i$  – Valor desejado e valor médio de ativação, respetivamente.

Como no caso de estudo os dados utilizados apresentavam relações complexas entre as características, foi necessário recorrer-se a um empilhamento de dois SAE's, conforme já foi sendo utilizado em alguns trabalhos [107]. Geralmente, a arquitetura de um SSAE envolve o treino de cada *autoencoder* individualmente em camadas sucessivas, com as saídas do *autoencoder* anterior a servir de entrada para o próximo, conforme se apresenta na Figura 31.

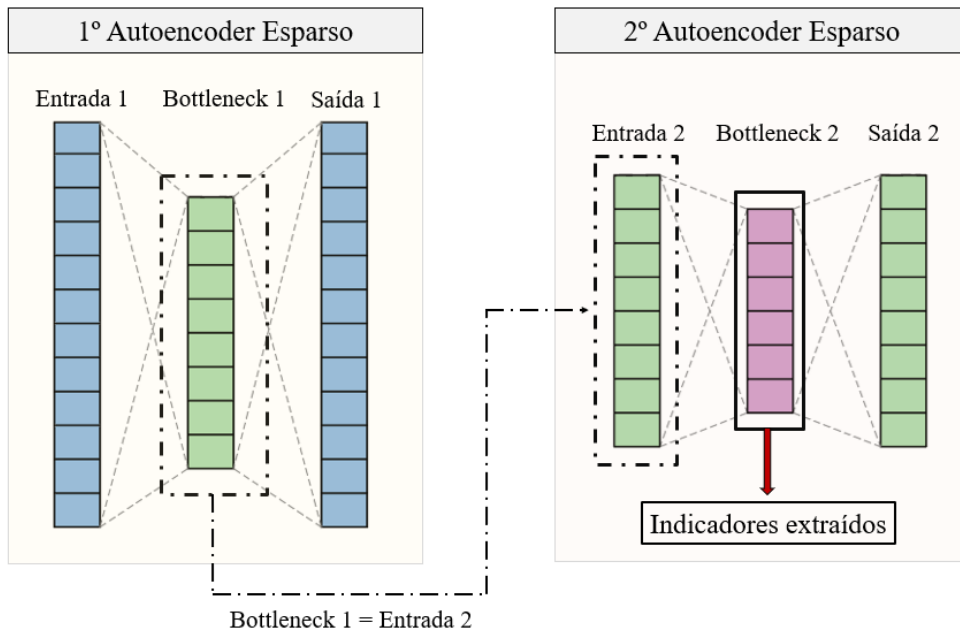


Figura 31 – Arquitetura adotada no SSAE utilizado no caso de estudo.

De forma muito sucinta, após os vários *inputs* serem incluídos no treino do SSAE, estes são incorporados na chamada “Entrada 1”. À medida que é realizado o treino, é efetuada uma compressão dos dados e são obtidas as características principais dos sinais, ficando retidos no “Bottleneck 1”. Por sua vez, estas características são utilizadas como entrada no treino do 2º autoencoder esparso, que após nova compressão possibilita a obtenção dos indicadores finais que são incorporados depois na etapa da fusão de dados.

Dependendo do tipo de defeito em análise, o nível de compressão varia, razão pela qual foram realizados vários testes com o intuito de otimizar esses valores. Para além dos tamanhos das camadas, é ainda comum definir-se outros hiperparâmetros para o treino do SSAE, que são eles o número de épocas, as funções transferência de codificação e de decodificação, o termo de regularização L2 ( $\lambda$ ), termo de regularização de esparsidade ( $\beta$ ) e a proporção da esparsidade, conforme foi realizado no Capítulo 4 e 5.

Na Tabela 12, são apresentados os vários hiperparâmetros adotados no SSAE do caso de estudo, juntamente com uma breve explicação de cada um deles.



Tabela 12 – Hiperparâmetros utilizados para o treino no SSAE referentes ao caso de estudo.

<i>Hiperparâmetro</i>	<i>O que é?</i>	<i>Opção adotada</i>	<i>Influência</i>
Função de transferência de codificação	Refere-se à função de transferência utilizada na camada de codificação do <i>autoencoder</i>	"logsig"	Função não linear
Função de transferência de decodificação	Refere-se a função de transferência utilizada na camada de decodificação do <i>autoencoder</i>	"logsig"	Função não linear
Tamanho da camada oculta	Refere-se à dimensão do vetor que representa a informação oculta em cada célula da rede	Calibrado conforme o tipo de defeito e da metodologia adotada	Maior valor permite reter mais informações complexas. Por outro lado, um menor valor pode limitar a capacidade do modelo.
Número máximo de épocas	Refere-se ao número de vezes que todo o conjunto de treino passa pela rede durante o processo de treino	Calibrado conforme o tipo de defeito e da metodologia adotada	Mais épocas permite melhorar a convergência e refinar os pesos, melhorando a capacidade de aprendizagem das características dos dados.
Regularização de pesos L2	É um termo adicionado à função de custo durante o treino para evitar o sobreajustamento dos pesos da rede	Calibrado conforme o tipo de defeito e da metodologia adotada	Maior valor incentiva menores valores dos pesos
Regularização de esparsidade	É um termo adicionado à função de custo para incentivar a esparsidade na camada de codificação	Calibrado conforme o tipo de defeito e da metodologia adotada	Maior valor incentiva a ativação de apenas uma pequena fração de neurónios da camada oculta
Proporção de esparsidade	Refere-se à taxa de esparsidade desejada	Calibrado conforme o tipo de defeito e da metodologia adotada	Maior valor incentiva uma codificação mais esparsa
Função de custo/perda	Utilizada para avaliar a qualidade da reconstrução dos dados de entrada	"MSEPARSE"	Tipo de função pode ser melhor ou pior consoante o tipo de dados
Algoritmo de treino	O algoritmo a ser usado para treinar o <i>autoencoder</i>	"Trainscg" ( <i>Scaled conjugate gradient backpropagation</i> )	Tipo de algoritmo de treino pode ser melhor ou pior consoante o tipo de dados

### 2.7.3 Fusão dos dados

A fusão de dados é o processo de integração e combinação de dados de múltiplas fontes para produzir informações mais exatas, completas e consistentes do que seria possível utilizando apenas fontes individuais [108]. Existem vários métodos para a fusão de dados, incluindo sistemas baseados em regras, modelos estatísticos (Distância de *Mahalanobis*) e técnicas de aprendizagem automática, como redes

neuronal e árvores de decisão, sendo que a escolha do método depende da aplicação específica e das características das fontes de dados envolvidas.

Em termos de deteção e classificação automática de danos, tem sido bastante frequente a utilização da Distância de *Mahalanobis* (MD) para a fusão, visto que é uma técnica relativamente simples e que permite que dados multivariados sejam reduzidos a um único indicador de dano (DI), reduzindo o volume dos dados extraídos preservando a informação mais relevante [17, 20, 26, 27].

Em termos práticos, esta calcula a distância entre os indicadores de dano e sem dano por forma a definir semelhanças entre eles, permitindo que curtas distâncias representem grandes diferenças. Dado que se trata de uma distância bastante genérica, ela é capaz de detetar qualquer cenário de dano, enquanto fornece uma ponderação não supervisionada, sendo assim independente de intervenções humanas, do tipo da estrutura/sistema e das ações atuantes na mesma [48]. Esta distância, designada aqui como *DI*, é calculada para cada sensor e para cada simulação *i*, pela seguinte expressão:

$$DI = \sqrt{(x - \bar{z})^T C^{-1} (x - \bar{z})} \quad (2.33)$$

onde:

$x$  – Vetor de  $m$  indicadores representativos de potenciais danos;

$\bar{z}$  – Matriz da média dos indicadores estimados no cenário de base;

$S_x$  – Matriz da covariância das simulações dos cenários de base;

$C^{-1}$  – Matriz de covariância inversa das simulações de base.

#### 2.7.4 Discriminação de indicadores

A discriminação de indicadores, como o nome o indica, consiste na distinção dos indicadores obtidos como sendo correspondentes a um caso de dano ou sem dano. É nesta etapa que é realizada uma análise dos valores atípicos que permitirão monitorizar automaticamente o estado das rodas do veículo e ainda classificar esses valores em diferentes classes.

A discriminação dos indicadores pode ser dividida em duas etapas, a análise *Outlier* e a análise de *Clusters*, as quais serão detalhadas de seguida.

##### 2.7.4.1 Análise Outlier

Um dado discordante ou também chamado de *outlier* é uma observação que é diferente dos restantes dados em algum sentido e, portanto, acredita-se que foi gerada por um mecanismo diferente dos outros dados [109].

A análise *Outlier*, enquadrada na metodologia utilizada, consiste em agrupar, por processos de distribuição probabilística, os dados de condições normais e verificar se os novos dados extraídos têm características semelhantes aos do grupo anteriormente criado, sendo assim possível estabelecer uma comparação entre ambos e determinar se existe uma reação anômala por parte do sistema ou não [48]. É de salientar que os dados devem ser avaliados em relação a um modelo construído a partir de uma condição normal do sistema de interesse, ou seja, sem dano. Sendo assim, com o intuito de possibilitar a distinção de forma automática de cada indicador de dano é utilizado um limite de confiança estatístico (CB – *confidence boundary*), baseado na função de distribuição cumulativa inversa gaussiana. Devido à sua simplicidade e efetividade, esta técnica tem sido utilizada em alguns estudos [109].

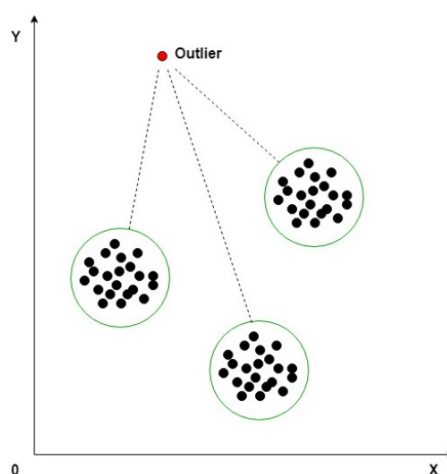


Figura 32 – *Outlier*.

Normalmente, a literatura assume que a distância quadrada de *Mahalanobis* pode ser aproximada por uma distribuição qui-quadrada no espaço n-dimensional. Portanto, a distância de *Mahalanobis* pode ser aproximada por uma distribuição gaussiana e a análise de *outliers* pode ser realizada com base num limiar estatístico. Sob esta hipótese, os limites de confiança (CB) para detecção de DI composto por *outliers* podem ser estimados usando a função de distribuição cumulativa gaussiana inversa (ICDF), considerando a média  $\bar{\mu}$  e o desvio padrão  $\sigma$  do autovetor da linha de base, e alcançando um nível de consistência significativo, na forma:

$$CB = invF_x(1 - \alpha) \quad (2.34)$$

$$F(x \setminus \bar{\mu}, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\alpha}^x e^{-\frac{1}{2}\left(\frac{x-\bar{\mu}}{\sigma}\right)^2} dy, \text{ for } x \in \mathbb{R} \quad (2.35)$$

Em que:

CB – Limite de confiança;

$\alpha$  – Nível de significância;

$\bar{\mu}$  – Valor médio;

$\sigma$  – Desvio padrão dos vetores dos indicadores relativos aos cenários de base.

Sempre que um valor de DI de um indicador é igual ou superior a CB, considera-se que estamos perante um *outlier* (a hipótese nula é rejeitada), representando assim que estamos perante um caso de dano.

### 2.7.4.2 Análise de Clusters

O *clustering* é uma técnica comum de análise de dados que permite dividir um conjunto de dados em grupos, grupos esses que apresentam as mesmas semelhanças. Para que não existam classificações erradas, estes grupos devem ser tão compactos e separados quanto possível, ou seja, o ideal será ter uma distância *intra-cluster* (distância entre os pontos de dados dentro do *cluster*) pequena e uma distância *inter-cluster* (distância entre os pontos de dados em diferentes *clusters*) grande.

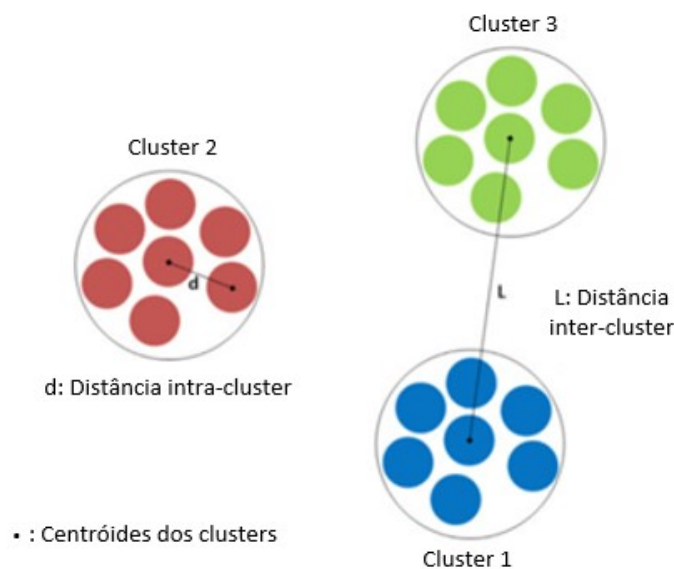


Figura 33 – Distâncias maximizadas e minimizadas na análise de *clusters* [110].

Considerando uma determinada amostra com  $k$  clusters,  $P_k = (C_1, \dots, C_k)$ , a diferença global dentro do mesmo cluster,  $W(P_k)$ , e a diferença global entre os clusters,  $OD$ , podem ser definidas através das seguintes equações:

$$W(P_k) = \frac{1}{2} \sum_{k=1}^k \sum_{c(i)=k} \sum_{c(j)=k} d_{ij} \quad (2.36)$$

$$OD = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N d_{ij} \quad (2.37)$$

Onde:

$N$  – Número total de indicadores;

$c(i)$  – Regra de alocação dos diferentes indicadores que possam existir num só *cluster*, no qual é feita atribuição para um determinado indicador  $i$  um *cluster*  $k$ , com base na diferença medida  $d_{ij}$  definida por cada par de indicadores  $i$  e  $j$ .

Este agrupamento de dados pode ser feito, por exemplo, baseado em densidade, em distribuição e em centróides. Este último, de que é exemplo o *k-means*, tem apresentado resultados surpreendentes [26]. Basicamente, o agrupamento por *k-means* é um método de clusterização baseado em centróides, em que cada elemento do *cluster* tem como base a sua distância média de um dos centros determinados no momento de instanciar o modelo. A principal limitação dos algoritmos do tipo *k-means* é que eles exigem que o número de clusters  $k$  seja especificado com antecedência, contudo em trabalhos de detecção de danos, é uma tarefa quase impossível saber-se a priori este número, daí ser imprescindível arranjar uma solução que possa conseguir estimar este valor de  $k$ .

Em Meixedo et al. [111] foi possível, através de um índice de silhueta global (SIL), obter-se o número de *clusters* de forma automática, sendo possível combater uma grande lacuna deste tipo de *clustering*. Para isso, foi utilizada a conhecida diferença média entre os aglomerados,  $DC$ , cuja formulação é definida abaixo:

$$DC = \frac{1}{K(K-1)} \sum_{k=1}^K \sum_{\substack{c=1 \\ c \neq k}}^K d_{ck} \quad (2.38)$$

Onde:

$K$  – Número de clusters pertencentes às partições maiores índices SIL;

$c$  e  $k$  – Dois de  $K$  clusters;

$d_{ck}$  – Diferença medida entre dois centróides.

Para situações em que não haja dano, os clusters gerados pelo *k-means* são semelhantes e os valores  $DC$  são bastante reduzidos. Caso contrário, em situação de dano, serão gerados clusters diferentes e separados pelo *k-means* e são gerados grandes valores de  $DC$ . Para facilitar na compreensão da técnica, é apresentado um exemplo na Figura 34, no qual é possível entender como é definido o número de *clusters* ótimo, como são criados os centróides, e como é feita a classificação dos dados conforme o número de *clusters*.

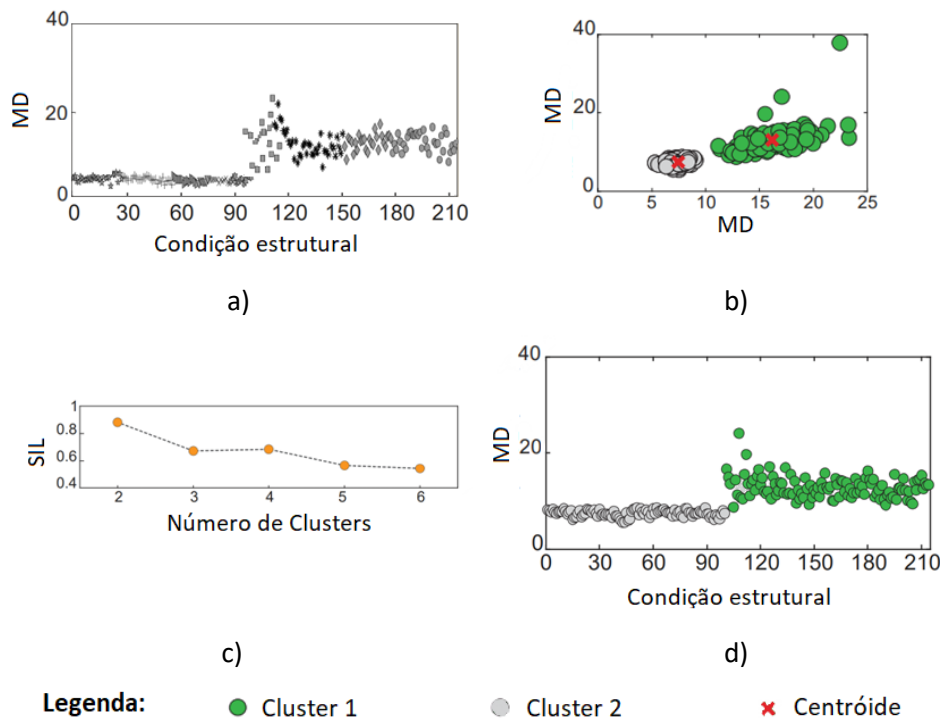


Figura 34 – Alocação dos indicadores em diferentes clusters (adaptado de [25]).

Após a fusão dos dados através da distância de *Mahalanobis*, é obtido um conjunto de pontos (Figura 34 a)). Com base nestes pontos e nas distâncias entre estes, são criados centróides iniciais que vão sendo reposicionados com base nos pontos que estão à sua volta (Figura 34 b)). A partir do momento em que há uma convergência e os centróides não se movem significativamente entre as iterações, o processo acaba e é obtido o valor máximo do índice SIL que permite obter o número ótimo de clusters que melhor descreve o conjunto de indicadores analisados (Figura 34 c)). Por fim, são agrupados os pontos de forma automática de acordo com o número de cluster obtido (Figura 34 d)). Para além do *k-means*, existem diversas outras técnicas. Na Tabela 13 são apresentados alguns estudos que utilizaram *clustering* para agrupar os dados.

Tabela 13 – Estudos relacionados com a análise cluster e respetivas conclusões.

<i>Estudos relacionados com a análise cluster</i>	<i>Tipo de cluster utilizado</i>	<i>Trabalho e conclusões</i>
Estratégia automatizada de deteção de danos em tempo real utilizando medições dinâmicas por Cardoso et al. [112].	<i>k-medoids clustering</i>	A metodologia proposta utiliza um agrupamento <i>k-medoids</i> e utiliza um índice de valor único para indicar se está presente dano ou não em estruturas.
Abordagens lineares para a modelação de não linearidades na monitorização de pontes por Figueiredo et al.[113].	<i>Gaussian mixture models</i>	Este trabalho apresenta um algoritmo baseado em modelos de mistura gaussiana para a deteção de danos em pontes. Compara o algoritmo com algoritmos lineares bem conhecidos e um algoritmo não linear com a capacidade de modelar não linearidades.

## CAPÍTULO 3

### CASO DE ESTUDO: MODELAÇÃO DINÂMICA DO SISTEMA VEÍCULO-VIA

#### 3.1 DESCRIÇÃO DO MODELO PROPOSTO

A análise dinâmica do sistema veículo-via é um problema desafiador devido à natureza tridimensional da infraestrutura e ao desenvolvimento longitudinal que esta possui. Além disso, a interação entre os veículos e a via gera cargas móveis de alta magnitude, influenciadas pelos efeitos inerciais das massas dos veículos. Essa interação ocorre num ponto de contato que leva em consideração a geometria da roda e do carril, sendo que para isso é necessário modelar o perfil da roda sobreposto ao perfil do carril [114]. Ao resolver o problema dinâmico de interação do sistema veículo-via, é possível determinar as forças de contato e, conseqüentemente, obter as respostas dinâmicas na via.

Assim, um modelo numérico do sistema dinâmico calibrado e validado é uma ferramenta essencial, pois permite teoricamente reproduzir cenários equivalentes aos encontrados na realidade. Além disso, é ainda possível simular cenários que seriam impraticáveis de reproduzir no sistema real, como a simulação de diferentes tipos de danos e a observação do comportamento do sistema diante dessas situações.

É de salientar que, neste caso de estudo, todas as modelações foram realizadas num nível bastante complexo, visando simular casos reais, possibilitando a validação da metodologia adotada para a identificação de danos nas rodas e desequilíbrio de cargas, antes do seu teste em ambiente real. Ainda assim, existem níveis de complexidade de modelação mais simples, que para casos em que não seja necessário ir ao pormenor também poderão ser suficientes.

#### 3.2 MODELAÇÃO DO VEÍCULO

Para o desenvolvimento do presente trabalho, foi necessária a utilização de um modelo numérico de um veículo para se conseguir simular os registos das passagens com dano e sem dano. Para isso, foi utilizado um modelo numérico de um veículo de mercadorias de dois eixos, sem *bogie*, constituído por cinco vagões do tipo *Laagrss*, previamente desenvolvido no trabalho de Bragança et al. [115].

O veículo *Laagrss*, Figura 35 a), é um tipo específico de vagão destinado ao transporte de contentores, operando em pares com uma ligação articulada, podendo atingir velocidades até 120 km/h, segundo a classificação da UIC. Este vagão tem uma tara de 27.100 kg e uma capacidade máxima de carga de 52.000 kg, acomodando contentores de 45 pés (13,716 m). Cada secção do veículo é constituída por um vagão de dois eixos com um único nível de suspensão assegurado por uma mola de lâminas. Com este tipo de elementos é possível tirar-se partido da rigidez à flexão progressiva do conjunto de lâminas de aço para controlar os deslocamentos da carroçaria.

Em termos de dimensões, o *Laagrss* tem um comprimento de carga de cerca de 14 m e uma distância entre eixos de 10 m, o que lhe permite transportar um único contentor de 45 pés (13,716 m) ou dois contentores mais pequenos de 20 pés cada (6,096 m cada).

Atualmente, faz parte do material circulante dos operadores ferroviários *Medway* e *COMSA*. A rota do comboio é observável na Figura 35 b), partindo da estação do Entroncamento, em Portugal, seguindo pela linha do Norte até chegar à estação da Pampilhosa. A partir daí, muda para a linha internacional da Beira Alta que o conduz a Espanha.



a)



b)

Figura 35 – Vagão de carga: a) vista geral do vagão b) itinerário de viagem (adaptado de [115]).

O modelo dinâmico tridimensional (3D) foi desenvolvido no software ANSYS® [116]. O modelo emprega elementos de mola-amortecedor para simular as suspensões em todas as direções, enquanto os elementos de massa pontual são utilizados para representar a massa e os efeitos de inércia no centro de gravidade de cada componente dos vagões, como por exemplo a plataforma e os rodados. Os componentes mencionados são conectados por meio de elementos de vigas rígidas. A Figura 36 apresenta o modelo dinâmico do vagão de carga *Laagrss*, mostrando a localização da suspensão, dos rodados e dos eixos.



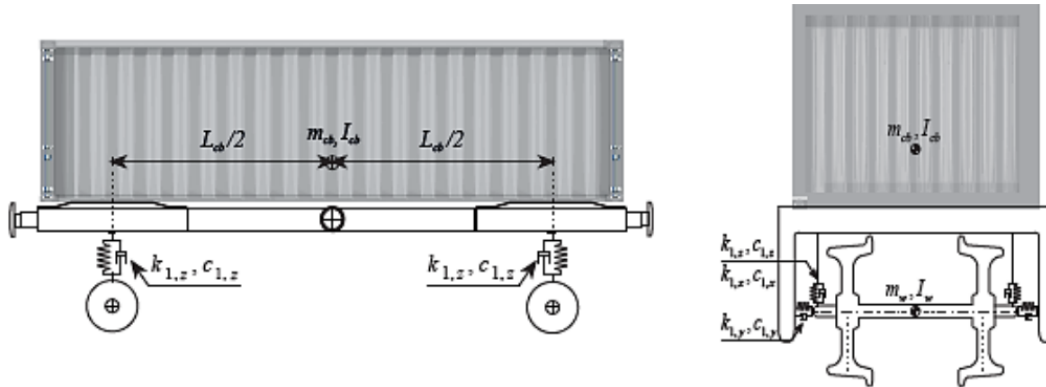


Figura 36 – Modelo dinâmico do vagão de carga *Laagrss* (adaptado de [21]).

Em relação ao modelo numérico do veículo encontra-se apresentado na Figura 37. O modelo foi desenvolvido de forma a representar a realidade, mas de forma simplificada, pelo que se considerou um conjunto de apenas 5 vagões.

As propriedades mecânicas do veículo, nomeadamente a rigidez ( $k$ ), o amortecimento das suspensões ( $c$ ), as inércias ( $I$ ) e as massas ( $m$ ), indicadas na Figura 36 e Figura 37, encontram-se apresentadas na Tabela 14. Esta tabela apresenta ainda as propriedades geométricas, nomeadamente as distâncias longitudinais ( $L$ ), transversais ( $W$ ) e verticais ( $H$ ). A referência a cada um dos elementos principais do veículo, plataforma e rodados, realiza-se pelos índices  $cb$  e  $w$ , respetivamente.

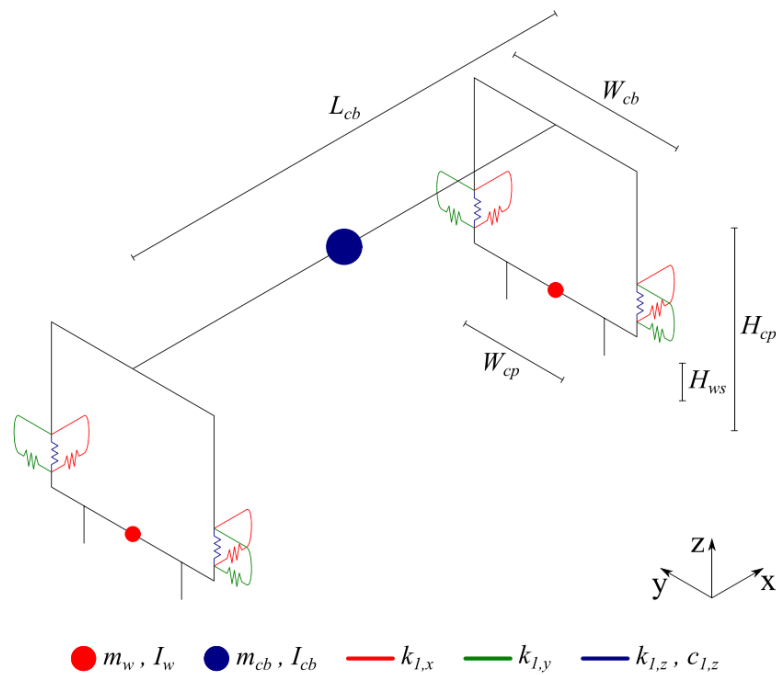


Figura 37 – Modelo numérico do vagão de carga *Laagrss* [115].

Tabela 14 – Propriedades geométricas e mecânicas do modelo numérico do vagão *Laagrss* [115].

<i>Parâmetro</i>	<i>Símbolo (unid.)</i>	<i>Limite de valores (mais baixo/mais alto)</i>	<i>Valor adotado</i>
<b>Carroçaria</b>			
Massa	$m_{cb}$ (ton)	33,2/49,8	41
Momento de inércia xx	$I_{cb,x}$ (t.m <sup>2</sup> )	39,2/58,8	49
Momento de inércia yy	$I_{cb,y}$ (t.m <sup>2</sup> )	337/674	673
Momento de inércia zz	$I_{cb,z}$ (t.m <sup>2</sup> )	-	665
Comprimento	$L_{cb}$ (m)	-	10
Largura	$W_{cb}$ (m)	-	2,170
Altura acima do solo	$H_{cp}$ (m)	-	2,297
<b>Rodado</b>			
Massa	$m_w$ (kg)	1250/2875	1247
Momento de inércia xx	$I_{w,x}$ (kg.m <sup>2</sup> )	-	312
Momento de inércia zz	$I_{w,z}$ (kg.m <sup>2</sup> )	-	312
Bitola	$W_{cp}$ (m)	-	1,668
Altura acima do solo	$H_{ws}$ (m)	-	0,450
<b>Suspensões</b>			
Rigidez longitudinal	$k_{1,x}$ (kN/m)	22500/67500	44981
Rigidez lateral	$k_{1,y}$ (kN/m)	15450/46350	30948
Rigidez vertical	$k_{1,z}$ (kN/m)	1560/2780	1860
Amortecimento vertical	$c_{1,z}$ (kN.s/m)	-	16,7

### 3.3 MODELAÇÃO DA VIA

À semelhança da modelação do veículo, a modelação da via foi realizada em ANSYS®. O modelo desenvolvido por Montenegro et al. [81], baseia-se num esquema de múltiplas camadas, simulando o balastro, as travessas e os carris. As palminhas, apoiadas nas travessas sob o carril, são simuladas como elementos de mola fazendo a ligação entre os dois elementos, assim como todas as restantes ligações entre elementos são também simuladas por elementos elásticos, representando o comportamento mecânico de interface. Através de elementos de pontos de massa discretos, é possível representar-se a massa do balastro, enquanto os carris e as travessas são modelados através de elementos de viga, adotando-se as propriedades dos materiais adequadas a cada um deles. Ambos os carris foram modelados com propriedades equivalentes às secções do tipo UIC60 e a fundação modela-se com recurso a uma série de molas independentes com comportamento elástico e linear. A Figura 38 apresenta de forma esquemática a via.

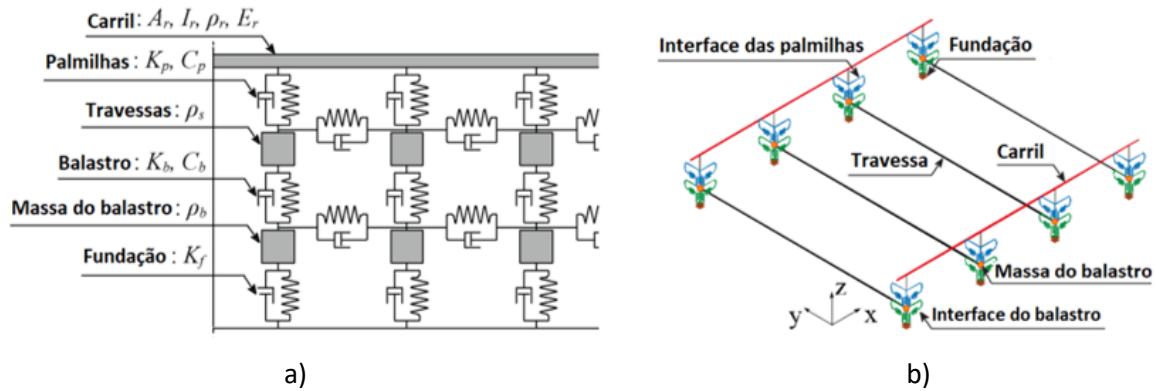


Figura 38 – Modelo numérico da via: a) representação multicamadas b) modelo de elementos finitos (adaptado de [17]).

As propriedades do modelo da via, incluindo a descrição das variáveis, são apresentadas na Tabela 15.

Tabela 15 – Propriedades mecânicas da via (adaptado de [48]).

Parâmetro	Símbolo (unid.)	Valor adotado
Carril	$A_r$ (m <sup>2</sup> )	$7,67 \times 10^{-4}$
	$\rho_r$ (kg.m <sup>3</sup> )	7850
	$I_r$ (m <sup>4</sup> )	$30,38 \times 10^{-6}$
	$\nu_r$	0,28
	$E_r$ (N/m <sup>2</sup> )	$210 \times 10^9$
Palmilhas, dir.longitudinal	$k_p$ (N/m)	$20 \times 10^6$
	$C_p$ (N.s/m)	$50 \times 10^3$
Palmilhas, dir.lateral	$k_p$ (N/m)	$20 \times 10^6$
	$C_p$ (N.s/m)	$50 \times 10^3$
Palmilhas, dir.vertical	$k_p$ (N/m)	$500 \times 10^6$
	$C_p$ (N.s/m)	$200 \times 10^3$
Travessas	$\rho_s$ (N/m)	2590
	$\nu_s$	0,2
	$E_s$ (N.s/m)	$40,9 \times 10^9$
Balastro, dir.longitudinal	$k_{b,x}$ (N/m)	$900 \times 10^3$
	$C_{b,x}$ (N.s/m)	$15 \times 10^3$
Balastro, dir.lateral	$k_{b,y}$ (N/m)	$2250 \times 10^3$
	$C_{b,y}$ (N.s/m)	$15 \times 10^3$
Balastro, dir.vertical	$k_{b,z}$ (N/m)	$30 \times 10^6$
	$C_{b,z}$ (N.s/m)	$15 \times 10^3$
Fundação, dir.longitudinal	$k_{f,x}$ (N/m)	$20 \times 10^6$
	$C_{f,x}$ (N.s/m)	$5,01 \times 10^2$
Fundação, dir.lateral	$k_{f,y}$ (N/m)	$20 \times 10^6$
	$C_{f,y}$ (N.s/m)	$5,01 \times 10^2$
Fundação, dir.vertical	$k_{f,z}$ (N/m)	$20 \times 10^6$
	$C_{f,z}$ (N.s/m)	$5,01 \times 10^2$

### 3.4 MODELAÇÃO DO PERFIL DE IRREGULARIDADES DA VIA

Num cenário real é inviável encontrar-se vias perfeitamente regulares, dado que os carris apresentam certas irregularidades que desempenham um papel crucial na criação de cargas dinâmicas durante os movimentos dos veículos sobre a via. Se essas irregularidades forem ignoradas, as cargas transmitidas pelos rodados à via seriam simplificadas como cargas quase estáticas, contudo, para uma aproximação adequada ao cenário real, é essencial modelar a irregularidade da via [48].

Numa primeira fase são obtidas curvas de densidade espectral de potência (PSD) e, em seguida, são criados perfis artificiais de irregularidades usando o software MATLAB®. Estes desvios são gerados quer na direção vertical quer na direção transversal, tendo por base condições reais na Linha do Norte dos Caminhos-de-Ferro Portugueses.

Na Norma Europeia EN 13848-2 (2006), são definidos os comprimentos de amostragem que devem ser avaliados. De forma a abranger-se os intervalos de comprimento de onda D1 (3 m a 25 m) e D2 (25 m a 70 m), foram selecionados comprimentos de amostragem de 1 m a 75 m, com uma discretização de amostragem de 0,25 m. As curvas de densidade espectral de potência adotadas são as propostas pela Administração Ferroviária Federal (FRA), seguindo a seguinte função [117, 118]:

$$S(K_1) = \frac{10^{-7} \cdot A \cdot K_3^2 \cdot (K_1^2 + K_2^2)}{K_1^4 \cdot (K_1^2 + K_3^2)}$$

Onde:

$K_1$  – Número de onda que depende da frequência espacial cíclica da irregularidade, dado por  $2\pi/\lambda$ ;

$K_2$  – Constante que assume o valor de 0.1465 rad/m;

$K_3$  – Constante que assume o valor de 0.8168 rad/m.

O parâmetro  $A$  está associado à qualidade da via, conforme apresentado na Tabela 16, e depende da velocidade máxima permitida na via em estudo, conforme a Tabela 17.

Tabela 16 – Parâmetro de rugosidade para as classes da via [119].

Classe	1	2	3	4	5	6
$A$ ( $m^3/rad$ )	660,079	376,229	208,841	116,856	65,929	37,505

Tabela 17 – Coeficiente para a densidade espectral de potência [119].

Classe	Tipo de veículo	1	2	3	4	5	6
Velocidade máxima	Mercadorias (km/h)	16	40	64	97	129	177

Neste estudo os dados adotados para o parâmetro de irregularidade são baseados nos cálculos realizados por Mosleh et al. [119]. Nessas medições, o valor calculado para o parâmetro  $A$  é igual a  $1.02 \text{ m}^3/\text{rad}$ , gerado a partir de medições reais efetuadas por um veículo de inspeção, como já referido anteriormente, na Linha do Norte. Sendo assim, foram considerados quatro tipos de perfis de irregularidade em cada uma das direções, vertical e transversal, e em ambos os lados da via. A Figura 39 mostra quatro perfis de desnível do carril correspondentes às irregularidades laterais e verticais apenas do carril direito. Os perfis desenvolvem-se ao longo de um trecho reto de 170 m, sendo este o comprimento total utilizado em todas as simulações numéricas. Este comprimento tem de ser suficientemente grande para que seja possível o veículo manter contacto com a via durante a simulação do comboio a passar sobre a via.

