

**Faculdade de Engenharia da Universidade do Porto**



**FEUP**

**Utilização de Árvores de Regressão Lineares  
para Avaliação de Segurança Dinâmica de  
Sistemas Interligados com Elevada Integração de  
Produção Eólica**

João Manuel Dantas Monteiro da Rocha Barbosa

Dissertação realizada no âmbito do  
Mestrado Integrado em Engenharia Electrotécnica e de Computadores  
Major Energia

Orientador: Prof. Dr. Maria Helena Osório Pestana de Vasconcelos

Porto, Julho de 2010

© João Dantas, 2010

# Dedicatória

*Para a Helena,  
para a minha Mãe  
e para a minha irmã Maria...*



# Resumo

Nesta dissertação apresenta-se uma metodologia para a avaliação de segurança dinâmica em sistemas interligados com elevada integração de produção eólica e onde não tenham sido totalmente adoptados os requisitos de capacidade para sobreviver a cavas de tensão.

Através de um método de aprendizagem automática híbrida procura-se avaliar de forma rápida a robustez destes sistemas, utilizando os dados de uma rede interligada de teste criada para o efeito. Esta rede permitiu avaliar a qualidade dos resultados fornecidos pelas metodologias propostas quando o problema de segurança dinâmica colocado em análise é a ocorrência de sobrecargas em regime quasi-estacionário, nos ramos da rede de transmissão, que violem os limites definidos como aceitáveis para sobrecargas temporárias.

Avaliou-se o desempenho da previsão das estimativas, para o problema de segurança dinâmica em análise, a partir dos resultados obtidos pelo treino de modelos de regressão linear e de estruturas híbridas, que resultam da utilização de modelos de regressão linear nas folhas de uma árvore de regressão.

O trabalho realizado e que conduziu à elaboração desta dissertação decorreu na Faculdade de Engenharia da Universidade do Porto (FEUP).



# Abstract

In this dissertation is presented a methodology for the evaluation of dynamic security in interconnected systems with high integration of wind production and where there haven't been totally adopted the capacity requirements to survive voltage drops.

Through a hybrid automatic learning method one tries to quickly evaluate the strength of these systems using the data of a test interconnect network that is created for this purpose. This network allowed the quality evaluation of the results given by the proposed methodologies when the dynamic security problem put into analysis is the occurrence of overloads in almost stationary regimes in the branches of the transmission network that violate the defined limits as acceptable to temporary overloads.

The performance of the estimates' prediction was evaluated for the problem of dynamic security in analysis from the obtained results through the training of models of linear regression and hybrid structures that result from the use of linear regression models in the leaves of a regression tree.

The work done that has lead to the writing of this dissertation took place at Faculty of Engineering of University of Porto (FEUP).



# Abstrakt

In dieser Masterarbeit wird eine Methodologie für die Bewertung der dynamischen Sicherheit in verbundenen Systemen mit hoher Integration der Windproduktion und der Stellen, wo die notwendige Kapazität um einen Spannungsabfall zu überstehen nicht im Ganzen angewendet wird, vorgetragen.

Durch eine automatische hybride Lernmethode versucht man die Widerstandsfähigkeit solcher Systeme rasch zu bewerten und dabei werden die Daten von einem verbundenen Testnetz, das für diesen Zweck erschaffen wurde, benutzt. Dieses Netz ermöglicht, die Qualität der Resultate zu bewerten nachdem Methoden vorgetragen wurden wenn das Problem der dynamischen Sicherheit eine Hochspannung in quasi-stationärem Zustand in der Verzweigung des Übertragungsnetzes ist und wenn diese Hochspannung den Sicherheitsgrenzwert für temporäre Hochspannungen überschreitet.

Die Ausführung der vorgesehenen Schätzung für das Problem der analysierten dynamischen Sicherheit wurde von den Resultaten aus dem Training der linearen Regressionsmodelle und hybriden Strukturen erlangt und bewertet. Diese ergeben sich aus der Anwendung von linearen Regressionsmodellen in den Blättern eines Regressionsbaumes.

Das Projekt, dass zu dieser Masterarbeit führte, wurde in der Ingenieurfacultät der Universität Porto (FEUP) durchgeführt.



# Agradecimentos

Gostaria de agradecer a todas as pessoas que me acompanharam ao longo da realização desta dissertação, nomeadamente:

À Professora Doutora Maria Helena Osório Pestana de Vasconcelos pela orientação prestada ao longo destes meses, pela boa disposição e clareza no esclarecimento de todas as dúvidas levantadas.

À Doutora Helena Sofia Lopes da UOI que me dotou de algumas ferramentas para retomar as linhas mestras da minha vida e à Professora Margarida Ferreira pelas palavras de incentivo no decorrer da vida académica.

À Helena, que esteve sempre presente na minha vida e em mim depositou a sua confiança.

À minha mãe e irmã, sem as quais não seria possível ter alcançado tamanho objectivo, pela sua compreensão e paciência.

Ao meu grande amigo João Paupério pela sua alegria contagiante, comentários sempre pertinentes e pela paciência na leitura desta dissertação.

Aos grandes amigos, Ana Sofia Costa, Diana Paupério, Florência Valente, Susana Oliveira, João Paredes e Paulo Coelho.

Aos pais da Helena, Moisés e Palmira, à minha madrinha Dulce e à minha família materna.

Queria também agradecer a todos os colegas da Universidade do Porto com os quais tive o prazer de trocar impressões e colaborar.



# Índice

<b>CAPÍTULO 1 : INTRODUÇÃO</b>	<b>1</b>
1.1 Considerações Gerais	1
1.2 Objectivos da Dissertação	2
1.3 Estrutura da Dissertação	3
<b>CAPÍTULO 2 : SEGURANÇA DINÂMICA DE UM SISTEMA DE ENERGIA</b>	<b>5</b>
2.1 Introdução	5
2.2 Segurança de Operação em Sistemas Eléctricos de Energia	6
2.3 Problemas Associados à Segurança Dinâmica de Redes Interligadas com Produção Eólica	8
2.4 Conclusões	10
<b>CAPÍTULO 3 : MODELOS DE REGRESSÃO PARA AVALIAÇÃO DE SEGURANÇA DINÂMICA DE UM SEE</b>	<b>11</b>
3.1 Introdução	11
3.2 Estimativas do Erro de Previsão	12
3.3 Avaliação da Capacidade de Generalização	13
3.4 Modelos de Regressão Linear (LRM)	14
3.5 Árvores de Regressão Linear (LRT)	16
3.5.1 Treino de uma RT	17
3.6 Metodologia adoptada para obtenção de uma Estrutura de Avaliação Dinâmica de um SEE	19
3.7 Conclusões	21
<b>CAPÍTULO 4 : ESTRUTURA DA REDE INTERLIGADA EM ANÁLISE</b>	<b>23</b>
4.1 Introdução	23
4.2 Estrutura da Rede Interligada	23
4.3 Parametização e Resultados dos Dados Gerados para a Rede Interligada de Teste	25

<b>4.4 Conclusões</b>	<b>27</b>
<b><u>CAPÍTULO 5 : AVALIAÇÃO DE SEGURANÇA DINÂMICA PARA A REDE DE TESTE</u></b>	<b><u>29</u></b>
<b>5.1 Introdução</b>	<b>29</b>
<b>5.2 Treino do Modelo de Regressão Linear (LRM)</b>	<b>31</b>
<b>5.3 Treino das Árvores de Regressão Linear (LRT)</b>	<b>34</b>
5.3.1 Treino de Árvores de Regressão com duas folhas	36
5.3.2 Treino de Árvores de Regressão com quatro folhas	44
<b>5.4 Comparação das Estruturas de Treino</b>	<b>52</b>
5.4.1 Comparação entre LRM, LRT de 2 folhas e LRT de 4 folhas (alternativa A):	52
5.4.2 Comparação entre LRM, LRT de 2 folhas e LRT de 4 folhas (alternativa B):	53
5.4.3 Comparação entre LRM, LRT de 2 folhas e LRT de 4 folhas (alternativa C):	62
5.4.4 Comparação entre LRM, LRT de 2 folhas e LRT de 4 folhas (alternativa D):	63
<b>5.5 Conclusões</b>	<b>64</b>
<b><u>CAPÍTULO 6 : CONCLUSÕES</u></b>	<b><u>65</u></b>
<b><u>REFERÊNCIAS</u></b>	<b><u>67</u></b>

# Lista de Figuras

Figura 2.1 – Diagrama de operação e transição de estados do sistema (adaptado de [3] e [6])	7
Figura 4.1 – Esquema unifilar da rede interligada de teste (obtido de [6])	24
Figura 5.1 – Valor dos parâmetros dos LRM treinados para a rede de teste (parte 1/2)	31
Figura 5.2 – Valor dos parâmetros dos LRM treinados para a rede de teste (parte 2/2)	32
Figura 5.3 – Erros de regressão e de classificação fornecidos pelos LRM treinados, para a rede de teste	33
Figura 5.4 – Erros de regressão e de classificação fornecidos pelos LRT treinados, para a rede de teste, considerando a alternativa A de 2 folhas	36
Figura 5.5 – Estrutura em árvore das LRT que foram treinadas, para a rede de teste, considerando a alternativa A de 2 folhas	37
Figura 5.6 - Erros de regressão e de classificação fornecidos pelos LRT treinados, para a rede de teste, considerando a alternativa B de 2 folhas	38
Figura 5.7 - Estrutura em árvore das LRT que foram treinadas, para a rede de teste, considerando a alternativa B de 2 folhas	39
Figura 5.8 - Erros de regressão e de classificação fornecidos pelos LRT treinados, para a rede de teste, considerando a alternativa C de 2 folhas	40
Figura 5.9 - Estrutura em árvore das LRT que foram treinadas, para a rede de teste, considerando a alternativa C de 2 folhas	41
Figura 5.10 - Erros de regressão e de classificação fornecidos pelos LRT, treinados para a rede de teste, considerando a alternativa D de 2 folhas	42
Figura 5.11 - Estrutura em árvore das LRT que foram treinadas, para a rede de teste, considerando a alternativa D de 2 folhas	43
Figura 5.12 - Erros de regressão e de classificação fornecidos pelos LRT treinados, para a rede de teste, considerando a alternativa A de 4 folhas	44
Figura 5.13 - Estrutura em árvore das LRT que foram treinadas, para a rede de teste, considerando a alternativa A de 4 folhas	45

Figura 5.14 - Erros de regressão e de classificação fornecidos pelos LRT treinados, para a rede de teste, considerando a alternativa B de 4 folhas	46
Figura 5.15 - Estrutura em árvore das LRT que foram treinadas, para a rede de teste, considerando a alternativa B de 4 folhas	47
Figura 5.16 - Erros de regressão e de classificação fornecidos pelos LRT treinados, para a rede de teste, considerando a alternativa C de 4 folhas	48
Figura 5.17 - Estrutura em árvore das LRT que foram treinadas, para a rede de teste, considerando a alternativa C de 4 folhas	49
Figura 5.18 - Erros de regressão e de classificação fornecidos pelos LRT treinados, para a rede de teste, considerando a alternativa D de 4 folhas	50
Figura 5.19 - Estrutura em árvore das LRT que foram treinadas, para a rede de teste, considerando a alternativa D de 4 folhas	51
Figura 5.20 - Erros de regressão e de classificação fornecidos por LRM, LRT de 2 folhas e LRT de 4 folhas, para a alternativa A (parte 1/2)	52
Figura 5.21 – Erros de regressão e de classificação fornecidos por LRM, LRT de 2 folhas e LRT de 4 folhas, para a alternativa A (parte 2/2)	53
Figura 5.22 - Erros de regressão e de classificação fornecidos por LRM, LRT de 2 folhas e LRT de 4 folhas, para a alternativa B (parte 1/2)	53
Figura 5.23 – Erros de regressão e de classificação fornecidos por LRM, LRT de 2 folhas e LRT de 4 folhas, para a alternativa B (parte 2/2)	54
Figura 5.24 – Erros de regressão fornecidos por LRM em função do índice de segurança (avaliação da linha 15-16)	55
Figura 5.25 – Erros de regressão fornecidos por LRT, 2 folhas, em função do valor do índice de segurança (avaliação da linha 15-16)	55
Figura 5.26 – Erros de regressão fornecidos por LRT, 4 folhas, em função do valor do índice de segurança (avaliação da linha 15-16)	55
Figura 5.27 – Erros de regressão fornecidos por LRM em função do índice de segurança (avaliação da linha 15-17)	56
Figura 5.28 – Erros de regressão fornecidos por LRT, 2 folhas, em função do valor do índice de segurança (avaliação da linha 15-17)	56
Figura 5.29 – Erros de regressão fornecidos por LRT, 4 folhas, em função do valor do índice de segurança (avaliação da linha 15-17)	56
Figura 5.30 – Erros de regressão fornecidos por LRM em função do índice de segurança (avaliação da linha 16-4)	57
Figura 5.31 – Erros de regressão fornecidos por LRT, 2 folhas, em função do valor do índice de segurança (avaliação da linha 16-4)	57

Figura 5.32 – Erros de regressão fornecidos por LRT, 4 folhas, em função do valor do índice de segurança (avaliação da linha 16-4)	57
Figura 5.33 – Erros de regressão fornecidos por LRM em função do índice de segurança (avaliação da linha 16-6)	58
Figura 5.34 – Erros de regressão fornecidos por LRT, 2 folhas, em função do valor do índice de segurança (avaliação da linha 16-6)	58
Figura 5.35 – Erros de regressão fornecidos por LRT, 4 folhas, em função do valor do índice de segurança (avaliação da linha 16-6)	58
Figura 5.36 – Erros de regressão fornecidos por LRM em função do índice de segurança (avaliação da linha 14-18)	59
Figura 5.37 – Erros de regressão fornecidos por LRT, 2 folhas, em função do valor do índice de segurança (avaliação da linha 14-18)	59
Figura 5.38 – Erros de regressão fornecidos por LRT, 4 folhas, em função do valor do índice de segurança (avaliação da linha 14-18)	59
Figura 5.39 – Erros de regressão fornecidos por LRM em função do índice de segurança (avaliação da linha 16-18)	60
Figura 5.40 – Erros de regressão fornecidos por LRT, 2 folhas, em função do valor do índice de segurança (avaliação da linha 16-18)	60
Figura 5.41 – Erros de regressão fornecidos por LRT, 4 folhas, em função do valor do índice de segurança (avaliação da linha 16-18)	60
Figura 5.42 – Erros de regressão fornecidos por LRM em função do índice de segurança (avaliação da linha 20-21)	61
Figura 5.43 – Erros de regressão fornecidos por LRT, 2 folhas, em função do valor do índice de segurança (avaliação da linha 20-21)	61
Figura 5.44 – Erros de regressão fornecidos por LRT, 4 folhas, em função do valor do índice de segurança (avaliação da linha 20-21)	61
Figura 5.45 – Erros de regressão e de classificação fornecidos por LRM, LRT de 2 folhas e LRT de 4 folhas, considerando a alternativa C	62
Figura 5.46 – Erros de regressão e de classificação fornecidos por LRM, LRT de 2 folhas e LRT de 4 folhas, considerando a alternativa D	63



# Lista de Abreviaturas

AA - Aprendizagem Automática

ENTSO-E - *European Network of Transmission System Operators for Electricity*

EWEA - *The European Wind Energy Association*

LRM - *Linear Regression Models* (Modelos de Regressão Linear)

LRT - *Linear Regression Tree* (Árvore de Regressão Linear)

MAT - Muito Alta Tensão

PSS/E - *Power System Simulation for Engineering*

RT - *Regression Tree* (Árvore de Regressão)

SCADA - *Supervisory Control and Data Acquisition*

UCTE - *Union for the Co-ordination of Transmission of Electricity*



# Capítulo 1: Introdução

## 1.1 Considerações Gerais

A avaliação de segurança encontra-se relacionada com métodos e análises para determinação do estado de segurança de um sistema (seguro ou inseguro), de acordo com um conjunto de critérios pré-estabelecidos. Este tipo de avaliação de segurança pode ser dividido em segurança estacionária e dinâmica. A segurança estacionária refere-se ao equilíbrio do sistema de energia, onde se respeitam os limites de tensão e térmicos, isto é, para um dado ponto de operação um sistema considera-se em equilíbrio estacionário se as tensões nos barramentos se encontram dentro de uma gama de valores pré-determinados e se os trânsitos de potência nas linhas e nos transformadores se encontram dentro dos seus limites térmicos. A segurança dinâmica é um pouco mais complexa, pois contempla a estabilidade da frequência, transitória, de tensão e de pequeno sinal.

A evolução dos sistemas de energia, com as inúmeras redes emalhadadas e a introdução da produção descentralizada, têm levado a estudos de estabilidade dinâmica em grandes áreas interligadas. As necessidades económicas e de eficiência parecem levar à crescente exploração dos sistemas de energia e das interligações entre áreas para próximo dos seus limites operacionais. O interesse nestes estudos e na avaliação de segurança dinâmica é levantado pelo aumento das situações de operação que podem causar perturbações ao funcionamento do sistema de energia, pois é importante manter os sistemas de energia interligados a operar com margem de segurança e monitorizar regularmente a segurança dinâmica em condições de operação normal dos sistemas, garantindo a fiabilidade dos sistemas.

## 1.2 Objectivos da Dissertação

Constitui objecto desta dissertação, a avaliação da capacidade de previsão da estrutura de aprendizagem automática híbrida, resultante do treino de um modelo de regressão linear em cada folha de uma árvore de regressão, no sentido de se obter uma estrutura de segurança linear mais precisa do que o uso isolado do modelo de regressão linear.

Nos sistemas interligados com elevada integração de energia eólica podem ocorrer certas perturbações que conduzem a problemas de segurança dinâmica nessas redes. Mais concretamente, em zonas críticas da rede de transporte, podem surgir quedas de tensão provocadas pelo curto-circuito que podem levar à actuação dos relés de mínimo de tensão instalados em parques que se encontrem nas proximidades do defeito e que não disponham de capacidade para sobreviver a cavas de tensão. Tal como se explicará à frente, este tipo de perturbações pode levar à perda de segurança do sistema de energia, uma vez que podem ser ultrapassadas as margens de segurança habitualmente adoptadas. Dos problemas dinâmicos destacam-se os problemas de estabilidade transitória e as sobrecargas temporárias nos ramos.

O problema de segurança que se aborda neste trabalho é o da sobrecarga temporária em ramos da rede de transporte e são aplicadas técnicas de aprendizagem automática supervisionada para investigar este problema. O desempenho das estruturas é avaliado por aplicação a um problema de segurança de uma rede interligada de teste, a qual foi fornecida pela orientadora da presente Tese.

A precisão de estimação da estrutura híbrida que resulta da aplicação do treino de um modelo de regressão linear a cada folha de uma árvore de regressão linear foi testada por comparação com os resultados fornecidos por um modelo de regressão linear, para o problema de segurança dinâmica em análise. Para as estruturas híbridas foram ainda criadas um conjunto de alternativas, para o vector de variáveis de entrada, com restrições de segurança do tipo linear para avaliar a capacidade de obtenção de estruturas de segurança do tipo linear a partir de árvores de regressão linear.

### **1.3 Estrutura da Dissertação**

A estrutura da dissertação é a que se segue. No Capítulo 2 estabelece-se uma relação entre as tendências de exploração dos sistemas de energia e as directrizes globais para uma modificação das estratégias de gestão do parque energético. São definidos os estados em que o sistema pode residir e os problemas de segurança dinâmica que podem surgir, decorrentes da exploração de sistemas com elevada produção eólica.

No Capítulo 3 descreve-se a metodologia utilizada para obtenção de estruturas de segurança de aprendizagem automática e que realizam avaliação de segurança para o problema em análise. Descrevem-se, então, essas técnicas utilizadas para obtenção das estruturas que serão aplicadas à rede interligada de teste, e que se encontra caracterizada no Capítulo 4.

O Capítulo 5 é dedicado à apresentação dos resultados obtidos para avaliação do desempenho do treino de estruturas de avaliação de segurança por aplicação de técnicas de aprendizagem automática para a rede interligada de teste.

No Capítulo 6 são resumidas as principais conclusões que foram alcançadas com a realização do presente trabalho.



# Capítulo 2: Segurança Dinâmica de um Sistema de Energia

## 2.1 Introdução

Nos últimos anos tem-se assistido a um aumento progressivo da integração de tecnologias de produção de energia eléctrica que promovam simultaneamente a redução dos gases de efeito de estufa assim como a dependência energética externa. O protocolo de Kyoto veio introduzir um conjunto de limitações muito significativo na redução dessas emissões e, mais recentemente, assistiu-se à tentativa de reforço destas medidas na conferência de Copenhaga em 2009, ainda que desta conferência não tenham resultado novos acordos.

Num relatório da ENTSO-E (*European Network of Transmission System Operators for Electricity*) de 2009, [1], pode ler-se que a produção eólica irá desempenhar um papel de destaque nos próximos anos, de entre as restantes fontes renováveis. Neste documento aborda-se a necessidade de uniformização europeia para as regras de ligação dos geradores eólicos à rede. O destaque dado a esta fonte de energia renovável reside não só em ser uma fonte de energia limpa, livre e inesgotável, mas também por este recurso oferecer baixos custos de exploração e sem riscos associados à escalada de preço dos produtos petrolíferos. Ainda, a uniformização europeia das ligações dos geradores eólicos à rede, pode permitir uma redução de custos importante e promover o aumento da fiabilidade nesta área alargada.

No relatório anual de actividade de 2009 da EWEA (*The European Wind Energy Association*), [2], pode observar-se que a capacidade de produção eólica instalada no decorrer desse ano foi superior a qualquer outra tecnologia de produção de electricidade. Com efeito, a capacidade instalada na Europa era de cerca 76 GW, no final de 2009.

Verifica-se que as metas impostas para a produção de energia eléctrica seguem uma tendência de modificar as características de operação relativamente às provenientes dos grupos de geração convencionais.

Reconhece-se, assim, a necessidade de compreender o impacto da integração de produção de energia eólica em redes interligadas. Do ponto de vista da avaliação de segurança dinâmica, interessa analisar o comportamento dos sistemas eléctricos na sequência da ocorrência de perturbações que levem à perda de elevados volumes de produção eólica.

Neste capítulo são descritos os estados e transições em que o sistema de energia pode residir e os problemas de segurança dinâmica levantados por algumas perturbações associadas às redes interligadas.

## 2.2 Segurança de Operação em Sistemas Eléctricos de Energia

A segurança de um sistema eléctrico de energia encontra-se relacionada com a robustez desse sistema perante a ocorrência de perturbações e, portanto, depende das condições de operação. A avaliação de segurança, conforme se descreve em [3] e [4], consiste na análise da capacidade que um sistema de energia tem para suportar perturbações credíveis (perturbações com maior probabilidade de ocorrência), fornecendo medidas de controlo para contrariar essas perturbações sempre que necessário. Entende-se por necessário quando o sistema se encontra num regime de funcionamento definido como não seguro.

Para compreender melhor os estados em que um sistema de energia pode residir, adopta-se o diagrama de operação e transição de estados, descrito em [5] e adaptado em [6], na Figura 2.1.

De acordo com os modos de operação apresentados na figura para um sistema de energia, o Estado Preventivo realiza uma avaliação sobre a capacidade do sistema suportar uma perturbação credível. Na incapacidade de garantir esta condição, uma acção de controlo preventivo pode levar o sistema de novo à região de operação segura. Neste estado, o objectivo da avaliação de segurança reside no equilíbrio entre a integridade do sistema e os custos económicos associados à sua exploração, conforme se descreve em [3].

A detecção do Estado de Emergência permite avaliar se o sistema de energia se encontra próximo da perda de integridade. Dada a natureza crítica da evolução das perturbações e da necessidade em operar o sistema nas melhores condições de segurança possíveis, este estado tem grandes necessidades de um tempo de resposta muito rápido em detrimento das condições económicas.

Pode acontecer que as medidas de controlo, quer preventivas quer de emergência, não sejam capazes de garantir o cumprimento das restrições de desigualdade e de trazer o sistema para um ponto de operação seguro e, nesta situação, deverão entrar em funcionamento elementos de protecção automática local para preservar os diversos componentes do sistema de operarem em

condições de risco e provocar danos irreversíveis. Estes acontecimentos levam a novas perturbações que podem resultar da interrupção parcial ou total de serviço, ilustrado na figura como *In-Extremis*. Consequentemente, o sistema entra em modo de Restabelecimento e são realizadas acções pelo operador do sistema para minimizar os efeitos da energia não fornecida através da ressincronização da produção e, posteriormente, da religação de cargas.

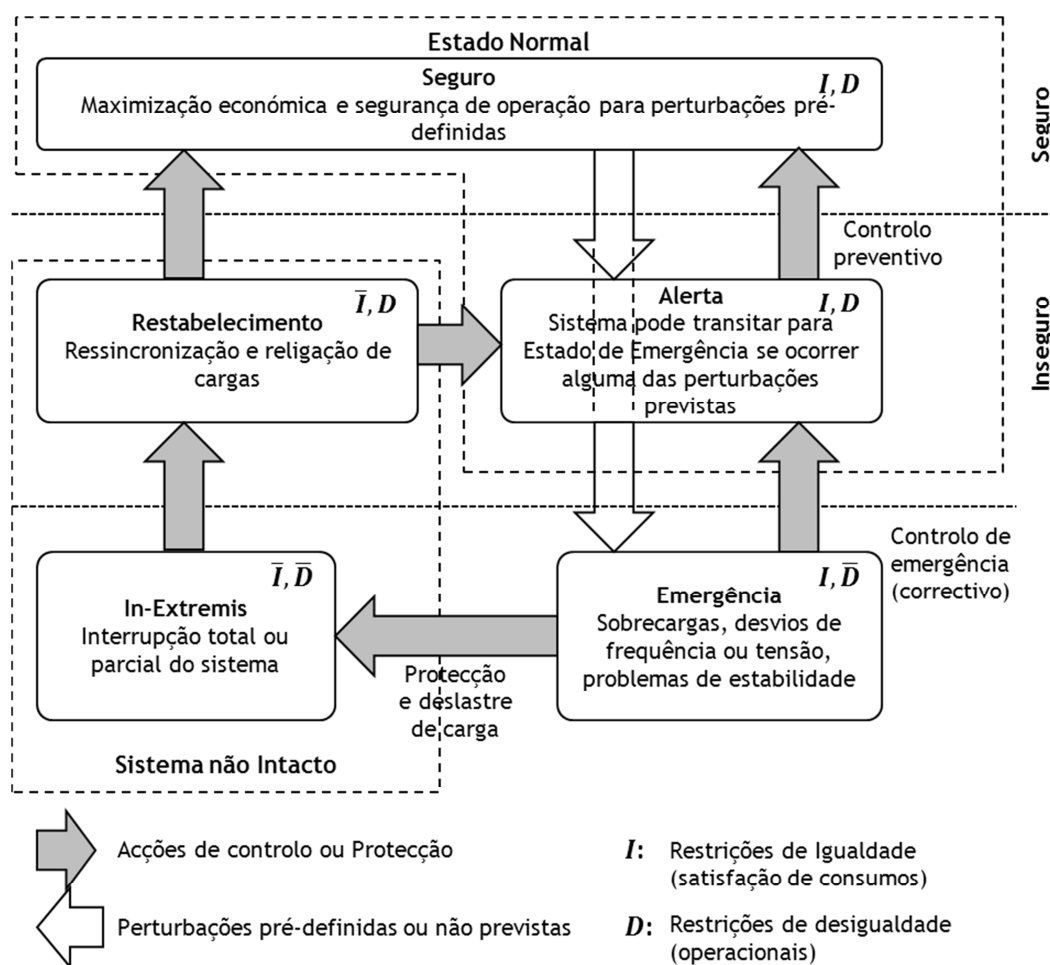


Figura 2.1 - Diagrama de operação e transição de estados do sistema (adaptado de [3] e [6])

Do ponto de vista dinâmico, os problemas de segurança que podem ocorrer para um sistema de energia dependem dos limites de operacionalidade definidos para a segurança de operação da rede de transporte. Esses limites encontram-se detalhados em [6], os quais se destacam a seguir:

- Perda de Estabilidade de Frequência: implica a violação da gama de valores admissíveis para o comportamento da frequência em regime transitório e dinâmico (de modo a garantir que não ocorre deslastre frequencimétrico de cargas e/ou grupos de geradores);

- Perda de Estabilidade Transitória: trata-se de um problema de estabilidade angular e resulta da perda de sincronismo do sistema eléctrico de energia após a ocorrência de uma grande perturbação;
- Perda de Estabilidade de Pequeno Sinal: é um problema de estabilidade angular que resulta da perda de sincronismo do sistema de energia após a ocorrência de uma pequena perturbação (oscilações pouco amortecidas e de baixa frequência);
- Perda de Estabilidade da Tensão: resultante da queda brusca da tensão (habitualmente associada ao aumento de carga em linhas de grande comprimento com produção reactiva deficiente na recepção);
- Sobrecargas temporárias nos ramos: resulta da violação das gamas de valores admissíveis das temperaturas máximas, definidas em projecto, para as linhas e transformadores e violação das distâncias mínimas ao solo, também definidas em projecto.

