



UNIVERSIDADE FEDERAL DE OURO PRETO  
ESCOLA DE MINAS  
DEPARTAMENTO DE ENGENHARIA CIVIL  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL



## **Estudo de novas estratégias para identificação de danos estruturais a partir de dados vibracionais**

**AUTOR: VINÍCIUS NICCHIO ALVES**

ORIENTADOR: Prof. Dr. Alexandre Abrahão Cury

Dissertação apresentada ao Programa de Pós-Graduação do Departamento de Engenharia Civil da Escola de Minas da Universidade Federal de Ouro Preto, como parte integrante dos requisitos para obtenção do título de Mestre em Ciências da Engenharia Civil, área de concentração: Construção Metálica.

Ouro Preto, outubro de 2012.

A474e

Alves, Vinicius Nicchio.

Estudo de novas estratégias para identificação de danos estruturais a partir de dados vibracionais [manuscrito] / Vinicius Nicchio Alves - 2012.

vi, 188f.: il. color.; graf. ; tabs.

Orientador: Prof. Dr. Alexandre Abrahão Cury.

Dissertação (Mestrado) - Universidade Federal de Ouro Preto. Escola de Minas. Departamento de Engenharia Civil. Programa de Pós-Graduação em Engenharia Civil.

Área de concentração: Construção Metálica.

1. Dinâmica estrutural - Teses. 2. Otimização estrutural - Teses. 3. Análise de dados simbólicos - Teses. 4. Análise de clusters - Aglomeração hierárquica - Teses. I. Universidade Federal de Ouro Preto. II. Título.

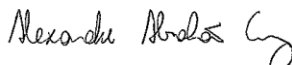
CDU: 624.014:004.93'14

Catálogo: [sisbin@sisbin.ufop.br](mailto:sisbin@sisbin.ufop.br)

**ESTUDO DE NOVAS ESTRATÉGIAS PARA IDENTIFICAÇÃO DE  
DANOS ESTRUTURAIS A PARTIR DE DADOS VIBRACIONAIS**

**AUTOR: VINICIUS NICCHIO ALVES**

Esta dissertação foi apresentada em sessão pública e aprovada em 23 de outubro de 2012, pela Banca Examinadora composta pelos seguintes membros:



Prof. Dr. Alexandre Abrahão Cury (Orientador / UFOP)



Prof. Dr. Marcílio Sousa da Rocha Freitas (UFOP)



Prof. Dr. Ney Roitman (UFRJ)

## RESUMO

O monitoramento da integridade de uma estrutura baseia-se na utilização de indicadores confiáveis e robustos que permitam detectar, localizar, quantificar e, se possível, prever a ocorrência de danos. Os estudos relacionados à detecção de danos em estruturas de Engenharia Civil são de notável interesse, visto que a detecção de uma modificação estrutural é de fundamental importância para evitar a ocorrência de graves consequências sociais, econômicas e ambientais.

Recentemente, muitas pesquisas têm se concentrado na avaliação dinâmica como parte do diagnóstico estrutural em um processo conhecido como identificação modal. O objetivo é extrair os parâmetros modais e/ou indicadores construídos a partir destes parâmetros. Estes indicadores têm mostrado as suas potencialidades, mas o problema de sua sensibilidade, a necessidade de um estado de referência e sua confiabilidade em termos de probabilidade de detecção de alarmes falsos permanece sem solução.

Neste contexto, é importante o uso de técnicas capazes de lidar não só com os dados brutos (sinais), mas também com os parâmetros modais de uma forma prática e relevante. Assim sendo, novas representações foram desenvolvidas para melhorar a manipulação e armazenamento desses dados. Essas representações são conhecidas como *dados simbólicos*. O desenvolvimento de novos métodos capazes de lidar com este tipo de dados é o objetivo da Análise de Dados Simbólicos (ADS).

Esta dissertação tem como interesse principal a utilização de diferentes métodos de classificação juntamente com a ADS para detectar danos estruturais. A ideia é utilizar procedimentos de particionamento de dados (aglomerações hierárquicas, nuvens dinâmicas e *clustering fuzzy*, por exemplo) para discriminar diferentes estados estruturais. Neste trabalho, a ADS é aplicada tanto às medições dinâmicas obtidas *in situ* (acelerações) quanto aos parâmetros modais identificados.

A fim de atestar a eficiência das abordagens propostas, estudos de sensibilidade considerando aplicações numéricas e um estudo experimental são realizados. Mostra-se que a ADS juntamente com os métodos de classificação são capazes de distinguir estados estruturais com adequados índices de acerto.

## ABSTRACT

The paradigm of structural health monitoring is based on the development of reliable and robust indicators able to detect, locate, quantify and predict damage. Studies related to damage detection in civil engineering structures have a noticeable interest for researchers in this area. Indeed, the detection of structural changes likely to become critical can avoid the occurrence of major dysfunctions associated with social, economic and environmental consequences.

Recently, many researchers have focused on dynamic assessment as part of structural diagnosis. Most of the studied techniques are based on time or frequency domain analyses to extract compressed information from modal characteristics or based on indicators built from these parameters. These indicators have shown their potentialities, but the problem of their sensitivity, the necessity of a reference state, and their reliability in terms of detection probability and false alarm, still remains.

In this context, it is important to use techniques capable of dealing not only with raw data but also modal parameters in a practical and relevant way. In order to give some insights to this problem, appropriate representations have been developed to improve both manipulation and storage of data. These representations are known as *symbolic data*. The development of new methods capable of dealing with this type of data is the goal of Symbolic Data Analysis (SDA).

This work has as its main interest the use of different methods coupled with the SDA to detect structural damage. The idea is to employ clustering procedures (e.g., hierarchy-agglomerative, dynamic clouds and fuzzy clustering) to discriminate different structural states. In this thesis, SDA is applied to dynamic measurements obtained on site (accelerations) and to the identified modal parameters.

In order to attest the efficiency of the proposed approaches, several sensibility studies considering numerical applications and an experimental investigation are carried out. It is shown that SDA coupled with classification methods is able to distinguish structural conditions with adequate rates.

## SUMÁRIO

1.	INTRODUÇÃO.....	1
1.1.	Considerações Iniciais .....	1
1.2.	Motivação .....	3
1.3.	Objetivos.....	3
1.4.	Apresentação.....	4
2.	REVISÃO BIBLIOGRÁFICA.....	5
2.1.	Estado da arte dos métodos de detecção de danos.....	7
2.1.1.	Métodos baseados nas variações dos parâmetros modais .....	9
2.1.2.	Métodos baseados na variação de indicadores evoluídos .....	11
2.1.3.	Métodos baseados em indicadores de anormalidades .....	13
3.	ANÁLISE DE DADOS SIMBÓLICOS .....	17
3.1.	Transformação de dados clássicos em dados simbólicos .....	18
4.	MÉTODOS DE CLASSIFICAÇÃO DE DADOS .....	25
4.1.	Cálculo das medidas de dissimilaridade .....	26
4.1.1.	Minkowski .....	27
4.1.2.	Cityblock.....	27
4.1.3.	Euclidiana.....	27
4.1.4.	Chebychev.....	27
4.1.5.	Cosine.....	28
4.1.6.	Correlation .....	28
4.1.7.	Spearman.....	28
4.2.	Método das Aglomerações Hierárquicas .....	29
4.2.1.	Critérios de agregação.....	30
4.2.1.1.	Critério do vizinho mais próximo – ‘Single’ (Single Linkage).....	31
4.2.1.2.	Critério do vizinho mais distante – ‘Complete’ (Complete Linkage).....	33
4.2.1.3.	Critério da média do cluster – ‘Average’ (Unweighted pair group average).....	35
4.2.1.4.	Critério da média simples – ‘Weighted’ (Weighted pair group average). .....	37
4.2.1.5.	Critério do centroide – ‘Centroid’ (Unweighted pair group centroid) .....	38
4.2.1.6.	Critério da mediana – ‘Median’ (Weighted pair group centroid).....	40
4.2.1.7.	Critério da mínima variância (Ward) .....	43
4.3.	Método <i>k-means</i> .....	43
4.4.	Método <i>c-means</i> .....	47

4.5. Metodologia.....	50
5. SIMULAÇÕES NUMÉRICAS .....	53
5.1. Análise Clássica.....	57
5.2. Análise Simbólica.....	60
5.2.1. Sinais.....	61
5.2.2. Frequências .....	65
5.2.3. Modos próprios .....	67
5.2.3.1. Primeiro modo próprio.....	68
5.2.3.2. Segundo modo próprio.....	71
5.2.3.3. Terceiro modo próprio .....	74
5.3. Estudo complementar .....	77
6. ENSAIOS EXPERIMENTAIS.....	81
6.1. Ensaios experimentais realizados no viaduto ferroviário PK 075+317.....	81
6.2. Identificação modal do viaduto Sens .....	84
6.3. Análise clássica.....	88
6.4. Análise simbólica.....	91
6.4.1. Sinais.....	92
6.4.2. Frequências .....	94
6.4.3. Modos Próprios .....	96
6.4.3.1. Primeiro modo.....	96
6.4.3.2. Segundo modo.....	98
6.4.3.3. Terceiro modo .....	100
6.4.3.4. Quarto modo.....	102
6.5. Estudo complementar .....	104
CONCLUSÕES .....	107
REFERÊNCIAS BIBLIOGRÁFICAS.....	113

# CAPÍTULO 1

## 1. INTRODUÇÃO

### 1.1. Considerações Iniciais

O monitoramento da integridade de uma estrutura baseia-se na utilização de indicadores confiáveis e robustos que permitam detectar, localizar, quantificar e, se possível, prever a ocorrência de danos. Os estudos relacionados à detecção de danos em estruturas de Engenharia Civil são de notável interesse, visto que a detecção de uma modificação estrutural é de fundamental importância para evitar a ocorrência de graves consequências sociais, econômicas e ambientais. Nos últimos anos, várias pesquisas têm sido realizadas utilizando ensaios dinâmicos para o desenvolvimento de diversos indicadores baseados nas características modais de estruturas, identificadas a partir de uma análise modal. Tais características estruturais são as frequências próprias, as taxas de amortecimento e os modos de vibração estruturais (DOEBLING et al., 1996; SALAWU, 1997; CRÉMONA, 2004). Esses parâmetros são diretamente afetados pela variação das propriedades físicas da estrutura, como por exemplo, de sua massa e, especialmente de sua rigidez. De modo geral, danos estruturais provocam uma redução da rigidez da estrutura e, como consequência, modificam suas características vibracionais (PANDEY; BISWAS, 1994).

A maioria dos indicadores de danos é baseada na análise no domínio do tempo ou no domínio da frequência dos sinais para extrair informações através das características modais ou de indicadores evoluídos construídos a partir dessas características. Tais indicadores se mostraram eficientes, porém ainda existem aspectos que necessitam ser melhorados (CURY, 2010). Muitos indicadores apresentam problemas de sensibilidade, necessitam de um estado de referência e não apresentam a probabilidade de detectar falsos alarmes, diminuindo sua confiabilidade. Além disso, a identificação dos parâmetros modais é, de certa forma, um procedimento de filtragem, o que pode conduzir a uma perda de informações se comparado à utilização dos sinais brutos. Essa perda de informação pode, portanto, mascarar a variação de pequenas amplitudes



causadas por uma modificação estrutural. Outro grande inconveniente relacionado à utilização dos parâmetros modais é o fato de eles descreverem um comportamento linear da estrutura pós-dano, o que nem sempre é correto.

Vários métodos de detecção de danos baseados na análise dos sinais da estrutura podem ser encontrados na literatura, mas eles geralmente falham na prática. A falta de ferramentas eficazes e a dificuldade de se gerenciar grandes quantidades de dados brutos podem explicar essa limitação. Portanto, apesar da capacidade de processamento oferecida atualmente por computadores de mesa, o custo computacional necessário para tratar grandes bases de dados continua sendo um problema. Além disso, a utilização de medições dinâmicas, especialmente quando vários canais de aquisição são utilizados, gera um grande volume de dados a ser armazenado. Medições dinâmicas podem facilmente conter milhares de valores, tornando o processo de análise difícil ou mesmo proibitivo em certos casos.

Nesse contexto, é importante empregar técnicas que permitam a utilização de dados brutos de maneira prática e pertinente. A mineração de dados (*Data Mining*) consiste em um conjunto de ferramentas que extraem características escondidas de grandes bases de dados brutos (HASTIE et al., 2009). Quanto maior a quantidade de dados armazenados, mais complexa é a exploração das informações contidas, o que justifica a utilização dessas ferramentas. Essa técnica é utilizada em várias áreas, tais como na economia, marketing, detecção de fraudes, etc. A fim de fornecer soluções que resolvam esse problema, representações adaptadas têm sido desenvolvidas para melhorar o armazenamento de dados. Essas representações são conhecidas como Dados Simbólicos. Esse tipo de dado permite caracterizar não só a variação, mas também a incerteza inerente a cada variável considerada. O desenvolvimento de novos métodos de análise adequados para tratar esse tipo de dado é o propósito da Análise dos Dados Simbólicos (ADS) (BILLARD; DIDAY, 2006). Atualmente, a maioria das técnicas desenvolvidas na ADS é extensão de métodos estatísticos existentes.

## **1.2. Motivação**

Poucos trabalhos têm se concentrado na análise direta da resposta de uma estrutura através de suas características sob excitações externas. De fato, os sinais provenientes de medições dinâmicas são raramente utilizados devido à falta de ferramentas apropriadas para sua manipulação e à dificuldade de se utilizá-los diretamente. Isso muitas vezes se deve à presença de ruídos e/ou à dificuldade de manipular esse tipo de dado corretamente.

A principal motivação deste trabalho é poder fornecer novas metodologias para a detecção de modificações (leia-se danos) estruturais. Este trabalho consiste em propor abordagens originais para a identificação de danos estruturais não só através de mudanças nos parâmetros modais (frequências e modos próprios), mas também diretamente a partir da resposta dinâmica da estrutura.

## **1.3. Objetivos**

O objetivo deste trabalho consiste em utilizar diferentes métodos de classificação de dados acoplados à Análise de Dados Simbólicos (ADS) para detectar modificações estruturais. A ADS permite lidar com grandes bases de dados, em que esses dados são compactados sem perder demasiadamente as informações originais. As técnicas de *clustering*, em contrapartida, são utilizadas para dividir os dados em grupos distintos.

A ideia é, portanto, utilizar as técnicas de *clustering* (por exemplo, divisões hierárquicas, k-means e *clustering fuzzy* ou c-means) a fim de discriminar os diferentes estados de integridade estrutural. Nesta dissertação, a metodologia proposta é aplicada tanto às medidas dinâmicas, ou seja, às acelerações, quanto aos parâmetros modais identificados, com o objetivo de detectar as modificações estruturais.

#### **1.4. Apresentação**

Esta dissertação é composta por seis capítulos, incluindo o presente, no qual foi feita uma breve introdução sobre o estudo desenvolvido e os objetivos a serem alcançados.

No segundo capítulo, apresenta-se uma revisão bibliográfica na qual são citados os principais métodos de detecção de danos em estruturas. Tais métodos são os métodos baseados nas variações dos parâmetros modais, os métodos baseados na variação de indicadores evoluídos e os métodos baseados em indicadores de anormalidade.

No terceiro capítulo, faz-se uma introdução aos conceitos da Análise de Dados Simbólicos, sendo explicitados alguns exemplos de como os dados podem ser representados de maneira simbólica e de como é realizada a transformação de dados clássicos em dados simbólicos.

O quarto capítulo descreve os métodos de classificação utilizados no presente trabalho. Alguns parâmetros importantes referentes aos métodos de classificação são apresentados, bem como diversas medidas de dissimilaridade, critérios de agregação e métodos de inicialização utilizados nas aplicações numéricas e experimentais deste trabalho.

No quinto capítulo, uma aplicação numérica de uma viga discretizada por elementos finitos é apresentada. Nela, são simuladas diversas configurações de danos e de ruído. Posteriormente, verifica-se a capacidade de os métodos de classificação propostos diferenciarem as diversas configurações de danos assim como a influência dos níveis de ruído sobre as classificações.

No sexto capítulo, uma aplicação experimental é apresentada. A aplicação em questão é baseada nos ensaios dinâmicos realizados no viaduto ferroviário PK 075+317, na França. O objetivo é verificar o desempenho dos métodos de classificação aplicados a uma estrutura real para classificar os ensaios antes, durante e depois de um procedimento de reforço estrutural nela realizado.

Finalmente, são apresentadas as conclusões e sugestões para trabalhos futuros.

# CAPÍTULO 2

## 2. REVISÃO BIBLIOGRÁFICA

Nos últimos anos, muitas pesquisas vêm sendo realizadas com o objetivo de desenvolver indicadores de danos confiáveis e robustos que permitam detectar, localizar, quantificar e até mesmo prever o tempo de vida restante de uma estrutura. Os estudos têm como principal interesse utilizar a avaliação dinâmica para desenvolver indicadores através das características modais da estrutura, identificadas a partir de uma análise modal. Nesse contexto, várias questões podem surgir. Saber quais são as causas do(s) dano(s) na estrutura e como detectar tal(is) dano(s), são as duas questões mais pertinentes.

Saber quais as causas do(s) dano(s) em uma estrutura, de maneira direta, é uma questão complexa. Várias razões podem explicar uma anomalia em uma estrutura. Em particular, o monitoramento estrutural necessita levar em conta diferentes etapas da vida útil da estrutura, tais como estudos já realizados, procedimentos passados de manutenção e reabilitação, etc. No entanto, é possível notar que mesmo estruturas novas podem apresentar erros de execução ou sofrer algum evento inesperado que possa causar uma falha localizada ou mesmo em escala global da estrutura. As aparições de danos estruturais podem estar relacionadas às modificações da capacidade de resistência das estruturas devido à degradação de seus materiais constitutivos (por exemplo, corrosão, fissuras, etc.) ou então, às diferentes solicitações às quais a estrutura pode ser submetida. De fato, estruturas são geralmente submetidas a sucessivos carregamentos que podem causar danos estruturais ao longo do tempo (fadiga na estrutura metálica de uma ponte devido ao tráfego de veículos, por exemplo) ou também de maneira descontínua (por exemplo, a ocorrência de um terremoto). Uma estrutura também pode apresentar anomalias devido à ações acidentais identificadas ou não identificadas.

Danos estruturais podem ser definidos como modificações estruturais que afetam o seu desempenho, tais como segurança, aptidão para serviços, etc. (EN1990, 1990). O

conceito de dano geralmente só é importante se dois estados estruturais forem identificados e comparados. Normalmente, considera-se o primeiro estado como sendo o de referência, sendo esse estado na maioria das vezes não-danificado. Para sistemas estruturais, os efeitos dos danos são essencialmente causados por modificações dos materiais e em suas propriedades geométricas. Além disso, mudanças nas condições de apoio e nas ligações podem alterar o desempenho da estrutura. Um processo de fissuração, por exemplo, pode causar alterações na geometria da estrutura, podendo levá-la ao colapso nos casos mais graves. Ademais, dependendo da dimensão e localização da fissura, e dos carregamentos aplicados à estrutura, os efeitos do dano sobre o desempenho da estrutura podem aparecer imediatamente ou após algum tempo. Em geral, dois conceitos distintos podem ser apresentados para descrever um dano estrutural. Para o conceito dimensional, todos os danos começam no nível do elemento. De acordo com o carregamento aplicado, esse dano pode progredir até a falha completa da estrutura. Para o conceito temporal, o dano pode se acumular gradualmente por longos períodos, como os efeitos devido à fadiga ou à corrosão, por exemplo. A figura 2.1 apresenta a fissuração de um dos pilares da ponte Honoré-Mercier em Montreal, no Canadá.



FIGURA 2.1 – Ponte Honoré-Mercier em Montreal e detalhes das fissuras em um de seus pilares. FONTE – CURY, 2010.

A resposta para “como detectar o(s) dano(s) estruturais” está geralmente relacionada à avaliação do desempenho das estruturas. Os procedimentos de inspeção visual servem de base, na maioria dos casos de detecção de danos, para definição da manutenção da estrutura em certo período. No entanto, uma simples inspeção visual pode não ser suficiente para detectar o dano, muito menos quantificá-lo. Além disso, o especialista responsável pela inspeção pode não ter acesso a determinadas partes da estrutura, ou ainda o dano pode estar localizado em uma região que não se encontra visível, o que limita o uso desse tipo de procedimento. Abordagens alternativas para substituir a inspeção visual foram uma forte motivação para o desenvolvimento de métodos para monitoramento da integridade estrutural. Em geral, duas técnicas de inspeção são utilizadas: técnicas não-destrutivas e técnicas destrutivas. Técnicas não-destrutivas são um conjunto de métodos utilizados para caracterizar o estado de integridade de uma estrutura sem danificá-la. Os exemplos mais conhecidos são os ensaios de emissão acústica, ultrassom, termografia, testes estáticos e dinâmicos, análise de espectros, etc.

Uma das técnicas de avaliação não-destrutiva mais apropriada para o monitoramento da integridade estrutural é aquela baseada na variação de suas características vibracionais. Os métodos vibracionais consideram a resposta dinâmica da estrutura como função de suas propriedades mecânicas. O princípio é que o dano altera a resposta dinâmica da estrutura pelo fato de sua presença alterar a rigidez, massa e as propriedades de dissipação de energia do sistema. Em outros termos, um dano geralmente ocasiona mudanças nos parâmetros modais da estrutura (frequências próprias, taxa de amortecimento, modos próprios). Mudanças nesses parâmetros são indicadores importantes da existência de danos na estrutura. Ademais, uma vantagem dessa abordagem é a possibilidade de se utilizar vibrações ambientes como fonte de excitação, o que a torna interessante em aplicações como pontes e viadutos pelo fato de o próprio tráfego ser a fonte de excitação.

## **2.1. Estado da arte dos métodos de detecção de danos**

A detecção de danos estruturais é baseada na identificação, localização e quantificação da perda de rigidez do sistema a partir da resposta dinâmica da estrutura. Os estudos de

indicadores de danos tornaram-se populares devido ao desenvolvimento de técnicas de avaliação dinâmica e de identificação modal (EWINS, 2000). Em geral, os efeitos de um dano estrutural podem ser classificados como lineares e não-lineares. O primeiro caso é uma situação em que a estrutura se encontra inicialmente no domínio elástico linear e permanece neste mesmo domínio após a ocorrência do dano. Este é o caso de problemas estruturais no estágio inicial. No segundo caso, o comportamento da estrutura torna-se não-linear após a ocorrência de dano. É possível notar que a não-linearidade ocorre para danos severos, como por exemplo, na formação de fissuras por fadiga devido aos ciclos de cargas em um ambiente de vibração, à plastificação de certos elementos devido a um choque, entre outros.

Os métodos de detecção de danos lineares são os mais utilizados. Eles são classificados em dois grupos: métodos baseados em modelos (MBM) e métodos não baseados em modelos (MNBM). Os métodos do primeiro grupo utilizam modelos numéricos discretizados pelo método dos elementos finitos (ZIENKIEWICZ; TAYLOR, 2000) para desenvolver os algoritmos de detecção de danos. Essas técnicas são, na maioria das vezes, baseadas na atualização dos modelos (matrizes de rigidez e massa) para descrever uma estrutura real. Como exemplo, é possível citar os algoritmos de atualização otimizada das matrizes (SMITH; BEATTIE, 1991), as técnicas de ajuste dos parâmetros modais (LIM; KASHANGAKI, 1994), os métodos de sensibilidade (ALVIN, 1997), as técnicas baseadas nas funções de danos (ABDEL WAHAB et al., 1999; TEUGHELIS et al., 2002), etc.

Os MNBM são essencialmente baseados nas mudanças dos parâmetros modais e suas derivadas, ou mesmo nas modificações de suas matrizes, como as de flexibilidade e de rigidez. Uma das vantagens dessas técnicas é que elas não necessitam de conhecer, a princípio, a localização do dano.

Os métodos de detecção de danos são utilizados para responder a quatro níveis de conhecimento (RYTTER, 1993):

- Nível 1: existência de dano estrutural;
- Nível 2: localização do dano;
- Nível 3: quantificação do dano;

- Nível 4: vida residual da estrutura.

Vários estudos utilizando os MNBM existentes na literatura são baseados na alteração das frequências próprias, das taxas de amortecimento e dos modos próprios. Além disso, indicadores evoluídos têm sido propostos por vários autores nos últimos anos (DOEBLING et al., 1996). O objetivo desses indicadores é melhorar e refinar a identificação de danos em relação aos métodos que levam em consideração apenas a variação dos parâmetros modais. Um vasto estudo dessas técnicas foi realizado por Alvandi (2004).

A seguir, serão apresentados de maneira sucinta os indicadores já existentes, dentre os quais podemos citar as técnicas que utilizam os parâmetros modais para a detecção de danos, alguns indicadores evoluídos e, por último, novos indicadores que são baseados nas técnicas de anormalidade.

### **2.1.1. Métodos baseados nas variações dos parâmetros modais**

Os primeiros estudos experimentais para o desenvolvimento de indicadores de danos, foram realizados no século XIX, quando os equipamentos eram pouco robustos e praticamente só as frequências naturais podiam ser identificadas com um nível adequado de precisão. Com isso, muitos estudos foram realizados utilizando as variações de tais parâmetros como indicadores de danos. Essas abordagens assumem que as frequências são indicadores sensíveis da integridade estrutural. De fato, mudanças nas propriedades estruturais causam alterações nas frequências da estrutura. Nesse contexto, o monitoramento desses parâmetros pode servir de base para um método de detecção de dano estrutural. Um dos primeiros a utilizar o método que relaciona a variação das frequências próprias a danos estruturais foram Cawley e Adams (1979), que aplicaram o método para localizar danos em uma viga bi-apoiada. Nesse estudo, observou-se que o dano ocasiona uma redução local na matriz de rigidez da estrutura. Embora a utilização da mudança da frequência possibilite a detecção do dano, sua quantificação não é possível. Melhorias foram propostas por Friswell (1994) e por Williams (1996). Messina (1998) propôs um método chamado MDLAC (*Multiple*



*Damage Location Assurance Criterion*), capaz de quantificar e detectar múltiplos danos.

Em todos os estudos, as frequências próprias decrescem com o aumento do nível de dano. Isto se deve ao fato de que o dano tende a reduzir a rigidez da estrutura. Vários estudos lidam com a detecção de danos através da análise da variação das frequências. No entanto, muitos desses estudos mostram que a análise da mudança desses parâmetros parece ter uso prático muito limitado. Kato e Shimada (1986) realizaram medições dinâmicas utilizando vibrações ambientais em uma ponte de concreto armado. Foi mostrado que o nível de carga estática aplicada deve estar próximo da carga máxima para detectar uma redução das frequências. Gudmundson (1982) também revelou que a diminuição dos valores das frequências próprias é menor em casos de fadiga do que em casos de uma fissura causada por uma carga permanente. Isso implica que a diminuição da frequência é afetada por fatores como cargas permanentes ou tensões residuais. De fato, as cargas permanentes podem manter as fissuras abertas e acentuar a variação das frequências. Utilizando os resultados numéricos e experimentais de uma viga, Fox (1992) mostrou que a variação das frequências próprias é um indicador pouco sensível no caso de uma viga com fissuras.

A variação dos parâmetros modais pode não ser idêntica para cada modo, pois ela depende de efeitos naturais, da localização e da severidade dos danos (DUGGAN et al. 1980). O uso da variação da frequência é mais viável quando os ensaios são realizados em um ambiente controlado, como os testes realizados em laboratório, por exemplo. Na maioria dos casos, é difícil estabelecer uma correlação entre uma variação na frequência e a ocorrência de danos. Nenhuma informação espacial é fornecida pelas frequências próprias. A única exceção é o caso de frequências elevadas, que geralmente estão associadas às respostas locais. No entanto, as limitações na excitação e na extração desses modos, tornam-lhes muito difíceis de serem identificados.

No caso das taxas de amortecimento, poucos estudos têm sido realizados nos últimos anos. Salane e Baldwin (1990) estudaram a influência dos danos utilizando as taxas de amortecimento e mostraram que esses coeficientes podem levar a indicadores não confiáveis, pois tendem a sofrer grandes oscilações, impedindo uma correlação entre os

efeitos dos danos com sua variação, por ser um parâmetro instável. No entanto, estudos realizados por Ndambi et al. (2000) e Kawiecki (2001) tendem a confirmar a utilidade do uso da taxa de amortecimento como indicador da integridade da estrutura. Contudo, a utilização da taxa de amortecimento como indicador de dano tem sido objeto de estudo de poucos trabalhos até o momento.

Em compensação, muitas pesquisas têm sido realizadas utilizando indicadores baseados na variação dos modos próprios. Esses métodos utilizam uma comparação direta entre os modos antes e depois da ocorrência do dano. Allemang e Brown (1982) propuseram um índice chamado MAC (*Modal Assurance Criterion*) que consiste em avaliar a correlação entre duas variáveis. Este índice varia entre 0 e 1, em que 0 indica a ausência de correlação entre as variáveis comparadas e 1, uma correlação perfeita. A avaliação do MAC para os modos danificados e não-danificados fornece um meio de quantificar a variação da correlação entre estes parâmetros. West (1984) foi o primeiro a utilizar os modos próprios para localizar danos sem recorrer a um modelo que utilize elementos finitos. Ele ainda utilizou o coeficiente MAC para determinar a correlação entre os modos próprios antes e depois da ocorrência de danos em um ônibus espacial.

Alguns anos mais tarde, Lieven e Ewins (1988) propuseram o índice COMAC (*COrdinate Modal Assurance Criterion*). Este índice é uma variação do MAC, porém a correlação é calculada para cada nó da estrutura, através dos modos. Ele fornece uma correlação entre os estados específicos medidos para cada grau de liberdade. Esse coeficiente é geralmente utilizado para identificar o local onde os estados de uma série de medidas não estão correlacionados uns com os outros. No caso de uma correlação perfeita entre os valores de uma coordenada, o coeficiente COMAC é igual a 1. Um grande desvio pode ser, portanto, interpretado como um dano estrutural.

### **2.1.2. Métodos baseados na variação de indicadores evoluídos**

Estudos sobre a utilização direta dos parâmetros modais têm colocado em evidência a baixa sensibilidade das frequências próprias e apresentado certo potencial na utilização dos modos próprios para detectar e/ou localizar danos. Para aumentar a capacidade dos

indicadores baseados nos modos próprios para detectar danos, diversas abordagens foram propostas (FARRAR; JAUREGUI, 1996).

Uma maneira de se utilizar tais parâmetros modais para obter informações sobre falhas estruturais é estudar suas curvaturas, que nada mais são do que as derivadas das formas modais. Há, de fato, uma relação entre a curvatura dos modos próprios e as deformações por flexão. Pandey et al. (1991) mostraram que a variação da curvatura dos modos próprios pode ser um bom indicador de dano em um modelo de elementos finitos. Utilizando o método das diferenças finitas, os valores da curvatura podem ser calculados a partir dos modos próprios. O valor absoluto da diferença da curvatura dos modos próprios sem dano e com dano da região danificada deve ser máximo. Quanto maior a redução da rigidez, correspondente a um nível elevado de dano, maior será a variação da curvatura. No entanto, na maioria dos ensaios, há geralmente muita dificuldade em se medir os graus de liberdade de rotação ou até mesmo diversos graus de liberdade de translação devido a limitações dos equipamentos, do número de sensores utilizados, etc. Além disso, Chance et al. (1994) mostraram que o cálculo da curvatura obtida a partir dos modos próprios pode levar a níveis de erros elevados.

Outra classe de métodos de detecção de danos para estimar mudanças no comportamento de uma estrutura é baseada na utilização da matriz de flexibilidade (PANDEY; BISWAS, 1994). Aktan et al. (1994) propuseram a utilização da flexibilidade como um índice de integridade de uma ponte. O método não exige um modelo analítico da estrutura, já que apenas algumas frequências e modos próprios antes e depois do dano são necessários. A formulação da matriz de flexibilidade por este método é aproximada, pois apenas um pequeno número de modos, geralmente os que estão associados às baixas frequências, são medidos. Como a presença de dano diminui a rigidez da estrutura, a sua flexibilidade aumenta, uma vez que a flexibilidade é o inverso da rigidez.

Através da combinação de certos aspectos do método da curvatura dos modos próprios e da flexibilidade, Zhang e Aktan (1995) desenvolveram outro método de detecção de danos. Da mesma forma que no método da curvatura, a ideia geral é que a perda da

rigidez localizada causa um aumento na curvatura no mesmo ponto. A mudança na curvatura é então obtida pela flexibilidade, ao invés dos modos próprios.

Kim e Stubbs (1993) deram um passo importante para a detecção e localização de danos com base na variação da energia de deformação à flexão das estruturas. Uma das vantagens desse método é que com poucos modos de vibração é possível obter bons resultados. Essa variação, expressa por meio de uma proporção, permite definir um indicador que indica ausência de danos quando seu valor é negativo e a presença de danos, quando positivo. Esse método foi desenvolvido para pequenos níveis de danos, pois ele depende de uma linearização das energias de deformação após o dano. Recentemente, Cury et al. (2011) propuseram uma abordagem híbrida para inicialmente localizar o dano e, depois, quantificá-lo de maneira independente.

### **2.1.3. Métodos baseados em indicadores de anormalidades**

Os métodos apresentados nas seções anteriores baseiam-se em uma análise essencialmente no domínio do tempo para extrair informações através de algumas características modais ou de indicadores evoluídos baseados nessas características. Esses indicadores mostraram-se eficazes, porém demonstraram alguns problemas como a falta de sensibilidade para determinados níveis de dano, a necessidade de se dispor de um estado de referência e a confiabilidade em termos da probabilidade de detecção de alarmes falsos. Um aspecto importante a ser considerado é que, para esses métodos, é assumida a hipótese de que o comportamento mecânico é linear após os danos, o que nem sempre acontece. Além disso, o processo de identificação modal é um procedimento de filtragem, que pode conduzir a uma perda de informações essenciais sobre o estado da estrutura. Paralelamente, pouco trabalho tem se concentrado na análise direta da resposta dinâmica de uma estrutura através de suas características sob excitações externas. De fato, os sinais de medições dinâmicas são raramente utilizados devido à falta de ferramentas apropriadas para sua manipulação e à dificuldade de se utilizá-los diretamente.

A fim de fornecer algumas respostas para essas questões, novas abordagens baseadas em análises estatísticas e de reconhecimento de padrões foram introduzidas, e técnicas

originais para detecção de danos foram propostas. Uma dessas técnicas é baseada no conceito de fusão de dados. Essa manipulação de dados foi desenvolvida pelas Forças Armadas dos Estados Unidos na década de 50. O objetivo era determinar as situações nos campos de batalha e avaliar as ameaças a partir de informações geradas por várias fontes diferentes (KLEIN, 1999). Esse conceito se espalhou rapidamente para outras áreas, como economia, biologia, geologia, etc. A fusão de dados consiste essencialmente em integrar e confrontar várias informações com o objetivo de reduzir as incertezas da informação resultante. Na engenharia civil, mais especificamente na área de detecção de falhas estruturais, poucos estudos têm sido realizados a respeito da gestão de grande número de dados coletados. Guo (2006) e Minor et al. (2007) propuseram diferentes metodologias para reunir informações obtidas por vários acelerômetros. Eles relataram que a robustez dessa abordagem para identificação de danos estruturais depende diretamente do tamanho da base de dados a ser tratada. Chun et al. (2005) também destacaram o problema da gestão de muitos dados adquiridos em uma campanha de ensaios de medições. Su et al. (2009) propuseram diversos esquemas de compactação de dados em que determinados sensores são escolhidos previamente e apenas suas medidas são exploradas. No entanto, os resultados obtidos mostraram que a perda de informações pode conduzir a resultados falsos na detecção de danos. Abordagens baseadas em métodos de subespaços também foram propostas para a detecção de comportamentos estruturais anormais (BASSEVILLE et al., 2006). O princípio é estimar os resíduos a partir de comparações entre a resposta dinâmica da estrutura intacta e da estrutura com dano. De maneira geral, a magnitude do resíduo dá uma ideia da localização e da quantificação do dano (MEVEL et al. 1999; BODEUX; GOLINVAL, 2003).

Uma abordagem aplicada ao monitoramento de danos estruturais que recentemente suscitou grande interesse é o reconhecimento de formas (*pattern recognition*), também conhecido como reconhecimento de padrões. O reconhecimento de padrões, aplicado à engenharia civil, consiste em um conjunto de métodos que permitem identificar as características intrínsecas das estruturas a partir de dados brutos. Essas características permitem identificar e discriminar comportamentos estruturais distintos ou diferentes estados de dano da estrutura. Existe uma variedade de métodos baseados no

reconhecimento de padrões utilizados para a detecção de danos descritos na literatura. Esses métodos são muitas vezes separados em duas famílias: os métodos ditos não-supervisionados (tais como o método de *clustering* ou particionamento de dados, testes de hipóteses, etc.) e os métodos ditos supervisionados (redes neurais, máquinas de vetor suporte, etc.). Os métodos não-supervisionados, alvo de interesse deste trabalho, consistem em agrupar um conjunto de informações desconhecidas em diferentes grupos. O objetivo é identificar os grupos de tal forma que as observações mais similares se encontrem num mesmo grupo e, as mais diferentes, sejam separadas em grupos distintos. Para o caso dos métodos supervisionados, uma série de rótulos previamente conhecidos é fornecida como dado de entrada para a determinação dos diferentes grupos. O problema consiste em dar um rótulo a uma nova observação, uma vez que os grupos já estão definidos. Alguns autores têm proposto abordagens baseadas nesses dois tipos de métodos. Trendafilova e Heylen (2003) e Iwasaki et al. (2004) utilizaram algoritmos de classificação não-supervisionada para detectar danos em uma viga engastada. Para isso, eles utilizaram diferentes respostas no domínio das frequências obtidas a partir de ensaios controlados. Os resultados obtidos mostraram que a abordagem considerada permite classificar corretamente os diferentes estados estruturais da viga. Haritos e Owen (2004) compararam várias técnicas estatísticas tais como a análise em componentes principais e utilização de testes de hipóteses em ensaios dinâmicos experimentais realizados em uma viga metálica. Os níveis de dano considerados foram cortes feitos em uma dada região da estrutura. Eles concluíram que os indicadores não foram suficientemente sensíveis para detectar pequenos cortes. Reda Taha e Lucero (2005) utilizaram uma abordagem baseada na lógica nebulosa (lógica *fuzzy*) para detecção de danos em um modelo numérico de uma ponte de concreto. A técnica mostrou-se eficaz para distinguir um estado intacto de um estado muito danificado. No caso de um estado de degradação moderado, a abordagem proposta não foi capaz de distingui-lo de um estado sem dano. Zhang (2007) usou índices estatísticos baseados na resposta no domínio da frequência de um modelo numérico de uma ponte metálica. Ele mostrou que, para baixos níveis de danos, esses índices podem não revelar a presença de problemas estruturais.

As respostas dinâmicas (sinais) da estrutura são raramente utilizadas. Vários autores relataram que as medidas dinâmicas são pouco úteis para a determinação dos comportamentos estruturais devido à presença de ruído ou a dificuldade de manipulá-las corretamente (KIM; LEE, 2000; YAN et al. 2007). Yeung e Smith (2005) propuseram um algoritmo de aprendizagem baseado nas características de um modelo numérico de uma ponte real, simulando diferentes níveis de dano e ruído. Observou-se que, para baixos níveis de ruído, o algoritmo pôde identificar corretamente as situações de dano impostas. No entanto, quando o nível de ruído aumenta, o desempenho da classificação deixa de ser satisfatório.

# CAPÍTULO 3

## 3. ANÁLISE DE DADOS SIMBÓLICOS

*Data Mining* (ou mineração de dados) utiliza ferramentas que possibilitam descobrir modelos e relações previamente desconhecidas ou escondidas em grandes conjuntos de dados. Essas ferramentas podem incluir métodos estatísticos, algoritmos matemáticos, métodos de aprendizagem, entre outros. O *data mining* não se limita em agrupar e estocar os dados. Ele também realiza uma análise crítica desses dados em um momento presente e traça previsões de tendências futuras. O *data mining* pode ser aplicado em dados representados de forma quantitativa, textual ou multimídia como, por exemplo, imagens e sons. As técnicas de extração de informações podem necessitar de vários parâmetros para caracterizar os dados. Elas incluem a associação (modelos em que um evento está relacionado ao outro), a ordem ou a análise de consequência (modelos em que um evento leva a outro), a classificação (identificação de novos modelos), o agrupamento (descoberta e visualização de grupos de dados) e previsão (descoberta de modelos a partir dos quais previsões futuras podem ser propostas). Uma grande contribuição do *data mining* está relacionada à utilização das abordagens chamadas de “descobertas”, no qual os algoritmos podem ser empregados para examinar simultaneamente várias relações multidimensionais de dados (HASTIE et al. 2009).

O *data mining* vem sendo cada vez mais utilizado atualmente tanto em setores públicos como em setores privados. Ele é muitas vezes utilizado como ferramenta que permite a gestão de informações de clientes e a redução da probabilidade de fraude, por exemplo. Apesar de os algoritmos de mineração de dados poderem ser utilizados como ferramentas poderosas, eles não são auto-suficientes. Nesse caso, a análise crítica do usuário faz-se sempre imprescindível.

As técnicas estatísticas utilizadas para a análise de dados têm sido empregadas em aplicações relativamente simples. A unidade de análise é geralmente um indivíduo, como por exemplo, uma pessoa, um objeto, um lugar, etc. Esse indivíduo é descrito por um tipo bem definido de variável (qualitativo ou quantitativo), associado a um valor



simples. Quando a unidade de estudo deixa de ser um indivíduo e passa ser um grupo de indivíduos, sua descrição torna-se um pouco mais complicada. Nesse caso, uma nova representação de dados é concebida. Esses dados contêm informações mais gerais, mas também mais complexas e são chamados de *dados simbólicos*. Para melhor compreender e extrair estas novas informações, a Análise de Dados Simbólicos (ADS) foi criada.

### **3.1. Transformação de dados clássicos em dados simbólicos**

Os recentes avanços técnicos e tecnológicos tornaram possível a aquisição contínua de diversas informações e sua estocagem em grandes bancos de dados. Porém, apesar da capacidade de processamento atual dos computadores, os recursos computacionais necessários para manipular e analisar grandes conjuntos de dados são geralmente limitados. Além disso, os métodos tradicionais de análise de dados foram desenvolvidos em uma época em que a quantidade de informações disponíveis era mais limitada que atualmente.

Ao contrário da análise clássica, que se concentra no estudo dos indivíduos, a análise simbólica manipula os conceitos, que representam uma forma mais rica e geral de dados. Os conceitos, considerados como novas unidades estatísticas são descritos por dados mais complexos do que aqueles habitualmente encontrados nas análises estatísticas. Eles são chamados de simbólicos por exprimir a variação interna inerente a cada variável, levando em conta as especificidades da base de dados, e não podem ser manipulados unicamente pela simples descrição numérica ou qualitativa. Com isso, eles necessitam de uma caracterização mais rica, mas também mais compacta. Nesse contexto, a extensão dos métodos clássicos a esses dados representam um grande desafio. Na tabela 3.1 é apresentado um exemplo simples de uma comparação das unidades de estudo de uma análise clássica e de uma análise simbólica.

TABELA 3.1 - Exemplos de unidades de estudo para uma análise clássica e uma análise simbólica.

<b>Dados clássicos:</b> <i>Indivíduos</i>	<b>Dados simbólicos:</b> <i>Conceitos</i>
Jogadores	Equipes
Orquídeas, rosas	Flores
Pombos, patos	Aves
<b>Pontos de medição</b>	<b>Ensaio dinâmico</b>

A primeira etapa da ADS é reunir e descrever os conceitos. Isso pode ser realizado para a utilização de variáveis que permitam não somente representar bem, mas também caracterizar os conceitos. A segunda etapa consiste em analisar a tabela de dados simbólicos construída na primeira etapa. As tabelas de dados simbólicos podem ter objetos de diferentes naturezas (por exemplo, um conjunto de vários valores), podem ser ponderados e relacionados entre eles por regras e/ou taxonomias. Exemplos dessas representações são listados abaixo (BILLARD; DIDAY, 2006):

- Apenas um valor quantitativo. Por exemplo, se a altura é a variável e  $z$  é um indivíduo: altura ( $z$ ) = 3,5 m.
- Apenas um valor categórico: cidade ( $z$ ) = Paris.
- Dados categóricos com vários valores. Preço ( $z$ ) = {alto, médio, baixo} indicando que o preço de um produto  $z$  pode ser alto, médio ou baixo.
- Dados do tipo intervalo: peso ( $z$ ) = [20; 180] significa que o peso de  $z$  varia no intervalo de 20 até 180.
- Dados do tipo modal com vários valores (histograma): altura ( $z$ ) = {[0; 1,20] (0,225); [1,20; 1,50] (0,321); [1,50; 1,80] (0,335); [1,80; 2,10] (0,119)} indica que 22,5% da população  $z$  tem a altura entre 0 e 1,20 m, 32,1% entre 1,20 m e 1,50 m, 33,5% entre 1,50 m e 1,80 m, e 11,9% entre 1,80 m e 2,10 m.

Para ilustrar a diferença entre uma tabela de dados clássicos e uma tabela de dados simbólicos, o exemplo de um grupo de pássaros (indivíduos) contendo 300 andorinhas, 100 pinguins e 100 avestruzes é considerado. Na tabela 3.2, pode-se observar uma tabela clássica em que cada linha descreve um pássaro (indivíduo) a partir de três

variáveis: espécie, capacidade de voar e peso (DIDAY; NOIRHOMME – FRAITURE, 2007).

TABELA 3.2 – Descrição clássica de 500 pássaros.

Pássaro	Espécie	Voa	Peso (Kg)
1	Pinguim	Não	35
2	Andorinhas	Sim	0,05
...	...	...	...
500	Avestruz	Não	100

Uma análise mais geral pode ser realizada considerando as espécies de pássaros como unidades de estudo, sendo então os conceitos. Na tabela 3.3, as espécies são classificadas de acordo com as características descritas anteriormente, ou seja, capacidade de voar e peso.