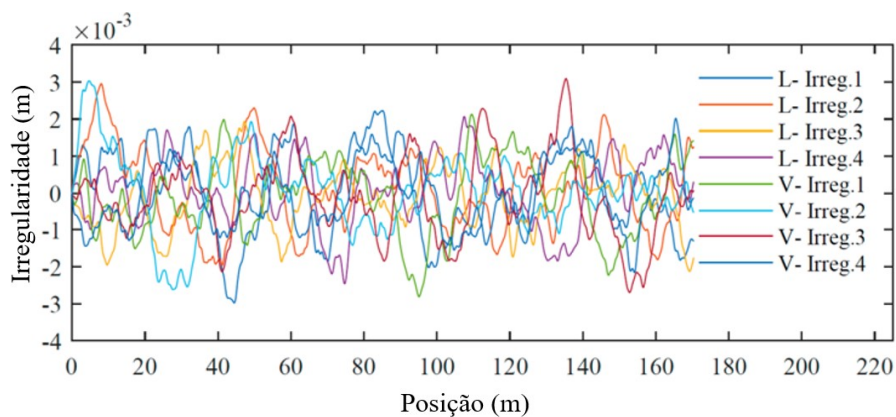


Figura 39 – Perfil de irregularidade da via (direção vertical, V e transversal, L) (adaptado de [27]).

### 3.5 MODELAÇÃO DO PERFIL DE IRREGULARIDADE DAS RODAS

Na medida que o caso-estudo a realizar neste trabalho engloba dois tipos de danos nas rodas (lisos e poligonização), serão abordados seguidamente detalhes da modelação dos perfis de irregularidade das rodas para cada tipo de dano.

#### 3.5.1 Modelação do perfil de irregularidade da roda poligonizada baseado no número de harmónicos

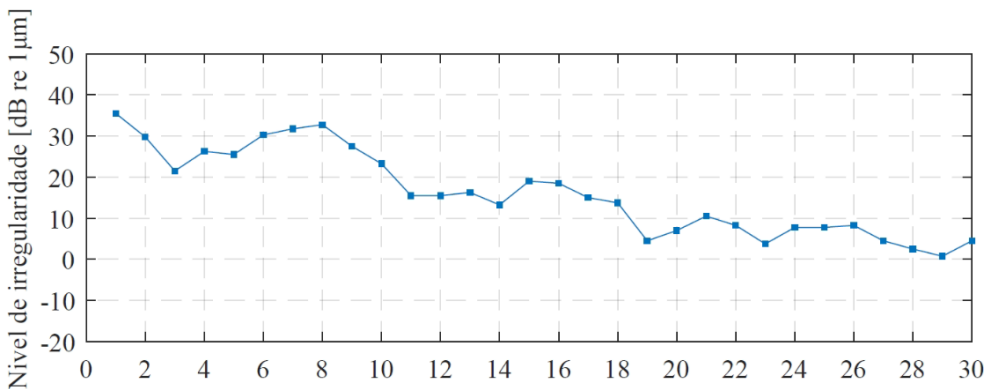
Para caracterizar as irregularidades das rodas poligonizadas do veículo de acordo com o número de harmónicos, foram considerados diferentes números de harmónicos dominantes de acordo com

medições experimentais recolhidas na literatura. Como a formação dos harmónicos depende de vários fatores com o tipo de comboio ou até da velocidade, a recolha de perfis de irregularidades incidu em estudos onde as condições fossem semelhantes às adotadas neste caso de estudo. Após alguma pesquisa, optou-se por considerar os intervalos de harmónicos que estão retratados na Tabela 18.

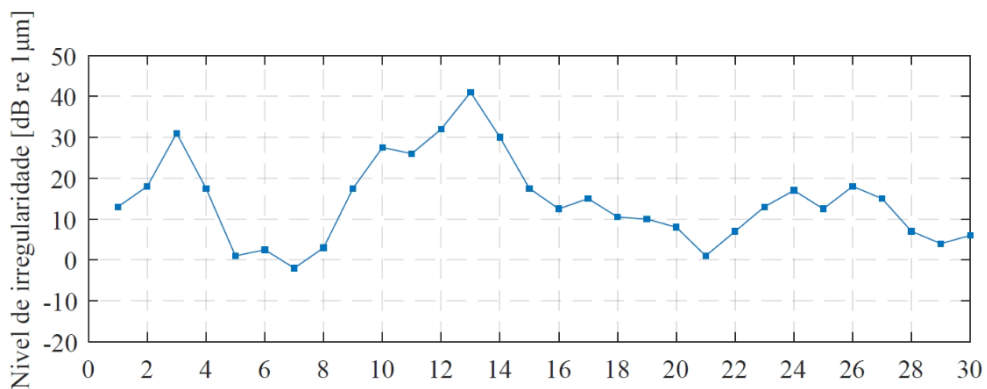
Tabela 18 – Intervalos das ordens dos harmónicos estudadas.

<b>Ordem do harmónico</b>	<b>Razão da escolha</b>
H 6-8	Harmónicos identificados em operações de serviço num veículo com velocidades semelhantes [78].
H 12-14	Caso intermédio e corrente num veículo com velocidades semelhantes [8].
H 17-18	Número de harmónicos elevado e veículo semelhante ao do caso de estudo [120].

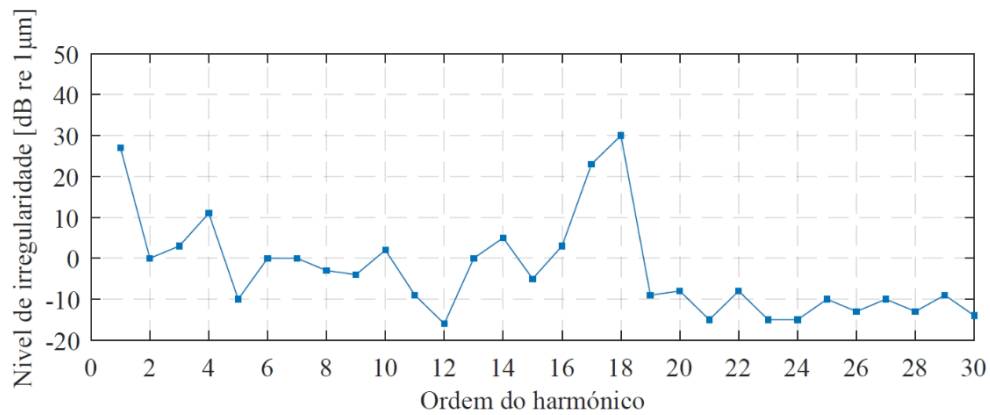
Com o intuito de gerar os perfis de irregularidade perimetrais, foram utilizados os espectros da irregularidade da roda ( $L_w$ ) reais apresentados na Figura 40, que foram obtidos através dos estudos experimentais reais. Estes perfis permitiram caracterizar a poligonização das rodas que serviram de base às análises numéricas deste trabalho.



a)



b)



c)

Figura 40 – Espectros da irregularidade da roda ( $L_w$ ) e ordem harmónica ( $\theta$ ) de rodas poligonais com ordens harmónicas dominantes: a) 6-8 harmónicos b) 12-14 harmónicos c) 17-18 harmónicos.

Para cada intervalo de harmónicos considerado (6-8; 12-14; 17-18), foram gerados 10 perfis aleatórios de irregularidades perimetrais no programa MATLAB®. Para a sua definição geométrica foram usadas as expressões apresentadas na secção 2.4.2.2.2.

Na Figura 41 é apresentado um exemplo de perfil de irregularidade para cada tipo de harmónico.

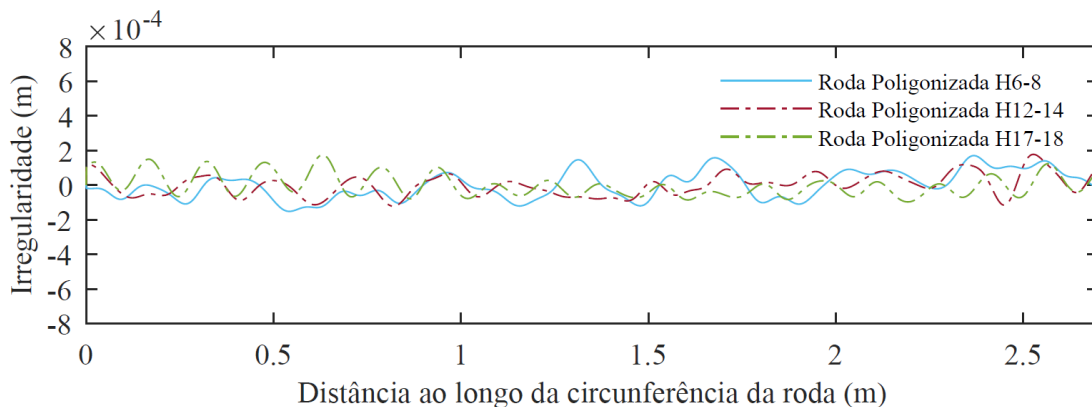


Figura 41 – Exemplos de perfis de danos poligonais para cada severidade de dano.

### 3.5.2 Modelação do perfil de irregularidade de liso na roda

À semelhança da poligonização, também nos lisos foi realizada uma recolha bibliográfica de forma a caracterizar o dano em diferentes severidades. Na Tabela 19 são apresentados os comprimentos dos lisos que serviram de base à geração dos perfis de irregularidade da roda.

Tabela 19 – Intervalos para comprimento dos lisos.

<b>Comprimento do liso</b>	<b>Razão da escolha</b>
10-20 mm	Verificar se a metodologia é suficientemente robusta para detetar o início de um liso [26].
25-50 mm	Intervalos entre 25-100 mm são correntemente encontrados neste tipo de veículos de mercadorias. Dividiu-se os grupos de forma a ter casos intermédios e casos severos [26].
55-100 mm	

Como um liso na roda é uma pancada pontual que ocorre de perímetro em perímetro, o valor da irregularidade aumenta apenas nessa zona. Sendo assim, quanto maior for o comprimento do liso, maior a irregularidade, logo mais severo será o liso. Na Figura 42 são apresentados os perfis de irregularidade característicos de cada severidade, obtidos através das equações da secção 2.4.1.3.2.

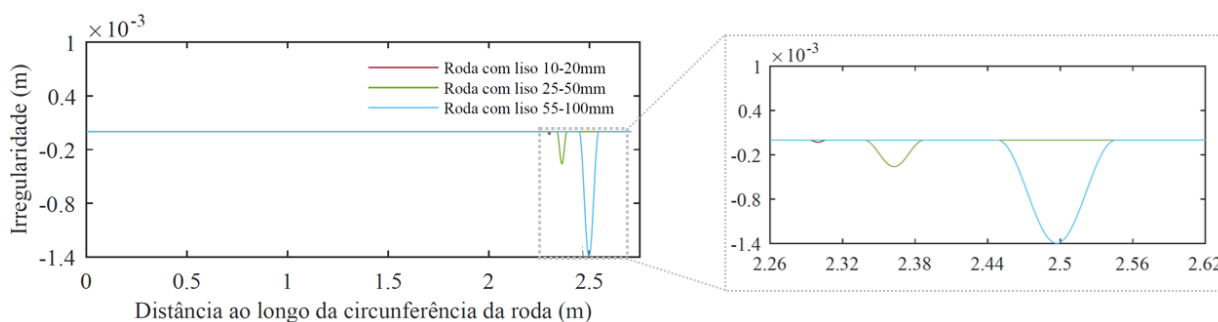


Figura 42 – Exemplos de perfis de lisos nas rodas de cada severidade.

### 3.6 INTERAÇÃO DINÂMICA VEÍCULO-VIA

No presente estudo, as simulações de interação dinâmica entre veículo e a via foram realizadas utilizando o software *VSI-Vehicle-Structure Interaction Analysis*, validado e descrito em pormenor em Montenegro et al. [121] e que vem sendo utilizado em diferentes aplicações [89, 122]. Este software foi inicialmente desenvolvido por Neves et al. [84] e posteriormente aprimorado por Montenegro et al. [121] e conta com cinco principais passos para a obtenção dos registos numéricos, que são posteriormente utilizados na metodologia de deteção e classificação de danos.

Resumidamente, o comboio (i) e a via (ii) são ligados através de um modelo 3D que representa o contacto roda-carril (iii). A teoria hertziana é aplicada para calcular as forças normais de contacto, enquanto a rotina USETAB permite calcular as forças tangenciais resultantes do fenómeno de rolamento. Para efetuar as simulações numéricas, utiliza-se o MATLAB® para importar as matrizes estruturais do veículo e da via, que foram previamente modeladas num software de elementos finitos (EF). É de notar que ambos os subsistemas são inicialmente modelados separadamente no ANSYS®. Uma vez que a via apresenta



irregularidades iniciais, são definidos perfis de irregularidade (iv) no MATLAB® com base na Norma Europeia EN 13848-2 [123]. Para os defeitos OOR, são definidos vários perfis com base em perfis representativos das diferentes severidades, que são sobrepostos posteriormente aos perfis da via. Tudo isto é integrado pelo software VSI (v) utilizando uma abordagem totalmente acoplada. Após a aplicação do VSI são obtidas respostas numéricas que simulam a leitura num sensor (por ex. acelerómetro) aquando da passagem de um comboio sobre esta zona.

Para mais pormenores sobre as propriedades da via e do comboio, incluindo informações adicionais sobre a interação comboio-via, podem ser consultadas as publicações anteriores dos autores [89, 121]. A Figura 43 mostra a esquematização de um modelo de interação veículo-via.

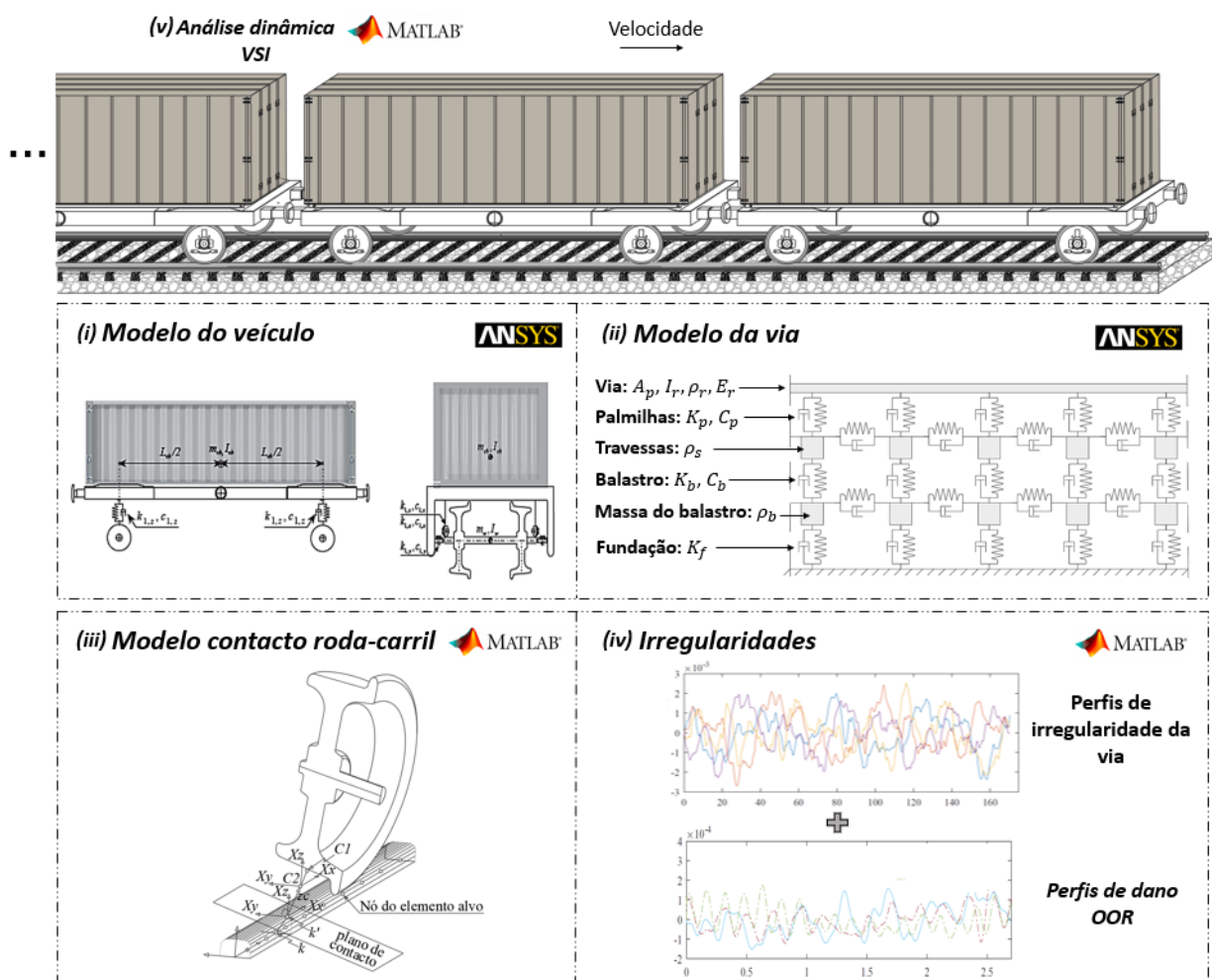


Figura 43 – Esquematização do modelo de interação veículo-via.

### 3.7 SIMULAÇÕES

Foram simulados diferentes tipos de cenários baseados em condições reais de operabilidade, quer para cenários de base (sem dano) quer para os cenários de dano. Como se pretende em trabalhos futuros a

aplicação da metodologia a respostas reais é imperativo a utilização das mesmas condições no estudo numérico.

### 3.7.1 Sistema de monitorização *Wayside* virtual

De acordo com os resultados demonstrados em [21, 26], a localização dos sensores mais promissora para a obtenção de bons resultados é no carril, daí ter sido adotada no caso de estudo. Com o intuito de reduzir os custos de instalação e manutenção, é importante reduzir o número de sensores instalados sem comprometer a qualidade dos resultados, pelo que foi definido que seriam 6 o número máximo de acelerómetros e 10 de extensómetros a aplicar no sistema de monitorização *wayside* virtual. Para perceber qual seria o número ótimo de sensores foi realizada uma análise de sensores, conforme se apresenta no Anexo 1.

É de salientar que estes sensores são apenas unidireccionais, avaliando apenas o movimento vertical. Na Figura 44 é apresentada a disposição dos acelerómetros e dos extensómetros.

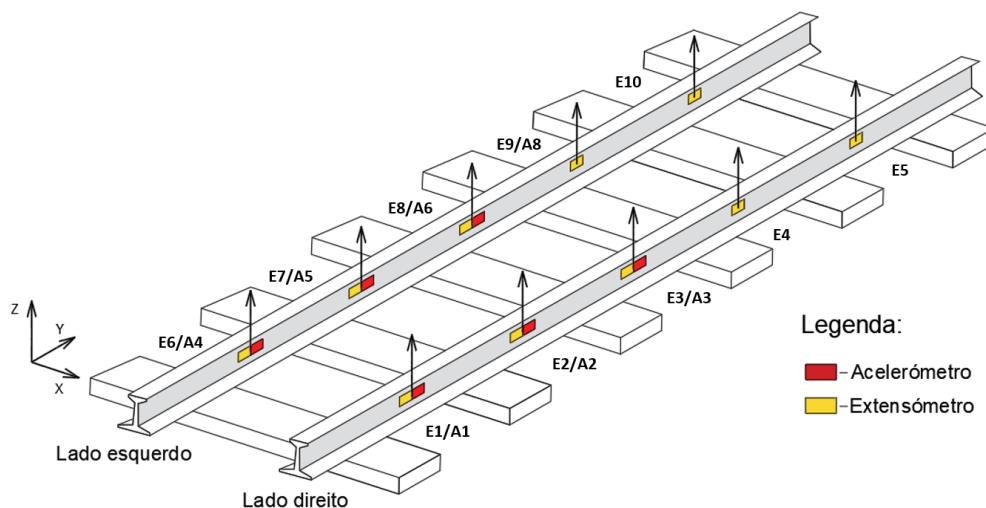


Figura 44 – Sistema de monitorização *wayside* virtual (em perspetiva 3D).

### 3.7.2 Cenários de base

Para definir uma base de referência robusta, são efetuadas várias simulações com diferentes configurações de cargas, irregularidades da via e velocidades dos veículos. A Figura 45 resume os pressupostos para cenários de referência, utilizando seis esquemas de carga, quatro perfis de irregularidade da via (1-4) e cinco velocidades diferentes (40 a 120 km/h em intervalos de 20 km/h). No total, são efetuadas 113 simulações. Os cenários de carga considerados são: vazio, meia carga e carga completa e três outros esquemas de carga (UNB1, UNB2 e UNB3) com cargas desequilibradas dentro dos limites prescritos pelo código UIC [79].

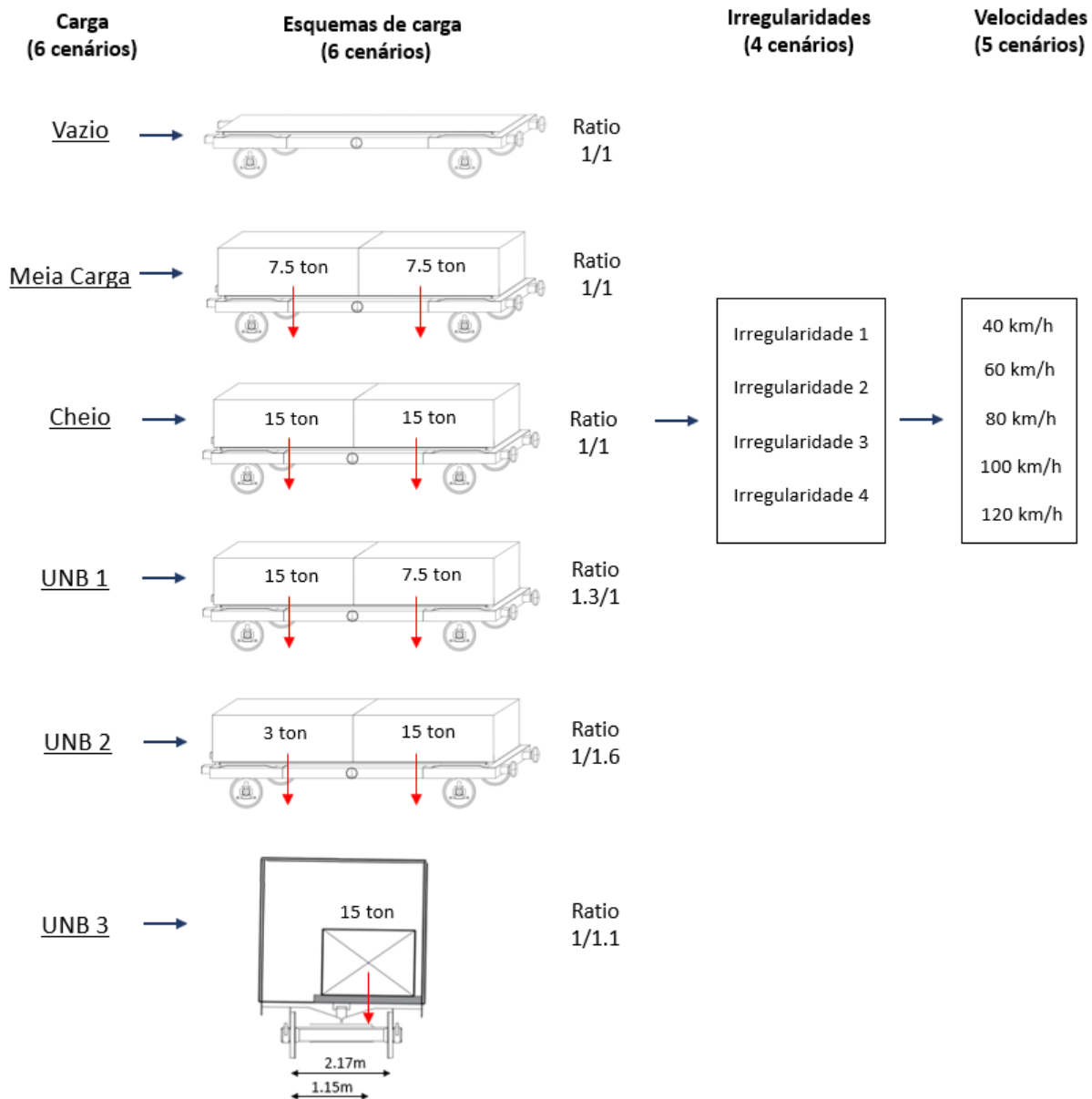


Figura 45 – Cenários de *baseline* (sem dano).

### 3.7.3 Cenários de dano

Os cenários de dano foram definidos com base nos perfis de irregularidades das rodas (secção 3.5.1) e em casos de desequilíbrios de carga (secção 2.5), e são apresentados esquematicamente na Figura 46. Para cada tipo de dano do rodado e desequilíbrio de carga, foram definidos, à semelhança dos tópicos abordados na *baseline*, esquemas de carga, localização, velocidades e as severidades características de cada dano, já abordadas anteriormente.

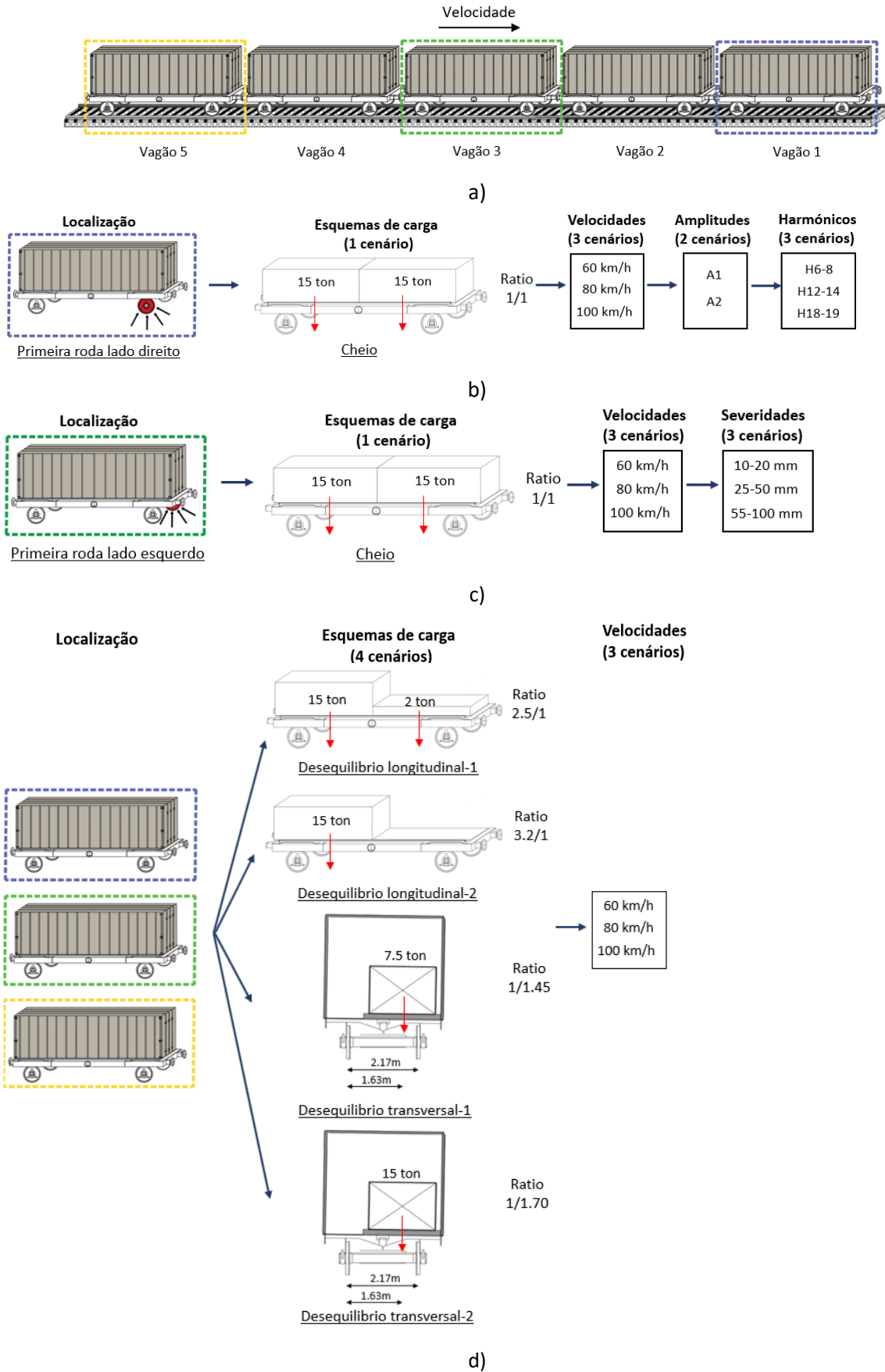


Figura 46 – Cenários de dano considerados: a) identificação dos vagões b) poligonização c) lisos d) desequilíbrios de carga.

Na Figura 46 a), é apresentado o veículo em estudo com a identificação dos 5 vagões do tipo *Laagrss*. Seguidamente, na Figura 46 b), é demonstrado um dos danos no rodado, a poligonização. Este dano foi simulado na primeira roda do lado direito do 1º vagão, com o veículo cheio (caso mais desfavorável). Foram consideradas três velocidades (60, 80 e 100 km/h), e para cada velocidade foram calculadas 10 análises para cada número de harmónicos. De entre as 10 análises, 5 destas são referentes a amplitudes mais baixas, entre 0,25-0,35 mm, designada por amplitude A1 e as outras 5 relativas a amplitudes mais altas, entre 0,65-0,75 mm, designadas por amplitude A2. Sendo assim, para cada velocidade foram obtidos 30 cenários, totalizando 90 cenários de dano de poligonização. A Figura 46 c), refere-se aos lisos nas rodas. Neste caso, o dano foi gerado no 3º vagão, na primeira roda do lado esquerdo. As condições adotadas nas cargas e nas velocidades são iguais à poligonização, pelo que neste caso também se obteve 90 cenários de dano.

Por fim, na Figura 46 d), são apresentados os cenários de desequilíbrio de cargas adotados. Foram definidos 4 esquemas de carga, em que dois se referem a desequilíbrios longitudinais e os outros dois a desequilíbrios transversais. Todos estes cenários representam desequilíbrios de carga acima dos rácios prescritos no código UIC. Ao contrário dos danos nos rodados, são definidas três localizações diferentes para os desequilíbrios, no 1º, 3º e 5º vagão. Para cada uma das localizações foram gerados os quatro esquemas de carga referentes a 3 diferentes velocidades (60, 80 e 100 km/h), totalizando assim 36 cenários.

### 3.7.4 Respostas dinâmicas da via

Por forma a aproximar as respostas obtidas a um contexto de medições experimentais na via, foi adicionado 5% de ruído artificial às respostas numéricas. Para além disto, os registos numéricos foram filtrados com base num filtro digital passa-baixo *Chebyshev* tipo II com uma frequência de corte igual a 1500 Hz nos danos OOR e 100 Hz para os desequilíbrios de carga, com o intuito de eliminar eventuais picos de altas frequências que não estão relacionados com os danos simulados. No caso dos danos OOR foi considerada a frequência de corte nos 1500 Hz para conseguir captar a influência dos lisos nas rodas, já que estes têm um conteúdo de frequência elevado. A variação desta frequência de corte na poligonização não tem grande influência pelo que a utilização desta frequência de corte pode também ser utilizada.

Já nos desequilíbrios de carga, como a unidade a ser medida é a extensão e esta não está dependente de fenómenos dinâmicos, a frequência de corte foi reduzida significativamente, para ser removida a influência de fenómenos indesejados para as análises.

Na Figura 47 é possível observar a influência de diferentes regimes de carga, velocidades dos comboios e perfis de irregularidade na resposta da via.

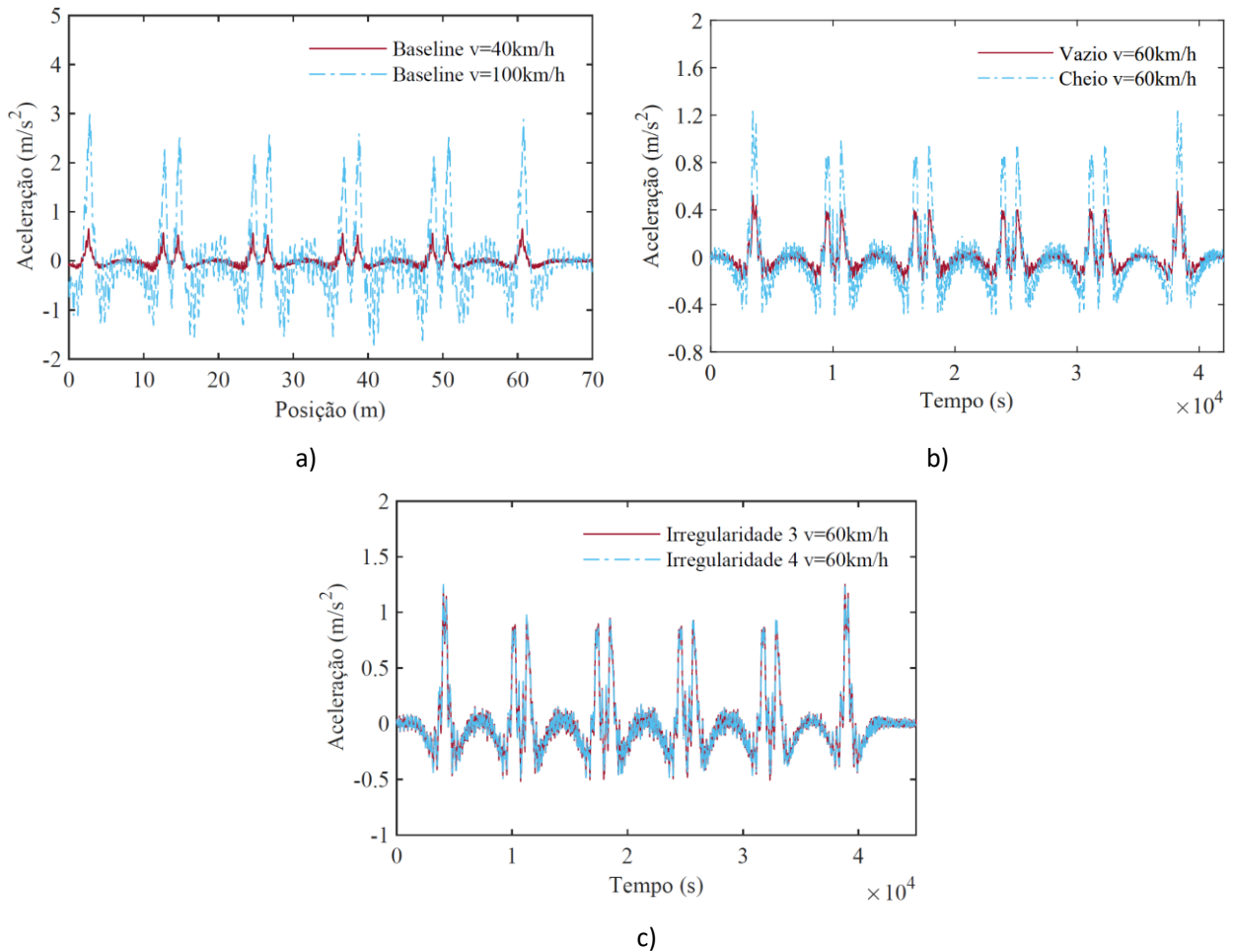


Figura 47 – Diferenças entre registos de acelerações no sensor 1 considerando cenários de base: a) influência da velocidade b) influência da carga c) influência da irregularidade da via.

A Figura 47 a) retrata a influência de diferentes velocidades (40 km/h e 100 km/h) nas respostas das acelerações medidas ao longo do tempo para um caso sem dano. Considerando o mesmo tipo de carregamento (cheio) e o mesmo perfil de irregularidade (irregularidade 2), verifica-se que quanto maior é a velocidade maior é a magnitude da resposta, daí a importância da adoção de técnicas não lineares. Na Figura 47 b), observa-se as diferenças na resposta variando o tipo de carga considerando a mesma velocidade (60 km/h) e o mesmo tipo de irregularidade (irregularidade 2), onde é possível concluir que quanto maior o carregamento maior é a magnitude da resposta. Por fim, na Figura 47 c) demonstra-se a influência dos perfis de irregularidades da via considerando o mesmo tipo de carga (cheia) e a mesma velocidade (60 km/h). Aqui conclui-se que não há uma grande diferença entre as respostas, sendo este o parâmetro menos sensível.

### 3.7.4.1 Poligonização

Na Figura 48 a), ilustra-se a comparação da resposta medida em relação à passagem de vagões com e sem dano poligonal. Nesta mesma figura é visível uma grande diferença entre as acelerações medidas num

cenário de dano com harmónicos 6-8H e um cenário *baseline*. Como o dano se encontra no primeiro vagão, o pico das acelerações é mais pronunciado numa fase inicial do registo e situa-se em redor de 20  $m/s^2$ . Já na Figura 48 b), é feita a comparação entre os dois grupos de harmónicos mais elevados (12-14H e 17-18H). Observa-se nestes casos que a resposta em aceleração aumenta consideravelmente para valores em redor de 45 e 60  $m/s^2$ , respetivamente.

