



ESCOLA DE CIÊNCIAS E TECNOLOGIA
DEPARTAMENTO DE QUÍMICA

Avaliação e Controlo da Qualidade da Água de Albufeiras com Base em Modelos Dinâmicos

Catarina Raquel Costa do Couto

Dissertação no âmbito do Mestrado em Análises Químicas Ambientais sob a orientação do Professor Doutor Henrique Agostinho Oliveira Moiteiro Vicente

Évora

2009

Avaliação e Controlo da
Qualidade da Água de Albufeiras
com Base em Modelos
Dinâmicos



Catarina Raquel Costa do Couto

Dissertação no âmbito do Mestrado em Análises Químicas Ambientais sob a orientação do Professor Doutor Henrique Agostinho Oliveira Moiteiro Vicente

Évora

2009

Avaliação e Controlo da Qualidade da Água de Albufeiras com Base em Modelos Dinâmicos

Catarina Raquel Costa do Couto

catarina_couto@hotmail.com

Universidade de Évora

Escola de Ciências e Tecnologia

Departamento de Química

Dissertação no âmbito do Mestrado em Análises Químicas Ambientais sob a
orientação do Professor Doutor Henrique Agostinho Oliveira Moiteiro Vicente

Évora

2009

AGRADECIMENTOS

Apesar de uma dissertação ser um trabalho individual, devido ao seu objectivo académico, há sempre contributos de natureza diversa que não podem deixar de ser mencionados, pelo que expresso aqui os meus sinceros agradecimentos a todos os que contribuíram para este trabalho.

Um agradecimento especial ao Professor Doutor Henrique Agostinho Oliveira Moiteiro Vicente, que desde logo aceitou a orientação deste trabalho, pelo apoio e partilha de saber, pelo estímulo e entusiasmo revelado por esta dissertação, pelas críticas e sugestões relevantes feitas durante a orientação, pela disponibilidade sempre revelada, pelo incansável apoio.

À minha família, em especial aos meus pais, pelo apoio incondicional que me deram.

Ao Luís Caeiro e à Carla Antunes o meu muito obrigada pelo incansável apoio e por toda a ajuda prestada.

A todos os meus amigos que me foram estimulando pela confiança em mim depositada, fazendo-me acreditar que era possível chegar ao fim com sucesso.

A todos, os meus sinceros agradecimentos.

RESUMO

A problemática relacionada com a modelação da qualidade da água de albufeiras pode ser abordada de diversos pontos de vista. Neste trabalho recorre-se a metodologias de resolução de problemas que emanam da Área Científica da Inteligência Artificial, assim como a ferramentas utilizadas na procura de soluções como as Árvores de Decisão, as Redes Neurais Artificiais e a Aproximação de Vizinhanças.

Actualmente os métodos de avaliação da qualidade da água são muito restritivos já que não permitem aferir a qualidade da água em tempo real. O desenvolvimento de modelos de previsão baseados em técnicas de Descoberta de Conhecimento em Bases de Dados, mostrou ser uma alternativa tendo em vista um comportamento pró-activo que pode contribuir decisivamente para diagnosticar, preservar e requalificar as albufeiras.

No decurso do trabalho, foi utilizada a aprendizagem não-supervisionada tendo em vista estudar a dinâmica das albufeiras sendo descritos dois comportamentos distintos, relacionados com a época do ano.

Palavras-chave: Qualidade da Água; Albufeiras; *Data Mining*; Descoberta de Conhecimento em Bases de Dados; Árvores de Decisão; Redes Neurais Artificiais; *k-means*

ABSTRACT

“Assessment and Monitoring of Water Quality of Reservoirs Based on Dynamic Models”

The problems related to the modelling of water quality in reservoirs can be approached from different viewpoints. This work resorts to methods of resolving problems emanating from the Scientific Area of Artificial Intelligence as well as to tools used in the search for solutions such as Decision Trees, Artificial Neural Networks and Nearest-Neighbour Method.

Currently, the methods for assessing water quality are very restrictive because they do not indicate the water quality in real time. The development of forecasting models, based on techniques of Knowledge Discovery in Databases, shows to be an alternative in view of a pro-active behavior that may contribute to diagnose, maintain and requalify the water bodies.

In this work, unsupervised learning was used to study the dynamics of reservoirs, being described two distinct behaviors, related to the time of year.

Keywords: Water Quality; Reservoirs; Data Mining, Knowledge Discovery in Databases, Decision Trees, Artificial Neural Networks, k-means

ÍNDICE GERAL

Agradecimentos	iii
Resumo	v
Abstract	vii
Índice Geral.....	ix
Índice de Figuras	xiii
Índice de Tabelas	xix
Notação	xxvii
Capítulo 1 – Introdução.....	1
1.1 A problemática da gestão da água	3
1.2 Monitorização da qualidade da água	5
1.3 Gestão do conhecimento.....	7
1.4 Modelação de ecossistemas aquáticos	9
1.5 Inovação e valor acrescentado do trabalho	13
1.6 Estrutura e organização do trabalho.....	14
Capítulo 2 – Caracterização do Local de Trabalho.....	17
2.1 Caracterização das albufeiras em estudo	18
2.1.1 Albufeira do Alvito.....	20
2.1.2 Albufeira de Odivelas	23
2.2 Caracterização da qualidade da água das albufeiras estudadas	26
2.2.1 Avaliação do estado trófico	26

2.2.2 Avaliação da qualidade da água de acordo com as suas características de qualidade para usos múltiplos	30
2.2.3 Qualidade das águas doces superficiais destinadas à produção de água para consumo humano	33
2.2.4 Qualidade mínima para as águas destinadas a rega.....	36
2.2.5 Qualidade mínima para as águas balneares.....	38
Capítulo 3 – Descoberta de Conhecimento em Base de Dados	41
3.1 Descoberta de conhecimento em base de dados.....	42
3.2 <i>Data mining</i>	45
3.3 Modelos e técnicas de <i>data mining</i>	47
3.3.1 Árvores de decisão	49
3.3.2 Indução de regras.....	52
3.3.3 Redes neuronais artificiais	53
3.3.4 Aproximação de vizinhanças.....	60
3.3.5 Avaliação dos modelos.....	62
3.3.5.1 Avaliação de modelos de classificação	63
3.3.5.2 Avaliação de modelos de regressão	64
Capítulo 4 – Procedimento Experimental	65
4.1 Controlo da qualidade da água	65
4.2 Amostragem e preservação das amostras.....	67
4.3 Procedimento analítico	70
4.4 Base de dados	70
Capítulo 5 – Apresentação, Interpretação e Análise de Resultados e Modelos.....	75
5.1 Breve descrição da ferramenta Weka	75

5.2 Modelos de previsão dos parâmetros da qualidade da água	76
5.2.1 Preparação dos dados	77
5.2.2 Modelação.....	79
5.3 Apresentação, discussão e avaliação dos modelos de classificação	82
5.3.1 Modelos de previsão da qualidade da água – albufeira do Alvito.....	83
5.3.2 Modelos de previsão da qualidade da água – albufeira de Odivelas	94
5.4 Apresentação, discussão e avaliação dos modelos de segmentação	106
5.4.1 Modelos de segmentação – albufeira do Alvito.....	107
5.4.2 Modelos de segmentação – albufeira de Odivelas.....	110
Capítulo 6 – Conclusões Finais e Perspectivas Futuras.....	115
Referências Bibliográficas.....	117
Anexo I	I
Anexo II.....	V
Anexo III.....	VII
Anexo IV.....	XXXI
Anexo V.....	XXXV
Anexo VI.....	XXXVII
Anexo VII	XXXIX
Anexo VIII	LXI
Anexo IX.....	LXV

ÍNDICE DE FIGURAS

Figura 1 – Ciclo biogeoquímico da água.....	2
Figura 2 – Elementos de um plano de monitorização da qualidade da água.....	6
Figura 3 – Esquema do desenvolvimento da modelação ambiental.....	11
Figura 4 – Localização das albufeiras do Alvito e de Odivelas.....	17
Figura 5 – Vista parcial da albufeira do Alvito.....	20
Figura 6 – Localização da bacia hidrográfica da albufeira do Alvito.....	21
Figura 7 – Usos do solo na área da bacia hidrográfica da albufeira do Alvito.....	22
Figura 8 – Fontes de poluição pontual na bacia hidrográfica da albufeira do Alvito.....	22
Figura 9 – Vista parcial da albufeira de Odivelas.....	23
Figura 10 – Localização da bacia hidrográfica da albufeira de Odivelas.....	24
Figura 11 – Usos do solo na área da bacia hidrográfica da albufeira de Odivelas.....	25
Figura 12 – Fontes de poluição pontual na bacia hidrográfica da albufeira de Odivelas.....	25
Figura 13 – Tipos de eutrofização.....	28
Figura 14 – Classificação da qualidade da água da albufeira do Alvito entre 1996 e 2008.....	33
Figura 15 – Classificação da qualidade da água da albufeira do Odivelas entre 2001 e 2008.....	33
Figura 16 – Praia fluvial localizada na albufeira de Odivelas.....	39
Figura 17 – Processo de descoberta de conhecimento em bases de dados.....	43
Figura 18 – Representação esquemática de um neurónio.....	55
Figura 19 – Estrutura geral de um nodo.....	56
Figura 20 – Alguns exemplos de funções de activação utilizadas em RNAs.....	57
Figura 21 – Arquitectura de uma RNU e de uma RNR.....	58
Figura 22 – Validação cruzada com k iterações.....	63
Figura 23 – Estações de qualidade, meteorológicas e hidrométricas das albufeiras do Alvito e Odivelas.....	69
Figura 24 – Excerto da base de dados utilizada.....	71

Figura 25 – Etapas de desenvolvimento do processo de DCBD agrupadas em tarefas.....	76
Figura 26 – Distribuição do número de casos por classe de qualidade para as albufeiras estudadas.....	78
Figura 27 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito.....	83
Figura 28 – RNA obtida para previsão da qualidade da água da albufeira do Alvito.....	85
Figura 29 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito utilizando o pH e os parâmetros hidrométicos e meteorológicos.....	92
Figura 30 – RNA obtida para previsão da qualidade da água da albufeira do Alvito utilizando o pH e os parâmetros hidrométicos e meteorológicos.....	93
Figura 31 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas.....	95
Figura 32 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas.....	97
Figura 33 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas utilizando o pH, o OD e os parâmetros hidrométicos e meteorológicos.....	103
Figura 34 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas utilizando o pH, o OD e os parâmetros hidrométicos e meteorológicos.....	105
Figura 35 – Associações entre os segmentos e os meses do ano em que foi realizada a amostragem.....	108
Figura 36 – Árvore de decisão explicativa do modelo de segmentação apresentado na Tabela 36, seguindo a estratégia 1 e a estratégia 2.....	109
Figura 37 – Associações entre os segmentos e os meses do ano em que foi realizada a amostragem.....	111
Figura 38 – Árvore de decisão explicativa do modelo de segmentação apresentado na Tabela 38 seguindo a estratégia 1 e a estratégia 2.....	112
Figura III.1 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro pH.....	VII

Figura III.2 – RNA obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro pH.....	IX
Figura III.3 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro CQO.....	XI
Figura III.4 – RNA obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro CQO.....	XII
Figura III.5 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro oxidabilidade.	XVI
Figura III.6 – RNA obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro oxidabilidade.	XVII
Figura III.7 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito, excluindo os parâmetros pH e CQO.....	XVIII
Figura III.8 – RNA obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro pH e CQO.	XIX
Figura III.9 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito, excluindo os parâmetros pH e oxidabilidade.....	XXI
Figura III.10 – RNA obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro pH e oxidabilidade.....	XXII
Figura III.11 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito, excluindo os parâmetros CQO e oxidabilidade.	XXIV
Figura III.12 – RNA obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro CQO e oxidabilidade.....	XXV
Figura III.13 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito, excluindo os parâmetros pH, CQO e oxidabilidade.....	XXVII
Figura III.14 – RNA obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro pH, CQO e oxidabilidade.....	XXVIII
Figura IV.1 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito, tendo em consideração a utilização de parâmetros meteorológicos e hidrométricos.	XXXI
Figura IV.2 – RNA obtida para previsão da qualidade da água da albufeira do Alvito, tendo em consideração a utilização de parâmetros meteorológicos e hidrométricos.	XXXIII