### **2.3 Problemas Associados à Segurança Dinâmica de Redes Interligadas com Produção Eólica**

À escala global, há uma tendência crescente de explorar os sistemas eléctricos com elevadas penetrações eólicas. Dada a flutuação da produção eólica, não é possível o escalonamento exacto da sua produção, quando se compara com a produção convencional, surgindo a necessidade de incluir nos cenários de exploração de um sistema de energia a previsão associada a este tipo de produção. Nos trabalhos realizados em [6] e [7], pode mesmo constatar-se que a influência das flutuações deste tipo de produção não é desprezável, influenciando a gestão das reservas secundárias (janela temporal entre os 10 a 15 minutos) e a gestão das reservas terciárias (a janela temporal pode exceder as horas). Portanto, a necessidade de previsões mais precisas para a produção eólica resulta em vantagens económicas e de estabilidade para o sistema de energia.

As flutuações de produção a que os aproveitamentos eólicos se encontram sujeitos têm origem nos seguintes tipos de fenómenos:

- Intermitência do vento, que implica flutuações da produção eólica devido às turbulências ou ocos de vento;
- Velocidade do vento que ultrapassa os limites mínimos ou máximos de operação dos geradores eólicos e implica a desligação destes;
- Ocorrência de curto-circuitos na rede que provoca a actuação das protecções de mínimo de tensão à saída dos aproveitamentos eólicos, implicando a desligação dos parques eólicos.

Em [6] são analisados trabalhos que permitem ao autor inferir que as variações de produção que resultam das flutuações de vento introduzidas num sistema interligado, não interferem com a gestão das reservas primárias. Este facto decorre de, na janela temporal referente à gestão dessas reservas, as flutuações serem tipicamente inferiores às flutuações provocadas pelo consumo e ainda muito inferiores às provocadas pela súbita saída de serviço do maior gerador convencional, sendo que o sistema de energia encontra-se já dimensionado para suportar esse tipo de perturbações. A perda de segurança dinâmica poderá estar em causa apenas em sistemas isolados onde a exploração tenha elevada penetração de produção eólica e esta se encontre concentrada em certas zonas geográficas.

Para o terceiro ponto definido no início da secção, a ocorrência de um curto-circuito em zonas mais sensíveis da rede traduz-se em quedas de tensão acentuadas podendo levar à actuação dos relés de mínimo de tensão, instalados em parques eólicos mais antigos e na proximidade do defeito. Estas perdas de produção podem ser consideravelmente grandes e ultrapassar as margens de segurança normalmente adoptadas.

Numa rede interligada, como a rede objecto de estudo deste trabalho, é necessário dispor de um conjunto de procedimentos para fazer face à ocorrência de curto-circuitos que levem à perda de grandes volumes de produção eólica, uma vez que, podem dar origem a problemas de origem dinâmica. Como se explica em [6], dos estudos efectuados para avaliação de segurança dinâmica, os operadores das redes de transporte passaram a incluir o requisito *ride-through fault capability*, um conjunto de especificações que requer do gerador eólico um tempo mínimo de permanência em serviço, onde esse tempo é função da cava de tensão provocada pelo curto-circuito.

Ainda no mesmo trabalho, [6], qualquer alteração brusca da produção em Portugal ou Espanha, resulta num acréscimo do trânsito na interligação França-Espanha, podendo colocar a rede Ibérica isolada da restante rede da UCTE. Deste isolamento pode resultar num défice energético e originar o colapso total ou parcial do sistema eléctrico Ibérico.

A tendência das redes de transporte é virem a ser exploradas mais próximo dos seus limites e, portanto, a elevada penetração de produção eólica irá provocar um maior congestionamento destas redes. Como consequência, é de esperar que se verifiquem súbitas perdas de produção e, em alguns casos, perdas de elevado volume, como consequência de curto-circuitos em determinados pontos da rede.

Uma das opções que se pode colocar à disposição dos operadores das redes de transporte é a inclusão, nos centros de controlo, de sistemas de monitorização de segurança que vigiem as perturbações eólicas que possam pôr em risco a segurança de operação do sistema. No sentido de

dotar o sistema de maior autonomia, é ainda importante que esses sistemas de monitorização possam fornecer, em tempo real, medidas de controlo preventivo para fazer face à eventual ocorrência de perturbações que levem à perda de grandes volumes de produção eólica.

Devido aos problemas que se levantam neste capítulo, para a avaliação de segurança de operação de um sistema interligado (nomeadamente, da rede de teste será apresentada à frente no Capítulo 4), define-se como problema de segurança em análise a ocorrência de sobrecargas em regime quasi-estacionário nos ramos de transmissão de uma determinada área de controlo de um sistema interligado, na sequência da ocorrência de uma perturbação que resulte numa súbita perda de produção eólica na área de controlo em análise, que viole os limites de operacionalidade especificados.

## 2.4 Conclusões

Neste capítulo realizou-se uma descrição do problema de segurança dinâmica em análise nesta dissertação. Começou por se definir o conceito de segurança dinâmica e de que factores depende, para melhor se compreender o contexto em que este trabalho se insere. Foi evidenciada a importância do aumento de precisão nas previsões de produção eólica e da especificação de requisitos mínimos para os parques eólicos para uma gestão mais eficiente do parque energético.

Foram referidas as perturbações que a produção eólica pode introduzir num sistema de energia e quais os problemas de segurança dinâmicos que se encontram associados. Refere-se que em redes isoladas, uma elevada penetração eólica pode levantar problemas de segurança dinâmica na gestão da reserva primária resultantes das flutuações do vento e da concentração geográfica dos aproveitamentos eólicos. A ocorrência de um curto-circuito em pontos críticos do sistema de energia que possa levar à perda de grandes volumes de produção eólica, corresponde à perturbação que pode colocar em risco a segurança das redes interligadas.

# Capítulo 3:

## Modelos de Regressão para Avaliação de Segurança Dinâmica de um SEE

### 3.1 Introdução

O trabalho que se descreve neste documento tem como objectivo a avaliação de desempenho da capacidade de previsão de uma estrutura de aprendizagem automática (AA) híbrida. Essa estrutura resulta da aplicação de um modelo de regressão linear (LRM) para cada folha de uma árvore de regressão (RT).

As funções que pretendem para as estruturas de AA consistem na análise da capacidade que um sistema de energia tem para suportar perturbações credíveis e fornecer medidas de controlo para contrariar essas perturbações. Recorre-se à aprendizagem automática supervisionada, [6], para a partir de um conjunto de exemplos de entrada-saída se obter um modelo funcional, a partir de inferência indutiva, que explique as relações de causa-efeito observadas entre as entradas e as saídas. A aprendizagem automática consiste, pois, na extracção de informação sintética de alto nível a partir de um conjunto de dados que contêm uma grande quantidade de informação de baixo nível, [3] e [6]. De [3] pode constatar-se que a extracção sintética de alto nível por processos de aprendizagem automática pode ela mesma ser complementar e fornecer informação mais importante que a extraída por processos clássicos de análise caso a caso.

Neste capítulo, é apresentada a metodologia seguida no presente trabalho, para a obtenção e avaliação de desempenho do modelo funcional de aprendizagem automática que realize a avaliação de segurança do problema dinâmico em análise. Na secção 3.4 explica-se em detalhe a técnica utilizada para obtenção de um LRM e na secção 3.5 descreve-se o processo de treino de uma árvore de regressão linear. Para obter a interpretabilidade que se deseja da aplicação do método de aprendizagem automática, na secção 3.2 são apresentados um conjunto de índices para estimação dos erros realizados na previsão. Na secção 3.3 descreve-se a técnica utilizada no presente trabalho para evitar problemas de sobre-adaptação das estruturas de AA treinadas.

### 3.2 Estimativas do Erro de Previsão

Para se obter uma estimativa do desempenho de previsão de um modelo, qualquer que seja o modelo funcional utilizado, é necessário conhecer o valor real da variável de saída  $y$ . A utilização deste verdadeiro valor permite a comparação com previsão do modelo e, assim, avaliar o seu desempenho, como indicado em [8]. Uma vez que a variável de um problema de regressão é numérica, os erros de medição, para cada elemento  $i$  da amostra, reflectem a diferença entre os valores previstos para o modelo funcional,  $\hat{y}_i(f)$ , e o valor real,  $y_i$ . Dado que a variável de saída é do tipo numérico, ao índice numérico que quantifica a estimativa para a precisão do modelo, dá-se o nome de erros de regressão.

Para o presente trabalho foram utilizados os seguintes índices:

- Erro Quadrático Médio (*Mean Squared Error*), calculando-se da seguinte forma:

$$MSE(f) = \frac{1}{ne} \sum_{i=1}^{ne} (y_i - \hat{y}_i(f))^2 \quad (3.1)$$

onde  $ne$  é o número de exemplos da amostra utilizada.

- Erro Absoluto Médio (*Mean Absolute Error*), determinado por:

$$MAE(f) = \frac{1}{ne} \sum_{i=1}^{ne} |y_i - \hat{y}_i(f)| \quad (3.2)$$

- Erro Quadrático Médio Relativo (*Relative Mean Squared Error*), determinado por:

$$RE(f) = \frac{MSE(f)}{MSE(\bar{y})} = \frac{\frac{1}{ne} \sum_{i=1}^{ne} (y_i - \hat{y}_i(f))^2}{\frac{1}{ne} \sum_{i=1}^{ne} (y_i - \bar{y})^2} \quad (3.3)$$

onde

$\bar{y} = \frac{1}{ne} \sum_{i=1}^{ne} y_i$ : valor médio amostral de  $y$ ;

$MSE(\bar{y})$ : desvio quadrado médio amostral de  $y$ .

O índice numérico  $MAE$  avalia a média dos desvios absolutos das previsões e, quando comparado com o  $MSE$ , é menos sensível à existência de casos excepcionais para os quais a previsão do modelo se afaste, consideravelmente, do verdadeiro valor da saída sendo, portanto, mais comum o uso do índice  $MSE$ . Como se pode verificar pela equação do  $RE$ , este índice estabelece uma razão entre o  $MSE$  e o valor que seria obtido se a função de previsão correspondesse ao valor médio amostral.

Para avaliação de segurança dos sistemas eléctricos, é necessário definir um limiar de segurança para um valor da saída  $y$ . Para o devido efeito, estima-se o erro do modelo funcional em relação à

sua capacidade de prever se o sistema é “seguro/inseguro”, a partir do cálculo dos seguintes erros de classificação:

- Erro de Classificação Global (*Global Classification Error*), determinado por:

$$GCE(f) = \frac{n_{CE}(f)}{ne} \quad (3.4)$$

onde  $n_{CE}(f)$  é o número de elementos da amostra que são mal classificados por  $f$ .

- Erro de Falso Alarme (*False Alarm Error*), determinado por:

$$FA(f) = \frac{n_{FA}(f)}{ne_S} \quad (3.5)$$

onde

$n_{FA}(f)$ : número de elementos seguros da amostra para os quais  $f$  forneceu uma classificação de “inseguro” (i.e., nº de falsos alarmes);

$ne_S$ : número de elementos seguros da amostra.

- Erro de Falha de Alarme (*Missed Alarm Error*), determinado por:

$$MA(f) = \frac{n_{MA}(f)}{ne_I} \quad (3.6)$$

onde

$n_{MA}(f)$ : número de elementos inseguros da amostra para os quais  $f$  forneceu uma classificação de “seguro” (i.e., nº de falha de alarmes);

$ne_I$ : número de elementos inseguros da amostra.

Pelas razões evidenciadas por Vasconcelos em [6], o  $MA$  é o erro de classificação que apresenta maior relevância, uma vez que estima a taxa de ocorrência de situações em que o modelo funcional não será capaz de detectar a perda de segurança do sistema eléctrico.

### 3.3 Avaliação da Capacidade de Generalização

À amostra utilizada para o processo de treino dá-se o nome de conjunto de treino (denominada, na literatura anglo-saxónica, *Learning Set* ou *Training Set*). Durante o processo de treino pode existir uma fase de avaliação do modelo funcional, pelo que essa avaliação é feita recorrendo a uma amostra da mesma população que não tenha sido usada para a construção do modelo, a essa amostra dá-se o nome de conjunto de validação. O conjunto de validação, diferente do conjunto de treino mas com base na mesma população, permite que se obtenha uma boa capacidade de generalização,

isto é, um modelo com capacidade de prever com precisão a saída de futuros exemplos pertencentes à mesma população do conjunto de treino, mas que não tenham sido usados para treino.

Durante o processo de treino são utilizados critérios de selecção para minimização dos erros de treino, pelo que as estimativas do erro de previsão a partir do conjunto de treino se traduzem em valores optimistas (denominado, na literatura anglo-saxónica, *biased estimates*). Deste modo, há um risco de a estrutura ficar sobre-adaptada ao conjunto de treino (denominado, na literatura anglo-saxónica, *overfitting*), perdendo, por isso, capacidade de generalização.

Assim, para avaliar correctamente a capacidade de generalização do modelo funcional resultante do processo de treino, é utilizada uma terceira amostra da mesma população, denominada conjunto de teste (na literatura anglo-saxónica, *Testing Set*).

Para aplicação da técnica de *Holdout*, a amostra disponível é aleatoriamente dividida em dois sub-conjuntos disjuntos, um dos sub-conjuntos é utilizado para o processo de treino e o restante (o *Holdout*) para a avaliação de desempenho final do modelo funcional treinado. Para partição da amostra, são utilizados cerca de 40% dos exemplos para o conjunto de teste (o *Holdout*) e os restantes 60% para o processo de treino.

### 3.4 Modelos de Regressão Linear (LRM)

Por definição, o termo regressão compreende a relação entre um conjunto de variáveis observadas, a amostra, com o valor de saída  $y$  de um modelo funcional. Com base num conjunto de  $n$  medidas (número de variáveis de entrada), é possível construir um modelo de regressão linear a partir de uma função paramétrica do tipo:

$$y_i = b_0 + b_1 \cdot x_{1i} + \dots + b_j \cdot x_{ji} + \dots + b_n \cdot x_{in} + \varepsilon_i \quad (3.7)$$

onde

$y_i$ : valor da variável de saída para o elemento  $i$ ;

$b_i$ : parâmetro  $i$  da relação linear procurada ( $i = 0, 1, \dots, n$ );

$x_{ip}$ : valor da variável de entrada  $p$  para o elemento  $i$  ( $p = 1, 2, \dots, n$ );

$\varepsilon_i$ : erro de previsão do modelo para o elemento  $i$ .

Para o modelo que se pretende criar, o problema de regressão linear consiste na minimização do quadrado dos erros de previsão, para estimar os valores dos parâmetros  $b_i$  ( $i = 0, 1, \dots, n$ ):

$$\text{mín: } \sum_{i=1}^{n_e} (y_i - \hat{y}_i(\text{LRM}))^2 \quad 3.8$$

onde

$ne$ : número de elementos da amostra utilizada para treino;

$\hat{y}_i(LRM)$ : valor fornecido pelo LRM para prever o valor de  $y_i$ .

Tal como se descreve em [9], é importante que os erros do modelo,  $\varepsilon$ , tenham média nula, não se encontrem correlacionados e apresentem variação constante. Para evitar que os estimadores tenham uma elevada variância, e causar problemas de instabilidade numérica [6], utiliza-se uma ferramenta informática com modelos programados para a obtenção de LRM que realiza a selecção dos regressores. No desenvolvimento dos trabalhos da presente dissertação, para construção de LRM, recorreu-se ao método passo a passo, da aplicação SPSS [10]. A selecção dos regressores é efectuada pelo método passo a passo. As variáveis de entrada são introduzidas no modelo de regressão linear sequencialmente, sendo um processo que engloba os seguintes passos:

- **Passo 1:** a primeira variável a dar entrada no modelo explica a maior influência sobre a variação total da saída  $y$  (isto é, quando se aplica um modelo relacional com apenas uma variável). Neste contexto, a explicação fornecida por cada um dos LRM definidos é obtida a partir de um teste de análise de variância (ANOVA, importância referida em [9], que segue uma distribuição Gaussiana) para cada um dos modelos.
- **Passo 2:** a variável de entrada a seleccionar que se segue é aquela que, entrando no modelo iniciado no passo anterior, melhor explica a maior influência seguinte sobre a variação total da saída  $y$ . Também aqui, a explicação é fornecida por um teste ANOVA, para cada um dos múltiplos LRM considerados. O ciclo de selecção de variáveis de entrada termina se nenhum dos novos modelos fornecer uma explicação adicional de variação total da saída  $y$  com significado estatístico. Pode ainda terminar o ciclo se já tiverem sido incluídos no modelo todos os regressores. Entende-se por regressores o conjunto das variáveis de entrada que entram para construção do modelo de regressão linear.

Conforme se descreve em [6], no método passo a passo, a entrada de um regressor para o modelo não é definitiva, já que em cada passo do processo é examinada a exclusão de regressores que tenham sido incluídos em passos anteriores. Deste modo, evita-se o risco de redundância e a inclusão de algum regressor menos relevante já adicionado ao conjunto de regressores do modelo.

A qualidade da relação entre a variável de saída  $y$  e os regressores  $x_1, x_2, \dots, x_n$  tende a aproximar-se de uma relação mais linear à medida que a maximização do coeficiente de correlação múltipla,  $R_{yx_1x_2\dots x_n}$ , se aproxima de 1 (se igual a 1 então a relação será linear).

A duração de um processo computacional de extracção de uma estrutura de LRM leva cerca de um minuto para um modelo de regressão linear com 63 variáveis candidatas, 2758 exemplos para o conjunto de treino e aplicando o método de selecção passo a passo (*stepwise selection*) para os regressores no *software* SPSS. No contexto global do trabalho, pode dizer-se que este processo é rápido.

### 3.5 Árvores de Regressão Linear (LRT)

Tal como descrito em [11], a estrutura de uma árvore de regressão (denominado na literatura anglo-saxónica *Regression Tree (RT)*) é em tudo semelhante a uma árvore de decisão, uma vez que também é formada por conjuntos de nós de decisão, contudo, diferem no resultado, que em vez de uma categoria fornece um escalar.

A árvore de regressão é uma ferramenta de decisão cuja estrutura treinada pode revelar um conjunto de regras interpretáveis e com razoável precisão, como se indica em [12]. Descreve-se na secção anterior que o problema de regressão consiste na criação de um modelo funcional que relacione o valor da saída  $y$  com os regressores (variáveis de entrada)  $x_1, x_2, \dots, x_n$ . Para aplicação desta técnica, assume-se que a variável de saída  $y$  a emular é um escalar, utilizando-se para a resolução de problemas de regressão. O processo de treino de uma RT consiste num algoritmo de partição do conjunto de treino, em cada nó da árvore, em dois subconjuntos disjuntos (tratando-se, então, de uma árvore binária). A divisão em dois novos nós segue o critério do erro quadrático médio (MSE, descrito na equação  $MSE(f) = \frac{1}{ne} \sum_{i=1}^{ne} (y_i - \hat{y}_i(f))^2$  (3.1)), isto é, resulta da minimização da variação do valor de  $y$  dos elementos de treino contidos em cada subconjunto. As árvores de regressão utilizam, como função de previsão para cada folha com um conjunto de elementos da amostra  $ne$ , o valor médio de  $y$ :

$$\hat{y}_i = \bar{y} = \frac{1}{ne} \sum_{i=1}^{ne} y_i \quad 3.9$$

Assim, o MSE dos valores de  $y$  em cada folha correspondem ao desvio quadrático médio de  $y$ :

$$MSE(\bar{y}) = \frac{1}{ne} \sum_{i=1}^{ne} (y_i - \bar{y})^2 \quad 3.10$$

Em [6] chama-se à atenção que se for considerado um valor constante para a previsão de  $y$  em cada folha, então o valor escolhido deverá ser  $\bar{y}$ , uma vez que corresponde à constante que minimiza o valor esperado do MSE. Descreve-se em [12] os tipos de abordagem realizados para melhorar a capacidade de previsão de uma RT. Esse mesmo trabalho, [12], sugere que por aplicação a diversos problemas de regressão e de forma empírica, se conclui que a utilização de LRM nas folhas de uma RT pode, em muitos casos, ter uma precisão superior, quando se compara com o treino isolado de um LRM ou de uma RT. Como o objectivo do presente trabalho é a avaliação de segurança por aplicação

de restrições de segurança do tipo linear mais precisas, foi utilizada a capacidade de previsão do modelo descrito em [12] que resulta do treino de modelos de regressão linear nas folhas de uma árvore de regressão. O processo foi denominado pelo autor em [6] de árvore de regressão linear e a estrutura da RT é detalhadamente explicada nesse trabalho.

Para o treino de cada RT foi utilizado um *software* desenvolvido nos trabalhos do autor de [6] e que classifica um nó como terminal se  $N(t) = 1$  ou se  $MSE(t) = 0$  e o mesmo aplica a técnica de podagem *Lowest Statistical Support* ( $N(t)$  é o número de exemplos de treino que pertencem ao nó  $t$ ;  $MSE(t)$  é o erro quadrático médio dos exemplos de treino pertencentes ao nó  $t$ ).

Do ponto de vista computacional, pode dizer-se que este processo é demorado, já que para o treino de uma RT com 90 variáveis de entrada e 2578 exemplos de treino o processo demorou cerca de 1 hora. Quando executado num computador com um processador Intel Core 2 Duo a 1,83 GHz, 1 GB de memória RAM e sistema operativo Windows 7.

### **3.5.1 Treino de uma RT**

Para a construção de uma árvore de regressão é necessário definir inicialmente o número de nós adequado para cada problema em análise. Em [6] regista-se que a definição de um número muito pequeno de nós conduz a RT a aproximar-se pouco do conhecimento funcional pretendido. Por outro lado, um número muito elevado pode levar à perda de generalização da RT.

Para o processo de treino de uma RT adoptou-se o critério de podagem definido em [6], no qual, partindo de um conjunto de treino (aplicado na raiz da árvore), o treino da RT é sequencial e engloba a execução das etapas que se seguem:

- Etapa 1: geração de uma RT de grande dimensão,  $T_{max}$ , resultante dos processos:
  - Divisão dos exemplos contidos em cada nó não-terminal da árvore em dois subconjuntos disjuntos;
  - Classificação de cada novo nó como sendo terminal ou não-terminal.

Esta etapa apenas divide a raiz da árvore e inicia a geração de  $T_{max}$ . A cada divisão de um nó, são criados dois novos nós e cada um desses nós é classificado em terminal/não-terminal, com base nas regras de paragem adoptadas. O critério de paragem para a geração de  $T_{max}$  estabelece-se quando todos os nós não-terminais da árvore já tiverem sido divididos.

- Etapa 2: geração de um conjunto de árvores podadas da estrutura  $T_{max}$  que resulta da etapa anterior.
- Etapa 3: recorrendo a um critério empírico de selecção em que se estabelece um compromisso entre a dimensão e a capacidade de generalização, medida pelo MSE(RT) fornecido para o conjunto de teste, é seleccionada a estrutura que, de entre o conjunto das árvores podadas, fornece melhores desempenhos.

A divisão de um nó realiza-se por aplicação de um teste de divisão óptimo do tipo:

$$\{atributo_k(exemplo) > u_k\} \quad (3.11)$$

onde

$atributo_k(exemplo)$ : é o valor que a variável de entrada  $k$  toma para um determinado exemplo do conjunto de treino;

$u_k$ : é o valor limite escolhido para o atributo  $k$ .

Assim, por aplicação deste teste a um nó  $t$ , são criados dois novos nós por sucessão,  $t_L$  e  $t_R$ , resultando os conjuntos:  $\{atributo_k(exemplo) > u_k\}$  e  $\{atributo_k(exemplo) \leq u_k\}$ .

O teste de divisão óptimo de cada nó  $t$ , é o que maximiza a redução do erro de previsão em relação ao do nó  $t$ . Em [12], sugere-se que para calcular o erro de previsão do nó  $t$ , o teste de divisão óptimo do nó  $t$  é aquele que maximiza:

$$\Delta MSE(t) = MSE(t) - \frac{N(t_L)}{N(t)} MSE(t_L) - \frac{N(t_R)}{N(t)} MSE(t_R) \quad 3.12$$

onde

$N(t)$ ,  $N(t_L)$  e  $N(t_R)$ : são o número de exemplos de treino que pertencem ao nó  $t$ ,  $t_L$  e  $t_R$ , respectivamente;

$MSE(t)$ ,  $MSE(t_L)$  e  $MSE(t_R)$ : são os erros quadráticos médios dos exemplos de treino pertencentes aos nós  $t$ ,  $t_L$  e  $t_R$ , respectivamente.

No algoritmo de treino de uma RT, o processo de divisão de um nó é o mais extenso em tempo, uma vez que realiza uma procura exaustiva por todos os testes de divisão possíveis do nó e para vários valores de  $u_k$ .