TABELA 3.3 – Descrição simbólica de 500 pássaros.

Espécie	Número de indivíduos	Voa	Peso (Kg)
Andorinha	300	Sim	[0,01; 0,08]
Pinguim	100	Não	[30; 100]
Avestruz	100	Não	[80; 160]

Analisando-se este exemplo, é possível obter duas respostas distintas para uma mesma questão: Quantos pássaros voam? Levando-se em consideração a descrição clássica, em que as unidades de estudo são os indivíduos, a resposta é  $300/500 = 60\%$ , já que apenas as 300 andorinhas do total de 500 pássaros voam. Entretanto, considerando-se a descrição simbólica, em que os conceitos são utilizados como unidades de estudo, a resposta é  $1/3=33,33\%$ , ou seja, apenas uma, das três espécies, voa.

Assim, é importante observar na tabela 3.3, que os valores da variável “peso” não são escalares, mas sim intervalos. Esse tipo de notação permite a construção de tabelas compactando as informações de forma mais geral que as tabelas clássicas, além da análise da dispersão das variáveis. A ADS permite organizar ou mesmo comprimir os dados originais de maneira controlada para possibilitar manipulá-los mais

adequadamente. O desafio é, portanto, saber qual o nível de compressão pode ser realizado sem distorcer a informação original. De fato, cada aplicação resultará na descrição e transformação de dados específicos e caberá ao usuário julgar qual aplicação é a mais adequada para determinado caso.

Um exemplo de um estudo do monitoramento dinâmico de uma estrutura, em que um ensaio hipotético *E* contendo 10000 valores medidos por 10 acelerômetros é agora considerado. Esse ensaio está ilustrado na figura 3.1.

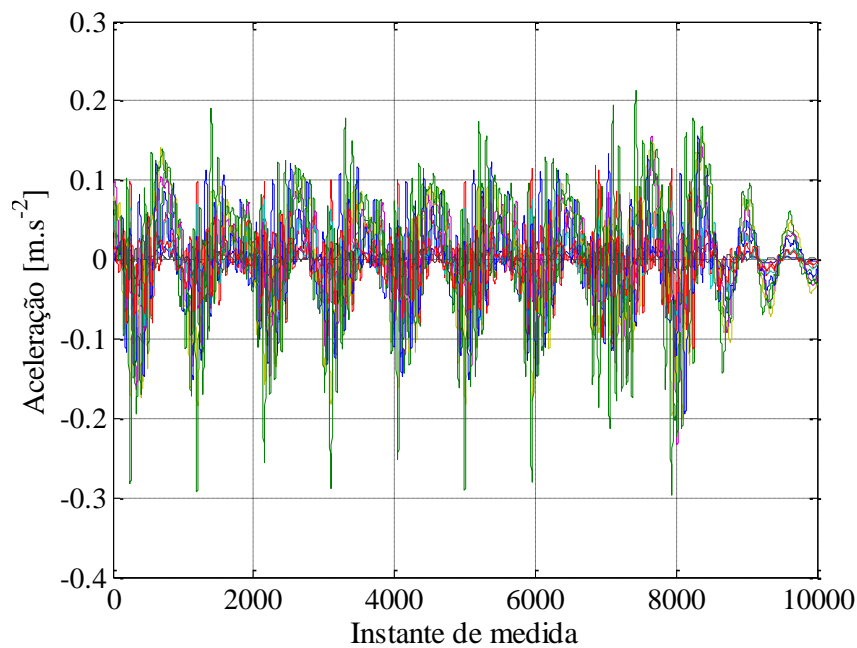


FIGURA 3.1 – Exemplo de um ensaio dinâmico. FONTE – CURY, 2010.

Este ensaio pode ser, evidentemente, representado de maneira simbólica. Nesse exemplo, os valores medidos são os indivíduos e os sensores são as variáveis que descrevem os ensaios dinâmicos, que são os conceitos. O sinal medido pelo acelerômetro sete é considerado para realização da transformação da representação clássica para representação simbólica. Nessa transformação, ilustrada na figura 3.2, cada valor medido é projetado no eixo das ordenadas. Um histograma é então construído, levando-se em conta a proporção dos valores incluídos entre os limites de cada divisão. Essas divisões são chamadas de *categorias*. O número de categorias é definido previamente para cada tipo de estudo realizado e determina o grau de refinamento para a descrição de cada conceito.

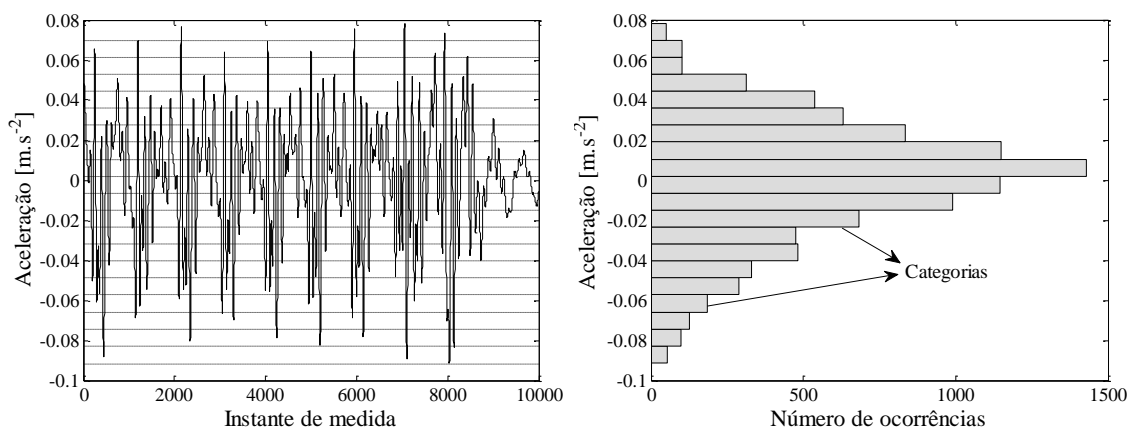


FIGURA 3.2 – Representação simbólica de um sinal clássico para um histograma de 20 categorias. FONTE – CURY, 2010.

A tabela 3.4 ilustra uma descrição em dados simbólicos para o ensaio *E* a partir de histogramas de 20 categorias. Nesta tabela, os dados entre parêntesis representam as proporções dos valores de cada sinal incluídos entre os limites de cada categoria.

TABELA 3.4 – Descrições simbólicas de um ensaio *E* representado pelos acelerômetros.

Acelerômetro	Histograma
<b>1</b>	[1 (0,002), ..., 10 (0,31), ..., 20 (0,001)]
<b>2</b>	[1 (0,001), ..., 10 (0,28), ..., 20 (0,005)]
...	...
<b>10</b>	[1 (0,003), ..., 10 (0,25), ..., 20 (0,004)]

Há outras maneiras de se representar ensaios dinâmicos simbolicamente. Os ensaios podem também ser descritos pelos parâmetros modais identificados, como as frequências próprias e os modos próprios. Nesse caso, os ensaios dinâmicos continuam sendo os conceitos, mas as variáveis agora são representadas pelas frequências próprias ou pelos modos próprios e não mais pelos sensores.

Para ilustrar tal situação, um exemplo de identificação modal de uma viga metálica bi-apoiada é considerado. As duas primeiras frequências próprias e modos próprios desta viga são calculados. Para alguns casos, é insuficiente encontrar um valor numérico adequado para os parâmetros modais, mas deve-se estimar também sua precisão e incertezas relacionadas ao processo de identificação modal. Em particular, quando o

número de ensaios é insuficiente, pontos de grande variação e erros podem aparecer. Para tratar esse problema, é possível gerar várias estimativas e produzir estatísticas a partir de amostras extraídas das respostas dinâmicas. Essas estimativas são chamadas de *realizações*. Na figura 3.3 são ilustradas as frequências e modos próprios de cinco realizações avaliadas para as duas primeiras frequências naturais e os dois primeiros modos próprios identificados. Esse procedimento produz uma distribuição de valores para cada parâmetro modal. A partir dessa distribuição, as frequências próprias e modos próprios são transformados em dados simbólicos, como detalhado nas tabelas 3.5 e 3.6, respectivamente.

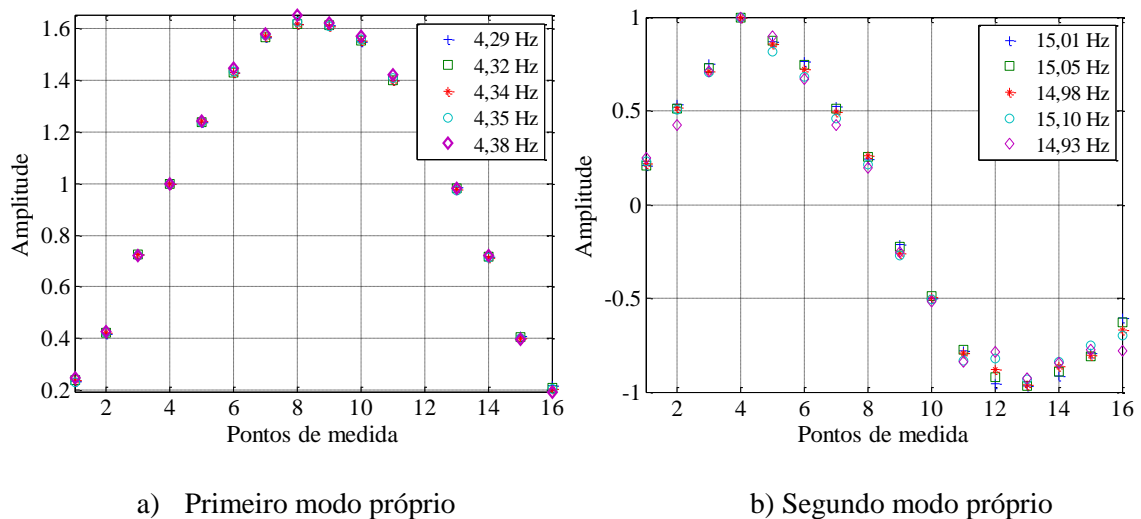


FIGURA 3.3 – Frequências e modos próprios identificados para 5 realizações de uma identificação modal. FONTE – CURY, 2010.

TABELA 3.5 – Descrições simbólicas de um ensaio *E* representado pelas frequências próprias.

Frequência	Histograma
1	[1 (0,01), ..., 10 (0,04), ..., 20 (0,001)]
2	[1 (0,02), ..., 10 (0,58), ..., 20 (0,005)]

TABELA 3.6 – Descrições simbólicas de um ensaio  $E$  representado pelos modos próprios.

<b>Modo</b>	<b>Histograma</b>
1	[1 (0), ..., 10 (0,40), ..., 20 (0)]
2	[1 (0,2), ..., 10 (0,12), ..., 20 (0,04)]

Através destes exemplos, é possível observar que várias representações simbólicas são possíveis. Nota-se que cada aplicação resulta em uma utilização de um tipo de dado simbólico apropriado. É importante sempre haver uma relação fidedigna entre os dados simbólicos e os dados clássicos. Ou seja, é necessário saber compactar os dados clássicos sem perder as propriedades originais, tais como os valores médios, a variância, a mediana, etc. Na medida em que o número de categorias de um histograma aumenta, as propriedades dos dados simbólicos tendem a se aproximar dos dados clássicos. Isso pode, entretanto, levar a criação de categorias vazias, que não contribuem em nada nos cálculos realizados posteriormente. A escolha da representação simbólica depende da aplicação estudada e cabe ao usuário a escolha da representação simbólica mais adequada.

Neste trabalho, a representação simbólica utilizada é a de histogramas com 12 categorias por apresentar melhor adequação aos estudos realizados.

# CAPÍTULO 4

## 4. MÉTODOS DE CLASSIFICAÇÃO DE DADOS

Neste capítulo são apresentados os diferentes métodos de classificação de dados, baseados no aprendizado não-supervisionado. Os métodos não-supervisionados consistem em agrupar um conjunto de informações desconhecidas (não-etiquetadas) em grupos distintos (*clusters*). Essas etiquetas são associadas aos *clusters* indiretamente a partir de cálculos realizados por diferentes algoritmos. Dessa forma, o objetivo é identificar os grupos de tal forma que as observações mais similares se encontrem em um mesmo grupo e as mais diferentes sejam separadas em grupos distintos.

O agrupamento de dados, também chamado de particionamento de dados ou *clustering*, é uma ferramenta utilizada durante muitos anos em várias áreas, tais como na economia, estudos estatísticos, medicina, etc. (BOCK; DIDAY, 2000). O objetivo, de maneira geral, é separar um conjunto de objetos em diferentes grupos em função de uma noção de similaridade, ou dissimilaridade. Os objetos considerados “similares” são agrupados em um mesmo *cluster* enquanto que os considerados “diferentes” são separados em *clusters* distintos. Assim, o objetivo do *clustering* é agrupar um conjunto de dados da forma mais natural possível.

Um dos pontos fortes dos métodos de *clustering* está relacionado à capacidade de organizar grandes conjuntos de dados para ajudar a simplificar as informações e generalizar a análise de diferentes componentes. Essas informações simplificadas podem também ser empregadas para realizar previsões e efetuar suposições com base nos *clusters* formados. Existem diferentes métodos de *clustering* na literatura e esses métodos podem ser classificados como métodos hierárquicos ou como métodos de partição. No primeiro caso, uma hierarquia de *clusters* é formada para caracterizar a similaridade entre os objetos. Existem dois tipos de métodos hierárquicos: os divisivos ou descendentes e os aglomerativos ou ascendentes. O método descendente começa com um único *cluster* no qual se encontram todos os objetos. Esse *cluster* inicial é então dividido sucessivamente em vários *clusters* de acordo com certo critério até que todos



os objetos estejam classificados em *clusters* diferentes, ou seja, cada *cluster* contenha apenas um objeto. Os métodos aglomerativos ou ascendentes começam com um objeto em cada *cluster*. Os *clusters* vão sendo aglomerados sucessivamente segundo certo critério, até que todos os objetos estejam agrupados em um único *cluster*.

No caso dos métodos de partição, os *clusters* também são definidos por um critério de similaridade, mas a noção de hierarquia não é aplicada nesse caso. Na seção seguinte serão apresentadas as medidas de distâncias para caracterizar as dissimilaridades entre os objetos e, após essas medidas terem sido definidas, três métodos de particionamento de dados serão detalhados. O primeiro deles é baseado no *clustering* hierárquico e os demais no agrupamento por partição.

#### **4.1. Cálculo das medidas de dissimilaridade**

Antes de introduzir os métodos de *clustering*, é necessário definir as medidas de distância para determinar a similaridade/dissimilaridade entre objetos para criar os *clusters*. As dissimilaridades entre os objetos são avaliadas por uma função que estima as distâncias que os separam. Se a distância entre dois objetos é pequena, os objetos são considerados como próximos, ou seja, similares (ESPOSITO et al. 2000). Inversamente, se a distância é grande, os objetos estão distantes, logo são classificados como diferentes. Uma vez que a distância é definida, a tarefa do método de *clustering* consiste em minimizar a soma das distâncias entre os objetos de um mesmo *cluster*, definido como inércia intra-cluster, e maximizar as distâncias entre os objetos de *clusters* diferentes, definido como inércia inter-cluster. O procedimento de *clustering* deve fornecer um conjunto de *clusters* com grande similaridade intra-clusters e grande dissimilaridade inter-cluster.

A escolha de uma medida de distância constitui uma etapa muito importante do método de *clustering*. Um objeto A pode ser considerado próximo de um objeto B considerando uma medida de distância, mas pode ser considerado distante se outra medida for utilizada. Pode-se encontrar na literatura vários tipos de medidas de distância e as mais comumente utilizadas serão apresentadas a seguir.

#### 4.1.1. Minkowski

A métrica Minkowski é a forma mais geral de se calcular distâncias entre dois objetos. Atributos com valores altos e alta variância tendem a dominar em relação a outros atributos. Para dois ensaios hipotéticos  $E_i$  e  $E_j$ , tem-se:

$$gs(E_i, E_j) = \sqrt[p]{\sum_{k=1}^n |E_{ik} - E_{jk}|^p} \quad (4.1)$$

em que  $p$  é o parâmetro de Minkowski e assume valores interiores.

#### 4.1.2. Cityblock

A distância Cityblock é um caso particular da distância de Minkowski para o caso em que  $p=1$ . Para dois ensaios hipotéticos  $E_i$  e  $E_j$ , tem-se:

$$gs(E_i, E_j) = \sum_{k=1}^n |E_{ik} - E_{jk}| \quad (4.3)$$

#### 4.1.3. Euclidiana

A distância Euclidiana é a métrica mais comumente utilizada. Ela é um caso especial da distância de Minkowski para o caso em que  $p=2$ . Sejam  $E_i$  e  $E_j$  ensaios dinâmicos hipotéticos e  $n$  o número de categorias dos ensaios. Então:

$$gs(E_i, E_j) = \sqrt{\sum_{k=1}^n (E_{ik} - E_{jk})^2} \quad (4.2)$$

#### 4.1.4. Chebychev

A distância Chebychev é um caso particular da distância de Minkowski para o caso em que  $n \rightarrow \infty$ . Para dois ensaios hipotéticos  $E_i$  e  $E_j$ , tem-se:

$$gs(E_i, E_j) = \max_k \{|E_{ik} - E_{jk}|\} \quad (4.4)$$

#### 4.1.5. Cosine

Também conhecida como similaridade do cosseno. Para dois ensaios hipotéticos  $E_i$  e  $E_j$ , tem-se:

$$gs(E_i, E_j) = 1 - \frac{\sum_{k=1}^n E_{ik} \times E_{jk}}{\sqrt{\sum_{k=1}^n (E_{ik})^2} \times \sqrt{\sum_{k=1}^n (E_{jk})^2}} \quad (4.5)$$

#### 4.1.6. Correlation

Também conhecida como correlação de Pearson, não é uma métrica. É utilizada na análise de dados que possuem distribuição normal. Não consegue detectar a magnitude da diferença de dois atributos. Sejam  $E_i$  e  $E_j$  os ensaios,  $\bar{E}_i$  e  $\bar{E}_j$  as médias dos ensaios e  $n$  o número de categorias dos ensaios, então:

$$gs(E_i, E_j) = 1 - \frac{\sum_{k=1}^n (E_{ik} - \bar{E}_i)(E_{jk} - \bar{E}_j)}{\sqrt{\sum_{k=1}^n (E_{ik} - \bar{E}_i)^2} \times \sqrt{\sum_{k=1}^n (E_{jk} - \bar{E}_j)^2}} \quad (4.6)$$

#### 4.1.7. Spearman

Este coeficiente de correlação é baseado no *rank* dos valores e é largamente utilizado como medida de correlação de dados não paramétricos. Sejam  $E_i$  e  $E_j$  os ensaios,  $\bar{E}_i$  e  $\bar{E}_j$  as médias dos ensaios e  $n$  o número de categorias dos ensaios, então:

$$gs(E_i, E_j) = \frac{\sum_{k=1}^n (E_{ik} - \bar{E}_i)(E_{jk} - \bar{E}_j)}{\sqrt{\sum_{k=1}^n (E_{ik} - \bar{E}_i)^2} \times \sqrt{\sum_{k=1}^n (E_{jk} - \bar{E}_j)^2}} \quad (4.7)$$

Todas estas medidas de distâncias encontram-se disponíveis no software Matlab® e foram utilizadas nas aplicações contidas neste trabalho.

## 4.2. Método das Aglomerações Hierárquicas

O método das aglomerações hierárquicas foi inicialmente proposto por Bertrand (1986) para classificar dados clássicos. A utilização desse método para classificar dados simbólicos foi proposta por Brito e Diday (1990) e por Brito (1994). Esse método consiste em um processo de particionamento do tipo ascendente (bottom-up) e os princípios para a determinação dos *clusters* são aplicados de maneira inversa aos descritos para o método das divisões hierárquicas, isto é, começam com um objeto em cada *cluster* que vão sendo aglomerados sucessivamente, até que todos os objetos estejam agrupados em um único *cluster*. No entanto, a complexidade computacional empregada no método descendente faz com que o método ascendente seja mais utilizado. Inicialmente, cada ensaio é representado por um único *cluster*. Esse método consiste em realizar sucessivos agrupamentos para produzir um único cluster que contem todos os ensaios. Os resultados desses agrupamentos são normalmente mostrados em uma árvore hierárquica ou dendograma.

Nesse método, os *clusters* são associados de acordo com o critério de distância para minimizar o quadrado da soma das distâncias entre os objetos de um mesmo *cluster* e maximizar o quadrado da soma das distâncias entre os membros de objetos diferentes. O método inicia agrupando os dois *clusters* mais próximos de acordo com um critério de agregação que determina o grau de similaridade entre estes objetos.

As etapas de um algoritmo de aglomeração hierárquica são listadas a seguir:

- 1) Inicia-se o procedimento de particionamento com  $k$  *clusters*, cada um contendo um único objeto.
- 2) Calcula-se a distância entre cada *cluster*.
- 3) A menor distância é escolhida e os dois *clusters* que fornecem essa menor distância são agrupados em um novo *cluster*.
- 4) As distâncias entre o novo *cluster* e os demais são calculadas considerando o critério de agregação adotado.
- 5) Repetem-se os dois passos anteriores até que todos os objetos se encontrem em um mesmo *cluster*.

#### 4.2.1. Critérios de agregação

Após a escolha da medida da distância, uma matriz de distâncias é construída. Como o método em questão é o agrupamento hierárquico, o primeiro passo consiste em determinar quais os ensaios mais próximos e com eles formar um novo *cluster*.

Após o primeiro passo, a matriz de distâncias é recalculada em passos sucessivos até a formação de um único *cluster*. Porém, é preciso estabelecer um critério de como medir a distância entre um ensaio e um *cluster*, ou entre dois *clusters*. Para isso, são estabelecidos os critérios de agregação.

Para ilustrar o procedimento do método da aglomeração hierárquica, é considerado um conjunto contendo 5 ensaios dinâmicos descritos por histogramas  $\Omega = \{E_1, \dots, E_n\}$ . Neste exemplo, o grau de similaridade utilizado entre dois ensaios hipotéticos  $E_i$  e  $E_j$  é calculado a partir da distância Euclidiana.

$$E_1 = \{[0-20](0,10);[20-40](0,15);[40-60](0,25);[60-80](0,30);[80-100](0,20)\}$$

$$E_2 = \{[0-20](0,05);[20-40](0,05);[40-60](0,40);[60-80](0,25);[80-100](0,25)\}$$

$$E_3 = \{[0-20](0,05);[20-40](0,15);[40-60](0,60);[60-80](0,15);[80-100](0,05)\}$$

$$E_4 = \{[0-20](0,10);[20-40](0,20);[40-60](0,45);[60-80](0,10);[80-100](0,05)\}$$

$$E_5 = \{[0-20](0,05);[20-40](0,10);[40-60](0,55);[60-80](0,20);[80-100](0,10)\}$$

Dessa forma, para os ensaios apresentados, o grau de similaridade entre os ensaios  $E_1$  e  $E_2$  é calculado pela equação:

$$gs(E_1, E_2) = \left( \sqrt{(0,10 - 0,05)^2} + \sqrt{(0,15 - 0,05)^2} + \sqrt{(0,25 - 0,40)^2} + \sqrt{(0,30 - 0,25)^2} + \sqrt{(0,20 - 0,25)^2} \right) = 0,20$$

De maneira similar, as outras distâncias são calculadas para todos os ensaios para a construção de uma matriz com os graus de similaridade referentes a cada par de ensaios. Para o exemplo considerado, a matriz com os graus de similaridade é apresentada na tabela 4.1.

TABELA 4.1 – Graus de similaridade calculados para cada par de ensaios.

E <sub>1</sub>	E <sub>2</sub>	E <sub>3</sub>	E <sub>4</sub>	E <sub>5</sub>	
0	0,20	0,41	0,32	0,34	E <sub>1</sub>
-	0	0,32	0,30	0,22	E <sub>2</sub>
-	-	0	0,17	0,10	E <sub>3</sub>
-	-	-	0	0,19	E <sub>4</sub>
-	-	-	-	0	E <sub>5</sub>

O par de ensaios com o menor grau de similaridade (nesse caso,  $gs(E_3, E_5) = 0,10$ ) é o primeiro a formar um novo *cluster*. No exemplo considerado, o novo *cluster* é o  $C^1 = \{E_3, E_5\}$ .

O segundo agrupamento é realizado utilizando o mesmo procedimento anterior, porém, os ensaios  $E_3$  e  $E_5$  são agora substituídos pelo *cluster*  $C^1$ . Para o cálculo do grau de similaridade entre um ensaio e o *cluster*  $C^1$ , é aplicado o critério de agregação. Os critérios de agregação são apresentados a seguir, juntamente com a continuação do exemplo anterior.

#### 4.2.1.1. Critério do vizinho mais próximo – ‘Single’ (Single Linkage)

Sejam  $n_{C^i}$  e  $n_{C^j}$  o número de objetos nos *clusters*  $C^i$  e  $C^j$ ,  $E_i$  o  $i^{\text{ésimo}}$  objeto do *cluster*  $C^i$  e  $E_j$  o  $j^{\text{ésimo}}$  objeto no *cluster*  $C^j$ , então:

$$gs(C^i, C^j) = \min \left( gs(E_i, E_j) \right) \quad i = 1, \dots, n_{C^i}; j = 1, \dots, n_{C^j} \quad (4.8)$$

Dados dois *clusters*  $C^i$  e  $C^j$ , a distância entre eles é dada pelo valor mínimo de todas as distâncias entre todos os ensaios de  $C^i$  e todos os ensaios de  $C^j$ .

Utilizando o critério de agregação do vizinho mais próximo (*single*), as distâncias entre o *cluster*  $C^1 = \{E_3, E_5\}$  e os demais ensaios do exemplo anterior são calculadas da seguinte forma:

$$gs(C^1, E_1) = \min(gs(E_3, E_1); gs(E_5, E_1)) = \min(0,41; 0,34) = 0,34$$

$$gs(C^1, E_2) = \min(gs(E_3, E_2); gs(E_5, E_2)) = \min(0,32; 0,22) = 0,22$$

$$gs(C^1, E_4) = \min(gs(E_3, E_4); gs(E_5, E_4)) = \min(0,17; 0,19) = 0,17$$

Uma nova matriz contendo os graus de similaridade é apresentada na tabela 4.2, porém, os ensaios  $E_3$  e  $E_5$  são substituídos pelo *cluster*  $C^1$ .

TABELA 4.2 - Graus de similaridade calculados para os pares de ensaios no segundo agrupamento.

E <sub>1</sub>	E <sub>2</sub>	E <sub>4</sub>	C <sup>1</sup>	
0	0,20	0,32	0,34	E <sub>1</sub>
-	0	0,30	0,22	E <sub>2</sub>
-	-	0	0,17	E <sub>4</sub>
-	-	-	0	C <sup>1</sup>

O segundo *cluster* a ser criado consiste em agrupar o par de ensaios com o menor grau de similaridade, levando em consideração a nova matriz. O novo *cluster* criado é, portanto, o  $C^2 = \{E_3, E_4, E_5\}$ .

Para realizar um terceiro agrupamento, uma nova matriz com os graus de similaridade é calculada, levando-se em conta o critério de agregação do vizinho mais próximo. Dessa vez, o critério de agregação vai definir as distâncias entre o *cluster*  $C^2 = \{E_3, E_4, E_5\}$  e os ensaios  $E_1$  e  $E_2$ . A nova matriz com os graus de similaridade é mostrada na tabela 4.3.

$$gs(C^2, E_1) = \min(gs(E_3, E_1); gs(E_4, E_1); gs(E_5, E_1)) = \min(0,41; 0,32; 0,34) = 0,32$$

$$gs(C^2, E_2) = \min(gs(E_3, E_2); gs(E_4, E_2); gs(E_5, E_2)) = \min(0,32; 0,30; 0,22) = 0,22$$

TABELA 4.3 - Grau de similaridade calculado para os pares de ensaios no terceiro agrupamento.

E <sub>1</sub>	E <sub>2</sub>	C <sup>2</sup>	
0	0,20	0,32	E <sub>1</sub>
-	0	0,22	E <sub>2</sub>
-	-	0	C <sup>2</sup>

O terceiro *cluster* criado agrupa os ensaios  $E_1$  e  $E_2$ . O único agrupamento que falta ser realizado é entre os *clusters*  $C^2$  e  $C^3$ . Para saber o grau de similaridade entre eles, é aplicado o critério de agregação da seguinte forma:

$$\begin{aligned} &gs(C^2, C^3) \\ &= \min(gs(E_3, E_1); gs(E_3, E_2); gs(E_4, E_1); gs(E_4, E_2); gs(E_5, E_1); gs(E_5, E_2)) \\ &= \min(0,41; 0,32; 0,32; 0,30; 0,34; 0,22) = 0,22 \end{aligned}$$

Após esses agrupamentos, a hierarquia está completamente definida. A figura 4.1 ilustra o dendograma final obtido e os *clusters* agrupados. É possível notar que esse método de particionamento mostra a proximidade entre os *clusters* através da diferença de altura entre eles.

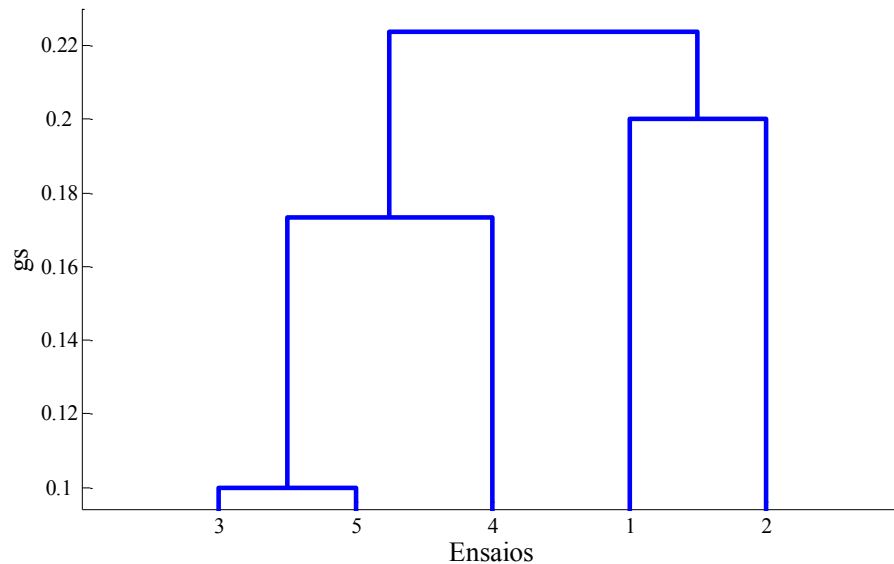


FIGURA 4.1 – Exemplo da aglomeração hierárquica utilizando o critério de agregação do vizinho mais próximo.

#### 4.2.1.2. Critério do vizinho mais distante – ‘Complete’ (Complete Linkage)

Sejam  $n_{C^i}$  e  $n_{C^j}$  o número de objetos nos *clusters*  $C^i$  e  $C^j$ ,  $E_i$  o  $i^{\text{ésimo}}$  objeto do *cluster*  $C^i$  e  $E_j$  o  $j^{\text{ésimo}}$  objeto no *cluster*  $C^j$ , então:

$$gs(C^i, C^j) = \max(gs(E_i, E_j)) \quad i = 1, \dots, n_{C^i}; j = 1, \dots, n_{C^j} \quad (4.9)$$



Dados dois *clusters*  $C^i$  e  $C^j$ , o critério de agregação entre eles é dada pelo valor máximo de todas as distâncias entre todos os ensaios de  $C^i$  e todos os ensaios de  $C^j$ .

Para o exemplo da aglomeração hierárquica anterior, utilizando o critério de agregação do vizinho mais distante (*complete*), as distâncias entre o *cluster*  $C^1 = \{E_3, E_5\}$  e os ensaios  $E_1, E_2$  e  $E_4$  são calculadas da seguinte forma:

$$gs(C^1, E_1) = \max(gs(E_3, E_1); gs(E_5, E_1)) = \max(0,41; 0,34) = 0,41$$

$$gs(C^1, E_2) = \max(gs(E_3, E_2); gs(E_5, E_2)) = \max(0,32; 0,22) = 0,32$$

$$gs(C^1, E_4) = \max(gs(E_3, E_4); gs(E_5, E_4)) = \max(0,17; 0,19) = 0,19$$

A nova matriz contendo os graus de similaridade é apresentada na tabela 4.4. Os graus de similaridade utilizados para a coluna do *cluster*  $C^1$ , dessa vez, são os obtidos através da aplicação do vizinho mais distante.

TABELA 4.4 - Graus de similaridade calculados para os pares de ensaios no segundo agrupamento.

E <sub>1</sub>	E <sub>2</sub>	E <sub>4</sub>	C <sup>1</sup>	
0	0,20	0,32	0,41	E <sub>1</sub>
-	0	0,30	0,32	E <sub>2</sub>
-	-	0	0,19	E <sub>4</sub>
-	-	-	0	C <sup>1</sup>

O segundo *cluster* criado contém o par de ensaios com menor grau de similaridade,  $C^2 = \{E_3, E_4, E_5\}$ .

Os demais agrupamentos são realizados considerando o menor grau de similaridade e o critério de agregação entre dois *clusters* continua a ser calculado através do vizinho mais distante, até que a hierarquia seja completamente definida. A figura 4.2 mostra o dendograma obtido. É possível notar que o agrupamento final é o mesmo que aquele obtido utilizando a regra do vizinho mais próximo, porém, os graus de similaridade entre os *clusters* formados, são diferentes. Por este motivo, os dendogramas obtidos utilizando os demais critérios de agregação serão omitidos nas seções seguintes.

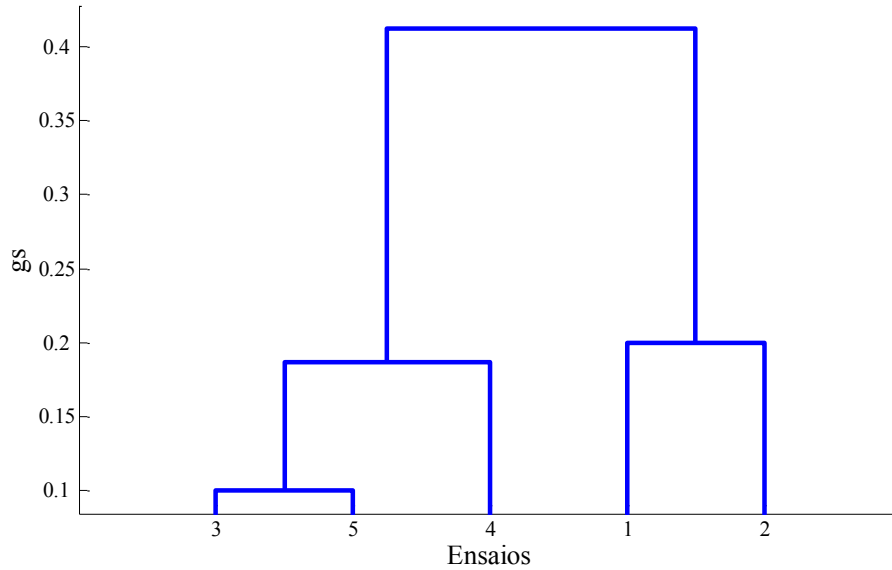


FIGURA 4.2 – Exemplo da aglomeração hierárquica utilizando o critério de agregação do vizinho mais distante.

#### 4.2.1.3. Critério da média do cluster – ‘Average’ (Unweighted pair group average)

Sejam  $n_{C^i}$  e  $n_{C^j}$  o número de objetos nos *clusters*  $C^i$  e  $C^j$ ,  $E_i$  o  $i^{\text{ésimo}}$  objeto do *cluster*  $C^i$  e  $E_j$  o  $j^{\text{ésimo}}$  objeto no *cluster*  $C^j$ , então:

$$gs(C^i, C^j) = \frac{1}{n_{C^i} n_{C^j}} \sum_{k=1}^{n_{C^i}} \sum_{l=1}^{n_{C^j}} gs(E_k, E_l) \quad (4.10)$$

A distância considerada entre os *clusters*  $C^i$  e  $C^j$  é a média de todas as distâncias entre todos os ensaios de  $C^i$  e todos os ensaios de  $C^j$ .

Continuando o exemplo anterior e utilizando o critério de agregação *average*, as distâncias entre o *cluster*  $C^1 = \{E_3, E_5\}$  e os ensaios  $E_1$ ,  $E_2$  e  $E_4$ , são calculadas da seguinte forma:

$$gs(E_1, C^1) = \frac{1}{1 \times 2} (gs(E_1, E_3) + gs(E_1, E_5)) = \frac{1}{2} (0,41 + 0,34) = 0,37$$

$$gs(E_2, C^1) = \frac{1}{1 \times 2} (gs(E_2, E_3) + gs(E_2, E_5)) = \frac{1}{2} (0,32 + 0,22) = 0,27$$

$$gs(E_4, C^1) = \frac{1}{1 \times 2} (gs(E_4, E_3) + gs(E_4, E_5)) = \frac{1}{2} (0,17 + 0,19) = 0,18$$

A nova matriz contendo os graus de similaridade é apresentada na tabela 4.5.

TABELA 4.5 - Graus de similaridade calculados para os pares de ensaios no segundo agrupamento.

E <sub>1</sub>	E <sub>2</sub>	E <sub>4</sub>	C <sup>1</sup>	
0	0,20	0,32	0,37	E <sub>1</sub>
-	0	0,30	0,27	E <sub>2</sub>
-	-	0	0,18	E <sub>4</sub>
-	-	-	0	C <sup>1</sup>

O segundo *cluster* criado,  $C^2 = \{E_3, E_4, E_5\}$ , agrupa o par de ensaios com menor grau de similaridade,  $gs(C^1, E_4) = 0,18$ . A seguir, o critério de agregação vai definir as distâncias entre o *cluster*  $C^2 = \{E_3, E_4, E_5\}$  e os ensaios  $E_1$  e  $E_2$ . A nova matriz com os graus de similaridade é mostrada na tabela 4.6.

$$gs(E_1, C^2) = \frac{1}{1 \times 3} (gs(E_1, E_3) + gs(E_1, E_4) + gs(E_1, E_5))$$

$$= \frac{1}{3} (0,41 + 0,32 + 0,34) = 0,36$$

$$gs(E_2, C^2) = \frac{1}{1 \times 3} (gs(E_2, E_3) + gs(E_2, E_4) + gs(E_2, E_5))$$

$$= \frac{1}{3} (0,32 + 0,30 + 0,22) = 0,28$$

TABELA 4.6 - Grau de similaridade calculado para os pares de ensaios no terceiro agrupamento.

E <sub>1</sub>	E <sub>2</sub>	C <sup>2</sup>	
0	0,20	0,36	E <sub>1</sub>
-	0	0,28	E <sub>2</sub>
-	-	0	C <sup>2</sup>

O terceiro *cluster* criado agrupa os ensaios  $E_1$  e o  $E_2$ . O grau de similaridade entre os *clusters*  $C^2$  e  $C^3$  é calculado da seguinte forma:

$$\begin{aligned} gs(C^2, C^3) &= \frac{1}{3 \times 2} (gs(E_3, E_1) + gs(E_3, E_2) + gs(E_4, E_1) + gs(E_4, E_2) \\ &\quad + gs(E_5, E_1) + gs(E_5, E_2)) \\ &= \frac{1}{6} (0,41 + 0,32 + 0,32 + 0,30 + 0,34 + 0,22) = 0,32 \end{aligned}$$

#### 4.2.1.4. Critério da média simples – ‘Weighted’ (Weighted pair group average).

Seja o *cluster*  $C^k = \{C^i, C^j\}$  e a distância entre o *cluster*  $C^k$  e o *cluster*  $C^l$  definida da seguinte forma:

$$gs(C^k, C^l) = \frac{(gs(C^i, C^l) + gs(C^j, C^l))}{2} \quad (4.11)$$

Utilizando o critério de agregação *weighted*, as distâncias entre o *cluster*  $C^1 = \{E_3, E_5\}$  e os ensaios  $E_1, E_2$  e  $E_4$  são calculadas da seguinte forma:

$$gs(C^1, E_1) = \frac{(gs(E_3, E_1) + gs(E_5, E_1))}{2} = \frac{0,41 + 0,34}{2} = 0,37$$

$$gs(C^1, E_2) = \frac{(gs(E_3, E_2) + gs(E_5, E_2))}{2} = \frac{0,32 + 0,22}{2} = 0,27$$

$$gs(C^1, E_4) = \frac{(gs(E_3, E_4) + gs(E_5, E_4))}{2} = \frac{0,17 + 0,19}{2} = 0,18$$

A nova matriz contendo os graus de similaridade é apresentada na tabela 4.7.

TABELA 4.7 - Graus de similaridade calculados para os pares de ensaios no segundo agrupamento.

E <sub>1</sub>	E <sub>2</sub>	E <sub>4</sub>	C <sup>1</sup>	
0	0,20	0,32	0,37	E <sub>1</sub>
-	0	0,30	0,27	E <sub>2</sub>
-	-	0	0,18	E <sub>4</sub>
-	-	-	0	C <sup>1</sup>

Como o menor grau de similaridade encontrado na tabela 4.7 é o  $gs(C^1, E_4) = 0,18$ , o novo *cluster* criado é o  $C^2 = \{E_3, E_4, E_5\}$ . A nova matriz com os graus de similaridade (tabela 4.8) é calculada utilizando o critério de agregação *weighted* para calcular as distâncias entre o *cluster*  $C^2$  e os ensaios  $E_1$  e  $E_2$ .

$$gs(C^2, E_1) = \frac{(gs(E_3, E_1) + gs(E_4, E_1) + gs(E_5, E_1))}{3} = \frac{0,41 + 0,32 + 0,34}{3} = 0,36$$

$$gs(C^2, E_2) = \frac{(gs(E_3, E_2) + gs(E_4, E_2) + gs(E_5, E_2))}{3} = \frac{0,32 + 0,3 + 0,22}{3} = 0,28$$

TABELA 4.8 - Grau de similaridade calculado para os pares de ensaios no terceiro agrupamento.

E <sub>1</sub>	E <sub>2</sub>	C <sup>2</sup>	
0	0,20	0,36	E <sub>1</sub>
-	0	0,28	E <sub>2</sub>
-	-	0	C <sup>2</sup>

O terceiro *cluster* criado é o  $C^3 = \{E_1, E_2\}$ . O grau de similaridade entre os *clusters*  $C^2$  e  $C^3$  é definido a seguir:

$$gs(C^3, C^2) = \frac{(gs(E_1, C^2) + gs(E_2, C^2))}{2} = \frac{0,36 + 0,28}{2} = 0,32$$

#### 4.2.1.5. Critério do centroide – ‘Centroid’ (Unweighted pair group centroid)

Sejam  $\bar{E}_{C^i}$ ,  $n_{C^i}$  e  $E_i$ , respectivamente, o centroide, o número de objetos do *cluster*  $i$  e o  $i^{\text{ésimo}}$  ensaio do *cluster*  $i$ , então:

$$\bar{E}_{C^i} = \frac{1}{n_{C^i}} \sum_{i=1}^{n_{C^i}} E_i \quad (4.12)$$

$$gs(C^i, C^j) = \sqrt{(\bar{E}_{C^i} - \bar{E}_{C^j})^2} \quad (4.13)$$

O grau de similaridade é medido através da distância euclidiana entre os centroides dos *clusters*  $C^i$  e  $C^j$ .

Calculando os graus de similaridade do exemplo entre o *cluster*  $C^1 = \{E_3, E_5\}$  e os ensaios  $E_1, E_2$  e  $E_4$ , obtém-se:

$$\begin{aligned} \bar{E}_{C^1} &= \frac{\vec{E}_3 + \vec{E}_5}{2} \\ &= \frac{(0,05 + 0,05)}{2}; \frac{(0,15 + 0,1)}{2}; \frac{(0,6 + 0,55)}{2}; \frac{(0,15 + 0,2)}{2}; \frac{(0,05 + 0,1)}{2} \\ &= 0,05; 0,125; 0,575; 0,175; 0,075 \end{aligned}$$

$$\begin{aligned} gs(E_1, C^1) &= \\ &= \sqrt{(0,1 - 0,05)^2 + (0,15 - 0,125)^2 + (0,575 - 0,25)^2 + (0,175 - 0,3)^2 + (0,075 - 0,2)^2} = \\ &= 0,37 \end{aligned}$$

$$\begin{aligned} gs(E_2, C^1) &= \\ &= \sqrt{(0,05 - 0,05)^2 + (0,05 - 0,125)^2 + (0,4 - 0,25)^2 + (0,25 - 0,3)^2 + (0,25 - 0,2)^2} \\ &= 0,27 \end{aligned}$$

$$\begin{aligned} gs(E_4, C^1) &= \\ &= \sqrt{(0,1 - 0,05)^2 + (0,2 - 0,125)^2 + (0,45 - 0,25)^2 + (0,1 - 0,3)^2 + (0,05 - 0,2)^2} \\ &= 0,17 \end{aligned}$$

A matriz com os graus de similaridade é mostrada na tabela 4.9.

TABELA 4.9 - Graus de similaridade calculados para os pares de ensaios no segundo agrupamento.

E <sub>1</sub>	E <sub>2</sub>	E <sub>4</sub>	C <sup>1</sup>	
0	0,20	0,32	0,37	E <sub>1</sub>
-	0	0,30	0,27	E <sub>2</sub>
-	-	0	0,17	E <sub>4</sub>
-	-	-	0	C <sup>1</sup>

O segundo cluster é o  $C^2 = \{E_3, E_4, E_5\}$ . Calculando os graus de similaridade, obtem-se:

$$gs(E_1, C^2) = 0,35$$

$$gs(E_2, C^2) = 0,27$$

A nova matriz com os graus de similaridade é mostrada na tabela 4.10.