Figura VII.1 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo o parâmetro CQO.....	XXXIX
Figura VII.2 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo o parâmetro CQO.....	XL
Figura VII.3 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo o parâmetro oxidabilidade.....	XLI
Figura VII.4 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo o parâmetro oxidabilidade.	XLII
Figura VII.5 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo o parâmetro SST.....	XLV
Figura VII.6 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo o parâmetro SST.....	XLVI
Figura VII.7 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo o parâmetro oxigénio dissolvido.....	XLVIII
Figura VII.8 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo o parâmetro oxigénio dissolvido.	XLIX
Figura VII.9 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo os parâmetros CQO e oxidabilidade.....	LI
Figura VII.10 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo os parâmetros CQO e oxidabilidade.	LII
Figura VII.11 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo os parâmetros SST, OD e oxidabilidade.....	LIII
Figura VII.12 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo os parâmetros SST, OD e oxidabilidade.....	LIV
Figura VII.13 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo os parâmetros SST, OD e CQO.....	LV
Figura VII.14 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo os parâmetros SST, OD e CQO.	LVII
Figura VII.15 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo os parâmetros SST, OD, oxidabilidade e CQO.....	LVIII

Figura VII.16 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo os parâmetros SST, OD, oxidabilidade e CQO..... LIX

Figura VIII.1 - Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas, tendo em consideração a utilização de parâmetros meteorológicos e hidrométricos..... LXI

Figura VIII.2 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas, tendo em consideração a utilização de parâmetros meteorológicos e hidrométricos. LXII

ÍNDICE DE TABELAS

Tabela 1 – Classificação de modelos por pares de tipos.....	9
Tabela 2 – Critério nacional para avaliação do estado trófico	29
Tabela 3 – Avaliação do estado trófico das albufeiras do Alvito e de Odivelas.	30
Tabela 4 – Usos permitidos por classes de qualidade da água.....	31
Tabela 5 – Grelha para classificação de massas de água superficiais de acordo com as suas características de qualidade para usos múltiplos.....	32
Tabela 6 – Grelha para classificação de massas de água superficiais destinadas à produção de água para consumo humano.	34
Tabela 7 – Frequência mínima de amostragem de águas superficiais.	35
Tabela 8 – Divisão dos parâmetros de qualidade de águas superficiais.	36
Tabela 9 – Grelha para verificação da conformidade das águas superficiais destinadas a rega.....	37
Tabela 10 – Parâmetros de qualidade da água.....	38
Tabela 11 – Etapas do processo de DCBD.....	44
Tabela 12 – Caracterização dos objectivos de <i>Data Mining</i>	47
Tabela 13 – Tarefas e técnicas usadas no processo de <i>Data Mining</i>	48
Tabela 14 – Algumas funções de activação utilizadas em RNAs.....	57
Tabela 15 – Matriz de coincidências 2x2.	63
Tabela 16 – Parâmetros determinados e métodos de análise.	70
Tabela 17 – Caracterização, em termos estatísticos, das variáveis utilizadas.....	72
Tabela 18 – Valores assumidos por defeito pelo algoritmo de implementação de RNAs.....	81
Tabela 19 – Valores assumidos por defeito pelo algoritmo de indução de árvores de decisão.....	82
Tabela 20 – Conjunto de regras relativo ao modelo apresentado na Figura 27.....	84
Tabela 21 – Matriz de coincidências relativa ao modelo apresentado na Figura 27.	85
Tabela 22 – Matriz de coincidências relativa ao modelo apresentado na Figura 28.	86
Tabela 23 – Síntese dos diversos modelos obtidos para previsão da qualidade da água da albufeira do Alvito.	89

Tabela 24 – Síntese dos modelos de previsão da qualidade da água da albufeira do Alvito utilizando os parâmetros hidrométicos e meteorológicos.....	91
Tabela 25 – Matriz de coincidências relativa ao modelo apresentado na Figura 29.....	93
Tabela 26 – Matriz de coincidências relativa ao modelo apresentado na Figura 30.....	94
Tabela 27 – Matriz de coincidências relativa ao modelo apresentado na Figura 31.....	95
Tabela 28 – Conjunto de regras relativo ao modelo apresentado na Figura 31.....	96
Tabela 29 – Matriz de coincidências relativa ao modelo apresentado na Figura 32.....	98
Tabela 30 – Síntese dos diversos modelos obtidos para previsão da qualidade da água da albufeira de Odivelas.....	99
Tabela 31 – Síntese dos modelos de previsão da qualidade da água da albufeira de Odivelas utilizando os parâmetros hidrométicos e meteorológicos.....	102
Tabela 32 – Matriz de coincidências relativa ao modelo apresentado na Figura 33....	104
Tabela 33 – Matriz de coincidências relativa ao modelo apresentado na Figura 34....	105
Tabela 34 – Centros de gravidade dos segmentos obtidos com $k = 2$	107
Tabela 35 – Matrizes de coincidências relativas aos modelos apresentados na Figura 36.....	109
Tabela 36 – Conjunto de regras relativo aos modelos apresentados na Figura 36....	110
Tabela 37 – Centros de gravidade dos segmentos obtidos com $k = 2$	111
Tabela 38 – Conjunto de regras relativo aos modelos apresentados na Figura 38....	113
Tabela 39 – Matrizes de coincidências relativas aos modelos apresentados na Figura 38.....	113
Tabela I.1 – Dados da qualidade da água da albufeira do Alvito.....	I
Tabela I.2 – Dados meteorológicos e hidrométricos referentes à albufeira do Alvito.....	II
Tabela I.3 – Dados da qualidade da água da albufeira de Odivelas.....	III
Tabela I.4 – Dados meteorológicos e hidrométricos referentes à albufeira de Odivelas.....	IV
Tabela II.1 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura 28.....	V
Tabela II.2 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura 28.....	V
Tabela II.3 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura 28.....	V

Tabela III.1 – Conjunto de regras relativo ao modelo apresentado na Figura III.1.	VIII
Tabela III.2 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura III.2.	IX
Tabela III.3 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura III.2.	X
Tabela III.4 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura III.2.	XI
Tabela III.5 – Conjunto de regras relativo ao modelo apresentado na Figura III.3.	XII
Tabela III.6 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura III.4.	XIII
Tabela III.7 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura III.4.	XIII
Tabela III.8 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura III.4.	XV
Tabela III.9 – Conjunto de regras relativo ao modelo apresentado na Figura III.5.	XVI
Tabela III.10 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 referente ao modelo apresentado na Figura III.6.	XVII
Tabela III.11 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 referente ao modelo apresentado na Figura III.6.	XVIII
Tabela III.12 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída referente ao modelo apresentado na Figura III.6.	XVIII
Tabela III.13 – Conjunto de regras relativo ao modelo apresentado na Figura III.7.	XIX
Tabela III.14 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 referente ao modelo apresentado na Figura III.8.	XX
Tabela III.15 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 referente ao modelo apresentado na Figura III.8.	XX
Tabela III.16 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída referente ao modelo apresentado na Figura III.8.	XX
Tabela III.17 – conjunto de regras relativo ao modelo apresentado na Figura III.9.	XXII
Tabela III.18 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 referente ao modelo apresentado na Figura III.10.	XXIII
Tabela III.19 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 referente ao modelo apresentado na Figura III.10.	XXIII

Tabela III.20 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída referente ao modelo apresentado na Figura III.10.....	XXIII
Tabela III.21 – Conjunto de regras relativo ao modelo apresentado na Figura III.11.....	XXIV
Tabela III.22 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 referente ao modelo apresentado na Figura III.12.....	XXVI
Tabela III.23 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 referente ao modelo apresentado na Figura III.12.....	XXVI
Tabela III.24 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída referente ao modelo apresentado na Figura III.12.....	XXVI
Tabela III.25 – Conjunto de regras relativo ao modelo apresentado na Figura III.13.....	XXVII
Tabela III.26 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 referente ao modelo apresentado na Figura III.14.....	XXIX
Tabela III.27 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 referente ao modelo apresentado na Figura III.14.....	XXIX
Tabela III.28 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída referente ao modelo apresentado na Figura III.14.....	XXIX
Tabela IV.1 – Conjunto de regras relativo ao modelo apresentado na Figura IV.1.....	XXXII
Tabela IV.2 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura IV.2.....	XXXIII
Tabela IV.3 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura IV.2.....	XXXIV
Tabela IV.4 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura IV.2.....	XXXIV
Tabela V.1 – Conjunto de regras relativo ao modelo apresentado na Figura 29.....	XXXV
Tabela V.2 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura 30.....	XXXVI
Tabela V.3 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura 30.....	XXXVI
Tabela V.4 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura 30.....	XXXVI

Tabela VI.1 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura 32.	XXXVII
Tabela VI.2 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura 32.	XXXVII
Tabela VI.3 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura 32.	XXXVIII
Tabela VII.1 – Conjunto de regras relativo ao modelo apresentado na Figura VII.1.	XL
Tabela VII.2 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura VII.2.	XL
Tabela VII.3 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura VII.2.	XLI
Tabela VII.4 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura VII.2.	XLI
Tabela VII.5 – Conjunto de regras relativo ao modelo apresentado na Figura VII.3.	XLII
Tabela VII.6 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura VII.4.	XLII
Tabela VII.7 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura VII.4.	XLIII
Tabela VII.8 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura VII.4.	XLIV
Tabela VII.9 – Conjunto de regras relativo ao modelo apresentado na Figura VII.5.	XLV
Tabela VII.10 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura VII.6.	XLVI
Tabela VII.11 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura VII.6.	XLVI
Tabela VII.12 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura VII.6.	XLVII
Tabela VII.13 – Conjunto de regras relativo ao modelo apresentado na Figura VII.7.	XLIX

Tabela VII.14 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura VII.8.	L
Tabela VII.15 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura VII.8.	L
Tabela VII.16 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura VII.8.	L
Tabela VII.17 – Conjunto de regras relativo ao modelo apresentado na Figura VII.9.	LI
Tabela VII.18 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura VII.10.	LII
Tabela VII.19 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura VII.10.	LII
Tabela VII.20 – pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura VII.10.	LIII
Tabela VII.21 – Conjunto de regras relativo ao modelo apresentado na Figura VII.11.	LIV
Tabela VII.22 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura VII.12.	LIV
Tabela VII.23 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura VII.12.	LV
Tabela VII.24 – pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura VII.12.	LV
Tabela VII.25 – conjunto de regras relativo ao modelo apresentado na Figura VII.13.	LVI
Tabela VII.26 – pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura VII.14.	LVII
Tabela VII.27 – pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura VII.14.	LVII
Tabela VII.28 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura VII.14.	LVIII
Tabela VII.29 – Conjunto de regras relativo ao modelo apresentado na Figura VII.15.	LVIII

Tabela VII.30 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura VII.16.	LIX
Tabela VII.31 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura VII.16.	LIX
Tabela VII.32 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura VII.16.	LIX
Tabela VIII.1 – Conjunto de regras relativo ao modelo apresentado na Figura VIII.1.	LXII
Tabela VIII.2 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura VIII.2.	LXIII
Tabela VIII.3 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo na Figura VIII.2.	LXIII
Tabela VIII.4 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura VIII.2.	LXIII
Tabela IX.1 – Conjunto de regras relativo ao modelo apresentado na Figura 33.	LXV
Tabela IX.2 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura 34.	LXVI
Tabela IX.3 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura 34.	LXVI
Tabela IX.4 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura 34.	LXVI

NOTAÇÃO

% Sat – Percentagem de saturação

AA – Aprendizagem Automática

AD – Árvores de Decisão

BD – Base de Dados

CBO₅ – Carência Bioquímica de Oxigénio

COT – Carbono Orgânico Total

CQO – Carência Química de Oxigénio

DCBD – Descoberta de Conhecimento em Base de Dados

DL – Decreto-Lei

DM – *Data Mining*

GC – Gestão do Conhecimento

HR₅ – Humidade Relativa (média dos 5 dias anteriores à amostragem)

HR₁₀ – Humidade Relativa (média dos 10 dias anteriores à amostragem)

HR_{am} – Humidade Relativa (média relativa ao período entre amostragens)

IA – Inteligência Artificial

INAG – INstituto da ÁGua

OCDE – Organização para a Cooperação e Desenvolvimento Económico

OD – Oxigénio Dissolvido

OMS – Organização Mundial de Saúde

PP₅ – Precipitação (média dos 5 dias anteriores à amostragem)

PP₁₀ – Precipitação (média dos 10 dias anteriores à amostragem)

PP_{am} – Precipitação (média relativa ao período entre amostragens)

Rad₅ – Radiação (média dos 5 dias anteriores à amostragem)

Rad₁₀ – Radiação (média dos 10 dias anteriores à amostragem)

Rad_{am} – Radiação (média relativa ao período entre amostragens)

RNAs - Redes Neurais Artificiais

SMEWW – Standard Methods for the Examination of Water and Wastewater

SNIRH – Sistema Nacional de Informação de Recursos Hídricos

SST – Sólidos Suspensos Totais

T₅ – Temperatura do ar (média dos 5 dias anteriores à amostragem)

T₁₀ – Temperatura do ar (média dos 10 dias anteriores à amostragem)

T_{am} – Temperatura do ar (média relativa ao período entre amostragens)

ufc°/100 mL – Unidades formadoras de colónias por 100 mililitros

UFES – Universidade Federal do Estado de São Paulo

VMA – Valor Máximo Admissível

VMR – Valor Máximo Recomendado

VV₅ – Velocidade do vento (média dos 5 dias anteriores à amostragem)

VV₁₀ – Velocidade do vento (média dos 10 dias anteriores à amostragem)

VV_{am} – Velocidade do vento (média relativa ao período entre amostragens)

WEKA - Waikato Environment for Knowledge Analysis

CAPÍTULO I – INTRODUÇÃO

A água é considerada um dos recursos naturais mais preciosos do nosso planeta, não só porque é indispensável à vida mas, também, porque dela dependem todos os processos associados ao desenvolvimento humano.