Por omissão, a classificação de um nó é não-terminal. Essa classificação muda para terminal quando se verifica uma de duas condições:

- Regra 1: quando não se consegue reduzir  $MSE(t)$  com significado estatístico, por o número de exemplos existentes num nó ser insuficiente. No caso da presente Tese, supôs-se o caso extremo de deixar que se atinja  $N = 1$ ;
- Regra 2: o  $MSE(t)$  já tem um valor suficientemente menor do que  $MSE(raiz)$ . No caso da presente Tese, supôs-se também o caso extremo de deixar que se atinja  $MSE(t) = 0$ .

Para a geração de um conjunto de árvores podadas, é aplicada a técnica *Lowest Statistical Support Algorithm* descrita em [6] e que consiste em definir o próximo nó  $t$  a podar como aquele que minimiza o número de exemplos de treino que contém, ou seja, que minimiza  $N(t)$ . Esta especificação permite eliminar o nó que tem menor influência na degradação do erro de previsão dessa estrutura, uma vez que quanto maior o número de exemplos utilizados, mais preciso se torna um estimador.

### **3.6 Metodologia adoptada para obtenção de uma Estrutura de Avaliação Dinâmica de um SEE**

A metodologia que se descreve no Capítulo 3 do trabalho [6], referente à obtenção de uma estrutura de avaliação dinâmica de um SEE, serviu de base para o presente trabalho. Os dados gerados no âmbito desse trabalho foram também os que serviram para treinar as estruturas de segurança, para a mesma perturbação de segurança, e que a seguir são descritos.

A Fase 1 e a Fase 2 presentes no trabalho em [6] encontram-se a seguir resumidas, sendo a Fase 3 e Fase 4 adaptadas ao contexto desta dissertação:

- Fase 1: identificação do problema de segurança a partir do tipo de perturbação, sendo no presente trabalho de *avaliação da ocorrência de sobrecargas em regime quasi-estacionário nos ramos de transmissão de uma determinada área de controlo de um sistema interligado, na sequência da ocorrência de uma perturbação que resulte numa súbita perda de produção eólica na área de controlo em análise, que viole os limites de operacionalidade especificados*; identificação das condições de operação a alterar entre os cenários do conjunto de dados, os quais devem ter uma influência relevante para o tipo de comportamento dinâmico em análise; definição do vector de variáveis de entrada candidatas, garantindo que não incluam características que estejam entre elas fortemente correlacionadas; e a adopção de índices de segurança para determinar se o cenário de operação será classificado como “inseguro” ou “seguro”.

- Fase 2: geração do conjunto de dados que reflecta correctamente o comportamento do sistema eléctrico, com base na técnica de sorteio das condições de operação do sistema, através do método de Monte Carlo.
- Fase 3: treino das estruturas de segurança de AA, caracterizadas na secção 3.4 e secção 3.5, tendo sido seguidas as seguintes etapas:

- utilização dos dados obtidos na Fase 2 e realização da normalização desses dados para que o conjunto de dados tenha valor médio nulo e variância unitária, de acordo com a equação:

$$a_N = \frac{a - \bar{a}}{s_a} \quad (3.13)$$

onde  $\bar{a}$  e  $s_a^2$  correspondem ao valor médio e variância amostral da variável  $a$  (variável original), respectivamente, e  $a_N$  (variável normalizada). Para o treino das LRM são utilizados valores normalizados dos dados, ao passo que para o treino das RT são utilizados dados não normalizados.

- treino de um LRM, para previsão da saída  $y$ , tendo sido incluída uma selecção automática de características, de entre o vector de variáveis de entrada candidatas, utilizando o método passo a passo (na literatura anglo-saxónica, *stepwise selection*) que se disponibiliza no *software* SPSS.

- construção de uma RT para emular cada comportamento em análise, com um número pré-definido de nós e na qual se consideram todas as características candidatas. Para cada divisão do conjunto de dados definida pela árvore treinada é posteriormente treinado um LRM, sendo aqui incluída a selecção de características fornecida pelo método passo a passo.

- Fase 4: avaliação de desempenho do processo de treino, crucial para se realizar a comparação entre as capacidades de previsão fornecidas pelas estruturas de segurança, recorrendo à validação da capacidade de previsão do modelo funcional com o conjunto de teste (aplicação da técnica de *Holdout*, descrita na secção 3.3).

### **3.7 Conclusões**

Neste capítulo foram apresentadas as técnicas de aprendizagem automática exploradas no presente trabalho, nomeadamente:

- Modelo de Regressão Linear;
- Estrutura híbrida que resulta da aplicação de um modelo de regressão linear em cada folha de uma árvore de regressão.

Foi ainda descrita a metodologia que foi seguida para a obtenção deste tipo de estruturas, para emular o problema de segurança em análise, bem como para avaliar a capacidade de previsão das estruturas treinadas. No âmbito desta metodologia, para as fases 1 e 2 (de definição do problema de segurança em análise e de geração de dados), foram utilizados os dados que resultaram dos trabalho descritos em [6].



## Capítulo 4: Estrutura da Rede Interligada em Análise

### 4.1 Introdução

Para implementar as metodologias de avaliação de segurança utilizadas neste trabalho, utiliza-se uma rede interligada de teste criada no âmbito dos trabalhos descritos em [6]. Esta rede é apresentada de forma sucinta na secção 4.2, para que se possa compreender a sua estrutura e organização.

O conjunto de dados utilizados para a rede interligada de teste criada resultou de um procedimento de geração de dados computacional (efectuado no âmbito dos trabalhos descritos em [6]), o qual traduz o comportamento dinâmico deste sistema eléctrico na sequência da ocorrência da perturbação em análise. Esta informação encontra-se detalhada em [6], pelo que não é objectivo do presente trabalho a sua análise. Importa referir que esses dados reflectem as especificações realizadas para a correcta caracterização do problema de segurança em análise - *avaliação da ocorrência de sobrecargas em regime quasi-estacionário em ramos de transmissão de redes interligadas que explorem elevadas penetrações de produção eólica*. Assim, para compreensão deste documento, serão apresentados os resultados mais relevantes, presentes do trabalho de investigação anteriormente referido.

### 4.2 Estrutura da Rede Interligada

A rede interligada criada baseia-se num sistema interligado real, no qual se prevê um grande volume de exploração eólica. Desta forma, pretende-se dispor de um caso de estudo que reflecta os procedimentos de dimensionamento habituais. Para compreender a estrutura da rede utilizada para teste, apresenta-se um esquema unifilar na Figura 4.1.

A rede pode ser descrita como um sistema com duas áreas de controlo: a Área de Controlo 1 corresponde a uma aproximação da rede de transporte portuguesa, já a Área de Controlo 2 pretende representar um equivalente dinâmico do sistema eléctrico espanhol e da restante rede da UCTE.

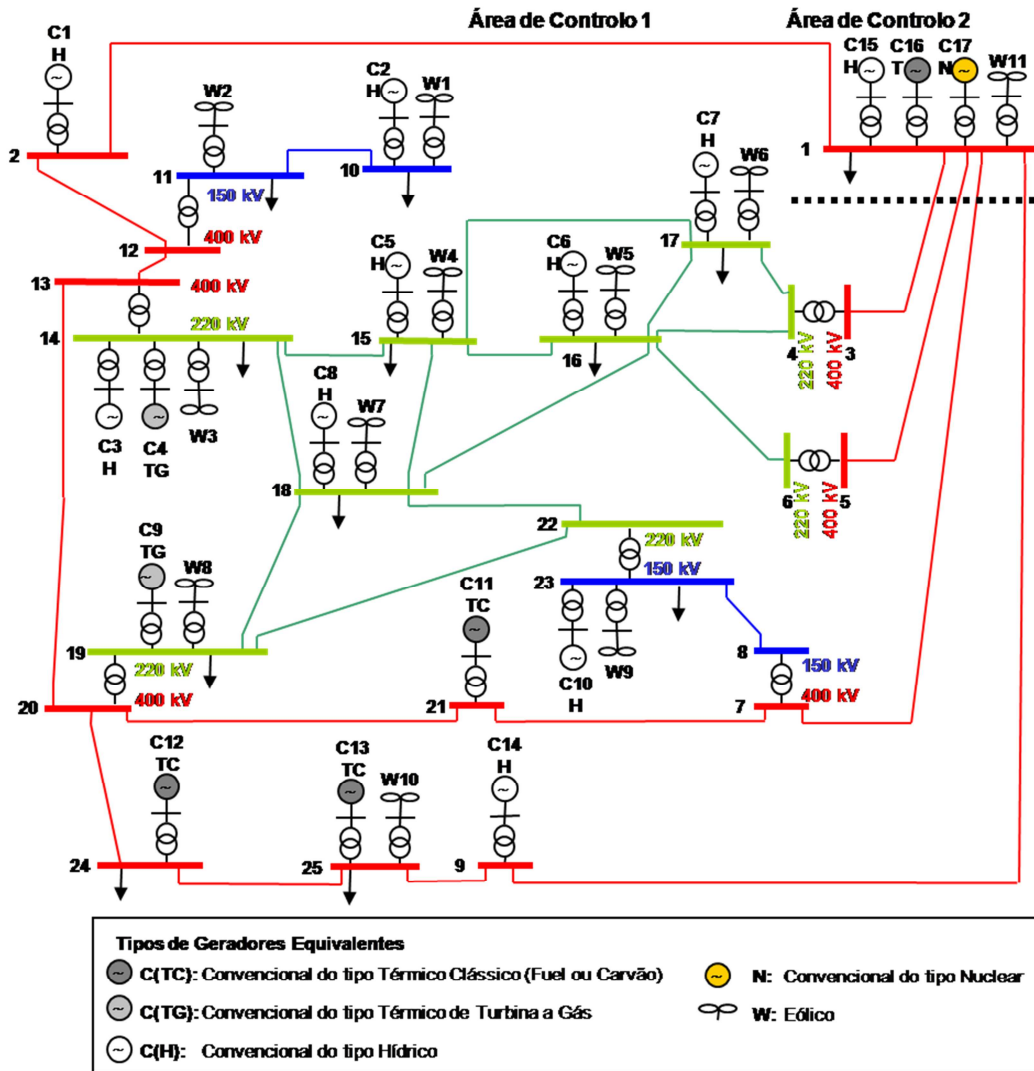


Figura 4.1 - Esquema unifilar da rede interligada de teste (obtido de [6])

Como já se referiu, a rede interligada de teste criada pretende reflectir a simplificação de um sistema interligado real. Este sistema é constituído por três níveis de tensão nominal (150, 220 e 400 kV) e contém:

- 24 barramentos, onde 11 abastecem a rede de distribuição (encontram-se modelizados como consumos através de uma carga equivalente) e 10 dispõem de baterias de condensadores (um por cada gerador eólico da área de controlo 1);
- 28 linhas de transmissão (podendo, cada uma, corresponder ao equivalente de linhas em paralelo);

- 7 transformadores de transmissão (podendo, cada um, corresponder ao equivalente de transformadores em paralelo);
- 14 geradores convencionais equivalentes (3 do tipo térmico clássico, 2 do tipo térmico de turbina a gás e 9 do tipo hídrico), correspondendo, cada um, ao modelo equivalente do conjunto de geradores com iguais características e que se encontrem em serviço na mesma central;
- 10 geradores eólicos equivalentes, correspondendo, cada um, ao modelo equivalente do conjunto de geradores eólicos com iguais características e que se encontrem em serviço e a partilhar o mesmo ponto de ligação à rede de transporte.

Quando se observa a Figura 4.1, verifica-se que o barramento 1 reflecte a modelização da área de controlo 2 por um único barramento no qual se incluem os geradores equivalentes de diversos tipos (1 térmico, 1 hídrico, 1 nuclear e 1 eólico), uma bateria de condensadores e o consumo total dessa área de controlo. As 5 linhas que ligam a área de controlo 1 à área de controlo 2, (linhas 1-2, 1-3, 1-5, 1-7 e 1-9) têm uma impedância equivalente vista do ponto de interligação da rede de transporte portuguesa com a área de controlo vizinha, sendo, portanto, pertencentes ao equivalente eléctrico da área de controlo 2, conforme referido em [6].

A redução do sistema em análise aos modelos equivalentes constitui uma simplificação útil para a criação da rede interligada de teste pois permite manter o foco na segurança de operação da área de controlo 1, sem que ocorra perda de informação relevante para a correcta caracterização da avaliação de segurança do sistema, isto é, redução a equivalentes que não interfiram na obtenção da correcta simulação do comportamento em regime estacionário e dinâmico da intensidade de corrente nos ramos da rede de transporte.

### **4.3 Parametrização e Resultados dos Dados Gerados para a Rede Interligada de Teste**

De forma sucinta, o comportamento dinâmico da rede interligada de teste foi obtida recorrendo a uma ferramenta analítica própria para a simulação do comportamento temporal de sistemas eléctricos. Conforme descrito em [6], essa ferramenta calcula o trânsito de potências pelo método de Newton-Raphson e resolve as equações diferenciais que caracterizam o comportamento dinâmico do sistema através do método de Runge-Kutta de 4ª ordem. O autor em [6] realiza a validação das escolhas efectuadas ao nível do valor para os parâmetros dos modelos analíticos da rede interligada criada, por um processo de comparação entre os resultados obtidos pela ferramenta de simulação e os resultados obtidos pelo programa PSS/E (este último programa utiliza um modelo completo da rede de transporte Ibérica e um modelo equivalente da rede francesa).

A perturbação considerada para simulação *consiste na ocorrência de um curto-circuito trifásico simétrico franco na extremidade mais próxima do barramento 15 de uma das duas linhas que se consideraram estar em paralelo a ligar os barramentos 15 e 16, tendo este defeito sido eliminado, após 300ms, pela saída de serviço da linha com defeito*. Foi ainda definido para simulação desta perturbação que as protecções de mínimo de tensão dos geradores eólicos equivalentes actuam instantaneamente sempre que a tensão no barramento de interligação do gerador à rede atingir um valor inferior a 0,9 p.u.

Com base nas especificações para as variáveis de entrada candidatas apresentadas em [6], o mesmo autor efectuou a geração de 4032 cenários de operação e 720 cenários de operação adicionais, que incluem ajustes específicos aos parâmetros de geração de modo a obter uma representação elevada de produção de potência eólica na área de controlo 1. Dos  $4032 + 720 = 4752$  cenários de operação, foram seleccionados 4596 cenários de operação, como sendo válidos, de modo a garantir que, durante o regime permanente de pré-perturbação, não ocorra qualquer uma das seguintes restrições:

- nenhuma das máquinas convencionais se encontra em sobrecarga;
- o trânsito de potências, que transita em qualquer ramo, viola a capacidade de transmissão em regime permanente desse mesmo ramo;
- a tensão em todos os barramentos de MAT encontra-se dentro da gama de valores, definida por  $[0.93; 1.1]p.u.$

Estabelece-se como critério de segurança, atendendo a sobrecargas temporárias nos ramos de transmissão na área de controlo 1, a violação de uma carga de 120% em relação ao valor máximo admissível da corrente em regime permanente para um dado ramo de transmissão, aos 120 segundos após a ocorrência da perturbação. O cenário é considerado inseguro se esta violação ocorrer, e seguro para os restantes casos.

A partir do critério de segurança descrito, foram identificadas 7 linhas de transmissão como críticas, uma vez que para determinadas condições de operação violam o critério de segurança adoptado. São, portanto, as linhas 15-16, 15-17, 16-4, 16-6, 14-18, 16-18 e 20-21.

## **4.4 Conclusões**

Neste capítulo realizou-se uma descrição da rede interligada de teste com duas áreas de controlo e alguns dos pontos considerados importantes para a geração de dados deste sistema eléctrico. A metodologia detalhada para a geração de dados encontra-se especificada em [6]. Caracterizou-se, também, o critério de segurança de operação relativamente às sobrecargas temporárias que podem ocorrer nos ramos de transmissão. Com este conjunto de dados gerado, foi possível realizar o treino das estruturas de segurança que se apresentam no Capítulo 5, sobre as quais se realizaram a avaliação de desempenho em termos de capacidade de previsão.



# Capítulo 5: Avaliação de Segurança Dinâmica para a Rede de Teste

## 5.1 Introdução

Os resultados obtidos, em termos da avaliação de desempenho das estruturas de segurança treinadas, são apresentados neste capítulo. O principal interesse é a rápida avaliação da robustez de exploração, em termos de segurança dinâmica.

No Capítulo 4 define-se o problema de segurança dinâmico em análise - *avaliação da ocorrência de sobrecargas em regime quasi-estacionário nos ramos de transmissão de uma determinada área de controlo de um sistema interligado, na sequência da ocorrência de uma perturbação que resulte numa súbita perda de produção eólica na área de controlo em análise*, para o qual se pretende realizar a avaliação do problema de segurança. Para o devido efeito, treinaram-se o seguinte tipo de estruturas:

- Modelos de Regressão Linear (LRM);
- Estrutura híbrida que resulta da aplicação de um LRM a cada folha de uma árvore de regressão, ou seja, uma árvore de regressão linear (LRT). Para este tipo de estrutura foram consideradas as hipóteses de a árvore ter 2 folhas e 4 folhas.

Estas técnicas foram descritas na secção 3.5 do Capítulo 3, com a apresentação das opções tomadas para efectuar o treino deste tipo de estruturas de avaliação de segurança.

Para a rede interligada de teste foram detectados 7 ramos de transmissão como sendo críticos para a perturbação em análise. Esses ramos consistem nas linhas da rede de transporte da área de controlo 1: 15-16, 15-17, 16-4, 16-6, 14-18, 16-18 e 20-21.

A partir de um conjunto de condições de operação que caracterizam o regime estacionário de pré-perturbação de cada cenário gerado para a rede interligada de teste é possível recriar quer o cálculo de trânsito de potências ou a simulação dinâmica para qualquer exemplo contido no conjunto

de dados gerado (de acordo com [6]). Esse conjunto é formado por 74 características que contêm a seguinte informação:

- $P_{load,ac1}(MW)$ : consumo total da área de controlo 1;
- $N_{c_i}$ : número de máquinas em operação em cada gerador convencional equivalente  $i$ ;
- $P_{c_i}(MW)$ : produção de potência activa em cada gerador convencional equivalente  $i$ ;
- $V_{c_i}(p.u.)$ : valor de tensão especificado aos terminais de cada gerador convencional equivalente  $i$ ;
- $N_{w_j}$ : número de máquinas em operação de cada gerador eólico equivalente  $j$ ;
- $P_{w_j}(MW)$ : potência mecânica desenvolvida por cada gerador eólico equivalente  $j$ .

Estas são as características que permitem a construção do vector de variáveis de entrada candidatas. Refere-se em [6], que para a função de monitorização de segurança o vector de variáveis de entrada candidatas para treino das estruturas de segurança apenas pode ser constituído por medidas que possam ser obtidas pelo sistema SCADA da área de controlo em análise. Atendendo a que é a área de controlo 1 que está em análise, excluíram-se, do vector de variáveis de entrada, as condições de operação referentes à área de controlo 2, saindo as variáveis  $N_{c_i}$ ,  $P_{c_i}$  e  $V_{c_i}$  para os 3 geradores convencionais equivalentes dessa área e as variáveis  $N_{w_j}$  e  $P_{w_j}$  para o gerador eólico equivalente aí considerado. É assim possível reduzir o vector de variáveis de entrada candidatas a  $74 - (3 \times 3 + 2) = 63$  características. O critério de selecção para as 63 características encontra-se bem documentado em [6].

Tal como referido no Capítulo 3, a técnica utilizada para o processo de avaliação de desempenho foi a técnica de *Holdout*. Dos 4596 cenários de operação, 1837 foram seleccionados aleatoriamente para formar o conjunto de teste (o *Holdout*), os restantes 2758 cenários foram utilizados para o conjunto de treino.

Nas secções que se seguem, são apresentados os resultados obtidos do treino de cada estrutura, sendo que a secção 5.2 se refere ao treino dos LRM e a secção 5.3 é referente ao treino das LRT, para duas e quatro folhas. Na secção 5.4 é feita uma comparação de desempenho entre a precisão fornecida por estes tipos de estruturas.

## 5.2 Treino do Modelo de Regressão Linear (LRM)

Para a realização do treino de um Modelo de Regressão Linear foram adoptados os procedimentos descritos na secção 3.4 do Capítulo 3. No seguimento do que já foi referido na introdução do presente capítulo, o vector de variáveis de entrada adopta o conjunto das 63 variáveis candidatas, sendo apenas necessário fazer referência à substituição do consumo total da área de controlo 1 ( $P_{load,ac1}$ ) pelo valor da importação da mesma área de controlo, na qual se desprezam as perdas de Joule ( $Imp$ ). Esta substituição não interfere com a capacidade de previsão de cada LRM treinado. Pelas razões evidenciadas em [6],  $P_{load,ac1}$  engloba grande parte da relevância conjunta das restantes variáveis de entrada, deteriorando a qualidade do valor encontrado para os parâmetros da regressão linear que estão associados às restantes variáveis de entrada. Ainda, considerar  $P_{load,ac1}$  poderia tornar a utilização dos LRM não eficaz para a obtenção de medidas de controlo preventivo.

Na Figura 5.1 e Figura 5.2 são apresentados os valores dos parâmetros dos LRM que se obtiveram pelo processo de treino. Esta representação é feita através da apresentação do valor encontrado para cada um dos parâmetros  $b_j$  de cada LRM, como formulado na secção 3.4 do Capítulo 3. Note-se que estas variáveis de entrada e de saída, que compõem o conjunto de dados, foram previamente sujeitas a uma normalização, de modo a terem um valor médio nulo e variância unitária.

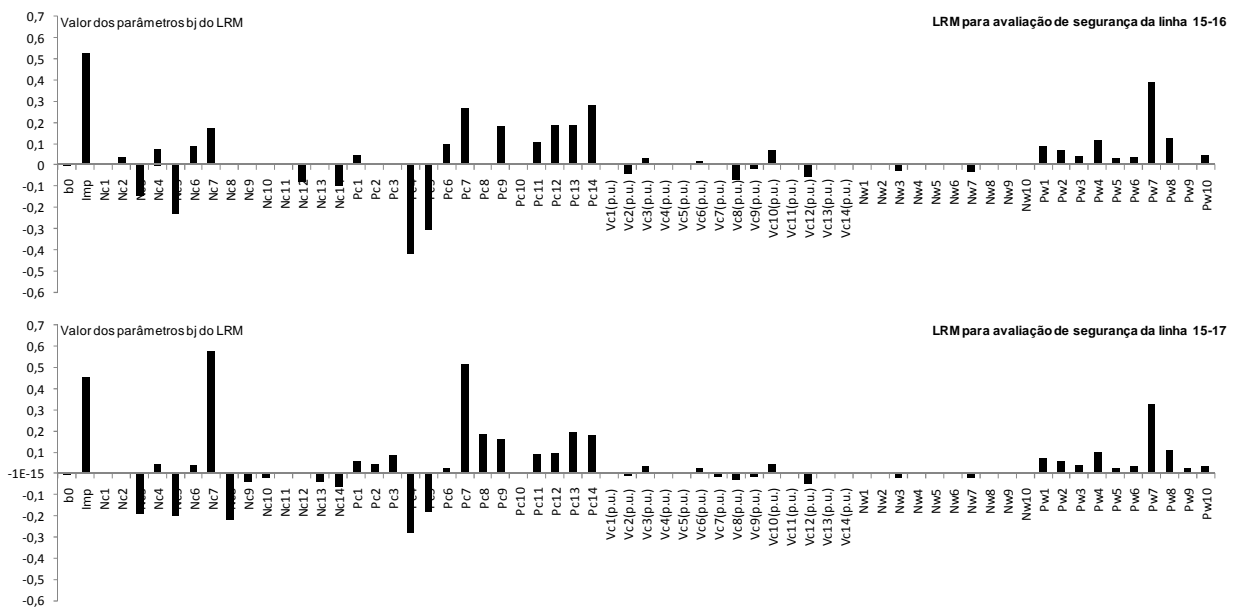


Figura 5.1 - Valor dos parâmetros dos LRM treinados para a rede de teste (parte 1/2)

Capítulo 5: Avaliação de Segurança Dinâmica para a Rede de Teste

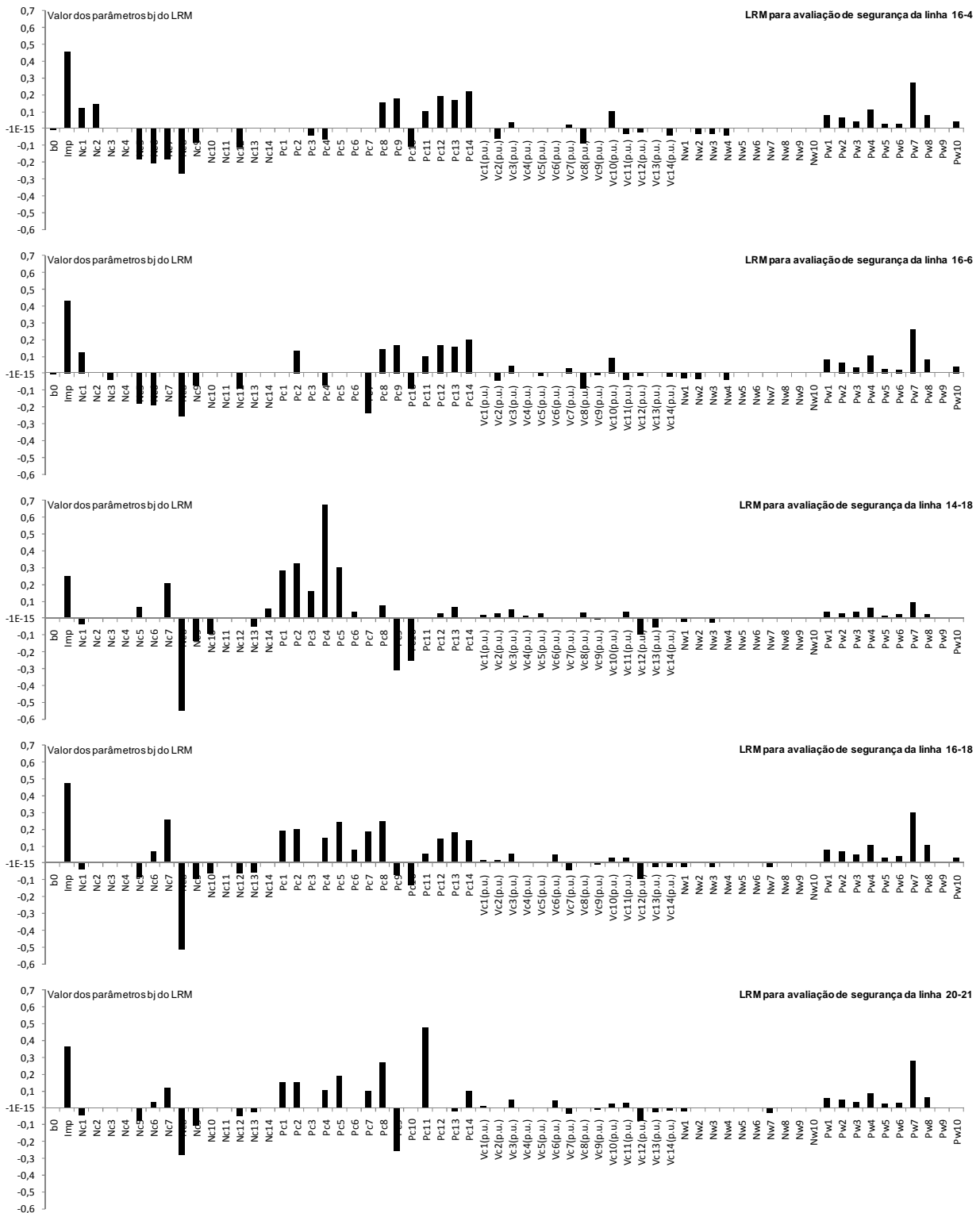


Figura 5.2 - Valor dos parâmetros dos LRM treinados para a rede de teste (parte 2/2)

Por aplicação de LRM ao conjunto de teste, obtiveram-se os erros de regressão e de classificação apresentados na Figura 5.3. Esses erros encontram-se já caracterizados na secção 3.2 do Capítulo 3.