TABELA 4.10 - Grau de similaridade calculado para os pares de ensaios no terceiro agrupamento.

E <sub>1</sub>	E <sub>2</sub>	C <sup>2</sup>	
0	0,20	0,35	E <sub>1</sub>
-	0	0,27	E <sub>2</sub>
-	-	0	C <sup>2</sup>

Para completar a hierarquia, o grau de similaridade entre os *clusters*  $C^2$  e  $C^3$  é definido a seguir:

$$gs(C^2, C^3) = 0,30$$

#### 4.2.1.6. Critério da mediana – ‘Median’ (Weighted pair group centroid)

Seja  $\tilde{C}^i$  o centroide do cluster  $C^i$  formado pelos clusters  $C^k$  e  $C^l$ , então:

$$\tilde{C}^i = \frac{1}{2}(\tilde{C}^k + \tilde{C}^l) \quad (4.14)$$

$$gs(C^i, C^j) = \sqrt{(\tilde{C}^i + \tilde{C}^j)^2} \quad (4.15)$$

A distância entre dois *clusters*  $C^i$  e  $C^j$  é a distância entre os centroides de  $C^i$  e  $C^j$ , mas, a obtenção do centroide do novo *cluster* não leva em consideração a quantidade de ensaios que cada um contém, sendo suas coordenadas a média simples dos centroides dos *clusters* que se uniram.

Utilizando o critério de agregação *median*, as distâncias entre os *clusters*  $C^1 = \{E_3, E_5\}$  e os demais ensaios do exemplo anterior são calculadas da seguinte forma:

$$\begin{aligned} \tilde{C}^1 &= \frac{1}{2}(\tilde{E}_3 + \tilde{E}_5) \\ &= \frac{1}{2}(0,05 + 0,05); (0,05 + 0,1); (0,4 + 0,55); (0,25 + 0,2); (0,25 \\ &\quad + 0,1) = (0,05; 0,125; 0,575; 0,175; 0,075) \end{aligned}$$

$$\begin{aligned} gs(C^1, E_1) &= \\ &= \sqrt{(0,05 - 0,1)^2 + (0,125 - 0,15)^2 + (0,575 - 0,25)^2 + (0,175 - 0,3)^2 + (0,075 - 0,2)^2} = \\ &= 0,37 \end{aligned}$$

$$\begin{aligned} gs(C^1, E_2) &= \\ &= \sqrt{(0,05 - 0,05)^2 + (0,125 - 0,05)^2 + (0,575 - 0,4)^2 + (0,175 - 0,25)^2 + (0,075 - 0,25)^2} = \\ &= 0,27 \end{aligned}$$

$$\begin{aligned} gs(C^1, E_4) &= \\ &= \sqrt{(0,05 - 0,1)^2 + (0,125 - 0,2)^2 + (0,575 - 0,45)^2 + (0,175 - 0,1)^2 + (0,075 - 0,05)^2} = \\ &= 0,17 \end{aligned}$$

A nova matriz com os graus de similaridade é apresentada na tabela 4.11.

TABELA 4.11 - Graus de similaridade calculados para os pares de ensaios no segundo agrupamento.

E <sub>1</sub>	E <sub>2</sub>	E <sub>4</sub>	C <sup>1</sup>	
0	0,20	0,32	0,37	E <sub>1</sub>
-	0	0,30	0,27	E <sub>2</sub>
-	-	0	0,17	E <sub>4</sub>
-	-	-	0	C <sup>1</sup>



O segundo *cluster* a ser criado é o  $C^2 = \{E_3, E_4, E_5\}$ . O novo centroide  $\tilde{C}^2$  e os novos graus de similaridade são calculados a seguir:

$$\begin{aligned}\tilde{C}^2 &= \frac{1}{2}(\tilde{C}^1 + \tilde{E}_4) \\ &= \frac{1}{2}(0,05 + 0,1); (0,125 + 0,2); (0,575 + 0,45); (0,175 + 0,1); (0,075 \\ &\quad + 0,05) = (0,075; 0,1625; 0,5125; 0,1375; 0,0625)\end{aligned}$$

$$gs(C^2, E_1) = 0,34$$

$$gs(C^2, E_2) = 0,27$$

A nova matriz com os graus de similaridade é mostrada na tabela 4.12.

TABELA 4.12- Grau de similaridade calculado para os pares de ensaios no terceiro agrupamento.

E <sub>1</sub>	E <sub>2</sub>	C <sup>2</sup>	
0	0,20	0,34	E <sub>1</sub>
-	0	0,27	E <sub>2</sub>
-	-	0	C <sup>2</sup>

O terceiro *cluster* a ser criado é o  $C^3 = \{E_1, E_2\}$ . O novo centroide  $\tilde{C}^3$  e o grau de similaridade entre os *clusters*  $C^2$  e  $C^3$ , são calculados a seguir:

$$\begin{aligned}\tilde{C}^3 &= \frac{1}{2}(\tilde{E}_1 + \tilde{E}_2) \\ &= \frac{1}{2}(0,1 + 0,05); (0,15 + 0,05); (0,25 + 0,4); (0,30 + 0,25); (0,20 \\ &\quad + 0,25) = (0,075; 0,1; 0,325; 0,275; 0,225)\end{aligned}$$

$$gs(C^2, C^3)$$

$$\begin{aligned}&= \sqrt{(0,075 - 0,075)^2 + (0,1625 - 0,1)^2 + (0,5125 - 0,325)^2 + (0,1375 - 0,275)^2 \\ &\quad + (0,0625 - 0,225)^2} \\ &= 0,29\end{aligned}$$

#### 4.2.1.7. Critério da mínima variância (Ward)

Este critério se baseia no pressuposto de que os *clusters* devem ter a máxima homogeneidade interna possível. Ward propôs uma técnica de minimização da soma dos quadrados das distâncias de cada ponto ao respectivo centroide do *cluster*. Em cada agrupamento, o novo *cluster* combina os dois *clusters* cuja união provoca o menor aumento da soma dos quadrados destas distâncias. Dado o elevado número de cálculos e a complexidade de se realizá-los manualmente, a demonstração deste critério será omitida. Maiores detalhes podem ser encontrados em (MATLAB, 2012).

### 4.3. Método *k-means*

O algoritmo *k-means* é classificado como um método não-hierárquico (ou por partição), em que o número de grupos ou *clusters* é considerado fixo e inicialmente definido. A diferença entre algoritmos de agrupamentos hierárquicos e não-hierárquicos é que os algoritmos por partição encontram todos os grupos simultaneamente e não impõem uma noção de hierarquia à estrutura de dados.

O algoritmo *k-means* também pode ser chamado de k-médias. O algoritmo *k-means* é bastante popular devido à sua facilidade de implementação (JAIN et al, 1999). Esse algoritmo utiliza o conceito de centroides que representam os *clusters*. Esses centroides são calculados a partir da média de todos os objetos do grupo (FONTANA; NALDI, 2009).

O objetivo deste algoritmo é encontrar a melhor divisão de  $n$  objetos em  $k$  grupos ( $C^i$   $i=1, \dots, k$ ) de forma que a distância total entre os objetos de um grupo e o seu respectivo centro, somada por todos os grupos, seja minimizada (PIMENTEL et al, 2003).

Em outras palavras, o algoritmo atribui aleatoriamente os  $n$  objetos a  $k$  grupos e calcula as médias dos objetos de cada grupo. Em seguida, cada objeto é deslocado para o grupo correspondente ao centroide do qual ele está mais próximo. Com este novo arranjo dos objetos em  $k$  grupos, novos centroides são calculados. O processo continua até que se chegue a uma situação em que todos os objetos estejam nos grupos dos seus centroides mais próximos.

O algoritmo k-means pode ser descrito pelas etapas a seguir (FONTANA; NALDI, 2009).

- 1) Seleciona-se o número  $k$  de *clusters* a serem formados.
- 2) Atribuem-se valores iniciais para os  $k$  centroides seguindo algum critério (sorteio aleatório desses valores dentro dos limites de domínio de cada objeto, por exemplo);
- 3) Atribui-se cada objeto ao grupo cujo centroide possua maior similaridade;
- 4) Recalcula-se o valor do centroide de cada grupo como sendo a média dos objetos atuais do grupo;
- 5) Repetem-se os dois passos anteriores até que os grupos se estabilizem, ou seja, os centroides se desloquem pouco em relação à sua posição anterior.

Para ilustrar o funcionamento do *k-means* considera-se um conjunto de dados em que se deseja dividi-los em três *clusters*.

Na figura 4.3 são gerados aleatoriamente três centroides, representados pelos círculos.

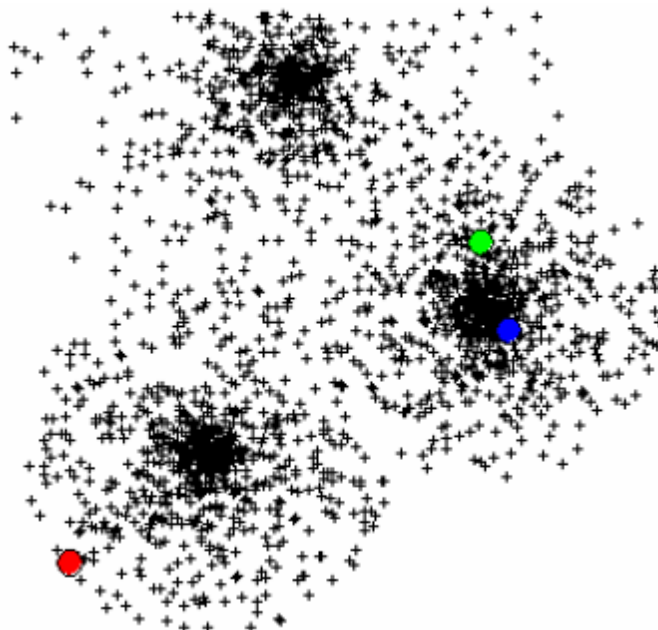


FIGURA 4.3 – Inicialização dos centroides.

Após a geração dos centroides, cada um dos objetos é atribuído ao grupo para o qual sua distância até o centroide seja a menor, como ilustrado na figura 4.4.

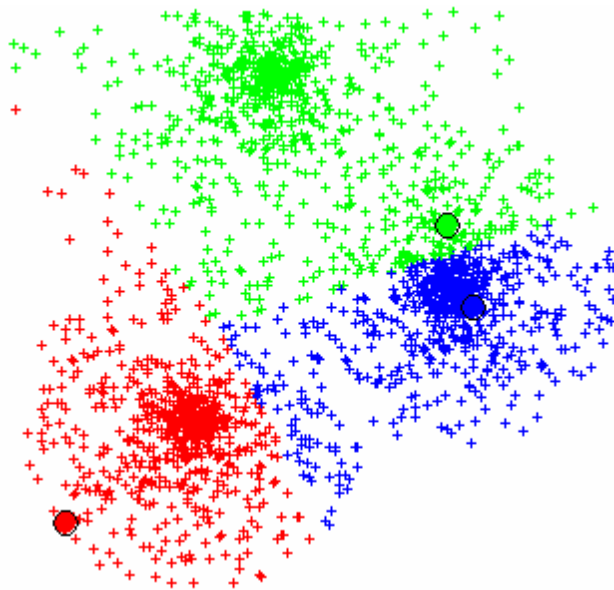


FIGURA 4.4 – Atribuição dos objetos aos *clusters*.

Em seguida, as posições dos centroides são recalculadas e os objetos são atribuídos aos *clusters* referentes aos novos centroides, como representado na figura 4.5. No momento em que as posições dos centroides não mais se alterarem, define-se que os objetos se encontram em seus *clusters* finais.

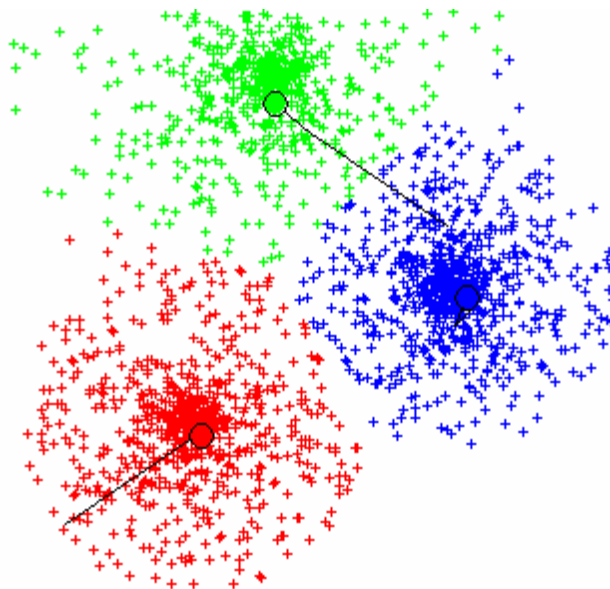


FIGURA 4.5 – Atualização dos centroides e nova atribuição dos objetos.

Um dos maiores problemas do algoritmo *k-means* deve-se à sensibilidade da seleção dos centroides iniciais, podendo convergir a um mínimo local caso os *clusters* iniciais não sejam devidamente escolhidos (JAIN et al, 1999).

Para exemplificar esse problema, a figura 4.6 mostra um exemplo com sete objetos. Se os centroides iniciais forem considerados os objetos A, B, C, então os *clusters* finais seriam  $\{(A), (B,C), (D,E,F,G)\}$  mostrados pelas elipses. Os objetos seriam melhor classificados, contudo, se os centroides iniciais escolhidos fossem, por exemplo, A, D e F. Nesse caso os *clusters* finais seriam  $\{(A,B,C), (D,E), (F,G)\}$  mostrados pelos retângulos.

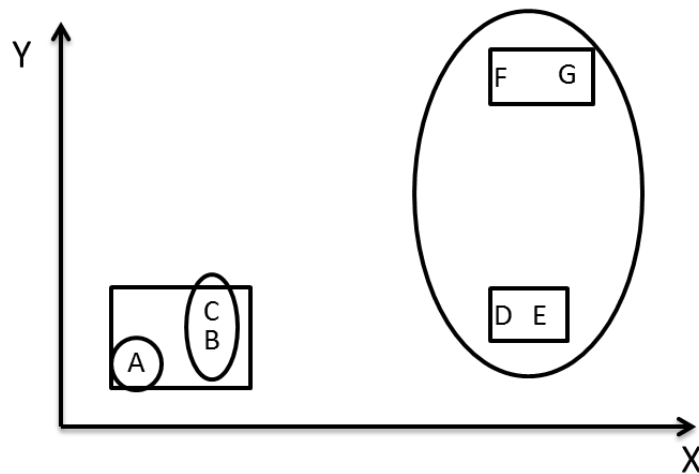


FIGURA 4.6 – Exemplo da sensibilidade do *k-means* à inicialização dos centroides.

FONTE – JAIN et al., 1999.

Diferentes formas de inicialização dos centroides podem ser utilizadas. Dentre elas, destacam-se:

**Amostrai** – Utiliza  $k$  objetos da base de dados escolhidas aleatoriamente.

**Uniforme** – Gera  $k$  objetos aleatórios distribuídos de maneira uniforme.

**Cluster** – são encontrados  $k$  subgrupos utilizando 10% do banco de dados. Os centroides resultantes serão utilizados como centroides iniciais para a classificação.

#### 4.4. Método *c-means*

Como visto nas técnicas anteriores, as abordagens tradicionais geram partições em que cada objeto (ensaio) pertence a somente um *cluster*, o que é conhecido como *clusterização hard*. Assim, nesse tipo de abordagem, os *clusters* são disjuntos. As técnicas de *clustering fuzzy* estende essa noção para permitir associar um objeto a todos os grupos usando uma função de pertinência (ZADEH, 1965). Esse tipo de técnica é útil em situações em que um objeto possui características de mais de um *cluster*. Um exemplo que pode ser citado é o da classificação de um anfíbio, que é uma espécie de animal que tem características aquáticas e terrestres. Neste caso, ele será parcialmente classificado aos dois *clusters*, através de um grau de pertinência que indica sua similaridade a determinado grupo.

Na figura 4.7 é ilustrada a diferença entre o *cluster hard* e o *cluster fuzzy* utilizando um mesmo exemplo em que 9 indivíduos são agrupados em dois *clusters*. Considerando o *cluster hard*, a partição seria  $H^1=\{1,2,3,4,5\}$  e  $H^2=\{6,7,8,9\}$  (representada pelos retângulos). Considerando o *clustering fuzzy*, os objetos são agrupados em dois *clusters*  $F^1$  e  $F^2$ , em que todos os indivíduos terão um grau de pertinência (probabilidade de pertencer ou não) definido no intervalo  $[0,1]$  para cada grupo. O *cluster*  $F^1$  pode ser descrito como  $\{(1; 1,0), (2; 1,0), (3; 1,0), (4; 0,70), (5; 0,65), (6; 0,40), (7; 0,30), (8; 0,0), (9; 0,0)\}$  e o *cluster*  $F^2$   $\{(1; 0,0), (2; 0,0), (3; 0,0), (4; 0,30), (5; 0,35), (6; 0,60), (7; 0,70), (8; 1,0), (9; 1,0)\}$ . Os pares ordenados  $(n,u_i)$  representam o objeto  $n$  e sua pertinência ao *cluster*  $i$ , respectivamente. Na figura 4.7, os *clusters fuzzy* são representados por elipses. Nota-se que os objetos 8 e 9 possuem grau de pertinência 0 e, por isso, encontram-se fora do *cluster*  $F^1$ . Já o objeto 6 encontra-se nos dois *clusters*, porém o seu grau de pertinência referente à  $F^2$  é maior que o referente à  $F^1$ , o que indica que ele possui mais semelhanças ao *cluster*  $F^2$  do que ao *cluster*  $F^1$ .

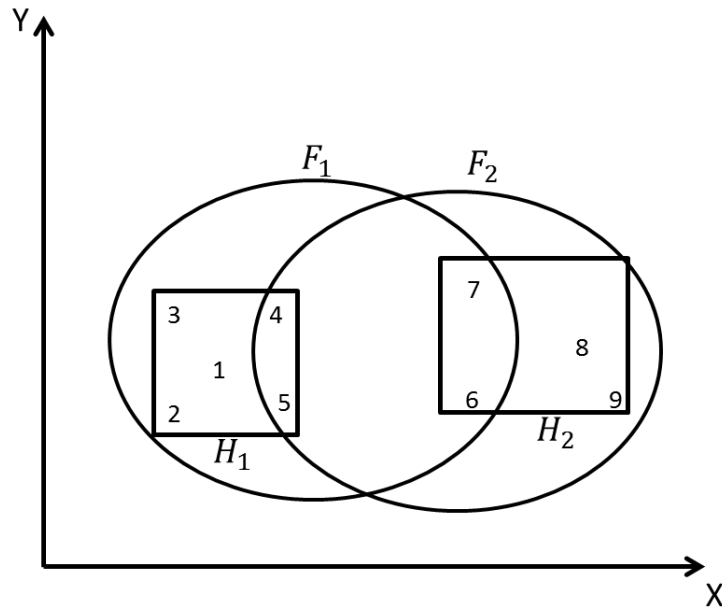


FIGURA 4.7 – Cluster hard x cluster fuzzy. FONTE – JAIN et al., 1999.

O algoritmo de *clusterização fuzzy* mais popular é o *c-means*. Considere-se um conjunto de objetos  $X = \{x_1, \dots, x_n\}$  em que se deseja organizá-los em  $k$  clusters,  $C = \{C^1, \dots, C^k\}$ . O *c-means* é um algoritmo não hierárquico cujo objetivo é fornecer uma partição de um conjunto de indivíduos em  $k$  clusters. Para isso, o *c-means* define uma função objetivo em que a ideia é minimizá-la para obter a melhor partição entre os indivíduos nos clusters. A função objetivo é definida como (BEZDEK, 1981):

$$J_2(U, G) = \sum_{i=1}^n \sum_{k=1}^c (u_{ik})^m \phi(x_i, g_k) \quad (4.16)$$

$$\phi(x_i, g_k) = \sum_{l=1}^p (x_{il} - g_{kl})^2 \quad (4.17)$$

onde  $U$  é a matriz de pertinência  $\{u_{ik}\}$  do indivíduo  $i$  ao cluster  $C^k$ ,  $G$  é o vetor dos centroides,  $m$  é o expoente *fuzzy* (grau de *fuzzyficação*) que varia de  $]1, \infty[$ , o que determina o grau de “confusão” entre os clusters criados e  $\phi(x_i, g_k)$  é a distância que mede a dissimilaridade entre um indivíduo  $i$  e um centroide de um cluster  $k$ .

Um exemplo do *fuzzy c-means* é apresentado utilizando uma base de dados composta por 140 dados a serem divididos em dois grupos, isto é, a ordem da matriz  $U$  é  $2 \times 140$ . Na figura 4.8 pode ser vista a distribuição de dados do exemplo.

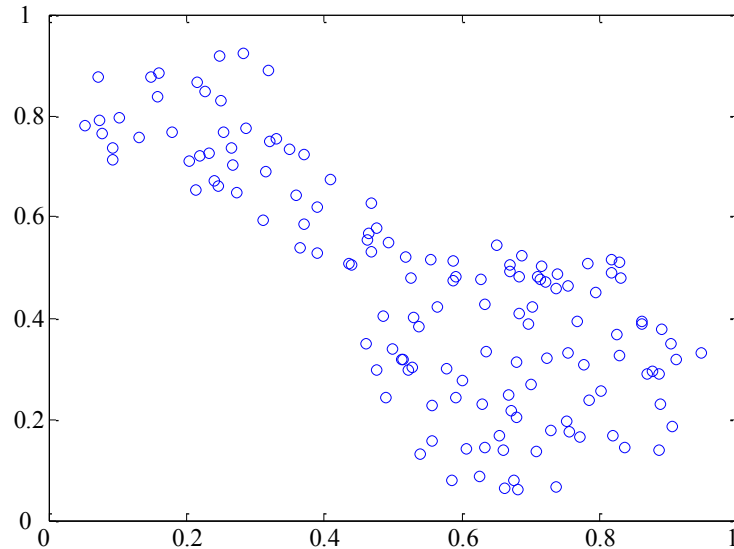


FIGURA 4.8 – Distribuição de dados. FONTE - Matlab®.

A figura 4.9 ilustra a convergência da função objetivo em função do número de iterações. A convergência da função objetivo varia com a partição inicial, gerada de forma aleatória. Em geral a função objetivo converge rápido. No caso do exemplo mostrado, a convergência ocorre a partir da quinta iteração.

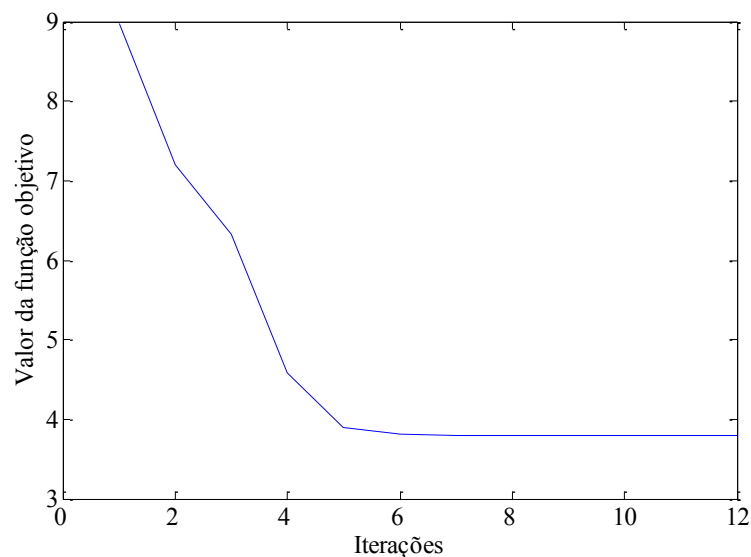


FIGURA 4.9 – Convergência da função objetivo. FONTE - Matlab®.



Finalmente, os dois grupos classificados pelo algoritmo fuzzy c-means são apresentados na figura 4.10.

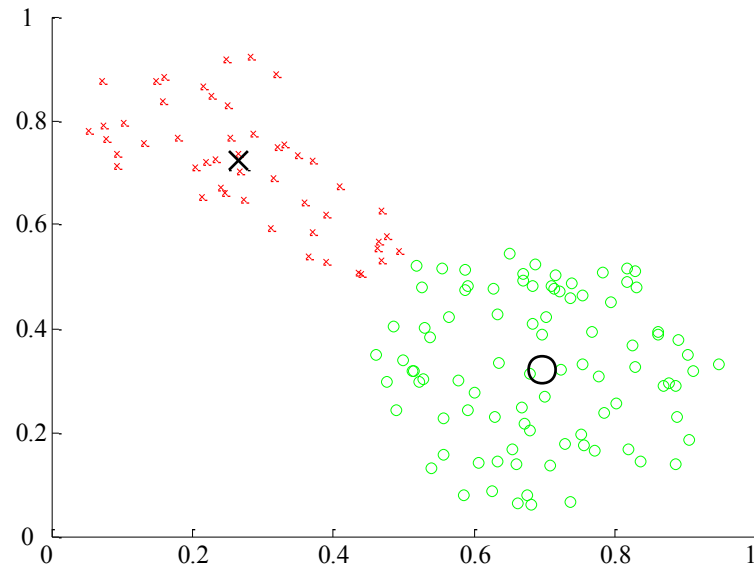


FIGURA 4.10 – Dados separados em dois agrupamentos. FONTE - Matlab®.

#### 4.5. Metodologia

A metodologia utilizada neste trabalho pode ser ilustrada através do fluxograma mostrado na figura 4.11.

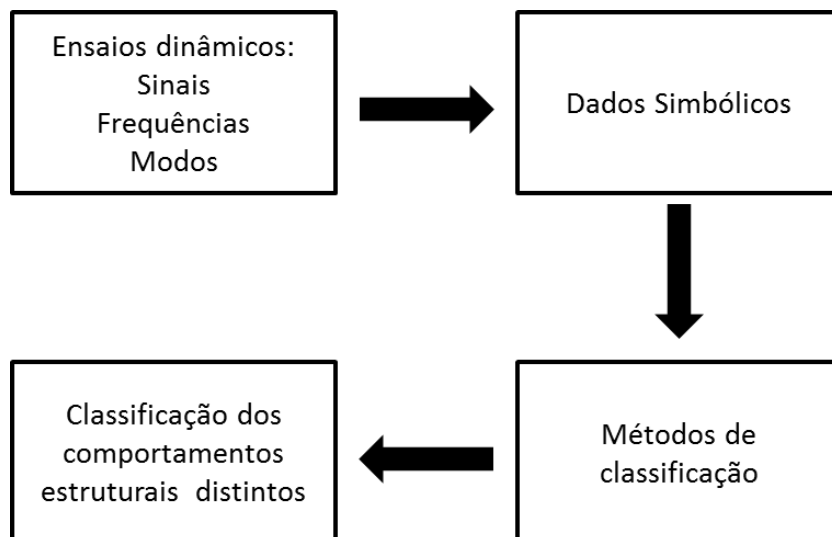


FIGURA 4.11 – Fluxograma da metodologia utilizada.

O princípio básico é de que os ensaios dinâmicos podem ser representados sejam pelos sinais, sejam pelos parâmetros modais (frequências naturais e modos próprios). Assim sendo, a fim de se descrever os ensaios de forma mais compacta, estes são representados de maneira simbólica (histogramas com 12 categorias). Para, de fato, diferenciar os estados estruturais, os ensaios dinâmicos representados de maneira simbólica são atribuídos aos métodos de classificação (aglomeração hierárquica, *k-means* e *c-means*). Após a aplicação dos métodos citados, obtém-se a classificação dos ensaios de acordo com os diferentes comportamentos estruturais.



# CAPÍTULO 5

## 5. SIMULAÇÕES NUMÉRICAS

O objetivo principal deste capítulo é apresentar, a partir de um estudo numérico, a aplicação dos métodos de classificação de dados descritos nas seções anteriores. A ideia é avaliar a robustez da abordagem baseada na combinação da ADS aos métodos de *clustering* que permitem discriminar os diferentes comportamentos estruturais. Para isso, diversas simulações de danos e de ruído são realizadas utilizando sinais, frequências naturais e modos de vibração, transformados em dados simbólicos.

Os estudos numéricos representam geralmente um primeiro passo para analisar o desempenho e a aptidão das técnicas de detecção de danos estruturais. Um vasto estudo estatístico relativo aos indicadores de danos ditos evoluídos foi realizado por (ALVANDI; CRÉMONA, 2006). As simulações numéricas permitem testar diferentes condições estruturais (diferentes níveis de danos, ruído, condições de contorno, materiais, etc). Portanto, considera-se relevante sua utilização para analisar a eficiência das abordagens propostas neste trabalho. Este estudo tem por objetivo a análise da sensibilidade dos métodos de *clustering* a partir de diversas simulações de danos e de ruído. Em particular, a ideia é observar se esses métodos são capazes de detectar um dano estrutural diante da existência de níveis de ruído nas medições dinâmicas (sinais).

Neste trabalho, um modelo numérico de uma viga bi-apoiada de aço de seção tipo I com 6 m de comprimento é estudado. A estrutura é modelada a partir de uma discretização utilizando o método dos elementos finitos utilizando-se um programa desenvolvido no Matlab® (ZIENKIEWICZ; TAYLOR, 2000). O modelo discretizado possui 200 elementos finitos, em que cada elemento é formado por dois nós com três graus de liberdade cada. A excitação é uma força aleatória aplicada no nó 178 (localizado a 0,69 m de distância do apoio direito) cuja amplitude e frequência variam aleatoriamente. Deve-se notar que essa posição permite melhor excitar as três primeiras frequências próprias (e modos próprios associados) da estrutura.

As respostas dinâmicas são medidas em dez pontos equidistantes da viga durante 100 segundos. A frequência de amostragem é de 1/100 s, o que corresponde a 10.000 instantes medidos por sensor. Na figura 5.1 é mostrada a representação esquemática da viga estudada, a posição da força de excitação e os pontos de medições. As propriedades relativas à geometria e à natureza do material da viga considerada são as seguintes:

- Módulo de elasticidade longitudinal: 210 GPa.
- Massa volumétrica:  $7850 \text{ kg.m}^{-3}$ .
- Seção transversal:  $2,81 \times 10^{-3} \text{ m}^2$ .
- Momento de inércia:  $1,845 \times 10^{-8} \text{ m}^4$ .

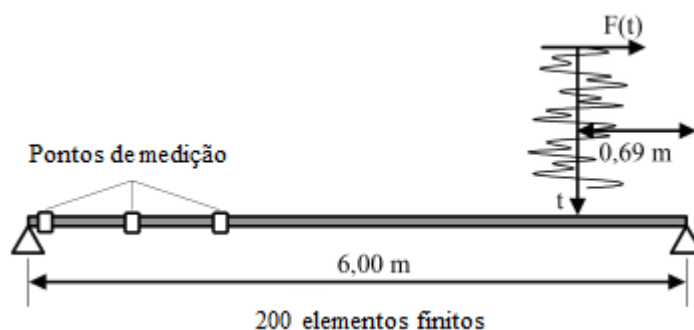


FIGURA 5.1 – Representação da viga estudada. FONTE – CURY, 2010.

Com o objetivo de analisar a robustez dos métodos de classificação considerando diferentes níveis de danos, diferentes localizações e ocorrências simultâneas, três configurações são propostas. Em todas as simulações realizadas, considera-se que os níveis de dano correspondem a uma redução do módulo de elasticidade longitudinal dos elementos danificados. O estado inicial da estrutura, denominado D0, representa a viga íntegra (sem danos). A primeira configuração com dano, denominada D1, consiste em uma redução de 20% do módulo de elasticidade longitudinal nos elementos 96 ao 105 (o que corresponde ao meio do vão da viga). Para a segunda configuração de dano, D2, além da redução de 20% nos elementos no meio da viga, também é considerada uma redução de 10% nos elementos 146 ao 155 (situados no quarto de vão da viga). A figura 5.2 detalha as configurações de danos impostos à estrutura.

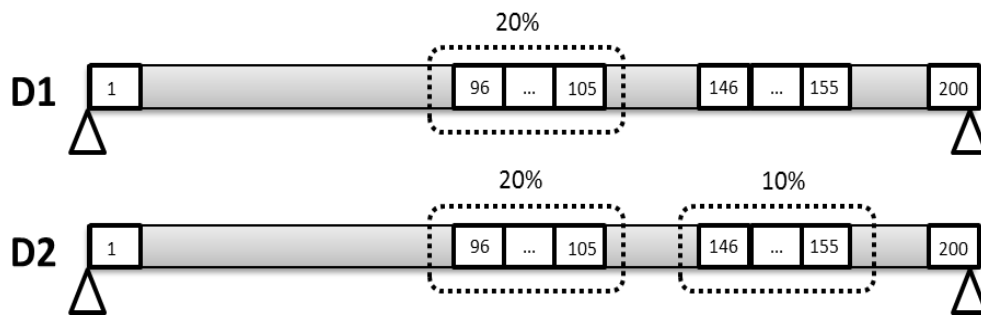


FIGURA 5.2 – Configurações de danos simulados.

Três níveis de ruído são considerados nas simulações. O primeiro deles, R0, indica as respostas dinâmicas sem ruído, R1 consiste em adicionar 5% de ruído às respostas dinâmicas e R2 consiste em adicionar 10% de ruído. Os ruídos são gerados através da multiplicação do nível de ruído pelo desvio-padrão de cada canal de medição (“sensor”) e por um vetor de valores aleatórios de uma distribuição gaussiana de média zero e desvio-padrão unitário (o que caracteriza um ruído branco). A resposta dinâmica com ruído é obtida adicionando a resposta sem ruído ao vetor de ruído, calculado da seguinte maneira:

$$X_{i,ruído} = X_i + \eta_{ruído} \cdot \sigma_{X_i} \cdot V \sim N(0,1) \quad (5.1)$$

em que  $X_i$  e  $X_{i,ruído}$  representam, respectivamente, as respostas dinâmicas da estrutura sem e com ruído medidas pelo sensor  $i$ ,  $\eta_{ruído}$  é o nível de ruído adicionado,  $\sigma_{X_i}$  o desvio-padrão de  $X_i$  e  $V \sim N(0,1)$  um vetor gaussiano de média nula e desvio-padrão unitário.

Para cada configuração de dano e cada nível de ruído, 10 ensaios diferentes foram simulados. Isso produz um total de 90 ensaios, como mostrado na tabela 5.1. Visto que para cada ensaio 10.000 pontos são medidos para cada um dos 10 sensores, existirão  $90 \times 10 \times 10.000 = 9.000.000$  (nove milhões) de valores a serem tratados. Esse cálculo dá uma visão geral do tamanho da base de dados a ser tratada e a necessidade de se utilizar ferramentas apropriadas, tais como a ADS, para manipular esses dados de maneira adequada.

TABELA 5.1 – Resumo dos ensaios simulados

		Níveis de Ruído		
		R0	R1	R2
Configurações de danos	D0	10	10	10
	D1	10	10	10
	D2	10	10	10

Os parâmetros modais da viga são identificados a partir do programa Sys-Ident desenvolvido no LCPC (Laboratoire Central des Ponts et Chaussées, Paris, França) em 2001 (BARBOSA; CRÉMONA, 2001). A identificação modal é realizada utilizando o método do decremento aleatório juntamente com o método de Ibrahim (COLE, 1968; IBRAHIM, 1977). O método do Decremento Aleatório é uma técnica de identificação modal não-paramétrica no domínio do tempo. Seu princípio é estimar as funções de decremento aleatório a partir das médias dos segmentos de medidas, obedecendo-se a condições precisas chamadas de desencadeamento. A partir destas funções aplicam-se os algoritmos clássicos de identificação modal, tais como o método de Ibrahim e é, então, possível extrair os parâmetros modais da estrutura (frequências naturais, taxas de amortecimento e modos próprios de vibração). Esse método propõe utilizar duas matrizes deduzidas da resposta transitória da estrutura, estimada através do método do decremento aleatório. Uma dessas matrizes é a matriz da resposta e a outra, defasada no tempo. O método de Ibrahim possui um inconveniente de necessitar de um número de canais de medidas superior ao número de parâmetros modais estimados. Anos depois, entretanto, foi proposto o conceito de pseudo-medidas (IBRAHIM, 1987). A ideia é criar artificialmente canais de medidas adicionais para a construção das respostas dinâmicas defasadas no tempo. Dessa forma, é possível estimar várias realizações das características modais a partir dessas respostas.

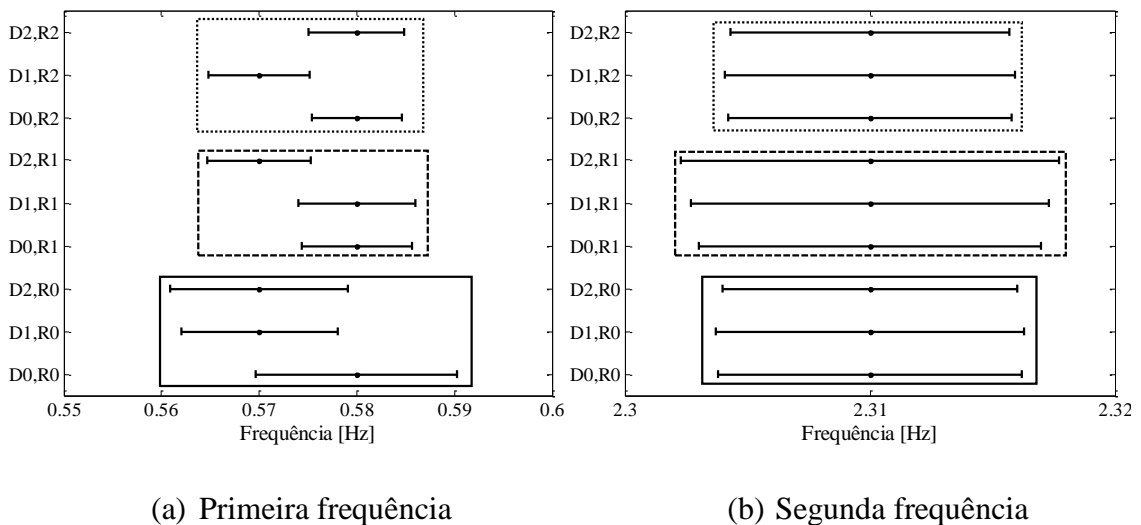
Nas simulações numéricas, 20 realizações são avaliadas para cada ensaio. Cada realização fornece um conjunto de frequências e modos próprios. A partir dessas realizações, os parâmetros modais são transformados em dados simbólicos (histogramas com 12 categorias, conforme explicado no capítulo 3). As simulações realizadas

consideram apenas as três primeiras frequências e modos próprios identificados a partir das respostas dinâmicas sem ruído (R0) e com ruído (R1 e R2).

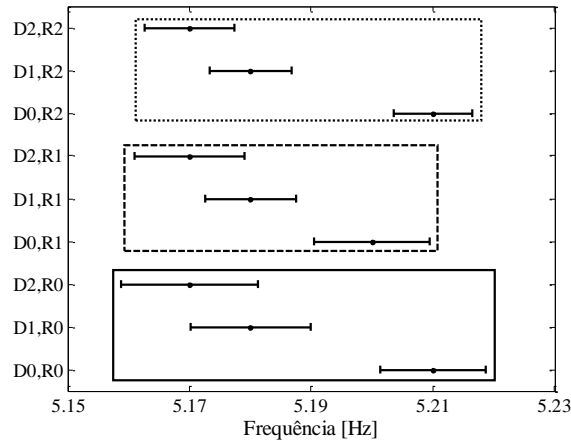
### 5.1. Análise Clássica

Antes de se passar à análise dos dados simbólicos, utilizando a transformação dos sinais e parâmetros modais em histogramas, um estudo simples baseado na variação das frequências e modos próprios é proposto. É sabido que as frequências naturais tendem a diminuir na medida em que o nível de dano na estrutura aumenta. Além disso, os valores das amplitudes modais das regiões afetadas pelo dano também tendem a diminuir (Cury, 2010).

A figura 5.3 apresenta as variações observadas para as três primeiras frequências naturais identificadas, regrupadas por nível de ruído. Nesta figura são representados os valores médios das 20 realizações, por frequência, além dos respectivos desvios-padrões. Com isso, intervalos de confiança são construídos a partir da expressão  $IC = [\bar{f} - \sigma_f; \bar{f} + \sigma_f]$ . Analisando a figura 5.3, pode-se observar que, para todas as simulações numéricas realizadas, os intervalos de confiança se superpõem. Isto é observado mesmo nos casos em que o nível de ruído é nulo. Conseqüentemente, torna-se muito difícil (até mesmo inviável) concluir que as variações das frequências naturais são devidas à presença de dano. Isto significa que este tipo de análise não é suficiente para detecção de danos na viga.







(c) Terceira frequência

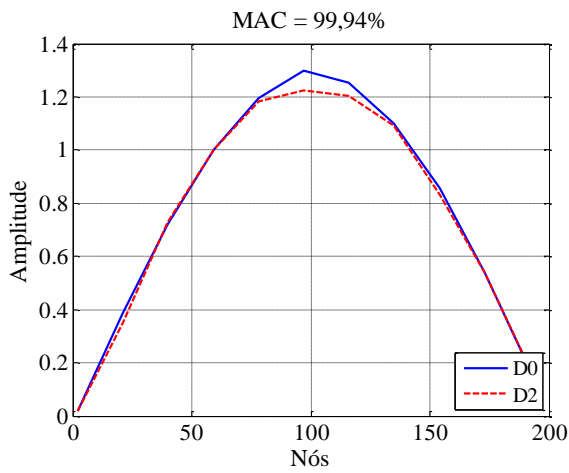
FIGURA 5.3. Comparativo das três primeiras frequências naturais (níveis de dano x níveis de ruído).

No que tange às deformadas modais, a figura 5.4 mostra o comparativo entre os valores médios das amplitudes dos três primeiros modos próprios identificados, considerando apenas os níveis de dano D0 e D2, sem ruído. É possível notar que os modos 1 e 3 são os mais afetados pela presença de dano. De qualquer forma, percebe-se que as diferenças em termos de amplitude são também muito pequenas, mesmo para níveis de dano em torno de 20%. Além disso, quando o nível de ruído de 5% é considerado, a influência do dano sobre as deformadas modais é ainda menos perceptível (figura 5.5).

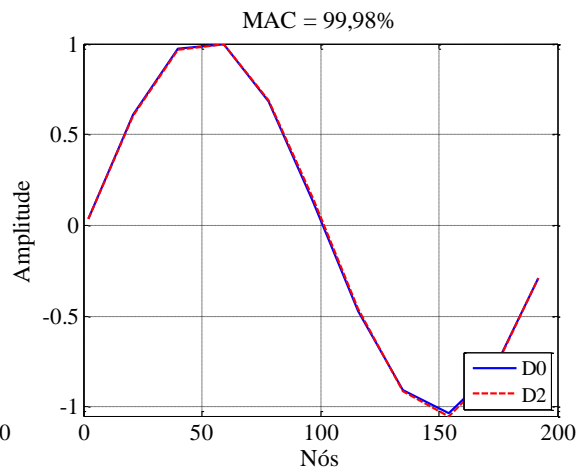
Uma maneira de se “medir” a correlação entre modos próprios é calculando o índice MAC para cada par de deformadas modais. Como citado no capítulo 2, este índice varia entre 0 e 1, onde 0 indica a ausência de correlação entre as variáveis comparadas e 1, uma correlação perfeita. A expressão que permite calcular o índice MAC para dois modos próprios é dada por (ALLEMANG; BROWN, 1982):

$$MAC_{ij} = \frac{|\sum_{k=1}^n [\phi_A]_k^i [\phi_B]_k^j|^2}{\sum_{k=1}^n ([\phi_A]_k^i)^2 \sum_{k=1}^n ([\phi_B]_k^j)^2} \quad (5.2)$$

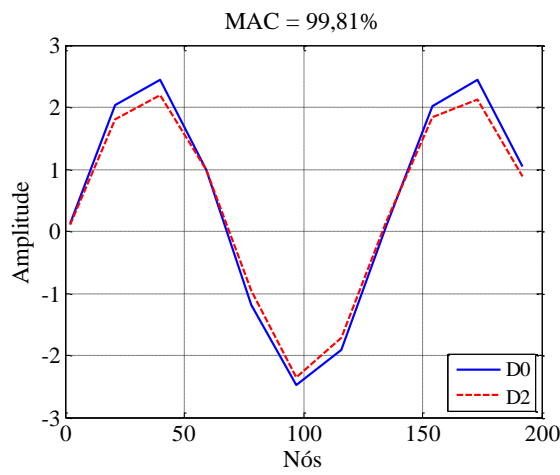
onde  $[\phi_A]_k^i$  e  $[\phi_B]_k^j$  são as k-ésimas amplitudes dos modos  $[\phi_A]^i$  e  $[\phi_B]^j$ , respectivamente.