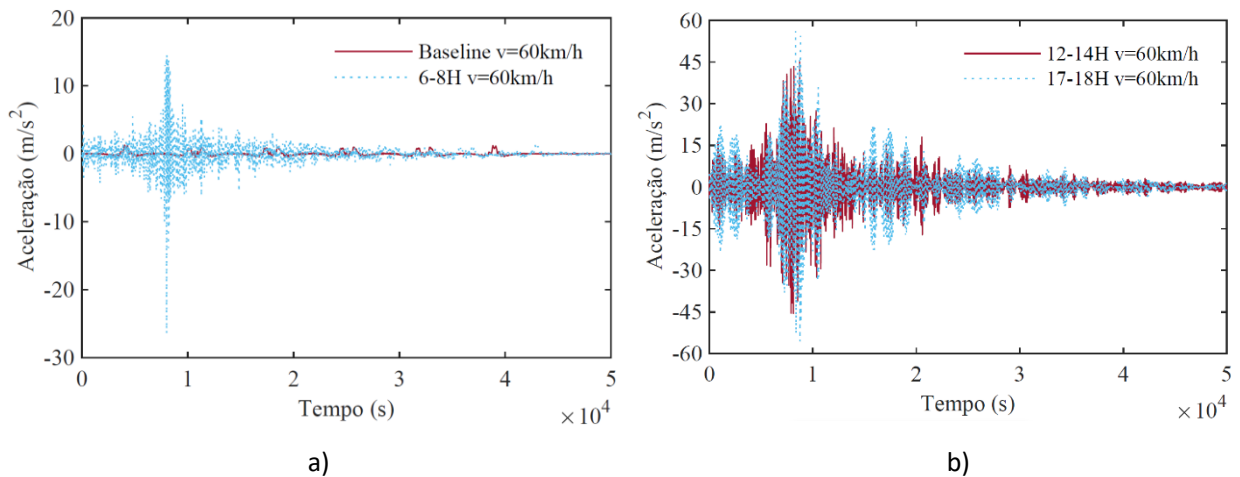
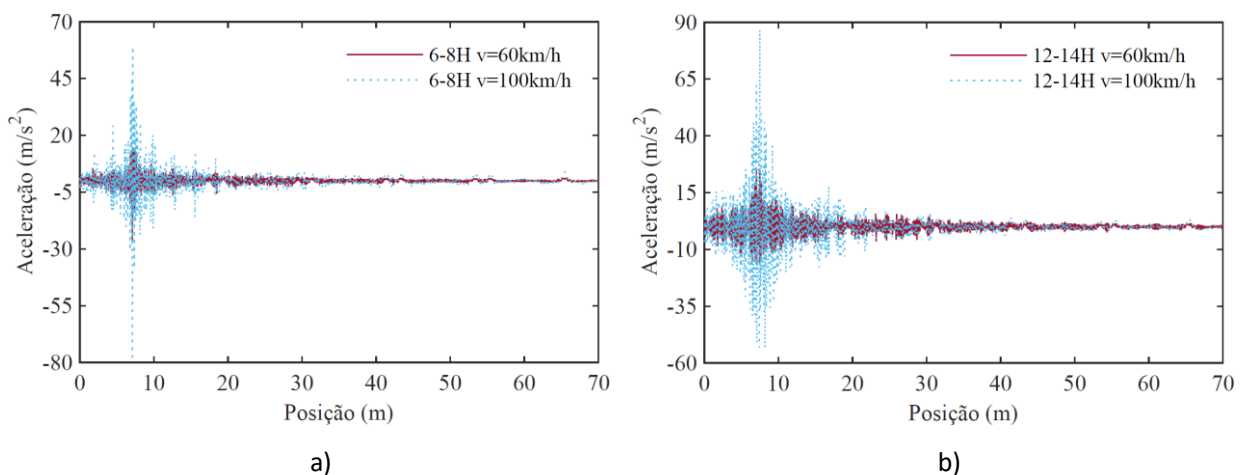


Figura 48 – Comparação de registos de aceleração com dano e sem dano e de diferentes severidades com presença de poligonização.

No caso da poligonização não é possível afirmar que uma resposta com maior aceleração seja mais prejudicial. Isso acontece porque este tipo de dano está diretamente relacionado com o tipo de veículo e com a velocidade. Para determinadas velocidades, veículos com ordens de harmónicos mais baixas podem entrar em ressonância com a via e, dessa forma, amplificar muito a resposta, tornando-a mais gravosa.

No sentido de classificar os danos independentemente da velocidade, é crucial compreender as discrepâncias que existem ao analisar-se a mesma severidade de dano, mas em velocidades diferentes de circulação, como se apresenta na Figura 49.



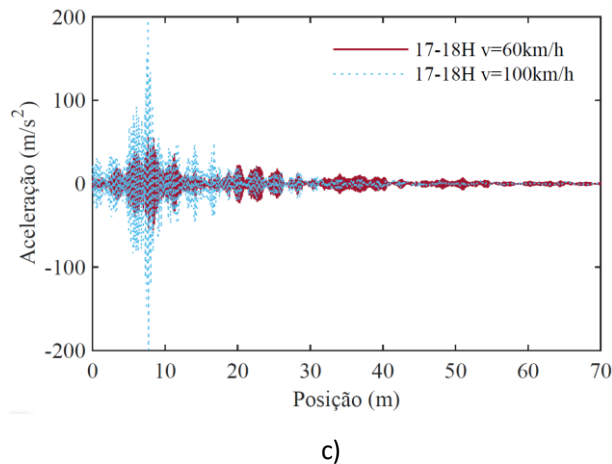
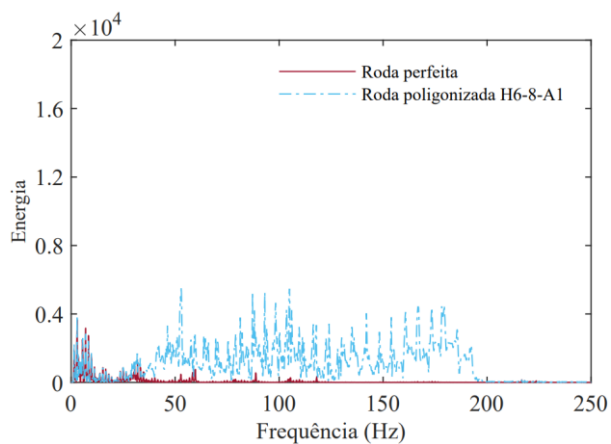


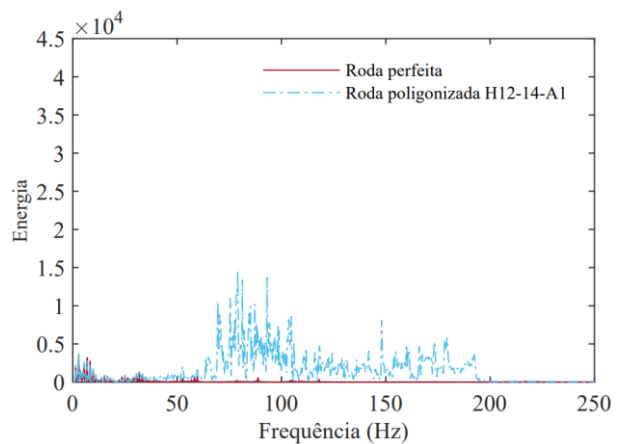
Figura 49 – Comparação de registos de aceleração de diferentes severidades e diferentes velocidades.

Começando pelo estudo das diferenças no número de harmónicos, é perceptível um acréscimo de acelerações na zona da ocorrência do dano aquando do aumento da velocidade. Como as diferenças são bastante notáveis entre velocidades e se pretende agrupar as severidades pelas características ligadas ao número de harmónicos e não por velocidades, a metodologia tem de ser robusta o suficiente para não confundir, por exemplo, um caso de 17-18H a uma velocidade de 60 km/h com um caso 6-8H a 100 km/h.

Após alguns testes, verificou-se que a aplicação da Transformada Rápida de Fourier (FFT) aos registos de aceleração, em que é feita a conversão de um sinal no domínio do tempo para o domínio da frequência, permitia a obtenção de melhores resultados em termos de classificação da severidade do dano. Nesse sentido deverá ser considerado este pré-processamento das acelerações com o FFT antes da aplicação da metodologia proposta. Na Figura 50, são apresentados alguns exemplos de respostas no domínio da frequência para o sensor 1, onde são comparadas rodas perfeitas com os vários grupos de harmónicos considerados para as rodas poligonizadas.

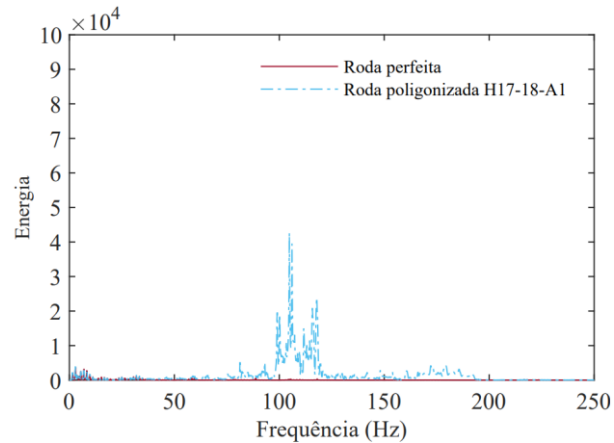


a)



b)





c)

Figura 50 – Respostas no domínio da frequência na posição 1 referente à poligonização para 60 km/h.

Após a análise dos vários espectros FFT, verifica-se que a poligonização apresenta conteúdo de frequência no intervalo de 0 a 200 Hz, para uma velocidade de 60 km/h. A partir da Figura 50 b) e da Figura 50 c), verifica-se que os picos máximos de frequência para os perfis H12-14 e H17-18 são em torno de 80 Hz e 110 Hz, respetivamente, enquanto que para o perfil H6-8, Figura 50 b), é obtido um espectro mais plano sem mostrar um pico de frequência bem evidente. Isso pode ser explicado pelo perfil de danos deste grupo (Figura 40), que tem um espectro de irregularidades com conteúdo significativo em vários harmónicos. Comparados os três grupos de dano, constata-se uma grande diferença em termos de energia. O grupo H6-8 apresenta um valor mais reduzido e vai aumentando progressivamente até ao grupo H17-18.

### 3.7.4.2 Lisos

Dado que o liso se manifesta como uma pancada pontual, o sinal de dano assemelha-se ao sinal sem dano, distinguindo-se apenas na zona da pancada, onde é observável um incremento significativo no valor das acelerações. Na Figura 51 a) é comparada a severidade mais baixa do dano (10-20 mm) com um caso sem dano, onde se verifica que as diferenças são muito reduzidas, mas ainda assim é possível identificar os picos de aceleração relativos às pancadas do liso. A escolha deste grupo de dano foi estratégica, com o objetivo de perceber se a metodologia conseguia desde logo detetar um dano inicial de liso na roda ou não. Na Figura 51 b), são comparadas as restantes severidades (25-50 mm – média e 55-100 mm – alta) e a diferença é notória. Enquanto a severidade média ronda valores de aceleração na ordem dos  $50 \text{ m/s}^2$ , a severidade alta atinge aos  $250 \text{ m/s}^2$ . Ainda assim, estas duas severidades apresentam valores muito superiores à de 10-20 mm que ronda os  $3 \text{ m/s}^2$ .

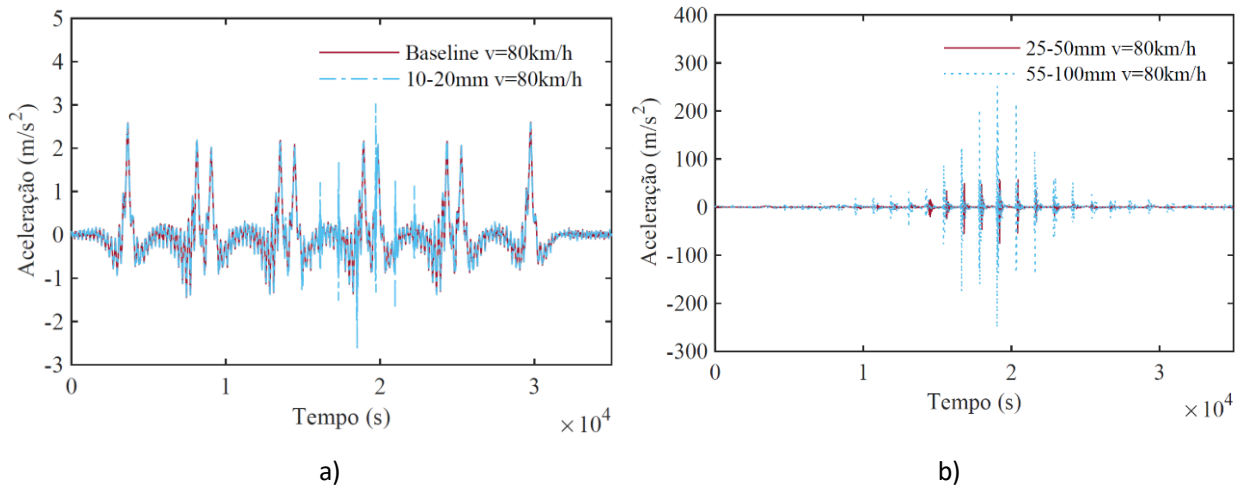
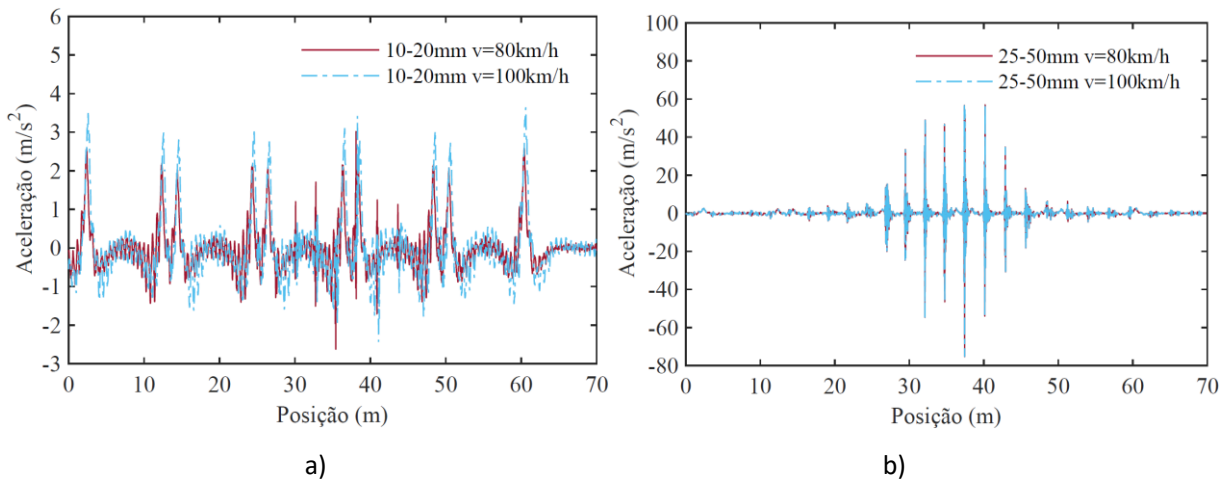
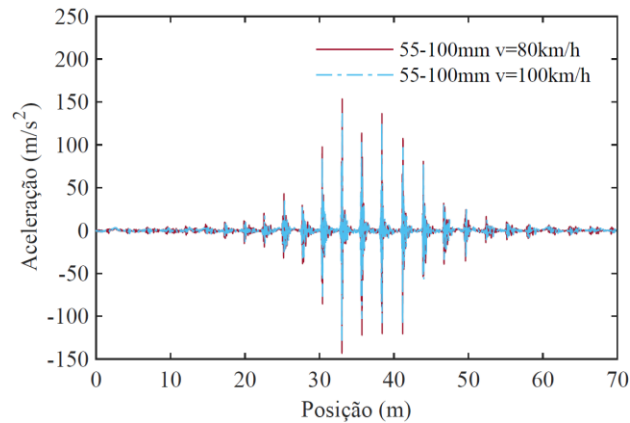


Figura 51 – Comparação de registros de aceleração com dano e sem dano e de diferentes severidades com presença de lisos.

A influência da velocidade nas respostas é também avaliada agora nos danos de lisos, ilustrada nos gráficos da Figura 52. Como as velocidades de circulação correntes de um veículo de mercadorias geralmente não variam muito (80 a 100 km/h) e os intervalos de velocidades que os separam são pequenos, os registros apresentam algumas diferenças, ainda que pouco significativas. Uma conclusão adicional que se pode retirar é que nem sempre velocidades mais elevadas resultam em acelerações mais altas, uma vez que essa relação não varia linearmente ao longo do tempo, como se observa na Figura 52.



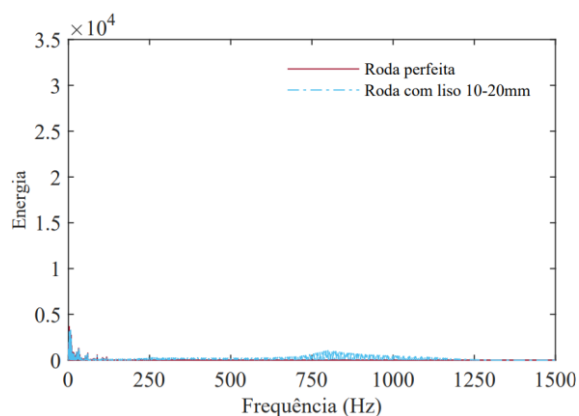


c)

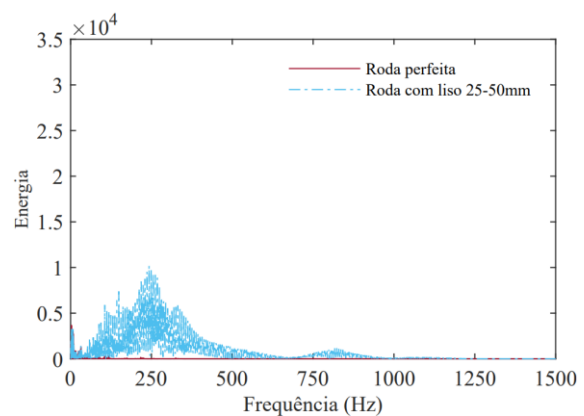
Figura 52 – Comparação de registos de aceleração das diferentes severidades de lisos com a alteração da velocidade.

À semelhança do caso da poligonização, a transformação do sinal de aceleração para o domínio da frequência usando a transformada FFT é também realizada no caso dos lisos, sendo apresentado na Figura 53 os respetivos espectros de frequência para cada severidade.

Neste caso dos lisos, o conteúdo de frequência está restrito na gama de 0 a 1500 Hz conforme a frequência de corte realizada. Verifica-se que a partir da frequência de 1250 Hz a energia tende a manter-se 0, pelo que se conclui que o corte a 1500 Hz é suficiente para captar todas as frequências necessárias para este tipo de defeito. Uma vez que um liso corresponde a uma força de impacto elevada aplicada num muito curto espaço de tempo, a frequência de vibração gerada por um liso será mais elevada do que a causada pela poligonização, daí estes mostrarem conteúdo em frequências mais altas, ao contrário das situações sem dano e poligonização, que mostram conteúdo predominantemente em frequências mais baixas. Em termos de severidade, o grupo mais severo (55-100 mm) é aquele com maior conteúdo de frequência, como esperado, com um pico a variar essencialmente entre 0 e 500 Hz. Por outro lado, o grupo menos severo (10-20 mm) é aquele que induz menos conteúdo de frequência e é muito semelhante à resposta de base.



a)



b)

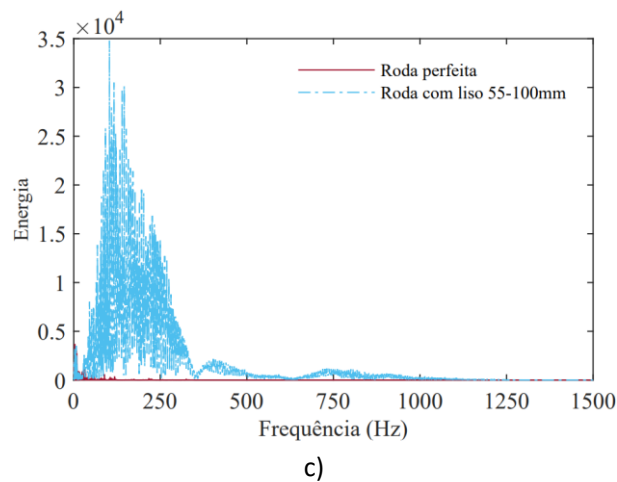
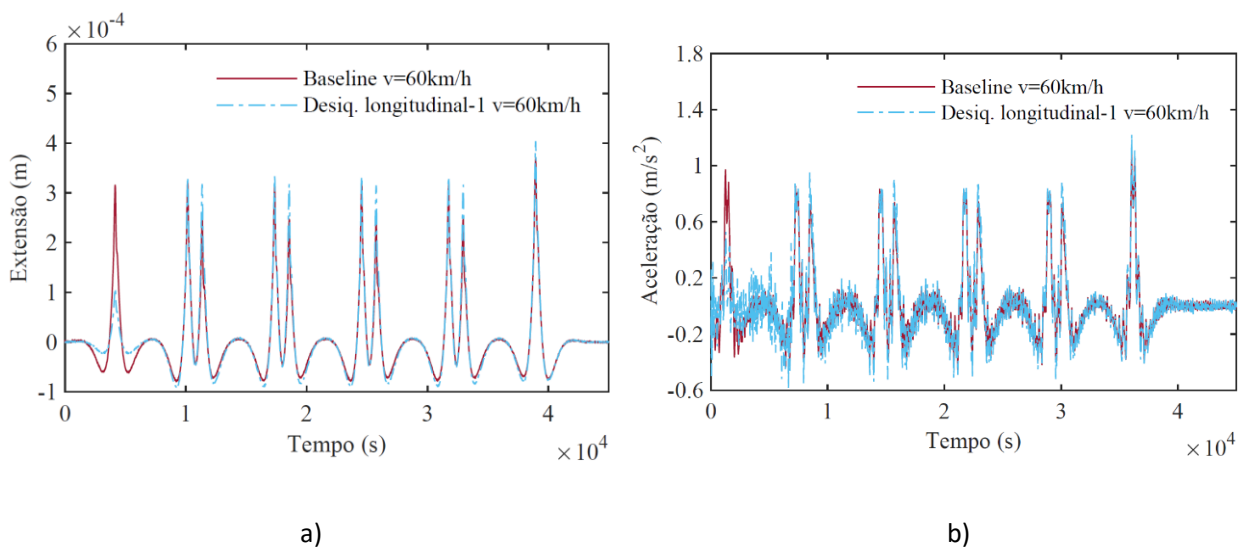


Figura 53 – Respostas no domínio da frequência na posição 1 referente aos lisos para 60 km/h.

### 3.7.4.3 Desequilíbrio de cargas

Ao contrário dos danos nos rodados atrás analisados (poligonização e lisos), no caso do desequilíbrio de cargas, as respostas de deformação são mais claras do que as respostas em aceleração e, visualmente, é possível identificar os desequilíbrios. De um modo geral, no caso dos desequilíbrios de carga as diferenças de aceleração ou extensão em relação aos cenários de base são muito reduzidas, e não há uma amplificação tão grande como no caso dos danos dos rodados, o que poderá antever uma maior dificuldade na identificação correta deste defeito.

Para compreender as principais diferenças na identificação de desequilíbrios de cargas através do uso de acelerações e extensões, e reconhecer as vantagens da utilização de registos em extensão, é exibida a Figura 54. Através da análise da Figura 54 a) e Figura 54 b), pode observar-se um cenário de desequilíbrio longitudinal e na Figura 54 c) e Figura 54 d), um cenário de desequilíbrio transversal.



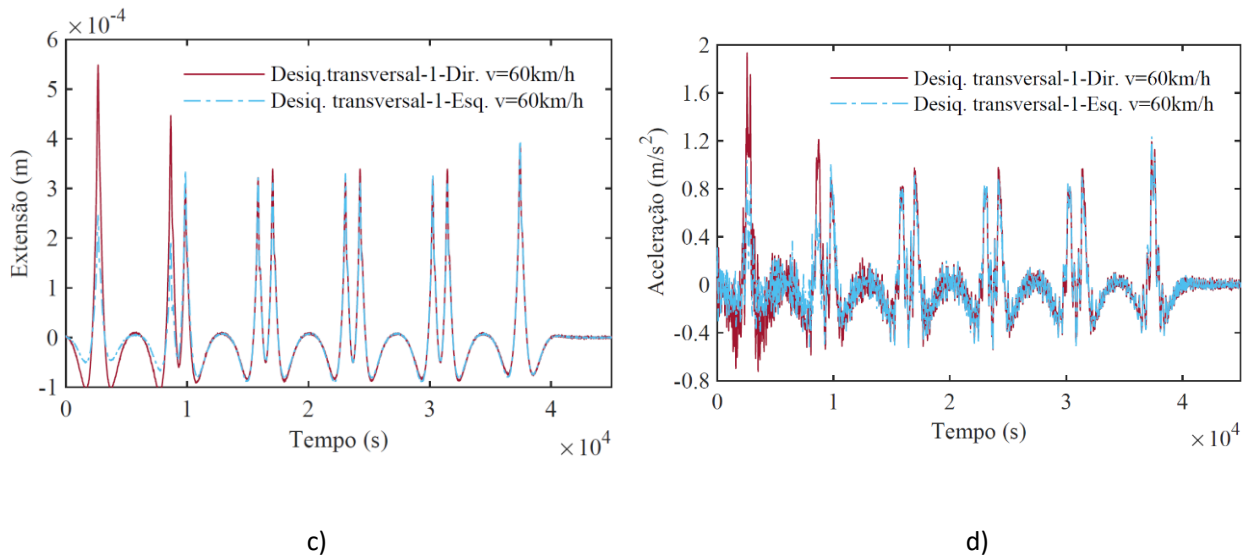


Figura 54 – Comparação de registos de extensões e acelerações dos diferentes cenários de desequilíbrios de carga, para a velocidade de 60 km/h.

Através da análise das várias figuras, pode concluir-se que a principal diferença observável entre o tipo de desequilíbrio longitudinal e a sua respetiva *baseline*, é o encurtamento do primeiro pico, permitindo afirmar que estamos perante um desequilíbrio no primeiro vagão. No caso dos desequilíbrios transversais, a análise tem de ser feita de maneira diferente. Como o desequilíbrio é transversal, as principais diferenças ocorrem entre o lado direito e o lado esquerdo do carril, pelo que, são apresentados dois registos, um do sensor do lado direito e outro do sensor do lado esquerdo. Os valores de pico da aceleração e da deformação são mais elevados para o carril direito do que para o carril esquerdo, em consequência da assimetria da carga.

É de salientar que no caso das respostas em extensão, o efeito da velocidade do comboio não tem significado, uma vez que as deformações não são sensíveis à componente dinâmica das ações do comboio.



## CAPÍTULO 4

# METODOLOGIAS PARA IDENTIFICAÇÃO DE DEFEITOS ISOLADOS BASEADAS NUMA REDE NEURONAL DO TIPO *AUTOENCODER*

### 4.1 INTRODUÇÃO

Neste capítulo, são propostas e comparadas duas metodologias que incluem uma rede neuronal do tipo *autoencoder* esparsa, apresentado em pormenor no Capítulo 2, combinando vários métodos e técnicas estatísticas multivariadas, que permitem a classificação de defeitos nos rodados e desequilíbrios de carga, com base em respostas dinâmicas induzidas pela passagem de veículos ferroviários. O princípio das metodologias teve como base o trabalho desenvolvido e testado por Meixedo et al. [25], em que foi utilizado para identificar danos em pontes ferroviárias, e aplicadas em diversos estudos [17, 20, 21, 27]. A diferença entre as duas metodologias prende-se com a utilização do treino na rede *autoencoder*, que numa das metodologias é usado numa fase inicial para a extração de indicadores e na outra metodologia como um passo posterior de normalização dos indicadores previamente extraídos com outras técnicas (PCA e CWT). Assim, será dada ênfase à otimização da classificação de danos isolados de poligonização, dos lisos e dos desequilíbrios de carga, com o intuito de avaliar a eficiência das duas metodologias.

Com recurso ao layout de sensores definidos subcapítulo 3.7.1 (6 acelerómetros e 10 extensómetros), serão abordados os 3 tipos de defeitos considerando apenas uma única velocidade (80 km/h), e posteriormente de forma a aumentar a complexidade do problema, serão consideradas as 3 velocidades em simultâneo (60, 80 e 100 km/h). Foi ainda realizada uma análise de sensibilidade ao número de sensores, que se apresenta no Anexo 1.

O estudo deste capítulo será importante numa fase seguinte, já que as conclusões deste estudo em danos isolados serão importantes para definir a estratégia para classificação dos três tipos de defeito em conjunto (multi-dano) no Capítulo 5.

## 4.2 METODOLOGIA 1

Na metodologia 1, o treino da rede neuronal *autoencoder* é alimentada com dados dos indicadores previamente extraídos com outras técnicas, neste caso, o PCA e o CWT, tendo aqui uma função de normalização desses indicadores. Na Figura 55 apresenta-se o fluxograma relativo à metodologia 1, que ilustra os seus principais passos.

Dependendo do tipo de defeito, os dados de entrada serão os valores transformados (após aplicação de um FFT) das acelerações ou o registo das extensões medidas. Sobre estes dados de entrada, são extraídos indicadores através das técnicas CWT e PCA, etapa i), permitindo uma compressão significativa dos dados iniciais. Estes indicadores extraídos dos sinais originais são posteriormente apresentados à rede dos vários *autoencoders* esparsos empilhados (SSAE). Nesta etapa, os indicadores são treinados pelo SSAE e são obtidos novos indicadores, etapa ii), que são provenientes da camada mais comprimida da rede, a “*Bottleneck layer*”. Na definição dos hiperparâmetros do SSAE é definido o tamanho da camada oculta que não é mais que o número de indicadores novos extraídos. Este número de indicadores foi testado e ajustado para cada um dos três tipos de problema, por forma a melhorar os resultados de cada um. Para além deste hiperparâmetro, também os restantes foram otimizados, como veremos mais a frente. Para concluir a ideia desta etapa, importa referir que a combinação destas técnicas (CWT e PCA) com o SSAE, surgiu com o intuito de tornar o processo de extração mais expedito. Com a combinação destas técnicas, é possível obter-se uma maior rapidez de cálculo e uma melhor remoção dos efeitos operacionais e ambientais, como é a velocidade.

A etapa iii) consiste na fusão dos dados. A fim de melhorar a sensibilidade ao dano, é aplicada a distância de *Mahalanobis* aos indicadores, permitindo uma fusão efetiva, primeiro dos indicadores de cada sensor e, numa segunda fase, dos vários sensores, num índice de dano (DI). É nesta etapa que é possível perceber se é ou não possível obter-se uma boa classificação, uma vez que é já bastante visível a tendência dos dados. É de salientar que nem sempre a fusão de todos os sensores se torna vantajosa, devendo ser avaliada caso a caso.

Por fim, na etapa iv), é aplicada a análise de *clusters*, na qual é utilizada a técnica *k-means*, permitindo efetuar um agrupamento dos dados baseado em centróides e classificar os 3 tipos de defeitos. Como é necessário definir previamente o número de clusters, foi utilizado o índice global de Silhueta (SIL) [40], que permitiu tornar o processo automático.



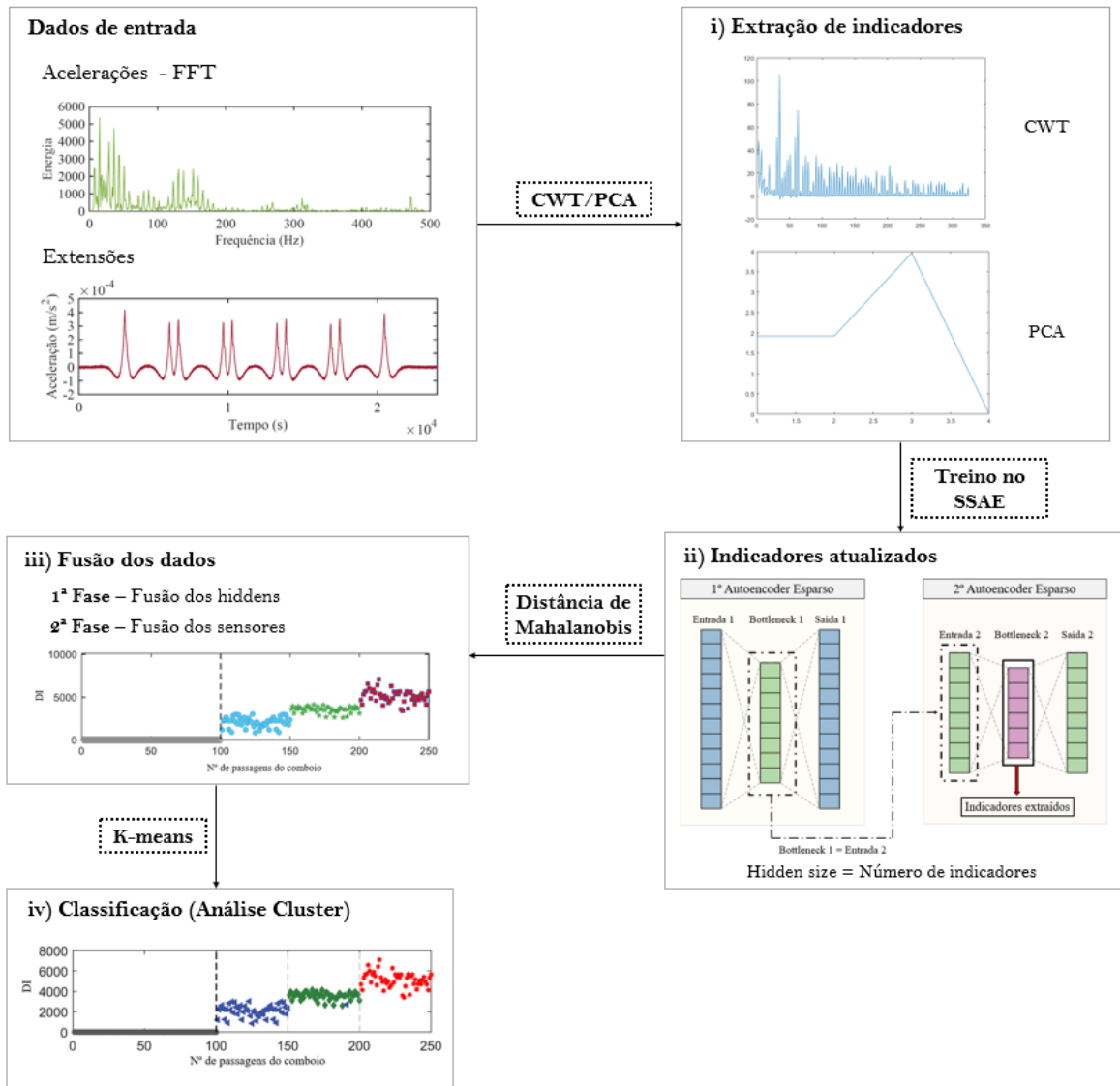


Figura 55 – Fluxograma referente à metodologia 1 com os principais passos da metodologia de identificação de danos isolados.

*i) Extração de indicadores*

Após efetuadas as transformações necessárias dos dados, aplicando no caso das acelerações uma transformada FFT, é realizada a primeira extração de indicadores, através das técnicas do CWT e do PCA. Ao realizar esta etapa de extração, é possível reduzir o número de dados e facilitar o treino na fase seguinte por parte do SSAE, quer em termos de tempo de cálculo quer em termos de interpretação de dados. Dependendo do tipo de defeito, a técnica a utilizar varia. Após realizadas várias análises de sensibilidade às técnicas, verificou-se que o CWT permite melhorar os resultados no caso dos danos nos rodados, enquanto o PCA é mais eficaz para os desequilíbrios de carga.

De forma a evitar a repetição de várias figuras e manter a objetividade, serão apenas apresentados os resultados obtidos para as etapas i), ii) e iii) para o caso dos lisos, sendo que na etapa final, etapa iv), serão

apresentados os resultados obtidos para os 3 defeitos. Sendo assim, na Figura 56, é possível observar um exemplo de um indicador extraído para o cenário de dano de lisos.