Para o cálculo dos erros de classificação foi ainda utilizada uma margem de segurança correspondente a 20% de sobrecarga em relação ao valor máximo admissível da corrente, em regime permanente, para os dias de Verão.

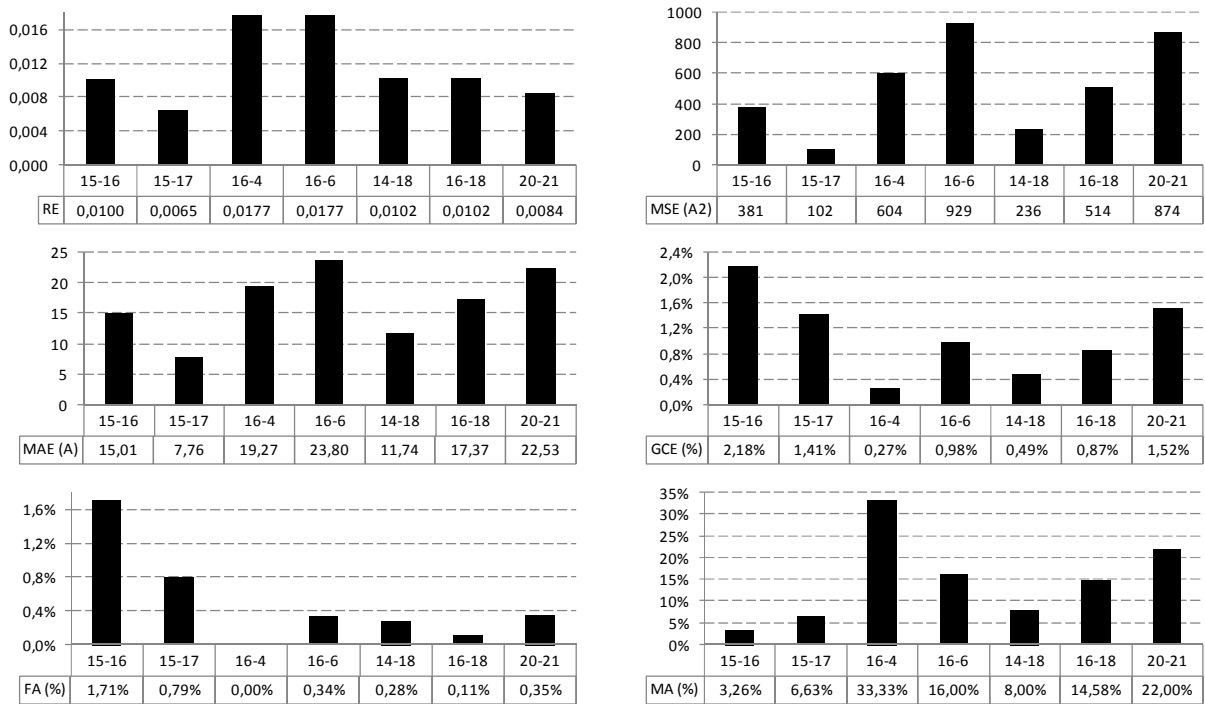


Figura 5.3 - Erros de regressão e de classificação fornecidos pelos LRM treinados, para a rede de teste

### 5.3 Treino das Árvores de Regressão Linear (LRT)

Em [6], levanta-se a hipótese de que a aplicação de modelos de regressão linear nas folhas de uma árvore de regressão (LRT) pode permitir a obtenção de uma estrutura mais precisa do que o treino isolado de um modelo de regressão linear (LRM). De modo a verificar esta hipótese, cada LRT foi treinada num processo de duas etapas (já descrito no Capítulo 3):

- Etapa 1: Treino de uma RT (descrito na secção 3.5, Capítulo 3);
- Etapa 2: Treino de um LRM em cada folha da RT (descrito na secção 3.4, Capítulo 3).

Para evitar a perda de capacidade de generalização da estrutura final treinada, isto é, um número de exemplos demasiado reduzido para cada folha, definiu-se que a dimensão de uma árvore de regressão (RT) não poderia exceder um número de nós superior a 7, correspondendo, portanto, a 4 folhas. Assim, ficou desde logo definido o interesse de avaliar duas situações para o treino de LRT: treino de uma RT para um número de nós igual a 3 (i.e., 2 folhas) e treino de uma RT para um número de nós igual a 7 (i.e., 4 folhas).

Para além das variáveis já referidas na introdução do presente capítulo, incluíram-se um conjunto de outras variáveis para o treino das RT. A inclusão destas variáveis não prejudica o treino das RT, uma vez que o treino deste tipo de estrutura não é prejudicado por as variáveis de entrada serem fortemente relacionadas, podendo mesmo sair o treino beneficiado em termos de precisão por se estar a alargar o conjunto de variáveis candidatas. Então, às 63 variáveis de entrada candidatas referidas na introdução, foram acrescentadas 27 novas condições de operação da área de controlo 1 (valores em MW):

- $Im p$ : potência activa importada (desprezando as perdas de Joule);
- $Pc(H)_{ac1}$  e  $Pc(T)_{ac1}$ : valor total da produção de potência activa dos geradores hídricos (H) e térmicos (T);
- $Pc_{ac1(N)}$  e  $Pc_{ac1(S)}$ : valor total da produção de potência activa dos geradores convencionais que se encontram em serviço na zona Norte (N) e Sul (S);
- $Pw_{ac1}$ : valor total da potência mecânica desenvolvida pelos geradores eólicos;
- $Pw_{ac1(N)}$  e  $Pw_{ac1(S)}$ : valor total da produção de potência activa dos geradores eólicos que se encontram em serviço na zona Norte (N) e Sul (S);
- $SR_i$ : reserva girante disponível em cada gerador convencional equivalente  $i$ ;
- $SR_{ac1}$ : valor total da reserva girante disponível;
- $SR(H)_{ac1}$  e  $SR(T)_{ac1}$ : valor total da reserva girante disponível nos geradores hídricos (H) e térmicos (T);
- $SR_{ac1(N)}$  e  $SR_{ac1(S)}$ : valor total da reserva girante disponível nos geradores convencionais que se encontram em serviço na zona Norte (N) e Sul (S).

Sendo que a estrutura de uma RT deve traduzir um conjunto de restrições lineares, ao conjunto das 90 variáveis (63 + 27) atrás apresentadas, importa referir que pode haver grandezas que sejam função de alguma das variáveis de decisão do problema de optimização e, nesse caso, as estruturas obtidas podem não traduzir restrições do tipo linear aquando da sua exploração para a obtenção de medidas de controlo preventivo. No entanto, constitui uma alternativa de treino sem restrições de linearidade e serve como comparação. Então, de modo a avaliar a capacidade de precisão de uma LRT que corresponda a uma estrutura do tipo linear, foram definidas as seguintes alternativas para o tipo de medidas de controlo preventivo a considerar:

- *Alternativa A:* sem restrições de linearidade. Corresponde ao vector de 90 variáveis de entrada definido no início desta secção.
- *Alternativa B:* são permitidas alterações ao nível do despacho das máquinas convencionais ou eólicas da área de controlo em análise.
- *Alternativa C:* são permitidas alterações ao nível do despacho e do escalonamento de grupos.
- *Alternativa D:* são permitidas alterações ao nível do despacho e do escalonamento de grupos, bem como do valor da importação.

As variáveis que resultam da definição destas restrições, são incluídas na aplicação de podagem de árvores de regressão. Para aplicação do modelo de regressão linear pelo método passo a passo do SPSS, são utilizadas as mesmas 63 variáveis que entram para o LRM. Como foi anteriormente referido, o modelo não é prejudicado pela inclusão destas restrições de linearidade.

Para que as alternativas B, C e D correspondam aos requisitos de linearidade, partindo do vector das 90 variáveis de entrada inicial, retiraram-se todas as grandezas que são função das variáveis de controlo. A *Alternativa B* sugere que  $Pc_i$  e  $Pw_j$  correspondem a variáveis de controlo, resultando um vector com 40 variáveis de entrada:  $P_{load,ac1}$ ,  $Nc_i$  e  $Vc_i$  (para  $i = 1, \dots, 14$ ),  $Nw_j$  (para  $j = 1, \dots, 10$ ) e  $Imp$ . A *Alternativa C* sugere que  $Pc_i$ ,  $Pw_j$ ,  $Nc_i$  e  $Nw_i$  correspondem a variáveis de controlo, pelo que o vector de variáveis de entrada se reduz a 16 variáveis:  $P_{load,ac1}$ ,  $Vc_i$  (para  $i = 1, \dots, 14$ ) e  $Imp$ . Por fim, a *Alternativa D* sugere que  $Pc_i$ ,  $Pw_j$ ,  $Nc_i$ ,  $Nw_i$  e  $Imp$  correspondem a variáveis de controlo, pelo que, neste caso, o vector de variáveis de entrada corresponde às seguintes 15 variáveis:  $P_{load,ac1}$  e  $Vc_i$  (para  $i = 1, \dots, 14$ ).

Após a definição das alternativas, seguiu-se o processo de treino das LRT com 3 nós (secção 5.3.1) e com 7 nós (secção 5.3.2), para emular cada um dos índices de segurança em análise, para cada linha crítica.

Na secção 5.4 serão apresentados os gráficos que permitem efectuar uma comparação directa entre os tipos de estruturas treinadas. Assim, será possível verificar quais as melhorias de precisão obtidas na utilização das LRT face ao LRM.

### 5.3.1 Treino de Árvores de Regressão com duas folhas

#### 5.3.1.1 Alternativa A

Da execução do processo de treino com 3 nós para emular cada um dos índices de segurança em análise, resultaram as estruturas em árvore que se apresentam na Figura 5.5, com a seguinte legenda:  $N$  é o número de exemplos de treino que pertencem ao nó,  $\bar{y}$  é a média amostral dos valores do índice de segurança em análise contidos nos exemplos de treino que pertencem ao nó. Verifica-se que o conjunto de treino a considerar para a avaliação de cada índice de segurança se encontra dividido em 2 subconjuntos disjuntos (um por cada folha de RT). A partir de cada um destes subconjuntos procedeu-se ao treino de um LRM, onde se consideraram as 63 características como variáveis de entrada já utilizadas no processo de treino isolado do LRM. Deste processo, resultaram 7 árvores de regressão linear, uma por cada linha crítica em análise, cujos erros de regressão e de classificação obtidos se apresentam na Figura 5.4.

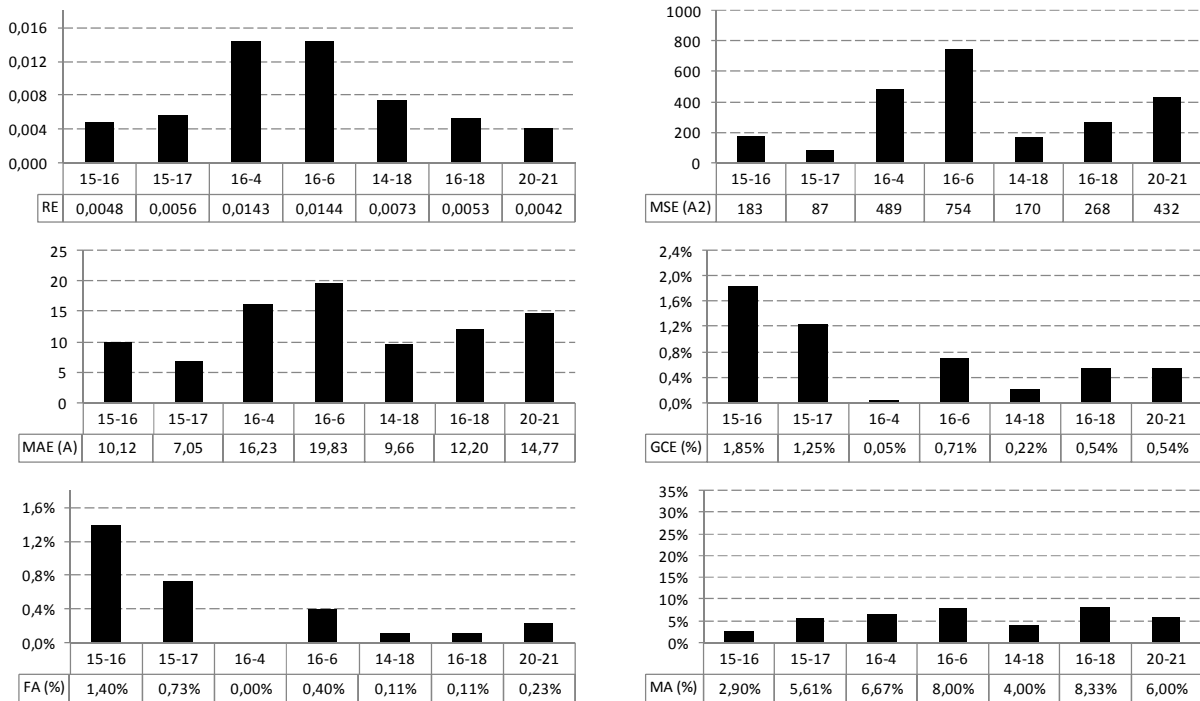


Figura 5.4 - Erros de regressão e de classificação fornecidos pelos LRT treinados, para a rede de teste, considerando a alternativa A de 2 folhas

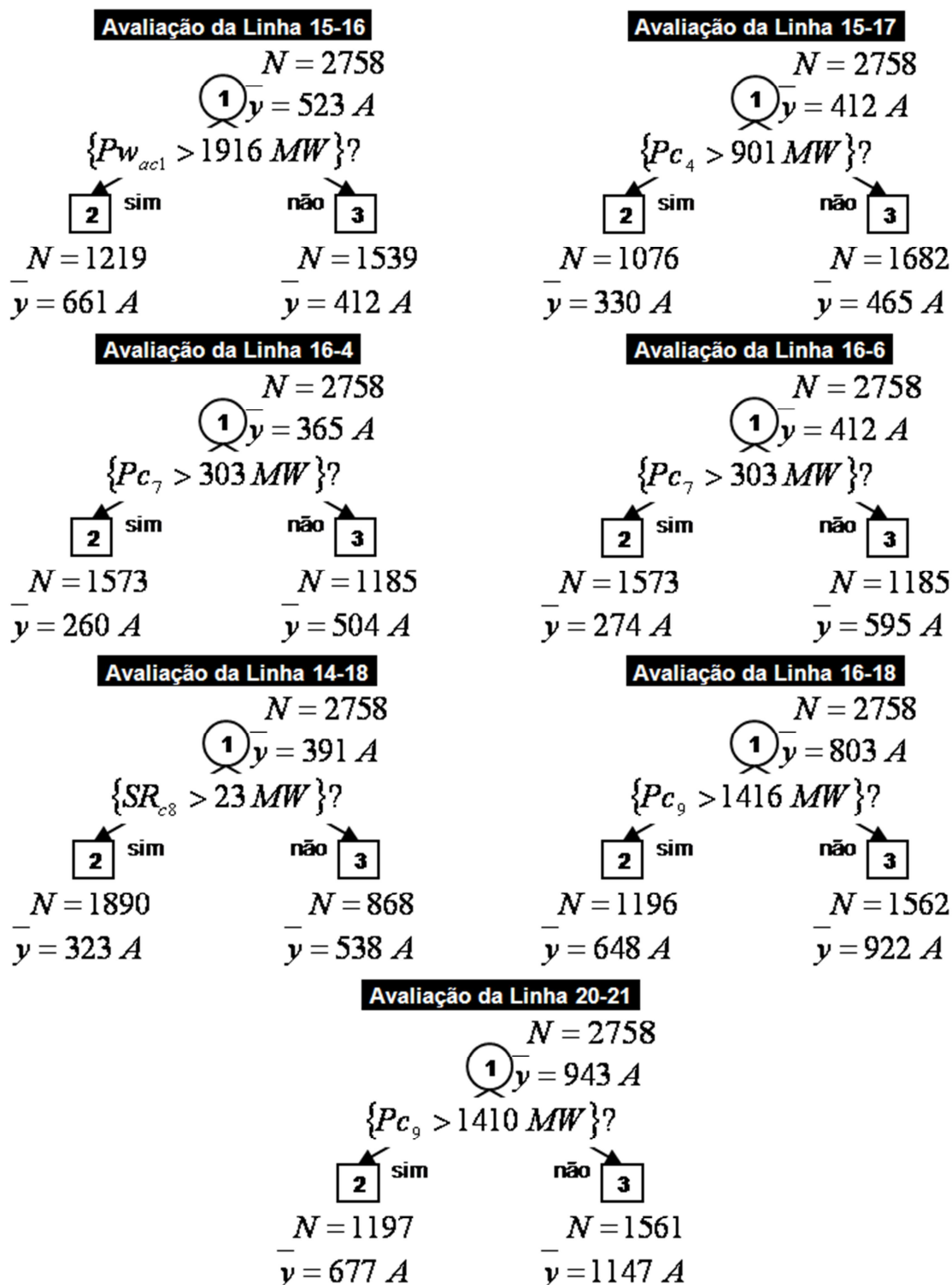


Figura 5.5 - Estrutura em árvore das LRT que foram treinadas, para a rede de teste, considerando a alternativa A de 2 folhas

Comparando os resultados apresentados na Figura 5.4 com os que se apresentaram na Figura 5.3, relativamente aos erros fornecidos pelo treino dos LRM e ao problema de segurança em análise,

consegue-se desde já verificar que esta estrutura de LRT permite obter uma avaliação de segurança mais precisa que os LRM.

### 5.3.1.2 Alternativa B

Para a Alternativa B, definida no início da secção 5.3, após a execução do processo de treino com 3 nós, resultaram as estruturas em árvore que se encontram representadas na Figura 5.7. Para cada um dos subconjuntos disjuntos obtidos, procedeu-se ao treino de um LRM, onde se consideraram as 63 características como variáveis de entrada. Do processo resultaram 7 árvores de regressão linear, uma por cada linha e os erros de regressão e de classificação são os que se apresentam na Figura 5.6.

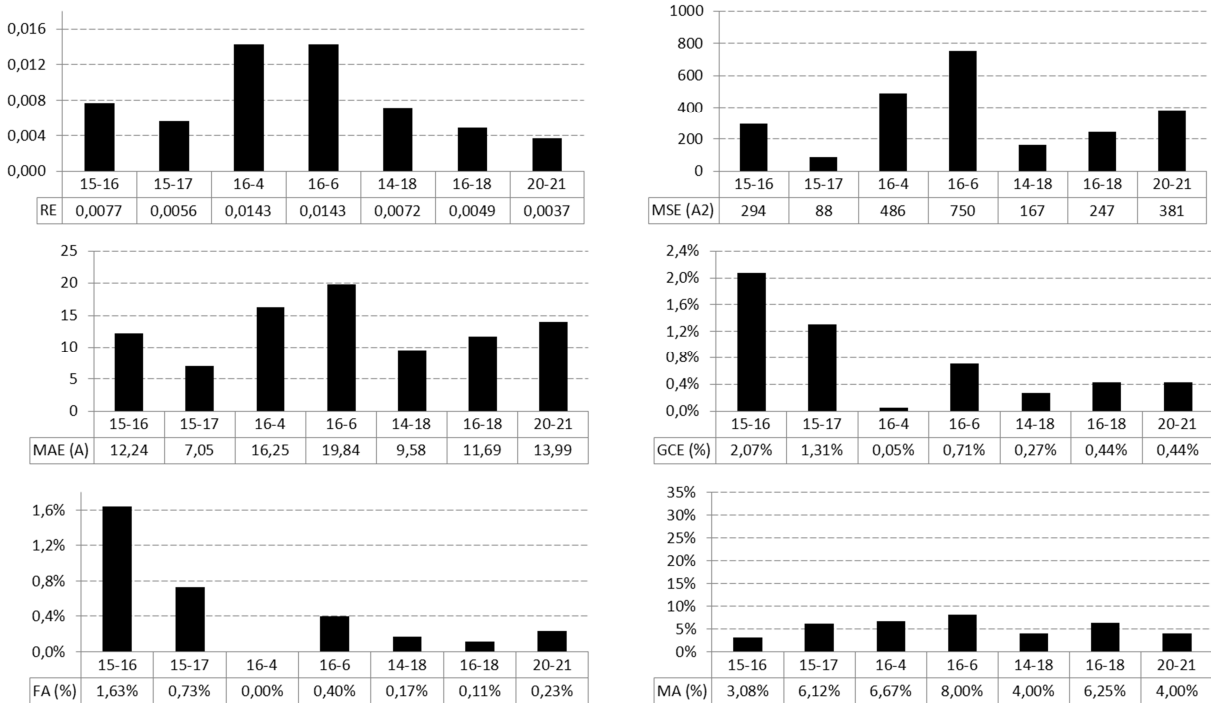


Figura 5.6 - Erros de regressão e de classificação fornecidos pelos LRT treinados, para a rede de teste, considerando a alternativa B de 2 folhas

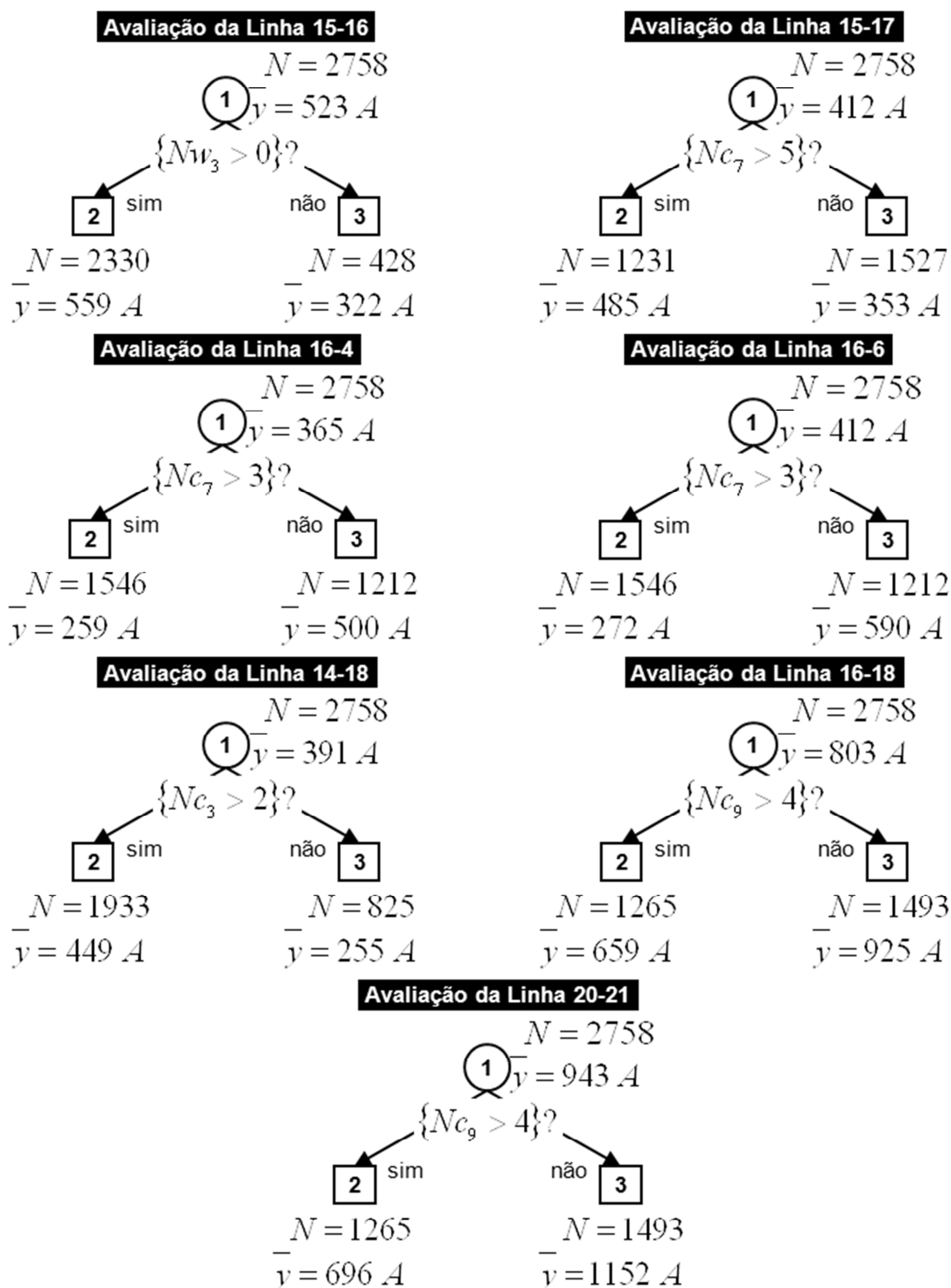


Figura 5.7 - Estrutura em árvore das LRT que foram treinadas, para a rede de teste, considerando a alternativa B de 2 folhas

Quando se comparam os resultados apresentados na Figura 5.6 com os que se apresentaram na Figura 5.3 para o LRM, verifica-se que este tipo de estrutura (LRT com a alternativa B de restrições

de linearidade) continua a permitir uma avaliação de segurança mais precisa do que os LRM, no entanto, quando se comparam estes mesmo resultados com os da Figura 5.4 (referentes às LRT sem restrições de linearidade), pode verificar-se que há uma diminuição da precisão da avaliação de segurança.

### 5.3.1.3 Alternativa C

Após a execução do processo de treino com 3 nós para a Alternativa C, definida no início da secção 5.3, resultaram as estruturas em árvore que se encontram representadas na Figura 5.9. Para cada um dos subconjuntos disjuntos obtidos, procedeu-se ao treino de um LRM, onde se consideraram as 63 características como variáveis de entrada. Do processo resultaram 7 árvores de regressão linear, uma por cada linha, e os erros de regressão e de classificação são os apresentados na Figura 5.8.