(a) Primeiro modo

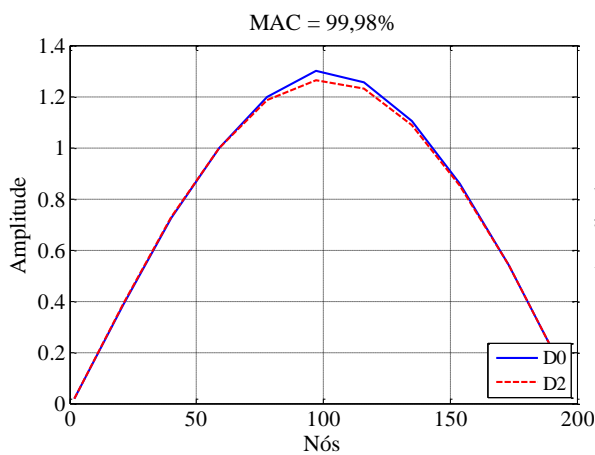


(b) Segundo modo

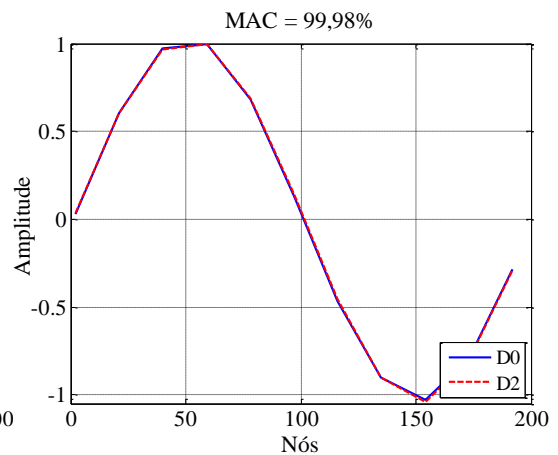


(c) Terceiro modo

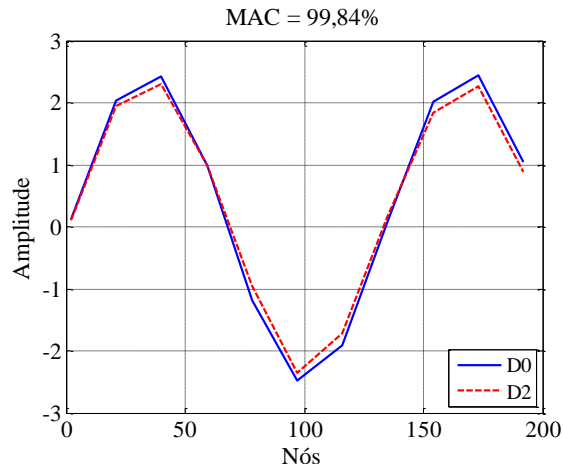
FIGURA 5.4 – Comparativo dos três primeiros modos de vibração (sem ruído).



(a) Primeiro modo



(b) Segundo modo



(c) Terceiro modo

FIGURA 5.5 – Comparativo dos três primeiros modos de vibração (ruído de 5%).

Em todos os casos, percebe-se que o índice MAC varia entre 99% e 100%, o que não permite concluir, de forma clara, se existe dano estrutural ou não.

De modo geral, as análises clássicas não são suficientemente sensíveis para discriminar os diferentes níveis de dano impostos à estrutura. Torna-se fundamental, portanto, aplicar novos tipos de análise que permitam detectar modificações estruturais, mas sempre considerando uma abordagem probabilística, isto é, levando-se em conta a variação interna de cada parâmetro modal (valor médio, desvio-padrão, mediana, etc.).

É importante ressaltar que, mesmo que os níveis de dano possuam valores considerados elevados de redução no módulo de elasticidade longitudinal (10% a 20%), isto não se traduz em diminuições abruptas nos valores das frequências naturais e amplitudes modais. Isto evidencia a natureza complexa do problema de detecção de danos em estruturas na Engenharia Civil.

## 5.2. Análise Simbólica

Os resultados obtidos a partir da utilização da ADS juntamente com os métodos de classificação são apresentados nesta seção. Os métodos são aplicados às representações simbólicas dos sinais, das frequências próprias e aos modos próprios para diferenciar os diversos estados estruturais simulados.

O objetivo da aplicação dos métodos de classificação é separar o conjunto inicial de 30 ensaios em 3 *clusters* que sejam mais homogêneos possíveis de forma que cada *cluster* contenha os ensaios de um estado estrutural diferente, ou seja, o primeiro grupo deve conter os 10 ensaios referentes a viga sem dano (D0), o segundo deve conter os 10 ensaios referentes a viga com nível de dano D1 e, o terceiro grupo, os outros 10 ensaios referentes a viga com nível de dano D2. As simulações serão realizadas para cada nível de ruído.

Dado o elevado número de parâmetros existentes na aplicação dos métodos de classificação (métricas, critérios de agregação e formas de inicialização), são apresentados somente os melhores resultados no corpo principal do texto. Os demais resultados são mostrados no anexo A.

### 5.2.1. Sinais

Primeiramente são apresentados os resultados correspondentes à aplicação do método da aglomeração hierárquica para os sinais transformados em histogramas. A figura 5.6 mostra o dendograma obtido para as simulações dos sinais sem ruído (R0) utilizando a distância *correlation* e o critério de agregação *complete*.

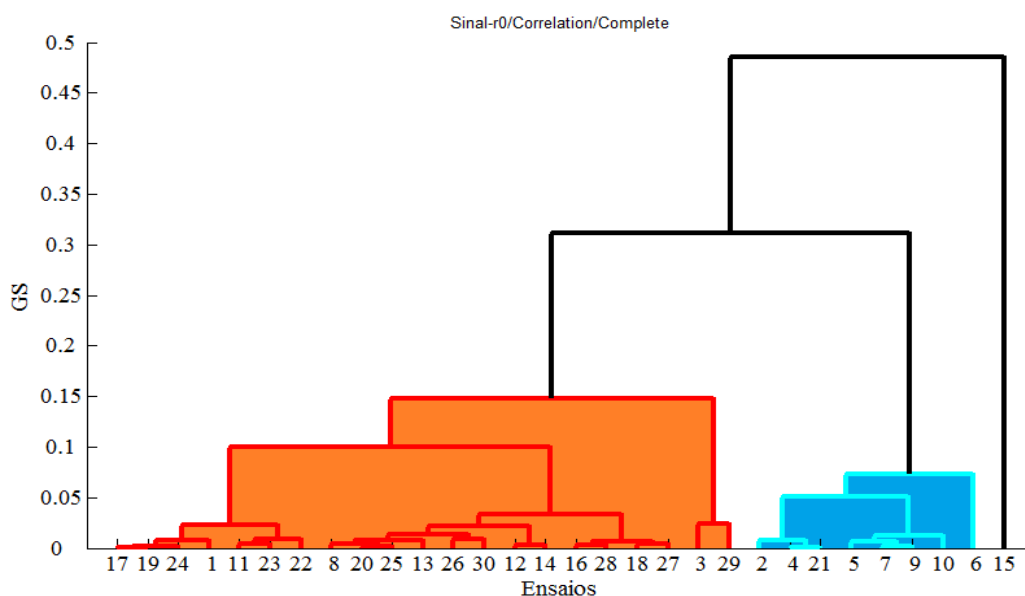


FIGURA 5.6 – Dendograma da classificação (sinais; ruído R0; dist. *correlation*; agreg. *complete*).

Os índices de acerto obtidos são relativamente satisfatórios, já que 70% dos ensaios referentes ao nível de dano D0 foram classificados corretamente, 10% dos ensaios referentes ao nível de dano D1 e 90% dos ensaios referentes ao nível D2. Apesar dos elevados índices de acerto obtidos para os níveis D0 e D2, o mesmo não ocorreu para o nível D1. Por ser um estado de dano intermediário, torna-se mais difícil sua perfeita classificação.

O mesmo procedimento de *clustering* é aplicado aos sinais com nível de ruído de 5% (R1). Na figura 5.7 é mostrada a partição obtida utilizando a distância *correlation* e o critério de agregação *complete*. Nesse caso, os índices de acerto para cada *cluster* não se altera com a presença de ruído intermediário considerado.

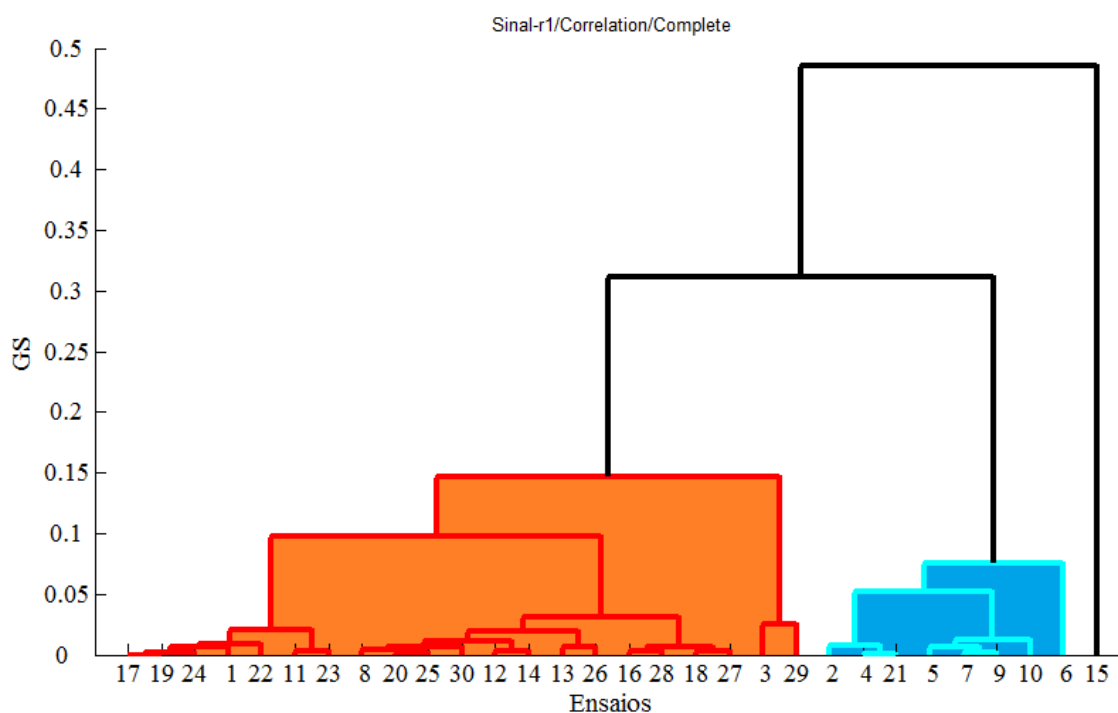


FIGURA 5.7 – Dendrograma da classificação (sinais; ruído R1; dist. *correlation*; agreg. *complete*).

A classificação para os sinais com nível de ruído de 10% (R2) é mostrada pelo dendrograma apresentado na figura 5.8. Para esse nível de ruído, os índices de acerto também não se alteraram.

Ainda que os resultados não sejam plenamente satisfatórios, é possível perceber que a abordagem adotada neste trabalho é robusta com relação à presença de ruído. Isto é, a adição de ruído aos sinais dinâmicos em nada influenciou nos resultados das classificações.

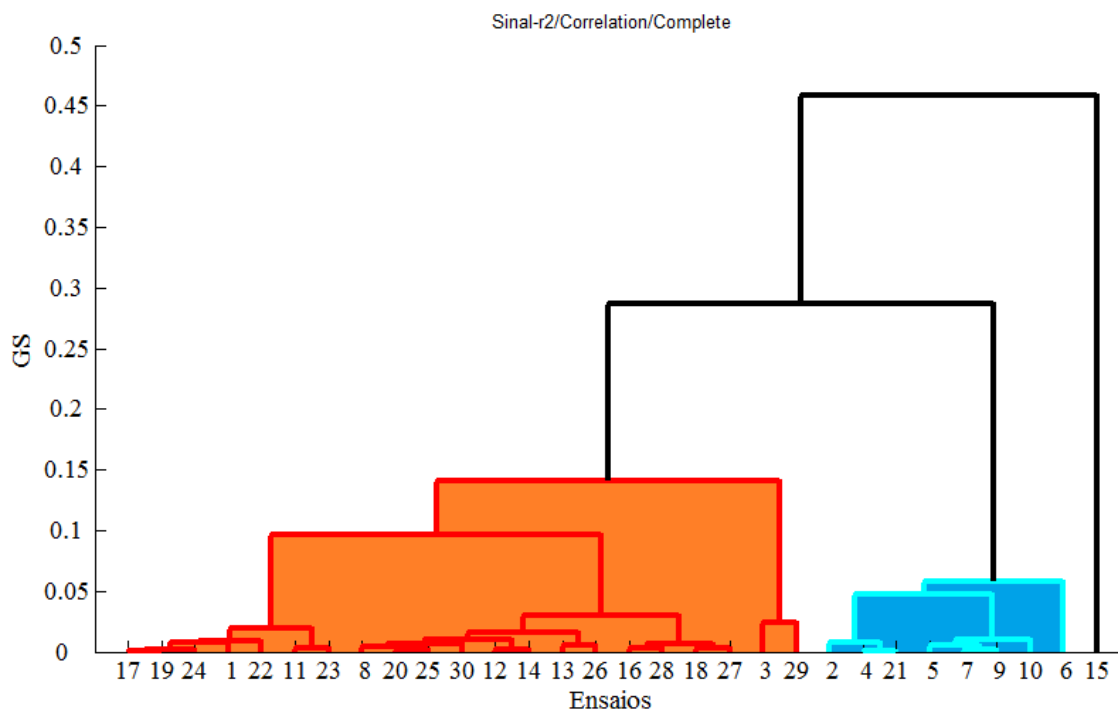


FIGURA 5.8 – Dendrograma da classificação (sinais; ruído R2; dist. *correlation*; agreg. *complete*).

Em seguida são mostrados os resultados correspondentes à aplicação do método k-means utilizando as respostas dinâmicas (sinais) transformadas em dados simbólicos. Na tabela 5.2 podem ser observados os resultados obtidos utilizando a distância euclidiana e o método de inicialização amostral, em que os centroides são inicializados de maneira aleatória. Percebe-se que os resultados são relativamente inferiores aos obtidos utilizando o método da aglomeração hierárquica.

TABELA 5.2 – Índices de acerto das classificações (k-means; sinais; dist. euclidiana; inicialização amostral).

	<b>R0</b>	<b>R1</b>	<b>R2</b>
<b>D0</b>	40%	40%	60%
<b>D1</b>	80%	80%	70%
<b>D2</b>	30%	30%	30%

Por último, o método c-means é aplicado para classificar as representações simbólicas dos sinais. Na aplicação do método c-means é utilizada a distância euclidiana, os centroides são inicializados de forma aleatória e o grau de *fuzzyficação* (o grau de “confusão” entres os *clusters* criados) considerado é igual a 2. Na tabela 5.3 encontram-se os índices de acerto da classificação, na qual é possível observar, para cada nível de ruído, a porcentagem de ensaios classificados corretamente para cada configuração de dano e o grau de certeza da classificação, encontrados entre parênteses.

TABELA 5.3 - Índices de acerto das classificações obtidas utilizando o método c-means para os sinais descritos de forma simbólica.

	<b>R0</b>	<b>R1</b>	<b>R2</b>
<b>D0</b>	40% (47%)	40% (47%)	50% (49%)
<b>D1</b>	70% (65%)	70% (66%)	70% (65%)
<b>D2</b>	30% (59%)	40% (59%)	30% (58%)

Os índices de acerto obtidos utilizando o c-means aplicado aos sinais não foram satisfatórios. Observando os graus de certeza da classificação, é possível notar que o método se confunde e acaba determinando que um resultado insatisfatório é bom. Isto pode ser observado na última linha da tabela 5.3. Para baixos índices de classificação correta (30 e 40%) o grau de certeza é alto (58-59%).

Na tabela 5.4 é apresentado um resumo dos índices de acerto obtidos para os três métodos de classificação utilizando os sinais como dados de entrada.

TABELA 5.4 – Resumo dos índices de acerto obtidos com todos os métodos de classificação utilizando os sinais.

	Aglomeración Hierárquica			k-means			c-means		
	R0	R1	R2	R0	R1	R2	R0	R1	R2
D0	70%	70%	70%	40%	40%	60%	40% (47%)	40% (47%)	50% (49%)
D1	10%	10%	10%	80%	80%	70%	70% (65%)	70% (66%)	70% (65%)
D2	90%	90%	90%	30%	30%	30%	30% (59%)	40% (59%)	30% (58%)

Em geral, o método das aglomerações hierárquicas fornece os melhores resultados. Todavia, percebe-se a dificuldade de se classificar corretamente os ensaios relativos ao nível de dano intermediário (D1). É interessante perceber que, de modo geral, todos os métodos são pouco sensíveis à presença de ruído, já que os resultados são similares para os níveis R0, R1 e R2.

### 5.2.2. Frequências

A utilização da ADS juntamente com as técnicas de *clustering* também é utilizada para o estudo das frequências próprias. Esse estudo consiste na transformação das três primeiras frequências próprias em dados simbólicos e na sua utilização como dados de entrada para os métodos de *clustering*. Aplicando o método da aglomeração hierárquica, utilizando a distância *correlation* e o critério de agregação *ward*, os índices de acerto foram de 100% para todas as configurações de dano, considerando todos os níveis de ruído. Esse resultado mostra que as frequências próprias transformadas em dados simbólicos podem ser utilizadas como indicadores confiáveis para detectar modificações estruturais. A figura 5.9 mostra o dendograma obtido através da aplicação do método de aglomeração hierárquica às frequências naturais considerando o nível máximo de ruído de 10% (R2).



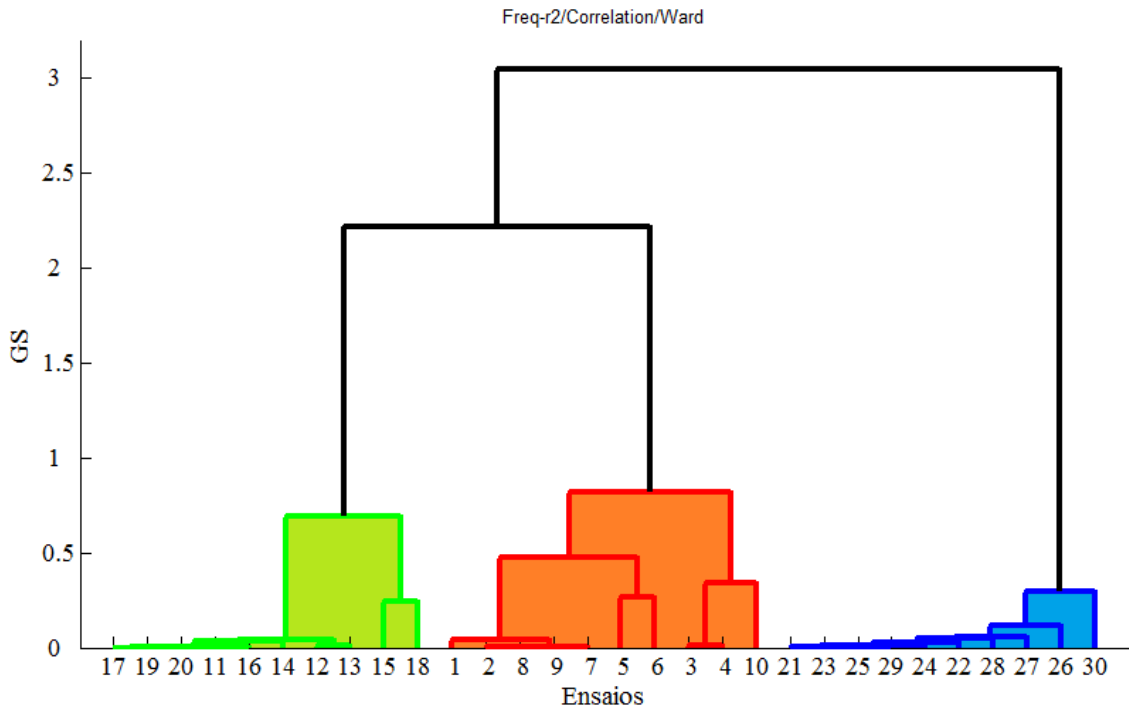


FIGURA 5.9 – Dendrograma da classificação (frequências; ruído R2; dist. *correlation*; agreg. *ward*).

Os resultados obtidos com o k-means são apresentados na tabela 5.5, no qual é possível observar que os índices de acerto também foram de 100% na classificação para os três níveis de ruído. A distância utilizada para a classificação é a euclidiana e o método de inicialização dos centroides utilizado é o uniforme.

TABELA 5.5 - Índices de acerto das classificações (k-means; frequências; dist. euclidiana; inicialização uniforme).

	R0	R1	R2
<b>D0</b>	100%	100%	100%
<b>D1</b>	100%	100%	100%
<b>D2</b>	100%	100%	100%

Em seguida, aplica-se o método de classificação c-means às frequências próprias. Para essa simulação, utiliza-se a distância euclidiana e o método de inicialização amostral. Os índices de acerto também foram de 100% e os graus de certeza de classificação

obtidos foram elevados, o que indica que a frequência própria é um parâmetro bastante confiável para ser utilizado na identificação de danos (tabela 5.6).

TABELA 5.6 - Índices de acerto das classificações obtidas utilizando o método c-means para as frequências descritas de forma simbólica.

	<b>R0</b>	<b>R1</b>	<b>R2</b>
<b>D0</b>	100% (85%)	100% (85%)	100% (78%)
<b>D1</b>	100% (87%)	100% (81%)	100% (86%)
<b>D2</b>	100% (82%)	100% (85%)	100% (93%)

A tabela 5.7 apresenta um resumo dos índices de acerto obtidos utilizando todos os métodos de classificação citados utilizando as frequências naturais como dados de entrada.

TABELA 5.7 – Resumo dos índices de acerto obtidos com todos os métodos de classificação utilizando as frequências.

	<b>Aglomerção Hierárquica</b>			<b>k-means</b>			<b>c-means</b>		
	<b>R0</b>	<b>R1</b>	<b>R2</b>	<b>R0</b>	<b>R1</b>	<b>R2</b>	<b>R0</b>	<b>R1</b>	<b>R2</b>
<b>D0</b>	100%	100%	100%	100%	100%	100%	100%(85%)	100%(85%)	100%(78%)
<b>D1</b>	100%	100%	100%	100%	100%	100%	100%(87%)	100%(81%)	100%(86%)
<b>D2</b>	100%	100%	100%	100%	100%	100%	100%(82%)	100%(85%)	100%(93%)

Como dito anteriormente, o objetivo é utilizar os parâmetros modais transformados em dados simbólicos para, através deles, aplicar os métodos de classificação e diferenciar os estados estruturais. Como pode ser observado na tabela 5.7, com a utilização das frequências próprias, os métodos classificaram todos os ensaios corretamente, o que indica que a frequência própria pode ser utilizada para a diferenciação de estados estruturais.

### 5.2.3. Modos próprios

Finalmente, os modos próprios transformados em dados simbólicos são utilizados como dados de entrada para a aplicação dos métodos de *clustering*. Diferentemente das

frequências, os três modos próprios são transformados em dados simbólicos separadamente. Estudos anteriores mostraram que a análise individual de cada modo próprio apresenta melhores resultados (CURY, 2010). Por esse motivo, os resultados serão apresentados separadamente, modo a modo.

### **5.2.3.1. Primeiro modo próprio**

O dendograma obtido através do método da aglomeração hierárquica aplicado aos primeiros modos próprios utilizando a distância euclidiana e o critério de agregação *ward* é mostrado nas figuras 5.10 a,b,c para os três níveis de ruído (R0, R1 e R2) respectivamente.

Os índices de acerto obtidos com a utilização dos primeiros modos próprios sem ruído (R0) foram satisfatórios, já que foram classificados corretamente 80% dos ensaios referentes ao *cluster* D0, 90% ao *cluster* D1 e 100% ao *cluster* D2. Para a classificação com o nível de ruído de 5% (R1) os resultados não foram afetados de maneira significativa visto que os índices de acerto foram de 80% para a configuração de dano D0, 80% para a configuração D1, e 100% para a D2. Apesar de a classificação para os dois níveis de danos não ter sido alterada significativamente, observando-se os dendogramas das respectivas classificações, é possível notar que a altura dos *clusters* referente ao nível de dano R1 diminuiu em relação ao dendograma referente ao nível R0. A altura do dendograma indica a similaridade entre os *clusters* e, com isso, é possível observar que a presença de ruído aumentou a similaridade entre os *clusters*, isto é, tornou menos óbvia a diferença entre os estados estruturais.

Os índices de acerto na classificação utilizando os primeiros modos próprios com nível de ruído de 10% (R2), não foram tão satisfatórios. Os índices de acerto foram de 80% para a configuração D0, 40% para a configuração D1 e 60% para a D2.

Percebe-se, portanto, que os modos de vibração são mais sensíveis à presença de ruído se comparados às frequências naturais.

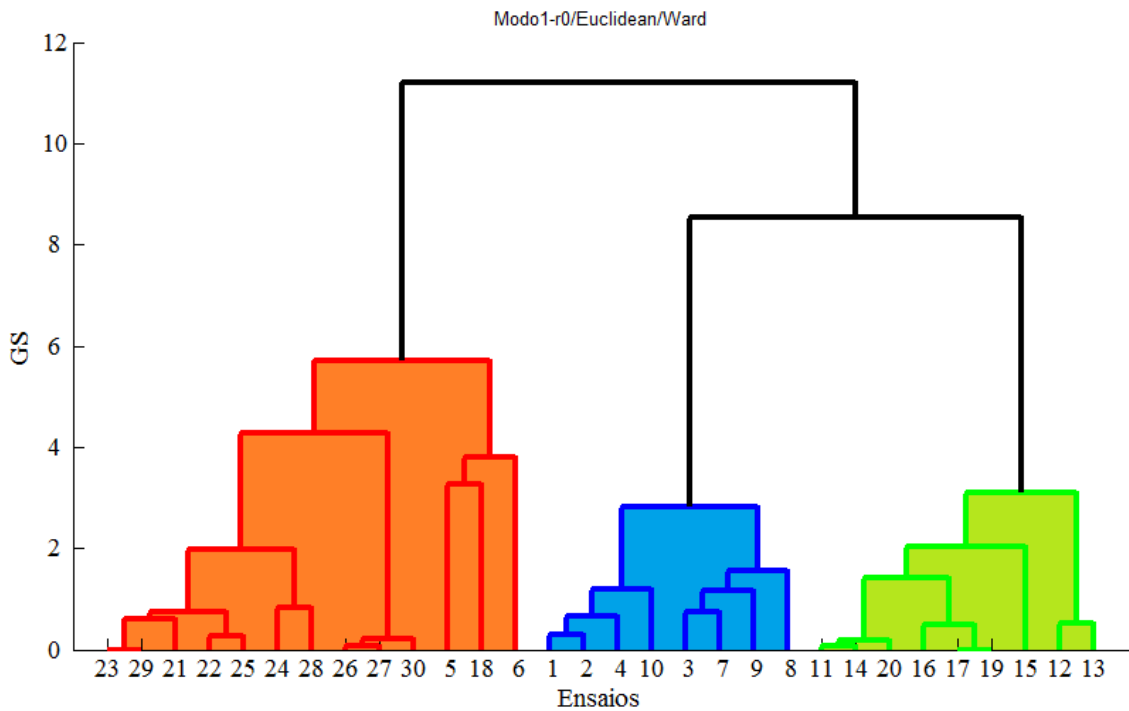


FIGURA 5.10a – Dendrograma da classificação (primeiro modo; ruído R0; dist. euclidiana; agreg. *ward*).

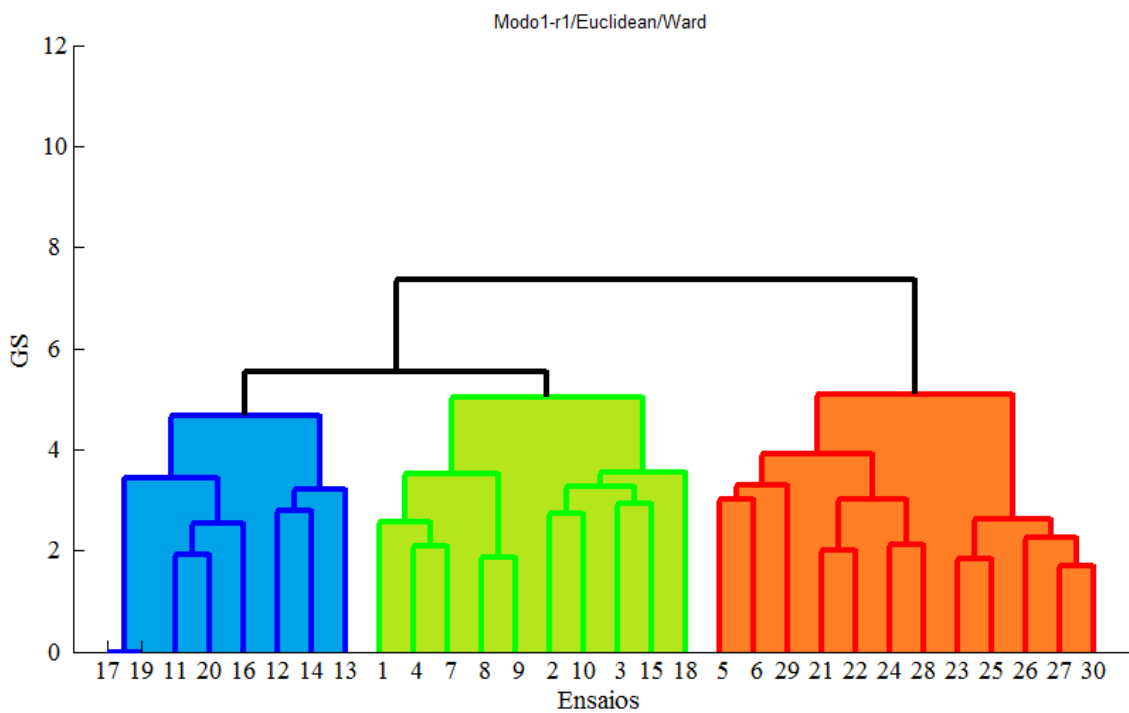


FIGURA 5.10b – Dendrograma da classificação (primeiro modo; ruído R1; dist. euclidiana; agreg. *ward*).

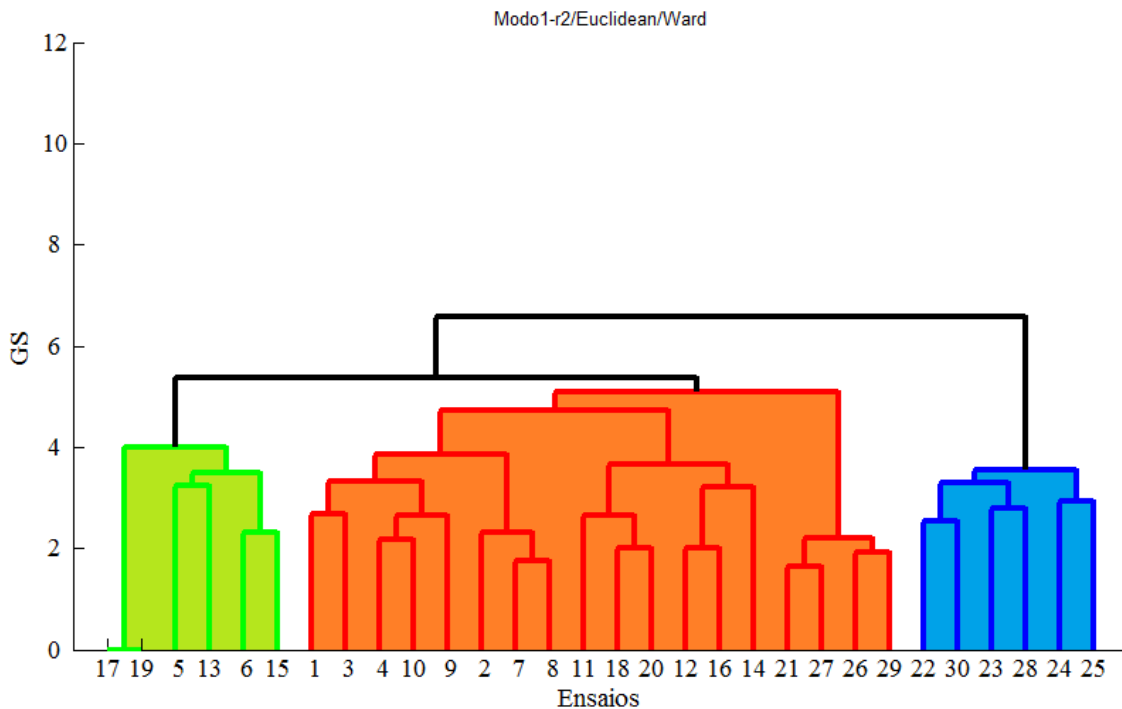


FIGURA 5.10c – Dendrograma da classificação (primeiro modo; ruído R2; dist. euclidiana; agreg. *ward*).

Aplicando-se o método de classificação k-means com a distância *correlation* e o método de inicialização uniforme, foram obtidos índices satisfatórios de acerto para a classificação dos ensaios. O método mostrou-se um pouco sensível ao ruído durante a classificação com nível de ruído de 10% (R2), porém os índices de acerto ainda continuaram elevados como mostrado na tabela 5.8.

TABELA 5.8- Índices de acerto das classificações (k-means; primeiro modo; dist. *correlation*; inicialização uniforme).

	R0	R1	R2
<b>D0</b>	90%	90%	90%
<b>D1</b>	100%	100%	100%
<b>D2</b>	100%	100%	80%

A classificação obtida aplicando o método de classificação c-means é apresentada na tabela 5.9. Como é possível observar, os índices de acerto obtidos são muito bons e também retornam adequados graus de certeza na classificação de cada configuração de

dano. Para o nível de dano de 5% (R1) os índices de acerto continuam elevados, porém o ruído introduziu dúvida à classificação. Já para o nível de ruído de 10% (R2), os índices de acerto da classificação diminuíram e o grau de certeza permaneceram baixos, o que significa que mesmo os ensaios classificados corretamente possuem características muito semelhantes aos ensaios dos outros *clusters*.

TABELA 5.9 - Índices de acerto das classificações obtidas utilizando o método c-means para os primeiros modos próprios descritos de forma simbólica.

	<b>R0</b>	<b>R1</b>	<b>R2</b>
<b>D0</b>	90%(75%)	90%(33%)	70%(33%)
<b>D1</b>	90%(82%)	90%(33%)	0%(33%)
<b>D2</b>	100%(85%)	100%(33%)	60%(33%)

A tabela 5.10 resume os índices de acerto encontrados para todos os métodos de classificação aplicados aos primeiros modos próprios.

TABELA 5.10 - Resumo dos índices de acerto obtidos com todos os métodos de classificação utilizando os primeiros modos próprios.

	<b>Aglomeraco Hierrquica</b>			<b>k-means</b>			<b>c-means</b>		
	<b>R0</b>	<b>R1</b>	<b>R2</b>	<b>R0</b>	<b>R1</b>	<b>R2</b>	<b>R0</b>	<b>R1</b>	<b>R2</b>
<b>D0</b>	80%	80%	80%	90%	90%	90%	90%(75%)	90%(33%)	70%(33%)
<b>D1</b>	90%	80%	40%	100%	100%	100%	90%(82%)	90%(33%)	0%(33%)
<b>D2</b>	100%	100%	60%	100%	100%	80%	100%(85%)	100%(33%)	60%(33%)

### 5.2.3.2. Segundo modo prprio

Utilizando os segundos modos prpios transformados em dados simblicos como dados de entrada para a aplicao do mtodo da aglomerao hierrquica e utilizando a distncia *spearman* e o critrio de agregao *complete*, foram obtidos os dendogramas mostrados nas figuras 5.11 a,b,c referentes aos nveis de rudo R0, R1 e R2, respectivamente.

Os resultados obtidos foram relativamente satisfatrios. Os índices de acerto para a classificao dos ensaios com nvel de rudo R0 foram de 100%, 0% e 90% para as

configurações de dano D0, D1 e D2 respectivamente, 90%, 40% e 90% para o nível de ruído R1 e 90%, 10%, 90% para o nível de ruído R2.

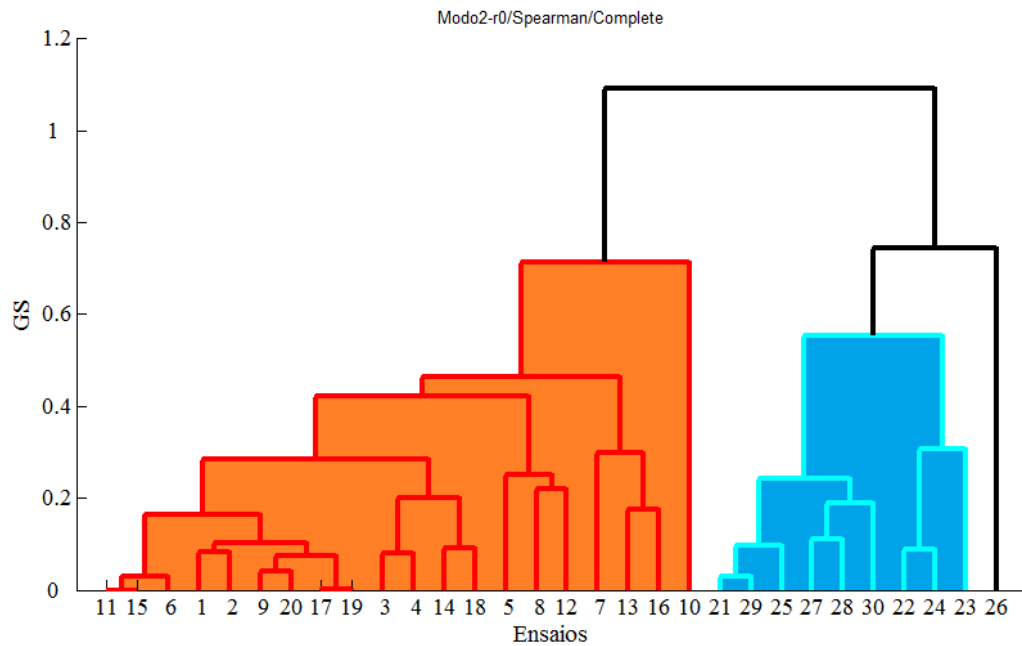


FIGURA 5.11a – Dendrograma da classificação (segundo modo; ruído R0; dist. *spearman*; agreg. *complete*).

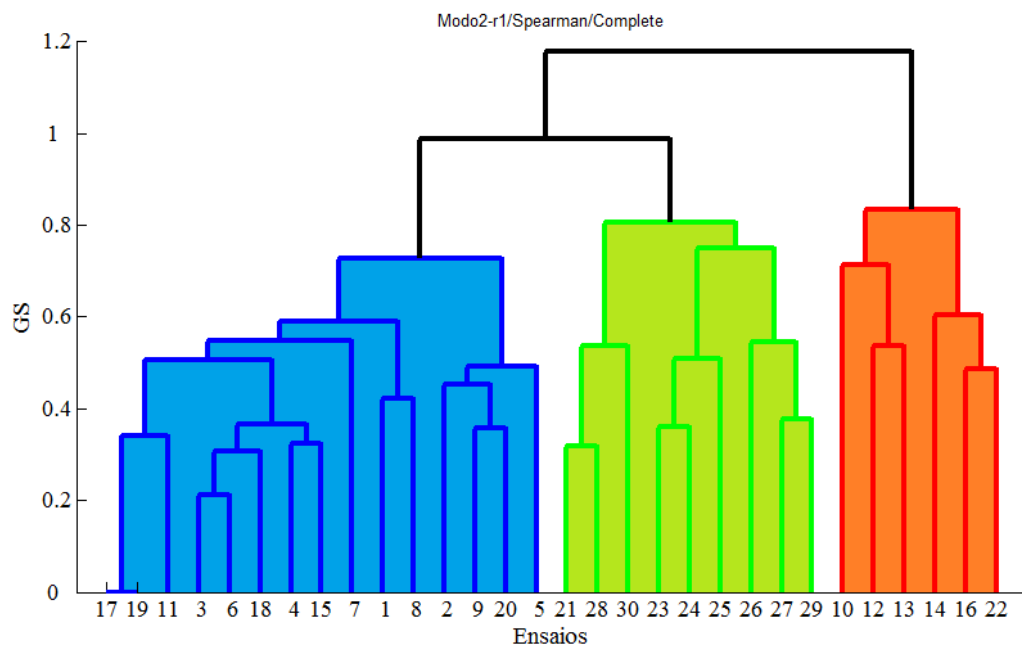


FIGURA 5.11b – Dendrograma da classificação (segundo modo; ruído R1; dist. *spearman*; agreg. *complete*).

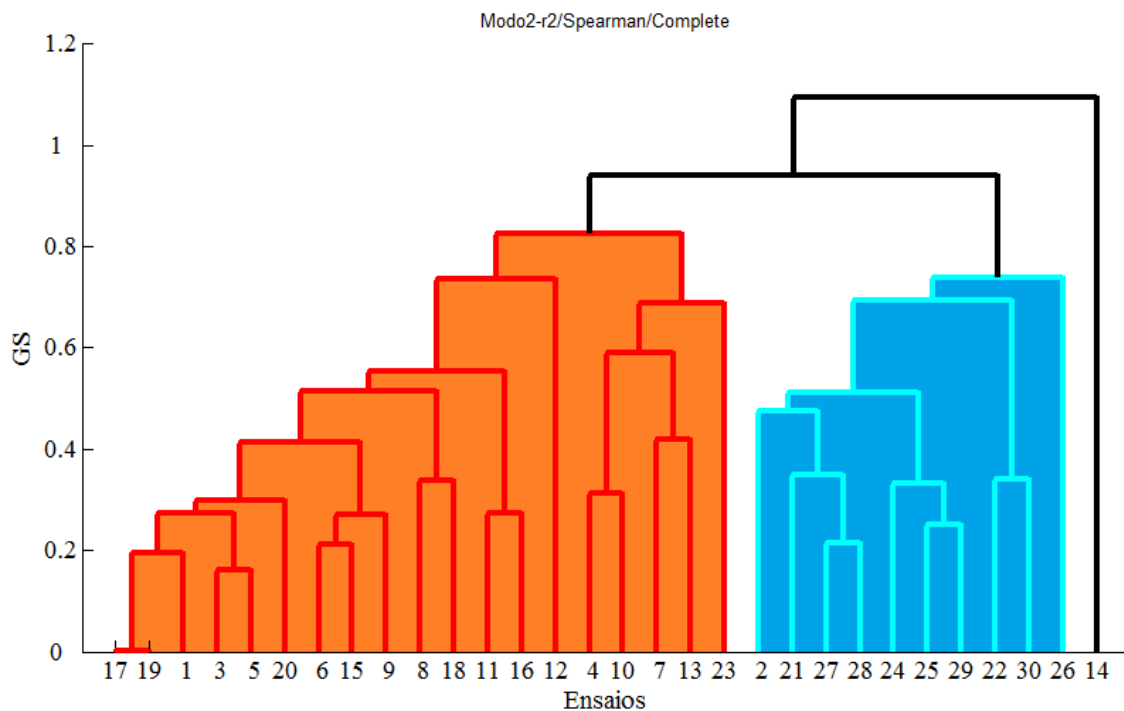


FIGURA 5.11c – Dendrograma da classificação (segundo modo; ruído R2; dist. *spearman*; agreg. *complete*).

Aplicando-se o método k-means, os melhores índices de acerto foram obtidos com a distância euclidiana e método de inicialização uniforme. Os índices de acerto para a primeira e terceira configurações de dano são razoáveis, porém os índices de acerto para as classificações dos danos intermediários não são tão bons, como pode ser observado na tabela 5.11.

TABELA 5.11- Índices de acerto das classificações (k-means; segundo modo; dist. euclidiana; inicialização uniforme).

	R0	R1	R2
D0	70%	60%	70%
D1	30%	50%	60%
D2	100%	100%	80%

As classificações obtidas pelo método c-means são apresentadas na tabela 5.12. É possível observar que a utilização dos segundos modos próprios como parâmetro para a



classificação não é confiável. Além de os índices de acerto não serem elevados, os graus de certeza fornecidos pelo c-means são baixos mesmo para a classificação sem ruído.

TABELA 5.12 - Índices de acerto das classificações obtidas utilizando o método c-means para os segundos modos próprios descritos de forma simbólica.

	<b>R0</b>	<b>R1</b>	<b>R2</b>
<b>D0</b>	10% (11%)	80% (33%)	10% (33%)
<b>D1</b>	100% (80%)	10% (33%)	80% (33%)
<b>D2</b>	70% (58%)	100% (33%)	100% (33%)

Um resumo dos índices de acerto quando utilizado os segundos modos próprios é mostrado na tabela 5.13.

TABELA 5.13 - Resumo dos índices de acerto obtidos com todos os métodos de classificação utilizando os segundos modos próprios.

	Aglomerção Hierárquica			k-means			c-means		
	<b>R0</b>	<b>R1</b>	<b>R2</b>	<b>R0</b>	<b>R1</b>	<b>R2</b>	<b>R0</b>	<b>R1</b>	<b>R2</b>
<b>D0</b>	100%	90%	90%	70%	60%	70%	10%(11%)	80%(33%)	10%(33%)
<b>D1</b>	0	40%	10%	30%	50%	60%	100%(80%)	10%(33%)	80%(33%)
<b>D2</b>	90%	90%	90%	100%	100%	80%	70%(58%)	100%(33%)	100%(33%)

### 5.2.3.3. Terceiro modo próprio

Os dendogramas apresentados nas figuras 5.12 a,b,c mostram as classificações para os terceiros modos próprios, utilizando a distância *correlation* e o critério de agregação *ward*. Os índices de acerto para o nível de ruído R0 foram de 100% para todas as configurações de dano. Para o nível de ruído R1, os índices de acerto foram de 100%, 80% e 100%, para as configurações de dano D0, D1 e D2, respectivamente e, para o nível de ruído R2, os índices foram de 100%, 80% e 40%.