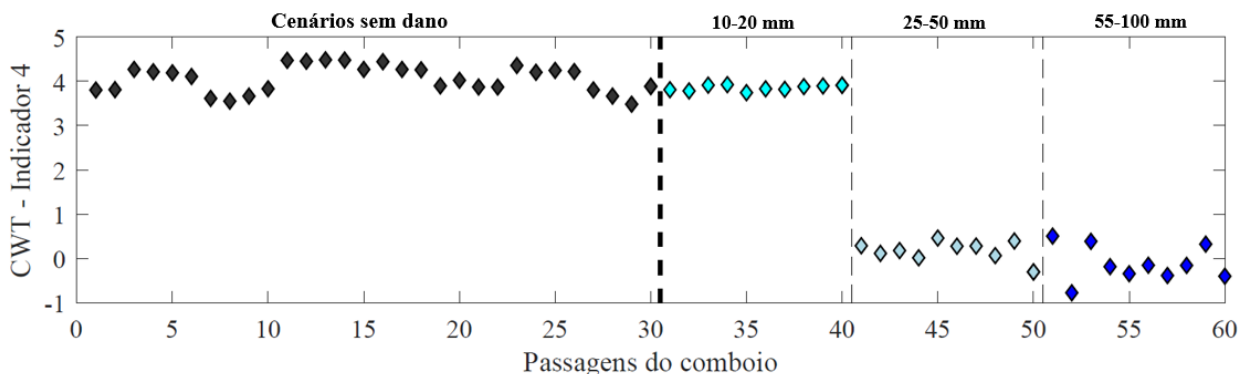


Figura 56 – Extração de indicadores para uma velocidade a partir do acelerómetro 1 referente aos cenários de dano de lisos.

Desde logo, após uma análise cuidadosa da figura, é evidente uma tendência de separação entre as diferentes severidades dos danos, embora ainda não exatamente da maneira desejada. É possível observar que a severidade intermédia e a mais severa se destacam em relação aos casos sem dano e ao grupo menos severo. A partir desta observação, é possível antever que a distinção entre o dano de 10-20 mm e a linha de base possa ser a mais desafiadora. No entanto, a tendência desejada está presente.

Para cada sensor, é criada uma matriz  $n \times m$ , em que  $n$  é o número de passagens de teste (cenários de base + cenários de danos) e  $m$  representa o número de características extraídas. Assim, considerando apenas a velocidade de 80 km/h, o tamanho das matrizes são:

- Matriz (Roda plana):  $60 \times 324$ ;
- Matriz (Roda Poligonal):  $60 \times 324$ ;
- Matriz (Desequilíbrio de carga):  $42 \times 4$ .

ii) *Indicadores atualizados*

Após realizada a extração de indicadores, estes dão depois entrada no treino do SSAE. A aplicação do SSAE introduz restrições através da utilização de hiperparâmetros de esparsidade (já mencionados no capítulo 2). Ao ajustar estes hiperparâmetros, é possível tornar o algoritmo mais eficiente e melhorar as capacidades de generalização, uma vez que apenas as características mais importantes são mantidas. A sua otimização é realizada através de um processo iterativo que requer vários testes e depende do tipo de dados em análise. Para este efeito, com base no treino de apenas cenários sem dano, foram obtidos os hiperparâmetros otimizados para a velocidade de 80 km/h. A Tabela 20 apresenta esses mesmos hiperparâmetros para o caso da poligonização, para os lisos e para os desequilíbrios de carga.

Tabela 20 – Hiperparâmetros utilizados no SSAE para 1 velocidade (80 km/h) e para os 3 tipos de dano.

Hiperparâmetros de treino no SSAE	Poligonização (Harmônicos)	Lisos	Desequilíbrios de carga
Função de transferência de codificação	logsig	logsig	logsig
Função de transferência de decodificação	logsig	logsig	logsig
Tamanho da camada oculta (1ª/2ª camadas)	500/250	750/500	650/400
Número máximo de épocas (1ª/2ª camadas)	50/300	2000/2000	2000/2000
Regularização de pesos L2 (1ª/2ª camadas) ( $\lambda$ )	$1 \times 10^{-4}/1 \times 10^{-6}$	$1 \times 10^{-4}/1 \times 10^{-6}$	$0.5/1 \times 10^{-2}$
Regularização de esparsidade (1ª/2ª camadas) ( $\beta$ )	$1 \times 10^{-1}/8$	$1 \times 10^{-1}/8$	$1 \times 10^{-4}/1$
Proporção de esparsidade (1ª/2ª camadas) ( $\rho$ )	$1 \times 10^{-2}/1 \times 10^{-3}$	$1 \times 10^{-2}/1 \times 10^{-3}$	$1 \times 10^{-4}/1 \times 10^{-3}$

Após a realização de vários testes em que foram alteados os hiperparâmetros baseados nos intervalos definidos pelo HelpCenter do MATLAB®, provou-se que a utilização de 2 camadas de autoencoder's no SSAE seria suficiente para os objetivos que se pretendiam atingir. As funções de transferência não lineares (logsig) revelaram-se mais eficazes do que as lineares, razão pela qual foram utilizadas em todos os casos estudados. Analisando a Tabela 20 ao pormenor, comparando os dois danos OOR (poligonização e lisos), verifica-se que as principais diferenças surgem no tamanho da camada oculta e no número máximo de épocas. Apesar de os registos em FFT dos lisos serem bastante diferentes entre as várias severidades, o SSAE não consegue discernir adequadamente a relação entre os dados, exigindo assim um maior período de aprendizagem e, conseqüentemente, apresentando valores mais elevados para os hiperparâmetros. semelhança dos lisos, também os desequilíbrios de carga requerem um treino mais rigoroso para uma boa classificação. Contudo, neste caso, uma vez que não estamos a lidar com um dano OOR e os dados de entrada são os registos de extensões e não valores do FFT, observa-se uma maior variação de valores em quase todos os hiperparâmetros. Desde logo, percebe-se que isso pode representar um ponto negativo desta metodologia, pois existem diversos tipos de danos que, se não forem estudados e calibrados na metodologia, podem ser classificados de forma inadequada. Após a definição dos hiperparâmetros e da realização do treino no SSAE, são obtidos indicadores atualizados, provenientes da segunda camada do SSAE, conforme mostra a Figura 57.

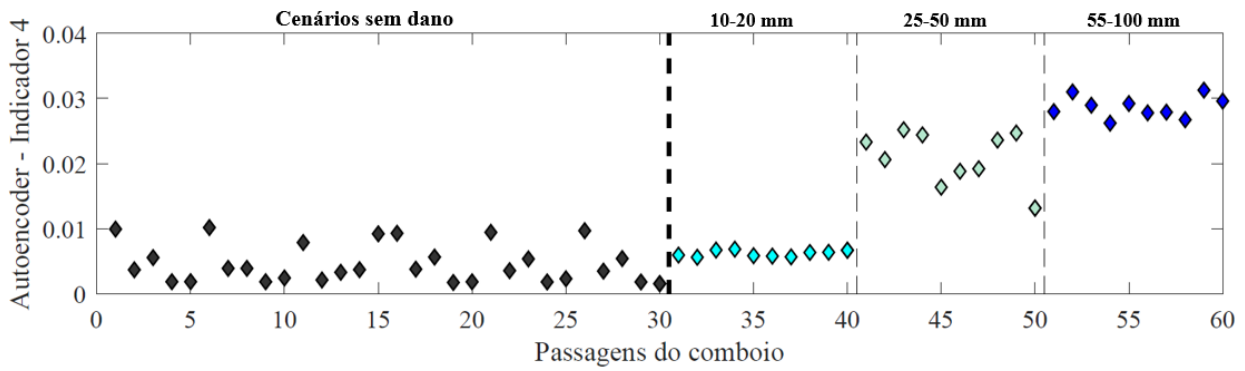
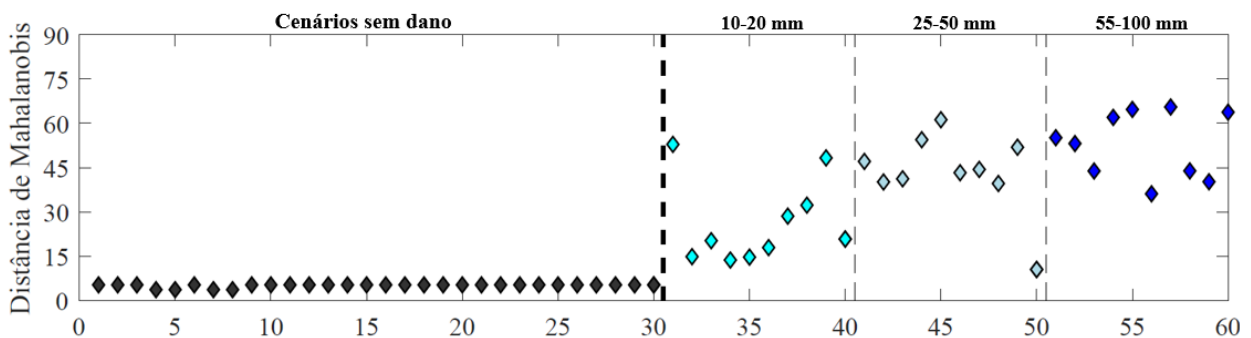


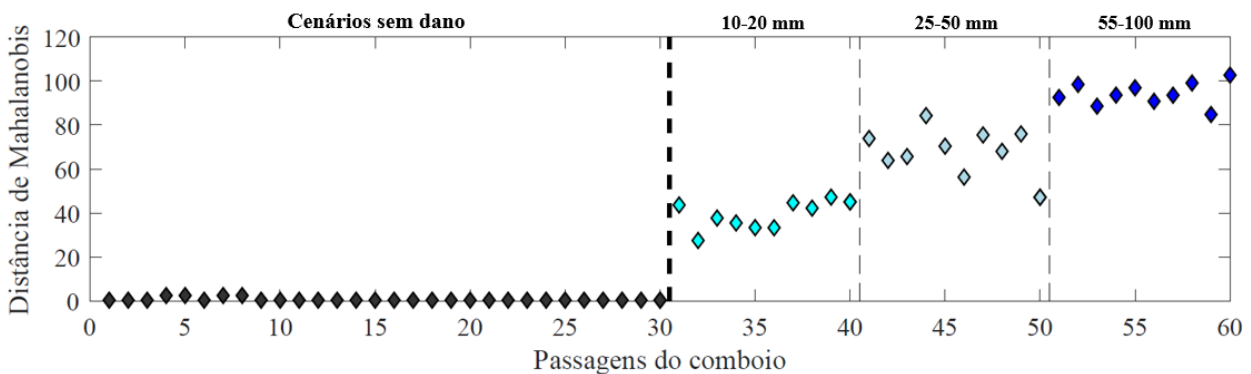
Figura 57 – Extração de indicadores para 1 velocidade a partir do acelerómetro 1 referente aos cenários de dano de lisos.

iii) Fusão dos dados

Para aumentar as distâncias entre os cenários de base e de dano e assim conseguir uma melhor separação entre as várias severidades dos defeitos, é realizada uma fusão de dados usando a distância de Mahalanobis, em dois níveis, de indicadores e de sensores. Na Figura 58 a), é apresentado o primeiro nível de fusão para o caso do sensor 2, com resultados favoráveis. No segundo nível de fusão, Figura 58 b), os resultados são significativamente melhorados. Nesta última figura, é possível identificar claramente três grupos distintos, sugerindo uma eficaz classificação.



a)



b)

Figura 58 – Fusão de dados para 1 velocidade referente aos lisos: a) 1º nível - fusão de características para o sensor 2 b) 2º nível - fusão de todos os sensores.

iv) Classificação (Análise Cluster)

Com recurso aos *outputs* da 2ª fusão dos dados, é realizada a classificação das severidades dos vários defeitos, recorrendo-se à análise de *clusters*. Com auxílio da técnica *k-means* foi realizada a classificação da poligonização em relação ao número de harmónicos (Figura 59), dos lisos em relação ao comprimento do liso (Figura 60) e dos desequilíbrios de carga em relação a desequilíbrios transversais e longitudinais (Figura 61). Com a Figura 59, conclui-se que a classificação em termos de número de harmónicos apresenta 6 falsas classificações, representando 80 % de taxa de acerto. A maioria das falsas classificações encontram-se no grupo H6-8, onde os casos deste grupo estão a ser confundidos com o grupo H12-14.

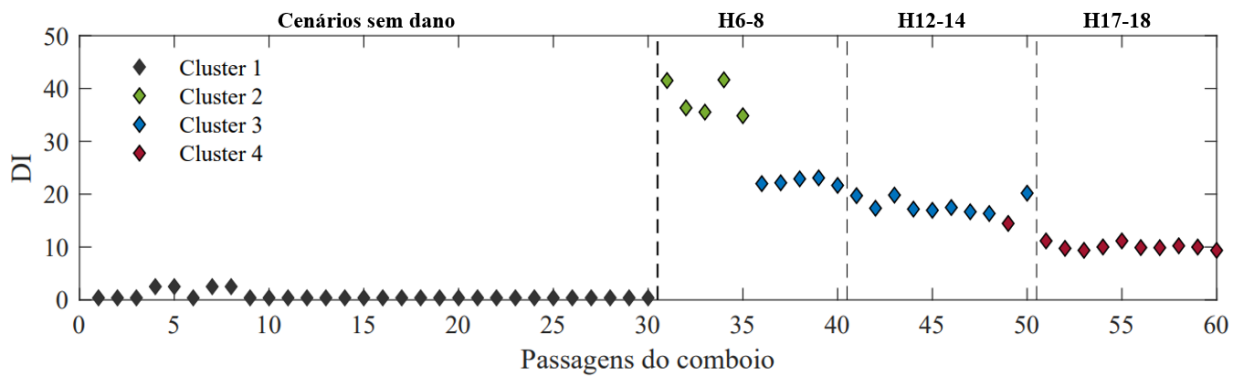


Figura 59 – Classificação automática da severidade da poligonização com base na análise de clusters para 1 velocidade, utilizando 6 acelerómetros (3 de cada lado).

Relativamente aos resultados associados aos lisos, Figura 60, são obtidas 2 falsas classificações no grupo intermédio, o que ainda assim representa mais de 95 % de acerto. Nesta figura é ainda possível constatar que o grupo menos severo apresenta um DI ainda distante da linha de base, apesar de estarmos perante um dano inicial, o grupo de 10-20 mm, o que prova que a metodologia é suficientemente robusta para distinguir este dano inicial. É também de salientar que a filtragem do sinal tem uma grande influência na identificação dos lisos, ao contrário do que acontece com a poligonização, uma vez que a gama de frequências de interesse para este tipo de dano está entre os 0 e 1500 Hz.

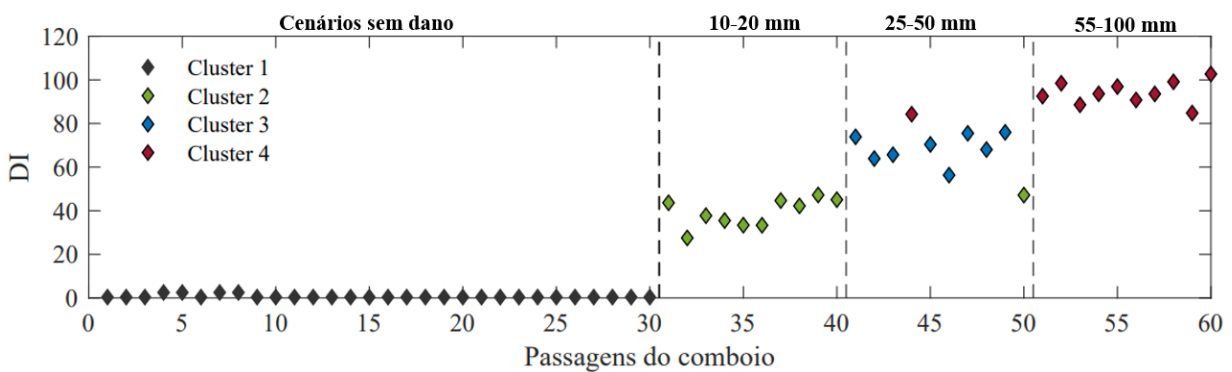


Figura 60 – Classificação automática da severidade dos lisos com base na análise de clusters para 1 velocidade, utilizando 6 acelerómetros (3 de cada lado).

Por fim, na Figura 61, consegue-se distinguir diferentes combinações de casos desequilibrados. É possível assim separar os cenários de desequilíbrio em dois grandes grupos: desequilíbrios longitudinais (DI entre  $0.5 \times 10^{13}$  e  $2 \times 10^{13}$ ) e desequilíbrios transversais (DI entre  $4 \times 10^{13}$  e  $8 \times 10^{13}$ ). No entanto, dentro dos desequilíbrios transversais, é ainda possível identificar dois subgrupos distintos, nomeadamente o desequilíbrio transversal 1 e o desequilíbrio transversal 2, sendo que apenas ocorre 1 classificação errada no caso do desequilíbrio transversal 2. É de salientar que no caso do desequilíbrio longitudinal, a distinção entre as duas severidades não é conseguida.

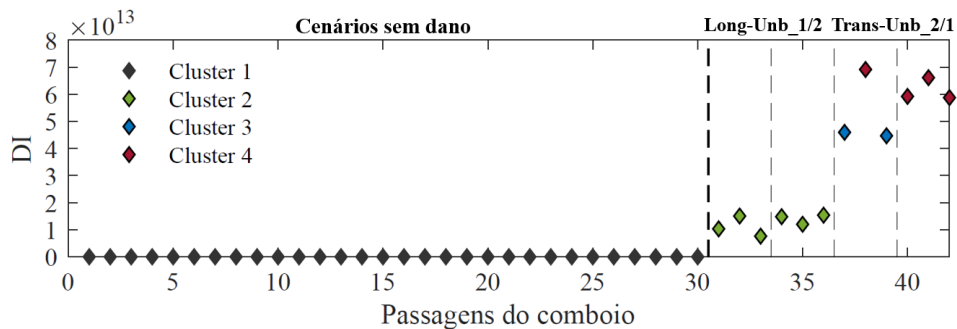


Figura 61 – Classificação automática da severidade dos desequilíbrios de carga com base na análise de clusters para 1 velocidade, utilizando 5 extensómetros do lado do desequilíbrio.

### 4.3 METODOLOGIA 2

Nesta segunda metodologia, a introdução do SSAE, ao contrário do que acontecia na metodologia 1, dá-se logo numa fase inicial como técnica de extração dos indicadores. Aqui, o SSAE é aplicado diretamente aos registos medidos, para extrair os indicadores. Visto que o treino é feito com base nos registos e não nos indicadores CWT e PCA, os indicadores obtidos nesta metodologia serão diferentes, assim como o tempo de cálculo, que é significativamente superior, devido ao elevado número de valores contido no sinal de entrada. Na Figura 62 apresenta-se o fluxograma relativo à metodologia 2.

De forma a facilitar a comparação entre as duas metodologias, a estrutura da apresentação dos resultados usada para a metodologia 1 é agora também adotada para a metodologia 2.

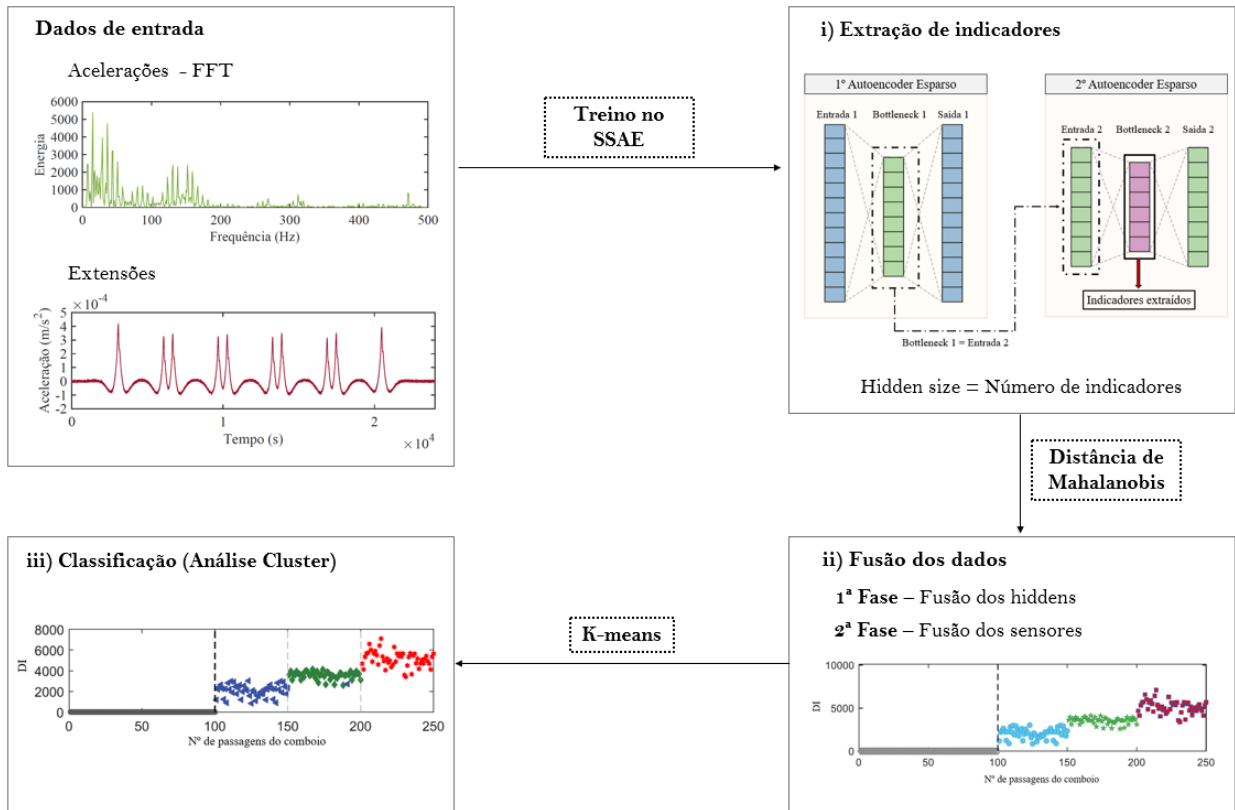


Figura 62 – Fluxograma referente à metodologia 2 com os principais passos da metodologia de identificação de danos isolados.

i) *Extração de indicadores*

Após a aplicação do SSAE diretamente nos FFT's e nas extensões, é realizada uma extração de indicadores, que são provenientes do "Bottleneck layer". Na Figura 63 é possível observar um exemplo de um desses indicadores obtidos. À semelhança do que acontecia na metodologia 1, o grupo 25-50 mm e 50-100 mm destacam-se do grupo menos severo e dos cenários sem dano. Ao comparar este indicador com o da Figura 56, verifica-se uma acentuada diminuição na escala e os casos de dano deslocam-se para posições superiores aos cenários sem dano.

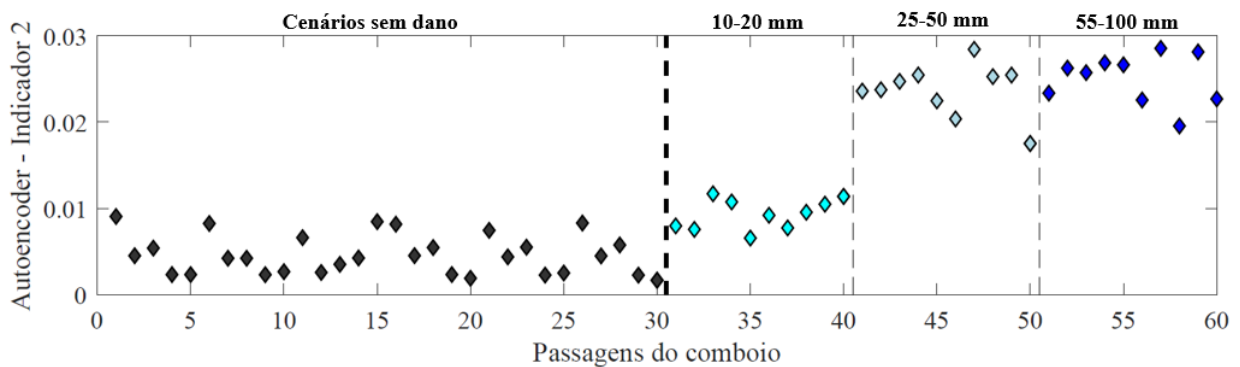


Figura 63 – Extração de indicadores para 1 velocidade a partir do acelerómetro 6 referente aos cenários de dano de lisos.

Como o tamanho da camada oculta equivale ao número de indicadores, o tamanho das matrizes obtidas após a extração dos indicadores são:

- Matriz (Roda plana):  $60 \times 500$ ;
- Matriz (Roda Poligonal):  $45 \times 500$ , para cada amplitude;
- Matriz (Desequilíbrio de carga):  $42 \times 50$ .

Para ocorrer este processo de extração de indicadores, são definidos os hiperparâmetros de treino. Como os dados de treino são diferentes dos da metodologia 1, os hiperparâmetros tiveram de ser atualizados e otimizados. A Tabela 21 apresenta os valores dos hiperparâmetros para o caso da poligonização, para os lisos e para o desequilíbrio de cargas.

Tabela 21 – Hiperparâmetros utilizados no SSAE para 1 velocidade (80 km/h) e para os 3 tipos de dano.

Hiperparâmetros de treino no SSAE	Poligonização (Harmônicos)	Lisos	Desequilíbrios de carga
Função de transferência de codificação	logsig	logsig	logsig
Função de transferência de decodificação	logsig	logsig	logsig
Tamanho da camada oculta (1ª/2ª camadas)	1500/500	1500/500	100/50
Número máximo de épocas (1ª/2ª camadas)	50/300	50/200	2000/2000
Regularização de pesos L2 (1ª/2ª camadas) ( $\lambda$ )	$1 \times 10^{-4}/1 \times 10^{-6}$	$1 \times 10^{-4}/1 \times 10^{-6}$	$1 \times 10^{-4}/1 \times 10^{-6}$
Regularização de esparsidade (1ª/2ª camadas) ( $\beta$ )	$1 \times 10^{-1}/8$	$1 \times 10^{-1}/1$	$1 \times 10^{-2}/8$
Proporção de esparsidade (1ª/2ª camadas) ( $\rho$ )	$1 \times 10^{-2}/1 \times 10^{-3}$	$1 \times 10^{-2}/1 \times 10^{-3}$	$1 \times 10^{-2}/1 \times 10^{-3}$

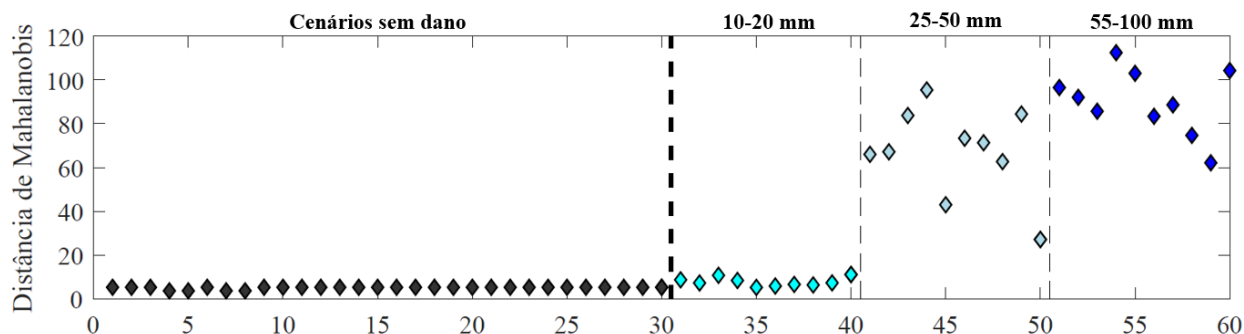
À semelhança da metodologia 1, o número de camadas ocultas utilizadas no SSAE são duas, assim como as funções transferência utilizadas são as não lineares (logsig). Analisando a Tabela 21, verifica-se uma semelhança entre os vários hiperparâmetros dos dois danos nos rodados, com uma ligeira alteração no tamanho da segunda camada oculta e no termo de regularização de esparsidade também da segunda camada. Por sua vez, comparando estes hiperparâmetros com os definidos na metodologia 1, verifica-se que neste caso, os tamanhos das camadas ocultas são superiores. Isso deve-se ao tamanho dos dados de entrada, que na metodologia tinham uma dimensão máxima de 324 pontos e nesta metodologia tem uma dimensão de 7000 pontos. Nesse sentido, é necessário aumentar o valor do tamanho da camada oculta para que o algoritmo consiga compreender todas as relações entre os dados de forma correta. No que toca aos desequilíbrios de carga, verificou-se que o tamanho da camada oculta não teve influência nos resultados. Os restantes parâmetros são bastante parecidos ao já adotados anteriormente.

#### ii) Fusão dos dados

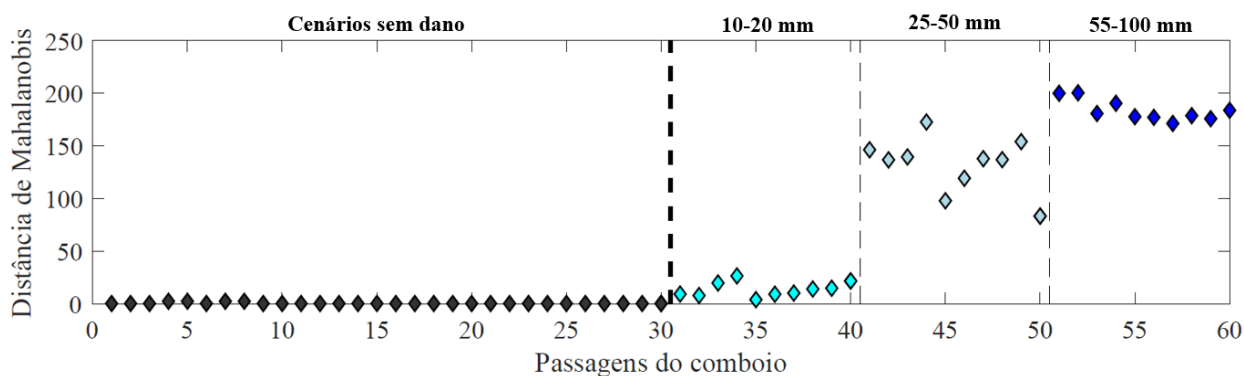
Com base nos indicadores obtidos através da segunda camada do SSAE é realizada a mesma fusão dos dados que foi realizada na metodologia 1. Na Figura 64 a), é apresentada a fusão de indicadores para o



caso do sensor 2, enquanto na Figura 64 b), é demonstrada a fusão de sensores. Os resultados obtidos mostram uma separação grande entre um dano inicial e os casos mais severos de lisos, mas a separação é menos efetiva entre a *baseline* e os danos iniciais e entre as duas maiores severidades.



a)

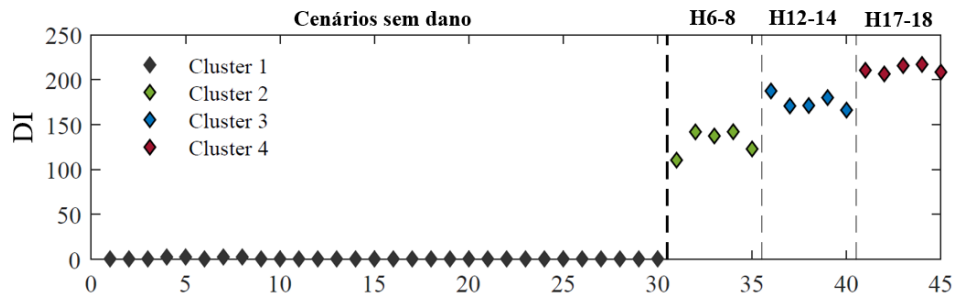


b)

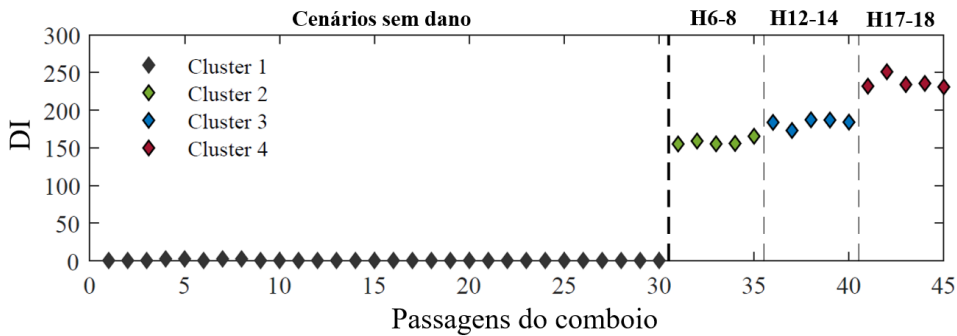
Figura 64 – Fusão de dados para uma velocidade referente aos lisos: a) 1º nível - fusão de características para o sensor 2 b) 2º nível - fusão de todos os sensores.

### iii) Classificação (Análise Cluster)

Sobre os dados obtidos na etapa anterior é realizada a classificação com base numa análise de *clusters*. Para os cenários de poligonização, Figura 65, a classificação em termos de harmónicos não é possível incorporando simultaneamente os dois grupos de amplitude, já que a metodologia acaba por se confundir. Este requisito acaba por ser uma desvantagem desta abordagem, uma vez que na metodologia 1, todas as amplitudes de vários tipos foram agrupadas e a classificação foi bem-sucedida. Ainda assim, se os casos de amplitudes mais baixas (A1) forem analisados separadamente, assim como os de amplitudes mais altas (A2), a classificação torna-se perfeita, como pode ser observado na Figura 65.



a)



b)

Figura 65 – Classificação automática da severidade da poligonização com base na análise de clusters para uma velocidade, utilizando 6 acelerómetros (3 de cada lado): a) Amplitude 1 b) Amplitude 2.

Comparado a Figura 65 com a Figura 59, verifica-se uma alteração na disposição dos grupos, já que na metodologia 1 o valor de DI é superior no grupo H6-8 e inferior no H17-18, e agora, a tendência é o inverso, com um valor superior no caso do grupo H17-18 e inferior no H6-8.

Relativamente à classificação dos danos de lisos, Figura 66, estes são muito semelhantes aos obtidos na primeira abordagem. São obtidas 2 falsas classificações, mas neste caso, devido à proximidade do grupo 10-20 mm dos cenários sem dano.

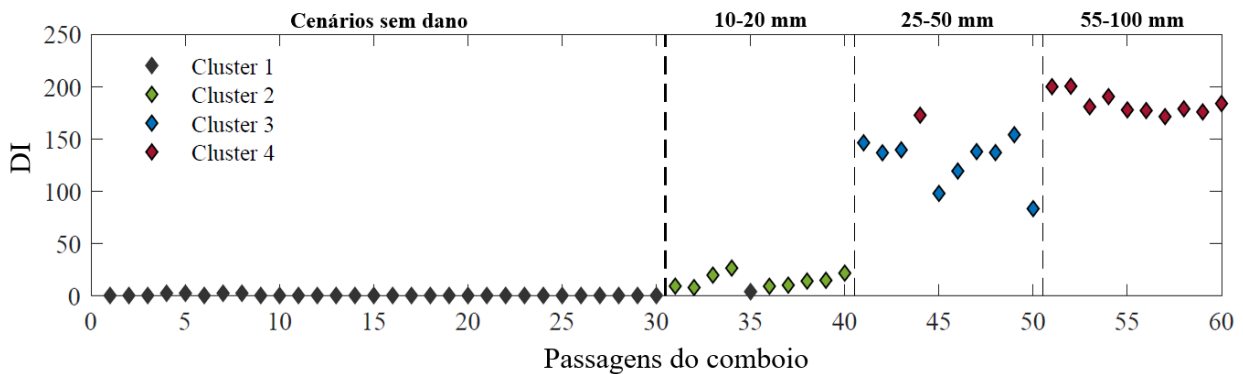


Figura 66 – Classificação automática da severidade dos lisos com base na análise de clusters para uma velocidade, utilizando 6 acelerómetros (3 de cada lado).

Na Figura 67 é apresentada a classificação relativa aos resultados obtidos para os desequilíbrios de carga. Á primeira vista, os resultados são muito semelhantes aos obtidos para a metodologia 1, com a exceção de que agora o valor de DI é superior nos casos dos desequilíbrios longitudinais. Apesar disso, com o cluster automático, são definidos 3 grupos de desequilíbrios, sendo um grupo os longitudinais e dois grupos para as duas severidades de desequilíbrios transversais.

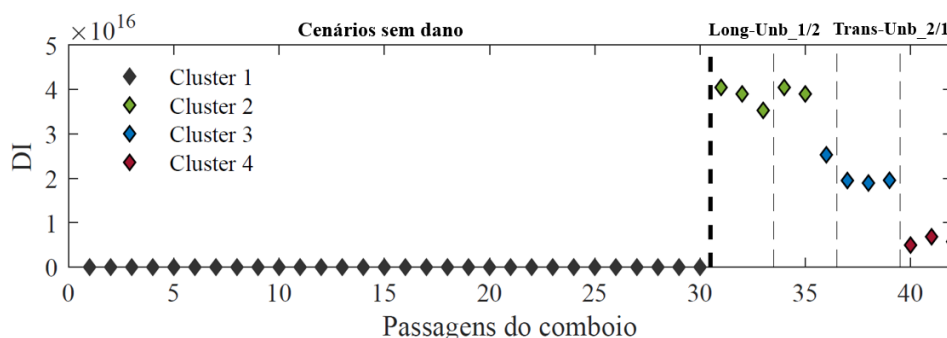


Figura 67 – Classificação automática da severidade dos desequilíbrios de carga com base na análise de clusters para uma velocidade, utilizando 10 extensómetros.

## 4.4 ANÁLISE DE SENSIBILIDADE

Foi realizada uma análise de sensibilidade à velocidade e ao número de sensores, contudo para facilitar na compreensão do documento, neste subcapítulo apenas se apresenta a análise para a velocidade. A análise ao número de sensores pode ser encontrada no Anexo 1.