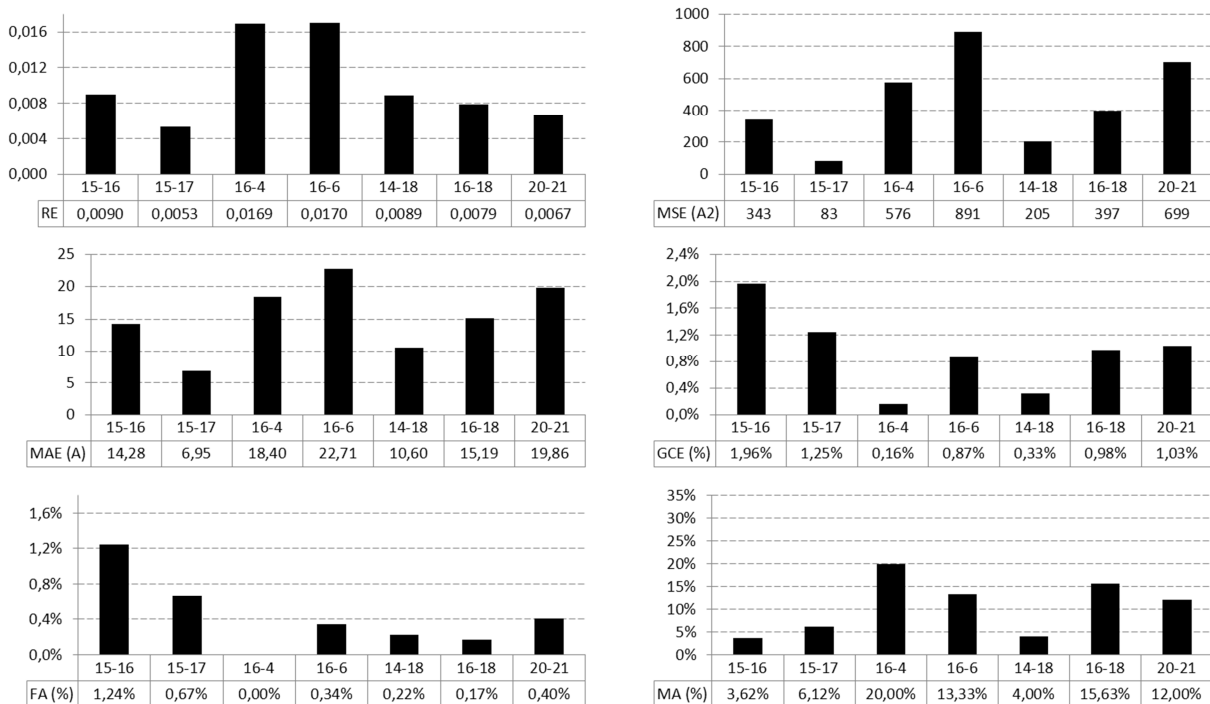


Figura 5.8 - Erros de regressão e de classificação fornecidos pelos LRT treinados, para a rede de teste, considerando a alternativa C de 2 folhas

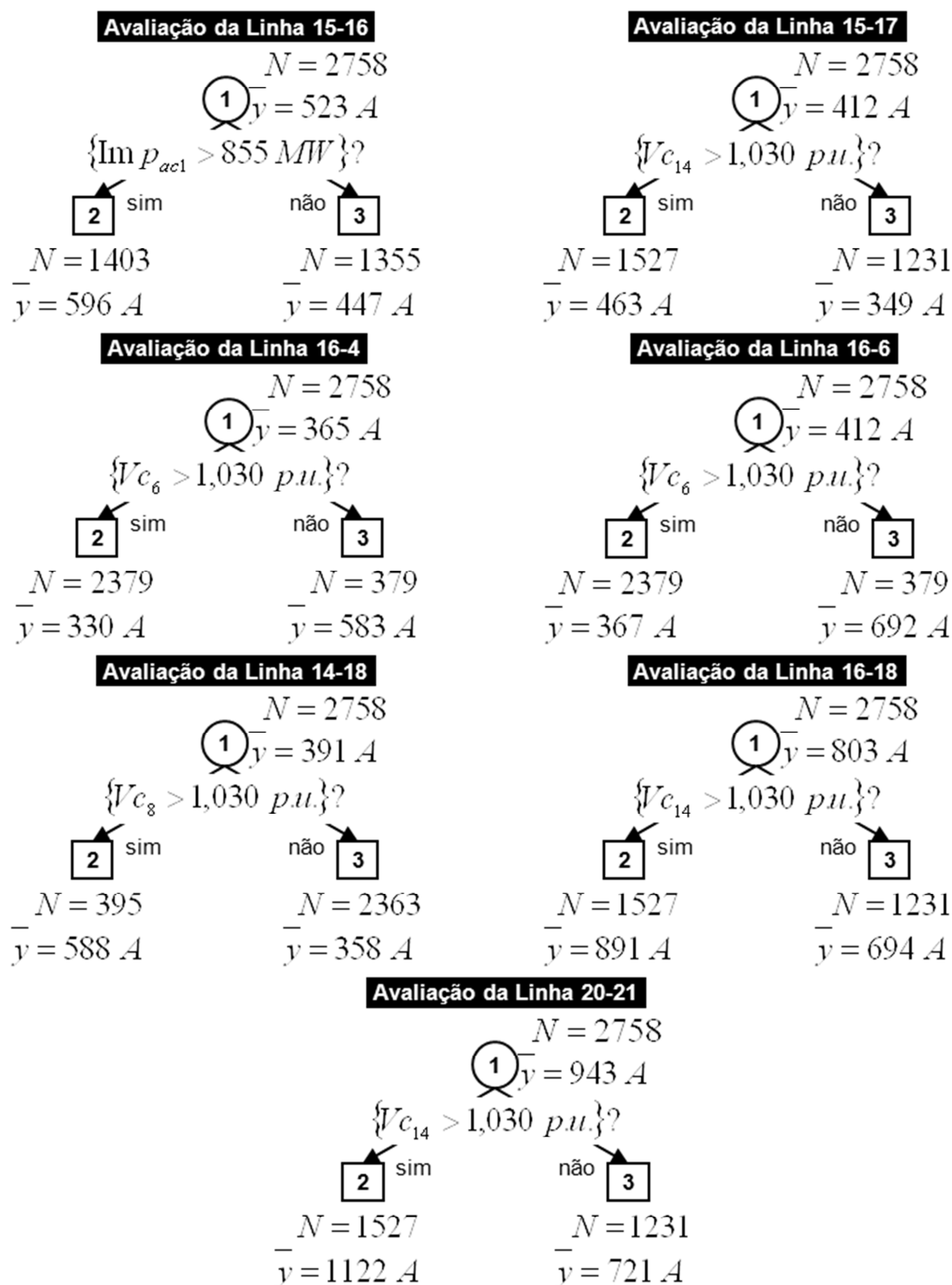


Figura 5.9 - Estrutura em árvore das LRT que foram treinadas, para a rede de teste, considerando a alternativa C de 2 folhas

A partir dos resultados da Figura 5.8 e comparando com os resultados que se apresentaram na Figura 5.3, esta estrutura de LRT também permite obter uma avaliação de segurança mais precisa

que o LRM, na mesma ordem que a estrutura obtida com a alternativa B, mas sem a mesma expressividade que a obtida na alternativa A.

5.3.1.4 Alternativa D

Para o último processo de treino com 3 nós, após a execução da Alternativa D, definida no início da secção 5.3, resultaram as estruturas em árvore que se encontram representadas na Figura 5.11. Para cada um dos subconjuntos disjuntos obtidos, procedeu-se ao treino de um LRM, onde se consideraram as 63 características como variáveis de entrada. Do processo resultaram 7 árvores de regressão linear, uma por cada linha, e os erros de regressão e de classificação são os apresentados na Figura 5.10.

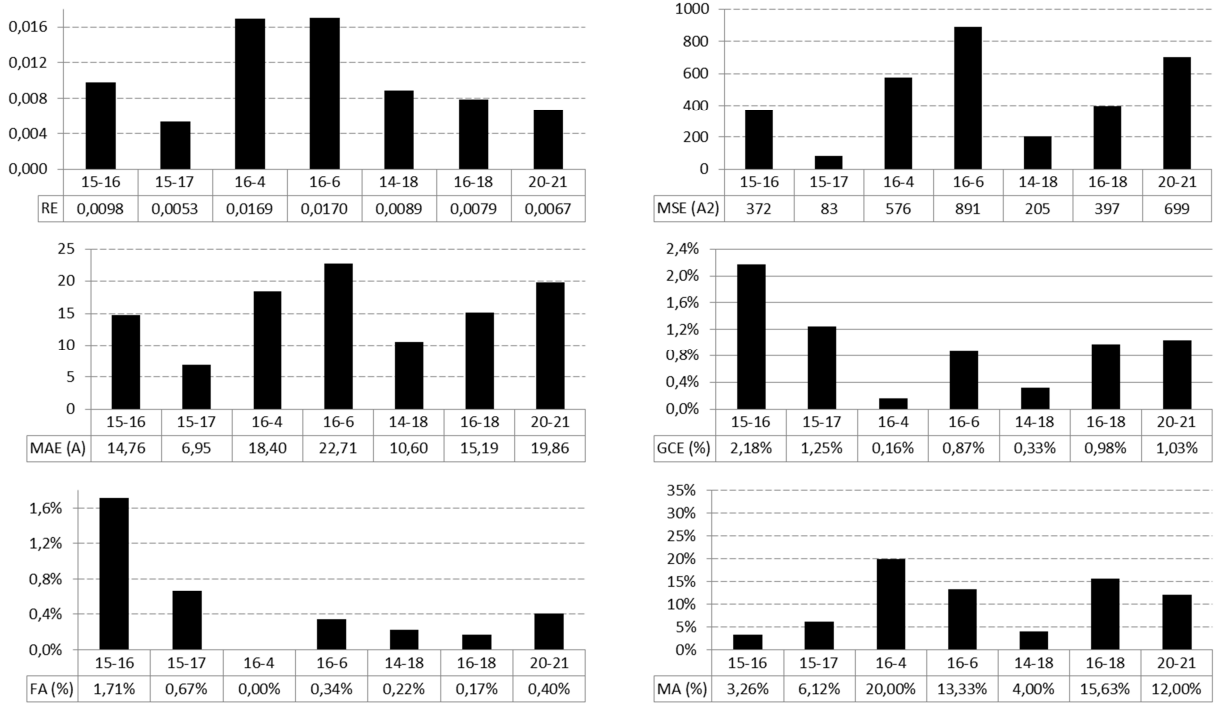


Figura 5.10 - Erros de regressão e de classificação fornecidos pelos LRT, treinados para a rede de teste, considerando a alternativa D de 2 folhas

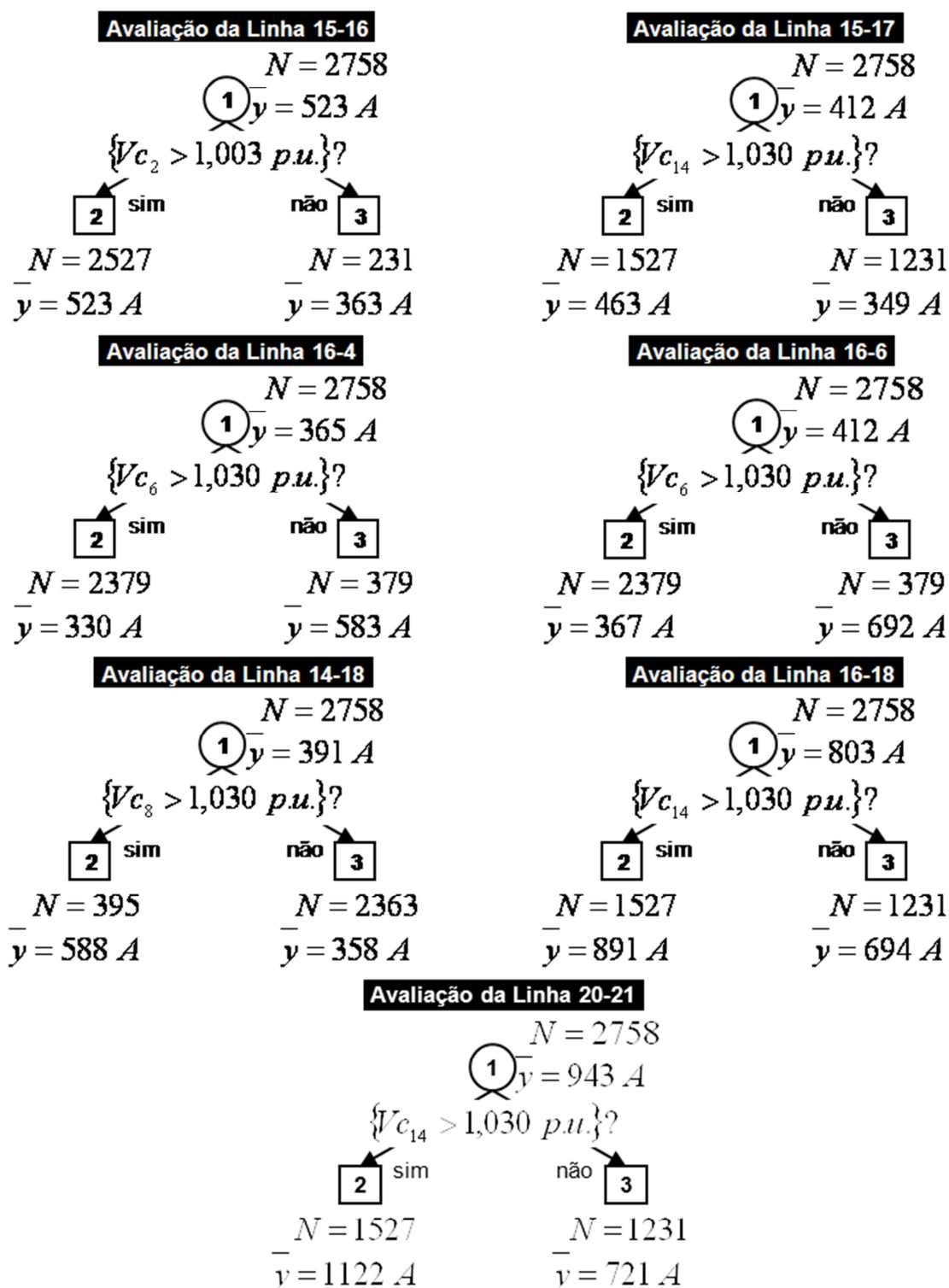


Figura 5.11 - Estrutura em árvore das LRT que foram treinadas, para a rede de teste, considerando a alternativa D de 2 folhas

Por comparação dos resultados obtidos na Figura 5.10 com os apresentados na Figura 5.3, os erros de regressão e classificação fornecidos para o treino da LRT são mais precisos que os obtidos pela

estrutura de LRM, ainda que a diferença seja menor e menos significativa que os já apresentados nas três alternativas anteriores. Observando a Figura 5.9 e a Figura 5.11, é possível verificar que as árvores obtidas são iguais às da alternativa C, com exceção das árvores para a linha 15-16 e linha 20-21. Verificou-se, tal como esperado, que os erros de previsão são os mesmos (a menos da linha 15-16 e da linha 20-21).

### 5.3.2 Treino de Árvores de Regressão com quatro folhas

#### 5.3.2.1 Alternativa A

Da execução do processo de treino com 7 nós para emular cada um dos índices de segurança em análise, resultaram as estruturas em árvore que se apresentam na Figura 5.13, com a seguinte legenda:  $N$  é o número de exemplos de treino que pertencem ao nó,  $\bar{y}$  é a média amostral dos valores do índice de segurança em análise contidos nos exemplos de treino que pertencem ao nó. Verifica-se que o conjunto de treino a considerar para a avaliação de cada índice de segurança se encontra dividido em 4 subconjuntos disjuntos (um por cada folha de RT). A partir de cada um destes subconjuntos procedeu-se ao treino de um LRM, onde se consideraram as 63 características como variáveis de entrada já utilizadas no processo de treino isolado do LRM. Deste processo, resultaram 7 árvores de regressão linear, uma por cada linha crítica em análise, cujos erros de regressão e de classificação obtidos se apresentam na Figura 5.12.

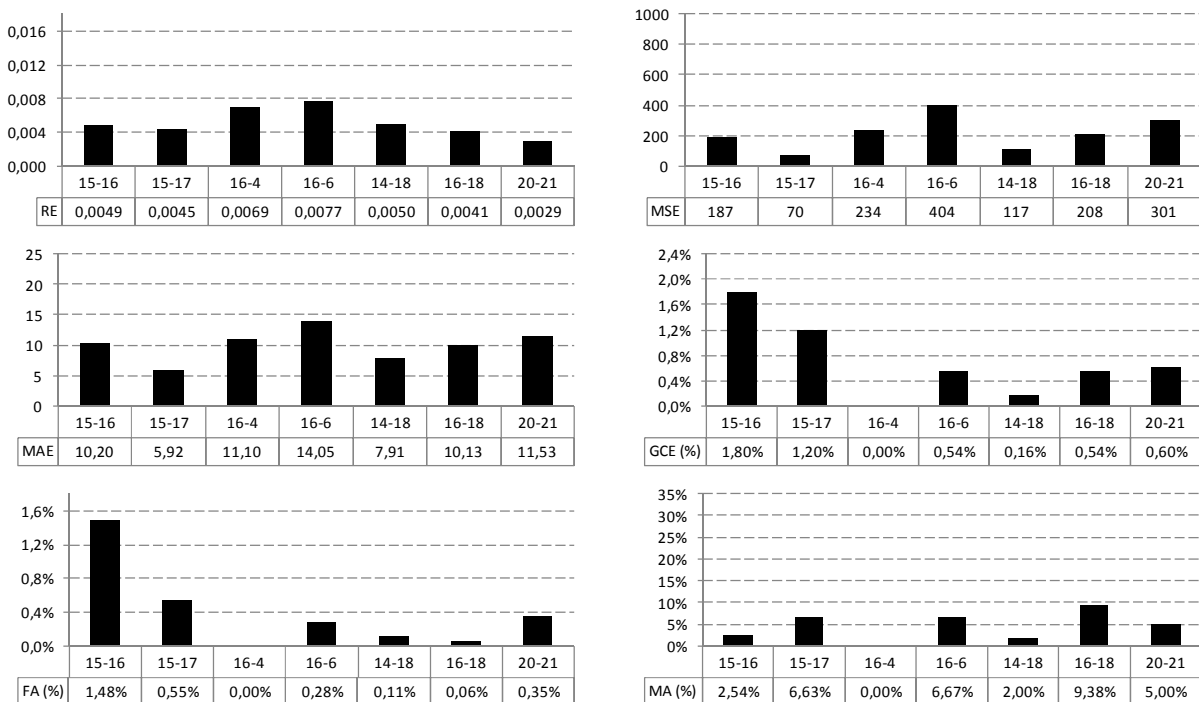


Figura 5.12 - Erros de regressão e de classificação fornecidos pelos LRT treinados, para a rede de teste, considerando a alternativa A de 4 folhas

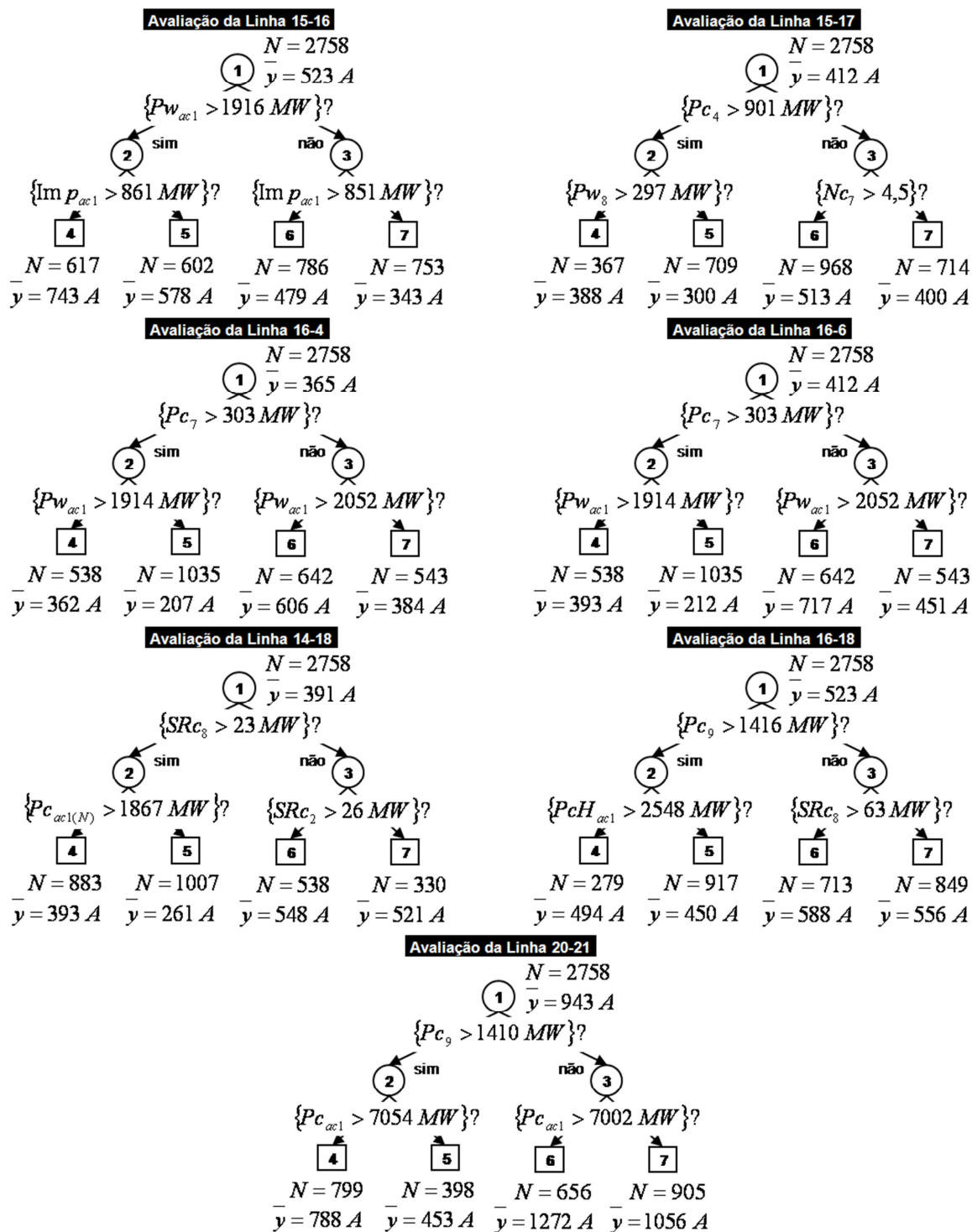


Figura 5.13 - Estrutura em árvore das LRT que foram treinadas, para a rede de teste, considerando a alternativa A de 4 folhas

Comparando os resultados obtidos na Figura 5.12 com os já apresentados para o LRM na Figura 5.3, relativamente aos erros fornecidos pelo treino dos LRM e ao problema de segurança em análise, pode verificar-se que com esta estrutura de LRT se obtém uma avaliação de segurança mais precisa.

### 5.3.2.2 Alternativa B

Para a Alternativa B, definida no início da secção 5.3, após a execução do processo de treino com 7 nós, resultaram as estruturas em árvore que se encontram representadas na Figura 5.15. Para cada um dos quatro subconjuntos disjuntos obtidos, procedeu-se ao treino de um LRM, onde se consideraram as 63 características como variáveis de entrada. Do processo resultaram 7 árvores de regressão linear, uma por cada linha crítica em análise, e os erros de regressão e de classificação são os apresentados na Figura 5.14.