A vida no planeta Terra teve origem na água, no chamado oceano primitivo. Todos os processos de evolução, desenvolvimento e mesmo de sobrevivência do nosso planeta estão dependentes deste líquido precioso. Esta assume um papel relevante, entre outros, na distribuição das populações, na modelação da superfície terrestre, na manutenção dos diferentes climas, na produção e fornecimento de energia, nas comunicações e transportes, nas construções, no tratamento e prevenção de doenças, na alimentação.

Desde as civilizações mais antigas que o homem aprendeu a trabalhar com a água, fixando-se junto às margens dos grandes rios, donde resultaram as primeiras experiências ao nível da actividade agrícola [Mendes & Oliveira, 2004; Lehr & Keeley, 2005a]. Este foi apenas um ponto de partida para a sedentarização humana.

Foi com os romanos que a água teve outro dos papéis fundamentais para o desenvolvimento humano, com a construção de milhares de quilómetros de aquedutos por toda a Europa e com o incentivo para a utilização adequada desta por parte das populações [Tebbutt, 1998].

A água, no nosso planeta, não constitui uma realidade estática. A água apresenta-se integrada num ciclo biogeoquímico, mantendo-se em circulação constante num sistema fechado (Figura 1) [Mendes & Oliveira, 2004; Lehr et al., 2005].

No mecanismo do ciclo biogeoquímico da água podem ser identificados os seguintes fenómenos:

- 1 – Transferência da água, na forma de vapor, da superfície para a atmosfera, por evaporação, a partir do estado líquido e por evapotranspiração;
- 2 – Transporte de água, entre as diversas fases, através de diferentes processos de circulação locais/regionais e pela circulação global;

- 3 – Condensação parcial do vapor de água presente na atmosfera sob a forma líquida e/ou sólida, ficando em suspensão na atmosfera, formando aerossóis (i.e. nuvens e nevoeiros);
- 4 – Transferência de água da atmosfera para a superfície, por precipitação, quer no estado líquido quer no estado sólido;
- 5 – escoamento superficial;
- 6 – Infiltração no solo.



Figura I – Ciclo biogeoquímico da água. Adaptado de Veolia [Veolia, 2009].

A água, enquanto composto dotado de características físico-químicas próprias, não se encontra na natureza no seu estado puro, i.e. não aparece como um composto incolor, inodoro e insípido. Encontra-se associada a diversas substâncias, que podem estar presentes em solução ou em suspensão e que resultam da sua capacidade de dissolver e transportar os materiais com que contacta. Deste modo, pode transportar substâncias que condicionam a sua utilização [Spellman, 2008].

A problemática da água, do seu uso e da gestão dos recursos hídricos disponíveis não é um problema meramente quantitativo, enquadrado no ciclo biogeoquímico. Na verdade, é, cada vez mais, um problema de qualidade.

A deterioração da qualidade da água e dos ecossistemas aquáticos está, de um modo geral, associada a descargas directas de efluentes urbanos, industriais ou agrícolas sem tratamento ou com tratamento deficiente, bem como com a entradas de nutrientes e de produtos químicos resultantes das práticas agrícolas [Lehr & Keeley, 2005].

1.1 A PROBLEMÁTICA DA GESTÃO DA ÁGUA

A água como recurso essencial à vida é, sem dúvida, um tema muito debatido à escala mundial quer seja ao nível económico, ambiental, social ou político.

O crescimento da população verificado nas últimas décadas, despoletou a intensificação da produção industrial e agrícola, originando maior produção de resíduos que, directa ou indirectamente, são descarregados na natureza e que afectam não só as massas de água superficiais como, também, as subterrâneas, poluindo ou degradando uma percentagem significativa da água doce disponível ao Homem [Varis et al., 2008].

A escassez de água doce com a qualidade necessária para os fins múltiplos que o Homem lhe dá é cada vez maior, de tal forma que começa a competir com o petróleo pelo pico de mediatização. Nas próximas décadas, os fornecimentos de água doce e de água potável serão objecto de pressões significativas. As mudanças climáticas que se reflectem no aumento do nível do mar, nos danos das tempestades e no agravamento dos efeitos sazonais como, por exemplo, as inundações no Inverno e as secas no Verão, irão reduzir a segurança e aumentar a vulnerabilidade dos recursos hídricos. Para além do exposto, os crescentes níveis de poluição, as ameaças de ruptura física dos fornecimentos e os possíveis ataques terroristas contra infra-estruturas críticas trará maior atenção às questões da segurança da água a nível mundial [OCDE, 2006].

Existe uma preocupação abrangente em relação ao facto de uma ineficaz gestão da água poder vir a constituir-se como um dos principais factores de limitação do desenvolvimento sustentável durante as próximas décadas [OCDE, 2003]. Neste contexto, a modelação da qualidade da água em albufeiras encontra-se na base da resolução de problemas de natureza ambiental, tendo vindo a afirmar-se como uma ferramenta relevante para o desenvolvimento sustentado e harmonioso das

populações. Por forma a dar resolução aos problemas de qualidade de água, tendo em vista prevenir problemas futuros, é necessário actuar de forma pró-activa recorrendo, para tal, a métodos de previsão, que possam ser integrados em sistemas de apoio à tomada de decisão [Rodrigues, 1992; Neves et al., 1998; Vicente, 2004; Santos et al., 2005].

No que respeita ao nosso país, na última década, foram criadas inúmeras empresas quer de abastecimento de água, quer de tratamento de águas residuais e de resíduos, visando minimizar o impacto negativo causado pelo crescimento populacional, industrial e pela má gestão dos recursos hídricos verificada durante décadas. No entanto, está-se muito aquém de outros países, tanto à escala europeia como mundial pois, apesar das muitas reformas que têm sido realizadas no sentido de valorizar os recursos hídricos, no que respeita à sensibilização das populações, principalmente ao nível das comunidades agrícolas, pouco ou nada foi feito [Eurostat, 2009a; Eurostat, 2009b].

No que concerne aos recursos hídricos, na última década, as diversas directivas nacionais e comunitárias, associadas à produção de legislação mais minuciosa e a reformas administrativas, levaram a uma malha mais apertada no que respeita à gestão da água em Portugal. O incremento da monitorização da qualidade das massas de água e a consequente verificação da conformidade relativamente ao estabelecido na legislação tem sido, do ponto de vista da avaliação das políticas de melhoria da qualidade da água, uma medida da sua eficácia [Rodrigues et al., 2001].

A Lei da Água (Lei n.º 58/2005, de 29 de Dezembro) transpôs, para o Direito Nacional, a Directiva Europeia n.º 2000/60/CE, de 23 de Outubro. Esta directiva estabelece um quadro de acção comunitário, no domínio da política da água, cujo principal objectivo visa criar um enquadramento para a protecção das águas de superfície interiores, das águas de transição, das águas costeiras e das águas subterrâneas. Neste sentido, obriga os Estados membros a protegerem e melhorarem todas as massas de água para que estas alcancem, até Dezembro de 2015, um estado considerado bom.

1.2 MONITORIZAÇÃO DA QUALIDADE DA ÁGUA

A preservação e o controlo da quantidade e da qualidade da água são temas cada vez mais exigentes, que obrigam a uma gestão cada vez mais apertada deste recurso sendo, para tal, necessário um conhecimento cada vez mais detalhado e atempado do estado e da possível evolução dos ecossistemas aquáticos [Weiner, 2000]. À escala de uma bacia hidrográfica, existe a necessidade de se estabelecerem metodologias que permitam a monitorização da qualidade da água e a caracterização do seu estado. Na verdade, a monitorização da qualidade da água é um dos passos necessários para uma gestão eficaz dos recursos hídricos que procure a melhoria da qualidade da água e, conseqüentemente, alargar os usos a que a mesma se destina. Um plano de monitorização da qualidade da água é, basicamente, gizado tendo em vista:

- ✓ Avaliar a disponibilidade de água;
- ✓ Caracterizar o estado da qualidade da água;
- ✓ Caracterizar as actividades humanas;
- ✓ Caracterizar as alterações, ao longo do tempo, das condições do meio hídrico;
- ✓ Avaliar a qualidade da água tendo em consideração os actuais e os futuros usos da água;
- ✓ Verificar o cumprimento dos objectivos de qualidade tendo, como base, o normativo nacional, comunitário e internacional; e
- ✓ Coleccionar dados que permitam avaliar a eficácia das normas em vigor, que suportem a fixação de novas normas e a elaboração de linhas orientadoras [Zhang, 2007].

Um plano de monitorização da qualidade da água compreende todas as actividades que vão desde a amostragem à obtenção, processamento, verificação e armazenamento dos dados [Bartram & Ballance, 1996]. Do exposto ressalta que a recolha dos dados não deve ser feita ao acaso. Pelo contrário, deve obedecer a um plano estratégico que envolva uma fase de planeamento, uma fase de implementação e, finalmente, uma fase de avaliação [Popek, 2006]. Assim, para que o plano de monitorização possa ser bem executado (i.e. de forma a minimizar erros e a possibilitar resultados fidedignos), deve compreender uma série de etapas, que se apresentam na Figura 2:

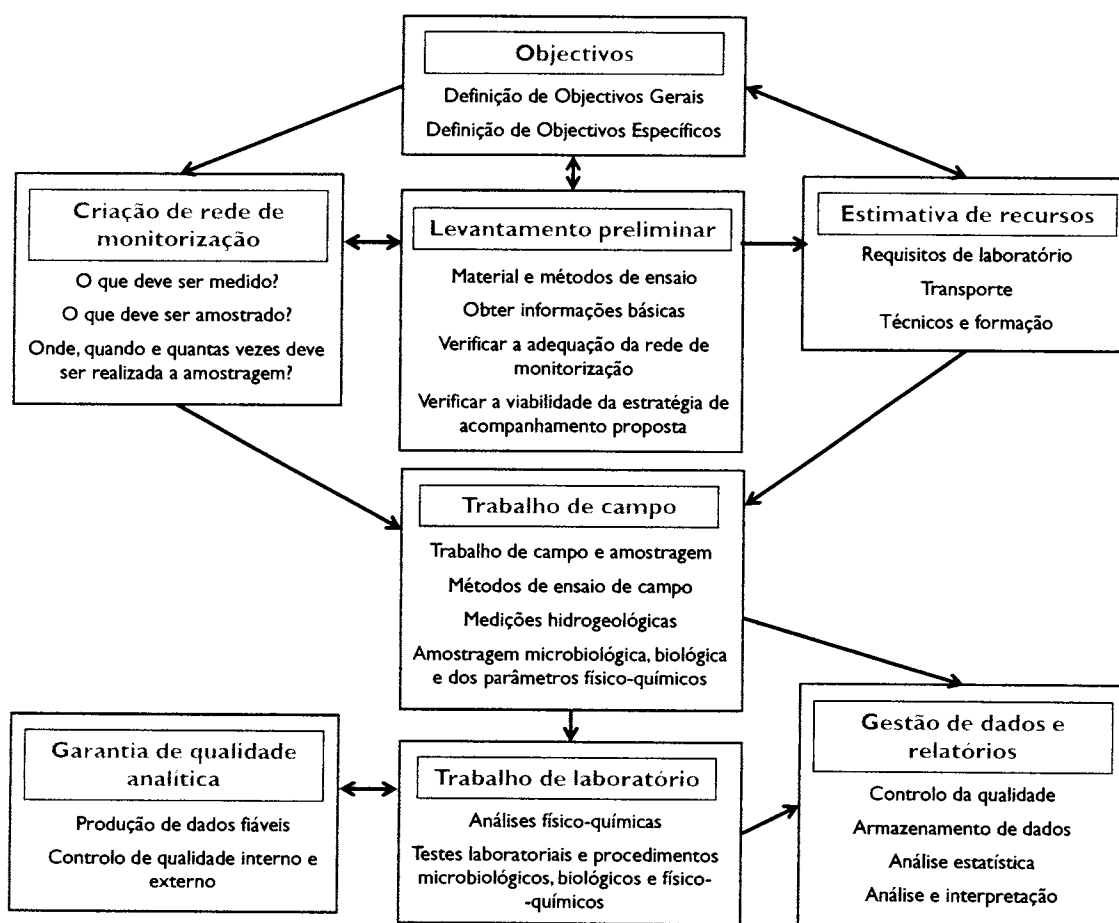


Figura 2 – Elementos de um plano de monitorização da qualidade da água. Adaptado de Bartram [Bartram & Ballance, 1996].