### 4.4.1 Velocidade

Com o objetivo de aumentar a complexidade do estudo, foram incluídas novas velocidades aos casos de dano (60 e 100 km/h). Como o treino é realizado com base em cenários sem dano, os dados de treino continuam a ser os mesmos, tendo-se apenas aumentado o número de dados de teste (cenários de dano).

#### 4.4.1.1 Metodologia 1

Considerando a metodologia 1 e os parâmetros definidos na Tabela 20, procedeu-se a aplicação dos passos definidos nas etapas ii) iii). Contudo, para isso, foram primeiramente definidas novas matrizes de dados e extraídos novos indicadores CWT e PCA, já que temos um maior volume de informação.

- Matriz (Roda plana):  $120 \times 324$ ;
- Matriz (Roda Poligonal):  $120 \times 324$ ;
- Matriz (Desequilíbrio de carga):  $66 \times 4$ .

iv) Classificação (Análise Cluster)

No caso da poligonização, a classificação apresentada na Figura 68 mostra que considerando as 3 velocidades se verifica que a metodologia é capaz de fazer uma distinção clara entre os três níveis de gravidade dos danos, à semelhança do que já se obtinha usando apenas uma única velocidade.

Dado os hiperparâmetros utilizados nas três velocidades serem os mesmos que os utilizados na análise de única velocidade, os valores de DI mantiveram-se semelhantes, conforme se apresenta na Figura 68. Para este tipo de defeito, registaram-se 4 classificações incorretas, sendo estas referentes à velocidade de 80 km/h, precisamente as mesmas que já eram mal classificadas na Figura 59.

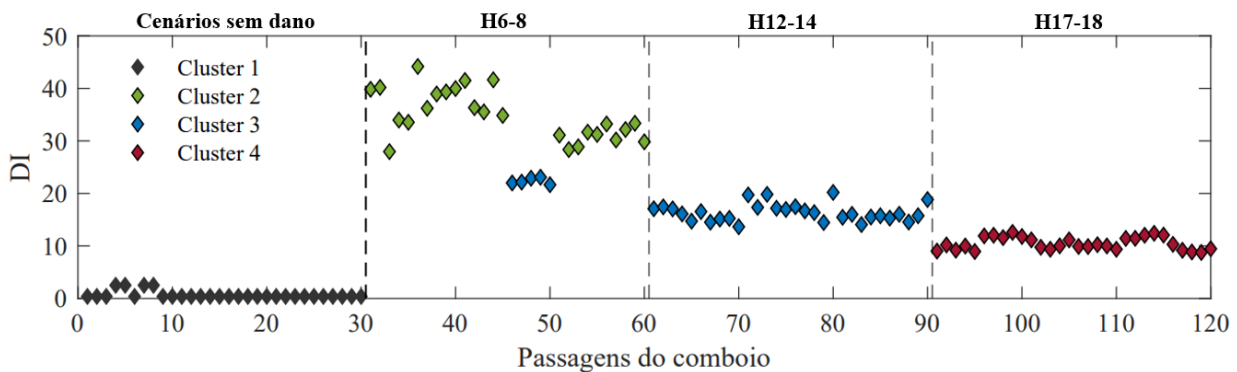


Figura 68 – Classificação automática da severidade da poligonização com base na análise de clusters para 3 velocidades, utilizando 6 acelerómetros (3 de cada lado).

Para os cenários de dano de lisos, a classificação obtida na Figura 69 mostra uma boa eficácia, com uma taxa de acerto superior a 90 %.

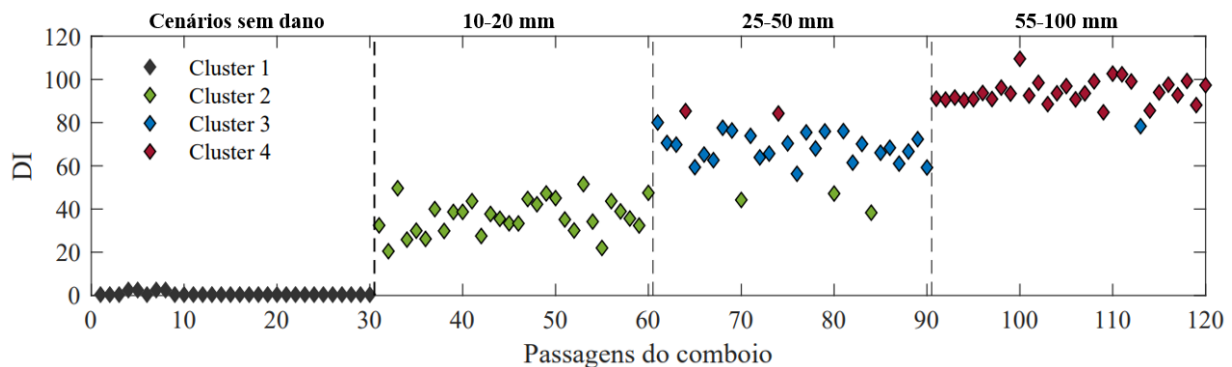


Figura 69 – Classificação automática da severidade dos lisos com base na análise de clusters para 3 velocidades, utilizando 6 acelerómetros (3 de cada lado).

Por fim, na Figura 70, os desequilíbrios de cargas são classificados em 3 grupos, com apenas 4 falsas classificações, em que a maioria das falhas é no desequilíbrio transversal 2. Ainda assim, se se considerar uma classificação apenas em 2 grupos, é obtida uma classificação quase perfeita entre desequilíbrio longitudinal e transversal, com apenas 1 falsa classificação.

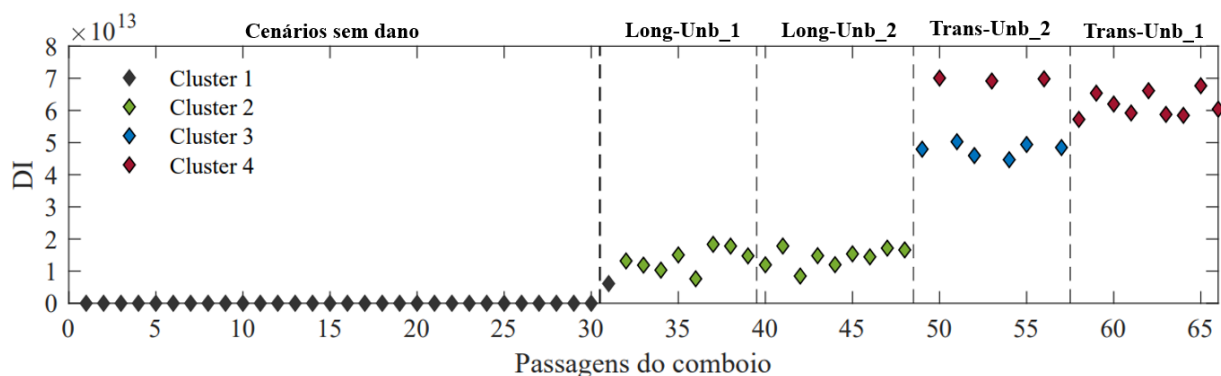


Figura 70 – Classificação automática da severidade dos desequilíbrios de carga com base na análise de clusters para 3 velocidades, utilizando 5 extensómetros do lado do desequilíbrio.

#### 4.4.1.2 Metodologia 2

Ao contrário do que acontecia anteriormente com a metodologia 1, em que os parâmetros de treino usados na análise anteriormente para uma única velocidade foram os mesmos, neste caso, na metodologia 2, isto já não acontece. No caso da poligonização e dos desequilíbrios de carga, a alteração dos hiperparâmetros não se mostrou vantajosa, daí permanecerem inalterados. Pelo contrário, no caso dos lisos, estes foram novamente otimizados.

Desta forma, na Tabela 22, são demonstrados os hiperparâmetros utilizados para esta análise considerando 3 velocidades para o caso dos lisos. As alterações foram mínimas e ocorreram apenas no número de épocas da segunda camada e na regularização da esparsidade, igualmente na segunda camada.

Tabela 22 – Hiperparâmetros utilizados no SSAE da metodologia 2 para 3 velocidades.

Hiperparâmetros de treino no SSAE	Lisos
Função de transferência de codificação	logsig
Função de transferência de decodificação	logsig
Tamanho da camada oculta (1ª/2ª camadas)	1000/500
Número máximo de épocas (1ª/2ª camadas)	50/300
Regularização de pesos L2 (1ª/2ª camadas) ( $\lambda$ )	$1 \times 10^{-4} / 1 \times 10^{-6}$
Regularização de esparsidade (1ª/2ª camadas) ( $\beta$ )	$1 \times 10^{-1} / 8$
Proporção de esparsidade (1ª/2ª camadas) ( $\rho$ )	$1 \times 10^{-2} / 1 \times 10^{-3}$

As matrizes com os dados novos podem ser definidas da seguinte forma:

- Matriz (Roda plana):  $120 \times 500$ ;
- Matriz (Roda Poligonal):  $75 \times 500$ , para cada amplitude;
- Matriz (Desequilíbrio de carga):  $66 \times 50$ .

iii) Classificação (Análise Cluster)

Em relação à poligonização, a Figura 71 evidencia uma boa classificação de todos os grupos de harmônicos, tanto para os casos de amplitude A1 como amplitude A2. A taxa de acerto é de 100 % considerando os casos de amplitude A1, e superior a 90 % considerando os casos de amplitude A2.

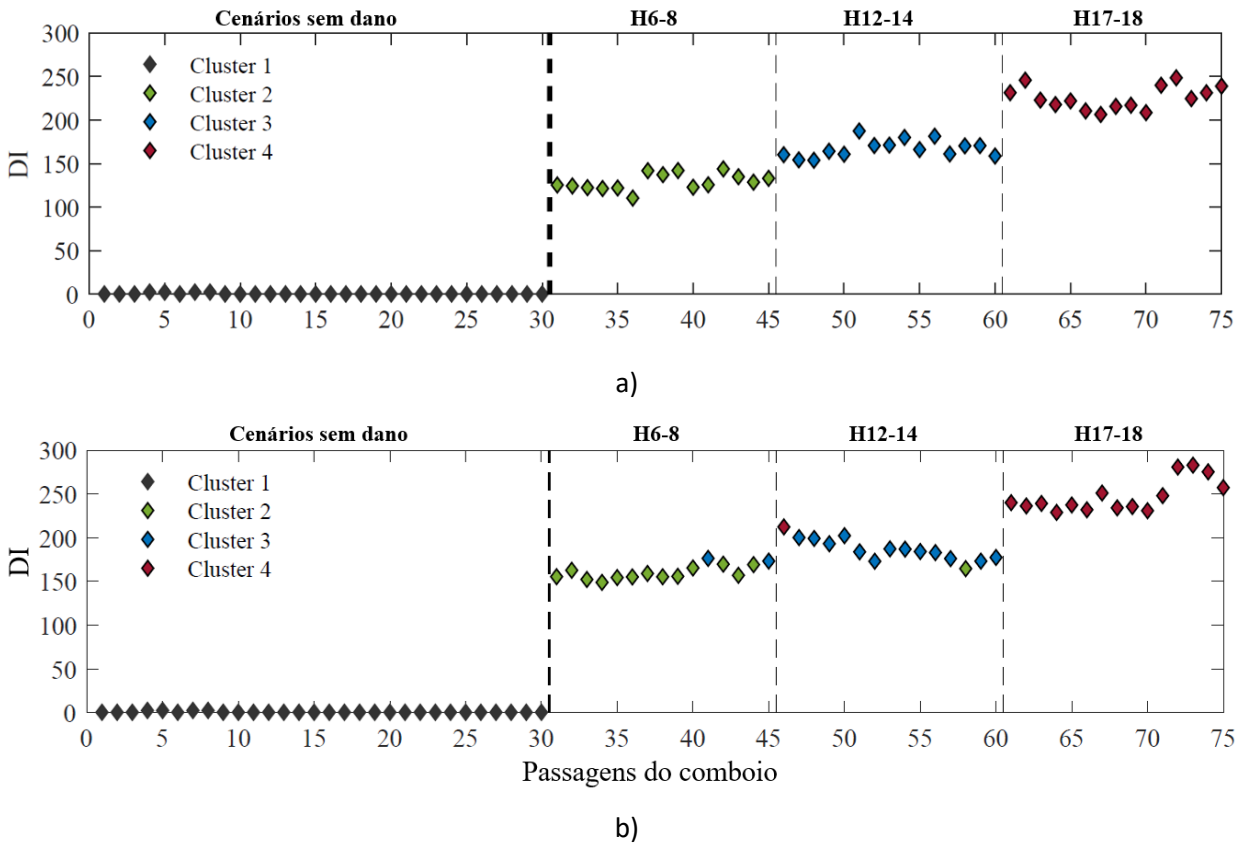


Figura 71 – Classificação automática da severidade da poligonização com base na análise de clusters para 3 velocidades, utilizando 6 acelerómetros (3 de cada lado).

Relativamente aos resultados dos lisos, ilustrados na Figura 72, são classificados incorretamente 24 casos (cerca de 75 % de acerto), sobretudo no grupo menos severo, devido à sua proximidade com os cenários sem dano. Comparando com a poligonização, verifica-se uma maior dificuldade na classificação dos lisos.

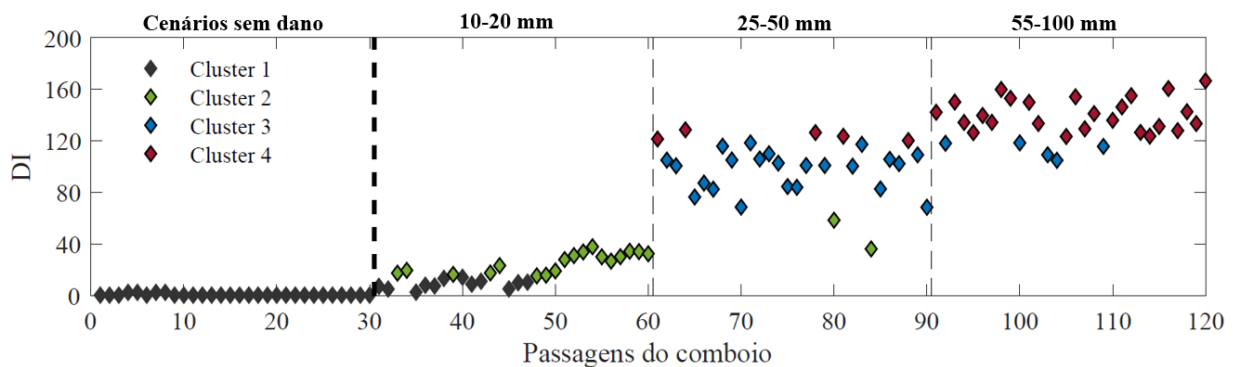


Figura 72 – Classificação automática da severidade dos lisos com base na análise de clusters para 3 velocidades, utilizando 6 acelerómetros (3 de cada lado).

Embora a velocidade não afete os registos em extensões, esta metodologia apresenta uma enorme dificuldade na classificação dos desequilíbrios de carga, conforme indicado na Figura 73. Uma explicação possível para a incapacidade dos indicadores extraídos do *autoencoder* em distinguir de forma nítida os distintos grupos, pode estar associada às diferenças de amplitudes medidas em extensões que são bastante pequenas, ao contrário do que acontece com os valores de FFT nos danos dos rodados. Como na primeira abordagem, o treino no SSAE é baseado em indicadores que já apresentavam uma tendência para a separação desses grupos, torna-se mais simples fazer a distinção, mesmo quando as diferenças são diminutas.

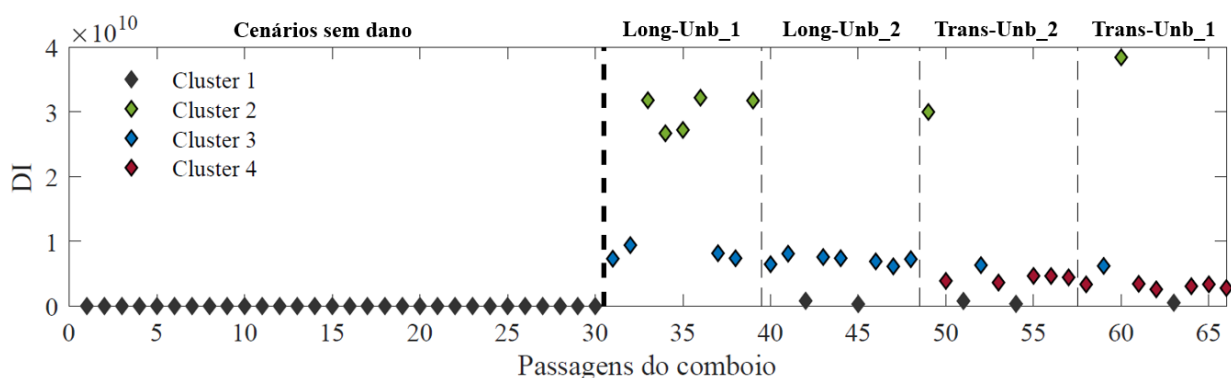


Figura 73 – Classificação automática da severidade dos desequilíbrios de carga com base na análise de clusters para 3 velocidades, utilizando 10 extensómetros.

De uma maneira geral, verifica-se que a metodologia 1 consegue na maioria dos casos apresentar melhores percentagens de acerto, exceto no caso da poligonização. Por sua vez, considerando as 3 velocidades, os resultados mantêm-se muito bons usando ambas as metodologias, embora em geral piorem ligeiramente. No caso específico dos desequilíbrios de carga, os resultados obtidos na metodologia 2 não são os desejados, com elevadas percentagens de falsas classificações.

Tabela 23 – Falsas classificações relacionadas com a análise de sensibilidade da velocidade.

Velocidade	Metodologia 1			Metodologia 2		
	Poligonização (Harmónicos)	Lisos	Desequilíbrio de cargas	Poligonização (Harmónicos)	Lisos	Desequilíbrio de cargas
80 km/h	20% (6/30)	6,7% (2/30)	8,3% (1/12)	0% (0/30)	10% (3/30)	8,3% (1/12)
60, 80, 100 km/h em simultâneo	5,6% (5/90)	6,7% (6/90)	11,1% (4/36)	4,5% (4/90)	26,7% (24/90)	-

## 4.5 DISCUSSÃO DOS RESULTADOS

Nesta secção é apresentada uma breve discussão dos resultados obtidos com as duas metodologias propostas usando uma rede neuronal *autoencoder* para a identificação dos três tipos de defeitos analisados separadamente. Numa primeira análise, as metodologias foram aplicadas considerando apenas cenários de dano para uma única velocidade (80 km/h), e tanto a metodologia 1 quanto a metodologia 2 demonstraram uma notável robustez na classificação dos defeitos. O nível de acerto aproxima-se dos 100 %, com a exceção da poligonização usando a metodologia 1, onde as percentagens baixam, mas ainda assim são resultados muito satisfatórios.

À medida que a complexidade aumentou com a incorporação de duas velocidades adicionais (60 km/h e 100 km/h), verificou-se que tanto a metodologia 1 quanto a metodologia 2 continuaram a exibir excelentes desempenhos na classificação de lisos e poligonização. No entanto, quando analisada a capacidade de lidar com desequilíbrios de carga, o comportamento de ambas as abordagens é distinto. A metodologia 1 demonstrou uma semelhante eficácia, ao passo que a metodologia 2 apresentou uma distribuição menos coerente dos casos, e conseqüentemente obtendo-se bastantes falhas na classificação.

Relativamente ao número de sensores utilizados, a análise indicou que o número ideal para a classificação de defeitos OOR seria de 6 acelerómetros, enquanto para a classificação de desequilíbrios de carga, 10 extensómetros. À medida que se reduz este número de sensores, verificou-se que ambas as metodologias demonstraram uma perda significativa de sensibilidade, impactando negativamente os resultados obtidos. Estes resultados podem ser consultados nas tabelas em anexo, tanto para a metodologia 1 como para a metodologia 2.

De uma forma geral, é possível constatar que ambas as metodologias propostas são adequadas para detetar defeitos nos rodados, mas no caso dos desequilíbrios de carga, a metodologia 1 é mais eficaz. Devido a esta limitação na metodologia 2, e considerando que instabilidades deste tipo afetam com bastante severidade a circulação dos veículos de mercadorias, podemos concluir que a metodologia 1 é mais abrangente e adequada para uma aplicação prática de identificação de defeitos em veículos de mercadorias. Os resultados mostram que a inclusão de uma rede neuronal SSAE numa metodologia de identificação de defeitos permite ter níveis de eficiência muito elevados para vários tipos de defeitos e condições operacionais, nomeadamente a questão da velocidade de circulação.

É ainda importante destacar que a metodologia 1 utiliza uma entrada de dados mais reduzida (indicadores previamente calculados com outras técnicas) para o treino do SSAE, ao contrário da metodologia 2, que utiliza os registos completos com milhares de pontos. Essa diferença tem um impacto considerável no



tempo de cálculo de cada metodologia, pelo que a metodologia 1 permite obter resultados em questão de horas, ao passo que a metodologia 2 pode exigir vários dias para concluir o processo.

Após todas essas análises, conclui-se que a metodologia 1 é a opção mais rápida e eficiente para a análise dos três defeitos alvo deste estudo. Nesse sentido, nas análises com múltiplos danos do Capítulo 5, optar-se-á por usar a metodologia 1 como um passo do algoritmo proposto para a identificação dos defeitos.



## CAPÍTULO 5

### ALGORITMO PARA DETEÇÃO E IDENTIFICAÇÃO DE MÚLTIPLOS DEFEITOS

#### 5.1 INTRODUÇÃO

Na análise de defeitos em veículos ferroviários é de extrema importância reconhecer a complexidade que caracteriza os cenários reais, nos quais poderá existir a ocorrência de vários tipos de danos numa única passagem. Podemos compreender que o termo multi danos abrange situações em que uma estrutura ou sistema sofre impacto de diversos tipos de danos ou então enfrenta múltiplos danos do mesmo tipo. Assim, torna-se essencial a implementação de estratégias dedicadas de identificação de dano que tenham em conta esse cenário. Num estudo conduzido por Rytter et al. [124], de forma a identificar os vários danos, foi proposta uma hierarquia que é composta por 4 níveis, e posteriormente, em 2004, Worden et al. [19] acrescentaram um quinto nível a essa hierarquia. Estas técnicas de monitorização da integridade estrutural são categorizadas de acordo com a sua capacidade de responder a questões específicas, estando organizadas na Tabela 24.

Tabela 24 - Classificação dos métodos de identificação de danos.

Nível	Objetivo
Nível 1	Deteção do dano - O dano está presente na estrutura (deteção)?
Nível 2	Localização do dano - Onde é que o dano está localizado (localização)?
Nível 3	Tipo do dano - Que tipo de dano está presente (tipo)?
Nível 4	Extensão do dano - Qual é a extensão dos danos (gravidade)?
Nível 5	Previsão - Qual é o tempo de vida residual (previsão)?

Importa referir que as respostas às várias perguntas, só podem ser dadas de forma sequencial, por exemplo, a resposta ao tipo de dano só pode ser dada com um conhecimento a priori da localização do dano. O nível 1 distingue-se no sentido em que pode ser realizado sem conhecimento prévio do comportamento do sistema quando danificado [19]. As respostas às perguntas sobre a localização, o tipo

e a gravidade dos danos podem ser consideradas como diagnóstico, já a quinta pergunta pode ser vista como um prognóstico.

No contexto deste estudo, uma das principais complexidades associadas à presença simultânea de vários danos, é a identificação individual dos danos, uma vez que uma única passagem de uma composição ferroviária pode conter múltiplos danos. Nesse sentido, é possível que a presença de um dano influencie os outros, especialmente quando há proximidade significativa entre eles. De forma a dar um passo em frente em relação às metodologias já existentes nesta área, as quais se focam exclusivamente apenas na identificação de danos individuais, foi desenvolvido um algoritmo para a análise multi dano descrito no subcapítulo 5.2, que foi calibrado em veículos ferroviários de mercadorias e para uma velocidade de 80 km/h. Depois de calibrado, este algoritmo é testado numericamente no subcapítulo 5.3, considerando passagens simuladas com várias combinações de tipo de defeitos e localizações.

### 5.2 ALGORITMO PARA ANÁLISE MULTI DANO

Com base nos primeiros 4 níveis da hierarquia de Rytter atrás referida, foi desenvolvido um algoritmo para a detecção e classificação de cenários multi dano. As principais etapas são: a detecção do dano (i); a localização do dano (ii); a identificação do tipo do dano (iii); e a classificação da severidade do dano (iv). Na Figura 74 é apresentada uma ilustração da sequência e principais etapas do algoritmo. De seguida, em cada sub-capítulo é descrito com maior detalhe em que consiste cada etapa e que métodos são usados.

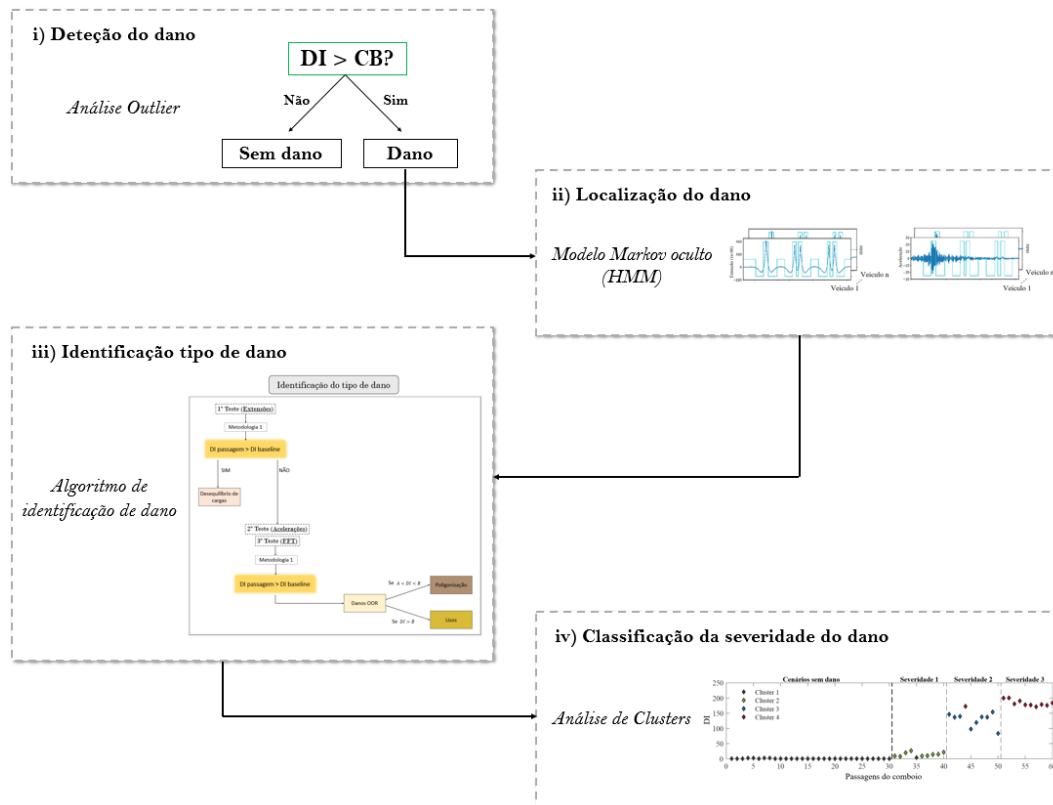


Figura 74 – Principais etapas do algoritmo proposto para a deteção e classificação de multi-dano.

### 5.2.1 Detecção do dano

Apesar da distinção visual entre casos sem dano e casos com dano, a aprendizagem não supervisionada deve utilizar estratégias capazes de detetar automaticamente os danos. Para a deteção do dano é usada uma análise *outlier* (secção 2.7.4.1), que consiste em definir um limite de confiança com um nível de significância de 1 % [25]. Para determinar se a passagem analisada tem defeito ou não, um índice de dano (DI) é comparado com o limite de confiança (CB). Nesta fase de deteção do dano é usado o registo completo sem qualquer tipo de segmentação dos defeitos, em que são extraídos indicadores PCA e seguidamente treinados no SSAE com os mesmos hiperparâmetros adotados para os lisos, apresentados na Tabela 20. Na Figura 75 é apresentada uma aplicação exemplificativa desta primeira etapa do algoritmo aos resultados obtidos pela metodologia 1 usando os registos dos três tipos de defeito isolados. Nesta mesma figura é possível detetar com sucesso a presença de danos OOR (poligonização, lisos) e desequilíbrios de carga sem falsos negativos, considerando uma velocidade de 80 km/h, uma vez que os valores do índice de dano estão todos acima da linha limite.

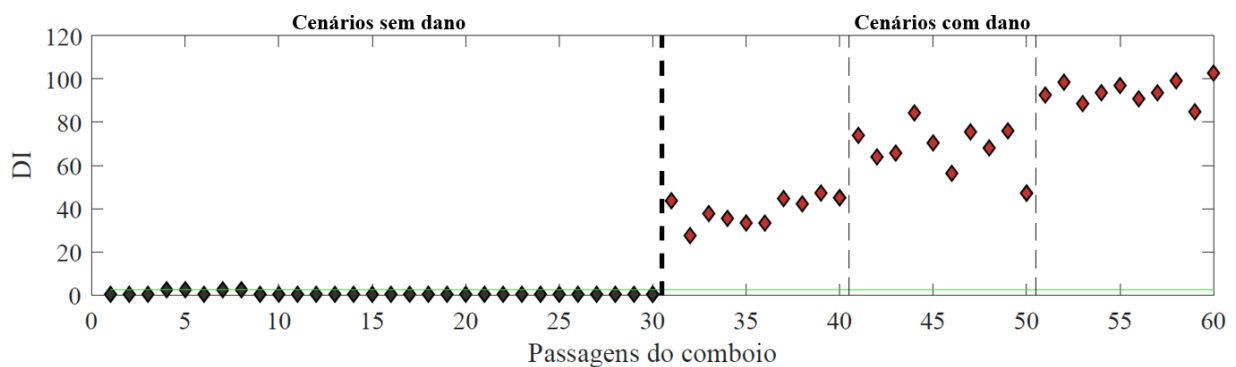


Figura 75 – Deteção automática de danos OOR e desequilíbrios de carga à velocidade de 80 km/h.

### 5.2.2 Localização do dano

Após a aplicação da primeira etapa e se o resultado for a presença de defeito, procede-se à sua localização. Esta etapa é realizada através de uma metodologia proposta por Lourenço et al. [13], em que é feita uma segmentação do sinal, permitindo obter segmentos distintos e fazendo com que estes sejam indexados automaticamente à roda correspondente. Isto permite uma análise mais eficiente e precisa da dinâmica roda-carril, identificando pontos de corte, de forma a possibilitar a comparação direta de padrões semelhantes em passagens individuais de rodas, Figura 76. Para isto, foi utilizado um Modelo de Markov Oculto [125], que permite extrair indicadores dos sinais que induzem a localização do dano.

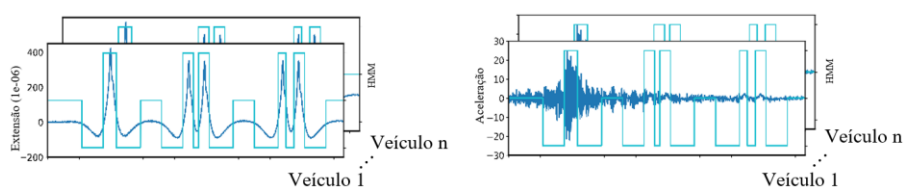
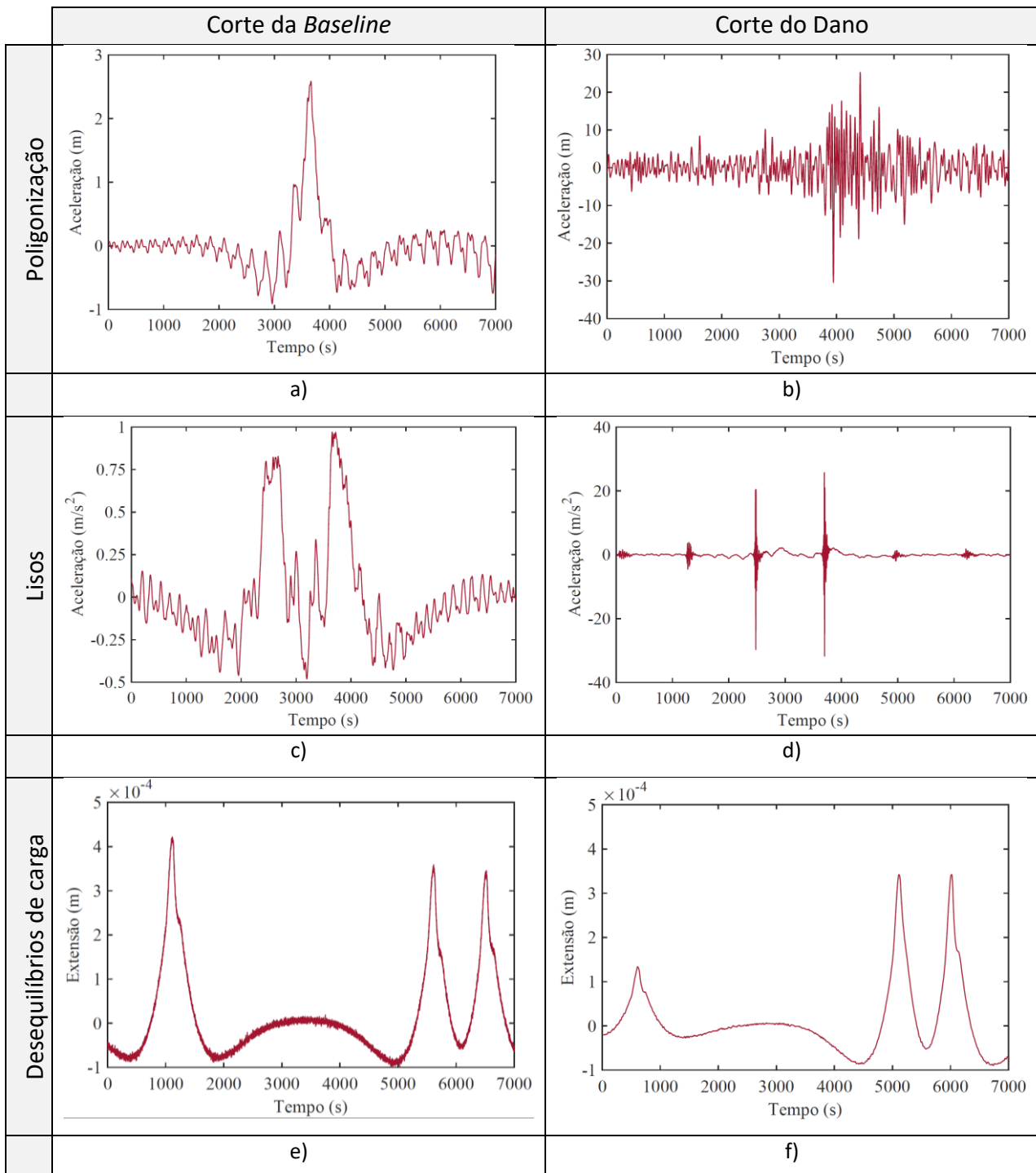


Figura 76 – Segmentação de sinais em extensão e aceleração (adaptado de [13]).

Após concluída a fase da segmentação, são definidas janelas de corte que serão importantes para isolar os vários danos da mesma passagem, permitindo depois analisá-los de forma individual, tornando o problema menos complexo. Resumidamente, a localização e a separação do dano por janelas, são um *input* importante no algoritmo representado na Figura 74, sendo imperativo a utilização das duas abordagens de forma conjunta. De seguida, na Tabela 25, são apresentadas as janelas de sinal cortado obtidas pela aplicação desta metodologia de segmentação do sinal medido, aos 3 defeitos estudados.

Tabela 25 – Exemplificação do corte dos sinais completos realizados através da segmentação.



Os cortes dos sinais são realizados tanto nos cenários sem dano quanto nos cenários com dano, mantendo a mesma dimensão na janela, para que na etapa da classificação seja possível comparar a mesma parte do registro. No caso da poligonização, figura a), como o dano se encontra no primeiro eixo do primeiro vagão, apenas é isolado um pico. Por sua vez, no caso do liso, o dano ocorre no primeiro eixo do terceiro vagão, resultando na influência de dois eixos de dois vagões consecutivos, o que leva à presença de 2 picos (figura b)). Ao contrário dos danos nos rodados em que apenas os eixos essenciais são selecionados, no caso dos desequilíbrios de carga é necessário isolar um vagão. Isto revela-se necessário para se conseguir efetivamente distinguir desequilíbrios tanto transversais quanto longitudinais.

### 5.2.3 Identificação do tipo de dano

Após a aplicação da segmentação ao sinal completo, são definidos  $n$  matrizes distintas em função do número de defeitos daquela passagem, com o sinal cortado na janela correspondente em acelerações, em acelerações no domínio da frequência (espectro FFT) e extensões.

Para a identificação do tipo de dano é realizada uma sequência de testes aplicados às matrizes obtidas da fase anterior. Nesta fase é utilizada a metodologia 1, e para isso são utilizados indicadores PCA para treinar o SSAE, para de seguida ser realizada a fusão dos dados. Na Figura 77 é apresentado o fluxograma que inclui os 3 testes aplicados sequencialmente aos registos até se concluir qual é o tipo de dano.