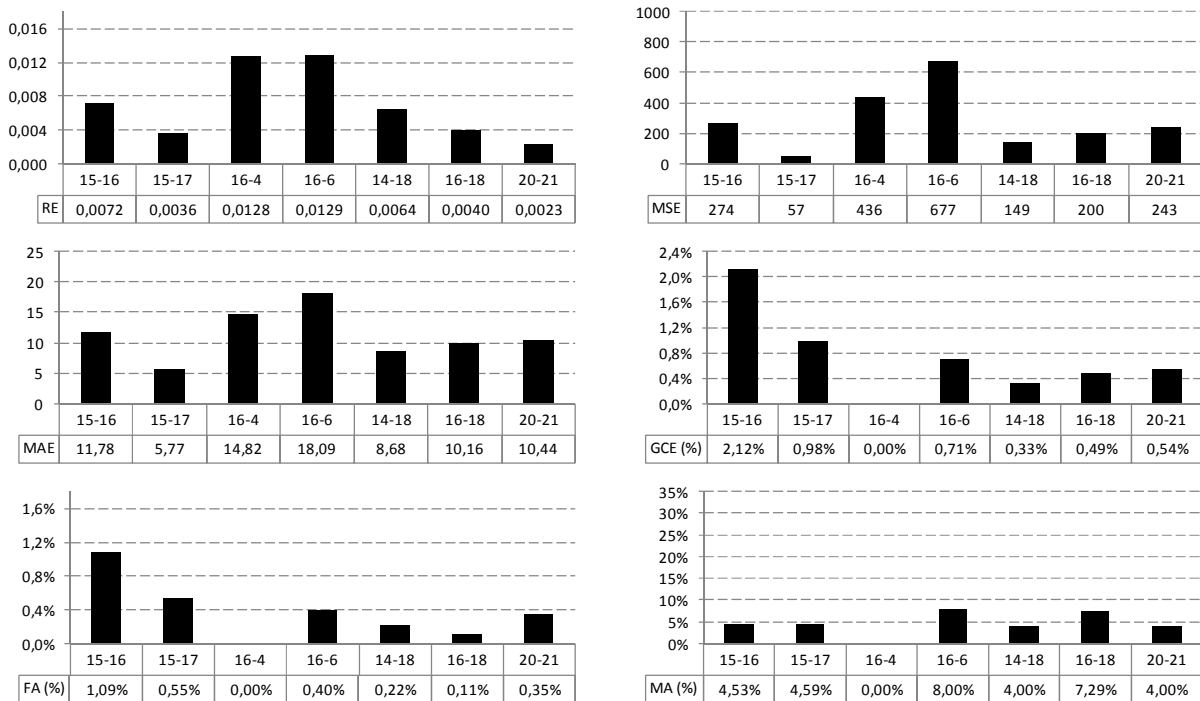


Figura 5.14 - Erros de regressão e de classificação fornecidos pelos LRT treinados, para a rede de teste, considerando a alternativa B de 4 folhas

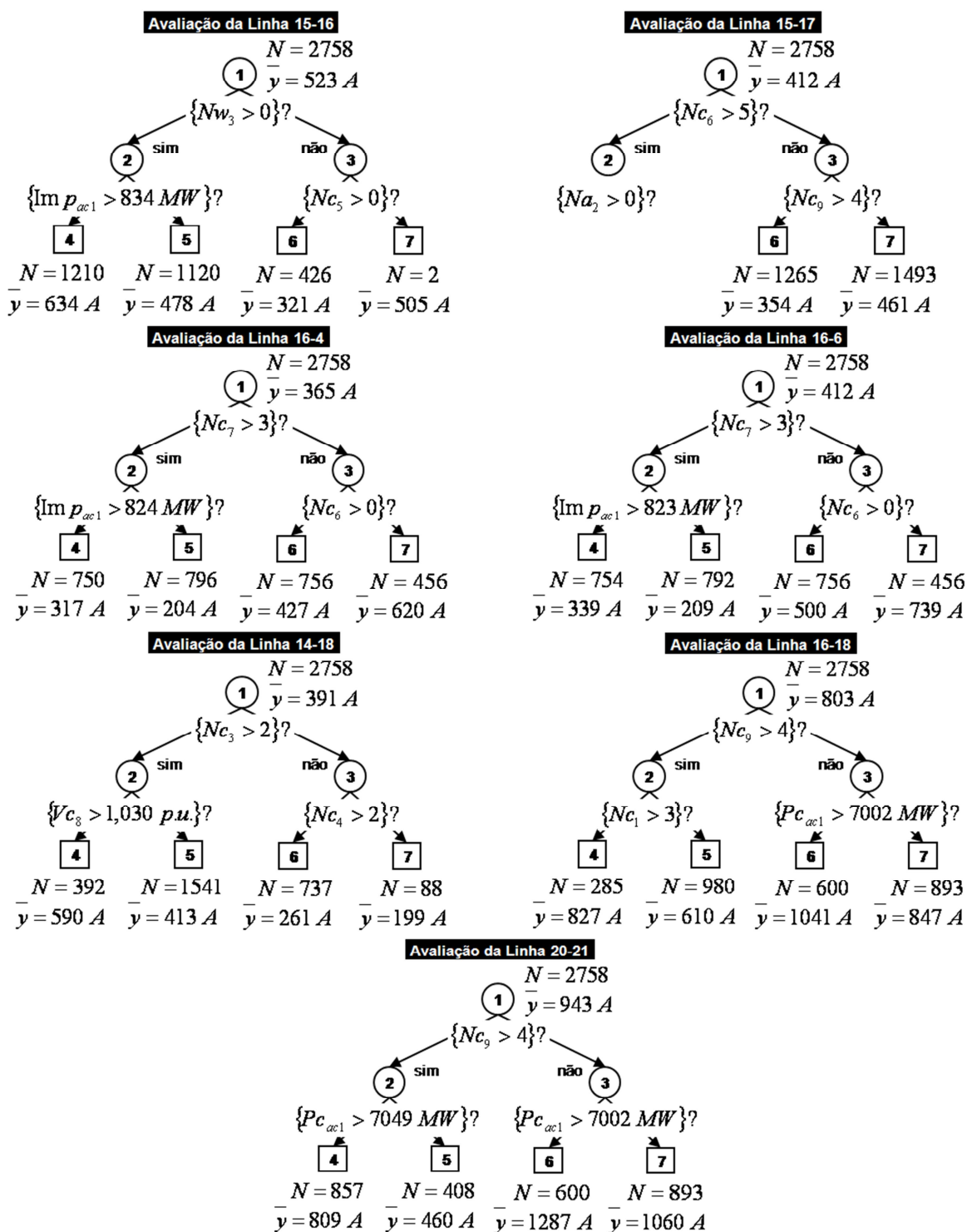


Figura 5.15 - Estrutura em árvore das LRT que foram treinadas, para a rede de teste, considerando a alternativa B de 4 folhas

A comparação dos resultados da Figura 5.14 com os resultados apresentados na Figura 5.3, revela que com esta estrutura de LRT se continua a obter uma maior precisão para a avaliação de segurança

do que com os LRM. Quando estes resultados são comparados com a Figura 5.12, é possível verificar que esta alternativa B já não apresenta uma avaliação tão precisa como a alternativa A.

### 5.3.2.3 Alternativa C

Após a execução do processo de treino com 7 nós para a alternativa C, resultaram as estruturas em árvore que se encontram representadas na Figura 5.17. Para cada um dos quatro subconjuntos disjuntos obtidos, procedeu-se ao treino de um LRM, onde se consideraram as 63 características como variáveis de entrada. Do processo resultaram 7 árvores de regressão linear, uma por cada linha crítica em análise, e os erros de regressão e de classificação são os apresentados na Figura 5.16.

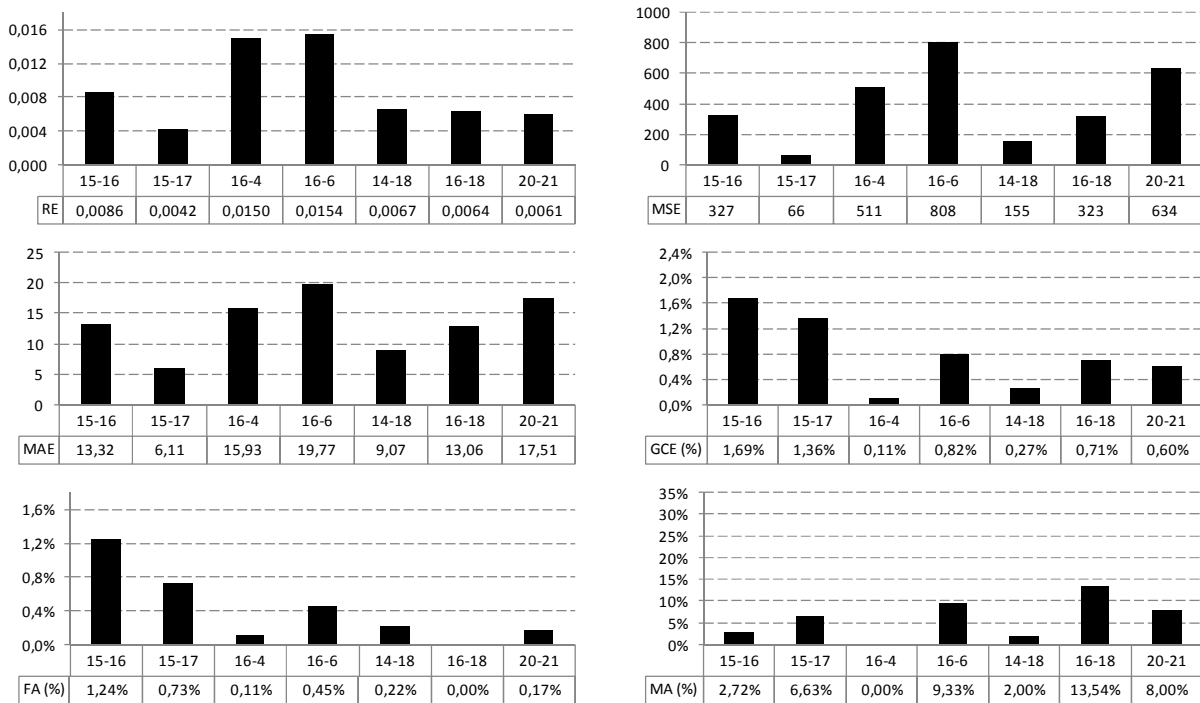


Figura 5.16 - Erros de regressão e de classificação fornecidos pelos LRT treinados, para a rede de teste, considerando a alternativa C de 4 folhas

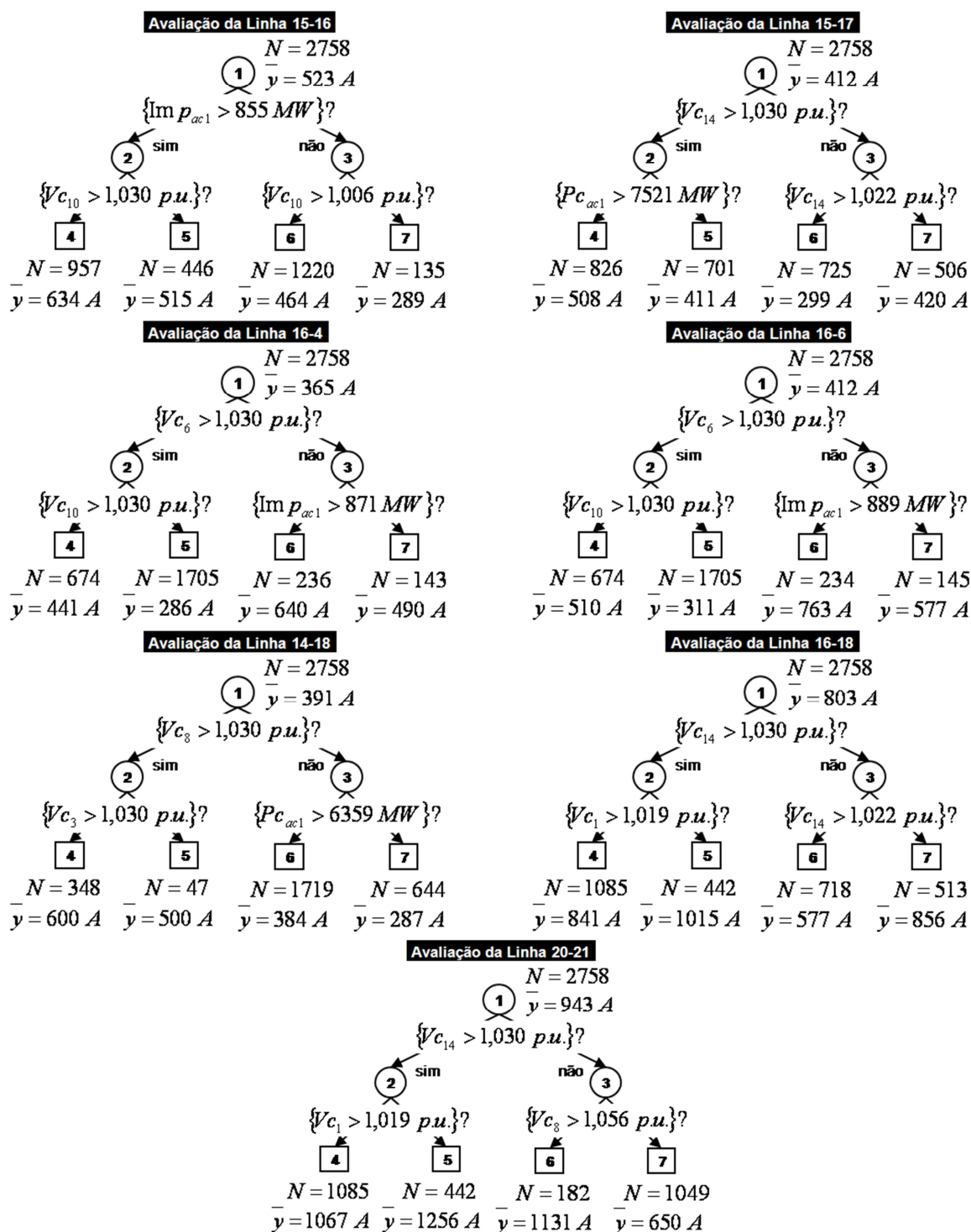


Figura 5.17 - Estrutura em árvore das LRT que foram treinadas, para a rede de teste, considerando a alternativa C de 4 folhas

A comparação dos resultados que se encontram na Figura 5.16 com os resultados da Figura 5.3, permite verificar que a estrutura obtida com esta alternativa é mais precisa que o LRM mas, ao comparar com ambas as alternativas A e B anteriores, Figura 5.12 e Figura 5.14, a precisão diminui.

5.3.2.4 Alternativa D

Para a Alternativa D, após a execução do processo de treino com 7 nós, resultaram as estruturas em árvore que se encontram representadas na Figura 5.19. Obtiveram-se 4 subconjuntos disjuntos e para cada um desses subconjuntos, procedeu-se ao treino de um LRM, onde se consideraram as 63 características como variáveis de entrada. Do processo resultaram 7 árvores de regressão linear, uma por cada linha crítica em análise, e os erros de regressão e de classificação são os apresentados na Figura 5.18.

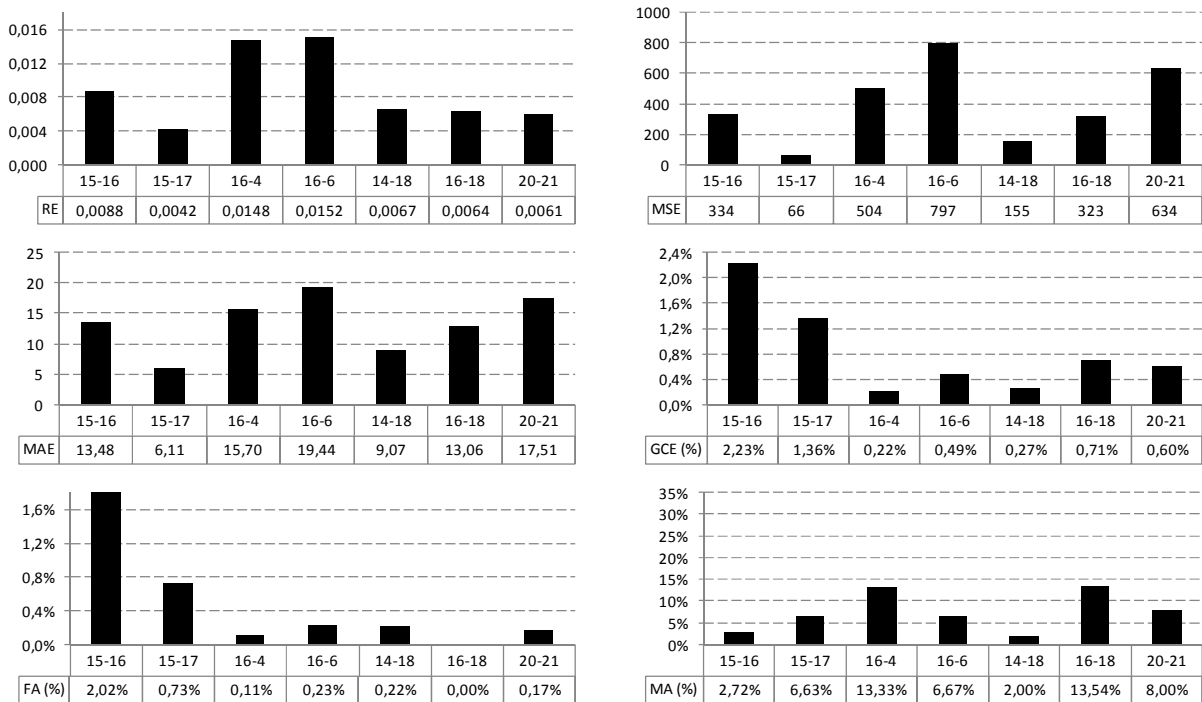


Figura 5.18 - Erros de regressão e de classificação fornecidos pelos LRT treinados, para a rede de teste, considerando a alternativa D de 4 folhas

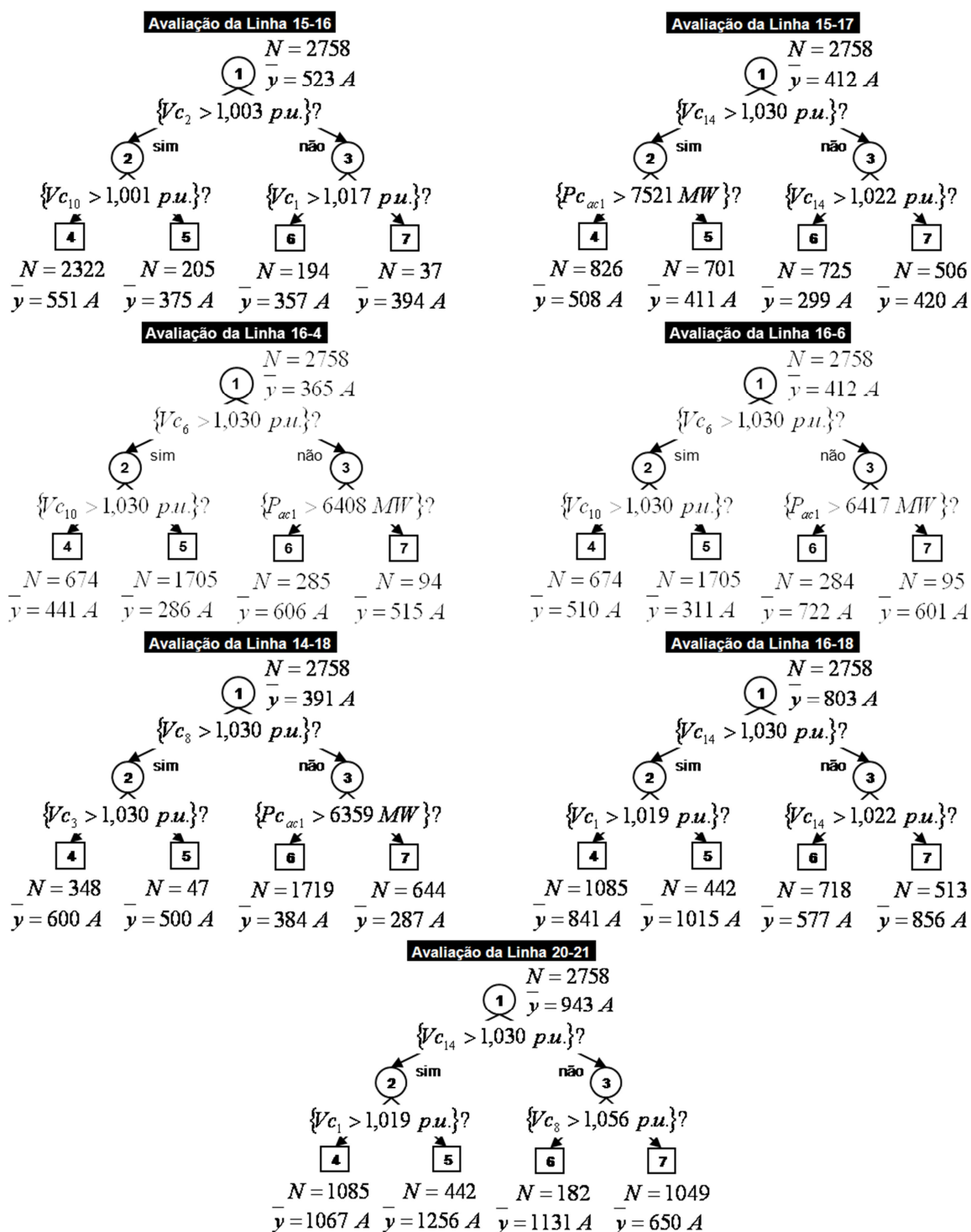


Figura 5.19 - Estrutura em árvore das LRT que foram treinadas, para a rede de teste, considerando a alternativa D de 4 folhas

A comparação da Figura 5.18 com os resultados que se apresentam na Figura 5.3 permitem concluir que se obtém uma estrutura de avaliação de segurança mais precisa de segurança que o LRM e semelhante à obtida na alternativa C, como se pode observar pelos resultados na Figura 5.16.

Comparando a Figura 5.17 com a Figura 5.19, pode-se verificar que as árvores obtidas para a linha 15-17, 14-18, 16-18 e 20-21, são iguais para as alternativas C e D. Logo, foram obtidos os mesmos valores de erros de previsão e de classificação.

## 5.4 Comparação das Estruturas de Treino

Após a execução de todos os processos de treino de cada estrutura e para cada restrição definida no Capítulo 3, nesta secção serão apresentados os erros de regressão e de classificação que foram obtidos, para cada um desses tipos de estruturas e restrições. Os dados foram agrupados em gráficos para melhor se avaliar a segurança de operação das 7 linhas de transmissão da rede de teste identificadas como críticas para a perturbação em análise.

Quaisquer dos erros apresentados resultaram da aplicação das estruturas ao conjunto de teste. Em todos os cenários de operação que compõem o conjunto de teste, considerou-se para o cálculo dos erros de classificação uma margem de segurança correspondente a 120% de carga em relação ao valor máximo admissível em regime permanente para os dias de Verão.

### 5.4.1 Comparação entre LRM, LRT de 2 folhas e LRT de 4 folhas (alternativa A):

Os erros que se apresentam na Figura 5.20 e Figura 5.21, para as estruturas do tipo LRT, correspondem aos que foram obtidos do treino de estruturas em árvore sem restrições de linearidade. Esses erros são comparados com os que foram fornecidos pelas LRM que se descrevem na secção 5.2. Como seria de esperar, da análise destas figuras, a LRM é a estrutura que apresenta os piores resultados, sendo que a aplicação do treino de LRM em cada folha da RT numa estrutura com 4 folhas foi a que forneceu maior precisão para emular o problema de segurança dinâmica em análise. Pode ainda observar-se o seguinte resultado curioso: na linha onde é simulado o curto-circuito (i.e., linha 15-16), o valor fornecido pelos dois tipos de estruturas LRT não diferem significativamente.

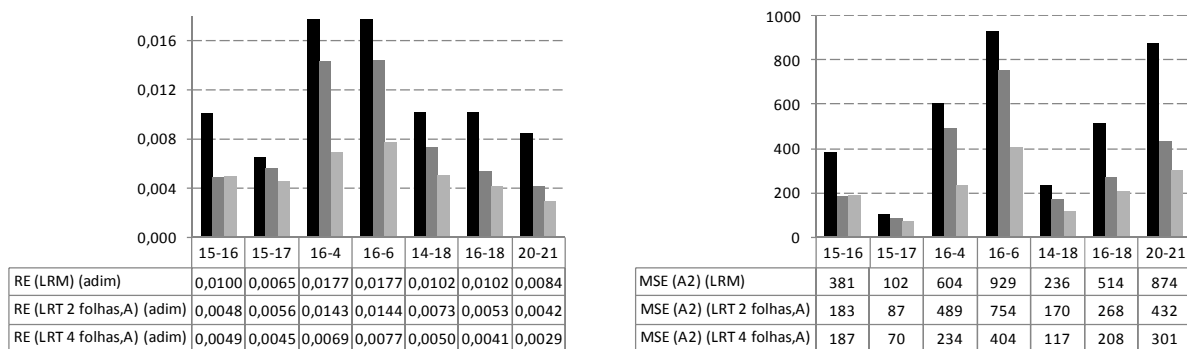


Figura 5.20 - Erros de regressão e de classificação fornecidos por LRM, LRT de 2 folhas e LRT de 4 folhas, para a alternativa A (parte 1/2)

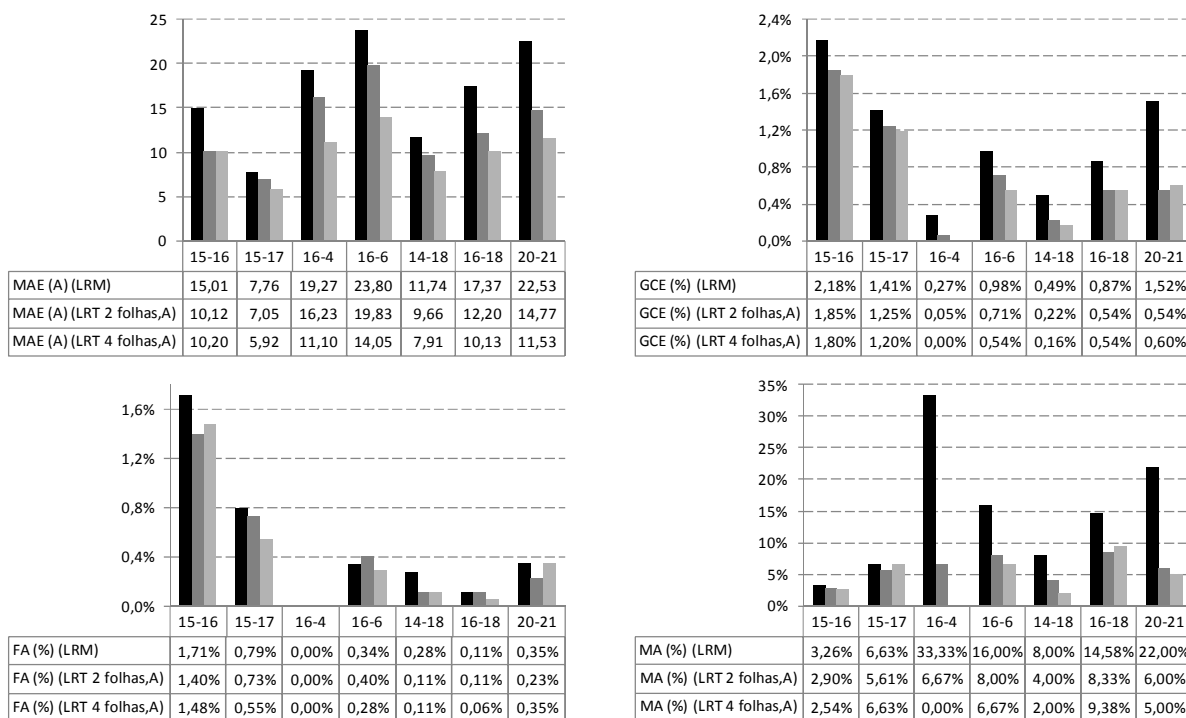


Figura 5.21 - Erros de regressão e de classificação fornecidos por LRM, LRT de 2 folhas e LRT de 4 folhas, para a alternativa A (parte 2/2)

### 5.4.2 Comparação entre LRM, LRT de 2 folhas e LRT de 4 folhas (alternativa B):

Relembre-se que para a alternativa B, as estruturas em árvore das LRT foram treinadas considerando restrições de linearidade que permitam alterações ao nível do despacho das máquinas. Destas figuras pode confirmar-se o que já se havia referido para a alternativa A, nomeadamente, que o treino de LRT com 4 folhas produz melhores resultados de precisão para avaliação de segurança dinâmica em análise.