No estabelecimento do plano estratégico apresentado na Figura 2 há várias etapas que se interligam e se influenciam mutuamente. É o caso da definição de objectivos e do levantamento preliminar. Pode-se, à partida, traçar as metas gerais que se procuram atingir. No entanto, não se podem definir objectivos específicos sem ter noção do que é possível executar e sem ter uma estimativa dos recursos existentes. Só assim pode ser criada a rede de monitorização, partir-se para o trabalho de campo e, depois, para o trabalho laboratorial. Os dados recolhidos devem ser fiáveis, devendo, para tal, estar sujeitos a controlo de qualidade tanto interno como externo.

A produção de dados visa a obtenção de mais-valias que se podem traduzir em sequências de acções que levem a uma melhoria contínua, indo ao encontro dos objectivos previamente delineados. É neste contexto que a gestão do conhecimento assume papel relevante no sentido de responder às necessidades de aquisição, partilha

e aplicação do conhecimento gerado pelos diversos colectivos que formam o todo, potenciando comportamentos pró-activos e atempados. Neste particular assume especial relevância a possibilidade de antecipar cenários e problemas.

1.3 GESTÃO DO CONHECIMENTO

O conceito de Gestão do Conhecimento (GC) surgiu na década de 90, associado à estratégia organizacional e à necessidade de produção de conhecimento que se traduzisse em mais-valias para o todo de uma instituição [Sveiby, 1998; Santos, 2001].

Nos últimos anos, a GC tem sido apresentada como uma importante abordagem para resolver os problemas de competitividade e inovação com que se defrontam as organizações utilizando, para tal, a captura, partilha e aplicação do conhecimento de um ou mais colectivos, de forma a tomar decisões em tempo real [Nonaka & Takeuchi, 1997].

Um dos problemas que afecta a maior parte das organizações de todo o mundo prende-se com o elevado volume de dados em bruto e a necessidade de, a partir deles, extrair conhecimento em tempo real para, assim, agir atempadamente. A GC está directamente relacionada com questões de eficiência, de produtividade, de inovação, de organização, de qualidade e do recurso à experiência do singular em articulação com o todo e as formas de meta-conhecimento. Uma abordagem integrada do conhecimento prevê diversas fases sequenciais [Kock, 1999]:

- ✓ A compreensão das relações entre unidades de dados, ou entre partes de dados gera informação;
- ✓ A compreensão dos padrões com elevado nível de fiabilidade e previsibilidade quanto à forma como irão evoluir ao longo do tempo, associado à informação e suas implicações gera conhecimento;
- ✓ A compreensão dos princípios fundamentais responsáveis pelos padrões que representam o conhecimento gera sabedoria;

Um dos argumentos que se encontra na base do desenvolvimento da GC reside no facto de se perspectivar o conhecimento como elemento central ao desempenho

organizacional, tendo por base a compreensão das características do ambiente competitivo e a percepção das necessidades individuais e colectivas, associadas aos processos de criação e de aprendizagem [Terra, 2005].

Na sociedade global da informação o conhecimento das organizações e dos indivíduos que as compõem tem ganho cada vez mais importância. Neste contexto, não é suficiente apostar somente na tecnologia e na gestão da inovação. Deste modo, a GC tem como objectivos melhorar a capacidade das empresas na aquisição, desenvolvimento, preservação, distribuição e uso do conhecimento desenvolvendo, para isso, algumas iniciativas que vão desde as técnicas de suporte criativo, aperfeiçoamento de competências até à implementação de sistemas *groupware* e de inteligência artificial. A partilha do conhecimento entre indivíduos, no seio de uma organização, é uma tarefa complexa, independentemente de estarem mais ou menos próximos quer do ponto de vista físico, quer do ponto de vista intelectual, para que, desta forma, possam:

- 1) Partilhar os seus conhecimentos;
- 2) Aplicar o conhecimento de terceiros;
- 3) Capturar e partilhar o conhecimento num formato que permita ser aplicado dentro da organização;
- 4) Melhorar os canais de comunicação, sejam estes locais ou remotos;
- 5) Interagir com os colegas, em qualquer local, em qualquer altura e de diferentes formas; e
- 6) Utilizar novas aplicações, tecnicamente avançadas, assim como o uso de ferramentas de suporte à decisão, como a mineração de dados, a extracção de conhecimento e a simulação [Scheer et al., 1998].

Na sociedade actual, designada por sociedade do conhecimento, este é encarado como o recurso mais importante. Deste modo, a capacidade de aprender de forma rápida e continuada é a chave para o sucesso das organizações.

I.4 MODELAÇÃO DE ECOSISTEMAS AQUÁTICOS

O Homem começou, desde muito cedo, a utilizar modelos como ferramentas para a compreensão e/ou resolução de problemas. Os modelos podem classificar-se com base em diversas características (e.g. a sua aplicação, as bases científicas, a formulação matemática) [Jørgensen & Bendoricchio, 2001]. A Tabela I mostra uma divisão dos modelos em pares e uma caracterização sumária de cada um deles.

Tabela I – Classificação de modelos por pares de tipos.

Tipo de Modelos	Caracterização
De investigação	Utilizados como ferramenta de investigação
De gestão	Utilizados como ferramenta de gestão
Deterministas	Os valores previstos são calculados de modo exacto
Estocásticos	Os valores previstos dependem da distribuição de probabilidades
Compartimentais	As variáveis que definem o sistema são quantificadas por equações diferenciais que dependem do tempo
Matriciais	Utilizam matrizes na formulação matemática
Reduccionistas	Incluem tantos pormenores relevantes quantos os possíveis
Holísticos	Utilizam princípios gerais
Estáticos	As variáveis que definem o sistema não dependem do tempo
Dinâmicos	As variáveis que definem o sistema dependem do tempo
Distribuídos	Os parâmetros são considerados função do tempo e do espaço
De parâmetros descontínuos	Os parâmetros são constantes e referem-se a localizações espaciais ou temporais
Lineares	Utilizam, consecutivamente, equações do primeiro grau
Não-lineares	Pelo menos uma das equações é de grau superior a um
Causais	As entradas, os estados e as saídas estão relacionadas por relações causais
De caixa negra	As perturbações nas entradas apenas se reflectem nas respostas das saídas. Não é necessário existir causalidade
Autónomos	As derivadas não são explicitamente dependentes da variável tempo
Não-autónomos	As derivadas são explicitamente dependentes da variável tempo

O processo de modelação começa normalmente pela definição do problema e pode ser desenvolvido seguindo um conjunto de etapas. A primeira, designada por conceptualização, consiste numa lista de variáveis de estado e de variáveis externas importantes para o problema em causa e, para além disso, deve ainda conter os processos que relacionam os diferentes componentes. Os modelos conceptuais são uma ferramenta para criar abstrações e delinear o nível de organização que melhor se ajusta aos objectivos.

Uma outra etapa importante é a verificação. A verificação implica um teste à consistência interna do modelo, promovem-se variações e verifica-se se a resposta está de acordo com o que se conhece sobre o sistema. Este exercício é decisivo na construção do modelo, devendo ser introduzidas as melhorias e correcções tidas como necessárias.

Em seguida, devem ser feitos testes à sensibilidade do modelo. Deste modo, tenta obter-se informação sobre quais são os componentes mais sensíveis do modelo (i.e. a importância dos parâmetros, das variáveis de estado, das variáveis externas e das equações utilizadas para descrever os processos). Esta análise fornece, portanto, indicações sobre quais os dados a recolher e sobre o cuidado que se deve ter na recolha de cada um deles.

No passo seguinte comparam-se os resultados fornecidos pelo modelo com um conjunto de dados experimentais independentes – está-se na fase de validação. São formulados critérios de validação com base nos objectivos do modelo e na qualidade dos dados disponíveis permitindo, desta forma, colher indicações sobre a confiança que se pode ter nas previsões.

A sequência apresentada não deve ser entendida como um esquema rígido, sequencial, mas sim como um processo iterativo que pode e deve ser repetido as vezes necessárias [Jørgensen, 1995].

A modelação de ecossistemas remonta aos trabalhos de Streeter e Phelps [Streeter & Phelps, 1925] sobre o balanço de oxigénio num rio e aos trabalhos de Lotka [Lotka, 1924] e de Volterra [Volterra, 1926] sobre as relações predador – presa. Na Figura 3 apresenta-se o esquema do desenvolvimento da modelação ambiental ao longo do tempo.

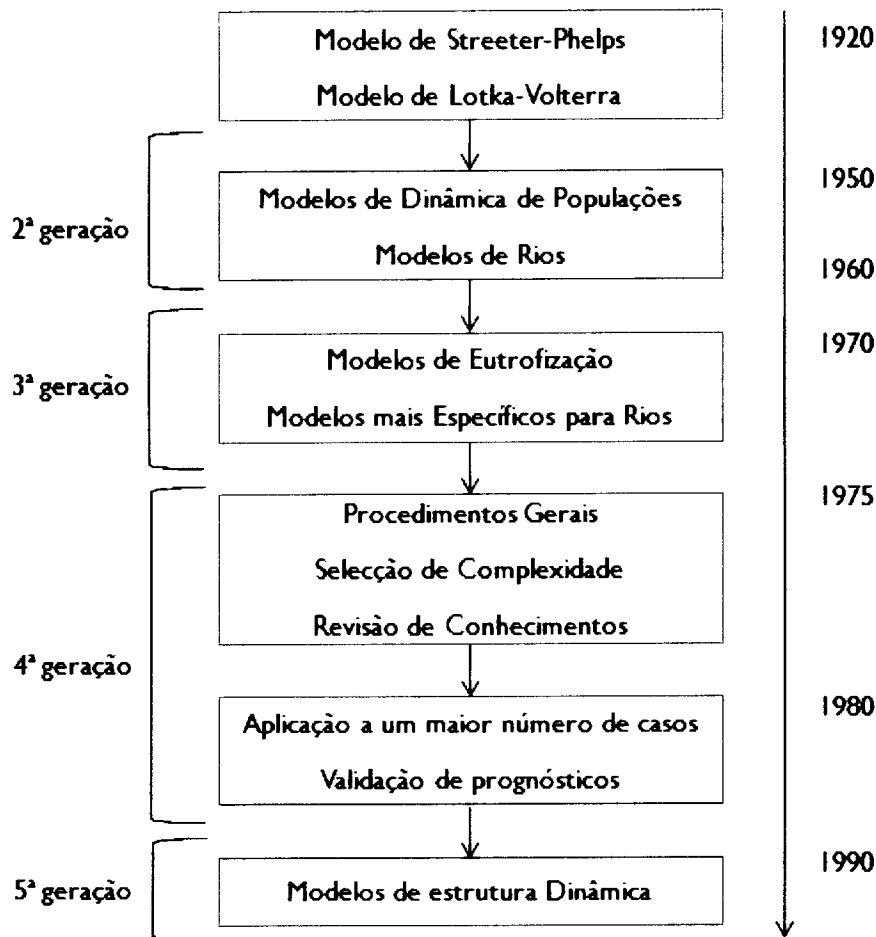


Figura 3 – Esquema do desenvolvimento da modelação ambiental.

Após os trabalhos anteriormente referidos, foi preciso esperar pelas décadas de cinquenta e de sessenta do século XX para assistir ao desenvolvimento de modelos de dinâmica de populações e também de modelos mais específicos para rios, actualmente designados por modelos de segunda geração. Os anos setenta trouxeram um aumento da aplicação de modelos como ferramentas de gestão ambiental, surgindo o primeiro modelo de eutrofização. Em certa medida, o artigo publicado em 1968 por Chen e Orlob intitulado “*A proposed ecological model for a eutrophic environment*” pode ser considerado como o início da terceira geração de modelos, caracterizados pela sua elevada complexidade. O desenvolvimento acelerado da tecnologia computacional, durante a década de 70, permitiu a construção de modelos cada vez mais complexos. No entanto, quer a quantidade/qualidade dos dados disponíveis, quer o conhecimento sobre os sistemas/processos a modelar não acompanharam o desenvolvimento dos meios de cálculo.