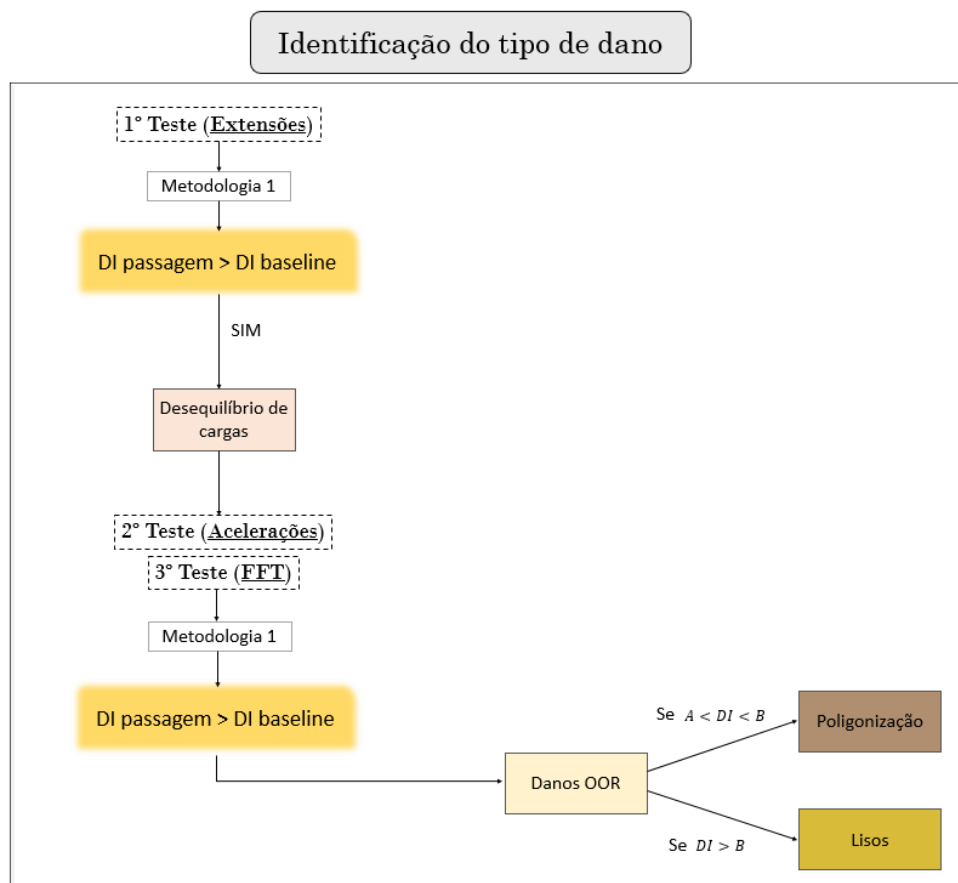


Figura 77 – Fluxograma proposto para a etapa da identificação do tipo de dano.

No primeiro teste, a metodologia é aplicada usando os registos de extensões. Como já foi referido anteriormente, quando comparados registos em extensões de um caso de desequilíbrio de carga com um caso de dano nos rodados ou até de um caso sem dano, as diferenças são identificáveis, pelo que o valor de DI será superior no caso dos desequilíbrios de carga em relação aos restantes. Sendo assim, nesta primeira fase, poderá desde logo isolar-se os casos de instabilidade de cargas de casos de danos nos rodados. Nesta fase são apenas utilizados 2 indicadores PCA (*kurtosis* e *assimetria*), visto que os restantes 2 não eram sensíveis e contribuíam negativamente para os resultados obtidos.

Num segundo teste, a metodologia é aplicada usando os registos em aceleração, onde os defeitos dos rodados se destacam em relação aos desequilíbrios de carga, já que estes últimos quando avaliados em termos de aceleração não apresentam diferenças significativas em comparação com casos sem dano. Nesta fase, poderá fazer-se a distinção entre casos de poligonização e casos de lisos. Como os lisos nas rodas apresentam uma maior excitação em termos de aceleração, estes tendem a ter um valor de DI superior ao da poligonização.

Por fim, num terceiro teste, a metodologia é aplicada usando como dados de entrada as acelerações no domínio da frequência (FFT). Nesta análise verifica-se igualmente que os defeitos dos rodados se destacam em relação aos desequilíbrios de carga, uma vez que no domínio da frequência a resposta dos desequilíbrios será semelhante à dos cenários sem dano. Este terceiro teste servirá, portanto, para comprovar os resultados em termos de aceleração, garantindo uma maior eficiência na distinção entre os dois danos nos rodados, lisos e poligonização.

Desse modo foi aplicada a segmentação e corte dos sinais completos das simulações numéricas conduzidas no capítulo 4 para cada tipo dano, e que são dispostos nas matrizes da seguinte forma: 30ª-60ª passagem - poligonização, 61ª-90ª passagem - lisos e 91ª-102ª passagem – desequilíbrios de carga. Com a informação devidamente organizada, são obtidos indicadores PCA a partir dos dados das três matrizes, em extensão, aceleração e FFT. Posteriormente, estes indicadores são incorporados no processo de treino do SSAE e obtidos DI aplicando a fusão de dados. Para uma correta identificação dos danos, é necessário proceder à definição dos hiperparâmetros de treino no SSAE, que por sua vez são apresentados na Tabela 26.

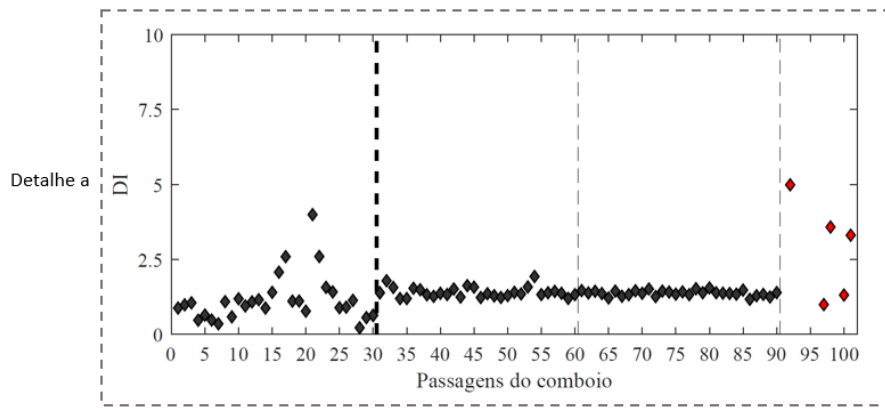
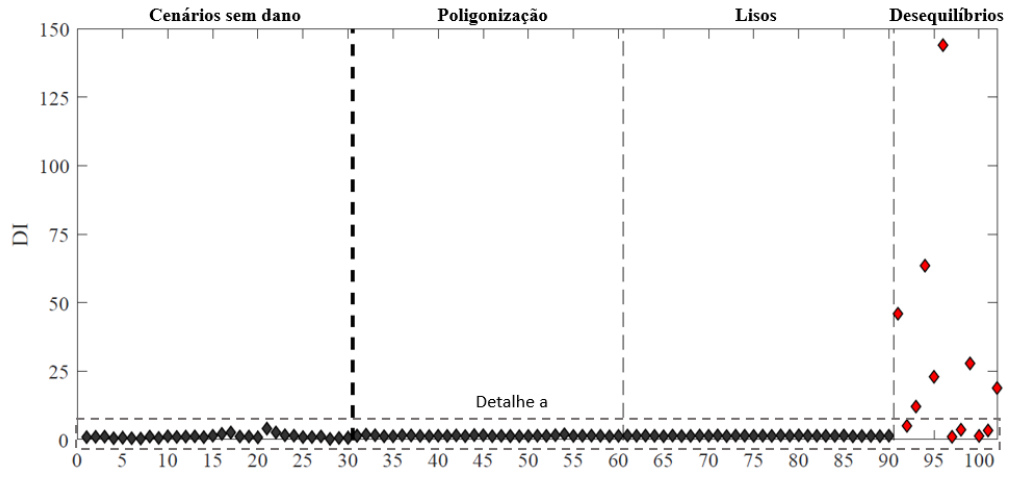


Tabela 26 – Hiperparâmetros utilizados no SSAE para a identificação do tipo de dano.

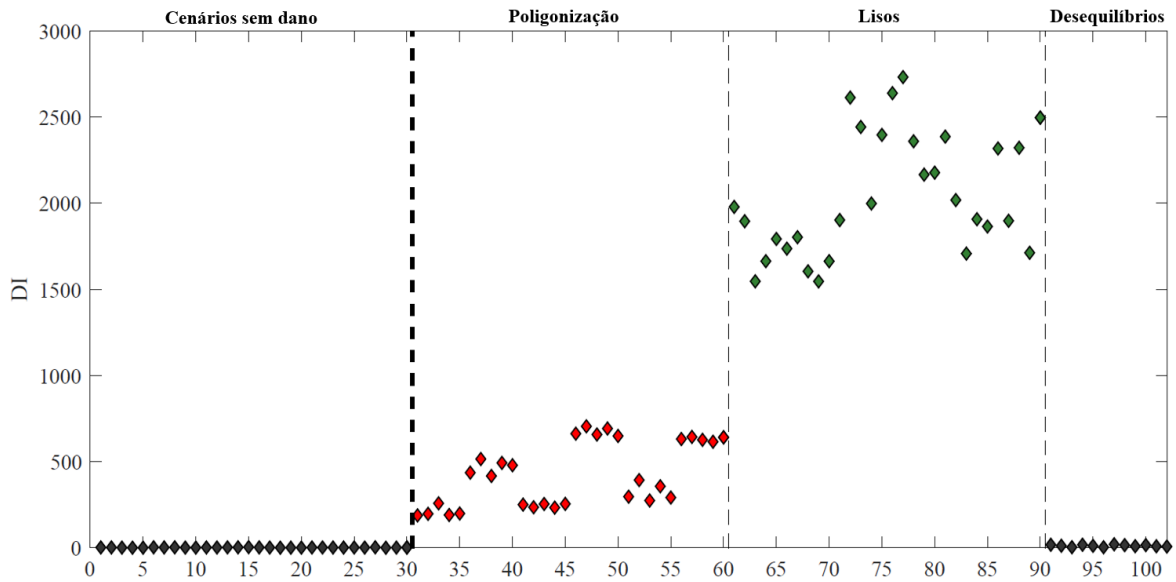
Hiperparâmetros de treino no SSAE	Extensões	Acelerações	FFT
Função de transferência de codificação	logsig	logsig	logsig
Função de transferência de decodificação	logsig	logsig	logsig
Tamanho da camada oculta (1ª/2ª camadas)	1000/500	1000/400	750/600
Número máximo de épocas (1ª/2ª camadas)	2000/2000	2000/2000	10000/5000
Regularização de pesos L2 (1ª/2ª camadas) ( $\lambda$ )	0.5/1×10 <sup>-6</sup>	1×10 <sup>-4</sup> /1×10 <sup>-6</sup>	1×10 <sup>-4</sup> /1×10 <sup>-6</sup>
Regularização de esparsidade (1ª/2ª camadas) ( $\beta$ )	1×10 <sup>-4</sup> /1	1×10 <sup>-4</sup> /0.05	1×10 <sup>-3</sup> /1
Proporção de esparsidade (1ª/2ª camadas) ( $\rho$ )	1×10 <sup>-4</sup> /1×10 <sup>-3</sup>	1×10 <sup>-3</sup> /1×10 <sup>-3</sup>	1×10 <sup>-2</sup> /1×10 <sup>-3</sup>

De uma maneira geral, verifica-se que a maioria dos parâmetros não estão sendo semelhantes aos já utilizados no Capítulo 4, no entanto verifica-se um aumento significativo no número máximo de épocas no caso dos dados de entrada em FFT, comparando com as extensões e as acelerações. Devido ao corte do sinal, que envolve a remoção de partes do mesmo, torna-se mais desafiador estabelecer relações nos dados no contexto dos FFT's, portanto, o SSAE requer mais iterações e passagens pela rede para efetuar as respectivas associações.

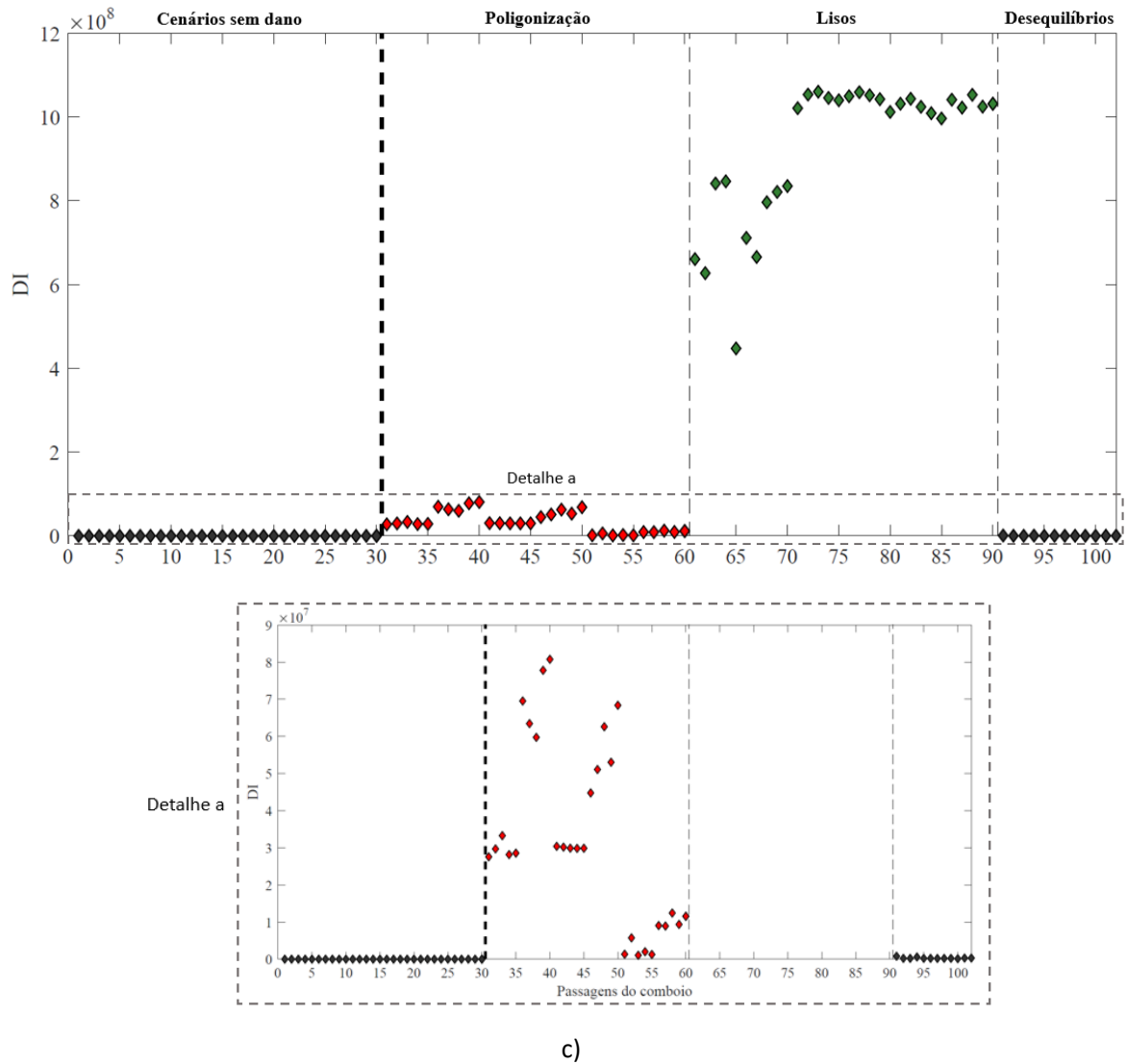
Com os hiperparâmetros definidos e o treino no SSAE concluído, os resultados obtidos após a fusão de dados são os que se apresentam na Figura 78. Na Figura 78 a), apresentam-se os resultados referentes às extensões, onde se observa uma tendência para destacar os desequilíbrios de carga em relação aos cenários sem dano e cenários de danos nos rodados. No entanto, é importante observar que existem quatro passagens em que essa tendência não se verifica, devido às suas características que se assemelham a registros sem dano, o que induz uma dificuldade na distinção. Na Figura 78 b), encontram-se os resultados para as acelerações. Aqui, é evidente que os desequilíbrios de carga estão agora próximos da *baseline*, com um valor de DI consideravelmente reduzido, em contraste com os danos nos rodados, que apresentam valores mais elevados de DI. A distinção entre a poligonização e os lisos é clara, permitindo estabelecer um intervalo de confiança sólido entre esses dois tipos de danos. Por último, na Figura 78 c), verifica-se que, em termos de FFT, os valores de DI aumentam significativamente em comparação com os valores das acelerações. No entanto, continua a existir uma boa distinção entre a poligonização e os lisos. Neste caso, as diferenças entre a *baseline* e o último grupo da poligonização são muito pequenas, embora ainda assim se encontrem acima dos cenários sem danos e dos desequilíbrios de carga.



a)



b)



c)

Figura 78 – Aplicação dos testes para identificação do tipo do dano: a) 1º teste: Extensões b) 2º teste: Acelerações c) 3º teste: FFT.

Com base nos resultados da Figura 78, são estabelecidos intervalos de valores de DI para categorizar cada tipo de dano. Por exemplo, para que um caso seja considerado como liso, este tem de atender aos critérios dos 3 testes, sendo que para isso tem de apresentar um valor de DI compreendido entre 0 e 5 no caso das extensões, um valor superior a 1000 no caso das acelerações e superior a  $2 \times 10^8$  no caso dos FFT's, conforme se apresenta na Tabela 27. Nesta mesma tabela são definidos os intervalos definidos para cada tipo de dano.

Tabela 27 – Intervalos de valores referentes a cada tipo de defeito para identificação do tipo de dano.

	Poligonização	Lisos	Desequilíbrio de cargas
Extensões	0 – 5	0 – 5	> 5
Acelerações	180 – 1000	> 1000	0 – 100
FFT	$0.1 \times 10^7 - 2 \times 10^8$	> $2 \times 10^8$	$0 - 0.1 \times 10^7$

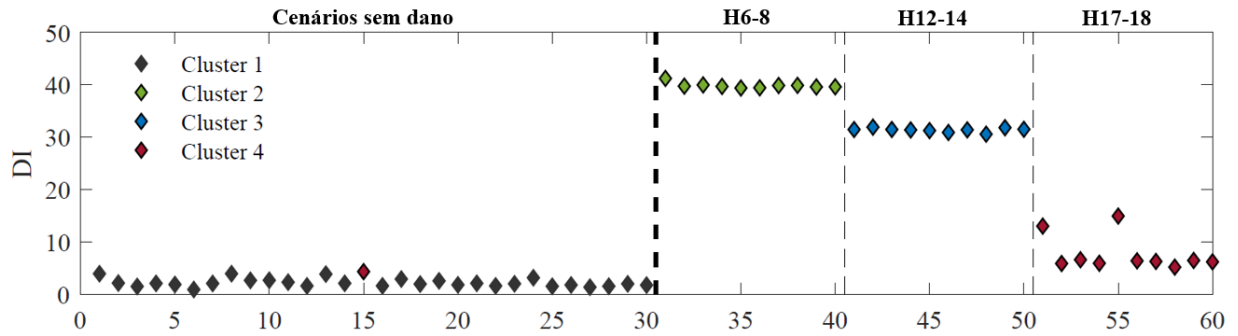
### 5.2.4 Classificação da severidade do dano

Após identificado o tipo de dano na etapa anterior, procede-se então à última etapa, a classificação da sua severidade. Para isso, é aplicada também a metodologia 1 discutida no Capítulo 4, com umas pequenas adaptações. Devido ao corte do sinal, uma parte é descartada, e apenas a parte que contém os danos é utilizada. Isso resulta numa perda significativa de sensibilidade, pelo que as técnicas utilizadas na etapa i) (extração de indicadores) da metodologia 1, requerem ajustes e também uma nova otimização dos hiperparâmetros utilizados no SSAE. Sendo assim, no contexto de múltiplos danos, a técnica CWT é usada para extrair os indicadores no caso dos danos de lisos. Pelo contrário, quer na poligonização quer nos desequilíbrios de carga é utilizado a técnica PCA. Na técnica PCA para a extração inicial de indicadores são apenas utilizados 2 indicadores, a raiz quadrada média e o desvio padrão para o caso da poligonização e o *kurtosis* e à assimetria no caso dos desequilíbrios de carga. Finalmente, na Tabela 28 são apresentados os valores otimizados do SSAE para a classificação da severidade dos 3 tipos de dano.

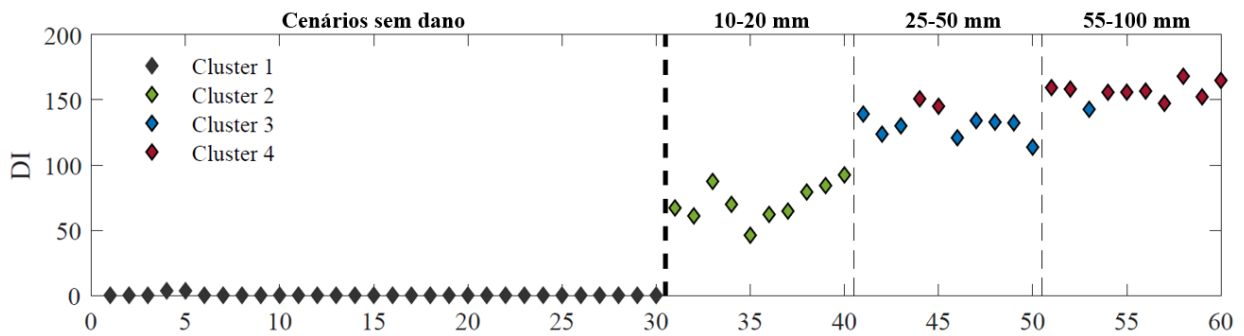
Tabela 28 – Hiperparâmetros utilizados no SSAE para 1 velocidade (80 km/h) e para os 3 tipos de defeitos.

Hiperparâmetros de treino no SSAE	Poligonização (Harmónicos)	Lisos	Desequilíbrios de carga
Função de transferência de codificação	logsig	logsig	logsig
Função de transferência de decodificação	logsig	logsig	logsig
Tamanho da camada oculta (1ª/2ª camadas)	500/250	750/500	1000/500
Número máximo de épocas (1ª/2ª camadas)	1000/1500	2000/2000	2000/2000
Regularização de pesos L2 (1ª/2ª camadas) ( $\lambda$ )	$1 \times 10^{-4}/2 \times 10^{-7}$	$1 \times 10^{-3}/1 \times 10^{-6}$	$0.5/1 \times 10^{-2}$
Regularização de esparsidade (1ª/2ª camadas) ( $\beta$ )	$1 \times 10^{-1}/8$	$1 \times 10^{-1}/8$	$1 \times 10^{-4}/1$
Proporção de esparsidade (1ª/2ª camadas) ( $\rho$ )	$1 \times 10^{-2}/1$	$1 \times 10^{-2}/1 \times 10^{-3}$	$1 \times 10^{-4}/1 \times 10^{-3}$

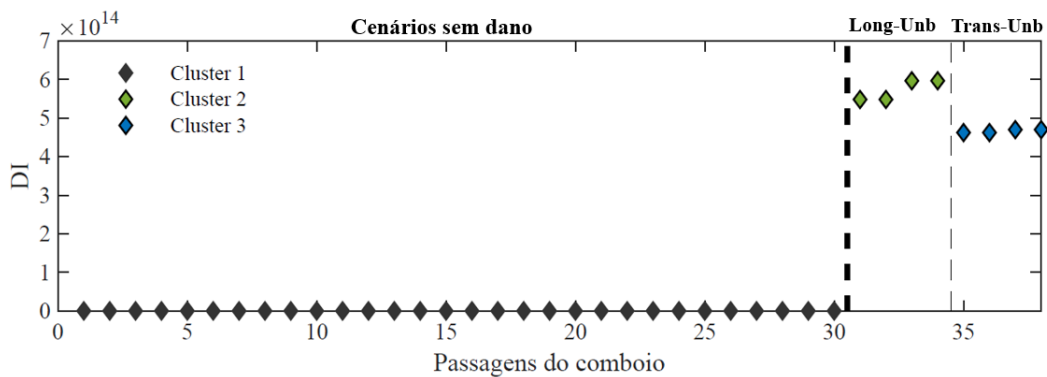
Na Figura 79 são apresentados os resultados da análise de clusters resultantes da aplicação da metodologia 1 para cada tipo de defeito, considerando a fusão de todos os acelerómetros no caso dos danos OOR, e de todos os extensómetros no caso dos desequilíbrios de carga. Com base nestes resultados conclui-se que não se verifica uma perda de sensibilidade ao dano, observando-se uma elevada eficiência na classificação de todos os tipos de dano, mesmo tendo em conta os cortes do sinal e consequente perda de uma quantidade significativa de informação dos dados.



a)



b)



c)

Figura 79 – Classificação automática da severidade dos 3 defeitos com base na análise de clusters para a velocidade de 80 km/h e com corte de sinal: a) poligonização b) lisos c) desequilíbrio de cargas.

No caso da poligonização, Figura 79 a), obtém-se uma classificação quase perfeita, com apenas uma classificação incorreta, devido à proximidade do grupo H17-18 da *baseline*. Observa-se também uma menor variabilidade dentro dos vários cenários do mesmo número de harmónicos. Neste caso, o grupo H17-18 apresenta um valor de DI muito próximo dos cenários de base, podendo dificultar uma correta classificação de todas as passagens.

Em relação ao dano de lisos, Figura 79 b), obtém-se uma boa classificação dos vários grupos de severidade, existindo apenas 3 falsas classificações. Finalmente, nos casos de desequilíbrios de carga, Figura 79 c), a distinção entre desequilíbrios longitudinais e transversais é perfeita.

Ao compararmos estes resultados com os obtidos com o sinal completo, observa-se uma tendência semelhante no que diz respeito aos danos nos rodados, contudo, verifica-se uma alteração nos casos dos desequilíbrios, onde agora os desequilíbrios transversais apresentam um valor de DI inferior.

Também aqui nesta etapa são definidos os intervalos de DI, neste caso para a classificação dos 3 defeitos. A Tabela 29 remete para esses mesmos valores.

Tabela 29 – Intervalos de valores referentes a cada tipo de defeito para classificação.

Poligonização		Lisos		Desequilíbrios de carga	
Severidade	Valores	Severidade	Valores	Severidade	Valores
H6-8	> 32	10-20 mm	4 – 100	Long-Unb	$> 5.0 \times 10^{14}$
H12-14	15 – 32	25-50 mm	100 – 140	Trans-Unb	$5.0 – 5.0 \times 10^{14}$
H17-18	6 – 15	55-100 mm	> 140	-	-

### 5.3 VALIDAÇÃO DO ALGORITMO

Após a calibração do algoritmo e a definição dos critérios em cada fase da sua aplicação, torna-se necessário a sua validação numérica com cenários simulados de passagens com múltiplos defeitos. Para esse efeito, foram gerados 9 cenários de multi dano tendo em atenção os vários pressupostos já definidos atrás. Foram definidas 3 configurações tipo para uma passagem de 5 veículos do tipo *Laagrss*, ilustrados na Figura 80, para avaliar a eficiência do algoritmo aos diferentes tipos e localizações dos danos. Na primeira configuração (designada de CT1), representada na Figura 80 a), os 3 danos foram definidos em vagões não consecutivos, existindo assim uma grande distância entre eles. Nesta configuração foram gerados 3 cenários diferentes, variando apenas as severidades dos danos. Por exemplo, no primeiro cenário, foi considerado um dano poligonal de 6-8H na primeira roda do lado direito do primeiro vagão, um liso menos severo (10-20 mm) na roda do lado direito do terceiro vagão, e um desequilíbrio longitudinal (desequilíbrio longitudinal 1) no quinto vagão. No cenário seguinte, com a mesma localização dos danos, simulou-se um dano poligonal de 12-14H, um liso de severidade intermédia (25-50 mm) e um desequilíbrio longitudinal (desequilíbrio longitudinal 1). No terceiro e último cenário desta configuração, simulou-se um dano poligonal de 17-18H, um liso mais severo (55-100 mm) e um desequilíbrio transversal (desequilíbrio transversal 1).

No que diz respeito à configuração 2 (CT2), representada na Figura 80 b), os dois tipos de danos nos rodados (poligonização e liso) são simulados em veículos adjacentes, para testar o algoritmo numa situação próxima de dois tipos diferentes de dano, em que poderá existir influência de dois tipos de danos no mesmo corte de sinal. Do mesmo modo, nesta configuração de danos foram gerados 3 cenários variando apenas a localização do dano liso, que passou do 3º vagão para o 2º vagão. Por fim, na configuração tipo 3 (CT3), ilustrada na Figura 80 c), optou-se por avaliar duas situações diferentes. Num primeiro momento, procurou-se compreender o impacto de danos em lados opostos, e numa segunda fase, procurou-se entender a relevância de ter dois tipos de defeitos no mesmo vagão, neste caso, um liso e um desequilíbrio de cargas. Também nesta configuração foram gerados 3 cenários com a variação das severidades, alterando a sua localização. Na primeira roda do lado direito do primeiro vagão, foi simulada uma poligonização, na primeira roda do lado esquerdo do terceiro vagão, um dano liso, e neste mesmo vagão, um desequilíbrio de cargas.

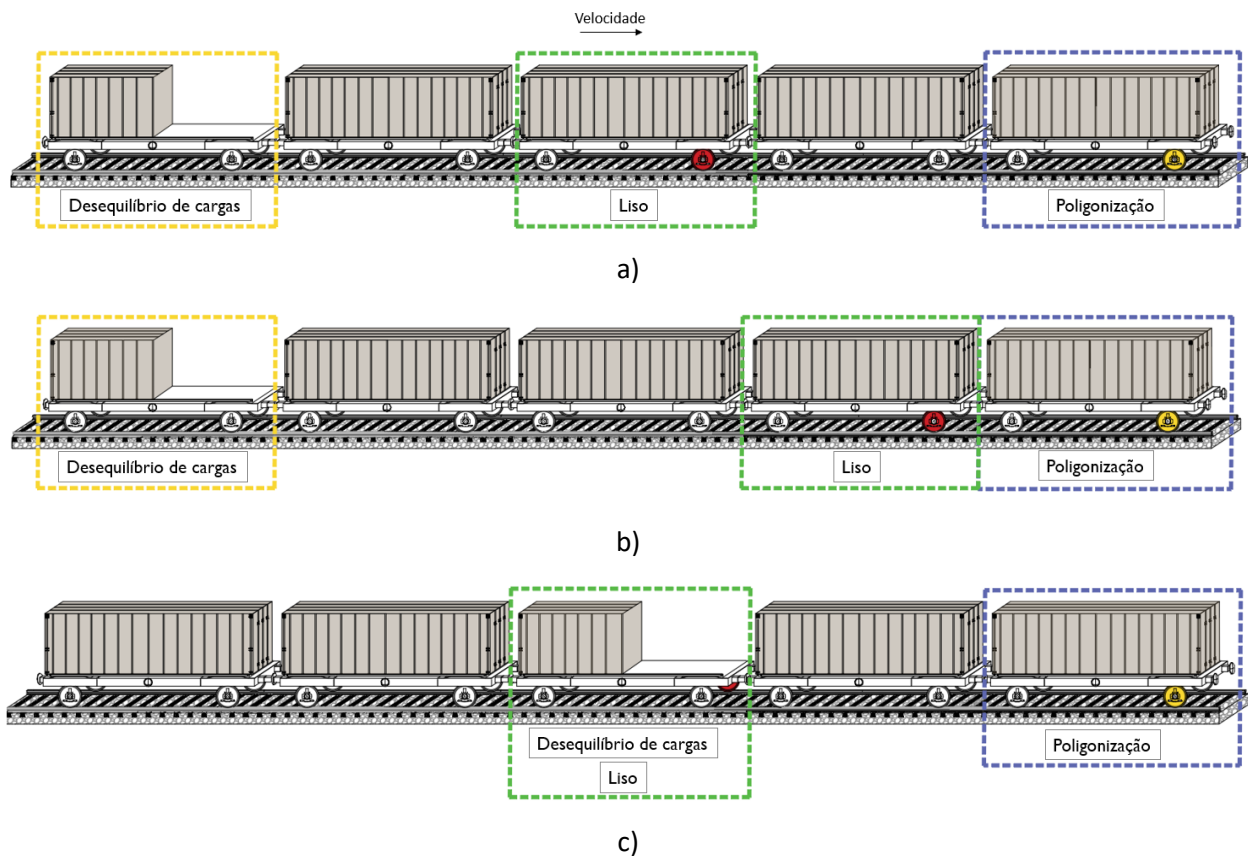


Figura 80 – Cenários multi dano considerados: a) configuração tipo 1 (CT1) b) configuração tipo 2 (CT2) c) configuração tipo 3 (CT3).

Na Figura 81 são apresentados os registos numéricos obtidos em acelerações e extensões para cada configuração definida anteriormente. Estes resultados são avaliados para o sensor 1 localizado no carril do lado direito.

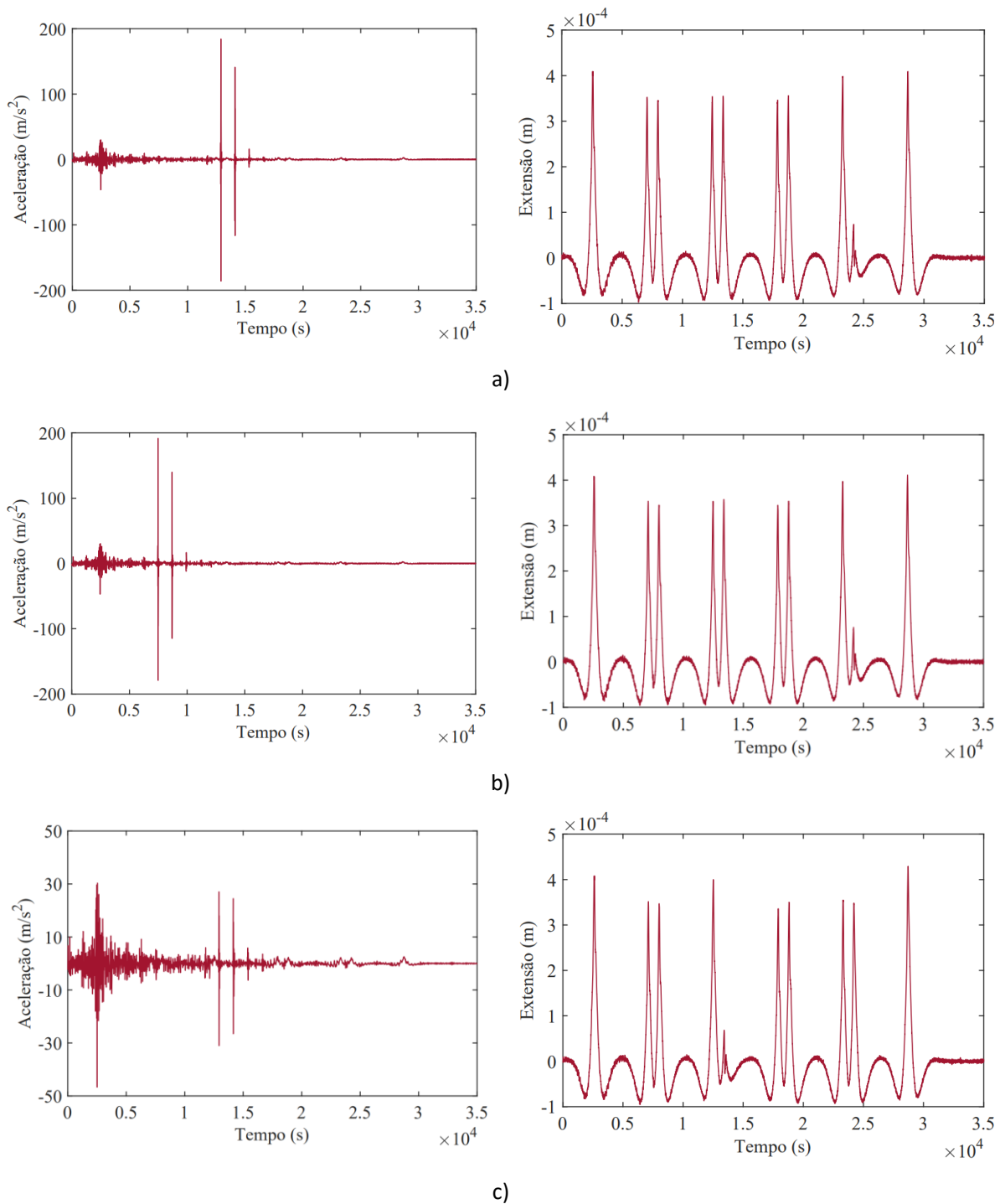


Figura 81 – Sinal completo com multi dano em acelerações e em extensões no sensor 1 (lado direito): a) CT1) b) CT2) c) CT3.