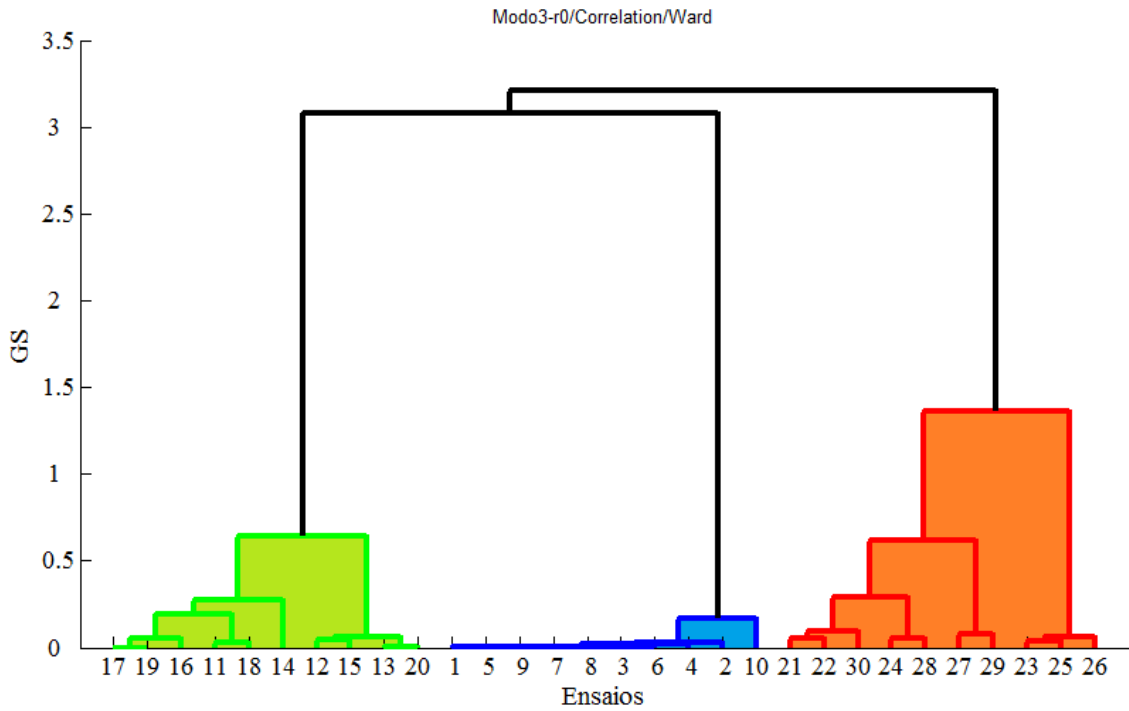


FIGURA 5.12a – Dendrograma da classificação (terceiro modo; ruído R0; dist. *correlation*; agreg. *ward*).

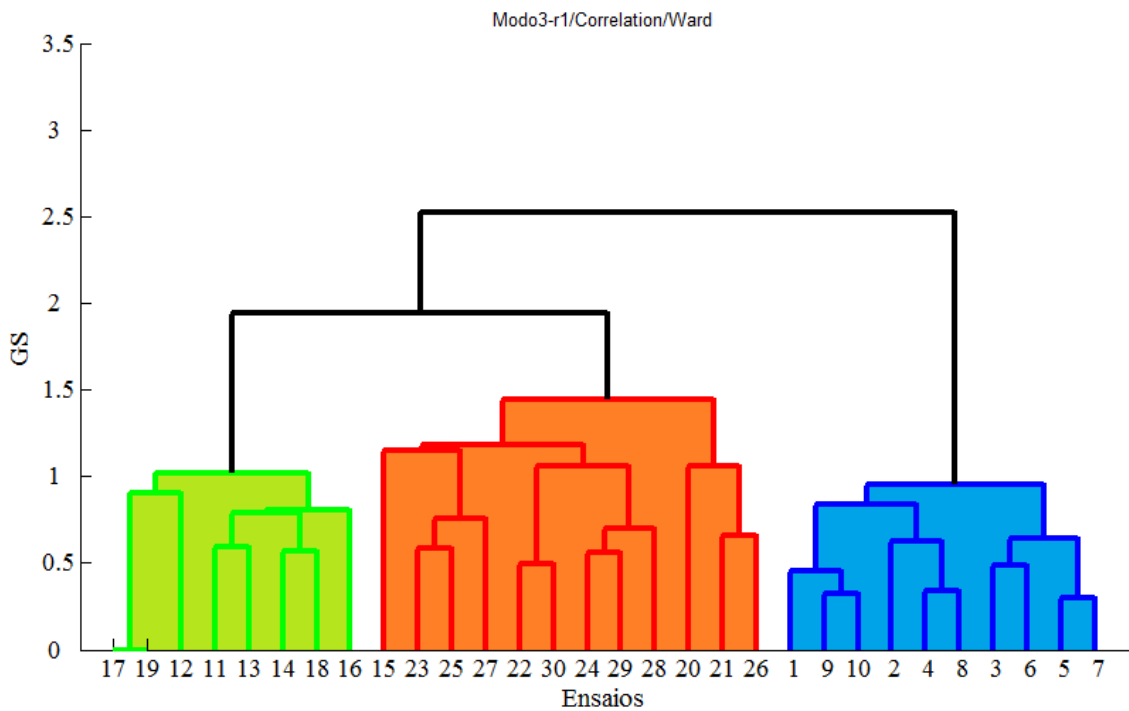


FIGURA 5.12b – Dendrograma da classificação (terceiro modo; ruído R1; dist. *correlation*; agreg. *ward*).

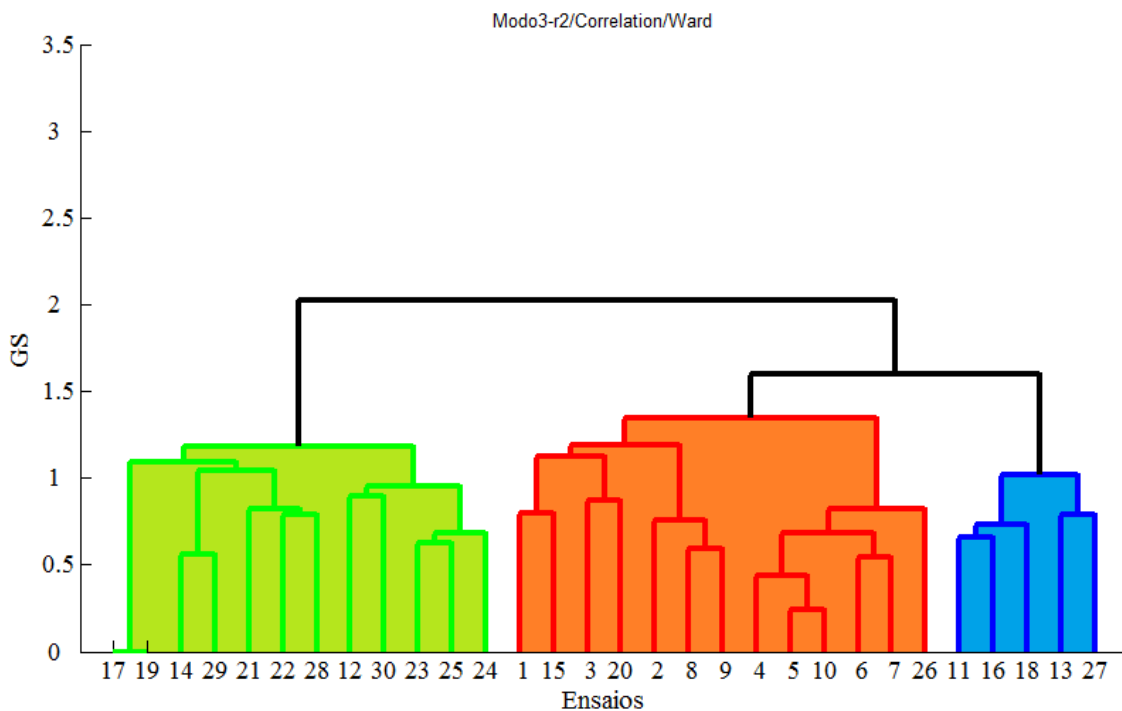


FIGURA 5.12c – Dendograma da classificação (terceiro modo; ruído R2; dist. *correlation*; agreg. *ward*).

Em seguida, o método k-means é considerado utilizando a distância euclidiana e o método de inicialização uniforme. Os índices de acerto são mostrados na tabela 5.14. Os índices obtidos são bons, porém a utilização dos terceiros modos mostrou maior sensibilidade para o nível de ruído de 10% (R2).

TABELA 5.14- Índices de acerto das classificações (k-means; terceiro modo; dist. euclidiana; inicialização uniforme).

	<b>R0</b>	<b>R1</b>	<b>R2</b>
<b>D0</b>	100%	100%	90%
<b>D1</b>	100%	100%	80%
<b>D2</b>	100%	90%	40%

Finalmente, são apresentados na tabela 5.15 os índices de acerto para as classificações utilizando o c-means. Os índices de acerto para as classificações sem ruído foram muito bons, com bons graus de certeza das classificações. Porém, quando se adiciona ruído aos dados, as classificações não se mostraram tão eficientes.

TABELA 5.15 - Índices de acerto das classificações obtidas utilizando o método c-means para os terceiros modos próprios descritos de forma simbólica.

	R0	R1	R2
D0	100%(97%)	100% (33%)	100% (33%)
D1	100% (84%)	0%(33%)	0%(33%)
D2	100% (75%)	100% (33%)	100% (33%)

Um resumo dos índices de acerto é mostrado na tabela 5.16, na qual é possível observar que a utilização dos terceiros modos próprios é mais confiável quando aplicado ao método da aglomeração hierárquica e ao método k-means. Todas as simulações apresentaram uma diminuição nos índices de acerto na presença de ruído (R2), porém o c-means apresentou diminuição dos índices de acerto já mesmo para o nível de ruído (R1).

TABELA 5.16 - Resumo dos índices de acerto obtidos com todos os métodos de classificação utilizando os terceiros modos próprios.

	Aglomeração Hierárquica			k-means			c-means		
	R0	R1	R2	R0	R1	R2	R0	R1	R2
D0	100%	100%	100%	100%	100%	90%	100%(97%)	100%(33%)	100%(33%)
D1	100%	80%	80%	100%	100%	80%	100%(84%)	0%(33%)	0%(33%)
D2	100%	100%	40%	100%	90%	40%	100%(75%)	100%(33%)	100%(33%)

### 5.3. Estudo complementar

Um estudo complementar é realizado para analisar a eficiência dos métodos de classificação. Para esse estudo, são considerados apenas os ensaios referentes à estrutura intacta (D0) e os ensaios com maior nível de dano (D2). Os dados utilizados foram os simulados com nível de ruído de 10% (R2). As distâncias, critérios de agregação e métodos de inicialização utilizados foram os mesmos utilizados para cada método de *clustering* nas aplicações considerando 3 *clusters*. Com essa abordagem e considerando-se agora apenas 2 *clusters*, é possível observar na tabela 5.17 que todos os métodos fornecem elevados índices de acerto.

TABELA 5.17 – Índices de acerto utilizando os métodos de classificação para os estados D0 e D2.

Aglomeración Hierárquica					
	Sinais	Frequências	Modo 1	Modo 2	Modo 3
D0	90%	100%	90%	90%	100%
D2	70%	100%	60%	90%	80%
k-means					
	Sinais	Frequências	Modo 1	Modo 2	Modo 3
D0	70%	100%	80%	90%	100%
D2	90%	100%	60%	100%	90%
c-means					
	Sinais	Frequências	Modo 1	Modo 2	Modo 3
D0	70% (67%)	100% (89%)	80% (50%)	90% (50%)	90% (50%)
D2	90% (80%)	100% (96%)	100% (50%)	100% (50%)	90% (50%)

É importante ressaltar que a aplicação dos métodos de classificação para diferenciar apenas dois estados estruturais (D0 e D2), fornece melhores resultados se comparada para diferenciação de três estados estruturais. Considerando apenas duas configurações de dano, os índices de acerto são todos superiores a 60%, enquanto que considerando três configurações de danos, os índices de acerto de certas classificações podem ser nulos.

Neste capítulo, foram realizadas diversas simulações numéricas, levando-se em consideração diversos conjuntos de testes dinâmicos. Os resultados mostraram que, em geral, a abordagem proposta é capaz de detectar danos estruturais. Além disso, observou-se que os métodos de *clustering* foram mais eficazes para a discriminação de diferentes cenários de danos, principalmente se as frequências naturais são usadas. Finalmente, pode-se observar que os sinais brutos também podem ser utilizados como parâmetros para detecção de comportamentos estruturais. Em alguns casos, os índices de classificação correta superaram 70%, o que representa um resultado expressivo. Todavia, foi também observado que, em certas circunstâncias, esses valores podem ser mais baixos, devido à presença de ruído nos sinais.

# CAPÍTULO 6

## 6. ENSAIOS EXPERIMENTAIS

Após as aplicações dos métodos de classificação em um modelo numérico de uma viga, um estudo experimental de um viaduto ferroviário é apresentado. O viaduto em questão é o PK 075+317 que liga Paris à Lyon. Inicialmente, um procedimento de identificação modal é realizado. Em seguida são realizadas análises clássicas a partir dos parâmetros modais identificados. Por último, a ADS é utilizada juntamente com os métodos de classificação para discriminar os diferentes estados estruturais de tal viaduto.

### 6.1. Ensaios experimentais realizados no viaduto ferroviário PK 075+317

O monitoramento dinâmico deste viaduto foi realizado com o objetivo de caracterizar e quantificar a eficiência de um procedimento de reforço estrutural nele realizado. O procedimento de reforço consistiu em "apertar" os mancais de sustentação do viaduto a fim de distanciar a primeira frequência de vibração da estrutura (em torno de 4,5 Hz) da frequência de excitação causada pela passagem do trem. O viaduto PK 075+317 está localizado no sudeste da França, perto das cidades de Sens e Soucy no condado de Yonne (figuras 6.1 e 6.2). Por este viaduto passam, diariamente, diversos trens de alta velocidade (TGVs) ligando Paris à Lyon.



FIGURA 6.1 – Localização do viaduto ferroviário PK 075+317. FONTE – CURY, 2010.



FIGURA 6.2 – Vista geral do viaduto. FONTE – CURY, 2010.

A estrutura possui 17,5 m de comprimento e é constituída de duas vigas metálicas revestidas de concreto (estrutura mista). A instrumentação e as medições são realizadas apenas no trilho 1, por onde circulam os trens que viajam no sentido Paris até Lyon. O trilho 2 não é instrumentado.

A instrumentação do viaduto foi realizada entre os dias 24 e 26 de junho de 2003 e foi composta pelos seguintes sensores:

- 3 sensores de deslocamento vertical localizados entre o parapeito e o trilho: cada sensor é montado em um suporte fixo no parapeito e mede o deslocamento relativo entre o parapeito supostamente fixo e o trilho.
- 8 acelerômetros verticais e 2 acelerômetros horizontais (longitudinal e lateral) sob o trilho: cada acelerômetro é montado em um suporte fixado sob o trilho no nível das vigas.
- 2 sensores de passagem de trens: cada sensor mede o carregamento no eixo na entrada e na saída da estrutura.

Os suportes são fixados mecanicamente por diferentes montagens que permitem o ajuste da orientação de cada sensor. Os sensores de deslocamento e os acelerômetros são conectados através de cabos sob o pavimento (figura 6.3) até um sistema de aquisição

de dados fixado em uma parede frontal a 3 m do solo, onde são instalados os canais de aquisição (figura 6.4).

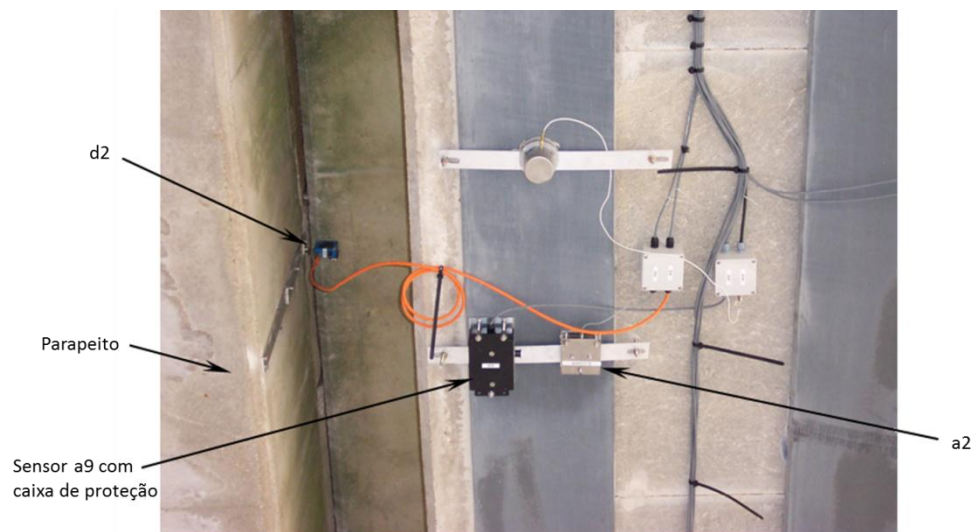


FIGURA 6.3 – Vista dos sensores d2, a2 e a9. FONTE – CURY, 2010.



FIGURA 6.4 – Vista do sistema de aquisição de dados e dos cabos. FONTE – CURY, 2010.

A figura 6.5 apresenta um esquema da localização dos sensores sob o pavimento e a numeração adotada para a determinação dos modos próprios experimentais. A fonte de



excitação do pavimento dos trilhos 1 é a passagem dos TGVs sobre os trilhos. Os registros são feitos continuamente durante um período de 3,5 s.

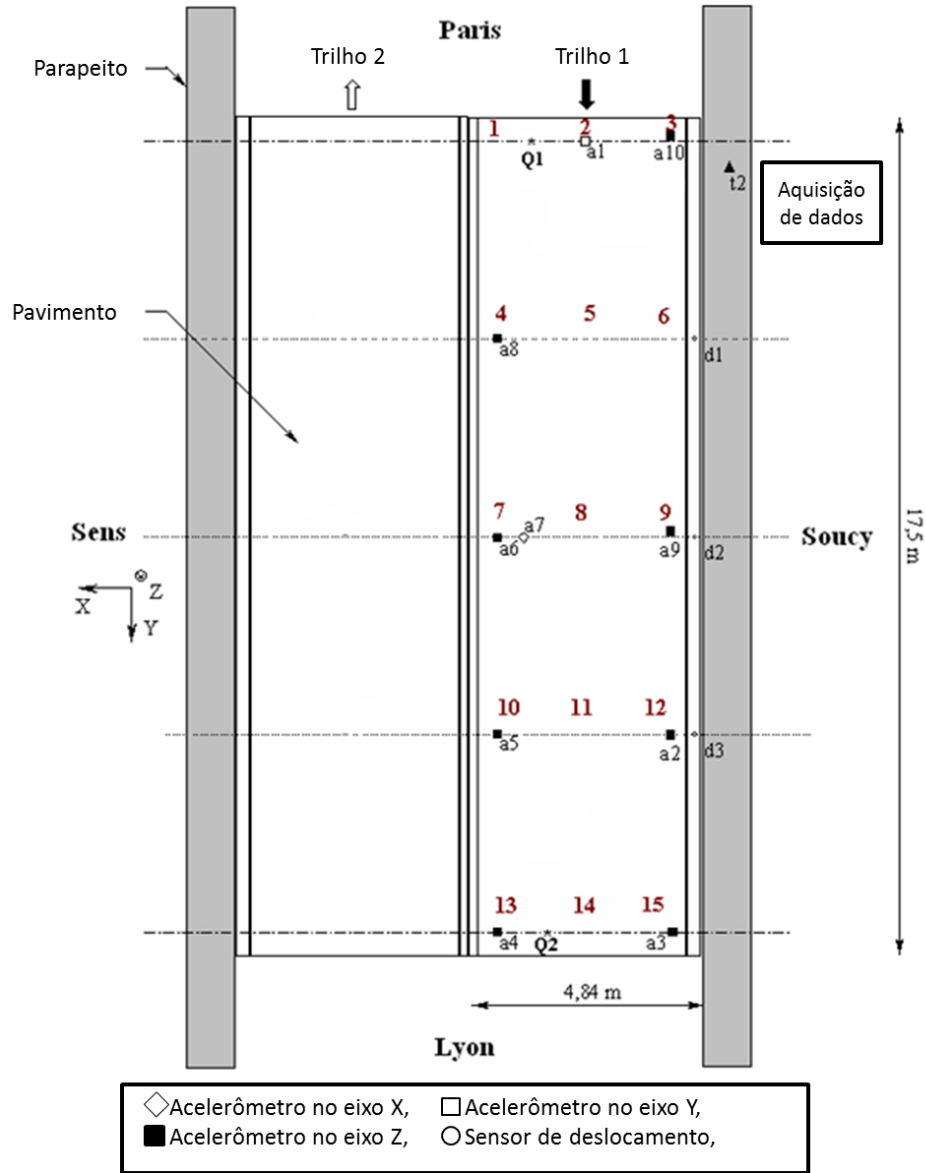


FIGURA 6.5 – Disposição dos sensores sob o pavimento. FONTE – CURY, 2010.

Os sensores de passagem Q1 e Q2 permitem determinar a frequência de excitação relacionada à passagem do TGV. De fato, a diferença entre o instante do primeiro pico da resposta dos sensores Q1 e Q2 permite obter a velocidade média da passagem do TGV. Por exemplo, a diferença de tempo entre os primeiros picos é de 0,223 s, o que

fornece uma velocidade média de passagem de  $17,5/0,223 = 78,47 \text{ m.s}^{-1}$ . Conhecendo-se a distância entre os eixos dos vagões (18,70 m), a frequência de excitação do trem é avaliada em 4,20 Hz.

Medições realizadas previamente na estrutura colocaram em evidência vibrações e deformações excessivas do pavimento durante a passagem do TGV. Essas constatações conduziram à modificações nas condições de apoio da estrutura. Inicialmente a estrutura estava simplesmente apoiada e, posteriormente, foram colocados engastes mecânicos sob os apoios. Este procedimento de reforço estrutural foi realizado a fim de afastar a frequência de excitação causada pela passagem dos TGVs, da primeira frequência própria do viaduto.

O risco de ressonância aumenta ainda mais durante a passagem do TGV, devido à massa adicional tender a diminuir a frequência própria do viaduto. Para esse procedimento de reforço, foi colocado um sistema de mancais nas extremidades da estrutura como ilustrado na figura 6.6 a,b,c. Esse sistema é apertado através de uma chave de torque.



Figura 6.6 a – Vista inferior dos mancais. FONTE – CURY, 2010.



FIGURA 6.6 b, c – Detalhe dos mancais e procedimento de reforço. FONTE – CURY, 2010.

## 6.2. Identificação modal do viaduto Sens

Antes da análise de classificação, um breve estudo das densidades espectrais foi realizado para estimar as frequências excitadas durante os ensaios. As figuras 6.7 e 6.8 mostram, respectivamente, a resposta dinâmica obtida através dos acelerômetros e a densidade espectral referente ao sensor 5.

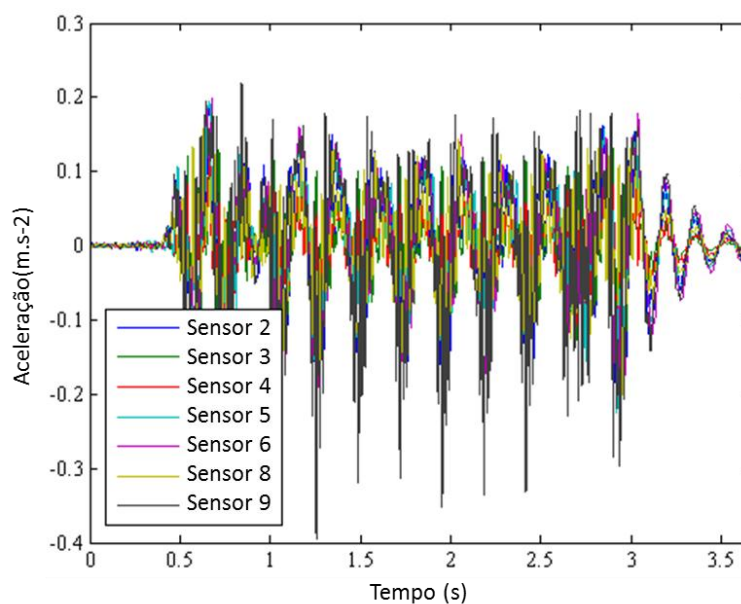


FIGURA 6.7 – Respostas temporais dos acelerômetros. FONTE – CURY, 2010.

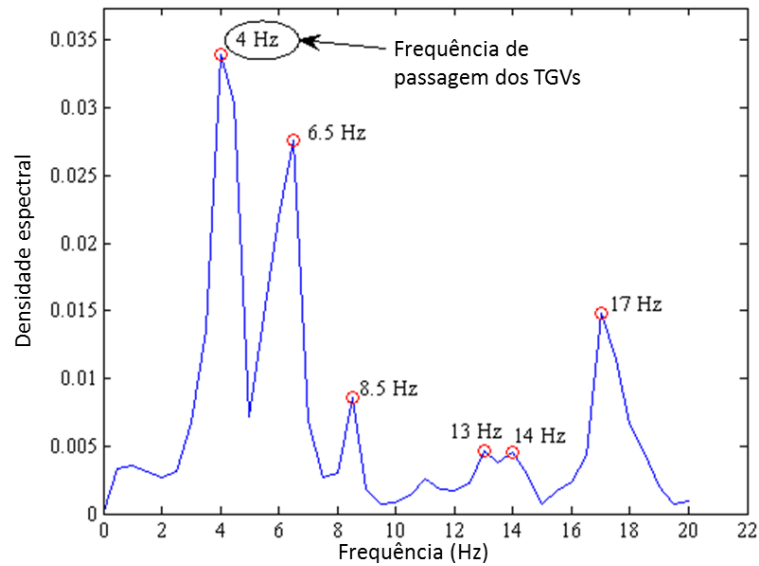


FIGURA 6.8 – Densidade espectral do acelerômetro 5. FONTE – CURY, 2010.

As medidas dos acelerômetros verticais permitem identificar as principais características modais da estrutura. Observando a densidade espectral avaliada, dois picos muito próximos podem ser observados. O primeiro pico, em torno de 4 Hz, corresponde à frequência de passagem dos TGVs. Os outros picos correspondem às quatro primeiras frequências próprias identificadas durante a passagem dos trens. De maneira geral, para todos os ensaios, as frequências estão compreendidas nos seguintes intervalos, representados por cada coluna:

$$\begin{bmatrix} 5 & | & 8 & | & 12 & | & 16 \\ 6 & | & 10 & | & 14 & | & 20 \end{bmatrix} \text{ Hz}$$

A ferramenta computacional Sys-Ident é utilizada para a identificação modal do viaduto. Após a filtragem em torno dos intervalos das frequências identificadas através da densidade espectral, o método do decremento aleatório é aplicado. A frequência de amostragem dos sinais é de 1/4096 s. A fim de se obter uma boa representação dos parâmetros modais identificados, 10 realizações são consideradas para cada ensaio. Nesta análise, cada passo do decremento é considerado como uma realização.

Os quatro primeiros modos próprios identificados são apresentados na figura 6.9 (CRÉMONA et al. 2004). Os pontos da malha não medidos, representados por quadrados, foram determinados a partir de uma interpolação linear.

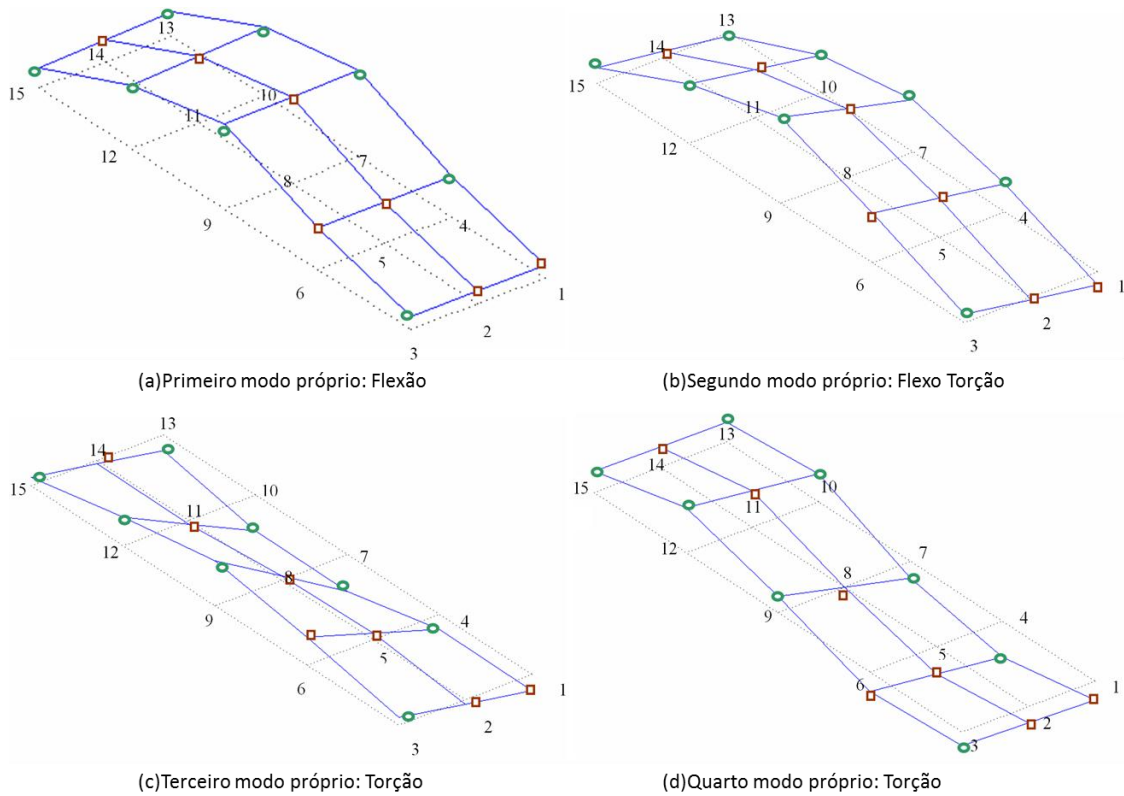


FIGURA 6.9 – Aspecto dos quatro primeiros modos próprios identificados. FONTE – CURY, 2010.

Como mencionado anteriormente, um dos objetivos das campanhas de medições dinâmicas realizadas neste viaduto é caracterizar e quantificar as melhorias geradas pelo reforço estrutural, a partir do comportamento dinâmico da estrutura. Para isso, três campanhas de medições foram realizadas. A primeira, realizada em 24 de junho de 2003 teve como objetivo avaliar as características modais do viaduto antes dos trabalhos de aperto dos mancais. Para essa campanha, 15 ensaios dinâmicos foram registrados, chamados de TGV1A a TGV15A.

Na tabela 6.1 são apresentados os valores médios calculados a partir das 10 realizações e os desvios-padrões globais, das frequências próprias e as taxas de amortecimento avaliadas.

TABELA 6.1 – Valores médios e desvios-padrões das frequências próprias e taxa de amortecimento antes do reforço. FONTE – CURY, 2010.

Modo	$\bar{f}$ [Hz]	$\sigma_f$ [Hz]	$\bar{\xi}$ (%)	$\sigma_{\xi}$ (%)
1	5,84	0,042	2,62	0,107
2	8,74	0,148	1,60	0,250
3	13,09	0,150	1,15	0,120
4	16,95	0,130	1,08	0,087

A segunda campanha de medições foi realizada no dia 25 de junho, no dia da realização do reforço. O procedimento de reforço foi efetuado em quatro fases. Durante a primeira fase, quatro mancais foram apertados. Em seguida, outros três mancais foram apertados. Nessas duas fases, foram apertados os mancais localizados no lado de Lyon. Durante a terceira fase foram apertados quatro mancais e, na quarta fase, mais três mancais foram apertados, desta vez do lado de Paris. No total, 13 ensaios dinâmicos foram registrados, chamados de TGV1R a TGV13R. Os valores médios e os desvios-padrões globais das frequências próprias são mostrados na tabela 6.2, para cada fase. Nessa tabela, a última linha corresponde aos valores médios levando-se em consideração as quatro fases do travamento. É possível notar que as frequências próprias aumentam significativamente na medida em que o procedimento vai sendo realizado.

TABELA 6.2 – Valores médios e desvios-padrões durante o travamento.

Fase	Frequência 1		Frequência 2		Frequência 3		Frequência 4	
	$\bar{f}$ (Hz)	$\sigma_f$	$\bar{f}$ (Hz)	$\sigma_f$	$\bar{f}$ (Hz)	$\sigma_f$	$\bar{f}$ (Hz)	$\sigma_f$
1	5,91	0,17	8,85	0,85	13,01	0,44	17,87	0,63
2	6,28	0,14	9,01	1,06	13,12	0,63	17,25	0,40
3	6,32	0,20	8,83	0,93	13,07	1,18	17,88	0,41
4	6,42	0,22	8,61	0,95	13,15	0,76	18,06	0,95
Média	<b>6,25</b>	0,18	<b>8,81</b>	0,95	<b>13,08</b>	0,75	<b>17,79</b>	0,63

Finalmente, a terceira campanha de medições ocorreu no dia 26 de junho de 2003. O objetivo era verificar os efeitos do reforço imediatamente após a conclusão do aperto dos mancais. Para a terceira campanha foram realizados 13 ensaios. A tabela 6.3 resume os valores médios e desvios-padrões globais das frequências próprias e das taxas de amortecimento avaliadas.

TABELA 6.3 – Valores médios e desvios-padrões das frequências próprias e taxas de amortecimento após o travamento. FONTE – CURY, 2010.

Modo	$\bar{f}$ [Hz]	$\sigma_f$ [Hz]	$\bar{\xi}$ (%)	$\sigma_{\xi}$ (%)
1	6,47	0,16	4,15	2,62
2	8,87	0,42	2,98	1,62
3	13,15	0,30	2,35	2,08
4	17,94	0,66	2,19	5,24

Na tabela 6.4 é apresentado um resumo das frequências próprias identificadas para cada campanha de medição. É possível constatar que as frequências naturais aumentaram, de fato, após a realização do reforço estrutural.

TABELA 6.4 – Comparativo das quatro primeiras frequências próprias identificadas para cada campanha de medição. FONTE – CURY, 2010.

Modo	Antes	Durante	Depois
	$\bar{f}$ [Hz]	$\bar{f}$ [Hz]	$\bar{f}$ [Hz]
1	5,84	6,25	6,47
2	8,74	8,81	8,87
3	13,09	13,08	13,15
4	16,95	17,79	17,94

Um dos objetivos desse capítulo é utilizar a abordagem baseada na ADS acoplada aos métodos de classificação a fim de discriminar os 3 estados estruturais do viaduto PK 075+317, ou seja, os estados antes, durante e depois do reforço estrutural.

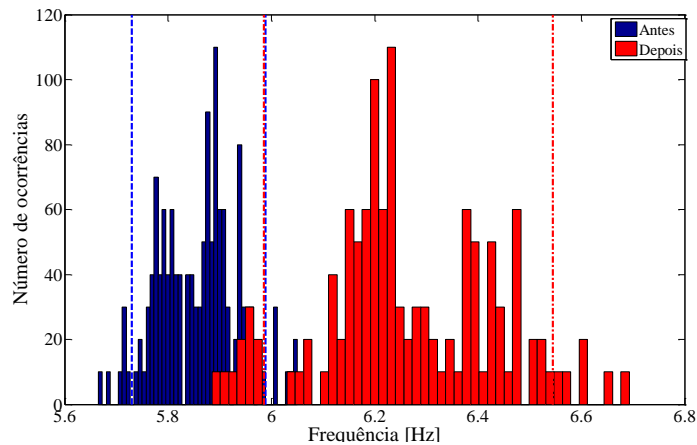
É importante salientar que, neste estudo considera-se que os TGV's possuem essencialmente as mesmas características físicas (peso, número de vagões, etc) em todos os ensaios realizados no viaduto.

### 6.3. Análise clássica

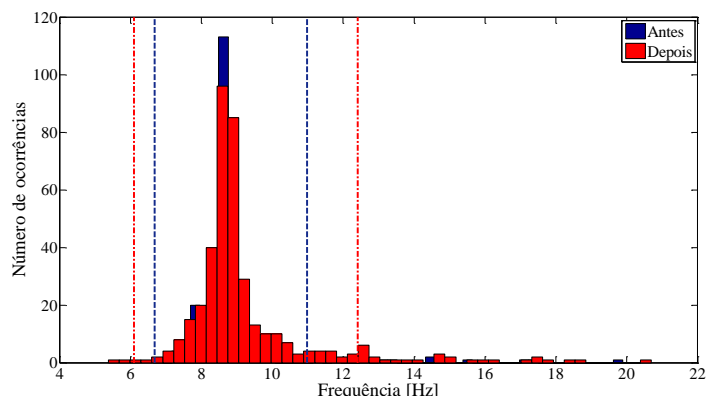
Assim como feito nas simulações numéricas, conduz-se, primeiramente, um breve estudo clássico dos parâmetros modais antes de se passar às abordagens propostas neste trabalho. O objetivo é mostrar que a análise clássica não permite, de forma geral, concluir com um adequado nível de confiança, que os três estados estruturais – antes,

durante e depois dos trabalhos de reforço – do viaduto PK 075, estão associados a um processo de modificação estrutural.

A figura 6.10 mostra os histogramas obtidos para as quatro primeiras frequências naturais identificadas para os estados estruturais “antes” e “depois” do reforço. Nesta figura, os tracejados verticais representam os intervalos de confiança de 95% para os valores médios das frequências próprias. Ainda que a diferença entre estes dois estados estruturais seja evidente para a primeira frequência (já que os tracejados verticais não se superpõem), o mesmo não pode ser observado para as demais frequências. Analisando unicamente estes histogramas, não é possível afirmar com 95% de confiança, que os estados estruturais são distintos.

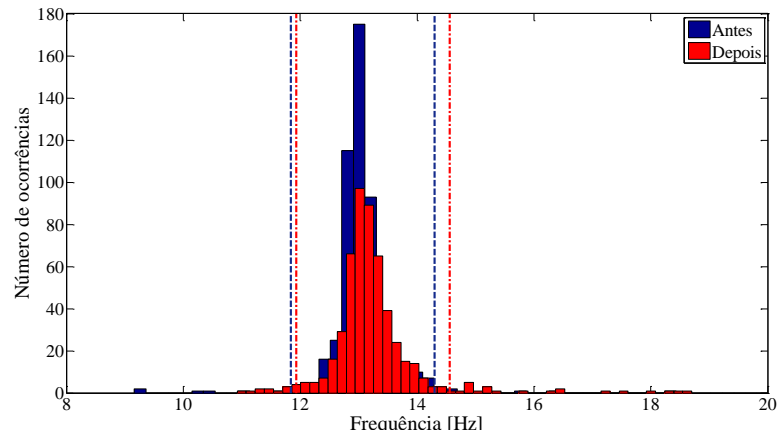


(a) Histogramas obtidos para a primeira frequência.

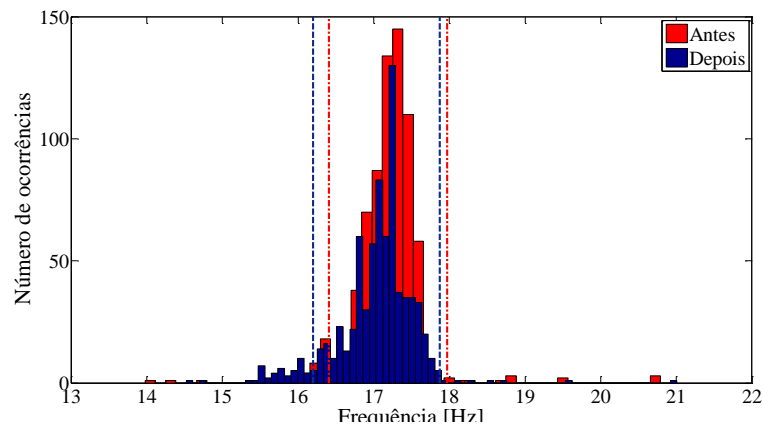


(b) Histogramas obtidos para a segunda frequência.





(c) Histogramas obtidos para a terceira frequência.



(d) Histogramas obtidos para a quarta frequência.

FIGURA 6.10 – Comparativo dos histogramas para as quatro primeiras frequências naturais, antes e depois do procedimento de reforço.

Ademais, uma análise considerando as deformadas modais é considerada. Similarmente ao proposto nas simulações numéricas, os valores do índice MAC são calculados para cada um dos quatro modos próprios identificados. Na figura 6.11, é possível notar que a variação dos modos é pequena, com exceção do modo 2 que se mostra um pouco mais sensível às modificações estruturais introduzidas pelos trabalhos de reforço. Esta observação se reflete diretamente sobre os índices MAC. Com efeito, para o primeiro, terceiro e quarto modos, os índices variam entre 76% e 100%, o que denota, para este indicador, pouca sensibilidade dos modos em relação ao reforço estrutural. Em

contrapartida, o segundo modo apresenta variações relativamente maiores, o que implica em coeficientes MAC mais baixos (entre 53 e 75%).

De toda sorte, os resultados obtidos a partir das análises clássicas não são suficientemente claros quanto à distinção dos estados estruturais. Assim como observado nas simulações numéricas, torna-se necessário aplicar técnicas mais robustas, que sejam capazes de discriminar, com adequado grau de confiança, os diferentes estados estruturais existentes no viaduto PK 075.

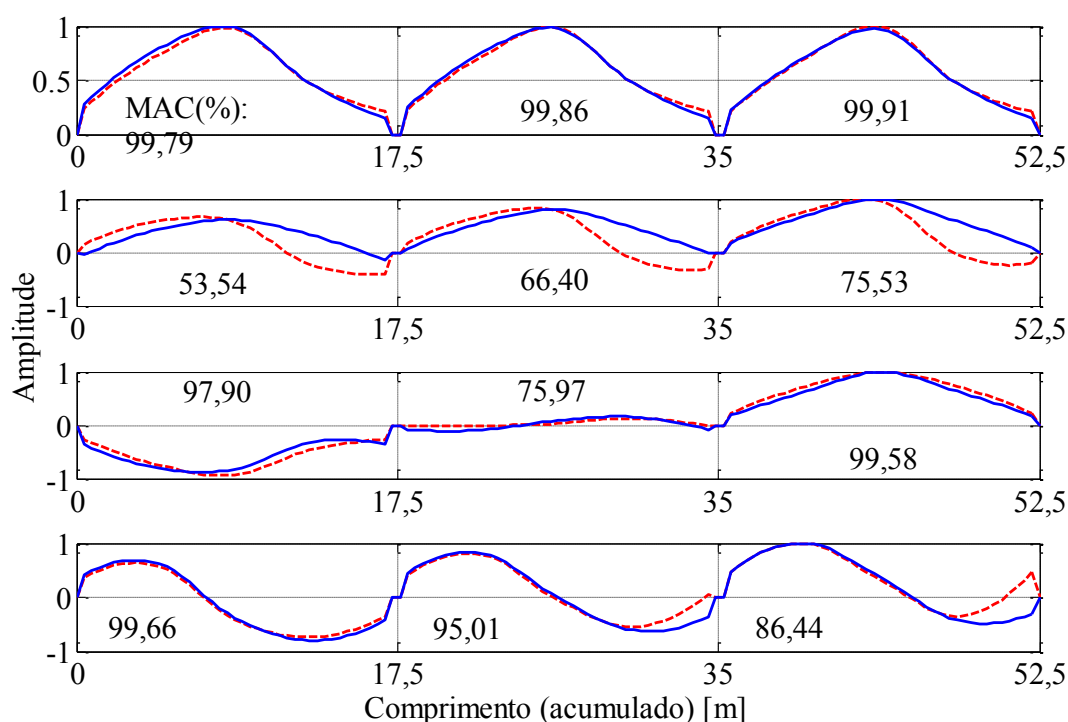


Figura 6.11 – Comparativo entre os quatro primeiros próprios identificados e respectivos valores MAC.

#### 6.4. Análise simbólica

Como realizado nas simulações numéricas, a ADS também é aplicada juntamente aos métodos de classificação aos ensaios realizados no viaduto PK 075+317. Durante as 3 campanhas de medições, 41 ensaios foram registrados, sendo 15 ensaios para o estado estrutural antes do procedimento de reforço na estrutura, 13 ensaios estruturais durante o procedimento de reforço estrutural e 13 ensaios depois da realização do reforço. O objetivo é utilizar as técnicas de *clustering* para separar os 41 ensaios dinâmicos em três

grupos distintos. Inicialmente, os métodos de classificação são aplicados às representações simbólicas dos sinais. Em seguida, consideram-se as representações simbólicas das frequências e, por último, as representações simbólicas dos modos próprios. Para todas as classificações realizadas na análise do viaduto PK 075+317, são consideradas as quatro primeiras frequências próprias e seus modos próprios associados. Neste estudo a representação simbólica considerada é a de histogramas com 12 categorias.

Dado o elevado número de parâmetros existentes na aplicação dos métodos de classificação (métricas, critérios de agregação e formas de inicialização), são apresentados somente os melhores resultados no corpo principal do texto. Os demais resultados são mostrados no anexo B.

#### **6.4.1. Sinais**

Os primeiros resultados correspondem à aplicação do método das divisões hierárquicas aplicados aos sinais descritos de maneira simbólica. O objetivo é conseguir distinguir os 3 grupos de ensaios de maneira homogênea, ou seja, conseguir separar os ensaios TGV1A até TGV15A em um *cluster* que represente os ensaios antes do reforço estrutural, os ensaios TGV1R até TGV13R em um *cluster* que represente os ensaios durante o reforço e os ensaios TGV1B até TGV13B em um *cluster* que represente os ensaios depois do reforço.

Na figura 6.12 é apresentado o dendograma obtido para os sinais transformados em dados simbólicos, utilizando-se o método da aglomeração hierárquica, a distância *cosine* e o critério de agregação *complete*. Os índices de acerto para o estado estrutural antes do reforço foi de 60%, para o durante de 100% e de 23,07% para depois do reforço.