A partir de meados dos anos setenta até à década de oitenta foram desenvolvidos modelos designados pela quarta geração, sendo caracterizados por procurarem dar ênfase à simplicidade e ao realismo. Tratam, na grande maioria dos casos, problemas relacionados com a eutrofização [Benndorf & Recknagel, 1982; Recknagel & Benndorf, 1982; Dejak et al., 1987e; Dejak et al., 1987a; Dejak et al., 1987b; Dejak et al., 1987c; Dejak et al., 1987d; Riley & Stefan, 1988; Dejak et al., 1989; Kmet & Straskraba, 1989; Thébault & Saleçon, 1993; Saleçon & Thébault, 1996].

A partir do final dos anos 90 foi dado início a uma nova era, na qual se reconheceu que os modelos incluídos nas quatro primeiras gerações eram demasiado rígidos. Ao assentarem num conjunto fixo de parâmetros os modelos são, muitas vezes, limitados no espaço e no tempo. Não se considerou a capacidade que os ecossistemas possuem de alterar e/ou substituir componentes de acordo com mecanismos auto-organização e de auto-regulação [Straskraba, 1994; Straskraba, 2001]. Deste modo, com o decorrer do tempo, surgiram os primeiros modelos de estrutura dinâmica ou de quinta geração [Jørgensen, 1999], os quais tentam considerar as alterações na composição das espécies bem como a capacidade de estas se adaptarem às condições que lhes são impostas. Estes modelos são designados por modelos de estrutura dinâmica para indicar a sua capacidade de prever as mudanças estruturais, sendo referidos como a quinta geração para distinguir e sublinhar que são diferentes dos anteriores, que podem fazer mais, ou seja, prever alterações na composição das espécies.

Desde essa altura até agora têm sido tentadas diversas vias como a utilização de funções objectivo [Salomonsen & Jensen, 1996; Wilhelm & Bruggemann, 2000; Ray et al., 2001; Chen & Ji, 2007], aplicação da teoria de caos [Hasting et al., 1993; Suárez, 1999; Vayenas & Pavlou, 2001], de modelos baseados na programação orientada ao objecto [Elshorbagy & Ormsbee, 2006] e modelos baseados em conhecimento difuso [Karmakar & Mujumdar, 2007]. Para além das vias referidas anteriormente, assistiu-se ao rápido crescimento de modelos com o recurso à Inteligência Artificial (IA) e a Sistemas Inteligentes (SI) [Dzeroski, 2002; Chau, 2006], onde se destaca a utilização de Redes Neurais Artificiais (RNAs) [Lek & Guégan, 1999; Karul et al., 2000; Jeong et al., 2001; Scardi, 2001; Wilson & Recknagel, 2001; Lee et al., 2003; Recknagel, 2003; Kuo et al., 2007; Tyagi et al., 2008; Yang et al., 2008], Algoritmos Genéticos (AG)

[Bobbin & Recknagel, 2001; Chen, 2003; Jeong et al., 2003; Liu et al., 2007], Árvores de Decisão (AD) e Indução de Regras (IR) [Santos et al., 2005; Atkins et al., 2007; Fernandes et al., 2009; Pinto et al., 2009] e, recentemente, combinações destas técnicas [Kuo et al., 2006; Chaves & Kojiri, 2007].

1.5 INOVAÇÃO E VALOR ACRESCENTADO DO TRABALHO

A importância de qualquer tipo de informação para a resolução de um dado problema está directamente relacionada com a frequência com que é considerada e com o modo como é aplicada nessas situações. A fonte de informação ideal seria aquela que facultasse apenas o conhecimento e os dados relevantes para a resolução do problema, e exigisse, do utilizador, apenas um esforço mínimo. As lacunas que as actuais fontes de informação apresentam tornam claro que algo mais é necessário e devem ser o ponto de partida para o desenvolvimento de alternativas.

Muito se tem dito sobre as ferramentas que irão transformar a Gestão Ambiental. Algumas das suas características, porém, estão finalmente a emergir como, por exemplo, o dar resposta a questões complexas em tempo útil. Deste modo, terão que estar ligadas a bases de dados e de conhecimento, terem como base a electrónica e a robótica, serem portáteis e fáceis de utilizar. As ferramentas não poderão ter uma atitude passiva, ou seja, limitarem-se a responder a questões colocadas pelos utilizadores, mas subscrever um comportamento pró-activo, o que passa por questionar o utilizador, explicar-lhe as suas decisões, procurando ajudar, encontrando uma solução técnica eficiente, desenvolvendo a alternativa mais amiga do ambiente e jamais se assumindo como uma entidade indesejada. São ferramentas com que os gestores ambientais irão tratar a informação quer no presente quer no futuro, o que passa por que se estabeleçam ligações entre as diversas fontes de dados e a difusão dos mesmos através da *Internet*, com o gestor a posicionar-se na sua intersecção, não tendo, contudo, que prestar atenção à tecnologia que está a ser utilizada.

Neste trabalho aplicam-se técnicas da Área Científica da Inteligência Artificial ao desenvolvimento de modelos para duas albufeiras, uma situada nos distritos de Évora e de Beja (albufeira do Alvito) e, a outra, situada no distrito de Beja (albufeira de

Odivelas). Com este trabalho pretende dar-se uma pequena contribuição para o esforço que há a fazer em Portugal, particularmente no Alentejo, na preservação e gestão adequada dos recursos hídricos.

Em resumo, os principais objectivos deste trabalho são os seguintes:

1. Descrever os problemas e fenómenos associados à qualidade da água em albufeiras;
2. Classificar as albufeiras em estudo com base na legislação aplicável e noutros critérios de classificação importantes como o utilizado pelo Instituto da Água;
3. Utilizar as ferramentas da Área Científica da Inteligência Artificial no desenvolvimento de modelos de previsão da qualidade da água para as albufeiras do Alvito e de Odivelas;
4. Realçar a importância de tais ferramentas e dos modelos na gestão e exploração sustentada dos recursos hídricos;
5. Realçar a capacidade dos modelos na aferição da qualidade da água em tempo real e na descrição de cenários futuros;
6. Alertar para a importância da obtenção de dados hidrológicos, meteorológicos, e de qualidade da água que possibilitem desenvolver e aplicar este tipo de modelos;
7. Conhecer quais os parâmetros fundamentais a determinar para efectuar a caracterização completa da qualidade da água nas albufeiras estudadas; e
8. Apresentar modelos para a resolução de problemas relacionados com a avaliação da qualidade da água de albufeiras.

I.6 ESTRUTURA E ORGANIZAÇÃO DO TRABALHO

Esta dissertação está organizada em seis capítulos, dos quais o primeiro é constituído pela presente introdução, que engloba o enquadramento do trabalho, apresenta os seus objectivos e descreve a sua estrutura.

No segundo capítulo é apresentado e caracterizado o local de trabalho. Neste capítulo faz-se ainda referência à legislação aplicada a cada uma das albufeiras em estudo, tendo em conta os usos a que se destinam.

No terceiro capítulo é abordada a problemática da Descoberta de Conhecimento em Base de Dados (DCBD), fazendo-se referência aos métodos e técnicas de *Data Mining* utilizados para a avaliação e previsão da qualidade da água numa albufeira destacando-se, entre outros aspectos, os conceitos, os objectivos e os tipos de abordagens.

Relativamente ao quarto capítulo apresenta-se, numa primeira parte, o procedimento experimental, nomeadamente as metodologias de trabalho no que concerne à monitorização da qualidade da água tendo em vista a caracterização da qualidade da água das albufeiras de Alvito e Odivelas. Na segunda parte descreve-se a construção da base de dados que irá servir de suporte aos modelos desenvolvidos.

No quinto capítulo fundamentam-se as opções tomadas durante o processo de DCBD e apresentam-se, discutem-se e avaliam-se quer os resultados, quer os modelos construídos.

O sexto e último capítulo apresenta as conclusões finais e perspectiva-se o trabalho que se pretende desenvolver no futuro.

CAPÍTULO 2 – CARACTERIZAÇÃO DO LOCAL DE TRABALHO

O presente trabalho tem por base o estudo de duas albufeiras, Alvito e Odivelas, que distam entre si cerca de 19 km. Estas albufeiras fazem parte da bacia hidrográfica do rio Sado e estão integradas na sub-bacia da ribeira de Odivelas. A bacia hidrográfica da albufeira do Alvito localiza-se nos distritos de Évora e Beja enquanto a bacia hidrográfica da albufeira de Odivelas se localiza apenas no distrito de Beja (Figura 4).

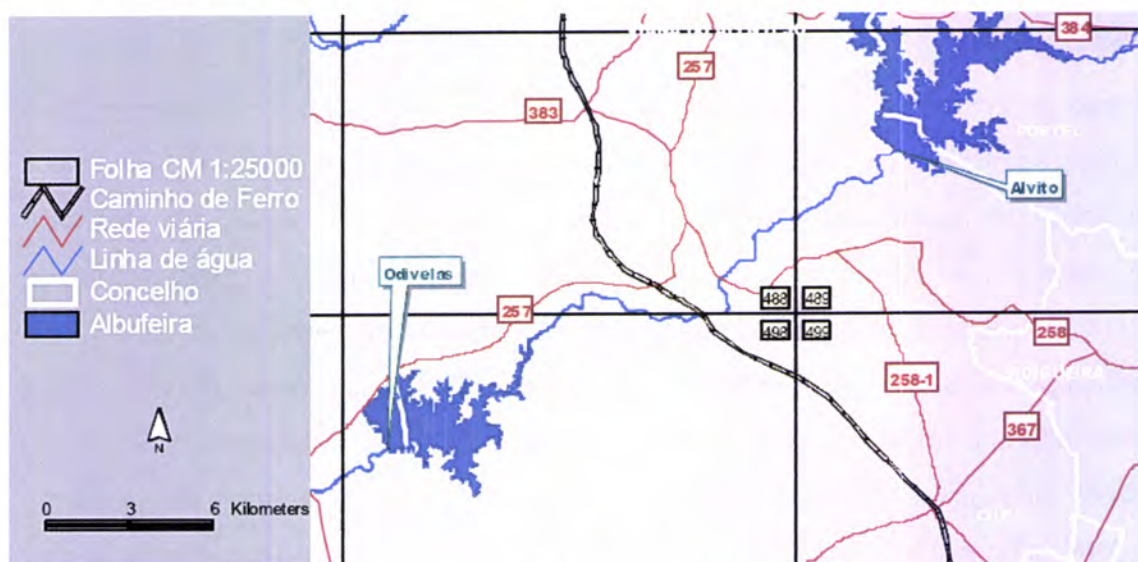


Figura 4 – Localização das albufeiras do Alvito e de Odivelas. Adaptada de Matoso [Matoso et al., 2004].

As albufeiras encontram-se localizadas na peneplanície do Baixo Alentejo, que constitui uma aplanção extensa que alterna, por vezes, com pequenas ondulações correspondentes a rugosidades residuais. A altitude varia entre os 250 metros na zona Norte e os 100 metros na zona a Sul, verificando-se uma gradual variação de cota [Fonseca et al., 1999b].

O clima desta zona apresenta forte feição mediterrânea traduzindo-se, nomeadamente, em precipitações relativamente baixas e concentradas no Inverno, temperaturas médias/altas, amplitudes térmicas elevadas, humidade relativamente baixa, nebulosidade baixa, insolação e radiações elevadas no Verão [Cunha et al., 1980]. A insolação média anual situa-se entre 2800 e 3000 horas de Sol, com as temperaturas e as precipitações

médias anuais a variarem, respectivamente, entre os 15 °C e os 17,5 °C e os 500 e 1000 mm [SNIRH, 2009].

A área de estudo insere-se num espaço de carácter fortemente rural, onde predominam os tradicionais sistemas agro-silvo-pastoris de montado de sobro e azinho ou olival, sobre rotação de cereais e pastagens [Matoso et al., 2004]. No que respeita à ecologia, regista-se a presença de habitats de importância comunitária, como o denominado Charco Temporário Mediterrâneo [Matoso et al., 2004].

Em termos socioeconómicos, a região é caracterizada por uma densidade populacional média de 20 hab/km², sendo o sector terciário o principal empregador da região [Silva, 2002].