No contexto do segundo cenário das 3 combinações tipo, em que são utilizadas as severidades intermédias dos danos OOR e desequilíbrios longitudinais, verifica-se que nas acelerações existe uma constante ligeira sobreposição dos efeitos da poligonização nos lisos. Ainda neste caso, comparando a Figura 81 a) e a Figura 81 b), verifica-se desde logo a mudança da localização do liso, do 3º para o 2º vagão. Nas extensões, apenas se destaca a ausência de um pico, não sendo visível o impacto dos danos OOR.



Depois de conhecidos os vários *inputs*, é agora o momento de apresentar os resultados finais da validação do algoritmo, com a exposição dos resultados de cada uma das várias etapas do processo.

### 5.3.1 Detecção do dano

Para a deteção do dano aplica-se a estratégia definida em secção 5.2.1, que resulta no gráfico da Figura 82. Pela análise do gráfico observa-se que a deteção de dano é realizada eficazmente para os 9 cenários multi dano simulados, obtendo-se com uma grande distância em relação ao limite de confiança. De realçar que nesta etapa são considerados os registos completos.

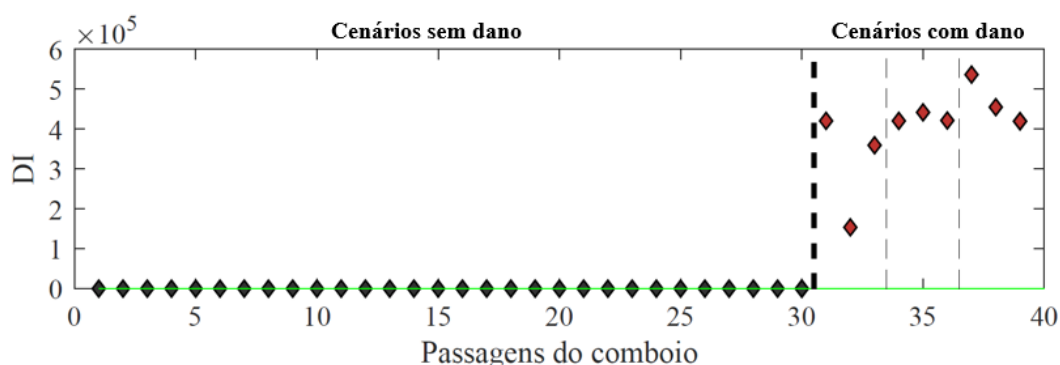


Figura 82 – Deteção automática de multi danos à velocidade de 80 km/h.

### 5.3.2 Identificação do tipo de dano

Após a deteção bem-sucedida do dano, procede-se à sua localização e subsequente corte do sinal, conforme explicado na secção 5.2.2. Com as janelas de análise devidamente estabelecidas, é aplicado o procedimento definido em 5.2.3, considerando os dados de entrada em extensão, em aceleração e em FFT, para assim ser possível obter valores de DI. Uma vez conhecidos esses valores, estes são comparados com os limites definidos desta mesma secção, o que possibilita a atribuição de um tipo de dano correspondente à janela em questão.

Na Tabela 30, é apresentada em forma de tabela a aplicação do procedimento para identificação do tipo de dano para o cenário 2 da configuração CT1. De igual forma, o mesmo procedimento é aplicado para os restantes 8 cenários, a fim de verificar a eficácia da identificação com um maior número de cenários. Esses resultados obtidos para as restantes configurações tipo estão inseridos no Anexo 2.

Tabela 30 – Identificação do tipo de dano com base nos limites de DI definidos na fase de teste (CT1).

CT1	Corte 1	Corte 2	Corte 3			
Loc.	1ª roda	5ª roda	5º vagão			
Janela de corte						
Identificação do dano	Extensões	DI=4.8	DI=4.8	DI=19.3		
	Acelerações	DI=380.4	DI=1010	DI=37.7		
	FFT	DI= 1.9 × 10 <sup>7</sup>	DI= 4.6 × 10 <sup>8</sup>	DI= 7.6 × 10 <sup>5</sup>		
	Extensões	0 – 5	Extensões	0 – 5	Extensões	> 5
	Acelerações	80 – 1000	Acelerações	> 1000	Acelerações	0 – 100
	FFT	0.1 × 10 <sup>7</sup> – 2 × 10 <sup>8</sup>	FFT	> 2 × 10 <sup>8</sup>	FFT	0 – 0.1 × 10 <sup>7</sup>
Decisão	Dano Poligonal	Dano Liso	Desequilíbrio de cargas			

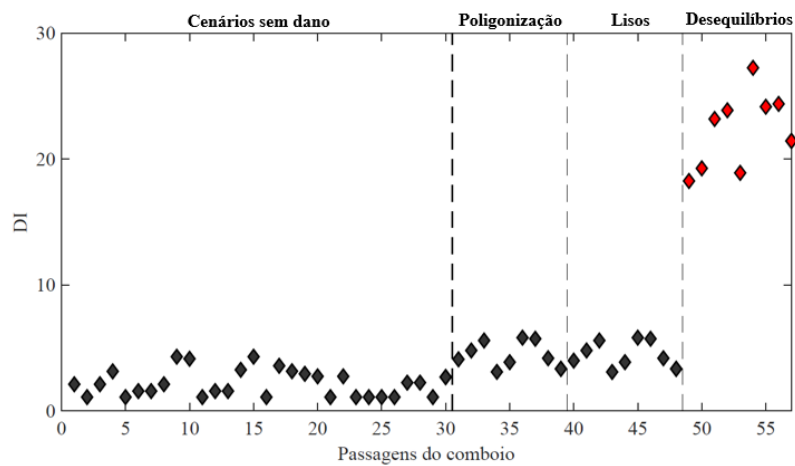
Os resultados considerando os 9 cenários são apresentados na Figura 83. Na Figura 83 a), é possível visualizar os resultados relativos à aplicação do 1º teste, usando o registo de extensões, onde é claramente visível o agrupamento da *baseline*, da poligonização e dos lisos, com valores de DI em torno de 6, e um agrupamento separado deste anterior, com valores de DI superiores a 15, os desequilíbrios de carga. É importante mencionar que esta fase é altamente sensível aos cortes, e uma execução incorreta

pode conduzir a resultados totalmente diferentes. No caso dos desequilíbrios de carga, como as diferenças entre os registos são mínimas, evidenciadas apenas por variações nos picos máximos do eixo com desequilíbrio, a distinção torna-se um desafio consideravelmente maior ao contrário do que acontece no caso das acelerações ou do FFT onde as diferenças são bem mais evidentes. Como o treino do SSAE é bastante sensível ao ruído presente nos sinais, é de extrema importância que os todos os sinais em extensão sejam filtrados a 100 Hz, para minimizar ao máximo quaisquer interferências indesejadas que possam interferir nos resultados.

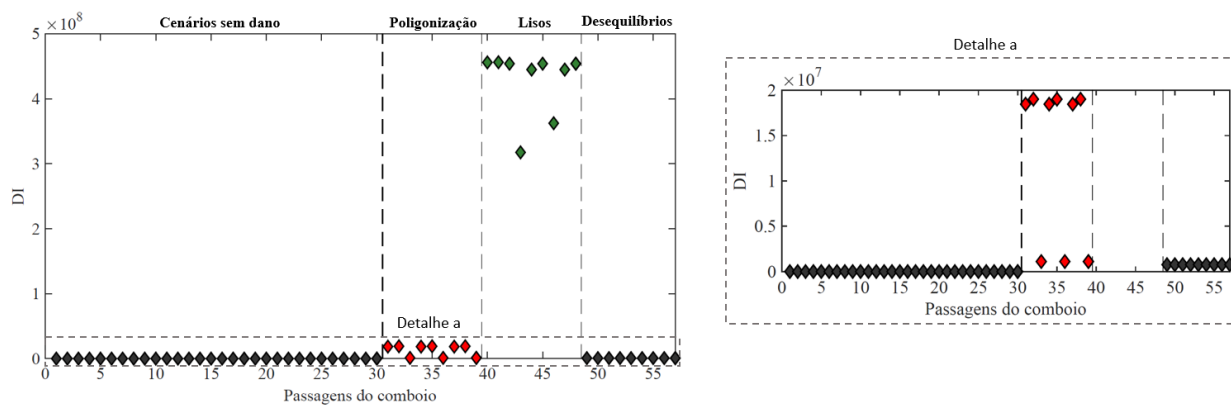
Na Figura 83 b) são apresentados os resultados da aplicação do 2º teste, usando o registo de acelerações. Os resultados mostram uma clara separação entre a poligonização e os lisos e a influência dos cortes neste caso não se mostrou tão influente. Como são defeitos que são facilmente distinguíveis com base no tipo de registo e nos valores que apresentam, permite que o SSAE consiga fazer a distinção entre os defeitos de forma mais eficiente. Neste caso, foram usados os sinais de aceleração filtrados a 1500 Hz.

Por fim, na Figura 83 c), são apresentados os resultados da aplicação do 3º teste, usando os valores do espectro FFT das acelerações. Os resultados obtidos provam que também aqui há uma grande separação entre os dois tipos de danos nos rodados.

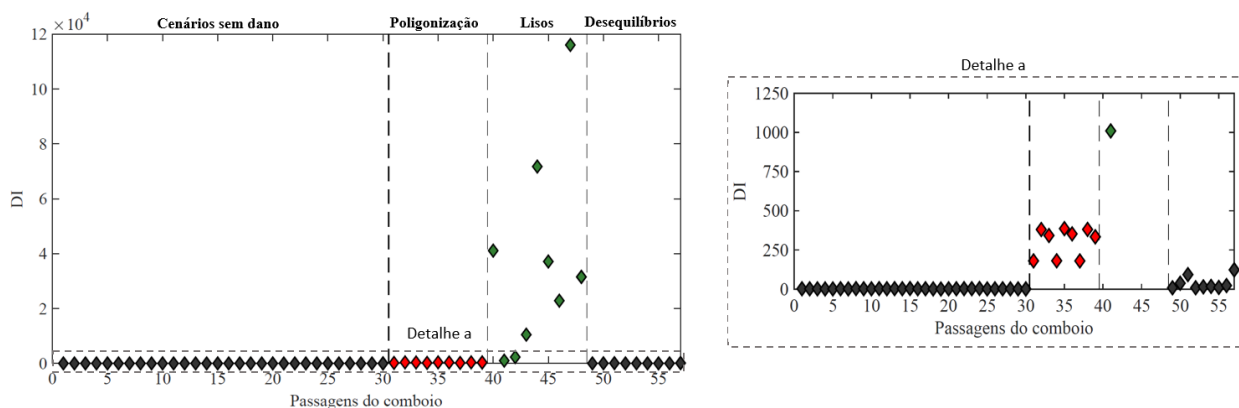
Dado que todos os valores de DI referentes às passagens cenários multi-dano estão enquadrados nos limites definidos nos casos de teste, prova-se que a identificação do dano pode ser realizada com êxito e com um grau razoável de confiança.



a)



b)



c)

Figura 83 – Etapas para identificação do tipo do dano para o caso de multi dano: a) 1º teste: Extensões b) 2º teste: Acelerações c) 3º teste: FFT.

### 5.3.3 Classificação da severidade do dano

Após a identificação do tipo de dano, procede-se à classificação individual da severidade de cada defeito apresentando a análise de cluster para cada tipo de dano na Figura 84.

Para o caso da poligonização, Figura 84 a), a classificação obtida é perfeita, agrupando em 4 grupos bem definidos. No caso dos lisos, Figura 84 b), a classificação obtida é eficaz, ocorrendo apenas uma falha na classificação devido à proximidade do grupo de severidade intermediária com o grupo mais severo. Por fim, no caso dos desequilíbrios de carga, Figura 84 c), a classificação é menos eficaz, uma vez que três casos de desequilíbrio longitudinais são agrupados erradamente como casos transversais.

Apesar da grande eficiência na classificação das severidades revelada pelos gráficos, torna-se necessário também verificar se os valores de DI obtidos para cada caso estão dentro dos limites previamente estabelecidos. Há a necessidade de avaliar estes limites visto que em situações reais onde temos apenas um único ponto, o *clustering* não pode ser utilizado. Embora o *clustering* não tenha sido completamente preciso nos três tipos de defeitos, é importante observar que a maioria das passagens, com exceção de

uma nos lisos e uma no desequilíbrio de cargas, estão de facto dentro dos limites de DI. Isto sugere que estes defeitos possam ser classificados usando essa estratégia.

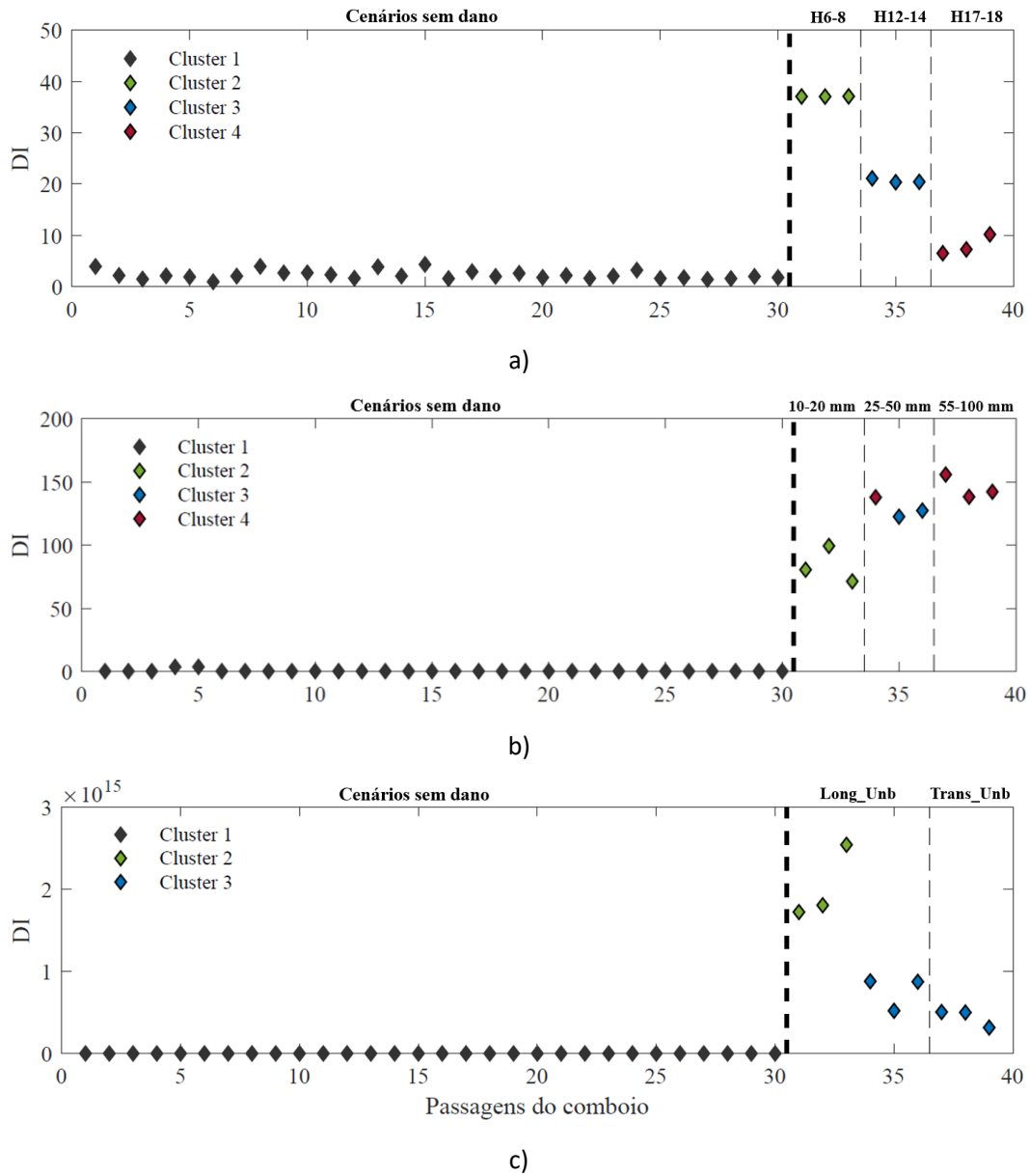


Figura 84 – Classificação automática da severidade dos 3 defeitos com multi dano para uma velocidade de 80 km/h: a) poligonização b) lisos c) desequilíbrio de cargas.



## CAPÍTULO 6

### CONCLUSÕES, DESENVOLVIMENTOS FUTUROS E PUBLICAÇÕES

#### 6.1 CONCLUSÕES

As técnicas convencionais de extração de indicadores lineares utilizadas nas várias metodologias de detecção e classificação de danos, não conseguem lidar eficazmente com uma variável operacional muito importante na engenharia ferroviária, a velocidade. A velocidade é fulcral nas respostas dinâmicas e o seu efeito amplificador pode induzir erros na identificação de danos. Por esta razão, este trabalho propõe uma metodologia automática baseada em inteligência artificial para detecção e classificação de defeitos nos rodados e desequilíbrios de carga, que inclui num dos seus passos principais a utilização de um novo método de extração de características baseado na teoria de aprendizagem profunda. O SSAE não é mais do que um empilhamento de *autoencoders* esparsos que permite a extração não linear de características, normalizando os dados intrinsecamente à medida que os dados são comprimidos através das várias camadas do *autoencoder*.

Resumidamente, no Capítulo 2, foi realizada uma comparação entre os dois tipos de sistemas de monitorização: *onboard* e *wayside*, destacando-se os prós e os contras. Investigou-se os tipos de veículos ferroviários de mercadorias e as suas características distintivas. De seguida, foi feita uma breve explicação das causas e tipos de defeitos nos rodados, e no caso dos desequilíbrios de carga, apresentou-se os limites convencionais para avaliar a estabilidade da carga. Foram ainda exploradas metodologias para simular o sistema dinâmico entre veículo e via, permitindo efetuar simulações mais realistas. Por fim, foram detalhadas as técnicas avançadas utilizadas nas várias etapas da metodologia.

No Capítulo 3, foi descrito o veículo em foco neste estudo, que é um vagão de carga do tipo *Laagrss*. Apresentou-se o modelo numérico desse vagão, criado no programa ANSYS®, que o representa digitalmente. Uma característica peculiar desses vagões é que eles possuem apenas um nível de suspensão, que consiste numa mola de lâminas de aço, e a ligação entre a carroçaria e os rodados é feita através dessa suspensão. Além disso, descreveu-se a modelação numérica da via, também realizada no programa ANSYS®, e realizou-se a caracterização geométrica das irregularidades tanto da via como das

rodas, utilizando o programa MATLAB®. Em seguida, gerou-se aleatoriamente perfis de irregularidades. Para obter dados das respostas numéricas, foi utilizada a ferramenta de análise e cálculo *VSI-Vehicle-Structure Interaction Analysis*. Nesta secção, foi também realizada uma breve descrição do funcionamento deste *software* interno.

No Capítulo 4 foram comparadas duas metodologias propostas de classificação automática de defeitos em veículos ferroviários baseadas em redes neuronais *autoencoder* do tipo SSAE. Para cada uma das abordagens foi feita a sua descrição e apresentados os resultados obtidos em cada uma das suas etapas. Foi ainda realizada uma análise de sensibilidade à velocidade e ao número de sensores em cada uma das metodologias, a fim de perceber qual das duas seria a mais eficaz.

Numa primeira fase, ambas as metodologias foram testadas considerando cenários só com velocidade de 80 km/h e ambas apresentaram resultados muito animadores. Os resultados alcançados mostram uma elevada eficiência na identificação dos 3 tipos de defeito. De seguida, foram incluídos cenários relativos a duas velocidades adicionais (60 e 100 km/h), e verificou-se que tanto a metodologia 1 quanto a metodologia 2 continuaram a exibir excelentes desempenhos na classificação de lisos e poligonização. No caso dos desequilíbrios de carga, a metodologia 1 demonstrou menos falsas identificações que a metodologia 2, sendo que esta última não conseguiu encontrar uma relação entre os dados como se previa.

Relativamente ao número de sensores utilizados, concluiu-se que o *setup* ótimo, seria 6 acelerómetros e 10 extensómetros. Com a redução do número de sensores verificou-se uma perda de sensibilidade, prejudicando os resultados. De uma forma geral, é possível constatar que ambas as metodologias são adequadas para detetar defeitos nos rodados, mas no caso dos desequilíbrios de carga, apenas a metodologia 1 é eficaz. Para além disso, o tempo de cálculo na metodologia 1 é bastante inferior ao da metodologia 2. Sendo assim, concluiu-se que a metodologia 1 será a opção mais eficiente para a classificação dos três tipos de defeitos de veículos ferroviários objeto de estudo neste trabalho, poligonização, lisos e desequilíbrio de cargas.

No Capítulo 5, é apresentado um algoritmo para a identificação de multi danos. Cada uma das etapas deste algoritmo foi calibrada usando registos completos dos três tipos de defeito isolados. Após esta calibração, foi realizada uma validação numérica com simulações de passagens com várias combinações dos três tipos de defeito a fim de testar a fiabilidade do algoritmo.

Este algoritmo incorpora a deteção do dano, a localização do dano, a identificação do tipo de dano e a severidade do dano. Com recurso a uma análise *outlier*, a deteção do dano foi realizada com 100 % de eficácia. Na parte da localização, em que é feita uma segmentação do sinal completo e que depois permite a aplicação dos restantes passos, estas devem ser realizadas quer para a *baseline* quer para os defeitos usando a mesma janela de corte de sinal, para assim possibilitar as comparações. Na etapa da



identificação do tipo de dano, são utilizados 3 testes aplicados ao sinal cortado, aplicando a metodologia de classificação definida anteriormente a extensões, acelerações e ao FFT das acelerações. Os valores de DI obtidos nas passagens de multi dano foram enquadrados de forma perfeita nos limites pré-estabelecidos com as passagens de teste. Por fim, na classificação da severidade, também os valores de DI obtidos foram enquadrados com sucesso nos limites estabelecidos, com apenas 2 casos de insucesso, 1 no caso dos lisos e 1 no caso dos desequilíbrios de carga.

De uma forma muito sucinta, concluiu-se que o algoritmo aqui proposto é bastante robusto, mas que ainda assim é necessário um aperfeiçoamento para os casos de desequilíbrio de cargas, uma vez que se mostrou uma perda grande de sensibilidade com os cortes de sinal. Para além disso, como estamos perante um algoritmo que é realizado de forma automática, onde é difícil de garantir os cortes adequados e perfeitos, podemos deduzir que será necessário adotar estratégias mais delicadas para se conseguir obterem melhores resultados. Ainda assim, de uma maneira geral, verificou-se uma excelente eficiência do algoritmo para a identificação e classificação dos 2 tipos de danos nos rodados, o que é bastante promissor para uma futura aplicação em contexto real destas ferramentas.

## 6.2 DESENVOLVIMENTOS FUTUROS

O trabalho descrito ao longo da presente tese levantou vários tópicos, questões e objetivos adicionais que não foram abordados. Neste contexto, são referidos nos parágrafos seguintes alguns tópicos que permitirão continuar a melhorar o trabalho já realizado.

- Aplicação do algoritmo a um maior número de simulações multi dano, de forma a abranger um maior número de cenários. Aqui, também poderão ser usados cenários que considerem vários danos do mesmo tipo, por exemplo 3 lisos.
- Aplicação do algoritmo a sinais reais que foram recolhidos através de uma campanha experimental em Estarreja. Apesar de os registos numericamente simulados tenham sido baseados em modelos calibrados experimentalmente e cenários muito semelhantes aos reais, a aplicação num contexto real poderá trazer desafios que obriguem a uma adaptação do algoritmo.
- Aplicação e otimização do algoritmo para outros tipos de veículos, por exemplo no veículo de mercadorias com *bogie* e nos veículos de passageiros, uma vez que as respostas dinâmicas e as velocidades de circulação são diferentes, podendo influenciar os resultados.
- Utilização de outros tipos de AE's, alternativamente ou complementarmente ao SAE. Com a continua criação de novos AE's, a incorporação de outro tipo pode ser uma mais-valia. Poderá utilizar-se o *autoencoder* variacional, o *autoencoder* convolucional e um mais recente, o *ensemble autoencoder*.

- De forma a reduzir os dados de entrada e reduzir consideravelmente o tempo de cálculo da rede de AE's, poderá considerar-se a aplicação de transformadas aos registos medidos, como por exemplo o *mel spectrogram*.

## 6.3 PUBLICAÇÕES

### 6.3.1 Artigos em conferências nacionais

- T. Jorge, J. Magalhães, R. Silva, A. Guedes, D. Ribeiro, A. Meixedo, A. Mosleh, C. Vale, P. Montenegro (2023) - "Unsupervised Feature Extraction using a Stacked Sparse Autoencoder for OOR Damage Wheels identification", Proceedings of XLIV Ibero-Latin American Congress on Computational Methods in Engineering (CILAMCE 2023), Porto, Portugal. (aceite para publicação).
- J. Magalhães, T. Jorge, R. Silva, A. Guedes, D. Ribeiro, A. Meixedo, A. Mosleh, C. Vale, P. Montenegro - "Application of a Sparse Autoencoder model for OOR Damage Wheels identification in a passenger train using Wayside track responses", Proceedings of XLIV Ibero-Latin American Congress on Computational Methods in Engineering (CILAMCE 2023), Porto, Portugal. (aceite para publicação).
- T. Jorge, J. Magalhães, R. Silva, A. Guedes, D. Ribeiro, A. Meixedo, A. Mosleh, C. Vale, P. Montenegro - "A Deep Autoencoder approach for OOR Damage Wheels Identification in a Railway Freight Vehicle", Proceedings of International Conference on Vehicle Body Engineering (VBE 2023), Porto, Portugal. (aceite para publicação).

### 6.3.2 Artigos em revistas internacionais

- T. Jorge, J. Magalhães, R. Silva, A. Guedes, D. Ribeiro, A. Meixedo, A. Mosleh, C. Vale, P. Montenegro (2023) - "Early Identification of Out of Roundness Damage Wheels in Railway Freight Vehicles using a Wayside System and a Stacked Sparse Autoencoder", Vehicle System Dynamics. (artigo em revisão).
- J. Magalhães, T. Jorge, R. Silva, A. Guedes, D. Ribeiro, A. Meixedo, A. Mosleh, C. Vale, P. Montenegro - "A strategy for out-of-roundness damage wheels identification in a passenger wagon based on sparse autoencoders", Railway Engineering Science. (artigo em revisão).

## REFERÊNCIAS BIBLIOGRÁFICAS

1. Shaikh, K., I. Hussain, and B.S. Chowdhry, *Wheel Defect Detection Using a Hybrid Deep Learning Approach*. *Sensors*, 2023. **23**(14): p. 6248.
2. Lagnebäck, R., *Evaluation of wayside condition monitoring technologies for condition-based maintenance of railway vehicles*. 2007.
3. Bian, J., Y. Gu, and M. Murray, *A dynamic wheel–rail impact analysis of railway track under wheel flat by finite element analysis*. *Vehicle System Dynamics*, 2013. **51**.
4. Yang, J., et al., *Influence of wheel flat on railway vehicle helical gear system under Traction/Braking conditions*. *Engineering Failure Analysis*, 2022. **134**: p. 106022.
5. Wu, X., et al., *Influence of a flexible wheelset on the dynamic responses of a high-speed railway car due to a wheel flat*. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 2017. **232**: p. 095440971770889.
6. Chen, Y.Z., et al., *The influence of wheel flats formed from different braking conditions on rolling contact fatigue of railway wheel*. *Engineering Failure Analysis*, 2018. **93**: p. 183-199.
7. Tao, G., et al., *Polygonisation of railway wheels: a critical review*. *Railway Engineering Science*, 2020. **28**(4): p. 317-345.
8. Tao, G., et al., *An investigation into the mechanism of high-order polygonal wear of metro train wheels and its mitigation measures*. *Vehicle System Dynamics*, 2021. **59**(10): p. 1557-1572.
9. Fröhling, R., U. Spangenberg, and E. Reitmann, *Root cause analysis of locomotive wheel tread polygonisation*. *Wear*, 2019. **432-433**: p. 102911.
10. Chang, Q. and M. Han, *Study on a combined method of Butterworth high-pass filter and Canny edge detector in the application of detecting cargo loading position on railway vehicles*. 2011. **3**.
11. Ding, Y. and L.R. Cheng. *High Speed Overload and Imbalance Load System in China*. in *2018 Joint Rail Conference*. 2018.
12. Mishra, S., P. Sharan, and K. Saara, *Real time implementation of fiber Bragg grating sensor in monitoring flat wheel detection for railways*. *Engineering Failure Analysis*, 2022. **138**: p. 106376.
13. Lourenço, A., et al., *Adaptive time series representation for out-of-round railway wheels fault diagnosis in wayside monitoring*. *Engineering Failure Analysis*, 2023. **152**: p. 107433.
14. Sun, Q., et al., *An on-board detection framework for polygon wear of railway wheel based on vibration acceleration of axle-box*. *Mechanical Systems and Signal Processing*, 2021. **153**: p. 107540.
15. Ye, Y., et al., *Multislice Time-Frequency image Entropy as a feature for railway wheel fault diagnosis*. *Measurement*, 2023. **216**: p. 112862.
16. Ye, Y., et al., *Shock detection of rotating machinery based on activated time-domain images and deep learning: An application to railway wheel flat detection*. *Mechanical Systems and Signal Processing*, 2023. **186**: p. 109856.

17. Silva, R., et al. *Early Identification of Unbalanced Freight Traffic Loads Based on Wayside Monitoring and Artificial Intelligence*. Sensors, 2023. **23**, DOI: 10.3390/s23031544.
18. Jianjun, P., L. Weilai, and D. Xin. *Train overload and unbalanced load detection based on FBG gauge*. in *Proc.SPIE*. 2013.
19. Worden, K. and J. Dulieu-Barton, *An Overview of Intelligent Fault Detection in Systems and Structures*. Structural Health Monitoring-an International Journal - STRUCT HEALTH MONIT, 2004. **3**: p. 85-98.
20. Mohammadi, M., et al. *An Unsupervised Learning Approach for Wayside Train Wheel Flat Detection*. Sensors, 2023. **23**, DOI: 10.3390/s23041910.
21. Mosleh, A., et al., *Early wheel flat detection: an automatic data-driven wavelet-based approach for railways*. Vehicle System Dynamics, 2023. **61**(6): p. 1644-1673.
22. Meixedo, A., et al., *Real-Time Unsupervised Detection of Early Damage in Railway Bridges Using Traffic-Induced Responses*. 2021. p. 117-142.
23. Xu, M., et al., *Structural damage detection by integrating robust PCA and classical PCA for handling environmental variations and imperfect measurement data*. Advances in Structural Engineering, 2022. **25**: p. 136943322210790.
24. Bisheh, H.B. and G.G. Amiri, *Structural damage detection based on variational mode decomposition and kernel PCA-based support vector machine*. Engineering Structures, 2023. **278**: p. 115565.
25. Meixedo, A.G., *Damage identification in railway bridges based on train induced dynamic responses*. 2021.
26. Mosleh, A., et al., *Automatic clustering-based approach for train wheels condition monitoring*. International Journal of Rail Transportation, 2022: p. 1-26.
27. Guedes, A., et al. *Detection of Wheel Polygonization Based on Wayside Monitoring and Artificial Intelligence*. Sensors, 2023. **23**, DOI: 10.3390/s23042188.
28. Javed, K., et al., *Enabling Health Monitoring Approach Based on Vibration Data for Accurate Prognostics*. IEEE Transactions on Industrial Electronics, 2014. **62**.
29. Cavadas, F., I. Smith, and J. Figueiras, *Damage detection using data-driven methods applied to moving-load responses*. Mechanical Systems and Signal Processing, 2013. **39**: p. 409-425.
30. McCulloch, W.S. and W. Pitts, *A logical calculus of the ideas immanent in nervous activity*. The bulletin of mathematical biophysics, 1943. **5**(4): p. 115-133.
31. Olshausen, B.A. and D.J. Field, *Sparse coding with an overcomplete basis set: A strategy employed by V1?* Vision Research, 1997. **37**(23): p. 3311-3325.
32. Ng, A., *Sparse autoencoder*.
33. Yu, Y., et al., *Automated damage diagnosis of concrete jack arch beam using optimized deep stacked autoencoders and multi-sensor fusion*. Developments in the Built Environment, 2023. **14**: p. 100128.
34. Yessoufou, F. and J. Zhu, *Deep autoencoder model for direct monitoring of bridges subjected to a moving vehicle load under varying temperature conditions*. Structures, 2023. **52**: p. 752-767.
35. Amaral, R., et al., *Structural novelty detection based on sparse autoencoders and control charts*. Structural Engineering & Mechanics, 2022. **81**: p. 647-664.
36. Finotti, R.P., et al., *Numerical and Experimental Evaluation of Structural Changes Using Sparse Auto-Encoders and SVM Applied to Dynamic Responses*. Applied Sciences, 2021. **11**(24): p. 11965.

37. Nick, W., et al., *A Study of Machine Learning Techniques for Detecting and Classifying Structural Damage*. International Journal of Machine Learning and Computing, 2015. **5**: p. 313-318.
38. Chen, Z., et al., *Acoustic emission analysis of crack type identification of corroded concrete columns under eccentric loading: A comparative analysis of RA-AF method and Gaussian mixture model*. Case Studies in Construction Materials, 2023. **18**: p. e02021.
39. Fiorini, L., et al., *Unsupervised emotional state classification through physiological parameters for social robotics applications*. Knowledge-Based Systems, 2019. **190**.
40. Meixedo, A., et al., *Online unsupervised detection of structural changes using train-induced dynamic responses*. Mechanical Systems and Signal Processing, 2022. **165**: p. 108268.
41. Osman, A., et al., *A Naïve-Bayes classifier for damage detection in engineering materials*. Materials & Design, 2007. **28**: p. 2379-2386.
42. Vitola, J., et al., *Distributed Piezoelectric Sensor System for Damage Identification in Structures Subjected to Temperature Changes*. Sensors, 2017. **17**(6): p. 1252.
43. Shafique, R., et al., *A Novel Approach to Railway Track Faults Detection Using Acoustic Analysis*. Sensors, 2021. **21**(18): p. 6221.
44. Sohn, H. and K.H. Law, *Bayesian probabilistic damage detection of a reinforced-concrete bridge column*. Earthquake Engineering & Structural Dynamics, 2000. **29**: p. 1131-1152.
45. Mosleh, A., et al., *10 - Approaches for weigh-in-motion and wheel defect detection of railway vehicles*, in *Rail Infrastructure Resilience*, R. Calçada and S. Kaewunruen, Editors. 2022, Woodhead Publishing. p. 183-207.
46. Ngigi, R.W., et al., *Modern techniques for condition monitoring of railway vehicle dynamics*. Journal of Physics: Conference Series, 2012. **364**.
47. Zhang, Z., et al., *Enhanced fault diagnosis of roller bearing elements using a combination of empirical mode decomposition and minimum entropy deconvolution*. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 2015. **231**(4): p. 655-671.
48. Guedes, A.M.C., *Deteção e identificação de dano em rodas de veículos ferroviários baseada na resposta dinâmica induzida pelo veículo na via*. 2022.
49. Português, P.C., *Privatização da CP Carga – atentado ao interesse nacional*. 2016.
50. WAY4SAFERAIL, ed. *Workshop - Data-driven damage detection in railway systems*. 2023.
51. Liao, D., *Assignment: Out-of-round railway wheels: A review*.
52. *EN 15313-2016 - Railway applications - In-service wheelset operation requirements - In-service and off-vehicle wheelset maintenance*. 2016.
53. Peng, B., *Mechanisms of Railway Wheel Polygonization*. 2020.
54. Wang, T., et al., *Research of Wheel-Rail Coupled Vibration Excited by Wheel Polygon for Heavy-Haul Locomotive*. IOP Conference Series: Earth and Environmental Science, 2018. **189**: p. 062026.
55. Brizuela, J., C. Fritsch, and A. Ibáñez, *Railway wheel-flat detection and measurement by ultrasound*. Transportation Research Part C: Emerging Technologies, 2011. **19**(6): p. 975-984.
56. Vale, C.M.N.A.S.d., *Influência da qualidade dos sistemas ferroviários no comportamento dinâmico e no planeamento da manutenção preventiva de vias de alta velocidade*. 2010.
57. Cao, W.-J., *Train wheel damage detection based on deep learning*. 2020.
58. Wu, T.X. and D.J. Thompson. *A hybrid model for wheel/track dynamic interaction and noise generation due to wheel flats*. 2001.