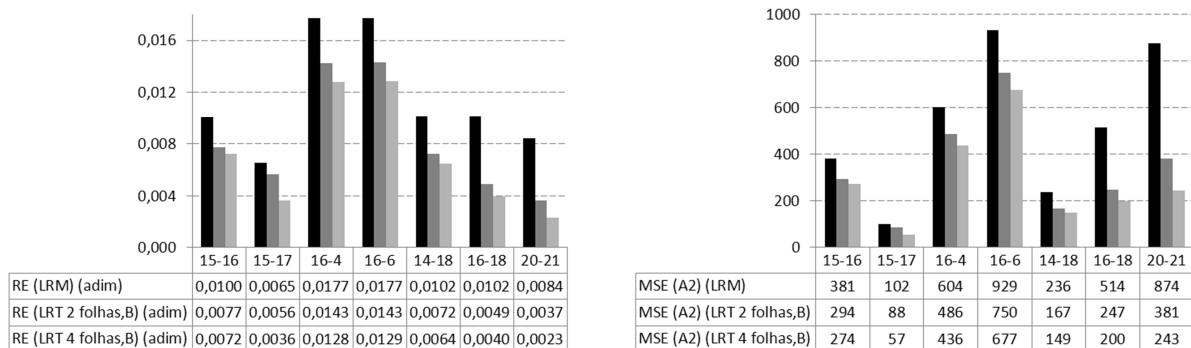


Figura 5.22 - Erros de regressão e de classificação fornecidos por LRM, LRT de 2 folhas e LRT de 4 folhas, para a alternativa B (parte 1/2)

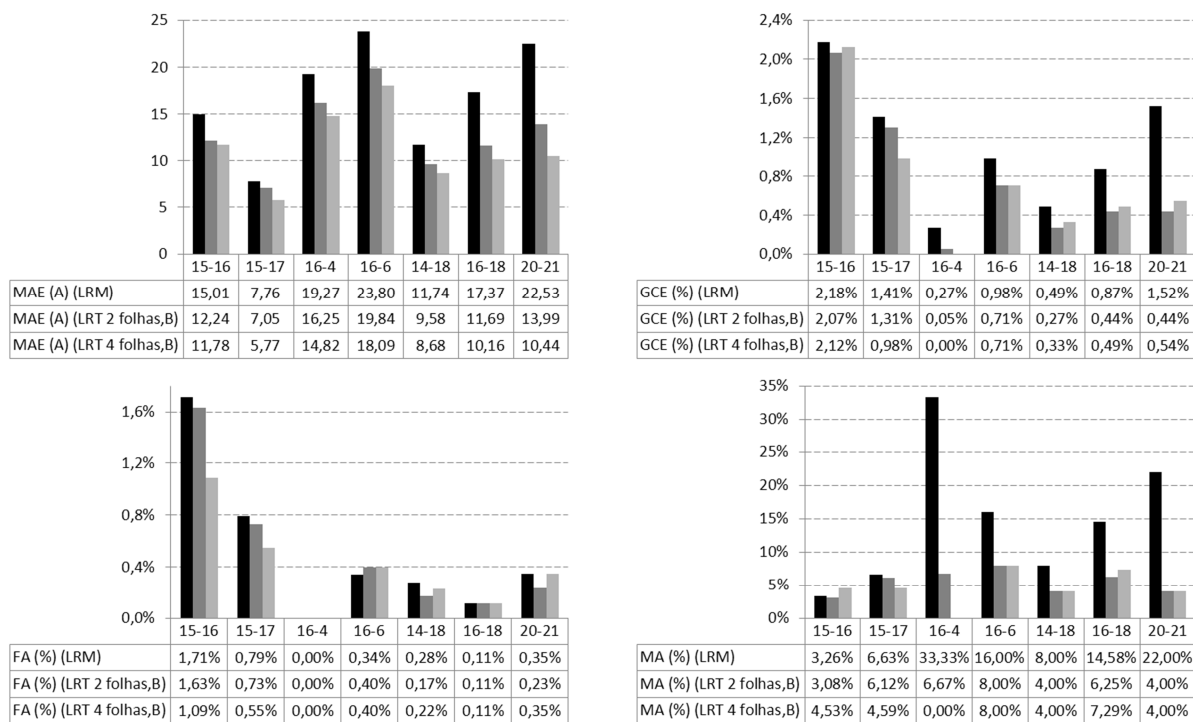


Figura 5.23 - Erros de regressão e de classificação fornecidos por LRM, LRT de 2 folhas e LRT de 4 folhas, para a alternativa B (parte 2/2)

De seguida, para a alternativa B, apresentam-se os gráficos que permitem identificar a dimensão dos erros de regressão fornecidos para cada cenário de operação que formou o conjunto de teste, para cada uma das estruturas treinadas.

Da Figura 5.24 à Figura 5.26, esta informação apresenta-se para a linha 15-16. Verifica-se que os erros de regressão fornecidos se situam dentro de uma margem de 15% de carga. O LRM foi o tipo de estrutura que forneceu uma maior dispersão dos erros, apresentando também o maior número de erros superiores a 10% de carga. Entre a estrutura LRT de 2 folhas e a estrutura LRT de 4 folhas não parece notar-se grande diferença. No entanto, tal como se pode observar pelos erros apresentados na Figura 5.22 e Figura 5.23, a LRT com 4 nós é, no geral, mais precisa do que a LRT com 2 nós. Observa-se, como desejado, que o valor dos erros de regressão não difere significativamente com a ordem de grandeza do índice de segurança.

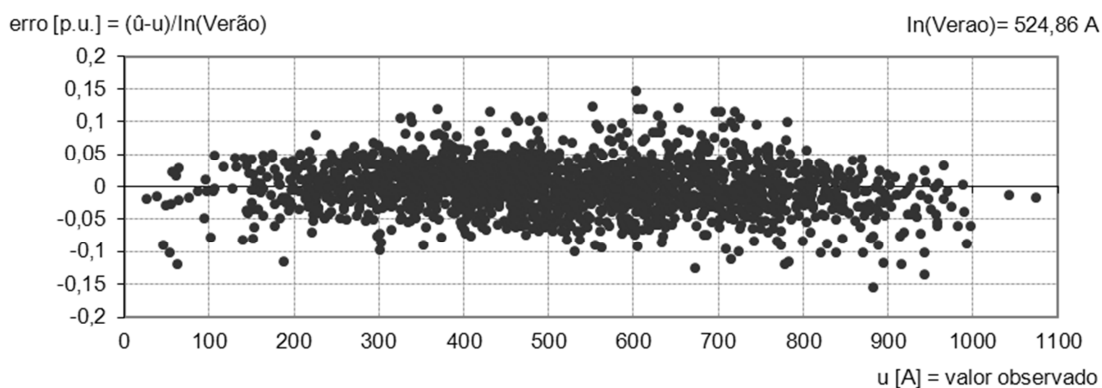


Figura 5.24 - Erros de regressão fornecidos por LRM em função do índice de segurança (avaliação da linha 15-16)

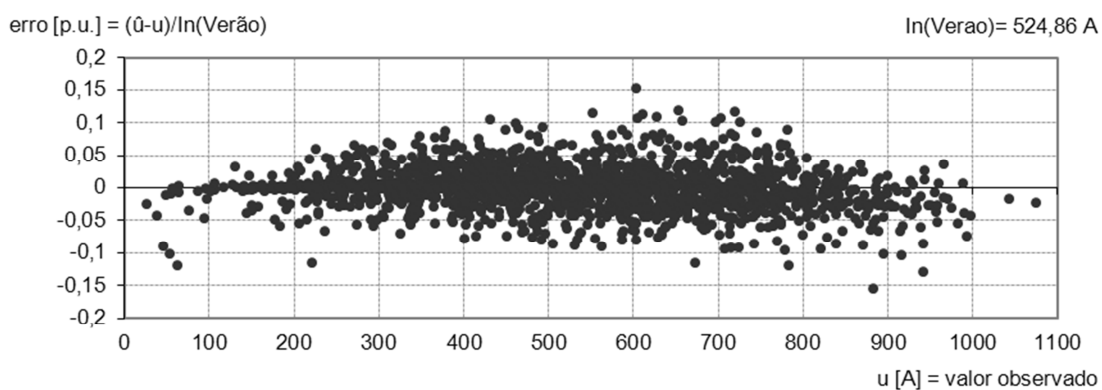


Figura 5.25 - Erros de regressão fornecidos por LRT, 2 folhas, em função do valor do índice de segurança (avaliação da linha 15-16)

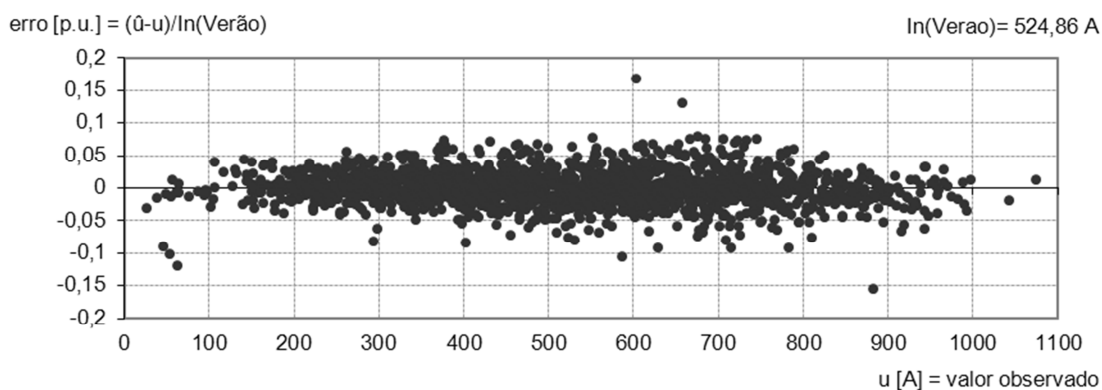


Figura 5.26 - Erros de regressão fornecidos por LRT, 4 folhas, em função do valor do índice de segurança (avaliação da linha 15-16)

Os resultados da Figura 5.27 à Figura 5.29 referem-se à linha 15-17. Da análise destes resultados, pode observar-se que para todas as estruturas treinadas os erros de regressão são inferiores a 0.1 p.u. (valor correspondente a 10% de carga).

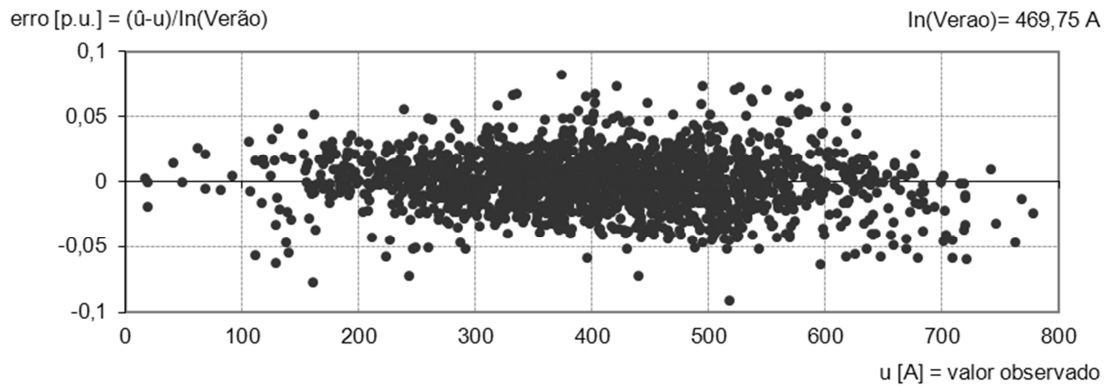


Figura 5.27 - Erros de regressão fornecidos por LRM em função do índice de segurança (avaliação da linha 15-17)

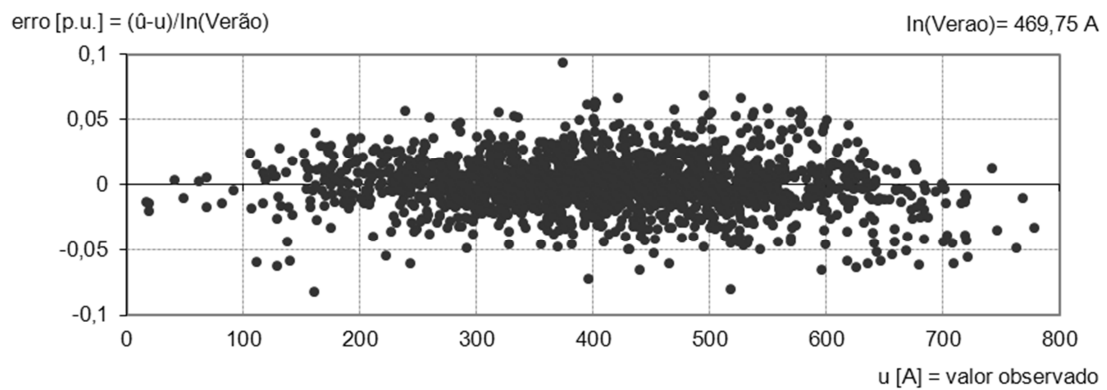


Figura 5.28 - Erros de regressão fornecidos por LRT, 2 folhas, em função do valor do índice de segurança (avaliação da linha 15-17)

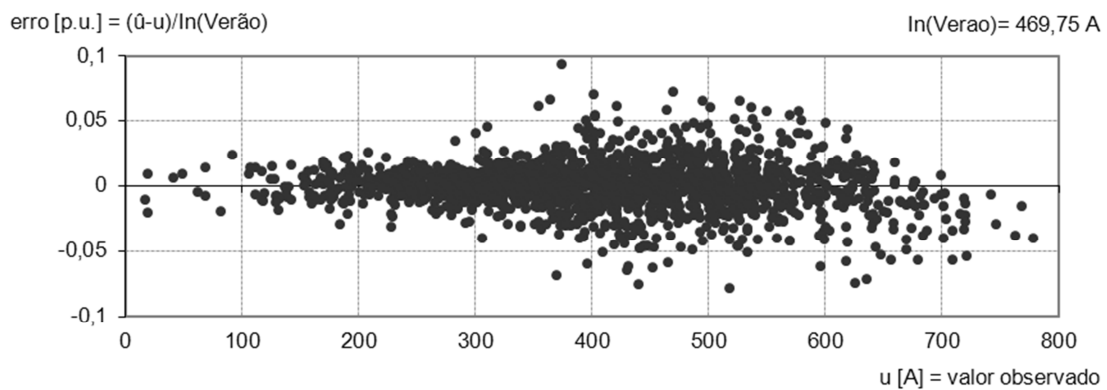


Figura 5.29 - Erros de regressão fornecidos por LRT, 4 folhas, em função do valor do índice de segurança (avaliação da linha 15-17)

Dos erros de regressão fornecidos para a linha 16-4, os quais se apresentam da Figura 5.30 à Figura 5.32, verifica-se que a maioria dos erros é inferior a 10% de carga, tendo ocorrido apenas algumas situações em que os erros foram superiores a este valor, sem que no entanto passassem os

15% de carga. Também aqui a estrutura LRM é a que apresenta mais casos em que o erro dos cenários ultrapassa os 10% de carga.

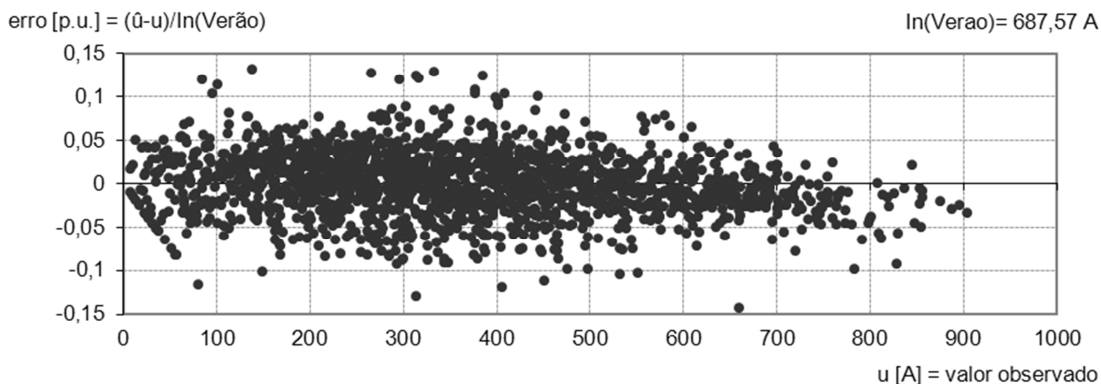


Figura 5.30 - Erros de regressão fornecidos por LRM em função do índice de segurança (avaliação da linha 16-4)

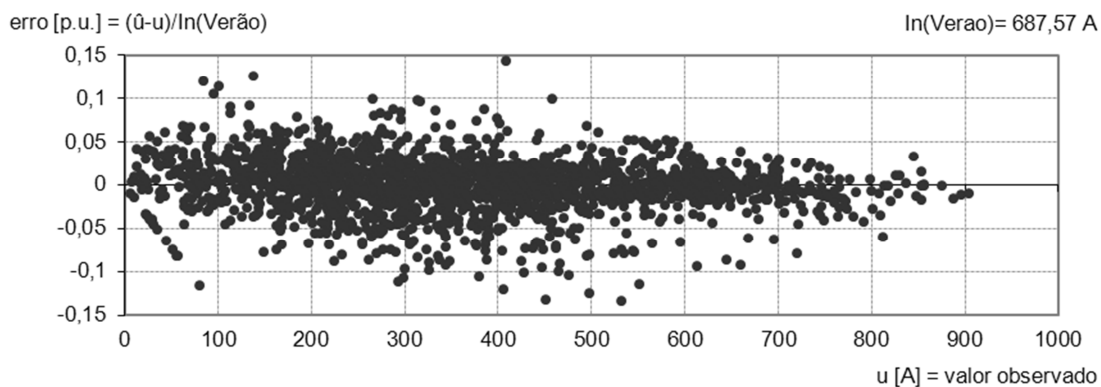


Figura 5.31 - Erros de regressão fornecidos por LRT, 2 folhas, em função do valor do índice de segurança (avaliação da linha 16-4)

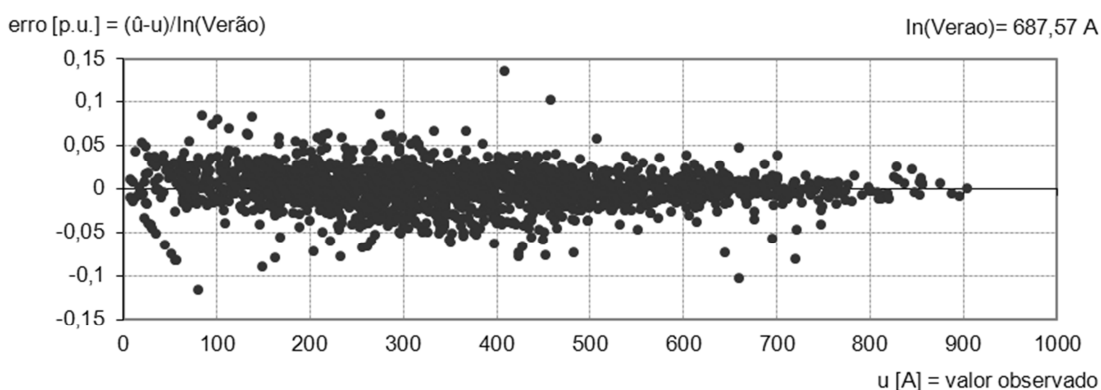


Figura 5.32 - Erros de regressão fornecidos por LRT, 4 folhas, em função do valor do índice de segurança (avaliação da linha 16-4)

Da análise da Figura 5.33 à Figura 5.35, relativas à linha 16-6, verifica-se que existem alguns casos em que os erros de regressão são superiores a 0,15 p.u. No entanto, essa situação corresponde

a um número reduzido quando comparado com o número de cenários de teste cujo erro de regressão surge dentro da margem dos 0,1 p.u. Isto pode dever-se a alguma perda de precisão levantada pela introdução das restrições de linearidade.

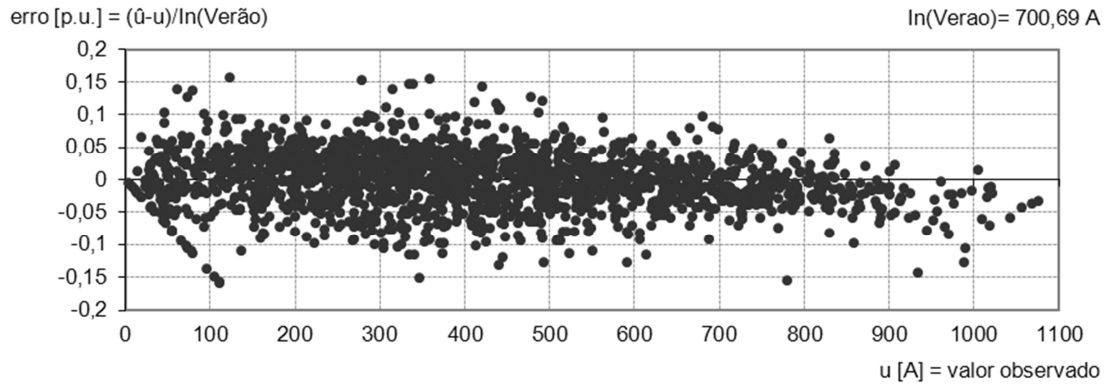


Figura 5.33 - Erros de regressão fornecidos por LRM em função do índice de segurança (avaliação da linha 16-6)

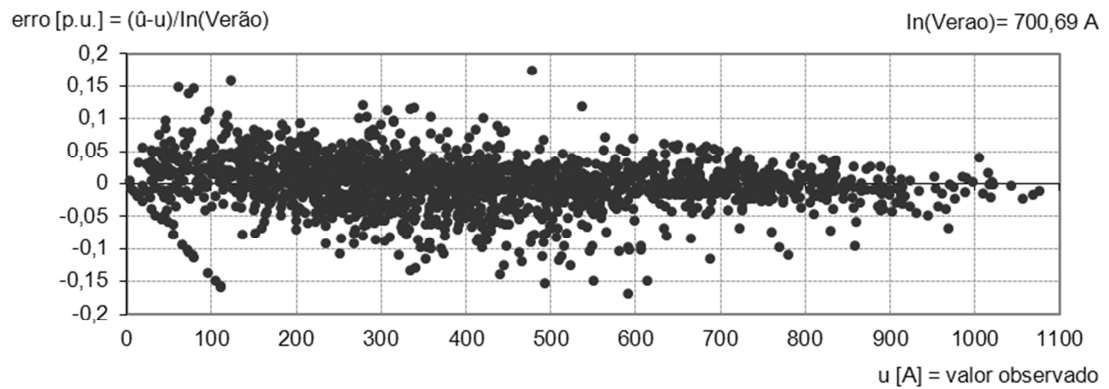


Figura 5.34 - Erros de regressão fornecidos por LRT, 2 folhas, em função do valor do índice de segurança (avaliação da linha 16-6)

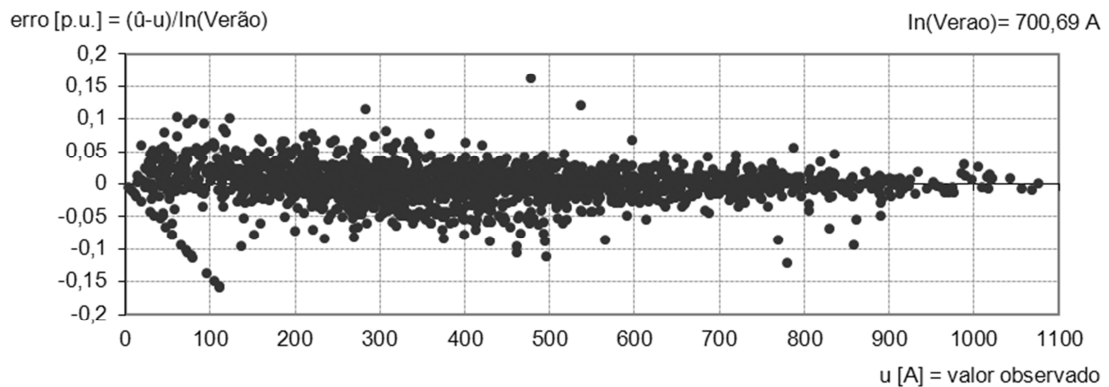


Figura 5.35 - Erros de regressão fornecidos por LRT, 4 folhas, em função do valor do índice de segurança (avaliação da linha 16-6)

Da Figura 5.36 à Figura 5.38, observa-se que os erros de regressão fornecidos para a linha 14-18 são maioritariamente inferiores a 0,1 p.u. Apenas em alguns erros fornecidos pelas LRM, ocorrem valores superiores a 10%, mas que não ultrapassam um valor de 15% de carga.

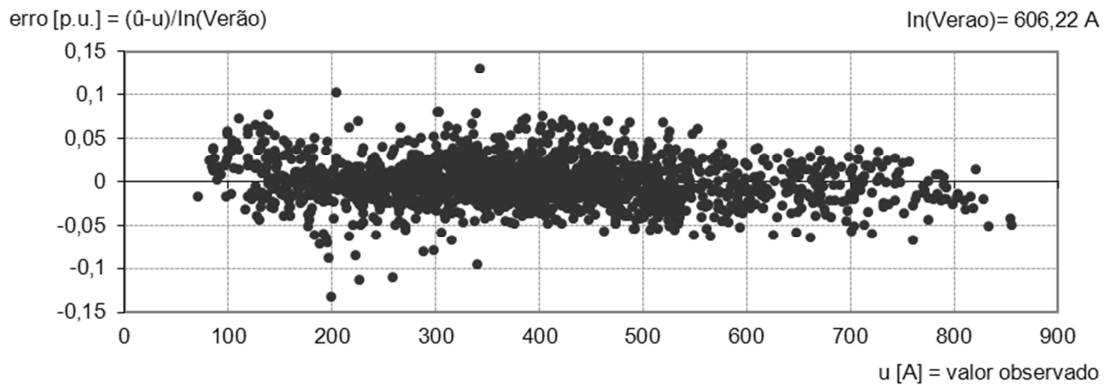


Figura 5.36 - Erros de regressão fornecidos por LRM em função do índice de segurança (avaliação da linha 14-18)

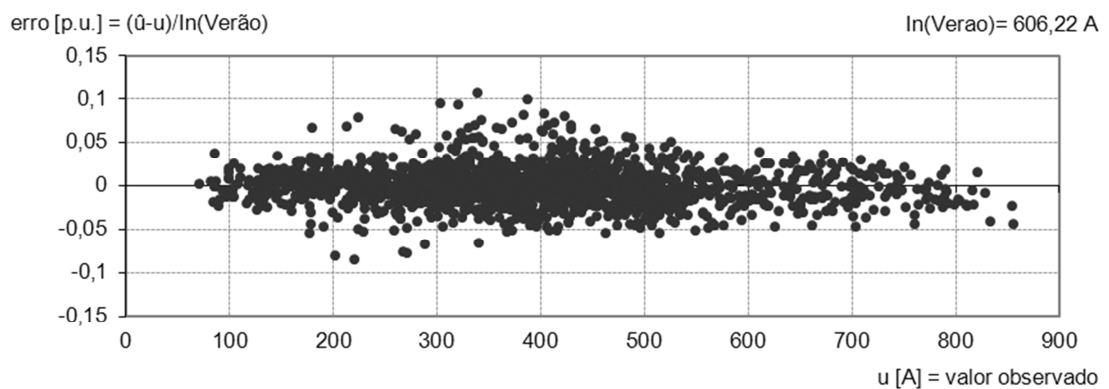


Figura 5.37 - Erros de regressão fornecidos por LRT, 2 folhas, em função do valor do índice de segurança (avaliação da linha 14-18)

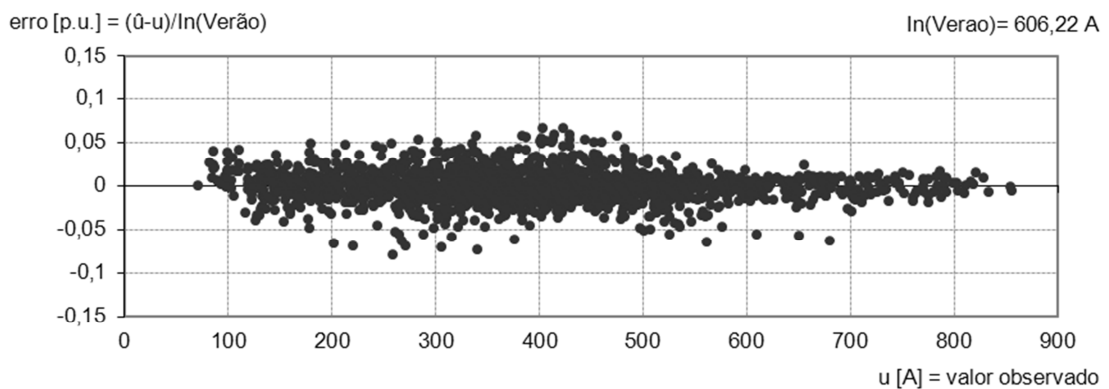


Figura 5.38 - Erros de regressão fornecidos por LRT, 4 folhas, em função do valor do índice de segurança (avaliação da linha 14-18)

Analisando a Figura 5.39 à Figura 5.41, observa-se que todos os erros de regressão fornecidos para a linha 16-18 são inferiores a 10% de carga. No caso das estruturas de LRT a maioria dos resultados situa-se mesmo abaixo dos 0,05p.u. (correspondente a 5% de carga).