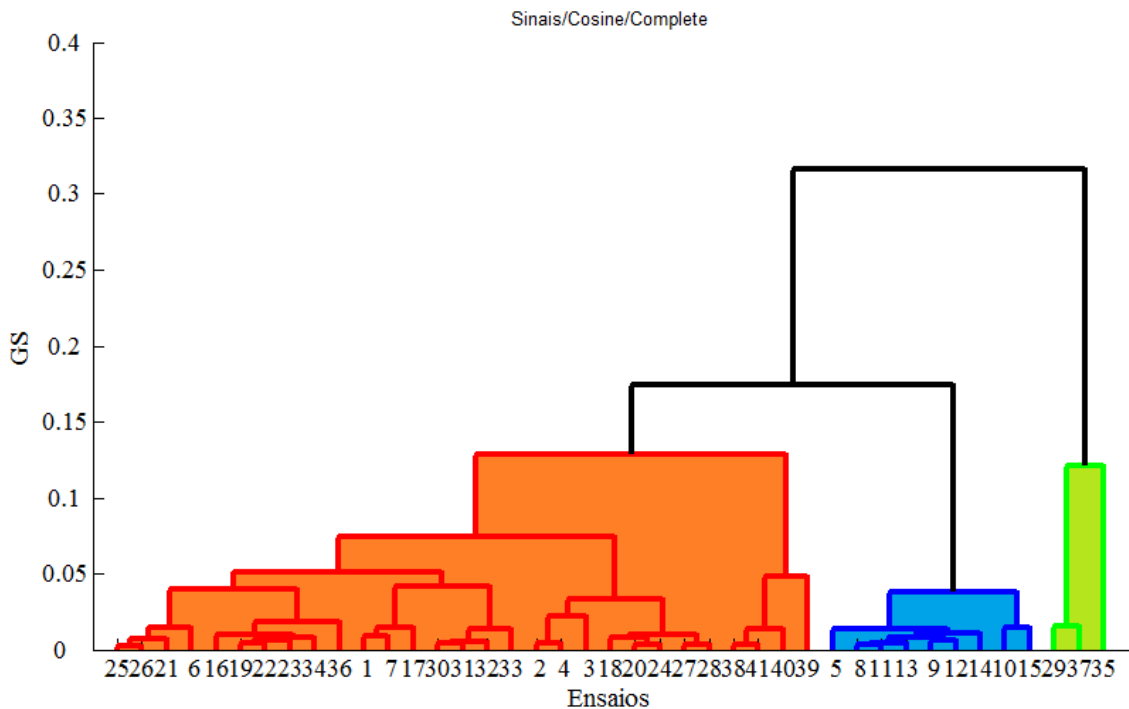


FIGURA 6.12 – Dendrograma da classificação (sinais; dist. *cosine*; agreg. *complete*).

Aplicando o método de classificação k-means, utilizando a distância *correlation* e o critério de inicialização amostral, são obtidos os índices de acerto apresentados na tabela 6.5.

TABELA 6.5 – Índices de acerto da classificação (k-means; sinais; dist. *correlation*; inicialização amostral).

Antes	73,33%
Durante	84,61%
Depois	30,77%

Utilizando o c-means, os índices de acerto são mais baixos (tabela 6.6). Entre parênteses são encontrados os graus de certeza da classificação. Como é possível perceber, os índices encontrados não são adequados.

TABELA 6.6 – Índices de acerto da classificação obtida utilizando o método c-means para os sinais descritos de forma simbólica.

Antes	66,67% (60%)
Durante	38,46% (42%)
Depois	53,84%(41%)

Na tabela 6.7 é apresentado um resumo com os resultados obtidos com os métodos de classificação utilizando os sinais representados simbolicamente. Assim como observado nas análises numéricas, os resultados não foram plenamente satisfatórios, mas indicam a capacidade de os sinais serem utilizados como parâmetros de detecção de danos estruturais.

TABELA 6.7 – Resumo dos índices de acerto obtidos com todos os métodos de classificação utilizando os sinais.

	Aglomeraco Hierrquica	k-means	c-means
Antes	60%	73,33%	66,67% (60%)
Durante	100%	84,61%	38,46% (42%)
Depois	23,07%	30,77%	53,84%(41%)

#### 6.4.2. Frequncias

Aps a utilizao dos sinais, as frequncias so transformadas em dados simblicos para serem utilizadas como dados de entrada nas tcnicas de *clustering*. Nesse estudo, as quatro primeiras frequncias prprias so utilizadas.

Aplicando o mtodo da aglomerao hierrquica com a distncia *spearman* e o critrio de agregao *ward*, obtm-se o dendograma apresentado na figura 6.13. Os índices de acerto obtidos foram de 92,30% antes do reforo estrutural, 100% durante o procedimento de aperto dos mancais e 92,30% aps a realizao do reforo.  possvel observar as frequncias proporcionam bons índices de acerto na classificao, assim como foi obtido na anlise numrica.

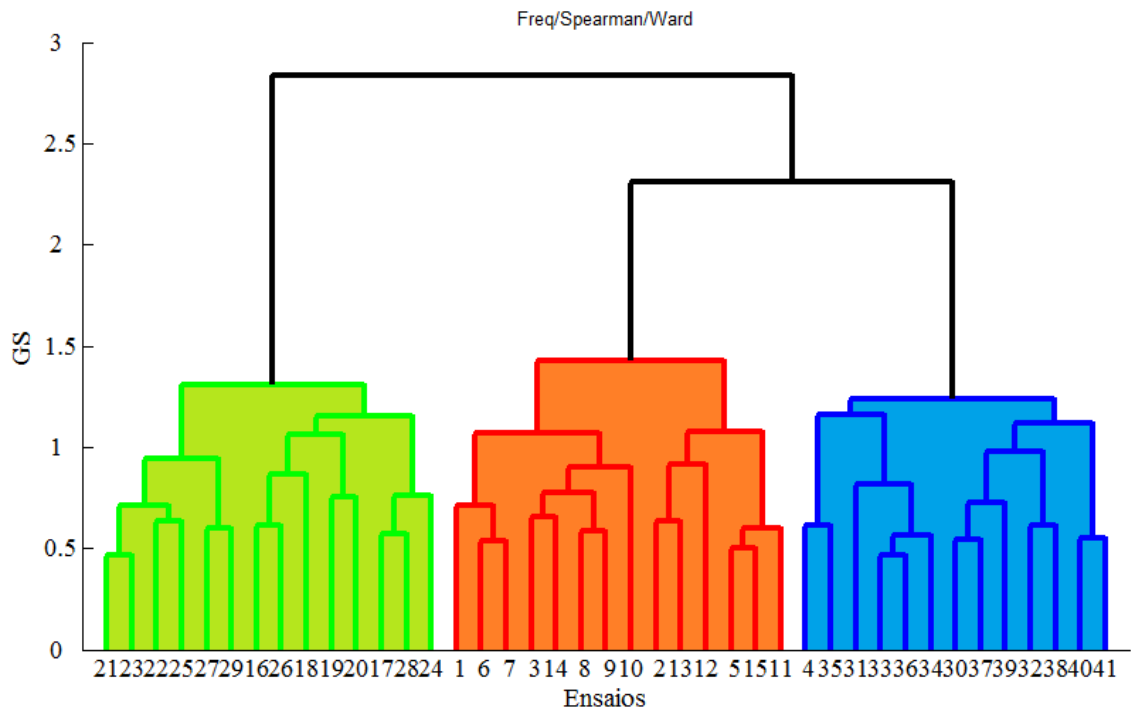


FIGURA 6.13 – Dendrograma da classificação (frequências; dist. *spearman*; agreg. *ward*).

Aplicando o método k-means, utilizando a distância *cityblock* e o método de inicialização amostral, os resultados são muito bons. Os índices de acerto são apresentados na tabela 6.8.

TABELA 6.8 – Índices de acerto da classificação (k-means; frequências; dist. *cityblock*; inicialização amostral).

Antes	100%
Durante	100%
Depois	100%

Finalmente, o método c-means é aplicado. Os índices de acerto obtidos são apresentados na tabela 6.9.

TABELA 6.9 – Índices de acerto da classificação obtida utilizando o método c-means para as frequências descritas de maneira simbólica.

Antes	86,67% (33%)
Durante	100% (33%)
Depois	92,3%(33%)

Na tabela 6.10 é apresentado um resumo dos índices de acerto obtidos aplicando todos os métodos de classificação utilizando as frequências. Os bons resultados obtidos na aplicação numérica foram repetidos na análise experimental, o que confirma que as frequências transformadas em dados simbólicos são um parâmetro confiável para identificação de comportamentos estruturais.

TABELA 6.10 – Resumo dos índices de acerto obtidos com todos os métodos de classificação utilizando as frequências.

	<b>Aglomeraco Hierrquica</b>	<b>k-means</b>	<b>c-means</b>
<b>Antes</b>	93,30%	100%	86,67%(33%)
<b>Durante</b>	100%	100%	100%(33%)
<b>Depois</b>	92,30%	100%	92,3%(33%)

### 6.4.3. Modos Prprios

Finalmente os modos prprios transformados em dados simblicos so utilizados para analisar os diferentes estados estruturais do viaduto. Como na anlise numrica, os modos prprios so transformados em dados simblicos separadamente. Com isso, as tcnicas de *clustering* so aplicadas para cada um dos quatro modos prprios considerados.

#### 6.4.3.1. Primeiro modo

A primeira classificao utilizando os primeiros modos prprios  obtida aplicando o mtodo da aglomerao hierrquica com a distncia *spearman* e o critrio de agregao *complete*. O dendograma obtido  mostrado na figura 6.14. Os índices de acerto foram de 100% para o estado estrutural “antes”, e de 92,30% para os estados “durante” e

“depois”. Os índices de acerto obtidos foram adequados, sendo que o método de classificação errou apenas a classificação dos ensaios 22 (TGV7R) e 39 (TGV11B).

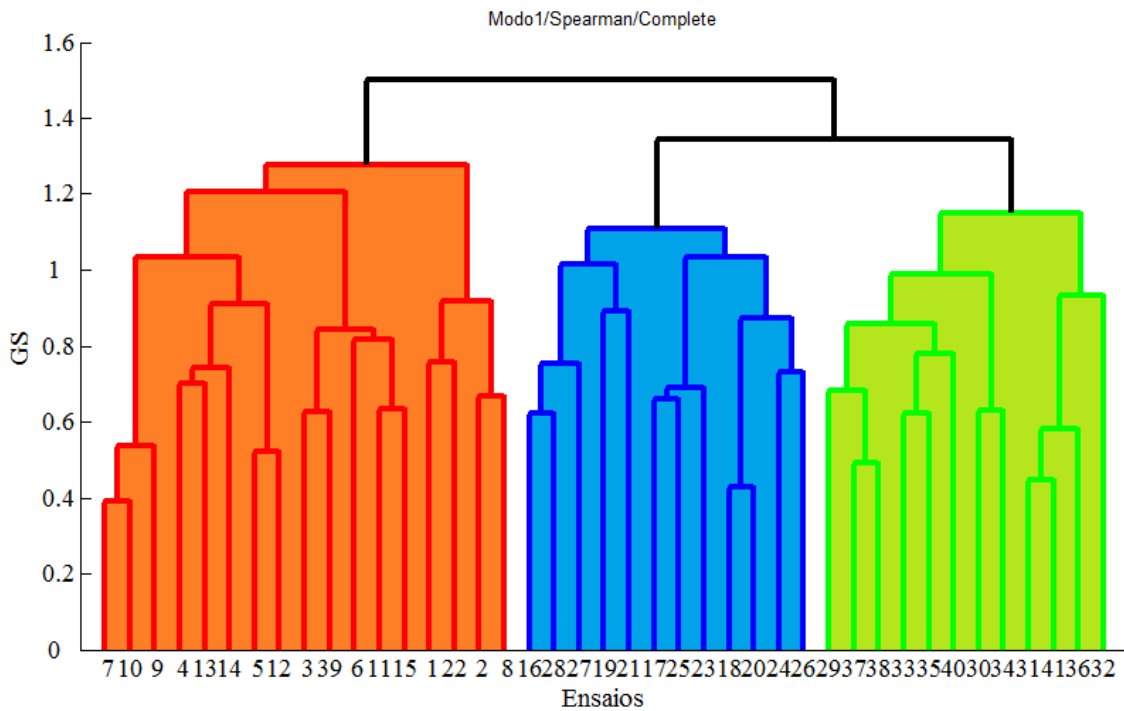


FIGURA 6.14 – Dendrograma da classificação (primeiro modo; dist. *spearman*; agreg. *complete*).

Aplicando o método k-means, utilizando a distância *cityblock* e o método de inicialização amostral, os resultados são muito bons. Os índices de acerto são apresentados na tabela 6.11.

TABELA 6.11 - Índices de acerto da classificação (k-means, primeiro modo; dist. *cityblock*; inicialização amostral).

Antes	100%
Durante	100%
Depois	100%

Para a classificação utilizando o c-means, os índices de acerto são apresentados na tabela 6.12.



TABELA 6.12 – Índices de acerto da classificação obtida aplicando o método c-means para os primeiros modos próprios.

Antes	86,67%(33%)
Durante	61,53%(33%)
Depois	100%(33%)

Na tabela 6.13 é apresentado o resumo dos índices de acerto obtidos com todos os métodos de classificação utilizando os primeiros modos próprios. Os índices de acerto foram satisfatórios, porém o método de classificação c-means se mostrou menos adequado em relação aos outros métodos. Ainda assim, a utilização dos modos próprios para essa abordagem mostrou-se bastante eficiente, já que taxas elevadas de classificação correta (100%) foram obtidas.

TABELA 6.13 - Resumo dos índices de acerto obtidos com todos os métodos de classificação utilizando os primeiros modos próprios.

	Aglomeraco Hierrquica	k-means	c-means
Antes	100%	100%	86,67%(33%)
Durante	92,30%	100%	61,53%(33%)
Depois	92,30%	100%	100%(33%)

#### 6.4.3.2. Segundo modo

Os segundos modos próprios transformados em dados simbólicos são agora utilizados como dados de entrada para os métodos de classificação. O dendograma obtido aplicando o método da aglomeração hierárquica é apresentado na figura 6.15. A distância utilizada é a *cityblock* e o critério de agregação utilizado é o *weighted*. Os índices de acerto obtidos foram de 86,66% para os ensaios referentes ao estado estrutural antes da realizao do procedimento de reforo, 100% para os ensaios referentes ao estado estrutural durante a realizao do procedimento de reforo, e 92,30% para os ensaios depois do reforo estrutural.

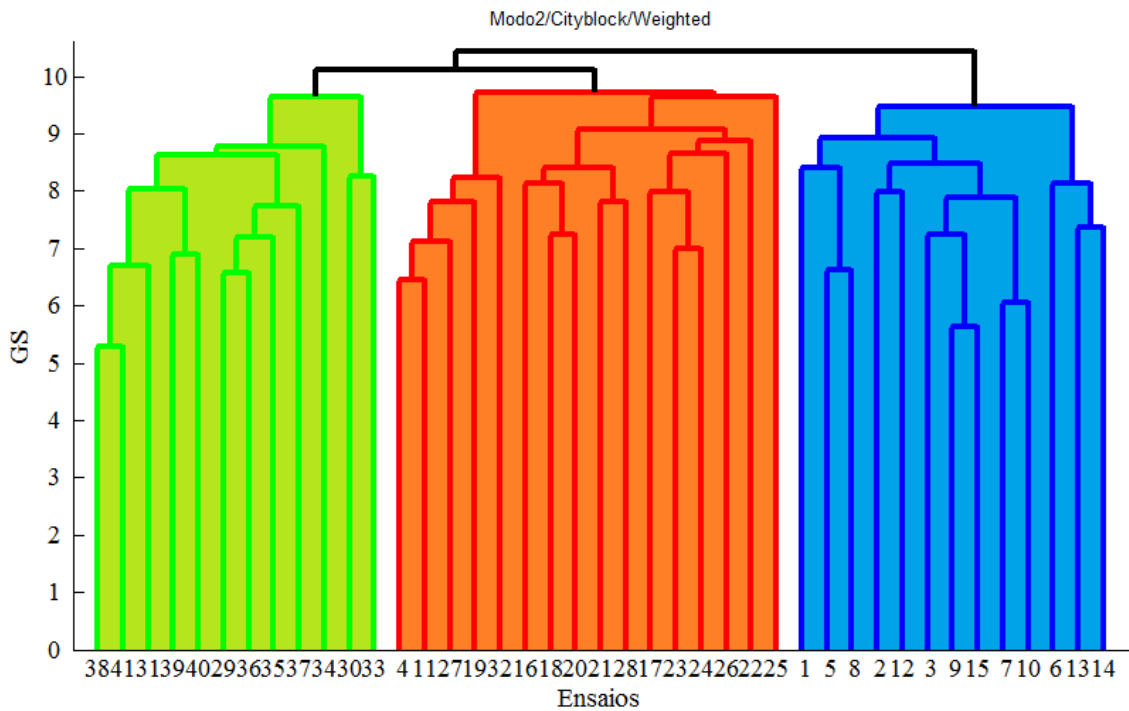


FIGURA 6.15 – Dendograma da classificação (segundo modo; dist. *cityblock*; agreg. *weighted*).

Para a classificação aplicando o método k-means utilizando os segundos modos próprios, considerando a distância *cityblock* e o método de inicialização amostral, são obtidos os índices de acerto apresentados na tabela 6.14.

TABELA 6.14 – Índices de acerto da classificação (k-means; segundo modo; dist. *cityblock*; inicialização amostral).

Antes	100%
Durante	76,92%
Depois	92,30%

Os índices de acerto obtidos para o método de classificação c-means utilizando os segundos modos próprios são apresentados na tabela 6.15.

TABELA 6.15 – Índices de acerto da classificação obtida utilizando o método c-means para os segundos modos próprios descritos de forma simbólica.

Antes	60% (33%)
Durante	76,92% (33%)
Depois	100% (33%)

O resumo dos índices de acerto do método de classificação utilizando os segundos modos próprios é apresentado na tabela 6.16. Observando os índices de acerto da tabela, é possível notar que foram obtidos bons resultados utilizando os três métodos de classificação. Os graus de certeza observados entre parênteses obtidos através do c-means indicam que, apesar dos bons resultados, o método possui certa dúvida quanto à classificação. O método de classificação que apresentou melhores resultados foi o método da aglomeração hierárquica.

TABELA 6.16 - Resumo dos índices de acerto obtidos com todos os métodos de classificação utilizando os segundos modos próprios.

	Aglomeração Hierárquica	k-means	c-means
Antes	86,66%	100%	60% (33%)
Durante	100%	76,92%	76,92% (33%)
Depois	92,30%	92,30%	100% (33%)

#### 6.4.3.3. Terceiro modo

Os dados aqui analisados são os referentes aos terceiros modos próprios. Primeiramente, é aplicado o método da aglomeração hierárquica com a distância *spearman* e o critério de agregação *ward*. Nessa classificação, 86,66% dos ensaios “antes” foram classificados corretamente, 61,53% dos ensaios “durante” e 100% dos ensaios “depois”. O dendograma referente a essa aplicação é mostrado na figura 6.16.

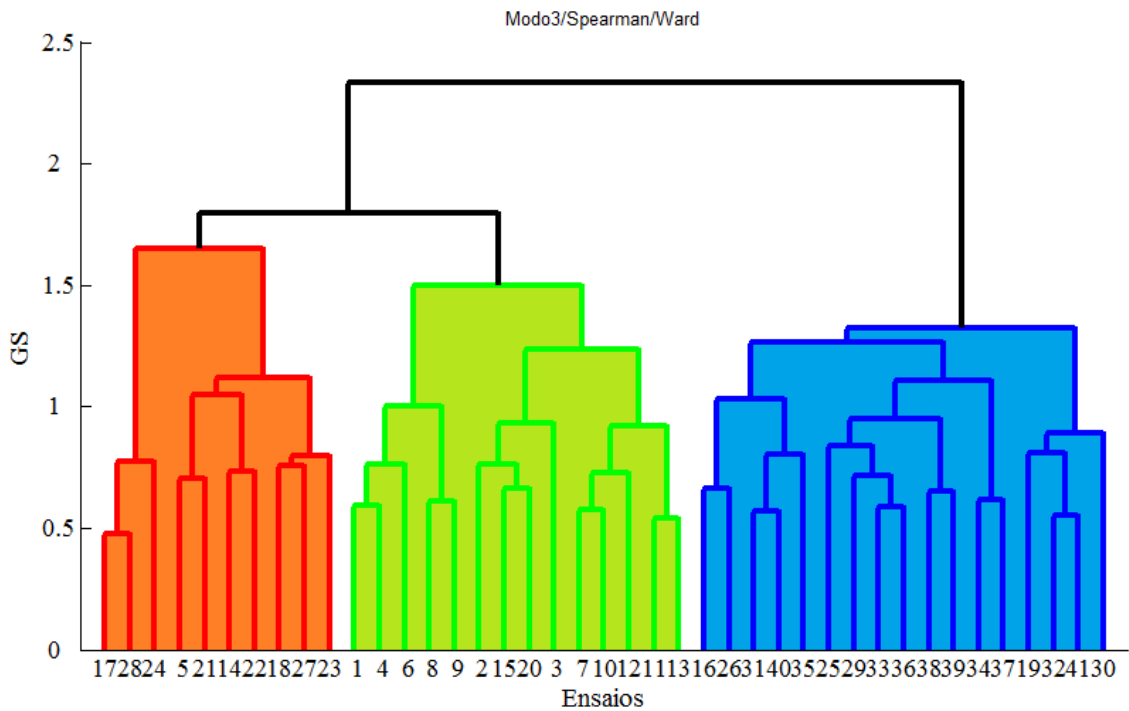


FIGURA 6.16 – Dendrograma da classificação (terceiro modo; dist. *spearman*; agreg. *ward*).

Em seguida, aplica-se o método k-means considerando a distância *cityblock* e o método de inicialização amostral. Os índices de acerto são apresentados na tabela 6.17.

TABELA 6.17 – Índices de acerto da classificação (k-means; terceiro modo; dist. *cityblock*; inicialização amostral).

Antes	100%
Durante	84,61%
Depois	92,30%

Finalmente, apresentam-se na tabela 6.18 os índices de acerto obtidos com o método c-means.

TABELA 6.18 – Índices de acerto da classificação obtida utilizando o método c-means para os terceiros modos próprios descritos de forma simbólica.

Antes	100%(33%)
Durante	38,46%(33%)
Depois	92,3%(33%)

Para melhor análise dos métodos de classificação, a tabela 6.19 apresenta um resumo dos índices de acerto obtidos com os três métodos de classificação. Os melhores resultados foram obtidos utilizando o método k-means. Os resultados obtidos utilizando os terceiros modos próprios foram melhores do que os resultados obtidos utilizando os segundo modos próprios para descrever os ensaios dinâmicos. Já para o método da aglomeração hierárquica e para o c-means, a utilização dos terceiros modos próprios forneceu índices de acerto menores do que utilizando os primeiros e segundos modos próprios.

TABELA 6.19 - Resumo dos índices de acerto obtidos com todos os métodos de classificação utilizando os terceiros modos.

	Aglomeração Hierárquica	k-means	c-means
Antes	86,66%	100%	100%(33%)
Durante	61,53%	84,61%	38,46%(33%)
Depois	100%	92,30%	92,3%(33%)

#### 6.4.3.4. Quarto modo

Finalmente, os métodos de classificação são aplicados para os quartos modos próprios. O método da aglomeração hierárquica é aplicado aos quartos modos próprios utilizando a distância *cityblock* e o critério de agregação *ward*. Nessa aplicação, os índices de acerto obtidos foram de 73,33%, 100% e 92,30%, para os ensaios referentes aos estados estruturais “antes”, “durante” e “depois”, respectivamente. Na figura 6.17 é apresentado o dendograma obtido para essa classificação.

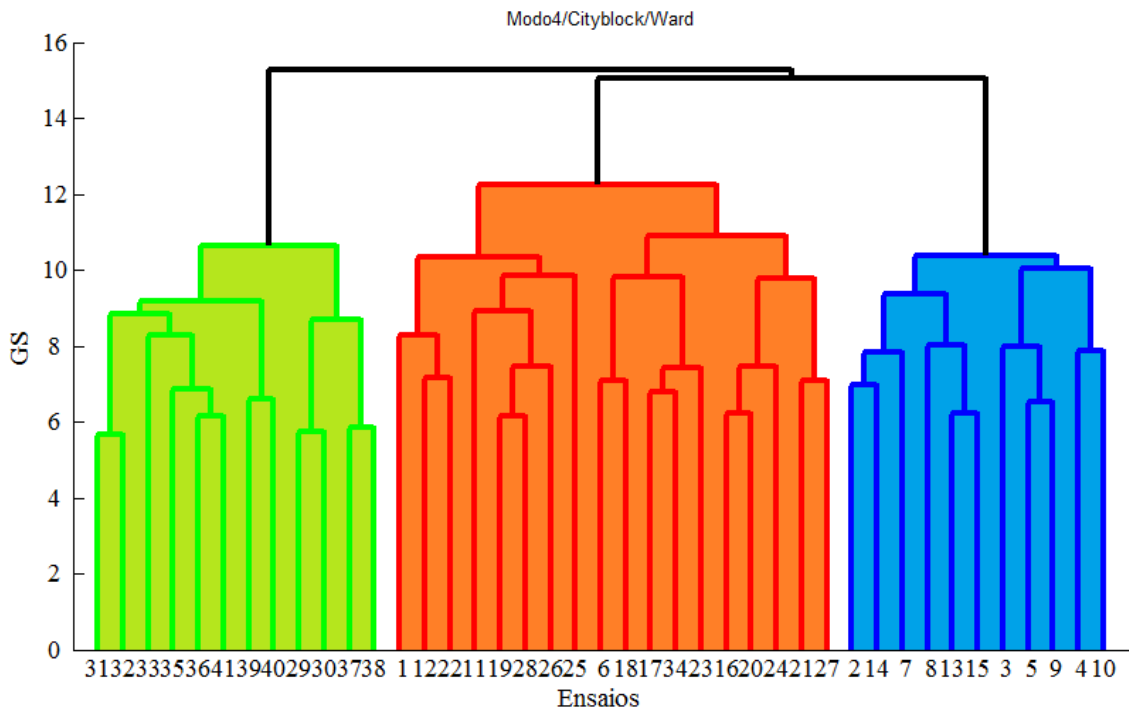


FIGURA 6.17 – Dendograma da classificação (quarto modo; dist. *cityblock*; agreg. *ward*).

Os índices de acerto obtidos através do k-means com a distância *correlation* e o método de inicialização amostral são mostrados na tabela 6.20.

TABELA 6.20 – Índices de acerto da classificação (k-means; quarto modo; dist. *correlation*; inicialização amostral).

Antes	60%
Durante	61,53%
Depois	92,30%

Por último, os índices de acerto aplicando o c-means são apresentados na tabela 6.21.

TABELA 6.21 – Índices de acerto da classificação obtida utilizando o método c-means para os quartos modos próprios descritos de forma simbólica.

Antes	73,33%(33%)
Durante	61,53%(33%)
Depois	84,61%(33%)

A tabela 6.22 apresenta um resumo com os índices de acerto dos métodos de classificação utilizados para os quartos modos próprios. O método k-means que apresentou melhores resultados nos outros modos próprios, não apresentou índices de acerto tão elevados utilizando os quartos modos próprios. Utilizando os quartos modos próprios, os melhores resultados foram obtidos utilizando o método da aglomeração hierárquica.

TABELA 6.22 - Resumo dos índices de acerto obtidos com todos os métodos de classificação utilizando os quartos modos próprios.

	Aglomeração Hierárquica	k-means	c-means
Antes	73,33%	60%	73,33%(33%)
Durante	100%	61,53%	61,53%(33%)
Depois	92,30%	92,30%	84,61%(33%)

Fazendo uma análise geral da utilização dos modos próprios, os ensaios dinâmicos descritos pelos primeiros modos próprios forneceram melhores resultados quando aplicados aos métodos de classificação. Para o primeiro e terceiro modo, o método de classificação que melhor discriminou os estados estruturais foi o k-means. Já para o segundo e quarto modo, o método de aglomeração hierárquica foi o que obteve melhores índices de acerto.

### 6.5. Estudo complementar

Um estudo complementar semelhante ao realizado para a análise numérica também é considerado para os dados experimentais. Nesse estudo são considerados somente os ensaios realizados antes e depois do reforço estrutural. As distâncias, critérios de agregação e métodos de inicialização utilizados são os mesmos utilizados para a classificação dos ensaios em 3 *clusters*. Como observado na tabela 6.23, os índices de acerto foram adequados, exceto para o método da aglomeração hierárquica e para o c-means aplicados aos sinais. Os graus de pertinência obtidos aplicando o c-means foram de 50% para todos os casos, o que indica que o método possui dúvidas em relação às classificações obtidas.

TABELA 6.23 - Índices de acerto (em %) utilizando os métodos de classificação para os estados antes e depois do reforço estrutural do viaduto PK 075+317.

Aglomeración Hierárquica						
	Sinais	Freq.	Modo 1	Modo 2	Modo 3	Modo 4
Antes	100	100	100	100	84,61	100
Depois	23,07	93,33	92,30	100	66,66	66,66
k-means						
	Sinais	Freq.	Modo 1	Modo 2	Modo 3	Modo 4
Antes	100	100	100	100	100	73,33
Depois	69,23	100	100	100	84,61	100
c-means						
	Sinais	Freq.	Modo 1	Modo 2	Modo 3	Modo 4
Antes	73,33(50%)	86,66(50%)	100(50%)	100(50%)	80(50%)	86,66(72%)
Depois	38,46(50%)	100(50%)	100(50%)	100(50%)	100(50%)	100(53%)

De maneira geral, quando apenas dois estados estruturais são considerados (antes e depois do reforço estrutural), os métodos de classificação fornecem resultados iguais ou superiores aos obtidos considerando-se três estados estruturais. Isto se deve ao fato de os estados “antes” e “depois” serem, na realidade, mais diferentes entre si, ao passo que o estado “durante” representa uma transição entre aqueles dois primeiros.

Deve-se destacar que os índices de acerto para o segundo modo de vibração foram de 100% para todos os métodos de classificação utilizados. Isto pode ser justificado pelo fato de que este modo é o que apresenta os menores valores de MAC; portanto, o mais sensível às modificações estruturais realizadas. Também deve-se destacar que os índices de acerto obtidos utilizando o primeiro modo de vibração foram muito bons, apesar de não ter sido possível observar uma grande variação na amplitude de vibração durante a análise clássica realizada.





# CONCLUSÕES

Estruturas de Engenharia Civil estão sujeitas à processos de deterioração que podem ocorrer gradualmente ou abruptamente. Esses processos de deterioração geralmente acarretam modificações nas respostas dinâmicas da estrutura, assim nas frequências naturais e modos de vibração.

Com a revisão bibliográfica realizada, foi possível perceber que as técnicas existentes são essencialmente baseadas na variação das frequências naturais, que não se mostravam indicadores de danos confiáveis, por serem pouco sensíveis à presença de danos; na variação dos modos de vibração, que mostravam certo potencial para serem utilizados na detecção de danos estruturais e na variação de indicadores evoluídos construídos a partir desses parâmetros. Um dos inconvenientes dos métodos encontrados na literatura é a incapacidade de utilizar diretamente as respostas dinâmicas. Isso muitas vezes se deve à presença de ruídos e a dificuldade de manipular esse tipo de dado corretamente.

Com o intuito de administrar grandes quantidades de dados, foi proposta uma abordagem baseada na Análise de Dados Simbólicos (ADS). Esta metodologia permite manipular os dados mais facilmente, de tal forma que os dados sejam representados de forma mais compacta, mas guardando sua variabilidade interna (média, desvio padrão, variância, etc). Nessa dissertação, a ADS foi utilizada juntamente com os métodos de classificação (aglomeração hierárquica, k-means e c-means) para discriminar diferentes estados estruturais. A utilização dos métodos de classificação de dados já é conhecida em áreas da matemática, biologia, economia, entre outras. Entretanto, aplicações em engenharia civil, como para a detecção de danos estruturais, representam uma alternativa inovadora.

Utilizando a abordagem proposta, observou-se que, diferentemente das técnicas de identificação de danos existentes, o uso das frequências naturais como indicadores de dano mostrou-se o mais confiável. Os modos de vibração, por sua vez, continuaram apresentando bons resultados, porém mostraram-se sensíveis à presença de ruído.

Finalmente, os sinais dinâmicos mostraram certo potencial para sua utilização como indicadores de danos, por não necessitarem de passar pelo processo de identificação modal e por apresentarem resultados relativamente satisfatórios.

De forma geral, o método de classificação que se mostrou mais eficiente foi o k-means. Dentre os diversos parâmetros utilizados para a aplicação dos métodos de classificação, os que apresentaram melhores resultados, na maioria das aplicações, foi a medida de dissimilaridade correlation, o critério de agregação ward e o método de inicialização uniforme.

Enfim, foi possível observar melhores resultados quando os métodos de classificação são utilizados para classificar dois estados estruturais essencialmente distintos, como foi o caso dos estudos complementares realizados nesse trabalho. Nesses estudos, os índices de acerto encontrados foram iguais ou superiores aos obtidos quando considerado todos os estados estruturais simulados/existentes.

A principal contribuição dessa dissertação foi o desenvolvimento de uma nova abordagem utilizando a ADS + técnicas de clustering capaz de discriminar diversos comportamentos estruturais a partir de dados descritos de forma mais compacta, mas, ao mesmo tempo, levando-se em consideração suas incertezas (i.e. valor médio e desvio-padrão).

## Sugestões para trabalhos futuros

Como sugestões para trabalhos futuros propõem-se:

- Redescrever os ensaios de forma simbólica utilizando uma representação global, ou seja, utilizar os sinais brutos, as frequências naturais e os modos de vibração para descrever os ensaios dinâmicos em um único histograma;
- Utilizar métodos para a classificação de novos ensaios, de forma que estes ensaios possam ser classificados segundo um estado estrutural de referência;
- Calcular índices de particionamento ótimo para os métodos de classificação. A utilização desses índices pode se mostrar importante por poder apontar um novo estado estrutural desconhecido *a priori*;
- Implementar o método c-means para as diversas distâncias e métodos de inicialização dos centroides;
- Utilizar diferentes representações simbólicas (intervalos interquartis, absolutos, etc.), para analisar suas influências sobre a classificação dos ensaios.
- Realizar simulações de danos com menores reduções do módulo de elasticidade.



# REFERÊNCIAS

## BIBLIOGRÁFICAS

Abdel Wahab, M.; De Roeck, G. e Peeters, B. (1999), Parametrisation of damage in reinforced concrete structures using model updating, *Journal of Sound and Vibration*, v. 228, n. 4, p. 717-730.

Aktan, A.; Lee, K.; Chuntavan, C. e Aksel, T. (1994). Modal testing for structural identification and condition assessment of constructed facilities, in: Proceedings of the International Modal Analysis Conference, 12, Honolulu, Havaí, p. 462-468.

Allemang, R. e Brown, D. (1982). A correlation coefficient for modal vector analysis, in: Proceedings of the International Modal Analysis Conference, 1, Orlando, Estados Unidos, p. 145-151.

Alvandi, A. e Crémona, C. (2006). Assessment of vibration-based damage identification techniques, *Journal of Sound and Vibration*, v. 292, n. 2, p. 179-202.

Alvandi, A. (2004). *Contribution à l'utilisation pratique de l'évaluation dynamique pour la détection d'endommagements dans les ponts*, Tese de doutorado, École Nationale des Ponts et Chaussées, França.

Alvin, K. (1997). *Finite element model update via bayesian estimation and minimization of dynamics residuals*, *AIAA Journal*, v. 35, n. 5, p. 879-886.

Barbosa, F. e Crémona, C. (2001). Identification modale de structures sous sollicitation ambiante, Relatório técnico, Laboratoire Central des Ponts et Chaussées, França.

Basseville, M.; Benveniste, A.; Goursat, M. e Mevel, L. (2006). Subspace-based algorithms for structural identification, damage detection, and sensor data fusion, *EURASIP Journal on Advances in Signal Processing*, v. 2007, n. 1, p. 1-13.

- Bertrand, P. (1986), *Etude de la Représentation Pyramidale*, Tese de doutorado, Université Paris-Dauphine, França.
- Bezdek, J.C. (1981). *Pattern recognition with fuzzy objective function algorithms*, Plenum Press, New York, 1981.
- Billard, L. e Diday, E. (2006). *Symbolic Data Analysis*, John Wiley and Sons, p. 321.
- Bock, H. e Diday, E. (2000). *Analysis of Symbolic Data. Exploratory Methods for Extracting Statistical Information from Complex Data*, Springer-Verlag, Berlin.
- Bodeux, J. B. e Golinval, J. C. (2003). Modal identification and damage detection using the data-driven stochastic subspace and armav methods, *Mechanical Systems and signal processing*, v. 17, n. 1, p. 83-89.
- Brito, P. e Diday, E. (1990). Knowledge, Data and Computer-Assisted Decisions, Springer-Verlag chapter Use of Pyramids in Symbolic Data Analysis, p. 3-16.
- Brito, P. (1994), *New Approaches in Classification and Data Analysis*, Springer-Verlag, chapter Use of Pyramids in Symbolic Data Analysis, p. 378-386.
- Cavalcanti, L. N. J. (2006). *Clusterização baseada em algoritmos fuzzy*, Dissertação de Mestrado, Centro de informática, Universidade Federal de Pernambuco, Recife.
- Cawley, P. e Adams, R. (1979), *The location of defects in structures from measurements of natural frequencies*, *Journal of Strain Analysis*, v. 14, n. 2, p. 49-57.
- Chance, J.; Tomlinson, J. e Worden, K. (1994). A simplified approach to the numerical and experimental modeling of the dynamics of a cracked beam, in: *Proceedings of the International Modal Analysis Conference*, 12, Honolulu, Hawaii, p. 778-785.
- Chun, X.; Weilian, Q. e Dongmei, T. (2005). An application of data fusion technology in structural health monitoring and damage identification, *in Proceedings of Smart Structures and Materials 2005: Smart Sensor Technology and Measurement Systems*, San Diego, Estados Unidos, p. 451-461.

- Cole, H. (1968). On-the-line analysis of random vibrations, AIAA/ASME, *Structural Dynamics and Materials Conference*, 9, p. 69-288.
- Crémona, C. (2004). *Évaluation dynamique des ponts*, Presses LCPC, Paris, France.
- Cury, A. (2010), *Techniques D'anormalité Appliquées à la surveillance de Santé Structurale*, Tese de doutorado, Université Paris-Est, França.
- Cury, A.; Borges, C. e Barbosa, F. (2011). A two-step technique for damage assessment using numerical and experimental vibration data, *Structural Health Monitoring*, v.10, p. 417-428.
- Diday, E. e Noirhomme-Fraiture, M. (2007). *Symbolic Data Analysis and the SODAS Software*, John Wiley and Sons.
- Doebbling, S.; Farrar, C.; Prime, M. e Shevitz, D. (1996). *Damage identification and health monitoring on structural and mechanical systems form changes in their vibration characteristics: a literature review*, Relatório técnico, Los Alamos National Laboratory, Estados Unidos.
- Dugan, D.; Wallance, E. e Caldwell, S. (1980). Measured and predicted vibrational behavior of gulf of Mexico platforms, in: Proceedings of the 12th Annual Offshore Technical Conference, Houston, Estados Unidos, p. 92-100.
- EN1990 (1990). *EN 1990 Eurocode: Basis of structural design*. European Comitee for standadization. Bruxelas.
- Esposito, F.; Malerba, D. e Tamma, V. (2000). Analysis of Symbolic Data: Exploratory Methods for Extracting Statistical Information from Complex Data, Springer-Verlag, chapter Dissimilarity Measures for Symbolic Data, p. 165-185.
- Ewins, D. (2000). *Modal Testing: Theory and Practice*, 2 ed., John Wiley and Sons, p. 562.
- Farrar, C. e Jauregui, D. (1996). Damage detection algorithms applied to experimental and numerical modal data from the i-40 brigde, relatório técnico, Los Alamos National Laboratory report, Estados Unidos.



Fontana, A. e Naldi, M. C. (2009). Estudo de Comparação de Métodos para Estimação de Números de Grupos em Problemas de Agrupamento de Dados, Relatório técnico, ICM, Universidade de São Paulo, São Carlos.

Fox, C. (1992). The location of defects in structures: A comparison of the use of natural frequency and mode shape data, in: Proceeding of the International Modal Analysis Conference, 10, San Diego, Estados Unidos, p. 522-528.

Friswell, M.; Penny, G. e Wilson, D. (1994). Using vibration data and statistical measures to locate damage in structures, *The International Journal of Analytical and Experimental Modal Analysis*, v. 9, n. 4, p. 239-254.

Gudmundson, P. (1982). Eigenfrequency changes of structures due to cracks, notches, or other geometrical changes, *Journal of the Mechanics and Physics of Solids*, v. 30, n. 5, p. 339-353.

Guiera, A. J. A. (2006). *Segmentação por agrupamentos Fuzzy c-means para identificação de linhas de transmissão de energia elétrica em imagens LiDAR*. Dissertação de Mestrado. Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial. Universidade Tecnológica Federal do Paraná. Curitiba.

Guo, H. (2006). Structural damage detection using information fusion technique, *Mechanical Systems and Signal Processing*, v. 20, n. 1, p. 1173-1188.

Haritos, N. e Owen, J. (2004). The use of vibration data for damage detection in bridges: a comparison of system identification and pattern recognition approaches, *Structural Health Monitoring*, v. 3, n. 2, p. 141-163.

Hastie, T.; Tibshirani, R. e Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2 ed., Springer Series in Statistics.

Ibrahim, S. (1977). Random decrement technique for modal identification of structures, *Journal of Spacecraft and Rockets*, v. 14, n. 11, p. 696-700.

Ibrahim, S. (1987). An upper hessenberg sparse matrix algorithm for modal identification on microcomputer, *Journal of sound and vibration*, v. 113, n. 1, p. 47-57.

- Iwasaki, A.; Todoroki, A.; Shimamura, Y. e Kobayashi, H. (2004). An unsupervised statistical damage detection method for structural health monitoring, *Smart Material Structures*, v. 13, p. 80-85.
- Jain, A. K.; Murty, M. N. e Flynn, P. J. (1999). Data Clustering: A Review. *ACM Computing Surveys*, v. 31, n. 31, p. 264-323.
- Kato, M. e Shimada, S. (1986), Vibration of pc bridge during failure process, *ASCE Journal of Structural Engineering*, v. 112, p. 1692-1703.
- Kawiecki, G. (2001), Modal damping measurement for damage detection, *Smart Materials and Structures*, v. 10, n. 3, p. 466-480.
- Kim, J. e Stubbs, N. (1993). Assessment of the relative impact of model uncertainties on the accuracy of global nondestructive damage detection in structures, relatório técnico, New Mexico University, Estados Unidos.
- Kim, S. e Lee, J. (2000). Use of modal testing to identify damage on steel members, *KSCCE Journal of Civil Engineering*, v. 4, n. 2, p. 75-82.
- Klein, L. (1999), *Sensor and Data Fusion: Concepts and Application*, 1ª ed., SPIE Optical Engineering Press.
- Lacerda, W. S. (2006). *Projeto e Implementação de Circuitos Classificadores Digitais com Controle da Generalização Baseado na Regra do Vizinho-mais-próximo Modificada*, tese de doutorado. Escola de Engenharia. Universidade Federal de Minas Gerais. Belo Horizonte.
- Lieven, N. e Ewins, D. (1998). Spatial correlation of mode shapes, in Proceeding of the International Modal Analysis Conference, 6, Kissimmee, Estados Unidos, p. 690-695.
- Lim, T. e Kashangaki, T. (1994). Structural damage detection of space truss structures using best achievable eigenvectors, *AIAA Journal*, v. 32, n. 5, p. 1049-1057.
- Lopes, S. M. O. F. (2008). *Técnicas Geométricas de Condensação para o Classificador k-nn*, dissertação de mestrado. Departamento de Matemática. Universidade de Aveiro.

Matlab® (2012). The Language of Technical Computing, versão R2012b, Mathworks.

Messina, A.; Williams, E. e Contursi, T. (1998). Structural damage detection by a sensitivity and statistical-based method, *Journal of Sound and Vibration*, v. 216, n. 5, p. 791-808.

Mevel, L.; Hermans, L. e Van der Auweraer, H. (1999). Application of a subspace-based fault detection method to industrial structures, *Mechanical systems and signal processing*, v. 13, n. 6, p. 823-838.

Minor, C.; Johnson, K.; Rose-Pehrsson, S.; Owrutsky, J.; Wales, S.; Steinhust, D. e Gottuk, D. (2007). Data fusion with a multisensor system for damage control and situational awareness, in: Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, Londres, Reino Unido, p. 1055-1068.

Ndambi, J.; Vantomme, J. e De Visscher, J. (2000). Modal damping as a damage detection parameter in reinforced concrete structures, Civil-Comp press, Edinburgh, Reino Unido, p. 1-7.

Pandey, A. e Biswas, M. (1994). Damage detection in structures using changes in flexibility, *Journal of Sound and Vibration*, v. 169, n. 1, p. 3-17.

Pandey, A.; Biswas, M. e Samman, M. (1991). Damage detection from changes in curvature mode shapes, *Journal of Sound and Vibration*, v. 145, n. 2, p. 321-332.

Pimentel, E. P.; França, V. F. e Omar, N. (2003). A identificação de grupos de aprendizes no ensino presencial utilizando técnicas de clusterização. Anais do Simpósio Brasileiro de Informática na Educação, Rio de Janeiro.

Reda Taha, M. e Lucero, J. (2005). Damage identification for structural health monitoring using fuzzy pattern recognition, *Engineering Structures*, v. 27, n. 1, p. 1774-1783.

Rytter, A. (1993). *Vibration based inspection of civil engineering structures*, tese de doutorado, University of Aalborg, Dinamarca.