59. Ren, Z., *Multi-point contact of the high-speed vehicle-turnout system dynamics*. Chinese Journal of Mechanical Engineering, 2013. **26**.
60. Zunsong, R., *An investigation on wheel/rail impact dynamics with a three-dimensional flat model*. Vehicle System Dynamics, 2019. **57**(3): p. 369-388.
61. Dukkipati, R.V. and R. Dong, *Impact Loads due to Wheel Flats and Shells*. Vehicle System Dynamics, 1999. **31**: p. 1-22.
62. Bian, J., Y. Gu, and M.H. Murray, *A dynamic wheel–rail impact analysis of railway track under wheel flat by finite element analysis*. Vehicle System Dynamics, 2013. **51**(6): p. 784-797.
63. Ye 叶运广, Y., et al., *Wheel flat can cause or exacerbate wheel polygonization*. Vehicle System Dynamics, 2019.
64. Steenbergen, M., *The role of the contact geometry in wheel–rail impact due to wheel flats: Part II*. Vehicle System Dynamics, 2008. **46**: p. 713-737.
65. Uzzal, R.U.A., A.K.W. Ahmed, and R.B. Bhat, *Modelling, validation and analysis of a three-dimensional railway vehicle–track system model with linear and nonlinear track properties in the presence of wheel flats*. Vehicle System Dynamics, 2013. **51**(11): p. 1695-1721.
66. Vernersson, T., *Thermally induced roughness of tread-braked railway wheels: Part 1: brake rig experiments*. Wear, 1999. **236**(1): p. 96-105.
67. Iwnicki, S., J.C.O. Nielsen, and G. Tao, *Out-of-round railway wheels and polygonisation*. Vehicle System Dynamics, 2023. **61**(7): p. 1787-1830.
68. Morys, B., *ENLARGEMENT OF OUT-OF-ROUND WHEEL PROFILES ON HIGH SPEED TRAINS*. Journal of Sound and Vibration, 1999. **227**(5): p. 965-978.
69. Langemann, D., *Numerical analysis of the polygonalization of railway wheels*. ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik, 2001. **81**(S3): p. 619-620.
70. Meywerk, M., *Polygonalization of railway wheels*. Archive of Applied Mechanics, 1999. **69**(2): p. 105-120.
71. Johansson, A., *Out-of-round railway wheels—assessment of wheel tread irregularities in train traffic*. Journal of Sound and Vibration, 2006. **293**: p. 795-806.
72. Johansson, A. and C. Andersson, *Out-of-round railway wheels—a study of wheel polygonalization through simulation of three-dimensional wheel–rail interaction and wear*. Vehicle System Dynamics, 2005. **43**(8): p. 539-559.
73. Tao, G., et al., *Measurement and assessment of out-of-round electric locomotive wheels*. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 2016. **232**.
74. Fröhling, R., U. Spangenberg, and E. Reitmann, *Root cause analysis of locomotive wheel tread polygonisation*. Wear, 2019. **432-433**.
75. Tao, G., et al., *Polygonisation of railway wheels: a critical review*. Railway Engineering Science, 2020. **28**: p. 1-29.
76. Nielsen, J., A. Johansson, and T. Vernersson, *Train-Track Interaction and Mechanisms of Irregular Wear on Wheel and Rail Surfaces*. Vehicle System Dynamics - VEH SYST DYN, 2003. **40**: p. 3-54.
77. Xuesong Jin, G.T., Xin Zhao, Wei Li, Zefeng Wen *Wear and RCF problems of metro wheel/rail systems: phenomena, causes and countermeasures in China*. 12th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems, 2022.

78. Cai, W., et al., *Experimental and Numerical Investigation into Formation of Metro Wheel Polygonalization*. Shock and Vibration, 2019. **2019**: p. 1538273.
79. (UIC), I.U.o.R., *Loading guideline - Code of practice for the loading and securing of goods on railway wagons*. 2022. **1**.
80. Ribeiro, D.R.F., *Efeitos dinâmicos induzidos por tráfego em pontes ferroviárias: modelação numérica, calibração e validação experimental*. 2012, Faculdade de Engenharia do Porto.
81. Almeida, P.A.M.M.e., *A methodology for the assessment of the train running safety on bridges*. 2015.
82. Frýba, L., *DYNAMICS OF RAILWAY BRIDGES*. 1996: Thomas Telford Publishing. -1.
83. Goicolea, J. and P. Antolín, *The Dynamics of High-Speed Railway Bridges: A Review of Design Issues and New Research for Lateral Dynamics*. International Journal of Railway Technology, 2012. **1**: p. 27-55.
84. Neves, S.G.M., A.F.M. Azevedo, and R. Calçada, *A direct method for analyzing the vertical vehicle-structure interaction*. Engineering Structures, 2012. **34**: p. 414-420.
85. Montenegro, P.A., et al., *Impact of the train-track-bridge system characteristics in the runnability of high-speed trains against crosswinds - Part I: Running safety*. Journal of Wind Engineering and Industrial Aerodynamics, 2022. **224**: p. 104974.
86. Gao, R., Q. He, and Q. Feng *Railway Wheel Flat Detection System Based on a Parallelogram Mechanism*. Sensors, 2019. **19**, DOI: 10.3390/s19163614.
87. Wang, R., et al., *Vibration-Based Detection of Wheel Flat on a High-Speed Train*. 2020. p. 159-169.
88. Mosleh, A., et al., *An approach for wheel flat detection of railway train wheels using envelope spectrum analysis*. Structure and Infrastructure Engineering, 2021. **17**(12): p. 1710-1729.
89. Mosleh, A., et al. *Railway Vehicle Wheel Flat Detection with Multiple Records Using Spectral Kurtosis Analysis*. Applied Sciences, 2021. **11**, DOI: 10.3390/app11094002.
90. Trilla, A., et al. *Integrated Multiple-Defect Detection and Evaluation of Rail Wheel Tread Images using Convolutional Neural Networks*. 2021.
91. Komorski, P., et al., *Advanced acoustic signal analysis used for wheel-flat detection*. Latin American Journal of Solids and Structures, 2021. **18**.
92. Song, Y., Z. Wang, and Y. Du, *Study on Train Wheel Out-of-Roundness Monitoring Method by PVDF Sensing Technology*. The Open Mechanical Engineering Journal, 2014. **8**: p. 77-80.
93. Fang, L., et al. *Method of Wheel Out-of-Roundness Detection Based on POVM and Multinuclear LS-SVM*. in *Proceedings of the 4th International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2019*. 2020. Singapore: Springer Singapore.
94. Xu, X., et al., *Detection Method for Polygonalization of Wheel Treads Based on Dynamic Response*, in *Proceedings of the 2020 2nd International Conference on Robotics Systems and Vehicle Technology*. 2021, Association for Computing Machinery: Xiamen, China. p. 16–20.
95. Huang, H., et al., *A Fault Diagnosis Method for Out-of-Round Faults of Metro Vehicle Wheels with Strong Noise*. Shock and Vibration, 2021. **2021**: p. 9257622.
96. Allotta, B., et al., *An innovative high speed Weigh in Motion system for railway vehicles*. 2014. 1-7.
97. Addison, P.S., *Introduction to redundancy rules: the continuous wavelet transform comes of age*. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2018. **376**(2126): p. 20170258.

98. Iberdrola. *'Deep learning': um conceito-chave para levar a inteligência artificial a um nível superior.*
99. Book, D.L., *Capítulo 58 – Introdução aos Autoencoders.*
100. Ceccon, D. *Os tipos de redes neurais.* 2020.
101. Ceccon, D., *A importância do bias nas redes neurais.* 2020.
102. M.Santos, J. *Redes neuronais artificiais - aplicações e potencialidades.*
103. Siqueira, R.F. *Redução de dimensionalidade em bases de dados de classificação hierárquica multirrótulo usando autoencoders* 2019.
104. Kazak, V., *Unsupervised feature extraction with autoencoder for the representation of parkinson's disease patients.* 2018.
105. Tech, D. *O que são Redes Neurais e Deep Learning?* 2020.
106. Yan, B. and G. Han, *Effective Feature Extraction via Stacked Sparse Autoencoder to Improve Intrusion Detection System.* IEEE Access, 2018. **6**: p. 1-1.
107. Yan, B. and G. Han, *Effective Feature Extraction via Stacked Sparse Autoencoder to Improve Intrusion Detection System.* IEEE Access, 2018. **6**: p. 41238-41248.
108. Klein, L., *Sensor and data fusion: A tool for information assessment and decision making: Second edition.* 2012. 1-474.
109. Worden, K., H. Sohn, and C.R. Farrar, *NOVELTY DETECTION IN A CHANGING ENVIRONMENT: REGRESSION AND INTERPOLATION APPROACHES.* Journal of Sound and Vibration, 2002. **258**(4): p. 741-761.
110. Bouzenad, A.E., et al. *A Semi-Supervised Based K-Means Algorithm for Optimal Guided Waves Structural Health Monitoring: A Case Study.* Inventions, 2019. **4**, DOI: 10.3390/inventions4010017.
111. Meixedo, A., et al., *Structural health monitoring strategy for damage detection in railway bridges using traffic induced dynamic responses.* 2022. p. 389-408.
112. de Almeida Cardoso, R., A. Cury, and F. Barbosa, *Automated real-time damage detection strategy using raw dynamic measurements.* Engineering Structures, 2019. **196**: p. 109364.
113. Figueiredo, E. and E. Cross, *Linear approaches to modeling nonlinearities in long-term monitoring of bridges.* Journal of Civil Structural Health Monitoring, 2013. **3**(3): p. 187-194.
114. Montenegro, P.A., et al., *Probabilistic numerical evaluation of dynamic load allowance factors in steel modular bridges using a vehicle-bridge interaction model.* Engineering Structures, 2020. **226**.
115. Bragança, C., et al., *Calibration and validation of a freight wagon dynamic model in operating conditions based on limited experimental data.* Vehicle System Dynamics, 2022. **60**(9): p. 3024-3050.
116. ANSYS®, *Academic Research, Release 19.2; Ansys Academic: Canonsburg, PA, USA.* 2018.
117. Fries, R.H. and B.M. Coffey, *A State-Space Approach to the Synthesis of Random Vertical and Crosslevel Rail Irregularities.* Journal of Dynamic Systems, Measurement, and Control, 1990. **112**(1): p. 83-87.
118. Hamid, A. and T.-L. Yang, *ANALYTICAL DESCRIPTIONS OF TRACK-GEOMETRY VARIATIONS.* Transportation Research Record, 1982.
119. Mosleh, A., P.A. Costa, and R. Calçada, *A new strategy to estimate static loads for the dynamic weighing in motion of railway vehicles.* Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 2020. **234**(2): p. 183-200.



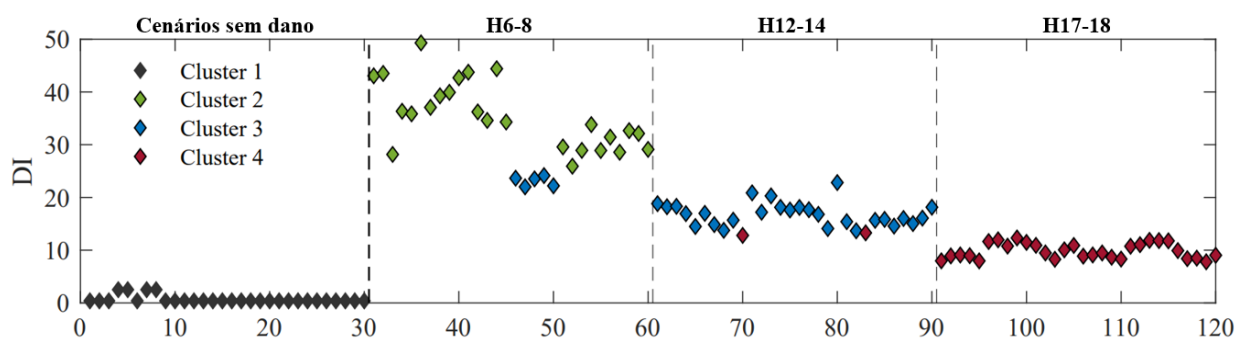
120. Wu, X., et al., *A study of formation of high order wheel polygonalization*. *Wear*, 2019. **424-425**: p. 1-14.
121. Montenegro, P.A., et al., *Wheel–rail contact formulation for analyzing the lateral train–structure dynamic interaction*. *Computers & Structures*, 2015. **152**: p. 200-214.
122. Montenegro, P.A., et al., *A comparative study on the running safety of trains subjected to crosswinds simulated with different wind models*. *Journal of Wind Engineering and Industrial Aerodynamics*, 2020. **207**: p. 104398.
123. *EN 13848-2: Railway Applications—Track—Track Geometry Quality—Part 2: Measuring Systems—Track Recording Vehicles*. 2006.
124. Rytter, A., R. Brincker, and L.P. Hansen. *Vibration Based Inspection of Civil Engineering Structures*. 1993.
125. Rabiner, L.R., *A tutorial on hidden Markov models and selected applications in speech recognition*. *Proceedings of the IEEE*, 1989. **77**(2): p. 257-286.



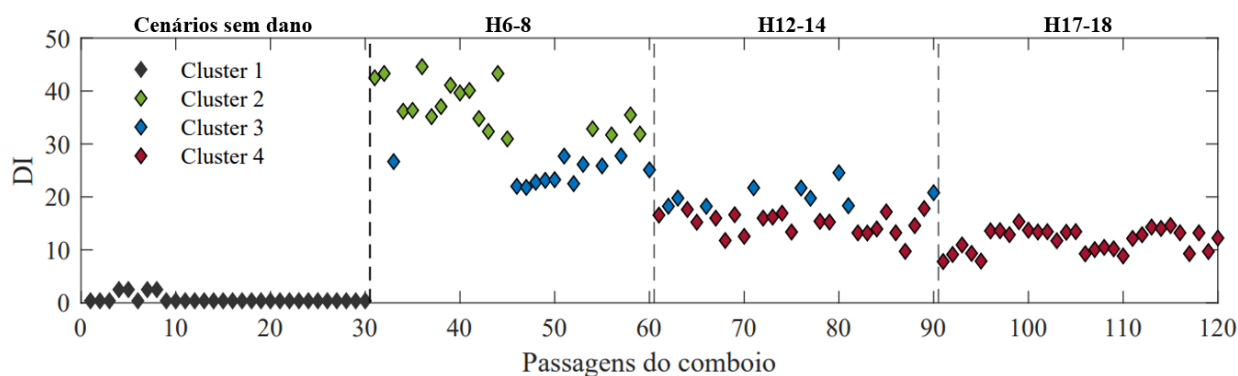
## ANEXO 1

Neste anexo são apresentados os resultados obtidos para a análise de sensibilidade de sensores realizada para os 3 tipos de defeitos, conforme apresentado no Capítulo 4. Na secção 1.1 serão apresentados os resultados para a Metodologia 1 e na secção 1.2 para a Metodologia 2. É ainda apresentada uma tabela que resume de forma sucinta os resultados obtidos.

### 1.1 Metodologia 1 – Análise de sensibilidade ao número de sensores para poligonização

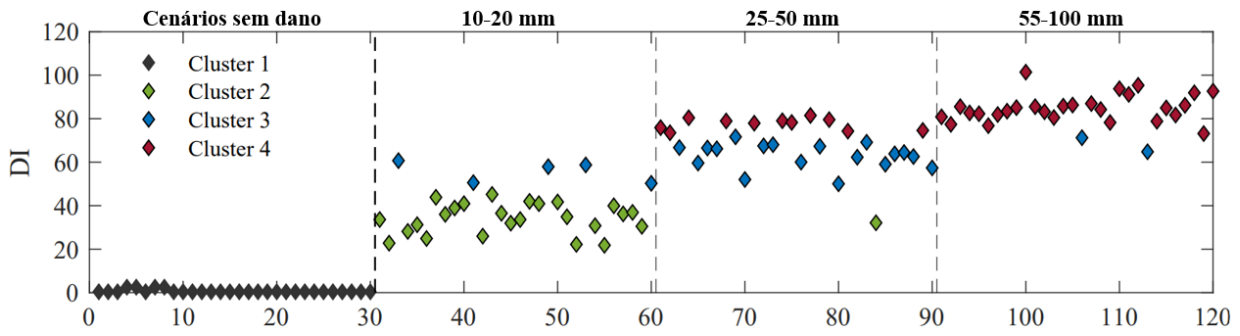


a)

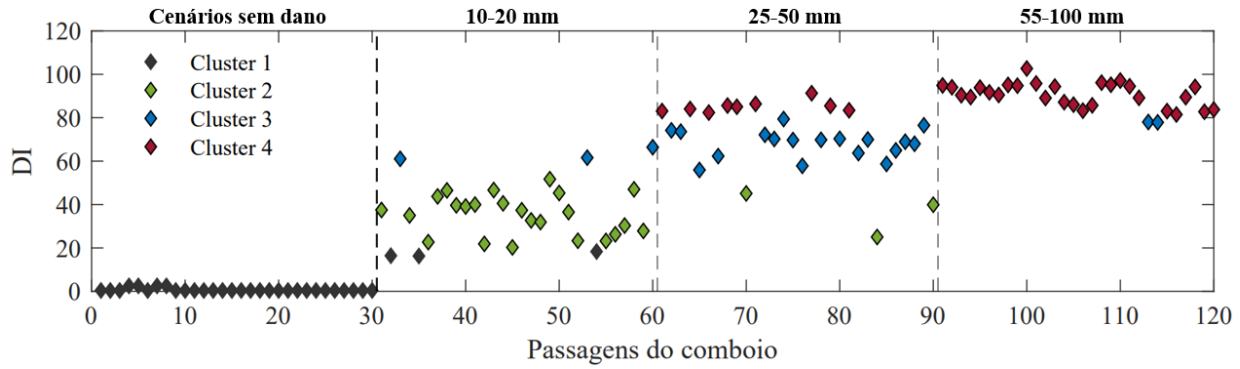


b)

Figura A.1.1.1 – Classificação automática da severidade da poligonização com base na análise de clusters para 3 velocidades, utilizando: a) 4 acelerómetros b) 2 acelerómetros.

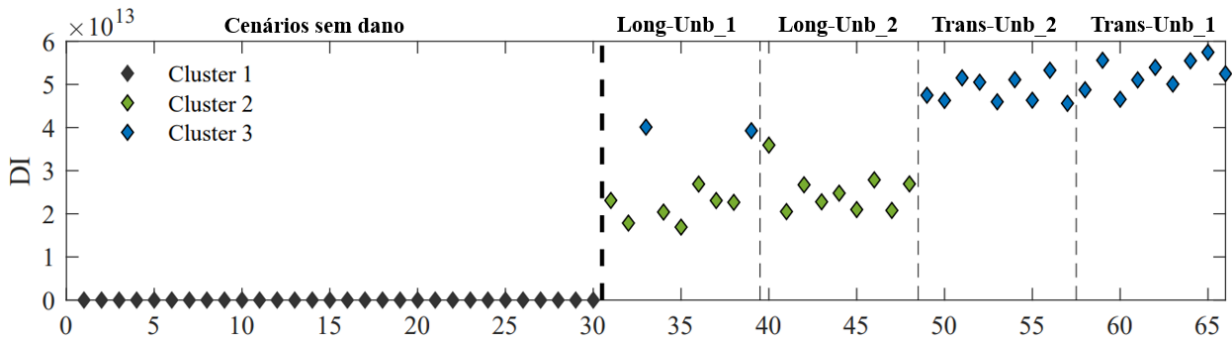


a)

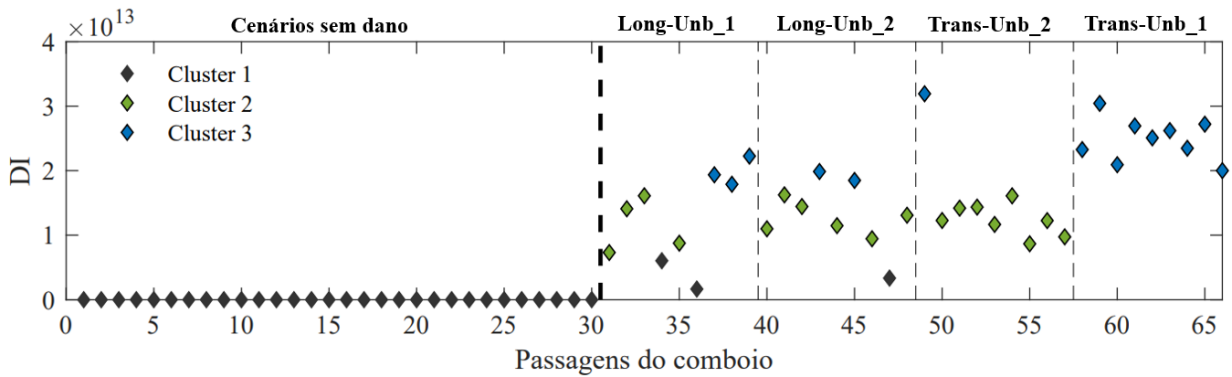


b)

Figura A.1.1.2 – Classificação automática da severidade dos lisos com base na análise de clusters para 3 velocidades, utilizando: a) 4 acelerómetros b) 2 acelerómetros.



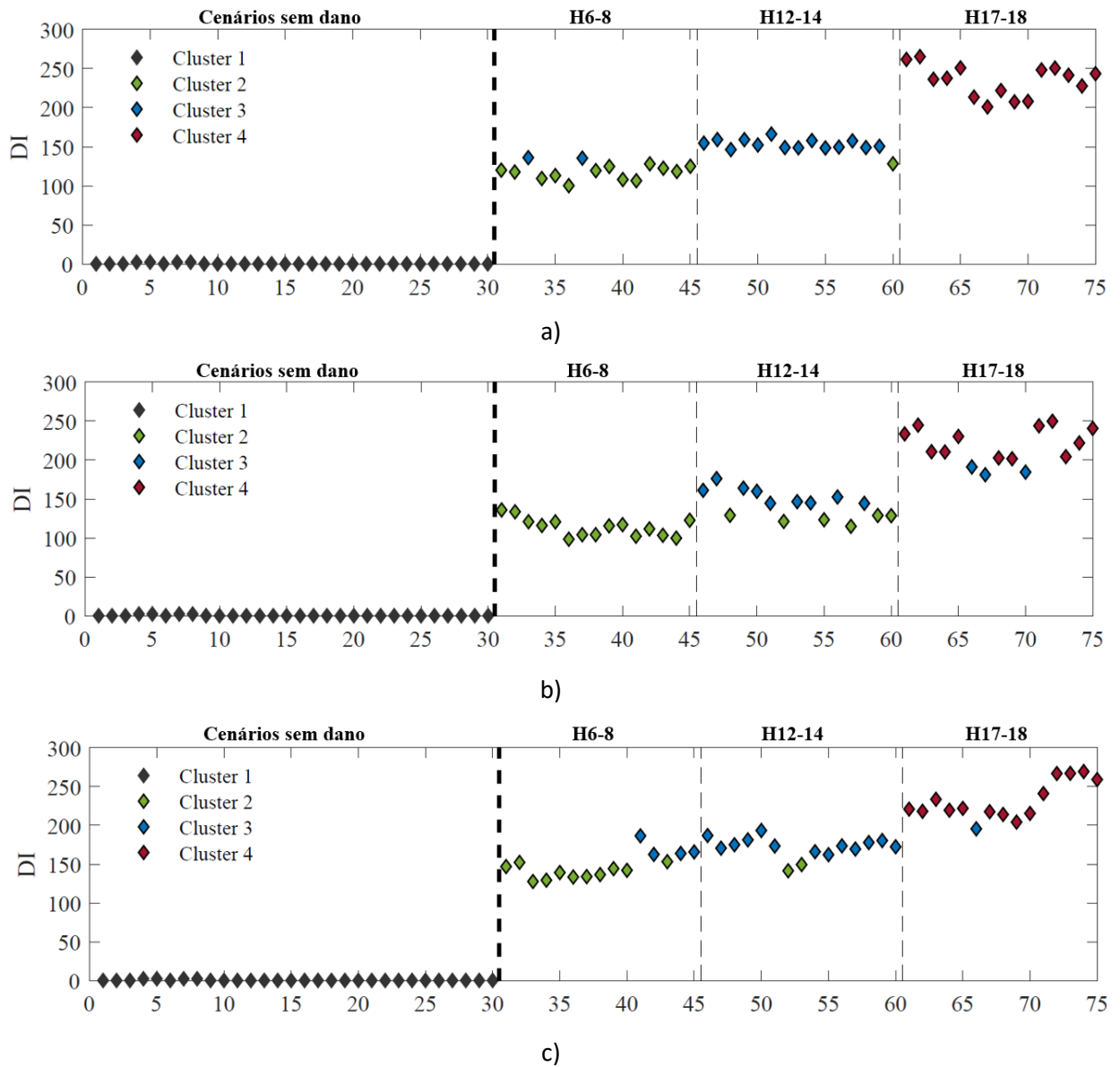
a)

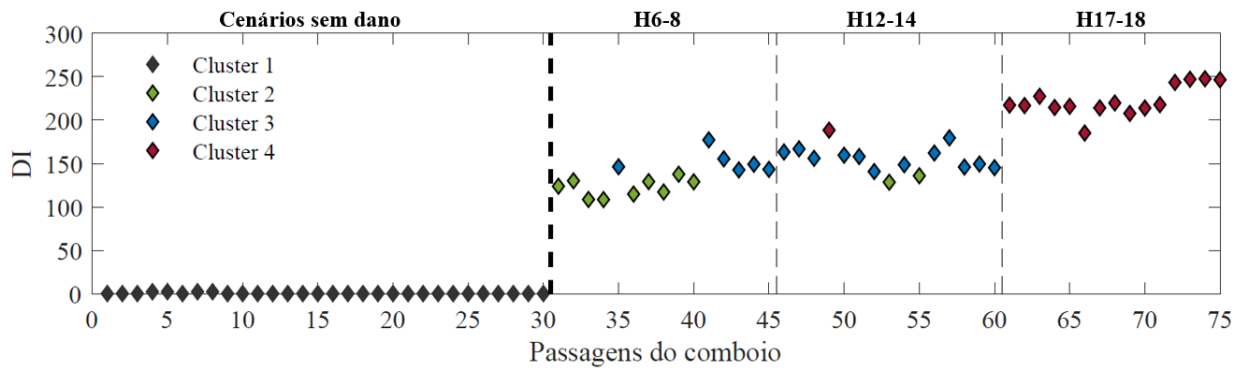


b)

Figura A.1.1.3 – Classificação automática da severidade de desequilíbrios de carga com base na análise de clusters para 3 velocidades, utilizando: a) 8 extensómetros b) 6 extensómetros.

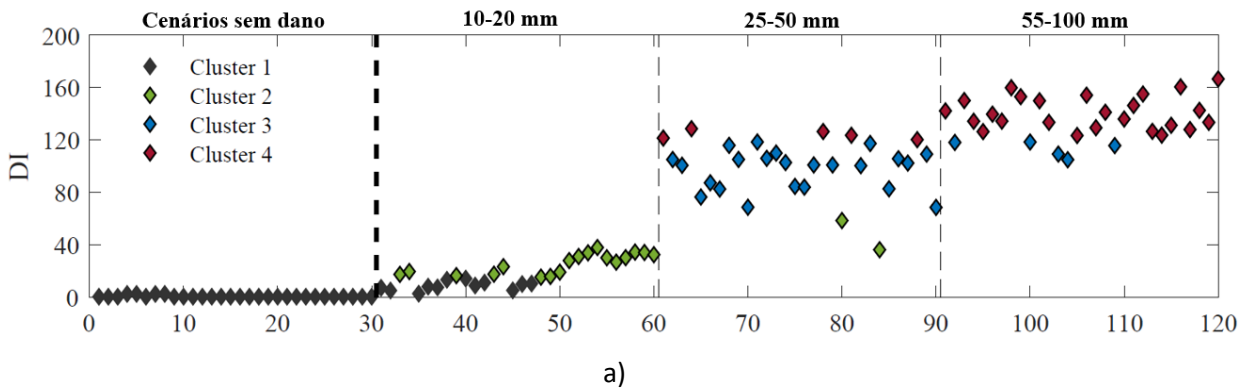
1.2 Metodologia 2 – Análise de sensibilidade ao número de sensores



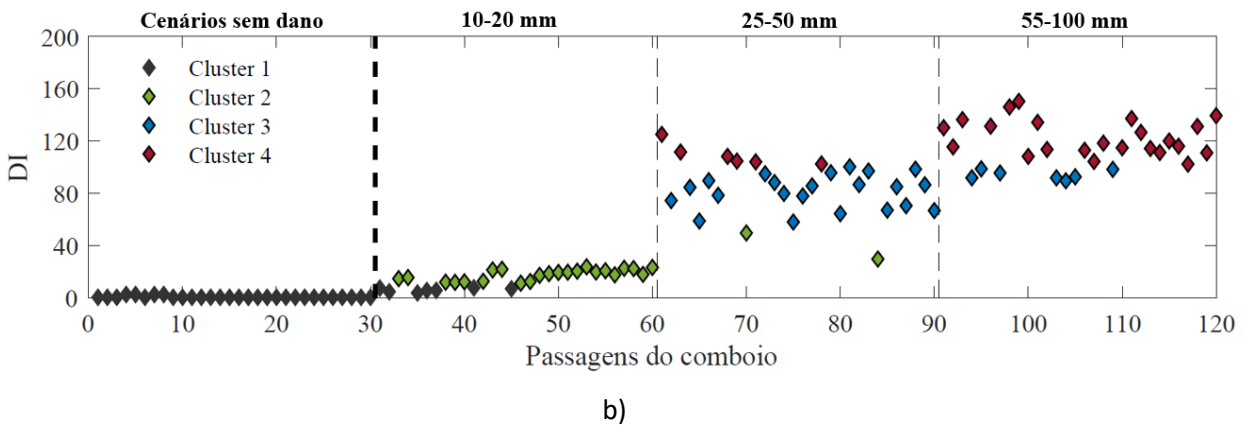


d)

Figura A.1.2.1 – Classificação automática da severidade da poligonização com base na análise de clusters para 3 velocidades, utilizando: a) c) 4 acelerómetros b) d) 2 acelerómetros.

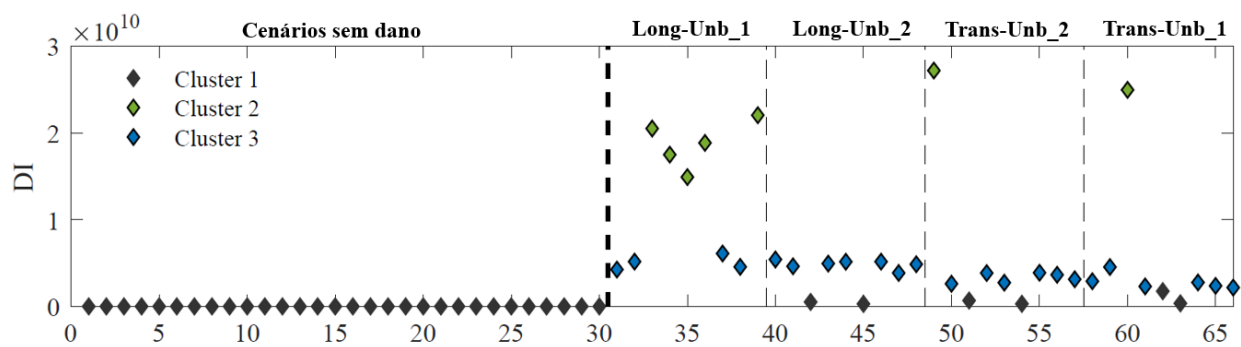


a)

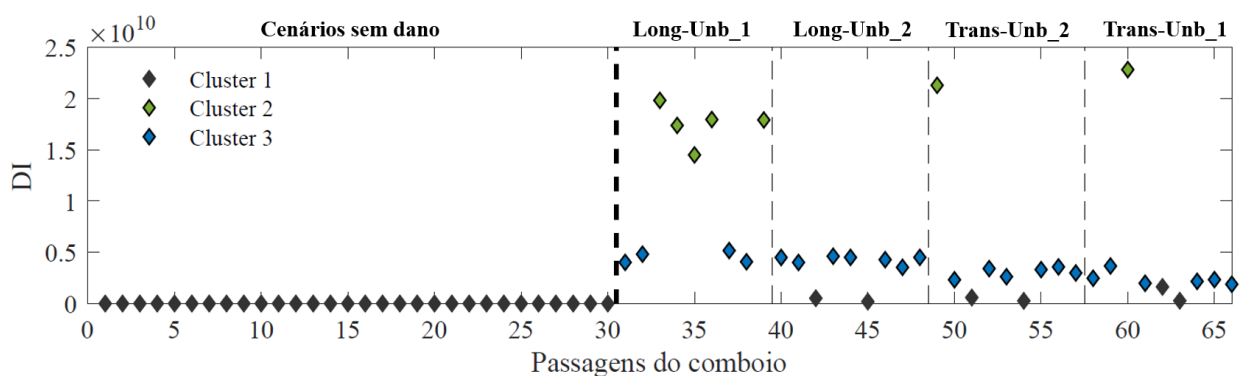


b)

Figura A.1.2.2 – Classificação automática da severidade dos lisos com base na análise de clusters para 3 velocidades, utilizando: a) 4 acelerómetros b) 2 acelerómetros.



a)



b)

Figura A.1.2.2 – Classificação automática da severidade de desequilíbrios de carga com base na análise de clusters para 3 velocidades, utilizando: a) 8 extensómetros b) 6 extensómetros.

Tabela A.1.2.3 – Falsas classificações relacionadas com a análise de sensibilidade dos vários sensores.

Sensores		Metodologia 1			Metodologia 2		
Acelerómetros	Extensómetros	Poligonização (Harmónicos)	Lisos	Desequilíbrio de cargas	Poligonização (Harmónicos)	Lisos	Desequilíbrio de cargas
6 (3+3)	10 (5+5)	6% (5/90)	7% (6/90)	9% (4/36)	5% (4/90)	27% (24/90)	-
4 (2+2)	8 (4+4)	11% (10/90)	21% (19/90)	6% (2/36)	11% (10/90)	20% (18/90)	-
2 (1+1)	6 (3+3)	37% (33/90)	22% (20/90)	22% (8/36)	20% (18/90)	25% (22/90)	-





## ANEXO 2

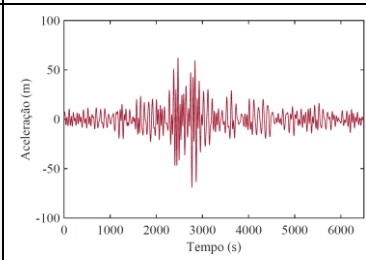
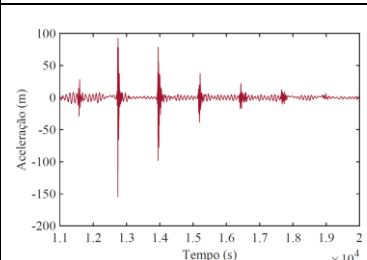
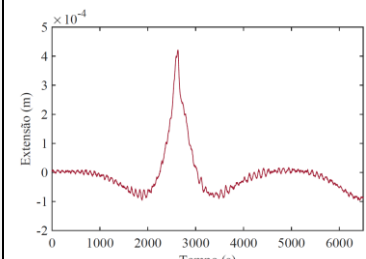
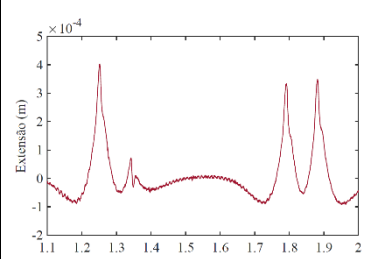
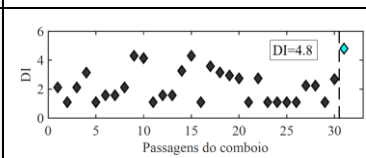
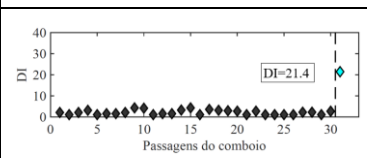
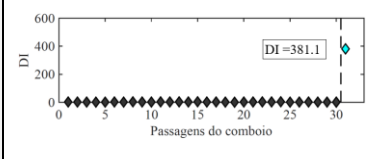
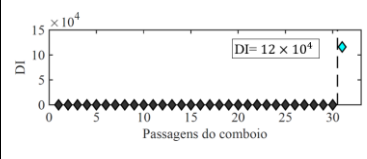
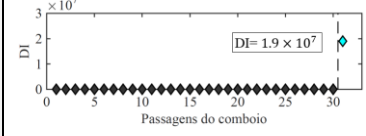
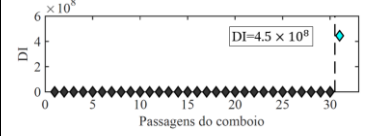
Na Tabela A.2.1 e Tabela A.2.2 são apresentadas as tabelas resumo relativas à etapa de identificação do tipo de dano para a CT2 e CT3, respectivamente.

Tabela A.2.1 – Identificação do tipo de dano com base nos limites de DI definidos na fase de teste (CT2).

CT2	Corte 1	Corte 2	Corte 3
Loc.	1ª roda	3ª roda	5º vagão
Janela de corte			
	Extensões	Extensões	Extensões
	0 – 5	0 – 5	> 5

	Acelerações	80 – 1000	Acelerações	> 1000	Acelerações	0 – 100
	FFT	$0.1 \times 10^7 - 2 \times 10^8$	FFT	$> 2 \times 10^8$	FFT	$0 - 0.1 \times 10^7$
Decisão	Dano Poligonal		Dano Liso		Desequilíbrio de cargas	

Tabela A.2.2 – Identificação do tipo de dano com base nos limites de DI definidos na fase de teste (CT3).

CT3	Corte 1	Corte 2		
Loc.	1ª roda	3º vagão		
Janela de corte				
				
Identificação do dano	Extensões			
	Acelerações			
	FFT			
	Extensões	0 – 5	Extensões	> 5
	Acelerações	80 – 1000	Acelerações	> 1000
	FFT	$0.1 \times 10^7 - 2 \times 10^8$	FFT	$> 2 \times 10^8$
Decisão	Dano Poligonal		Dano Liso Desequilíbrio de cargas	