A área de estudo insere-se na zona estrutural do maciço antigo, designada por Ossa Morena caracterizada, maioritariamente, por formações geológicas rochosas que à superfície se apresentam alteradas e cobertas, na zona mais a Sul, por depósitos sedimentares de idades mais recentes. Nesta zona identificam-se gneisses félsicos, derivados de metavulcanitos siliciosos com evidências de alto grau de metamorfismo, atribuídos ao topo do Proterozóico Superior e ao Paleozóico Médio [Piçarra, 2000], onde afloram estruturas em antifoma, preferencialmente orientadas NW-SE ou NNW-SSE [Carvalhosa, 1983; Carvalhosa, 1985]. Na estrutura de Viana do Alentejo – Alvito aflora um importante conjunto quartzo-feldspático, correspondente a uma intrusão subvulcânica de rochas alcalinas ricas em potássio [Fonseca, 1995; Gomes, 2000; Rosas, 2003].

2.1 CARACTERIZAÇÃO DAS ALBUFEIRAS EM ESTUDO

Tendo em consideração os Decretos Regulamentares n.º 2/88 de 20 de Janeiro e n.º 37/91 de 23 de Julho, actualmente revogados, as massas de água que se pretendem estudar eram classificadas como protegida (albufeira do Alvito) e de utilização limitada (albufeira de Odivelas) [Fonseca et al., 1999a]. De acordo com a referida legislação, era permitida a prática de actividades recreativas sem prejuízo dos fins para que foram planeadas, nomeadamente, a produção de água para abastecimento público (Alvito) e água para rega (Alvito e Odivelas). Posteriormente, com a entrada em vigor do

Decreto-Lei n.º 152/97, de 19 de Junho, também ele revogado, as albufeiras em estudo passaram a ser consideradas zonas sensíveis.

Mais recentemente o Decreto-Lei n.º 107/2009, de 15 de Maio veio reestruturar, reorganizar e, de certa forma, compactar toda a informação relativa à classificação de albufeiras de serviço público, estabelecendo três tipologias:

- a) Albufeiras de utilização protegida – Pertencem a esta classe as albufeiras que se destinam a abastecimento público ou se prevê que venham a ser utilizadas para esse fim. Para além destas estão incluídas, também, as albufeiras onde a conservação dos valores naturais exige um regime de protecção mais elevado, designadamente as que se encontram inseridas em áreas classificadas, tal como está definido na Lei da Água (Lei n.º 58/2005, de 29 de Dezembro);
- b) Albufeiras de utilização condicionada – São incluídas nesta classe as albufeiras que apresentem condicionamentos naturais que aconselham a imposição de restrições às actividades secundárias, nomeadamente as que tenham superfície reduzida, obstáculos submersos, margens declivosas, dificuldades de acesso ou quaisquer outras características que possam constituir risco para a sua utilização. Para além das albufeiras referidas anteriormente estão englobadas, também, as que se localizem em situação fronteiriça e aquelas que estejam sujeitas a variações significativas ou frequentes de nível, de potencial ecológico e de estado químico; e
- c) Albufeiras de utilização livre – Esta classe é constituída pelas albufeiras que não são susceptíveis de classificação nas tipologias descritas nas alíneas anteriores, apresentando outras vocações, designadamente turísticas e recreativas.

Os principais objectivos do referido diploma são a protecção e valorização dos recursos hídricos associados às albufeiras de serviço público, bem como da zona terrestre de protecção para a qual se manteve a largura de 500 m podendo, sempre que tal se justifique, ser estendida até à largura máxima de 1000 m.

De acordo com a legislação em vigor, as albufeiras do Alvito e de Odivelas são classificadas, respectivamente, como albufeira de utilização protegida e de albufeira de utilização livre (Portaria n.º 522/2009, de 15 de Maio).

2.1.1 ALBUFEIRA DO ALVITO

A albufeira do Alvito, Figura 5, localiza-se no concelho de Cuba, distrito de Beja, na ribeira de Odivelas, estando integrada na bacia hidrográfica do Rio Sado. Iniciou funcionamento no ano de 1977 destinando-se a rega e à produção de água para consumo humano, servindo cerca de 27000 habitantes dos concelhos de Alvito, Cuba, Vidigueira (pertencentes ao distrito de Beja), Portel e Viana do Alentejo (pertencentes ao distrito de Évora) [Silva, 2002].



Figura 5 – Vista parcial da albufeira do Alvito. Adaptado de Matoso [Matoso et al., 2004].

Quando o nível de pleno armazenamento é atingido (i.e. a cota de 197,5 m), a área inundada é de 1480 ha, apresentando uma altura de 48,5 m e um comprimento de coroamento de 1105 m. A sua capacidade máxima de armazenamento é $13,25 \times 10^7 \text{ m}^3$, sendo a sua capacidade útil de $13 \times 10^7 \text{ m}^3$ [Matoso et al., 2004].

A barragem do Alvito funciona como reservatório da albufeira de Odivelas libertando, quando necessário, caudais para a ribeira de Odivelas por intermédio de descargas de fundo.

A bacia hidrográfica ocupa 212 km^2 apresentando uma altitude média de 245 m e um declive médio de 4,6 %. Encontra-se localizada nos concelhos de Évora, Viana do Alentejo, Cuba, Portel, Alvito e Vidigueira (Figura 6).

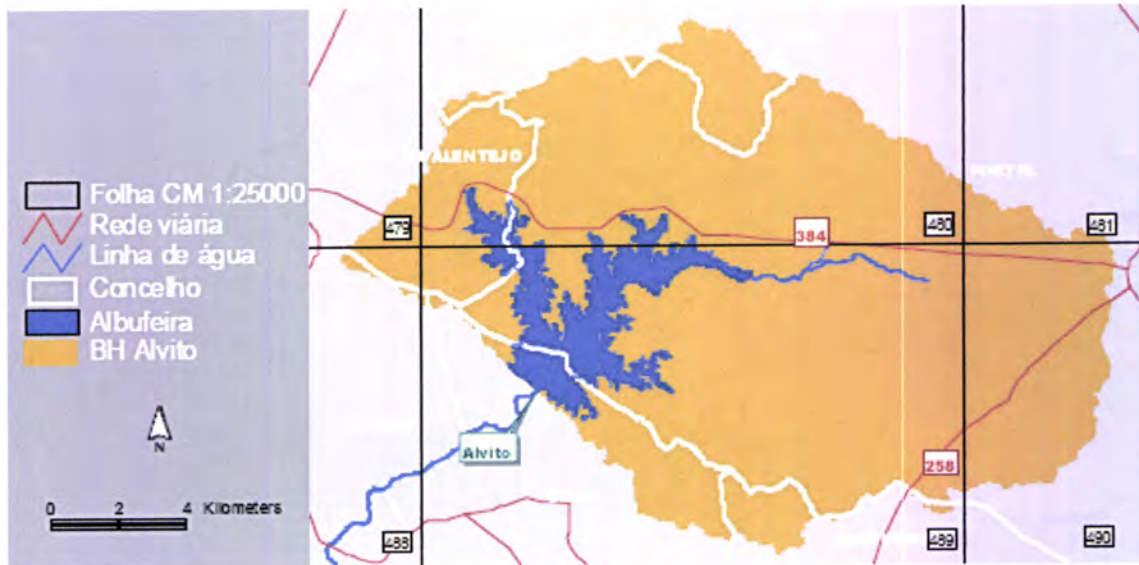


Figura 6 – Localização da bacia hidrográfica da albufeira do Alvito. Adaptada de Matoso [Matoso et al., 2004].

Os grandes grupos litológicos representados na área abrangida pela bacia hidrográfica encontram-se fortemente associados à unidade morfoestrutural sobre a qual a bacia assenta i.e. o Maciço Antigo. As formações litológicas predominantes na bacia hidrográfica da albufeira do Alvito são pertencentes à Era Paleozóica, nomeadamente ao Período do Câmbrio Inferior (calcários, dolomitos e rochas calco-silicatadas) e do Câmbrio-Ordovícico (rochas quartzo-feldspáticas orto-derivadas: gnaisses, leptinitos e metavulcânitos ácidos; micaxistos e rochas verdes), todas elas respeitantes às Séries Cristalofílicas Azóicas [Carvalhosa, 1971]. Os solos que predominam são, essencialmente, os Luvisolos (solos minerais, não hidromórficos, com argila em alta actividade no horizonte B) com pequenas ocorrências de alguns Cambissolos (solos pouco profundos, ainda em processo de desenvolvimento sem a presença de argila acumulada) [FAO, 2006].

Na área da bacia hidrográfica existe somente uma zona urbana e praticam-se, sobretudo, culturas arvenses de sequeiro e explorações agro-silvo-pastoris e florestais embora existam, também, culturas arbóreas-arbustivas e pequenas manchas de eucaliptais (Figura 7) [Matoso et al., 2004].

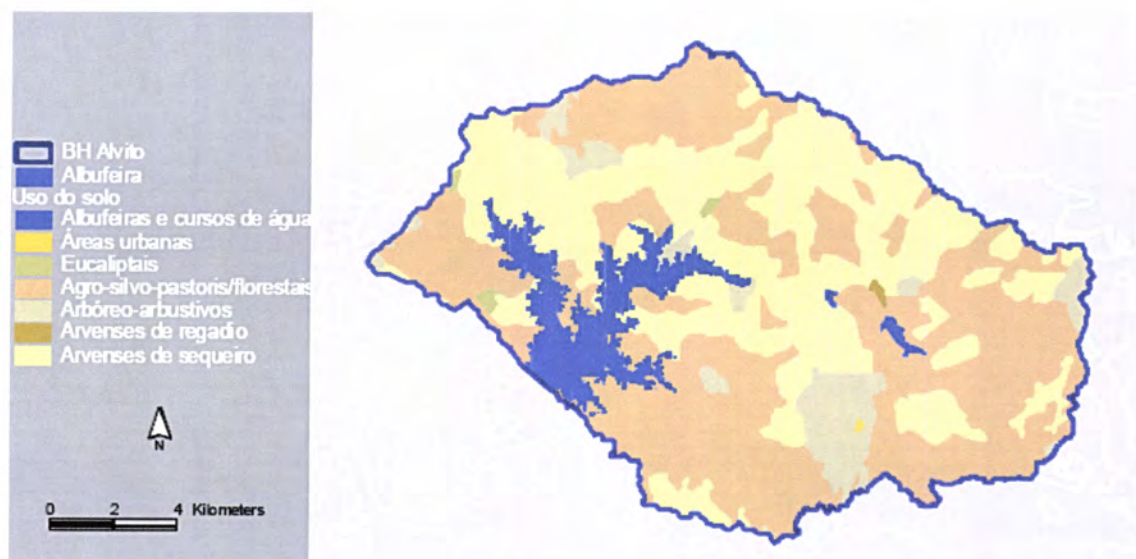


Figura 7 – Usos do solo na área da bacia hidrográfica da albufeira do Alvito. Adaptada de Matoso [Matoso et al., 2004].

As principais fontes de poluição pontual, Figura 8, resultam da suinicultura, da bovinicultura, das áreas urbanas, da queijaria e da fábrica de enchidos. O levantamento apresentado refere-se, exclusivamente, a fontes de poluição pontuais não incluindo, desta forma, as fontes de poluição difusa. As actividades pecuárias assinaladas dizem respeito apenas aos regimes intensivos e semi-intensivos. O regime extensivo, por seu lado, é considerado uma forma de poluição difusa não estando, portanto, incluído neste levantamento [DRAOTA, 1999].

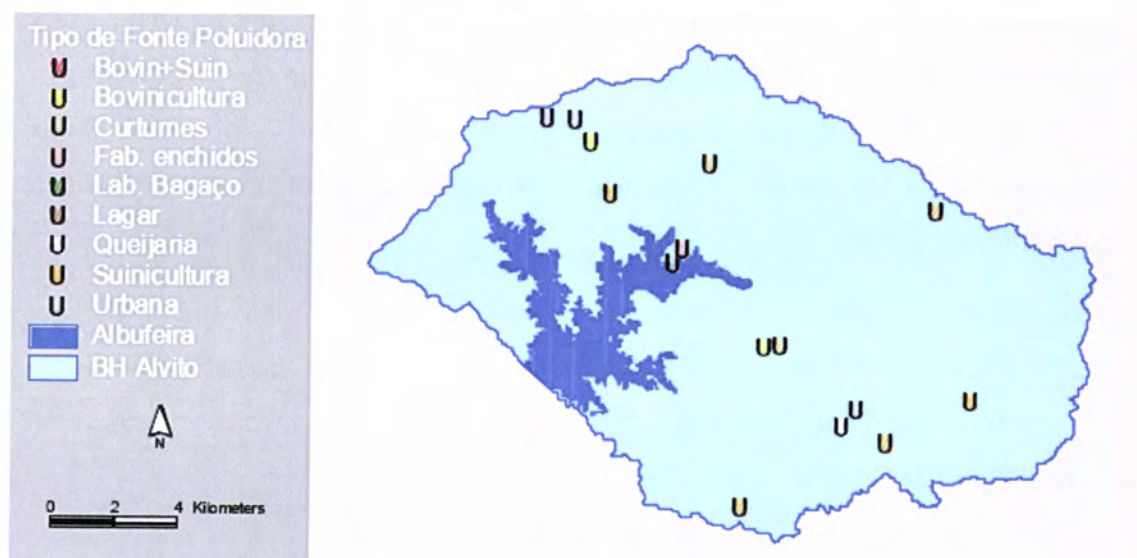


Figura 8 – Fontes de poluição pontual na bacia hidrográfica da albufeira do Alvito. Adaptada de Matoso [Matoso et al., 2004].