- Salane, H. e Baldwin, J. (1990). Identification of modal properties of bridges, *ASCE Journal of Structural Engineering*, v. 116, n. 7, p. 2008-2021.
- Salawu, O. (1997). Detection of structural damage through changes in frequency: a review, *Engineering Structures*, v.19, n.9, p.718-723.
- Sierra, B. (2000). *Aportaciones metodológicas a la clasificación supervisada*, tese de doutorado, Departamento de Ciências da Computação e Inteligência Artificial, Universidade del Pais Vasco.
- Smith, S. e Beattie, C. (1991). Secant-method adjustment for structural models, *AIAA Journal*, v. 29, n. 1, p. 119-126.
- Su, Z.; Wang, X.; Cheng, L.; Yu, L. e Chen, Z. (2009). On selection of data fusion schemes for structural damage evaluation, *Structural Health Monitoring*, v. 9, n. 3, p. 223-241.
- Teughels, A.; Maeck, J. e De Roeck, G. (2002), Damage assessment by fe model updating using damage functions, *Computers and Structures*, v. 80, n.25, p. 1869-1879.
- Trendafilova, I. e Heylen, W. (2003). Categorisation and pattern recognition methods for damage localisation from vibration measurements, *Mechanical Systems and Signal Processing*, v. 17, n. 4, p. 825-836.
- West, W. (1984), Illustration of the use of modal assurance criterion to detect structural changes in na orbiter test specimen, in: Proceedings of the Air Force Conference on Aircraft Structural Integrity, Estados Unidos, p. 1-6.
- Williams, E.; Contursi, T. e Messina, A. (1996). Damage detection and localisation using natural frequency sensivity, in: Identification in Engineering Systems, Reino Unido, p. 368-376.
- Yan, A. M.; Cheng, L.; Wu, Z. e Yam, L. (2007). Development in vibration-based structural damage detection technique, *Mechanical Systems and Signal Processing*, v. 21, p. 2198-2211.

Yeung, W. e Smith, J. (2005). Damage detection in bridges using neural networks for pattern recognition of vibration signatures, *Engineering Structures*, v. 27, p. 685-698.

Zadeh, L. A. (1965). Fuzzy Sets, *Information and Control*, vol. 3, n. 8, p. 338-353.

Zhang, Z. e Aktan, A. (1995). The damage indices for constructed facilities, in: *Proceedings of the International Modal Analysis Conference*, 13, Nashville, Estados Unidos, p. 1520-1529.

Zhang, Z. e Aktan, A. (2007). Statistical damage identification for bridges using ambiente vibration data, *Computers and structures*, v. 85, p. 476-485.

Zienkiewicz, O. e Taylor, R. (2000). *The Finite Element Method*, 3 ed., Butterworth-Hienemann.

# ANEXO A

## A. Aplicação Numérica.

Neste anexo, os resultados complementares das análises realizadas para o modelo de uma viga bi-apoiada são apresentados. Esses resultados são obtidos a partir dos sinais e dos parâmetros modais transformados em dados simbólicos. Aqui serão mostrados os resultados das diferentes distâncias, critérios de aglomeração e métodos de inicialização que não foram expostos no capítulo 5.

### A.1. Análise simbólica

#### A.1.1. Sinais

Inicialmente, são mostrados os resultados obtidos utilizando os sinais transformados em dados simbólicos para separar os três estados estruturais. A seguir, apresentam-se todos os índices de acerto obtidos aplicando os sinais ao método da aglomeração hierárquica. O método da aglomeração hierárquica é aplicado às distâncias euclidiana, *cityblock*, *minkowski*, *chebychev*, *cosine*, *correlation* e *spearman*. As distâncias citadas são aplicadas com os critérios de agregação *average*, *centroide*, *complete*, *median*, *single*, *ward* e *weighted*.

Na tabela A.1 apresentam-se os índices de acerto utilizando os sinais aplicando o método da aglomeração hierárquica com a distância euclidiana para todos os critérios de agregação.

TABELA A.1 – Índices de acerto (em %) obtidos utilizando os sinais aplicando o método da aglomeração hierárquica com a distância euclidiana.

Aglomeração Hierárquica								
Euclidiana	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	80	40	40	10	40	40
	D1	10	10	80	70	10	80	80
	D2	70	70	30	40	100	30	30
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	40	40	40	40	10	40	80
	D1	10	10	70	70	10	80	30
	D2	100	100	40	40	100	30	40
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	40	40	40	70	10	40	70
D1	10	10	80	10	10	80	10	
D2	100	100	30	90	100	30	90	

Na tabela A.2 são apresentados os índices de acerto do método da aglomeração hierárquica com a distância *cityblock* utilizando os sinais como dados de entrada.

TABELA A.2– Índices de acerto (em %) obtidos utilizando os sinais aplicando o método da aglomeração hierárquica com a distância *cityblock*.

Aglomeração Hierárquica								
Cityblock	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	80	80	70	10	40	40
	D1	10	10	10	10	10	80	60
	D2	70	70	70	90	100	30	50
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	80	80	40	10	40	40
	D1	10	10	10	70	10	80	60
	D2	70	70	70	40	100	30	50
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	80	80	40	10	40	40
D1	10	10	10	60	10	80	60	
D2	70	70	70	50	100	30	50	

A tabela A.3 mostra os índices de acerto utilizando os sinais aplicando o método da aglomeração hierárquica com a distância *minkowski*.

TABELA A.3– Índices de acerto (em %) obtidos utilizando os sinais aplicando o método da aglomeração hierárquica com a distância *minkowski*.

Aglomeração Hierárquica								
<i>Minkowski</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	80	40	40	10	40	40
	D1	10	10	80	70	10	80	80
	D2	70	70	30	40	100	30	30
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	40	40	40	40	10	40	40
	D1	10	10	70	70	10	80	80
	D2	100	100	40	40	100	30	30
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	40	40	40	70	10	40	70
	D1	10	10	80	90	10	80	10
	D2	100	100	30	10	100	30	90

Na tabela A.4 são apresentados os índices de acerto utilizando os sinais aplicando o método da aglomeração hierárquica com a distância *chebychev* para todos os critérios de agregação.

TABELA A.4 – Índices de acerto (em %) obtidos utilizando os sinais aplicando o método da aglomeração hierárquica com a distância *chebychev*.

Aglomeração Hierárquica								
<i>Chebychev</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	70	40	40	40	10	40	60
	D1	10	0	100	0	10	20	0
	D2	90	100	10	100	100	90	90
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	60	50	40	60	10	40	60
	D1	10	10	20	0	10	20	0
	D2	90	100	90	90	100	90	90
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	60	50	70	60	10	60	50



	D1	10	10	90	100	10	60	10
	D2	90	100	10	0	100	40	100

Os próximos índices de acerto mostrados na tabela A.5, são referentes à utilização dos sinais para o método da aglomeração hierárquica, utilizando a distância *cosine*.

TABELA A.5– Índices de acerto (em %) obtidos utilizando os sinais aplicando o método da aglomeração hierárquica com a distância *cosine*.

Aglomeração Hierárquica								
<i>Cosine</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	80	70	80	10	80	80
	D1	10	10	10	10	10	10	10
	D2	70	60	90	60	100	70	60
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	80	70	80	10	70	80
	D1	10	10	10	10	10	10	10
	D2	60	60	90	60	100	90	60
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	70	70	80	90	70	80
	D1	10	90	10	10	10	90	10
	D2	60	10	90	60	10	10	60

Em seguida, a distância *correlation* é considerada na aplicação do método da aglomeração hierárquica para separar os diversos estados estruturais simulados, utilizando as respostas dinâmicas. Os índices de acerto apresentam-se na tabela A.6.

TABELA A.6– Índices de acerto (em %) obtidos utilizando os sinais aplicando o método da aglomeração hierárquica com a distância *correlation*.

Aglomeração Hierárquica								
<i>Correlation</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	80	70	80	90	80	80
	D1	10	10	10	10	10	10	10
	D2	60	60	90	60	10	60	60
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	80	70	80	10	80	80
	D1	10	10	10	10	10	10	10

	D2	60	60	90	60	100	60	60
	R2							
	Average	Centroide	Complete	Median	Single	Ward	Weighted	
	D0	60	60	70	80	90	80	80
	D1	10	10	10	10	10	10	10
	D2	90	90	90	60	10	60	60

A última distância utilizada para o método da aglomeração hierárquica é a *spearman*. Os índices de acerto para os diferentes critérios de agregação são apresentados na tabela A.7.

TABELA A.7– Índices de acerto (em %) obtidos utilizando os sinais aplicando o método da aglomeração hierárquica com a distância *spearman*.

Aglomeração Hierárquica								
<i>Spearman</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	70	70	70	70	10	70	70
	D1	10	10	10	10	10	10	10
	D2	80	80	80	80	100	80	80
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	70	70	70	70	10	70	70
	D1	10	10	10	10	10	10	10
	D2	80	80	80	80	100	80	80
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	70	70	70	70	10	70	70
	D1	10	10	10	10	10	10	10
	D2	80	80	80	90	100	80	90

O segundo método de classificação de dados utilizado é o k-means. Aplicando-se o k-means para a separação dos sinais, as distâncias euclidiana, *cityblock*, *cosine* e *correlation* foram aplicadas para os métodos de inicialização amostral, e as distâncias euclidiana e *correlation* foram aplicadas com os métodos de inicialização uniforme e *cluster*. Na tabela A.8 são apresentados os índices de acerto referentes à aplicação do k-means para os sinais.

TABELA A.8– Índices de acerto (em %) obtidos utilizando os sinais aplicando o método k-means.

		k-means								
		Amostrai			Uniforme			Cluster		
		R0	R1	R2	R0	R1	R2	R0	R1	R2
Euclidiana	D0	40	40	60	40	40	40	40	40	40
	D1	80	80	70	60	60	60	60	60	60
	D2	30	30	30	30	40	30	30	40	30
Cityblock		R0	R1	R2	R0	R1	R2	R0	R1	R2
	D0	50	50	40						
	D1	60	60	60						
Cosine		R0	R1	R2	R0	R1	R2	R0	R1	R2
	D0	40	40	40						
	D1	60	60	60						
Correlation		R0	R1	R2	R0	R1	R2	R0	R1	R2
	D0	40	40	40	40	40	40	40	40	40
	D1	60	60	60	60	60	60	60	60	60
	D2	30	30	30	30	30	30	30	30	30

O último método de classificação aplicado para a separação dos sinais dos três estados estruturais é o c-means. Para essa simulação, apenas foi utilizada a distância euclidiana e o método de inicialização amostral. Os índices de acerto obtidos são mostrados na tabela A.9. A segunda coluna de cada nível de ruído indica o grau de certeza da classificação.

TABELA A.9– Índices de acerto (em %) obtidos utilizando os sinais aplicando o método c-means.

c-means							
		Sinais					
		R0	R1	R2	R0	R1	R2
D0		40	47	40	47	50	49
D1		70	65	70	65	70	64
D2		30	59	40	59	30	58

### A.1.2. Frequências

A utilização da ADS juntamente com as técnicas de *clustering* é aplicada para as frequências próprias previamente identificadas. Como pode ser observada nas aplicações a seguir, as simulações das frequências fornecem 100% de acerto na maioria das classificações.

Primeiramente apresentam-se os índices de acerto obtidos aplicando as frequências ao método da aglomeração hierárquica. Como realizado para os sinais, as diversas distâncias e critérios de agregação foram combinados para a obtenção das classificações utilizando as frequências.

A tabela A.10 mostra os índices de acerto utilizando as frequências aplicando o método da aglomeração hierárquica com a distância euclidiana para todos os critérios de agregação.

TABELA A.10 – Índices de acerto (em %) obtidos utilizando as frequências aplicando o método da aglomeração hierárquica com a distância euclidiana.

Aglomeração Hierárquica								
Euclidiana	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	90	100	100	90
	D1	100	100	100	100	100	100	100
	D2	100	100	100	0	100	100	0

Na tabela A.11 são apresentados os índices de acerto do método da aglomeração hierárquica com a distância *cityblock* utilizando as frequências como dados de entrada.

TABELA A.11 – Índices de acerto (em %) obtidos utilizando as frequências aplicando o método da aglomeração hierárquica com a distância *cityblock*.

Aglomeração Hierárquica								
Cityblock	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
D1	100	100	100	100	100	100	100	
D2	100	100	100	100	100	100	100	

A tabela A.12 mostra os índices de acerto utilizando as frequências aplicando o método da aglomeração hierárquica com a distância *minkowski*.

TABELA A.12 – Índices de acerto (em %) obtidos utilizando as frequências aplicando o método da aglomeração hierárquica com a distância *minkowski*.

Aglomeração Hierárquica								
Minkowski	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	90	100	100	90
D1	100	100	100	100	100	100	100	
D2	100	100	100	0	100	100	0	

Em seguida, a tabela A.13 mostra os índices de acerto obtidos para o método da aglomeração hierárquica com a distância *chebychev* para as frequências.

TABELA A.13 – Índices de acerto (em %) obtidos utilizando as frequências aplicando o método da aglomeração hierárquica com a distância *chebychev*.

Aglomeração Hierárquica								
<i>Chebychev</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	100	10	10	70	10
	D1	100	100	80	100	100	80	100
	D2	10	10	0	10	10	100	10
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	100	10	10	70	10
	D1	100	100	80	100	100	100	100
	D2	0	0	10	0	0	70	0
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	90	20	70	10	90	60	90
	D1	100	100	100	100	100	100	100
	D2	0	0	0	0	0	70	0

Na tabela A.14 são apresentados os índices de acerto utilizando as frequências aplicando o método da aglomeração hierárquica com a distância *cosine* para todos os critérios de agregação.

TABELA A.14 – Índices de acerto (em %) obtidos utilizando as frequências aplicando o método da aglomeração hierárquica com a distância *cosine*.

Aglomeração Hierárquica								
<i>Cosine</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100

	D1	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100

Os próximos índices de acerto apresentados na tabela A.15 são referentes à utilização das frequências para o método da aglomeração hierárquica, utilizando a distância *correlation*.

TABELA A.15 – Índices de acerto (em %) obtidos utilizando as frequências aplicando o método da aglomeração hierárquica com a distância *correlation*.

Aglomeração Hierárquica								
<i>Correlation</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100

Em seguida, a distância *spearman* é considerada na aplicação do método da aglomeração hierárquica para separar os diferentes estados estruturais utilizando as frequências. Os índices de acerto apresentam-se na tabela A.16.

TABELA A.16 – Índices de acerto (em %) obtidos utilizando as frequências aplicando o método da aglomeração hierárquica com a distância *spearman*.

Aglomeração Hierárquica								
<i>Spearman</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	70	70	70	70	10	70	70
	D1	10	10	10	10	10	10	10
	D2	80	80	80	90	100	80	90
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted

	D0	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R2							
	Average	Centroide	Complete	Median	Single	Ward	Weighted	
	D0	90	90	90	90	90	90	100
	D1	100	100	100	100	90	100	100
	D2	100	100	100	100	100	100	100

Os próximos índices de acerto mostrados na tabela A.17 são referentes à aplicação do método k-means para separar os ensaios descritos através das frequências próprias transformadas em dados simbólicos. Para a aplicação do k-means foram utilizadas as distâncias euclidiana, *cityblock*, *cosine* e *correlation* e os métodos de inicialização amostral, uniforme e *cluster*.

TABELA A.17 - Índices de acerto (em %) obtidos utilizando as frequências aplicando o método k-means.

k-means										
		Amostral			Uniforme			Cluster		
		R0	R1	R2	R0	R1	R2	R0	R1	R2
Euclidiana	D0	100	100	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100	100	100
		R0	R1	R2	R0	R1	R2	R0	R1	R2
<i>Cityblock</i>	D0	100	100	100						
	D1	100	100	100						
	D2	100	100	100						
		R0	R1	R2	R0	R1	R2	R0	R1	R2
<i>Cosine</i>	D0	100	100	100						
	D1	100	100	100						
	D2	100	100	100						
		R0	R1	R2	R0	R1	R2	R0	R1	R2
<i>Correlation</i>	D0	100	100	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100	100	100
		R0	R1	R2	R0	R1	R2	R0	R1	R2

O método c-means também é utilizado e os resultados são mostrados na tabela A.18.



TABELA A.18– Índices de acerto (em %) obtidos utilizando as frequências aplicando o método c-means.

c-means						
Frequências						
	R0		R1		R2	
D0	100	85	100	85	100	78
D1	100	87	100	82	100	87
D2	100	83	100	85	100	93

### A.1.3. Primeiro Modo

Os índices de acerto levando em conta a classificação dos primeiros modos próprios aplicando o método da aglomeração hierárquica utilizando a distância euclidiana são apresentados na tabela A.19.

TABELA A.19 - Índices de acerto aplicando o método da aglomeração hierárquica para os primeiros modos próprios utilizando a distância euclidiana.

Aglomeração Hierárquica									
R0									
	Average	Centroide	Complete	Median	Single	Ward	Weighted		
Euclidiana	D0	10	10	0	10	10	80	10	
	D1	90	10	90	10	90	90	90	
	D2	100	100	100	100	100	100	100	
	R1								
		Average	Centroide	Complete	Median	Single	Ward	Weighted	
	D0	20	0	50	100	100	80	50	
	D1	80	10	70	30	10	80	80	
	D2	100	100	100	10	0	100	100	
	R2								
		Average	Centroide	Complete	Median	Single	Ward	Weighted	
	D0	80	100	80	100	10	80	80	
	D1	10	10	60	10	10	40	10	
D2	50	10	0	10	100	60	50		

Os índices de acerto utilizando a distância *cityblock* para a classificação dos primeiros modos, aplicando o método da aglomeração hierárquica, são mostrados na tabela A.20.

TABELA A.20 - Índices de acerto aplicando o método da aglomeração hierárquica para os primeiros modos próprios utilizando a distância *cityblock*.

Aglomeração Hierárquica								
Cityblock	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	10	10	10	80	10
	D1	90	90	90	90	90	90	90
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	50	10	10	30	30
	D1	80	10	50	10	10	80	80
	D2	100	100	100	100	100	90	90
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	10	10	10	70	80
D1	80	10	80	10	10	30	10	
D2	60	100	60	100	100	60	60	

Na tabela A.21 são apresentados os índices de acerto utilizando os primeiros modos próprios aplicando o método da aglomeração hierárquica com a distância *minkowski* para todos os critérios de agregação.

TABELA A.21 - Índices de acerto aplicando o método da aglomeração hierárquica para os primeiros modos próprios utilizando a distância *minkowski*.

Aglomeração Hierárquica								
Minkowski	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	0	10	10	10	10
	D1	90	10	90	10	90	90	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	20	100	50	100	100	80	50
	D1	80	10	70	30	10	80	80
	D2	100	0	100	10	0	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	100	80	100	10	80	80
D1	10	10	0	10	10	40	10	
D2	50	10	60	10	100	60	50	

Os próximos índices de acerto mostrados na tabela A.22 são referentes à utilização dos primeiros modos próprios para o método da aglomeração hierárquica, utilizando a distância *chebychev*.

TABELA A.22 - Índices de acerto aplicando o método da aglomeração hierárquica para os primeiros modos próprios utilizando a distância *chebychev*.

Aglomeração Hierárquica								
<i>Chebychev</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	10	40	10	10	100	0
	D1	10	10	100	10	10	60	20
	D2	100	100	0	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	10	100	100	100	40	100
	D1	10	100	0	10	10	80	10
	D2	0	0	20	0	0	20	0
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	10	10	10	10	70	100
	D1	0	100	100	100	100	20	10
	D2	10	10	10	10	10	70	10

Em seguida, a distância *cosine* é considerada na aplicação do método da aglomeração hierárquica para separar os diferentes estados estruturais utilizando os primeiros modos próprios. Os índices de acerto são mostrados na tabela A.23.

TABELA A.23 - Índices de acerto (em %) obtidos utilizando os primeiros modos próprios aplicando o método da aglomeração hierárquica com a distância *cosine*.

Aglomeração Hierárquica								
<i>Cosine</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	10	10	10	80	10
	D1	90	90	90	90	90	90	90
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	30	10	10	80	0
	D1	80	10	80	100	100	50	80
	D2	100	100	90	0	100	100	90
	R2							

	Average	Centroide	Complete	Median	Single	Ward	Weighted
D0	10	10	10	10	10	70	80
D1	90	10	80	10	10	30	10
D2	60	100	60	100	100	60	60

Os próximos índices de acerto mostrados na tabela A.24, são referentes aos primeiros modos próprios utilizando o método da aglomeração hierárquica utilizando a distância *correlation*.

TABELA A.24 - Índices de acerto (em %) obtidos utilizando os primeiros modos próprios aplicando o método da aglomeração hierárquica com a distância *correlation*.

Aglomeração Hierárquica								
<i>Correlation</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	10	10	10	90	10
	D1	90	90	90	90	90	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	10	10	10	30	10
	D1	100	10	50	100	100	80	80
	D2	90	100	90	0	100	90	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	80	10	10	70	80
	D1	10	10	10	10	10	30	10
	D2	100	100	60	100	100	60	60

Na tabela A.25 são apresentados os índices de acerto utilizando os primeiros modos próprios utilizando a método da aglomeração hierárquica com a distância *spearman*.

TABELA A.25 - Índices de acerto (em %) obtidos utilizando os primeiros modos próprios aplicando o método da aglomeração hierárquica com a distância *spearman*.

Aglomeração Hierárquica								
<i>Spearman</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	10	10	10	80	10
	D1	100	100	10	10	0	100	10
	D2	0	0	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	10	10	10	80	10

	D1	10	100	60	100	100	90	100
	D2	100	0	90	0	0	90	0
	R2							
	Average	Centroide	Complete	Median	Single	Ward	Weighted	
	D0	10	10	10	10	10	80	10
	D1	10	10	10	10	10	10	10
	D2	100	100	100	100	100	50	100

Em seguida, o método da classificação utilizado para a distinção dos ensaios através dos primeiros modos próprios é o k-means. Os índices de acerto para essa aplicação são apresentados na tabela A.26.

TABELA A.26 - Índices de acerto (em %) obtidos utilizando os primeiros modos próprios aplicando o método k-means.

k-means										
		Amostrai			Uniforme			Cluster		
		R0	R1	R2	R0	R1	R2	R0	R1	R2
<i>Euclidiana</i>	D0	80	80	50	100	80	80	100	80	80
	D1	90	80	60	100	100	100	100	100	100
	D2	100	100	60	100	100	80	100	100	80
		R0	R1	R2	R0	R1	R2	R0	R1	R2
<i>Cityblock</i>	D0	90	70	50						
	D1	100	100	30						
	D2	100	50	100						
		R0	R1	R2	R0	R1	R2	R0	R1	R2
<i>Cosine</i>	D0	90	80	90						
	D1	100	100	100						
	D2	100	100	80						
		R0	R1	R2	R0	R1	R2	R0	R1	R2
<i>Correlation</i>	D0	90	90	90	90	90	90	90	90	90
	D1	100	100	100	100	100	100	100	100	100
	D2	100	100	80	100	100	80	100	100	80
		R0	R1	R2	R0	R1	R2	R0	R1	R2

A tabela A.27 mostra os índices de acerto obtidos com o método de classificação c-means.

TABELA A.27 - Índices de acerto (em %) obtidos utilizando os primeiros modos próprios aplicando o método c-means.

	c-means					
	Modos 1					
	R0		R1		R2	
D0	90	75	90	33	70	33
D1	90	81	90	33	0	33
D2	100	85	100	33	60	33

#### A.1.4. Segundo Modo

Agora, os segundos modos próprios transformados em dados simbólicos são utilizados para classificar os diversos estados estruturais. Na primeira abordagem utilizando os segundos modos próprios, o método de aglomeração hierárquica é utilizado com a distância euclidiana. Os índices de acerto obtidos são apresentados na tabela A.28.

TABELA A.28 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica com a distância euclidiana.

Aglomeração Hierárquica								
Euclidiana	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	20	100	100	20	100
	D1	0	0	100	0	0	100	0
	D2	90	90	100	90	90	60	90
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	100	10	10	100	50	10
	D1	100	0	100	100	0	40	100
	D2	90	10	70	0	10	100	90
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	90	100	10	10	100	30	10
	D1	90	0	60	100	10	80	100
	D2	20	10	70	10	10	90	80

Em seguida, os índices de acerto utilizando os segundos modos próprios aplicando o método de aglomeração hierárquica com a distância *cityblock* são apresentados na tabela A.29.

TABELA A.29 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica com a distância *cityblock*.

Aglomeração Hierárquica								
Cityblock	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	20	100	100	100	100
	D1	0	0	100	0	0	0	0
	D2	90	90	100	90	90	70	90
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	90	10	10	10	100	90	90
	D1	10	100	100	100	10	40	20
	D2	70	10	50	50	10	70	70
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	90	100	90	100	10	40	90
D1	20	10	20	10	100	60	70	
D2	90	10	70	10	0	70	20	

Na tabela A.30 são mostrados os índices de acerto utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica com a distância *minkowski*.

TABELA A.30 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica com a distância *minkowski*.

Aglomeração Hierárquica								
Minkowski	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	20	100	100	20	100
	D1	0	0	100	0	0	100	0
	D2	90	90	100	90	90	60	90
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	100	10	10	100	50	10
	D1	100	0	100	0	0	40	100
	D2	90	10	70	100	10	100	90
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	90	100	10	10	100	30	10
D1	20	0	60	100	10	80	100	
D2	90	10	70	10	10	90	80	

Na tabela A.31 são apresentados os índices de acerto utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica com a distância *chebychev* para todos os critérios de agregação.

TABELA A.31 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica com a distância *chebychev*.

Aglomeração Hierárquica								
Chebychev	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	70	10	10	70	10
	D1	100	100	30	100	100	50	100
	D2	100	10	100	10	90	60	60
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	10	10	10	80	10
	D1	100	100	100	100	100	20	100
	D2	10	10	10	10	10	30	10
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	20	10	10	90	10
D1	10	10	90	10	100	20	10	
D2	100	100	30	100	0	50	100	

Os próximos índices de acerto apresentados na tabela A.32 são referentes á utilização dos segundos modos próprios para o método da aglomeração hierárquica, utilizando a distância *cosine*.

TABELA A.32 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica com a distância *cosine*.

Aglomeração Hierárquica								
Cosine	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	10	100	100	100	100
	D1	0	0	100	0	0	0	0
	D2	90	90	100	90	90	70	90
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	10	100	10	100	80	90
	D1	0	100	0	100	10	40	40
	D2	70	10	50	0	10	70	60
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	90	100	80	100	100	40	90
D1	10	10	70	10	10	60	90	
D2	90	10	30	10	10	90	20	



A distância *correlation* é considerada na aplicação do método da aglomeração hierárquica utilizando os segundos modos próprios. A tabela A.33 mostra os índices de acerto obtidos.

TABELA A.33 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica com a distância *correlation*.

Aglomeração Hierárquica								
Correlation	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	0	0	0	0	0	0	0
	D2	90	90	70	90	90	70	90
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	10	90	10	100	80	90
	D1	10	100	30	100	10	40	40
	D2	70	10	90	0	10	70	70
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	90	100	90	90	100	40	90
	D1	10	10	20	10	10	60	20
	D2	90	10	70	10	0	90	90

A última distância utilizada para o método da aglomeração hierárquica utilizando os segundo modos próprios é a *spearman*, onde os índices de acerto são apresentados na tabela A.34.

TABELA A.34 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica com a distância *spearman*.

Aglomeração Hierárquica								
Spearman	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	0	100	0	0	0	0	0
	D2	90	10	90	90	90	70	90
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	90	10	90	0	10	90	90
	D1	30	100	40	30	10	20	20
	D2	90	10	90	90	100	60	60
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	90	100	100	90	90
	D1	10	10	10	10	10	10	10

	D2	10	10	90	10	0	90	90
--	----	----	----	----	----	---	----	----

Agora, o método de classificação de dados utilizado é o k-means para identificar os diversos estados estruturais utilizando os segundos modos próprios. Diferentes distâncias e métodos de inicialização foram utilizados. Os índices de acerto para essa aplicação são apresentados na tabela A.35.

TABELA A.35 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método k-means.

		k-means								
		Amostrai			Uniforme			Cluster		
		R0	R1	R2	R0	R1	R2	R0	R1	R2
Euclidiana	D0	20	60	60	70	60	70	70	60	70
	D1	100	50	40	30	50	60	30	50	60
	D2	60	100	80	100	100	80	100	100	80
Cityblock		R0	R1	R2	R0	R1	R2	R0	R1	R2
	D0	70	20	30						
	D1	30	30	80						
	D2	100	100	50						
Cosine		R0	R1	R2	R0	R1	R2	R0	R1	R2
	D0	70	60	80						
	D1	30	50	20						
	D2	100	100	60						
Correlation		R0	R1	R2	R0	R1	R2	R0	R1	R2
	D0	70	60	80	70	60	20	70	60	80
	D1	30	50	20	30	50	80	30	50	20
	D2	100	100	60	100	100	60	100	100	60

Os índices de acerto aplicando o método c-means para os segundos modos próprios são apresentados na tabela A.36.

TABELA A.36 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método c-means.

		c-means					
		Modos 2					
		R0		R1		R2	
D0		10	81	80	33	10	33
D1		100	80	10	33	80	33
D2		70	58	100	33	100	33

### A.1.5. Terceiro Modo

Em seguida, os terceiros modos próprios transformados em dados simbólicos são utilizados juntamente com os métodos de classificação a fim de distinguir os diferentes estados estruturais.

Na tabela A.37 são apresentados os índices de acerto utilizando os terceiros modos próprios aplicando o método de aglomeração hierárquica com a distância euclidiana para todos os critérios de agregação.

TABELA A.37 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica com a distância euclidiana.

Aglomeração Hierárquica								
Euclidiana	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	80	0	80	20	20	90	70
	D2	100	10	10	10	10	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	90	100	10	50	30
	D1	20	10	20	10	100	70	20
	D2	40	100	60	10	0	70	80

Na tabela A.38 são apresentados os índices de acerto obtidos aplicando o método da aglomeração hierárquica com a distância *cityblock* utilizando os terceiros modos próprios como dados de entrada.

TABELA A.38 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica com a distância *cityblock*.

Aglomeração Hierárquica								
<i>Cityblock</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100

	D1	10	10	10	10	10	80	80
	D2	100	0	80	0	100	100	10
R2								
	Average	Centroide	Complete	Median	Single	Ward	Weighted	
	D0	100	100	100	100	100	80	80
	D1	10	10	40	10	10	60	20
	D2	90	0	90	10	0	70	80

Na tabela A.39 são apresentados os índices de acerto utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica com a distância *minkowski*.

TABELA A.39 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica com a distância *minkowski*.

Aglomeração Hierárquica								
<i>Minkowski</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100
	D2	100	100	70	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	100	0	10	10	10	90	70
	D2	80	10	80	20	20	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	90	100	10	50	30
	D1	20	10	20	10	100	70	20
	D2	100	100	60	10	0	70	80

Em seguida, a distância *chebychev* é utilizada na aplicação do método da aglomeração hierárquica para os terceiros modos próprios. A tabela A.40 mostra os índices de acerto obtidos para essa simulação.

TABELA A.40 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica com a distância *chebychev*.

Aglomeração Hierárquica								
<i>Chebychev</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	90	100
	D1	100	10	40	10	100	100	100
	D2	100	10	100	10	100	100	100

R1							
	Average	Centroide	Complete	Median	Single	Ward	Weighted
D0	20	100	20	100	20	20	20
D1	50	10	40	10	90	100	50
D2	70	10	80	10	10	20	70
R2							
	Average	Centroide	Complete	Median	Single	Ward	Weighted
D0	10	10	90	10	10	100	10
D1	100	100	20	100	100	20	100
D2	0	0	0	0	0	20	0

Na tabela A.41 são apresentados os índices de acerto utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica com a distância *cosine* para todos os critérios de agregação.

TABELA A.41 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica com a distância *cosine*.

Aglomeração Hierárquica							
R0							
	Average	Centroide	Complete	Median	Single	Ward	Weighted
D0	100	100	100	100	100	100	100
D1	100	100	100	100	100	100	100
D2	100	100	70	100	100	100	100
R1							
	Average	Centroide	Complete	Median	Single	Ward	Weighted
D0	100	100	100	100	100	100	100
D1	10	10	20	10	10	80	100
D2	100	0	50	0	100	100	90
R2							
	Average	Centroide	Complete	Median	Single	Ward	Weighted
D0	90	100	100	100	100	100	90
D1	10	10	40	10	10	40	10
D2	80	0	80	20	0	80	80

Os próximos índices de acerto apresentados na tabela A.42 são referentes à utilização dos terceiros modos próprios para o método da aglomeração hierárquica, utilizando a distância *correlation*.

TABELA A.42 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica com a distância *correlation*.

Aglomeração Hierárquica								
Correlation	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	10	10	10	10	10	80	100
	D2	100	0	50	0	0	100	90
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	90	10	100	100	20
	D1	10	10	20	10	10	80	10
	D2	80	0	80	100	0	40	100

Em seguida, a distância *spearman* é considerada na aplicação do método da aglomeração hierárquica para separar os diferentes estados estruturais utilizando os terceiros modos próprios. Os índices de acerto são apresentados na tabela A.43.

TABELA A.43 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica com a distância *spearman*.

Aglomeração Hierárquica								
Spearman	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D1	10	10	20	10	10	100	10
	D2	100	0	80	0	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	90	100	100	90	90
	D1	20	10	20	10	10	20	20
	D2	0	0	90	0	0	90	10

O próximo método de classificação a ser aplicado para os terceiros modos próprios transformados em dados simbólicos é o método k-means. A tabela A.44 mostra os índices de acerto obtidos.

TABELA A.44 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método k-means.

		k-means								
		Amostrai			Uniforme			Cluster		
		R0	R1	R2	R0	R1	R2	R0	R1	R2
Euclidiana	D0	100	100	70	100	100	90	100	100	90
	D1	100	80	60	100	100	80	100	100	80
	D2	100	90	80	100	90	40	100	90	40
Cityblock		R0	R1	R2	R0	R1	R2	R0	R1	R2
	D0	100	100	100						
	D1	100	40	30						
Cosine		R0	R1	R2	R0	R1	R2	R0	R1	R2
	D0	100	100	90						
	D1	100	100	70						
Correlation		R0	R1	R2	R0	R1	R2	R0	R1	R2
	D0	100	100	90	100	100	90	100	100	90
	D1	100	100	70	100	100	70	100	100	70
	D2	100	90	40	100	90	40	100	90	40

O método c-means é utilizado e os resultados são apresentados na tabela A.45.

TABELA A.45 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método c-means.

		c-means					
		Modos 3					
		R0		R1		R2	
D0		100	97	100	33	100	33
D1		100	83	0	33	0	33
D2		100	75	100	33	100	33

## A.2. Estudos complementares

A seguir, serão mostrados os resultados complementares das análises realizadas considerando apenas os ensaios referentes à estrutura intacta (D0) e os ensaios com nível de dano (D2).

### A.2.1. Sinais

Na tabela A.46, são mostrados os resultados obtidos utilizando os sinais transformados em dados simbólicos aplicando o método da aglomeração hierárquica com a distância euclidiana.

TABELA A.46 – Índices de acerto (em %) obtidos utilizando os sinais aplicando o método da aglomeração hierárquica com a distância euclidiana.

Aglomeração Hierárquica								
Euclidiana	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	40	40	70	80	10	80	80
	D2	100	100	90	60	100	70	60
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	40	40	70	80	10	80	80
	D2	100	100	90	60	100	70	60
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	40	70	70	10	80	70
	D2	60	100	90	90	100	70	90

Na tabela A.47 são apresentados os índices de acerto do método da aglomeração hierárquica com a distância *cityblock* utilizando os sinais como dados de entrada.

TABELA A.47 – Índices de acerto (em %) obtidos utilizando os sinais aplicando o método da aglomeração hierárquica com a distância *cityblock*.

Aglomeração Hierárquica								
<i>Cityblock</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	80	80	40	10	80	80
	D2	60	60	60	100	100	60	60



R1							
	Average	Centroide	Complete	Median	Single	Ward	Weighted
D0	80	80	80	40	10	80	40
D2	60	60	60	100	100	60	100
R2							
	Average	Centroide	Complete	Median	Single	Ward	Weighted
D0	80	80	80	80	10	70	80
D2	60	60	60	60	100	90	60

Na tabela A.48 são mostrados os índices de acerto utilizando os sinais aplicando o método da aglomeração hierárquica com a distância *minkowski*.

TABELA A.48 – Índices de acerto (em %) obtidos utilizando os sinais aplicando o método da aglomeração hierárquica com a distância *minkowski*.

Aglomeração Hierárquica								
<i>Minkowski</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	40	40	70	80	10	80	80
	D2	100	100	90	60	100	70	60
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	40	40	70	80	10	80	80
	D2	100	100	90	60	100	70	60
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	40	70	70	10	80	70
	D2	60	100	90	90	100	70	90

Na tabela A.49 são apresentados os índices de acerto utilizando os sinais aplicando o método da aglomeração hierárquica com a distância *chebychev* para todos os critérios de agregação.

TABELA A.49– Índices de acerto (em %) obtidos utilizando os sinais aplicando o método da aglomeração hierárquica com a distância *chebychev*.

Aglomeração Hierárquica								
<i>Chebychev</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	70	70	60	70	10	60	70
	D2	90	90	90	90	100	90	90
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted

	D0	70	70	60	70	10	60	70
	D2	90	90	90	90	100	90	90
R2								
	Average	Centroide	Complete	Median	Single	Ward	Weighted	
	D0	70	70	70	70	10	70	70
	D2	90	90	90	90	100	90	90

Os próximos índices de acerto apresentados na tabela A.50 são referentes à utilização dos sinais para o método da aglomeração hierárquica, utilizando a distância *cosine*.

TABELA A.50– Índices de acerto (em %) obtidos utilizando os sinais aplicando o método da aglomeração hierárquica com a distância *cosine*.

Aglomeração Hierárquica								
<i>Cosine</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	70	70	70	80	90	70	80
	D2	90	90	90	60	10	90	60
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	70	70	70	80	90	70	80
	D2	90	90	90	60	10	90	60
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	70	70	70	80	90	70	80
	D2	90	90	90	60	10	90	60

Em seguida, a distância *correlation* é considerada na aplicação do método da aglomeração hierárquica para separar os diferentes estados estruturais utilizando as respostas dinâmicas. Os índices de acerto são apresentados na tabela A.51.

TABELA A.51– Índices de acerto (em %) obtidos utilizando os sinais aplicando o método da aglomeração hierárquica com a distância *correlation*.

Aglomeração Hierárquica								
<i>Correlation</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	70	70	70	70	90	70	80
	D2	90	90	90	90	10	90	60
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	70	70	70	80	90	70	80
	D2	90	90	90	60	10	90	60

		R2						
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	70	70	70	70	90	70	80
	D2	90	90	90	90	10	90	60

A última distância utilizada para o método da aglomeração hierárquica é a *spearman*, onde os índices de acerto para os diferentes critérios de agregação são apresentados na tabela A.52.

TABELA A.52– Índices de acerto (em %) obtidos utilizando os sinais aplicando o método da aglomeração hierárquica com a distância *spearman*.

Aglomeração Hierárquica									
		R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted	
<i>Spearman</i>	D0	70	70	80	70	10	70	70	
	D2	90	90	70	80	100	90	80	
			R1						
			Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	70	70	70	70	10	70	70	
	D2	90	90	90	90	100	90	90	
			R2						
			Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	70	70	70	70	10	70	70	
	D2	90	90	90	90	100	90	90	

Na tabela A.53 são apresentados os índices de acerto referentes à aplicação do k-means para os sinais.

TABELA A.53– Índices de acerto (em %) obtidos utilizando os sinais aplicando o método k-means.

k-means				
		Sinais		
		Amostrai	Uniforme	Cluster
Euclidiana	D0	70	70	70
	D2	90	90	90
<i>Cityblock</i>	D0	80	80	80
	D2	70	70	70
<i>Cosine</i>	D0	70	70	70
	D2	90	90	90
<i>Correlation</i>	D0	70	70	70
	D2	90	90	90

O último método de classificação aplicado para a separação dos sinais dos dois estados estruturais é o c-means. Os índices de acerto obtidos são mostrados na tabela A.54.

TABELA A.54– Índices de acerto (em %) obtidos utilizando os sinais aplicando o método c-means.

c-means		
	Sinais	
D0	70	67
D2	90	79

### A.2.2. Frequências

Primeiramente são apresentados os índices de acerto obtidos aplicando as frequências ao método da aglomeração hierárquica. As diversas distâncias e critérios de agregação foram combinados para a obtenção das classificações utilizando as frequências.

Na tabela A.55 são apresentados os índices de acerto utilizando as frequências aplicando o método da aglomeração hierárquica com a distância euclidiana para todos os critérios de agregação.

TABELA A.55 – Índices de acerto (em %) obtidos utilizando as frequências aplicando o método da aglomeração hierárquica com a distância euclidiana.

Aglomeração Hierárquica								
Euclidiana	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100

Na tabela A.56 são apresentados os índices de acerto do método da aglomeração hierárquica com a distância *cityblock* utilizando as frequências como dados de entrada.

TABELA A.56 – Índices de acerto (em %) obtidos utilizando as frequências aplicando o método da aglomeração hierárquica com a distância *cityblock*.

Aglomeração Hierárquica								
Cityblock	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
D0	100	100	100	100	100	100	100	
D2	100	100	100	100	100	100	100	

Na tabela A.57 são apresentados os índices de acerto utilizando as frequências aplicando o método da aglomeração hierárquica com a distância *minkowski*.

TABELA A.57 – Índices de acerto (em %) obtidos utilizando as frequências aplicando o método da aglomeração hierárquica com a distância *minkowski*.

Aglomeração Hierárquica								
Minkowski	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
D0	100	100	100	100	100	100	100	
D2	100	100	100	100	100	100	100	

Em seguida, na tabela A.58, são apresentados os índices de acerto obtidos para o método da aglomeração hierárquica com a distância *chebychev* para as frequências.

TABELA A.58 – Índices de acerto (em %) obtidos utilizando as frequências aplicando o método da aglomeração hierárquica com a distância *chebychev*.

Aglomeração Hierárquica								
Chebychev	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	10	10	10	10	70	80
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	10	10	80	70	80
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	50	100	100	100	100
	D2	100	100	100	100	100	100	100

Na tabela A.59 são apresentados os índices de acerto utilizando as frequências aplicando o método da aglomeração hierárquica com a distância *cosine* para todos os critérios de agregação.

TABELA A.59 – Índices de acerto (em %) obtidos utilizando as frequências aplicando o método da aglomeração hierárquica com a distância *cosine*.

Aglomeração Hierárquica								
Cosine	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100

Os próximos índices de acerto apresentados na tabela A.60 são referentes à utilização das frequências para o método da aglomeração hierárquica, utilizando a distância *correlation*.

TABELA A.60 – Índices de acerto (em %) obtidos utilizando as frequências aplicando o método da aglomeração hierárquica com a distância *correlation*.

Aglomeração Hierárquica								
<i>Correlation</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100

Em seguida, a distância *spearman* é considerada na aplicação do método da aglomeração hierárquica para separar os diferentes estados estruturais utilizando as frequências. Os índices de acerto são apresentados na tabela A.61.

TABELA A.61 – Índices de acerto (em %) obtidos utilizando as frequências aplicando o método da aglomeração hierárquica com a distância *spearman*.

Aglomeração Hierárquica								
<i>Spearman</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100

Os próximos índices de acerto mostrados na tabela A.62 são referentes à aplicação do método k-means.

TABELA A.62 - Índices de acerto (em %) obtidos utilizando as frequências aplicando o método k-means.

k-means			
	Frequências		
	Amostral	Uniforme	Cluster
D0	100	100	100
D2	100	100	100
D0	100	100	100
D2	100	100	100
D0	100	100	100
D2	100	100	100
D0	100	100	100
D2	100	100	100

O método c-means também é utilizado e os resultados são apresentados na tabela A.63.

TABELA A.63- Índices de acerto (em %) obtidos utilizando as frequências aplicando o método c-means.

c-means	
Frequências	
100	89
100	97

### A.2.3. Primeiro Modo

Os índices de acerto levando em conta a classificação dos primeiros modos próprios aplicando o método da aglomeração hierárquica utilizando a distância euclidiana são apresentados na tabela A.64.

TABELA A.64 - Índices de acerto aplicando o método da aglomeração hierárquica para os primeiros modos próprios utilizando a distância euclidiana.

Aglomeração Hierárquica								
Euclidiana	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	10	80	10	10	80	80
	D2	100	100	100	100	100	100	100
	R1							
	Average	Centroide	Complete	Median	Single	Ward	Weighted	
D0	80	10	80	100	10	80	80	



	D2	100	100	100	10	100	100	100
	R2							
	Average	Centroide	Complete	Median	Single	Ward	Weighted	
	D0	10	100	90	100	10	90	90
	D2	100	10	40	10	100	60	50

Os índices de acerto utilizando a distância *cityblock* para a classificação dos primeiros modos aplicando o método da aglomeração hierárquica são mostrados na tabela A.65.

TABELA A.65 - Índices de acerto aplicando o método da aglomeração hierárquica para os primeiros modos próprios utilizando a distância *cityblock*.

Aglomeração Hierárquica								
<i>Cityblock</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	80	10	10	80	10
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	50	10	10	80	10
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	10	10	10	80	10
	D2	100	100	100	100	100	60	100

Na tabela A.66 são apresentados os índices de acerto utilizando os primeiros modos próprios aplicando o método da aglomeração hierárquica com a distância *minkowski* para todos os critérios de agregação.

TABELA A.66 - Índices de acerto aplicando o método da aglomeração hierárquica para os primeiros modos próprios utilizando a distância *minkowski*.

Aglomeração Hierárquica								
<i>Minkowski</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	10	80	10	10	80	80
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	10	80	100	10	80	80
	D2	100	100	100	10	100	100	100
	R2							

		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	100	90	100	10	90	90
	D2	100	10	40	10	100	60	50

Os próximos índices de acerto apresentados na tabela A.67 são referentes à utilização dos primeiros modos próprios para o método da aglomeração hierárquica, utilizando a distância *chebychev*.

TABELA A.67 - Índices de acerto aplicando o método da aglomeração hierárquica para os primeiros modos próprios utilizando a distância *chebychev*.

Aglomeração Hierárquica								
<i>Chebychev</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	40	10	10	100	10
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	100	10	10	40	10
	D2	100	100	10	100	100	80	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	10	70	100
	D2	10	10	10	10	100	70	10

Em seguida, a distância *cosine* é considerada na aplicação do método da aglomeração hierárquica para separar os diferentes estados estruturais utilizando os primeiros modos próprios. Os índices de acerto são apresentados na tabela A.68.