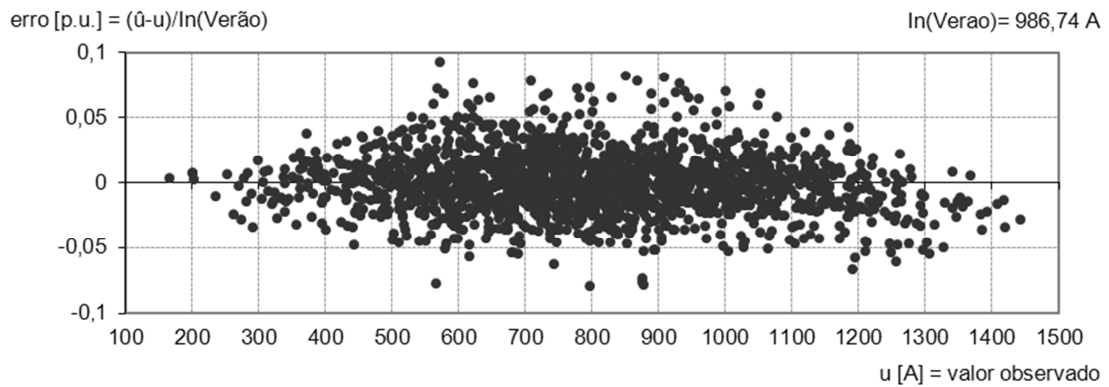


Figura 5.39 - Erros de regressão fornecidos por LRM em função do índice de segurança (avaliação da linha 16-18)

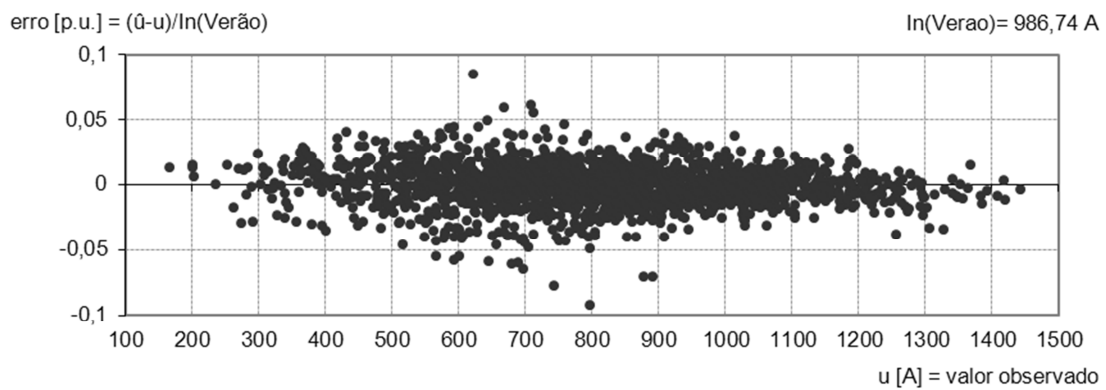


Figura 5.40 - Erros de regressão fornecidos por LRT, 2 folhas, em função do valor do índice de segurança (avaliação da linha 16-18)

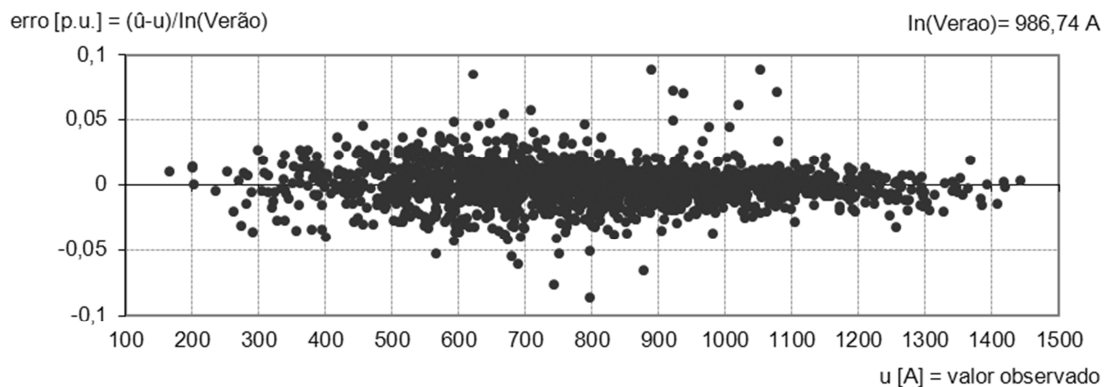


Figura 5.41 - Erros de regressão fornecidos por LRT, 4 folhas, em função do valor do índice de segurança (avaliação da linha 16-18)

Da Figura 5.42 à Figura 5.44, pode observar-se que a maior parte dos erros de regressão fornecidos para a linha 20-21 apresentam um valor inferior a 0,05p.u. Apenas alguns cenários de teste resultaram em erros que se encontram dentro da gama de valores compreendida entre 0,05p.u. e 0,15p.u.

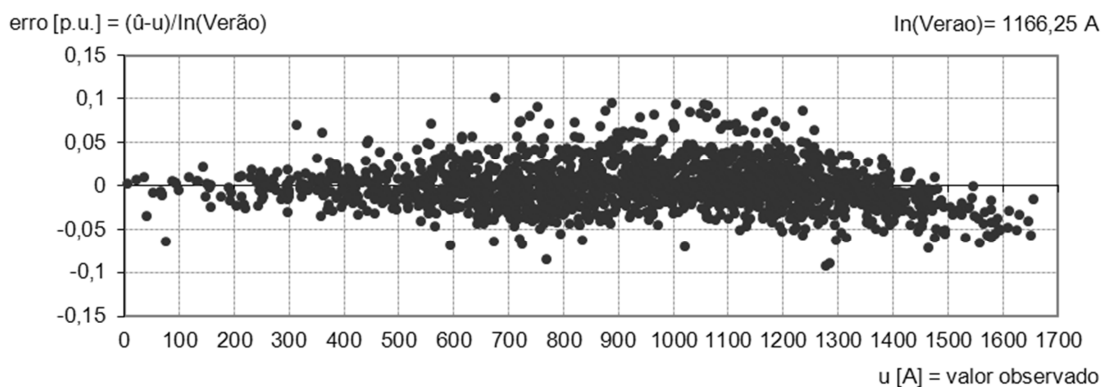


Figura 5.42 - Erros de regressão fornecidos por LRM em função do índice de segurança (avaliação da linha 20-21)

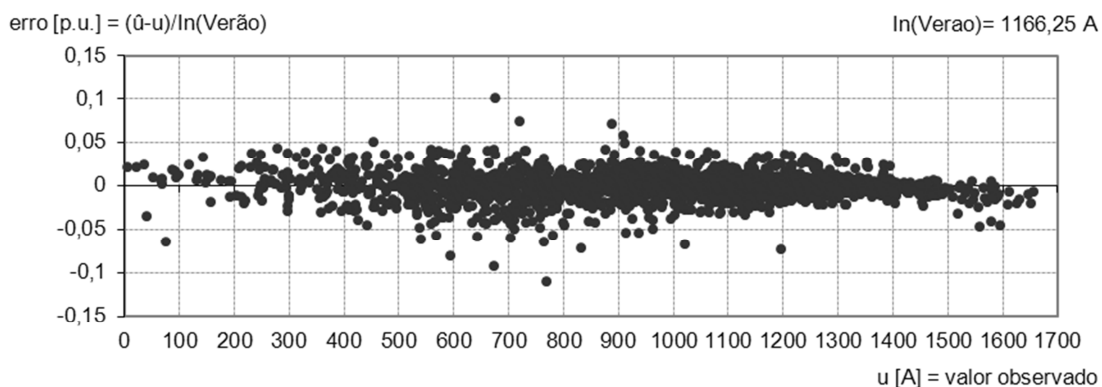


Figura 5.43 - Erros de regressão fornecidos por LRT, 2 folhas, em função do valor do índice de segurança (avaliação da linha 20-21)

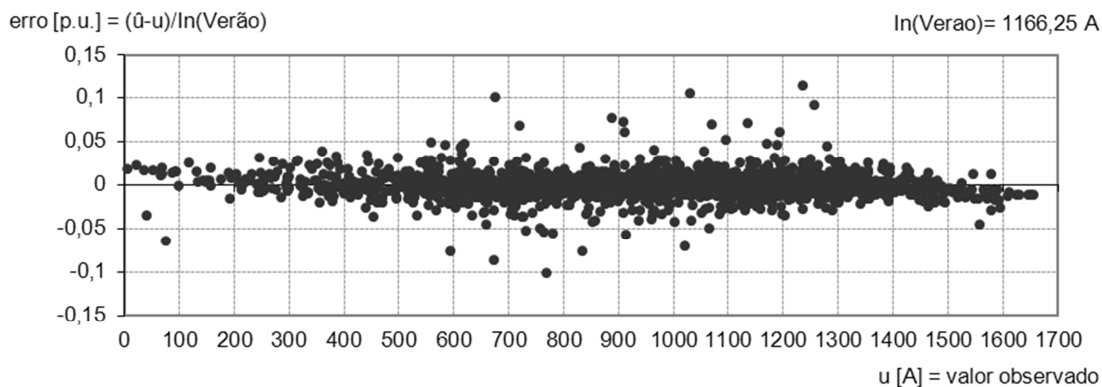


Figura 5.44 - Erros de regressão fornecidos por LRT, 4 folhas, em função do valor do índice de segurança (avaliação da linha 20-21)

Em termos gerais, observa-se uma melhoria dos erros de regressão, quando estes resultam de estruturas do tipo LRT.

### 5.4.3 Comparação entre LRM, LRT de 2 folhas e LRT de 4 folhas (alternativa C):

Recorde-se que, para a alternativa C, as estruturas em árvore dos LRT foram treinadas com a consideração de restrições de linearidade que permitam alterações ao nível do despacho e escalonamento dos grupos. Parte-se do pressuposto que as restrições de linearidade introduzidas para esta situação degradam a qualidade da previsão efectuada ao nível das estruturas LRT. Ainda assim, verifica-se que os resultados fornecidos pelas LRT continuam a ser melhores do que as fornecidas pelos LRM.

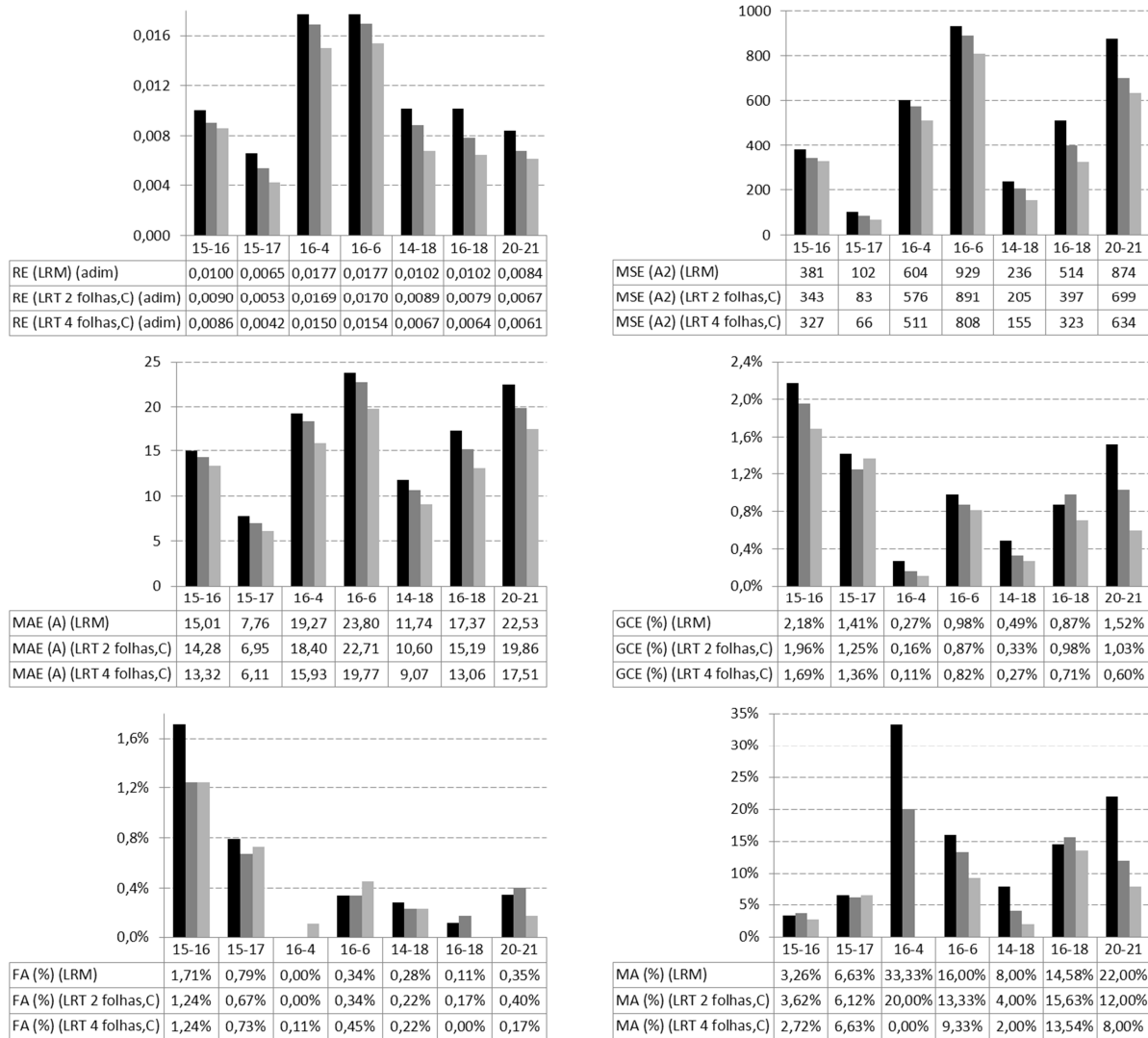


Figura 5.45 - Erros de regressão e de classificação fornecidos por LRM, LRT de 2 folhas e LRT de 4 folhas, considerando a alternativa C

#### 5.4.4 Comparação entre LRM, LRT de 2 folhas e LRT de 4 folhas (alternativa D):

Com a consideração de restrições de linearidade, no treino de LRT, que permitam alterações ao nível do despacho, escalonamento dos grupos e importação, os erros de regressão e de classificação obtidos foram os que se apresentam na Figura 5.46. Estas restrições degradam, tal como no caso anterior, a qualidade da previsão efectuada ao nível das estruturas LRT. A estrutura de 4 folhas é superior, em termos de qualidade de previsão, em relação aos restantes tipos de estrutura treinados. O tipo de estrutura que fornece os piores resultados em termos de previsão é o LRM.

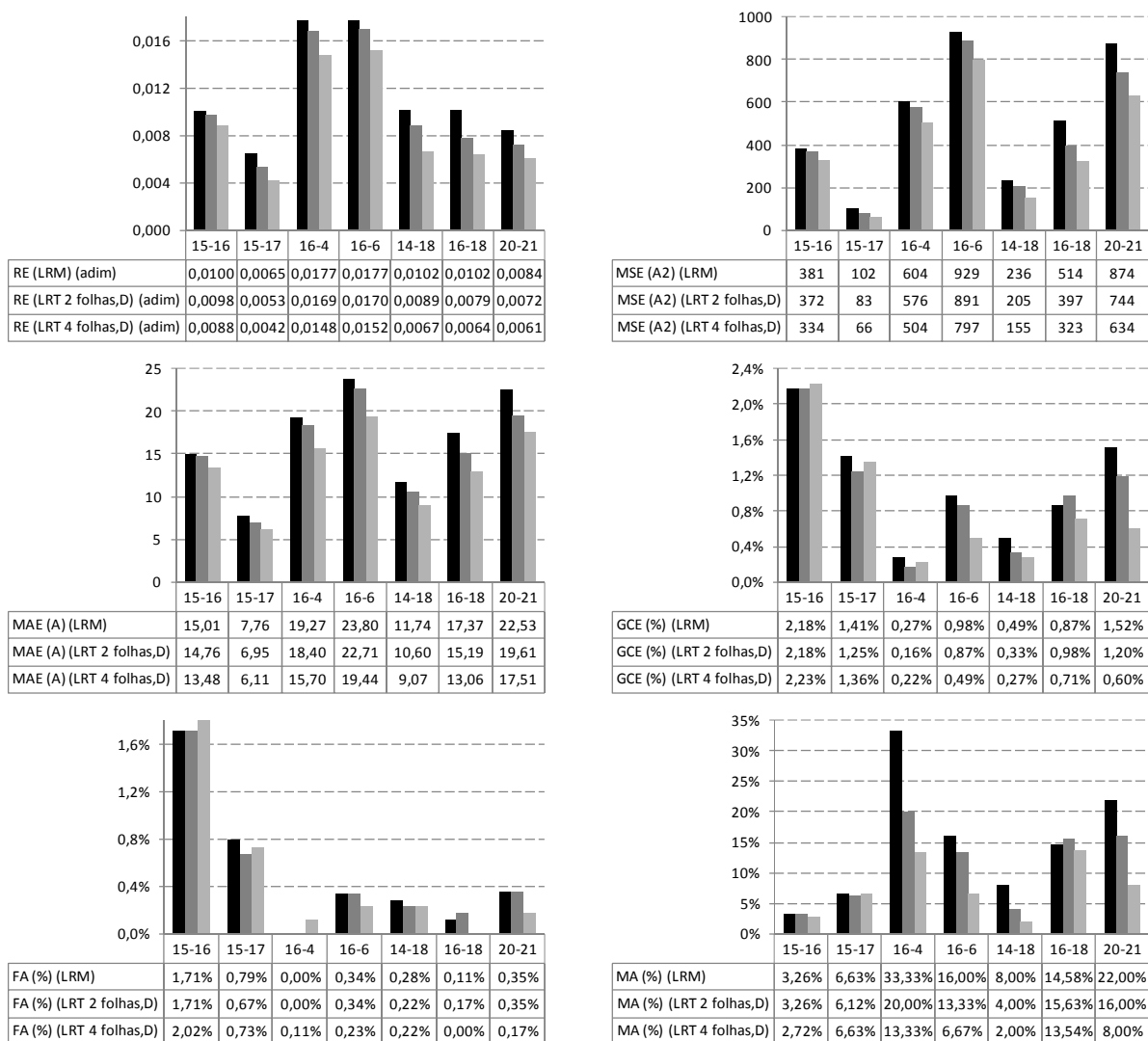


Figura 5.46 - Erros de regressão e de classificação fornecidos por LRM, LRT de 2 folhas e LRT de 4 folhas, considerando a alternativa D

## 5.5 Conclusões

O presente capítulo foi dedicado à apresentação dos resultados obtidos do treino e avaliação de desempenho das estruturas de aprendizagem, abordadas e exploradas neste trabalho para realizar a avaliação de segurança dinâmica da rede interligada de teste.

Para abordar a temática das árvores de regressão linear (LRT), antecedeu-se o estudo e o treino de modelos de regressão linear (LRM) e verificou-se que para se obterem estruturas eficazes para a obtenção de medidas de controlo preventivo será conveniente proceder à validação dos LRM treinados, para que os valores dos parâmetros obtidos não violem as relações matemáticas e físicas existentes entre as variáveis que figuram num problema de segurança.

Do treino efectuado para as árvores de regressão linear (LRT), foi possível verificar que os resultados obtidos permitem a este tipo de estrutura superar os resultados obtidos pelos LRM em termos de precisão para efectuar a avaliação do problema de segurança em análise. As suposições consideradas para o tipo de medidas de controlo a realizar, ilustrado pelas alternativas A a D, exigiram que as LRT se apresentassem com restrições de segurança do tipo linear sem, no entanto, deixarem de ser superiores em relação às LRM. Os resultados obtidos permitiram inferir sobre o superior desempenho, em termos de precisão, das estruturas de LRT sobre a aplicação isolada de LRM, para avaliação do problema de segurança em análise.

A comparação dos erros de regressão e de classificação permitiram a estimação de desempenho entre as estruturas de LRM e de LRT, para a variante de 2 folhas e para a variante de 4 folhas. A estrutura de aprendizagem LRT na variante de 4 folhas apresentou, no geral, a melhor precisão na avaliação do problema de segurança em análise.

## Capítulo 6: Conclusões

Na presente dissertação foram implementadas técnicas de aprendizagem automática, nomeadamente os modelos de regressão linear (LRM) e uma estrutura de aprendizagem híbrida que resulta do treino de um modelo de regressão linear em cada folha de uma árvore de regressão (LRT). Foi apresentada a metodologia para a aplicação destas ferramentas no sentido de permitir a rápida avaliação de segurança de operação de áreas de controlo de redes interligadas onde sejam explorados elevados volumes de produção eólica e onde não tenham adoptado totalmente a capacidade de sobreviver a cavas de tensão.

Compreendeu-se que os sistemas de energia tendem a ser explorados cada vez mais próximos dos seus limites e que a precisão de estimação das previsões se revela muito importante para uma gestão eficaz do sistema de energia. A crescente integração da produção eólica na rede levanta alguns problemas de segurança às redes interligadas, sendo que foi abordado o problema da ocorrência de sobrecargas em regime quasi-estacionário, em ramos de transmissão, que ultrapassem os limites de operacionalidade usualmente adoptados, e cuja ocorrência de defeitos leve à perda significativa de produção eólica.

As características das árvores de regressão permitiram a introdução de restrições de segurança do tipo linear, descritas na secção 5.3 pelas alternativas A a D.

Dos resultados obtidos, Capítulo 5, foi possível observar que as LRT permitem superar os LRM em termos de precisão para todas as experiências realizadas, mesmo nas condições em que foram adoptadas condicionantes de linearidade no treino das LRT. Essas condicionantes trouxeram alguma degradação na qualidade da precisão das LRT, não afectando, no entanto, o resultado final de o desempenho destas estruturas ser superior ao dos LRM.

Comparando os LRT relativamente à dimensão da árvore com o desempenho obtido, pode concluir-se que, para a quase totalidade dos casos, a estrutura com 7 nós oferece superioridade de precisão face à estrutura de 3 nós, ainda que não seja tão expressiva como quando comparando com os desempenhos fornecidos pelos LRM.

Do ponto de vista da análise em que o tempo de treino é um factor chave, pode ser mais interessante a utilização de uma estrutura híbrida de aprendizagem automática com 3 nós, uma vez que os tempos de execução para este tipo de estrutura são consideravelmente inferiores e rondam a casa dos 5 minutos. O tempo de treino para 7 nós, por cada árvore, ronda os 25 minutos. Quando executado num computador com um processador Intel Core 2 Duo a 1,83 GHz, 1 GB de memória RAM e sistema operativo Windows 7.

Quando o critério de desempenho é a qualidade da precisão, importante do ponto de vista de gestão eficiente do sistema de energia, então deverá optar-se por uma estrutura híbrida de aprendizagem automática com 7 nós.

Sob perda de generalização, não se deve aumentar indefinidamente o número de nós, já que o número de exemplos de treino pertencentes a cada nó pode-se tornar muito reduzido.

## Referências

1. *ENTSO-E Work Program 2009 Through December 2010*. ENTSO-E. 2009.
2. *Wind in Power 2009 European statistics*. EWEA. February 2010.
3. *Machine-Learning Approaches to Power-System Security Assessment*. Wehenkel, Louis. 1997, IEEE Expert, pp. 60-72.
4. *Power System Stability and Control*. Kundur, Prabha. New York : McGraw Hill, 1993.
5. *Operating Under Stress and Strain*. Fink, L. H. e Carlsen, K. s.l. : IEEE Spectrum, March 1978, Vols. 15, pp. 48-53.
6. *Avaliação e controlo de segurança de redes interligadas com grande penetração eólica com base em métodos de aprendizagem automática*. Vasconcelos, Maria Helena Osório Pestana de. Porto : Tese de Doutoramento, FEUP, 2007.
7. *Application of Hybrid Automatic Learning Techniques for Fast Dynamic Security Assessment of Isolated Power Systems with Wind Power Production*. Vasconcelos, Maria Helena Osório Pestana de. Porto : Tese de Mestrado, FEUP, 1999.
8. *Inductive Learning of Tree-Based Regression Models*. Torgo, Luís. Porto : Tese de Doutoramento, FCUP, 1999.
9. *Manual Operacional para a Regressão Linear*. Matos, Manuel António. Porto : FEUP, 1995.
10. SPSS Inc. [Manual do SPSS] <http://www.spss.com>.
11. *Árvores de Regressão 'R'*. Vieira, Pedro Cosme da Costa. Porto : FEP, 2002.
12. *Functional Models for Regression Tree Leaves*. Torgo, Luís. s.l. : Morgan Kaufmann Publishers, 1997. Proceedings of the International Machine Learning Conference.

## Referências

13. *Power System Security Assessment*. Morison, Kip, Wang, Lei e Kundur, Prabha. September/October 2004, IEEE power & energy magazine, pp. 30-39.
14. *On-line Dynamic Security Assessment of Isolated Networks Integrating Large Wind Power Production*. J. A. Peças Lopes. 1999, Wind Engineering Review vol. 23, nº 2, pp. 107-117.
15. *A Comparative Study of Reliable Error Estimators for Pruning Regression Trees*. Torgo, Luís. 1998. Proceedings of the Iberoamerican Conference on AI.
16. *Gestão de Sistemas Eléctricos com Elevada Integração de Geração Eólica - Apontamentos da Disciplina Energia Eólica do Curso de Mestrado em Engenharia Electrotécnica e de Computadores*. J. A. Peças Lopes. s.l. : FEUP, 2006.
17. *Real-time dynamic security assessment: fast simulation and modeling applied to emergency outage security of the electric grid*. R. Schainker, P. Miller, W. Dubbelday, P. Hirsh, G. Zhang. IEEE Power & Energy Magazine, vol 4, nº2, pp. 51-58, March-April 2006.
18. *Large scale integration of wind energy in the European power supply: analysis, issues and recommendations*. EWEA. December 2005