Das quinze fontes de poluição, identificadas na bacia hidrográfica da albufeira do Alvito, duas têm ligação à rede de esgotos urbanos (fábrica de enchidos e queijaria), cinco possuem sistemas de retenção (uma bovinicultura e quatro suiniculturas) e sete não estão providas de qualquer sistema de tratamento (quatro suiniculturas e três urbanas). A carga produzida pelas referidas fontes de poluição é de $556,4 \text{ Kg}_{\text{CBOS}} \text{ dia}^{-1}$, o que corresponde a 9273 equivalentes de população. A carga rejeitada, que entra nas linhas de água é, contudo, $215,1 \text{ Kg}_{\text{CBOS}} \text{ dia}^{-1}$, o que corresponde a 3585 equivalentes de população [DRAOTA, 1999].

2.1.2 ALBUFEIRA DE ODIVELAS

A albufeira de Odivelas, Figura 9, localiza-se no distrito de Beja, concelho de Ferreira do Alentejo, na localidade de Odivelas, estando integrada na bacia hidrográfica do rio Sado, sub-bacia da Ribeira de Odivelas. Iniciou funcionamento no ano de 1972 e destina-se à rega, servindo uma área de cerca de 13640 ha. Nesta albufeira evidencia-se, ainda, a presença de uma praia fluvial.

Quando o nível de pleno armazenamento é atingido (i.e. a cota de 103 m), a área inundada ocupa 973 ha, apresenta a altura de 55 m e um comprimento de coroamento de 544 m. A capacidade máxima de armazenamento é $9,6 \times 10^7 \text{ m}^3$, a que corresponde uma capacidade útil de $7 \times 10^7 \text{ m}^3$ [Matoso et al., 2004].



Figura 9 – Vista parcial da albufeira de Odivelas. Adaptado de Matoso [Matoso et al., 2004].

A bacia hidrográfica está localizada nos concelhos de Alvito, Ferreira do Alentejo, Cuba e Vidigueira, Figura 10, ocupando uma área de 430 km² e apresentando uma altitude média de 179 m com um declive médio de 3,8 %.

A Ribeira de Odivelas, curso de água responsável pela transição da água da albufeira do Alvito até à albufeira de Odivelas apresenta, entre as duas albufeiras, um declive médio de 0,7 % e um comprimento de 23,1 km.

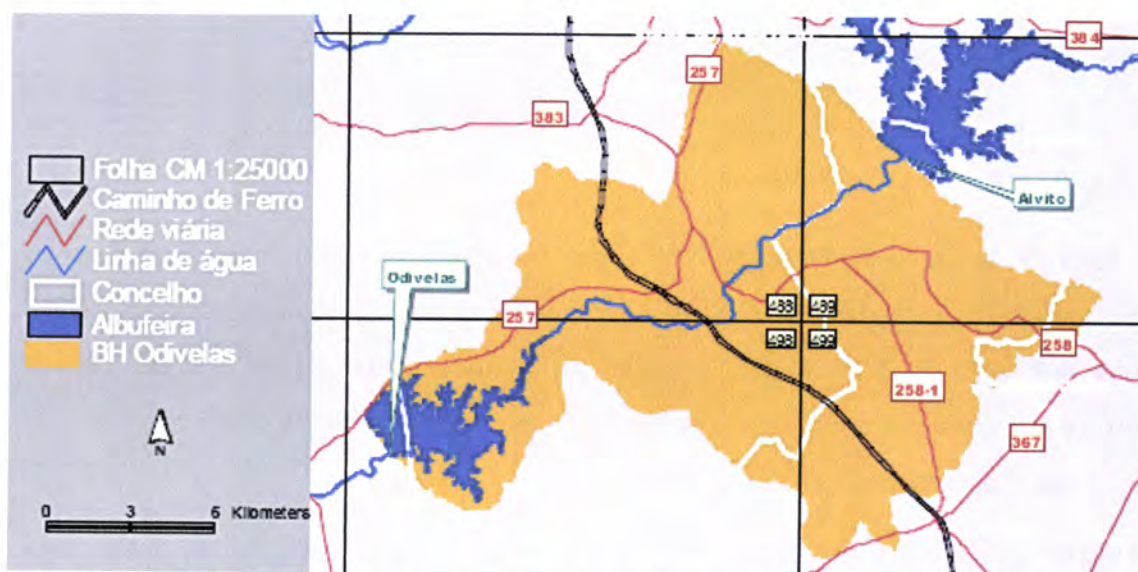


Figura 10 – Localização da bacia hidrográfica da albufeira de Odivelas. Adaptada de Matoso [Matoso et al., 2004].

As litologias predominantes na bacia hidrográfica da referida albufeira são as rochas eruptivas da Idade Hercínica (pórfiros da região de Beja, gabros e dioritos) e os depósitos de detritos grosseiros da Era Cenozóica, nomeadamente dos Períodos do Paleogénico e Neogénico indiferenciados [Carvalhosa, 1971]. Na área da bacia hidrográfica predominam os Luvissoles com pequenas alternâncias de Vertissolos (solos argilosos pesados com elevada percentagem de argilas expansivas) [FAO, 2006]. As culturas agro-silvo-pastoris e florestais, arvenses de sequeiro e arbóreas-arbustivas são as predominantes existindo, no entanto, uma pequena mancha de eucaliptos e uma outra de culturas arvenses de regadio. Podem, ainda, identificar-se duas áreas urbanas, localizadas a montante da albufeira (Figura 11).

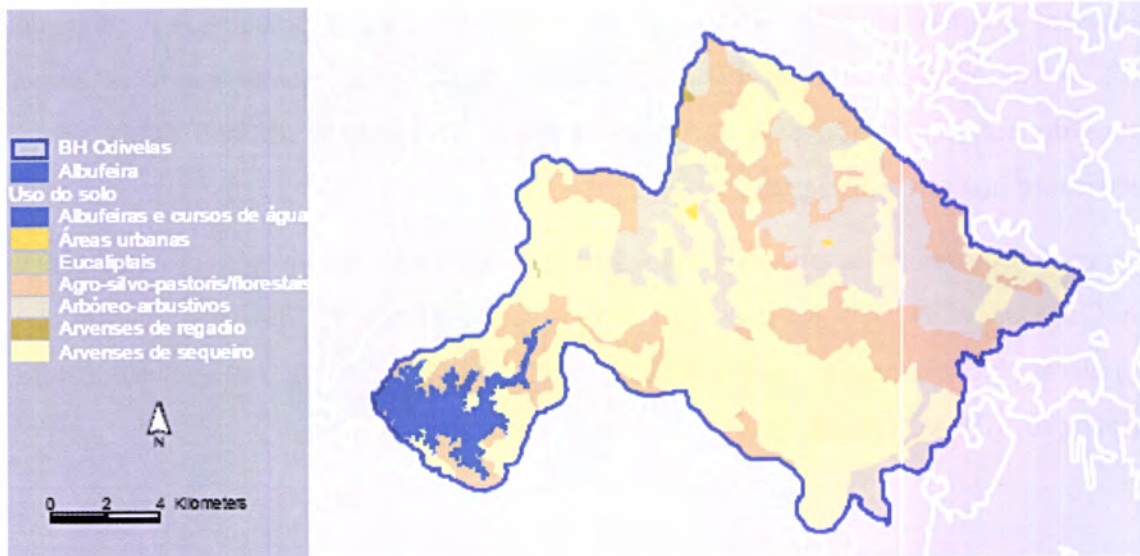


Figura 11 – Usos do solo na área da bacia hidrográfica da albufeira de Odivelas. Adaptada de Matoso [Matoso et al., 2004].

Na Figura 12 estão identificadas as fontes de poluição pontual situadas na bacia hidrográfica da albufeira de Odivelas, sendo cinco de origem urbana e apenas uma proveniente de uma suinicultura.

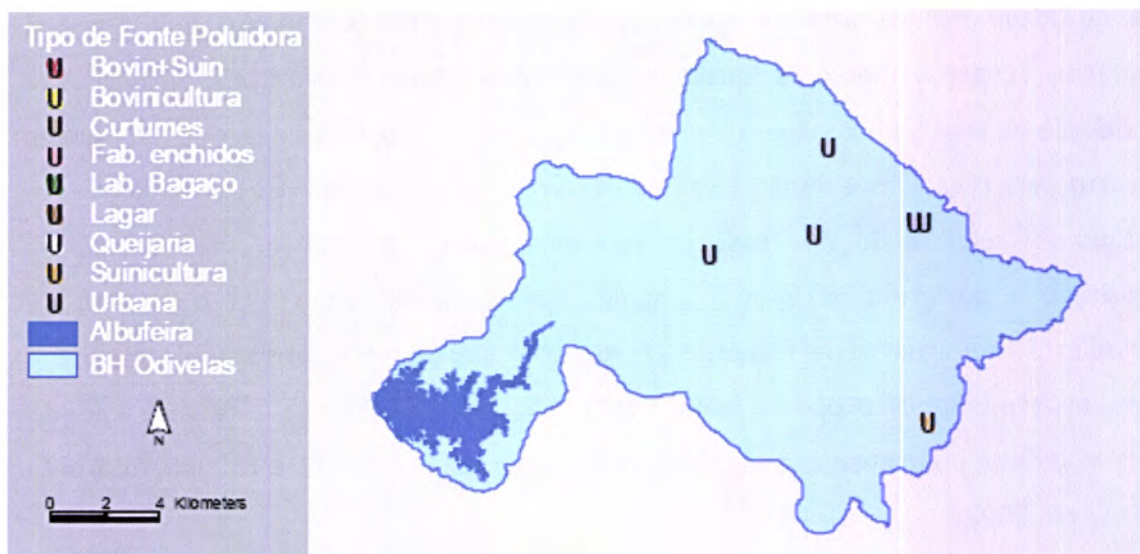


Figura 12 – Fontes de poluição pontual na bacia hidrográfica da albufeira de Odivelas. Adaptada de Matoso [Matoso et al., 2004].

As fontes de poluição identificadas possuem sistemas de tratamento de águas residuais que apesar de deficientes, fazem baixar as cargas lançadas nas linhas de água. As fontes de poluição urbana possuem sistemas de tratamento, nomeadamente, fossas sépticas,

biodiscos e lamas activadas sendo capazes de reduzir a carga poluente em cerca de 65%. A exploração suinícola possui um sistema de retenção, comportando-se como fonte de poluição difusa, uma vez que a matéria orgânica recolhida é usada como fertilizante nos terrenos agrícolas.

A carga produzida pelas fontes de poluição existentes na bacia hidrográfica da albufeira de Odivelas é de $357 \text{ Kg}_{\text{CBO5}} \text{ dia}^{-1}$, o que corresponde a 5950 equivalentes de população. No entanto, a carga rejeitada é de $42,6 \text{ Kg}_{\text{CBO5}} \text{ dia}^{-1}$ (i.e. 710 equivalentes de população) [DRAOTA, 1999].

2.2 CARACTERIZAÇÃO DA QUALIDADE DA ÁGUA DAS ALBUFEIRAS ESTUDADAS

O Decreto-Lei nº 236/98, de 1 de Agosto estabelece normas, critérios e objectivos de qualidade com a finalidade de proteger o meio aquático e melhorar a qualidade das águas em função dos seus principais usos. O anexo XXI define os objectivos ambientais de qualidade mínima para as águas superficiais enquanto os anexos I, XV e XVI definem, respectivamente, a qualidade das águas doces superficiais destinadas à produção da água para consumo humano, a qualidade das águas balneares e a qualidade da água para rega. Deste modo, a caracterização da qualidade da água das albufeiras em estudo foi feita tendo por base os referidos anexos. Para além do disposto na legislação a qualidade da água foi, ainda, avaliada tendo em conta o critério de classificação de águas superficiais de acordo com as suas características de qualidade para usos múltiplos, proposto pelo INAG. No que concerne à caracterização do estado trófico utilizou-se o critério definido pela OCDE em 1982 e reformulado pelo INAG em 2002.