TABELA A.68 - Índices de acerto (em %) obtidos utilizando os primeiros modos próprios aplicando o método da aglomeração hierárquica com a distância *cosine*.

Aglomeração Hierárquica								
<i>Cosine</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	80	10	10	80	10
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	60	10	10	80	10
	D2	100	100	100	100	100	100	100
	R2							

	Average	Centroide	Complete	Median	Single	Ward	Weighted
D0	10	10	90	10	10	80	10
D2	100	100	60	100	100	60	100

Os próximos índices de acerto apresentados na tabela A.69, são referentes aos primeiros modos próprios utilizando o método da aglomeração hierárquica utilizando a distância *correlation*.

TABELA A.69 - Índices de acerto (em %) obtidos utilizando os primeiros modos próprios aplicando o método da aglomeração hierárquica com a distância *correlation*.

Aglomeração Hierárquica								
<i>Correlation</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	80	10	10	80	10
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	60	10	10	80	10
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	90	10	10	80	10
	D2	100	100	60	100	100	60	100

Na tabela A.70 são apresentados os índices de acerto utilizando os primeiros modos próprios utilizando a método da aglomeração hierárquica com a distância *spearman*.

TABELA A.70 - Índices de acerto (em %) obtidos utilizando os primeiros modos próprios aplicando o método da aglomeração hierárquica com a distância *spearman*.

Aglomeração Hierárquica								
<i>Spearman</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	10	10	10	80	10
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	70	10	10	90	10
	D2	100	100	90	100	100	90	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	10	10	10	80	10

	D2	100	100	100	100	100	50	100
--	----	-----	-----	-----	-----	-----	----	-----

Em seguida, o método da classificação utilizado para a distinção dos ensaios através dos primeiros modos próprios é o k-means. Os índices de acerto para essa aplicação são apresentados na tabela A.71.

TABELA A.71 - Índices de acerto (em %) obtidos utilizando os primeiros modos próprios aplicando o método k-means.

k-means				
		Modos 1		
		Amostrai	Uniforme	Cluster
Euclidiana	D0	80	80	80
	D2	60	60	60
<i>Cityblock</i>	D0	100	100	100
	D2	40	40	40
<i>Cosine</i>	D0	80	80	80
	D2	60	60	60
<i>Correlation</i>	D0	80	80	80
	D2	60	60	60

Os índices de acerto obtidos com o método de classificação c-means são mostrados na tabela A.72.

TABELA A.72 - Índices de acerto (em %) obtidos utilizando os primeiros modos próprios aplicando o método c-means.

c-means	
Modos 1	
80	50
100	50

#### A.2.4. Segundo modo

Na primeira abordagem utilizando os segundos modos próprios, o método de aglomeração hierárquica é utilizado com a distância euclidiana. Os índices de acerto obtidos são apresentados na tabela A.73.

TABELA A.73 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica com a distância euclidiana.

Aglomeração Hierárquica								
Euclidiana	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	80	100
	D2	100	100	100	10	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	100	80	10	10	80	90
	D2	100	10	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	90	100	90	100	100	90	90
	D2	100	10	100	10	10	100	10

Em seguida, os índices de acerto utilizando os segundos modos próprios aplicando o método de aglomeração hierárquica com a distância *cityblock* são apresentados na tabela A.74.

TABELA A.74 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica com a distância *cityblock*.

Aglomeração Hierárquica								
Cityblock	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	90	10	10	80	10
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	90	100	100	90	100
	D2	30	10	70	10	10	100	30

Na tabela A.75 são mostrados os índices de acerto utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica com a distância *minkowski*.

TABELA A.75 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica com a distância *minkowski*.

Aglomeração Hierárquica								
<i>Minkowski</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	80	100
	D2	100	100	100	10	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	80	100	80	10	10	80	90
	D2	100	10	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
D0	90	100	90	100	100	90	90	
D2	100	10	100	10	10	100	10	

Na tabela A.76 são apresentados os índices de acerto utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica com a distância *chebychev* para todos os critérios de agregação.

TABELA A.76 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica com a distância *chebychev*.

Aglomeração Hierárquica								
<i>Chebychev</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	90	10	10	80	10
	D2	100	100	10	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	100	10	100	100	40	10
	D2	100	10	100	10	10	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
D0	10	10	100	10	20	90	10	
D2	100	100	30	100	100	50	100	

Os próximos índices de acerto, apresentados na tabela A.77, são referentes á utilização dos segundos modos próprios para o método da aglomeração hierárquica, utilizando a distância *cosine*.

TABELA A.77 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica com a distância *cosine*.

Aglomeração Hierárquica								
Cosine	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	90	10	10	80	10
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
D0	100	100	90	100	100	90	100	
D2	30	10	60	10	10	100	30	

A distância *correlation* é considerada na aplicação do método da aglomeração hierárquica utilizando os segundos modos próprios, onde seus índices são mostrados na tabela A.78.

TABELA A.78 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica com a distância *correlation*.

Aglomeração Hierárquica								
Correlation	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	100	10	10	80	10
	D2	100	100	50	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
D0	100	100	100	100	100	90	100	
D2	30	10	30	10	10	100	30	

A última distância utilizada para o método da aglomeração hierárquica utilizando os segundo modos próprios, é a *spearman*. Os índices de acerto são apresentados na tabela A.79.

TABELA A.79 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica com a distância *spearman*.

Aglomeração Hierárquica								
<i>Spearman</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	10	10	100	10	10	100	100
	D2	100	100	90	100	100	60	60
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	90	100	100	90	100
	D2	10	10	90	10	10	90	10

Agora, o método de classificação de dados utilizado é o k-means para identificar os diferentes estados estruturais utilizando os segundos modos próprios. Diversas distâncias e métodos de inicialização foram utilizados. Os índices de acerto para essa aplicação são apresentados na tabela A.80.

TABELA A.80 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método k-means.

k-means				
		Modos 2		
		Amostrai	Uniforme	Cluster
Euclidiana	D0	90	90	90
	D2	100	100	100
<i>Cityblock</i>	D0	50	50	50
	D2	100	100	100
<i>Cosine</i>	D0	90	90	90
	D2	100	100	100
<i>Correlation</i>	D0	90	90	90
	D2	100	100	100

Os índices de acerto aplicando o método c-means para os segundos modos próprios são apresentados na tabela A.81.



TABELA A.81 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método c-means.

c-means		
Modos 2		
D0	90	50
D2	100	50

### A.2.5. Terceiro Modo

Em seguida, os terceiros modos próprios transformados em dados simbólicos são utilizados juntamente com os métodos de classificação a fim de distinguir os diferentes estados estruturais.

Na tabela A.82 são apresentados os índices de acerto utilizando os terceiros modos próprios aplicando o método de aglomeração hierárquica com a distância euclidiana para todos os critérios de agregação.

TABELA A.82 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica com a distância euclidiana.

Aglomeração Hierárquica								
Euclidiana	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	10	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	10	100	100
	D2	90	10	20	10	100	80	80

Na tabela A.83 são apresentados os índices de acerto obtidos aplicando o método da aglomeração hierárquica com a distância *cityblock* utilizando os terceiros modos próprios como dados de entrada.

TABELA A.83 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica com a distância *cityblock*.

Aglomeração Hierárquica								
Cityblock	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	10	100	100	100	100	100
	D2	90	100	60	10	10	90	90

Na tabela A.84 são apresentados os índices de acerto utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica com a distância *minkowski*.

TABELA A.84 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica com a distância *minkowski*.

Aglomeração Hierárquica								
Minkowski	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	10	100	100
	D2	90	10	20	10	100	80	80

Em seguida, a distância *chebychev* é utilizada na aplicação do método da aglomeração hierárquica para os terceiros modos próprios. Os índices de acerto são mostrados na tabela A.85.

TABELA A.85 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica com a distância *chebychev*.

Aglomeração Hierárquica								
<i>Chebychev</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	10	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	10	100	10	10	100	100
	D2	30	100	20	100	100	30	30
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	10	100	10	10	100	100
	D2	30	100	20	100	100	20	30

Na tabela A.86 são apresentados os índices de acerto utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica com a distância *cosine* para todos os critérios de agregação.

TABELA A.86 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica com a distância *cosine*.

Aglomeração Hierárquica								
<i>Cosine</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	10	100	100	100	100	20
	D2	80	100	40	10	80	80	100

Os próximos índices de acerto apresentados na tabela A.87 são referentes á utilização dos terceiros modos próprios para o método da aglomeração hierárquica, utilizando a distância *correlation*.

TABELA A.87 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica com a distância *correlation*.

Aglomeração Hierárquica								
<i>Correlation</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	10	100	100	100	100	20
	D2	80	100	90	10	80	80	100

Em seguida, a distância *spearman* é considerada na aplicação do método da aglomeração hierárquica para separar os diferentes estados estruturais utilizando os terceiros modos próprios. Os índices de acerto são apresentados na tabela A.88.

TABELA A.88 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica com a distância *spearman*.

Aglomeração Hierárquica								
<i>Spearman</i>	R0							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	100	100	100	100	100
	R1							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	100	100	80	100	100	100	100
	R2							
		Average	Centroide	Complete	Median	Single	Ward	Weighted
	D0	100	100	100	100	100	100	100
	D2	10	10	40	20	10	90	90

O próximo método de classificação a ser aplicado para os terceiros modos próprios transformados em dados simbólicos é o método k-means. Os índices de acerto podem ser observados na tabela A.89.

TABELA A.89 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método k-means.

		k-means		
		Modos 3		
		Amostral	Uniforme	Cluster
Euclidiana	D0	100	100	100
	D2	90	90	90
<i>Cityblock</i>	D0	50	50	50
	D2	100	100	100
<i>Cosine</i>	D0	100	100	100
	D2	90	90	90
<i>Correlation</i>	D0	100	100	100
	D2	90	90	90

O método c-means é utilizado e os resultados são apresentados na tabela A.90. na segunda coluna apresentam-se os graus de certeza das classificações.

TABELA A.90 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método c-means.

c-means	
Modos 3	
90	50
90	50



# ANEXO B

## B. Viaduto PK 075+317

Os resultados complementares das análises utilizando os ensaios registrados a partir das campanhas de medições realizadas para o viaduto PK 075+317 são apresentados a seguir. Esses resultados são obtidos a partir dos sinais e dos parâmetros modais transformados em dados simbólicos. Aqui serão mostrados os resultados das diferentes distâncias, critérios de aglomeração e métodos de inicialização que não foram expostos no capítulo 6.

### B.1. Análise simbólica

#### B.1.1. Sinais

Inicialmente, são mostrados os resultados obtidos utilizando os sinais transformados em dados simbólicos para separar os três estados estruturais (antes, durante e depois do reforço estrutural). A seguir são apresentados todos os índices de acerto obtidos aplicando os sinais ao método da aglomeração hierárquica. O método da aglomeração hierárquica é aplicado utilizando as distâncias euclidiana, *cityblock*, *minkowski*, *chebychev*, *cosine*, *correlation* e *spearman*. As distâncias citadas são aplicadas com os critérios de agregação *average*, *centroide*, *complete*, *median*, *single*, *ward* e *weighted*.

Na tabela B.1 são apresentados os índices de acerto utilizando os sinais aplicando o método da aglomeração hierárquica.

TABELA B.1 – Índices de acerto (em %) obtidos utilizando os sinais aplicando o método da aglomeração hierárquica.

Aglomeração Hierárquica							
	Average	Centroid	Complete	Median	Single	Ward	Weighted
	Euclidiana						
Antes	100	100	100	100	100	80	80
Durante	0	0	0	0	0	38,46	38,46
Depois	30,76	15,38	15,38	30,76	30,76	7,69	7,69

	<i>Cityblock</i>						
Antes	80	100	80	100	100	80	80
Durante	38,46	0	38,46	0	0	38,46	38,46
Depois	7,69	7,69	7,69	15,38	7,69	7,69	7,69
	<i>Minkowski</i>						
Antes	100	100	100	100	100	80	80
Durante	0	0	0	0	0	38,46	38,46
Depois	30,76	15,38	15,38	30,76	30,76	7,69	7,69
	<i>Chebychev</i>						
Antes	100	100	100	100	100	66,66	100
Durante	0	0	0	0	0	76,92	0
Depois	30,76	30,76	30,76	30,76	30,76	30,76	30,76
	<i>Cosine</i>						
Antes	100	100	60	100	100	60	100
Durante	0	0	100	0	0	100	0
Depois	30,76	30,76	23,07	46,15	30,76	7,69	30,76
	<i>Correlation</i>						
Antes	100	100	100	66,66	100	80	60
Durante	0	0	0	76,92	0	69,23	100
Depois	30,76	30,76	30,76	7,69	30,76	30,76	7,69
	<i>Spearman</i>						
Antes	100	100	80	100	100	100	100
Durante	0	0	84,61	0	0	0	0
Depois	30,76	30,76	7,69	30,76	30,76	30,76	30,76

O segundo método de classificação de dados utilizado é o k-means. Aplicando o k-means para a separação dos sinais, as distâncias euclidiana, *cityblock*, *cosine* e *correlation* foram aplicadas para o método de inicialização amostral, e a distância euclidiana e *correlation* foram aplicadas com os métodos de inicialização *uniforme* e *cluster*. Na tabela B.2 são apresentados os índices de acerto referentes à aplicação do k-means para os sinais.

TABELA B.2– Índices de acerto (em %) obtidos utilizando os sinais aplicando o método k-means.

		k-means		
		Amostral	Uniforme	Cluster
		Sinais		
Euclidiana	Antes	73,33	73,33	73,33
	Durante	76,92	76,92	76,92
	Depois	7,69	7,69	7,69
<i>Cityblock</i>	Antes	80	80	80
	Durante	38,46	38,46	38,46
	Depois	61,53	61,53	61,53



<i>Cosine</i>	Antes	73,33	73,33	73,33
	Durante	76,92	76,92	76,92
	Depois	30,76	30,76	30,76
<i>Correlation</i>	Antes	73,33	73,33	73,33
	Durante	84,61	84,61	84,61
	Depois	30,76	30,76	30,76

O terceiro método de classificação aplicado para a separação dos sinais dos três estados estruturais é o c-means. Apenas foi utilizado a distância euclidiana e o método de inicialização amostral. Os índices de acerto são mostrados na tabela B.3.

TABELA B.3– Índices de acerto (em %) obtidos utilizando os sinais aplicando o método c-means.

c-means		
	Sinais	
Antes	66,67	60
Durante	38,46	42
Depois	53,84	41

### B.1.2. Frequências

A utilização da ADS juntamente com as técnicas de *clustering* é aplicada para as frequências próprias previamente identificadas.

Na tabela B.4 são apresentados os índices de acerto utilizando as frequências aplicando o método da aglomeração hierárquica.

TABELA B.4 – Índices de acerto (em %) obtidos utilizando as frequências aplicando o método da aglomeração hierárquica.

Aglomeração Hierárquica							
	Average	Centroid	Complete	Median	Single	Ward	Weighted
<i>Euclidiana</i>							
Antes	13,33	6,66	0	6,66	6,66	61,53	7,69
Durante	100	100	69,23	100	100	92,30	100
Depois	7,69	7,69	93,33	0	7,69	46,66	60
<i>Cityblock</i>							
Antes	84,61	0	86,66	100	6,66	93,33	84,61
Durante	100	100	92,30	7,69	100	92,30	92,30
Depois	92,30	7,69	53,84	6,66	7,69	84,61	93,33
<i>Minkowski</i>							

Antes	13,33	6,66	0	6,66	6,66	61,53	7,69
Durante	100	100	69,23	100	100	92,30	100
Depois	7,69	7,69	93,33	0	7,69	46,66	60
<i>Chebychev</i>							
Antes	6,66	6,66	33,33	6,66	7,69	86,66	6,66
Durante	100	100	100	100	100	38,46	100
Depois	7,69	7,69	7,69	7,69	7,69	0	7,69
<i>Cosine</i>							
Antes	93,33	7,69	15,38	7,69	7,69	69,23	86,66
Durante	100	100	93,33	100	100	80	84,61
Depois	7,69	7,69	61,53	0	0	61,53	7,69
<i>Correlation</i>							
Antes	15,38	7,69	92,30	7,69	7,69	46,66	0
Durante	76,92	100	38,46	100	100	46,15	80
Depois	80	0	80	0	0	92,30	92,30
<i>Spearman</i>							
Antes	86,66	0	100	6,66	6,66	92,30	76,92
Durante	92,30	100	84,61	7,69	100	93,33	93,33
Depois	100	7,69	86,66	100	7,69	100	100

Os próximos índices de acerto mostrados na tabela B.5 são referentes à aplicação do método k-means para separar os ensaios descritos através das frequências próprias transformadas em dados simbólicos.

TABELA B.5 - Índices de acerto (em %) obtidos utilizando as frequências aplicando o método k-means.

		k-means		
		Amostral	Uniforme	Cluster
		Frequências		
Euclidiana	Antes	86,66	86,66	86,66
	Durante	92,30	92,30	92,30
	Depois	69,23	69,23	69,23
Cityblock	Antes	100	100	100
	Durante	100	100	100
	Depois	100	100	100
Cosine	Antes	80	80	80
	Durante	76,92	76,92	76,92
	Depois	92,30	92,30	92,30
Correlation	Antes	80	80	80
	Durante	76,92	76,92	76,92
	Depois	92,30	92,30	92,30

O método c-means também é utilizado e os resultados são apresentados na tabela B.6.

TABELA B.6– Índices de acerto (em %) obtidos utilizando as frequências aplicando o método c-means.

c-means		
	Frequências	
Antes	86,67	33
Durante	100	33
Depois	92,3	33

### B.1.3. Primeiro modo

Os índices de acerto levando em conta a classificação dos primeiros modos próprios aplicando o método da aglomeração hierárquica são apresentados na tabela B.7.

TABELA B.7- Índices de acerto das classificações obtidas aplicando o método da aglomeração hierárquica para os primeiros modos próprios.

Aglomeração Hierárquica							
	Average	Centroid	Complete	Median	Single	Ward	Weighted
<i>Euclidiana</i>							
Antes	100	100	100	100	100	76,92	100
Durante	15,38	7,69	53,33	7,69	15,38	66,67	13,33
Depois	7,69	6,67	38,46	6,67	7,69	61,54	7,69
<i>Cityblock</i>							
Antes	100	100	100	100	100	100	100
Durante	100	6,67	86,67	23,08	7,69	100	93,33
Depois	84,62	0	69,23	6,67	7,69	76,92	7,69
<i>Minkowski</i>							
Antes	100	100	100	100	100	76,92	100
Durante	15,38	7,69	53,33	7,69	15,38	66,67	13,33
Depois	7,69	6,67	38,46	6,67	7,69	61,54	7,69
<i>Chebychev</i>							
Antes	100	100	100	100	100	84,62	100
Durante	7,69	7,69	7,69	7,69	7,69	33,33	7,69
Depois	6,67	6,67	6,67	6,67	6,67	30,77	6,67
<i>Cosine</i>							
Antes	100	100	76,92	100	100	84,62	100
Durante	33,33	15,38	66,67	7,69	15,38	53,85	66,67
Depois	15,38	0	23,08	0	7,69	53,33	15,38
<i>Correlation</i>							
Antes	100	100	92,31	100	100	84,62	100
Durante	53,33	15,38	73,33	7,69	15,38	73,33	23,08
Depois	38,46	0	69,23	7,69	7,69	53,85	6,67
<i>Spearman</i>							
Antes	100	100	100	100	100	100	92,31

Durante	86,67	7,69	92,31	7,69	86,67	92,31	92,31
Depois	61,54	0	92,31	0	7,69	86,67	86,67

Em seguida, o método da classificação utilizado para a distinção dos ensaios através dos primeiros modos próprios é o k-means. Os índices de acerto para essa aplicação são apresentados na tabela B.8.

TABELA B.8 - Índices de acerto (em %) obtidos utilizando os primeiros modos próprios aplicando o método k-means.

		k-means		
		Amostral	Uniforme	Cluster
		Modos 1		
Euclidiana	Antes	60	60	60
	Durante	53,84	53,84	53,84
	Depois	100	100	100
<i>Cityblock</i>	Antes	100	100	100
	Durante	100	100	100
	Depois	100	100	100
<i>Cosine</i>	Antes	60	60	60
	Durante	53,84	53,84	53,84
	Depois	100	100	100
<i>Correlation</i>	Antes	60	60	60
	Durante	46,15	46,15	46,15
	Depois	100	100	100

Os índices de acerto obtidos com o método de classificação c-means é mostrado na tabela B.9.

TABELA B.9 - Índices de acerto (em %) obtidos utilizando os primeiros modos próprios aplicando o método c-means.

c-means		
Modos 1		
Antes	86,67	33
Durante	61,53	33
Depois	100	33

#### B.1.4. Segundo modo

Agora, os segundos modos próprios transformados em dados simbólicos são utilizados para classificar os diferentes estados estruturais. Na primeira abordagem utilizando os

segundos modos próprios, o método de aglomeração hierárquica é utilizado. Os índices de acerto obtidos são apresentados na tabela B.10.

TABELA B.10 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica.

Aglomeração Hierárquica							
	Average	Centroid	Complete	Median	Single	Ward	Weighted
<i>Euclidiana</i>							
Antes	100	100	84,62	100	100	92,31	92,31
Durante	7,69	7,69	80	7,69	7,69	76,92	30,77
Depois	6,67	6,67	7,69	7,69	6,67	53,33	6,67
<i>Cityblock</i>							
Antes	100	100	100	100	100	100	100
Durante	100	7,69	69,23	7,69	7,69	100	92,31
Depois	76,92	6,67	30,77	6,67	0	69,23	86,67
<i>Minkowski</i>							
Antes	100	100	84,62	100	100	92,31	92,31
Durante	7,69	7,69	80	7,69	7,69	76,92	30,77
Depois	6,67	6,67	7,69	7,69	6,67	53,33	6,67
<i>Chebychev</i>							
Antes	100	100	93,33	100	100	84,62	100
Durante	7,69	7,69	15,38	7,69	7,69	30,77	7,69
Depois	7,69	7,69	7,69	6,67	7,69	20	7,69
<i>Cosine</i>							
Antes	92,31	100	69,23	100	100	92,31	100
Durante	84,62	7,69	61,54	6,67	7,69	76,92	76,92
Depois	13,33	6,67	53,33	0	6,67	66,67	0
<i>Correlation</i>							
Antes	92,31	100	84,62	100	100	84,62	84,62
Durante	61,54	7,69	46,67	7,69	7,69	76,92	33,33
Depois	6,67	6,67	38,46	6,67	7,69	60	15,38
<i>Spearman</i>							
Antes	92,31	100	93,33	100	100	92,31	92,31
Durante	92,31	7,69	84,62	7,69	7,69	92,31	80
Depois	6,67	7,69	61,54	7,69	7,69	46,67	0

Agora, o método de classificação de dados utilizado para identificar os diversos estados estruturais utilizando os segundos modos próprios, é o k-means. Os índices de acerto para essa aplicação são apresentados na tabela B.11.

TABELA B.11 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método k-means.

		k-means		
		Amostral	Uniforme	Cluster
		Modos 2		
Euclidiana	Antes	6,67	6,67	6,67
	Durante	76,92	76,92	76,92
	Depois	76,92	76,92	76,92
Cityblock	Antes	100,00	100,00	100,00
	Durante	76,92	69,23	69,23
	Depois	92,31	100,00	100,00
Cosine	Antes	13,33	13,33	13,33
	Durante	76,92	76,92	76,92
	Depois	69,23	69,23	69,23
Correlation	Antes	73,33	73,33	73,33
	Durante	53,85	53,85	53,85
	Depois	100,00	100,00	100,00

Os índices de acerto aplicando o método c-means para os segundos modos próprios são apresentados na tabela B.12.

TABELA B.12 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método c-means.

c-means		
Modos 2		
Antes	60	33
Durante	76,92	33
Depois	100	33

### B.1.5. Terceiro modo

Em seguida, os terceiros modos próprios transformados em dados simbólicos são utilizados juntamente com os métodos de classificação a fim de distinguir os estados estruturais, antes, durante e depois o reforço estrutural.

Na tabela B.13 são apresentados os índices de acerto utilizando os terceiros modos próprios aplicando o método de aglomeração hierárquica.

TABELA B.13 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica.

Aglomeração Hierárquica							
	Average	Centroid	Complete	Median	Single	Ward	Weighted
<i>Euclidiana</i>							
Antes	20,00	100,00	40,00	100,00	6,67	26,67	46,67
Durante	7,69	7,69	0,00	7,69	7,69	38,46	92,31
Depois	100,00	0,00	92,31	15,38	100,00	92,31	0,00
<i>Cityblock</i>							
Antes	20,00	6,67	73,33	100,00	6,67	73,33	53,33
Durante	23,08	100,00	53,85	0,00	100,00	23,08	92,31
Depois	100,00	0,00	69,23	7,69	0,00	100,00	84,62
<i>Minkowski</i>							
Antes	20,00	100,00	40,00	100,00	6,67	26,67	46,67
Durante	7,69	7,69	0,00	7,69	7,69	38,46	92,31
Depois	100,00	0,00	92,31	15,38	100,00	92,31	0,00
<i>Chebychev</i>							
Antes	100,00	93,33	100,00	26,67	100,00	93,33	100,00
Durante	7,69	7,69	7,69	7,69	7,69	7,69	7,69
Depois	30,77	30,77	30,77	100,00	0,00	30,77	30,77
<i>Cosine</i>							
Antes	20,00	100,00	73,33	100,00	6,67	40,00	46,67
Durante	30,77	7,69	23,08	7,69	7,69	38,46	0,00
Depois	92,31	0,00	76,92	0,00	100,00	92,31	92,31
<i>Correlation</i>							
Antes	80,00	100,00	53,33	100,00	100,00	53,33	40,00
Durante	30,77	15,38	53,85	15,38	7,69	38,46	38,46
Depois	92,31	0,00	69,23	0,00	7,69	76,92	92,31
<i>Spearman</i>							
Antes	100,00	6,67	86,67	6,67	100,00	86,67	60,00
Durante	30,77	23,08	30,77	23,08	7,69	61,54	23,08
Depois	100,00	100,00	84,62	100,00	0,00	100,00	100,00

O próximo método de classificação a ser aplicado para os terceiros modos próprios transformados em dados simbólicos é o método k-means. Os índices de acerto são mostrados na tabela B.14.

TABELA B.14 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método k-means.

		k-means		
		Amostrai	Uniforme	Cluster
		Modos 3		
Euclidiana	Antes	46,67	46,67	46,67
	Durante	38,46	38,46	38,46
	Depois	92,31	92,31	92,31
Cityblock	Antes	100,00	66,67	100,00
	Durante	84,62	61,54	92,31
	Depois	92,31	92,31	76,92
Cosine	Antes	53,33	53,33	53,33
	Durante	38,46	38,46	38,46
	Depois	84,62	84,62	84,62
Correlation	Antes	53,33	53,33	53,33
	Durante	38,46	38,46	38,46
	Depois	84,62	84,62	84,62

O método c-means é utilizado e os resultados são apresentados na tabela B.15.

TABELA B.15 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método c-means.

c-means		
Modos 3		
Antes	100	33
Durante	38,46	33
Depois	92,3	33

### B.1.6. Quarto modo

Em seguida, os quartos modos próprios transformados em dados simbólicos são utilizados juntamente com os métodos de classificação a fim de distinguir os estados estruturais antes, durante e depois do reforço estrutural.

Na tabela B.16 são apresentados os índices de acerto utilizando os quartos modos próprios aplicando o método de aglomeração hierárquica.



TABELA B.16 - Índices de acerto (em %) obtidos utilizando os quartos modos próprios aplicando o método da aglomeração hierárquica.

Aglomeração Hierárquica							
	Average	Centroid	Complete	Median	Single	Ward	Weighted
<i>Euclidiana</i>							
Antes	6,67	6,67	46,67	6,67	6,67	46,67	6,67
Durante	7,69	7,69	100,00	7,69	100,00	46,15	100,00
Depois	100,00	100,00	0,00	100,00	0,00	76,92	0,00
<i>Cityblock</i>							
Antes	13,33	6,67	66,67	6,67	6,67	73,33	6,67
Durante	7,69	7,69	7,69	100,00	7,69	100,00	7,69
Depois	100,00	100,00	92,31	7,69	100,00	92,31	100,00
<i>Minkowski</i>							
Antes	6,67	6,67	46,67	6,67	6,67	46,67	6,67
Durante	7,69	7,69	100,00	7,69	100,00	46,15	100,00
Depois	100,00	100,00	0,00	100,00	0,00	76,92	0,00
<i>Chebychev</i>							
Antes	6,67	13,33	6,67	6,67	6,67	46,67	13,33
Durante	100,00	100,00	15,38	7,69	100,00	84,62	0,00
Depois	0,00	0,00	92,31	100,00	0,00	0,00	100,00
<i>Cosine</i>							
Antes	93,33	6,67	20,00	6,67	6,67	46,67	20,00
Durante	38,46	7,69	46,15	7,69	7,69	84,62	92,31
Depois	0,00	100,00	100,00	100,00	100,00	61,54	0,00
<i>Correlation</i>							
Antes	6,67	6,67	33,33	6,67	6,67	40,00	6,67
Durante	30,77	7,69	38,46	100,00	7,69	69,23	30,77
Depois	100,00	100,00	53,85	7,69	100,00	100,00	100,00
<i>Spearman</i>							
Antes	73,33	6,67	46,67	6,67	6,67	66,67	0,00
Durante	30,77	7,69	46,15	7,69	7,69	61,54	30,77
Depois	30,77	100,00	53,85	100,00	100,00	84,62	100,00

O próximo método de classificação aplicado para aos quartos modos próprios transformados em dados simbólicos é o método k-means. Os índices de acerto podem ser observados na tabela B.17

TABELA B.17 - Índices de acerto (em %) obtidos utilizando os quartos modos próprios aplicando o método k-means.

k-means				
		Amostrai	Uniforme	Cluster
Modos 4				
Euclidiana	Antes	53,33	53,33	53,33
	Durante	53,85	53,85	53,85
	Depois	69,23	69,23	69,23
Cityblock	Antes	60,00	60,00	60,00
	Durante	61,54	46,15	30,77
	Depois	92,31	61,54	76,92
Cosine	Antes	53,85	53,85	53,85
	Durante	69,23	69,23	69,23
	Depois	0,00	0,00	0,00
Correlation	Antes	46,67	46,67	46,67
	Durante	53,85	53,85	53,85
	Depois	61,54	61,54	61,54

O método c-means é utilizado e os resultados são apresentados na tabela B.18.

TABELA B.18 - Índices de acerto (em %) obtidos utilizando os quartos modos próprios aplicando o método c-means.

c-means		
	Modos 4	
Antes	73,33	33
Durante	61,53	33
Depois	84,61	33

## B.2. Estudos complementares

A seguir, serão mostrados os resultados complementares das análises realizadas considerando apenas os ensaios referentes à estrutura antes do reforço estrutural e os ensaios depois do reforço estrutural.

### B.2.1. Sinais

Na tabela B.19, são mostrados os resultados obtidos utilizando os sinais transformados em dados simbólicos aplicando o método da aglomeração hierárquica.

TABELA B.19 – Índices de acerto (em %) obtidos utilizando os sinais aplicando o método da aglomeração hierárquica.

Aglomeración Hierárquica							
Euclidiana							
	Average	Centroide	Complete	Median	Single	Ward	Weighted
Antes	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Depois	7,69	7,69	7,69	7,69	7,69	7,69	7,69
<i>Cityblock</i>							
Antes	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Depois	7,69	7,69	7,69	7,69	7,69	7,69	7,69
<i>Minkowski</i>							
Antes	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Depois	7,69	7,69	7,69	7,69	7,69	7,69	7,69
<i>Chebychev</i>							
Antes	100,00	100,00	100,00	100,00	100,00	40,00	100,00
Depois	7,69	7,69	7,69	7,69	7,69	61,54	7,69
<i>Cosine</i>							
Antes	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Depois	7,69	7,69	23,08	7,69	7,69	23,08	7,69
<i>Correlation</i>							
Antes	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Depois	23,08	23,08	23,08	7,69	7,69	53,85	7,69
<i>Spearman</i>							
Antes	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Depois	7,69	7,69	7,69	7,69	7,69	7,69	7,69

Na tabela B.20 são apresentados os índices de acerto referentes à aplicação do k-means para os sinais.

TABELA B.20– Índices de acerto (em %) obtidos utilizando os sinais aplicando o método k-means.

k-means				
		Sinais		
		Amostrai	Uniforme	Cluster
Euclidiana	Antes	100	100	100
	Depois	23	23	23
<i>Cityblock</i>	Antes	73	73	73
	Depois	84	84	84
<i>Cosine</i>	Antes	100	100	100
	Depois	69	69	69
<i>Correlation</i>	Antes	100	100	100
	Depois	69	69	69

O último método de classificação aplicado para a separação dos sinais dos dois estados estruturais é o c-means. Os índices de acerto obtidos são mostrados na tabela B.21.

TABELA B.21– Índices de acerto (em %) obtidos utilizando os sinais aplicando o método c-means.

c-means		
	Sinais	
Antes	73,33	73
Depois	38,46	54

### B.2.2. Frequências

Na tabela B.22 são apresentados os índices de acerto utilizando as frequências aplicando o método da aglomeração hierárquica.

TABELA B.22 – Índices de acerto (em %) obtidos utilizando as frequências aplicando o método da aglomeração hierárquica.

Aglomeração Hierárquica							
<i>Euclidiana</i>							
	Average	Centroide	Complete	Median	Single	Ward	Weighted
Antes	53,33	6,67	40,00	100,00	6,67	53,33	53,33
Depois	100,00	100,00	100,00	7,69	100,00	100,00	100,00
<i>Cityblock</i>							
Antes	93,33	100,00	93,33	6,67	6,67	93,33	93,33
Depois	100,00	7,69	100,00	100,00	100,00	100,00	100,00
<i>Minkowski</i>							
Antes	53,33	6,67	40,00	100,00	100,00	53,33	53,33
Depois	100,00	100,00	100,00	7,69	6,67	100,00	100,00
<i>Chebychev</i>							
Antes	100,00	6,67	13,33	6,67	6,67	73,33	100,00
Depois	15,38	100,00	100,00	100,00	100,00	76,92	23,08
<i>Cosine</i>							
Antes	100,00	100,00	66,67	6,67	100,00	60,00	53,33
Depois	15,38	7,69	69,23	100,00	7,69	100,00	100,00
<i>Correlation</i>							
Antes	100,00	100,00	93,33	100,00	100,00	60,00	86,67
Depois	7,69	7,69	92,31	7,69	7,69	100,00	46,15
<i>Spearman</i>							
Antes	86,67	100,00	93,33	100,00	100,00	93,33	93,33
Depois	100,00	7,69	92,31	7,69	7,69	100,00	23,08

Os próximos índices de acerto mostrados na tabela B.23 são referentes à aplicação do método k-means.

TABELA B.23 - Índices de acerto (em %) obtidos utilizando as frequências aplicando o método k-means.

k-means				
		Frequências		
		Amostrai	Uniforme	Cluster
Euclidiana	Antes	53	53	53
	Depois	100	100	100
<i>Cityblock</i>	Antes	100	100	100
	Depois	100	100	100
<i>Cosine</i>	Antes	100	100	100
	Depois	100	100	100
<i>Correlation</i>	Antes	100	100	100
	Depois	100	100	100

O método c-means também é utilizado e os resultados são apresentados na tabela B.24.

TABELA B.24- Índices de acerto (em %) obtidos utilizando as frequências aplicando o método c-means.

c-means		
		Frequências
Antes	86,66	50
Depois	100	50

### B.2.3. Primeiro Modo

Os índices de acerto levando em conta a classificação dos primeiros modos próprios aplicando o método da aglomeração hierárquica são apresentados na tabela B.25.

TABELA B.25 - Índices de acerto das classificações obtidas aplicando o método da aglomeração hierárquica para os primeiros modos próprios.

Aglomeração Hierárquica							
Euclidiana							
	Average	Centroide	Complete	Median	Single	Ward	Weighted
Antes	66,67	100,00	86,67	13,33	100,00	100,00	66,67
Depois	100,00	7,69	100,00	100,00	7,69	92,31	100,00
<i>Cityblock</i>							

Antes	100,00	6,67	100,00	6,67	100,00	100,00	100,00
Depois	100,00	100,00	100,00	100,00	7,69	100,00	100,00
<i>Minkowski</i>							
Antes	66,67	100,00	86,67	13,33	100,00	100,00	66,67
Depois	100,00	7,69	100,00	100,00	7,69	92,31	100,00
<i>Chebychev</i>							
Antes	6,67	6,67	6,67	6,67	100,00	33,33	6,67
Depois	100,00	100,00	100,00	100,00	7,69	92,31	100,00
<i>Cosine</i>							
Antes	100,00	100,00	100,00	100,00	100,00	100,00	66,67
Depois	100,00	7,69	92,31	7,69	7,69	100,00	100,00
<i>Correlation</i>							
Antes	100,00	6,67	100,00	100,00	100,00	100,00	100,00
Depois	100,00	100,00	92,31	7,69	7,69	92,31	100,00
<i>Spearman</i>							
Antes	86,67	93,33	100,00	100,00	86,67	86,67	86,67
Depois	100,00	100,00	92,31	7,69	100,00	100,00	100,00

Em seguida, o método da classificação utilizado para a distinção dos ensaios através dos primeiros modos próprios é o k-means. Os índices de acerto para essa aplicação são apresentados na tabela B.26.

TABELA B.26 - Índices de acerto (em %) obtidos utilizando os primeiros modos próprios aplicando o método k-means.

k-means				
		Modos 1		
		Amostrai	Uniforme	Cluster
Euclidiana	Antes	86	86	86
	Depois	100	100	100
<i>Cityblock</i>	Antes	100	100	100
	Depois	100	100	100
<i>Cosine</i>	Antes	86	86	86
	Depois	100	100	100
<i>Correlation</i>	Antes	86	86	86
	Depois	100	100	100

Os índices de acerto obtidos com o método de classificação c-means é mostrado na tabela B.27.

TABELA B.27 - Índices de acerto (em %) obtidos utilizando os primeiros modos próprios aplicando o método c-means.

c-means		
	Modos 1	
Antes	100	50
Depois	100	50

#### B.2.4. Segundo Modo

Na primeira abordagem utilizando os segundos modos próprios, o método de aglomeração hierárquica é utilizado. Os índices de acerto obtidos são apresentados na tabela B.28.

TABELA B.28 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método da aglomeração hierárquica.

Aglomeración Hierárquica							
Euclidiana							
	Average	Centroide	Complete	Median	Single	Ward	Weighted
Antes	6,67	6,67	100,00	6,67	100,00	80,00	86,67
Depois	100,00	100,00	61,54	100,00	7,69	69,23	38,46
Cityblock							
Antes	100,00	6,67	100,00	6,67	100,00	100,00	100,00
Depois	100,00	100,00	100,00	100,00	7,69	100,00	100,00
Minkowski							
Antes	6,67	6,67	100,00	6,67	100,00	80,00	86,67
Depois	100,00	100,00	61,54	100,00	7,69	69,23	38,46
Chebychev							
Antes	100,00	100,00	93,33	6,67	100,00	53,33	6,67
Depois	7,69	7,69	7,69	100,00	7,69	61,54	100,00
Cosine							
Antes	53,33	100,00	100,00	100,00	6,67	86,67	46,67
Depois	100,00	7,69	69,23	7,69	100,00	76,92	61,54
Correlation							
Antes	93,33	100,00	73,33	100,00	100,00	60,00	93,33
Depois	15,38	7,69	53,85	7,69	7,69	69,23	76,92
Spearman							
Antes	93,33	100,00	93,33	100,00	100,00	93,33	100,00
Depois	100,00	7,69	100,00	7,69	7,69	100,00	92,31

Agora, o método de classificação de dados utilizado é o k-means para identificar os diferentes estados estruturais utilizando os segundos modos próprios. Os índices de acerto para essa aplicação são apresentados na tabela B.29.

TABELA B.29 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método k-means.

k-means				
		Modos 2		
		Amostrai	Uniforme	Cluster
Euclidiana	Antes	80	80	80
	Depois	92	92	92
Cityblock	Antes	100	100	100
	Depois	100	100	100
Cosine	Antes	80	80	80
	Depois	92	92	92
Correlation	Antes	73	73	73
	Depois	84	84	84

Os índices de acerto aplicando o método c-means para os segundos modos próprios são apresentados na tabela B.30.

TABELA B.30 - Índices de acerto (em %) obtidos utilizando os segundos modos próprios aplicando o método c-means.

c-means		
		Modos 2
Antes	100	50
Depois	100	50

### B.2.5. Terceiro Modo

Em seguida, os terceiros modos próprios transformados em dados simbólicos são utilizados juntamente com os métodos de classificação a fim de distinguir os diferentes estados estruturais.

Na tabela B.31 são apresentados os índices de acerto utilizando os terceiros modos próprios aplicando o método de aglomeração hierárquica.



TABELA B.31 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método da aglomeração hierárquica.

Aglomeração Hierárquica							
Euclidiana							
	Average	Centroide	Complete	Median	Single	Ward	Weighted
Antes	13,33	6,67	80,00	6,67	6,67	80,00	86,67
Depois	100,00	100,00	76,92	100,00	100,00	76,92	76,92
<i>Cityblock</i>							
Antes	100,00	6,67	73,33	100,00	6,67	100,00	66,67
Depois	84,62	100,00	100,00	7,69	100,00	69,23	100,00
<i>Minkowski</i>							
Antes	13,33	6,67	80,00	6,67	6,67	80,00	86,67
Depois	100,00	100,00	76,92	100,00	100,00	76,92	76,92
<i>Chebychev</i>							
Antes	93,33	93,33	100,00	6,67	100,00	93,33	93,33
Depois	30,77	30,77	30,77	100,00	7,69	30,77	30,77
<i>Cosine</i>							
Antes	80,00	6,67	73,33	100,00	6,67	73,33	93,33
Depois	76,92	100,00	76,92	7,69	100,00	76,92	69,23
<i>Correlation</i>							
Antes	86,67	6,67	33,33	6,67	6,67	86,67	100,00
Depois	92,31	100,00	76,92	100,00	100,00	76,92	15,38
<i>Spearman</i>							
Antes	100,00	6,67	46,67	100,00	6,67	66,67	66,67
Depois	100,00	100,00	69,23	7,69	100,00	84,62	100,00

O próximo método de classificação a ser aplicado para os terceiros modos próprios transformados em dados simbólicos é o método k-means. Os índices de acerto são mostrados na tabela B.32.

TABELA B.32 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método k-means.

k-means				
		Modos 3		
		Amostrai	Uniforme	Cluster
Euclidiana	Antes	93	93	93
	Depois	76	76	76
<i>Cityblock</i>	Antes	100	100	100
	Depois	84	84	84
<i>Cosine</i>	Antes	93	93	93
	Depois	76	76	76
<i>Correlation</i>	Antes	80	80	80
	Depois	92	92	92

O método c-means é utilizado e os resultados são apresentados na tabela B.33.

TABELA B.33 - Índices de acerto (em %) obtidos utilizando os terceiros modos próprios aplicando o método c-means.

c-means		
	Modos 3	
Antes	80	50
Depois	100	50

### B.2.6. Quarto Modo

Os índices de acerto levando em conta a classificação dos quartos modos próprios aplicando o método da aglomeração hierárquica são apresentados na tabela B.34.

TABELA B.34- Índices de acerto das classificações obtidas aplicando o método da aglomeração hierárquica para os quartos modos próprios.

Aglomeração Hierárquica							
Euclidiana							
	Average	Centroide	Complete	Median	Single	Ward	Weighted
Antes	6,67	6,67	46,67	6,67	6,67	53,33	6,67
Depois	100,00	100,00	53,85	100,00	100,00	76,92	100,00
<i>Cityblock</i>							
Antes	13,33	6,67	66,67	6,67	6,67	66,67	6,67
Depois	92,31	100,00	92,31	100,00	100,00	100,00	100,00
<i>Minkowski</i>							
Antes	6,67	6,67	46,67	6,67	6,67	53,33	6,67
Depois	100,00	100,00	53,85	100,00	100,00	76,92	100,00
<i>Chebychev</i>							
Antes	6,67	6,67	6,67	6,67	6,67	6,67	6,67
Depois	100,00	100,00	100,00	100,00	100,00	100,00	100,00
<i>Cosine</i>							
Antes	6,67	6,67	40,00	6,67	6,67	66,67	13,33
Depois	100,00	100,00	100,00	100,00	100,00	84,62	100,00
<i>Correlation</i>							
Antes	6,67	6,67	53,33	93,33	6,67	26,67	86,67
Depois	100,00	100,00	53,85	7,69	100,00	76,92	15,38
<i>Spearman</i>							
Antes	60,00	6,67	100,00	13,33	6,67	73,33	66,67
Depois	84,62	100,00	46,15	92,31	100,00	84,62	61,54

Em seguida, o método da classificação utilizado para a distinção dos ensaios através dos quartos modos próprios é o k-means. Os índices de acerto para essa aplicação são apresentados na tabela B.35.

TABELA B.35 - Índices de acerto (em %) obtidos utilizando os quartos modos próprios aplicando o método k-means.

k-means				
		Modos 4		
		Amostrai	Uniforme	Cluster
Euclidiana	Antes	60	60	60
	Depois	76	76	76
<i>Cityblock</i>	Antes	73	73	73
	Depois	100	100	100
<i>Cosine</i>	Antes	60	60	60
	Depois	76	76	76
<i>Correlation</i>	Antes	60	60	60
	Depois	46	46	46

Os índices de acerto obtidos com o método de classificação c-means é mostrado na tabela B.36.

TABELA B.36 - Índices de acerto (em %) obtidos utilizando os quartos modos próprios aplicando o método c-means.

c-means		
		Modos 4
Antes	86,66	50
Depois	100	50