2.2.1 AVALIAÇÃO DO ESTADO TRÓFICO

O enriquecimento excessivo em nutrientes e a consequente degradação dos ecossistemas aquáticos é um fenómeno cada vez mais comum nas águas superficiais. Na verdade, as actividades humanas afectam directamente e indirectamente os

equilíbrios dos ecossistemas aquáticos e deterioram a qualidade da água, limitando a sua utilização [Metcalf & Eddy, 2003].

A eutrofização pode ser definida como um aumento da quantidade de nutrientes disponíveis num ecossistema aquático, sobretudo azoto e fósforo, que provoca o crescimento acelerado (“*blooms*”) de algas verdes, de cianobactérias e de formas superiores de plantas aquáticas, perturbando o equilíbrio biológico e a qualidade das águas em causa [Ambiente, 1997; Scholten et al., 2005].

A eutrofização das águas superficiais é geralmente reconhecida como uma questão ambiental de preocupação mundial já que acelera o processo de envelhecimento dos corpos de água doce, deteriora a qualidade da água e inibe o desenvolvimento quer de plantas quer de animais que vivam nesses ecossistemas [Stewart & Howell, 2003]. Tendo em vista compreender as causas de eutrofização dos ecossistemas de água doce, é útil fazer uma distinção entre “eutrofia da água” (classificação da água de acordo com o seu estatuto intrínseco de nutrientes, i.e. eutrofização no sentido estrito) e “águas eutróficas” (percepção dos problemas de qualidade da água relacionados com o mau funcionamento ecológico, como água turva, mau cheiro e alta densidade algal, i.e. eutrofização no sentido lato) [Scholten et al., 2005].

A eutrofização, como foi referido anteriormente, pode ter causas naturais ou pode ser provocada pelas actividades antropogénicas [Esteves & Barbosa, 1986]. Quando as causas da eutrofização são naturais as alterações são lentas, permitindo que o ecossistema se adapte e se mantenha em equilíbrio. Deste modo, as águas exibem qualidade aceitável e a comunidade biológica permanece saudável e diversificada. Quando, pelo contrário, a eutrofização resulta de actividades humanas há um aceleração do processo (Figura 13). Os ciclos biogeoquímicos podem ser interrompidos e, muitas vezes, o sistema evolui para um estado de latência [Esteves, 1988; Wetzel, 1993].

A eutrofização artificial evolui muito rapidamente, comparativamente com a eutrofização natural devido, fundamentalmente, às entradas de nutrientes provenientes das actividades humanas tais como a utilização de fertilizantes na agricultura, as descargas de efluentes e a desflorestação. Todas estas actividades libertam para os ecossistemas aquáticos grandes quantidades de nutrientes que ficam disponíveis para o

crescimento de fitoplâncton¹. À medida que a produtividade do fitoplâncton aumenta a transparência da água e a penetração da luz diminuem e, conseqüentemente, a comunidade biótica é afectada. A capacidade de auto-depuração do sistema diminui e a acumulação de detritos e sedimentos aumenta [Lampert & Sommer, 2007]. Nestas condições, a vida de muitos organismos torna-se insustentável e a diminuição da qualidade da água reflecte-se nalguns parâmetros, tais como o oxigénio dissolvido e o pH. O quadro anteriormente descrito pode, nalguns casos, agravar-se devido à acumulação de toxinas, produzidas pelas cianobactérias, tendo conseqüências negativas ao nível da saúde pública [Scholten et al., 2005].

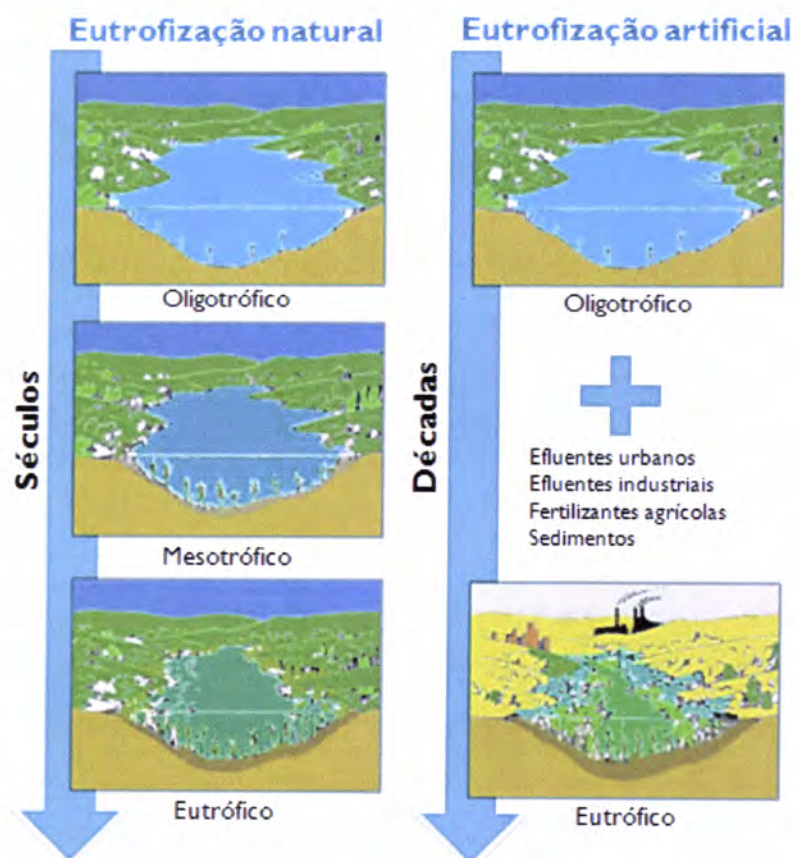


Figura 13 – Tipos de eutrofização (adaptado de laboratório de Limnologia e Planejamento Ambiental – UFES).

O conceito de estado trófico é multidimensional na medida em que envolve aspectos relacionados com o transporte de nutrientes, a relação da concentração de nutrientes

¹ Conjunto de algas microscópicas com pouco ou nenhum poder de locomoção, que se deslocam segundo o movimento da água, que inclui as algas verdes e as cianobactérias

no meio, a produtividade, a quantidade/qualidade da biota² e a morfologia da massa de água [Schwarzenbach et al., 2003; Mota, 2006]. Existem diversos critérios para avaliar o estado trófico das massas de água. Actualmente, em Portugal, o critério utilizado baseia-se no critério definido pela OCDE em 1982 e reformulado pelo INAG em 2002. São considerados três estados tróficos (Oligotrófico, Mesotrófico e Eutrófico) definidos pelas concentrações de fósforo total, clorofila-a e oxigénio dissolvido (Tabela 2), sendo atribuída à massa de água a classificação do parâmetro mais desfavorável.

Tabela 2 – Critério nacional para avaliação do estado trófico em albufeiras e lagoas (INAG 2002).

	Oligotrófico	Mesotrófico	Eutrófico
Fósforo total (µg/L)	<10	10 - 35	>35
Clorofila-a (µg/L)	<2,5	2,5 - 10	>10
Oxigénio dissolvido (% sat)	-	-	<40

Nota: os valores correspondem a médias geométricas

No que concerne à avaliação das albufeiras em estudo, foram determinadas as médias geométricas dos parâmetros fósforo total, clorofila-a e oxigénio dissolvido utilizando, para o efeito, os dados disponibilizados pelo INAG através do portal do SNIRH [SNIRH, 2009]. No estudo foram considerados os anos hidrológicos de 1999/2000 a 2007/2008 e 2001/2002 a 2007/2008, respectivamente para as albufeiras do Alvito e de Odivelas. A Tabela 3 mostra que, nos anos hidrológicos de 2000/2001 e 2003/2004, a albufeira do Alvito obteve a classificação mesotrófico tendo, em todos os outros, obtido a classificação eutrófico. A classificação deve-se, essencialmente, ao parâmetro fósforo total muito embora nos anos hidrológicos de 1999/2000 e 2001/2002 seja devida ao parâmetro clorofila-a. No que concerne à albufeira de Odivelas a classificação obtida foi eutrófico em todos os anos hidrológicos considerados. O parâmetro fósforo total foi o responsável pela classificação excepto no ano hidrológico de 2001/2002 onde a classificação foi devida parâmetro clorofila-a. O parâmetro oxigénio dissolvido não influenciou, nos anos hidrológicos considerados, a classificação obtida pelas albufeiras consideradas no estudo.

² Conjunto de seres vivos de um ecossistema

Tabela 3 – Avaliação do estado trófico das albufeiras do Alvito e de Odivelas.

Ano hidrológico	P total	Clorofila-a	OD	Estado Trófico
	(µg/L)	(µg/L)	(% Sat)	
Albufeira do Alvito				
1999/2000	30,8	18,7	74,3	Eutrófico
2000/2001	35	7,8	87,1	Mesotrófico
2001/2002	24,3	11,6	110,9	Eutrófico
2002/2003	58,7	16,8	99,7	Eutrófico
2003/2004	27,7	4,2	106,9	Mesotrófico
2004/2005	40,6	3	117,5	Eutrófico
2005/2006	69,3	4,7	101,5	Eutrófico
2006/2007	36,7	7,2	94,6	Eutrófico
2007/2008	59,3	6,4	87,7	Eutrófico
Albufeira de Odivelas				
2001/2002	34,4	10,9	76,4	Eutrófico
2002/2003	49,2	7,3	84,8	Eutrófico
2003/2004	41	1,9	80,6	Eutrófico
2004/2005	71,1	3,6	76,7	Eutrófico
2005/2006	80,5	7,5	78	Eutrófico
2006/2007	60,3	13,6	81,2	Eutrófico
2007/2008	99,2	6,8	82,2	Eutrófico

Nota: A cor do fundo reflecte a classificação do parâmetro. A cor verde, amarela e rosa denota, respectivamente, a classificação Oligotrófico, Mesotrófico e Eutrófico.

2.2.2 AVALIAÇÃO DA QUALIDADE DA ÁGUA DE ACORDO COM AS SUAS CARACTERÍSTICAS DE QUALIDADE PARA USOS MÚLTIPLOS

O critério de classificação de águas superficiais utilizado pelo INAG divide-as em cinco classes de acordo com as suas características de qualidade para usos múltiplos. Na Tabela 4 apresentam-se as várias classes, bem como o uso permitido para cada uma delas. De acordo com este critério a classificação é realizada parâmetro a parâmetro, sendo a classificação da massa de água dada pelo segundo parâmetro mais desfavorável, tendo como base a grelha paramétrica apresentada na Tabela 5 [SNIRH, 2009].

Na Figura 14 apresenta-se a classificação da albufeira do Alvito tendo em consideração o critério exposto anteriormente, ao longo dos últimos anos. A sua análise mostra que, excepto nos anos de 1999, 2000 e 2003, a água da referida albufeira obteve classificação C ou superior permitindo, deste modo, a sua utilização para os fins previstos, nomeadamente irrigação e produção de água para consumo humano.

Relativamente à albufeira de Odivelas apresenta-se, na Figura 15, a classificação obtida entre 2001 e 2008 verificando-se que, a água da albufeira de Odivelas obteve a classificação C permitindo, assim, a sua utilização para os fins previstos, nomeadamente irrigação. A única excepção ocorreu no ano de 2003 em que a classificação obtida foi D.

Tabela 4 – Usos permitidos por classes de qualidade da água.

Classe	Nível de qualidade
A – Sem poluição	Águas consideradas como isentas de poluição, aptas a satisfazer potencialmente as utilizações mais exigentes em termos de qualidade.
B – Fracamente poluída	Águas com qualidade ligeiramente inferior à classe A mas podendo, também, satisfazer potencialmente todas as utilizações.
C – Poluída	Águas com qualidade “aceitável”, suficiente para irrigação, uso industrial e produção de água potável após tratamento rigoroso. Permite a existência de vida piscícola (espécies menos exigentes) mas com reprodução aleatória; apta para recreio sem contacto directo.
D – Muito poluída	Águas com qualidade “mediocre”, apenas potencialmente aptas para irrigação, arrefecimento e navegação. A vida piscícola pode subsistir, mas de forma aleatória.
E – Extremamente poluída	Águas ultrapassando o valor máximo da Classe D, para um ou mais parâmetros. São consideradas como inadequadas para a maioria dos usos e podem ser uma ameaça para a saúde pública e ambiental.

Tabela 5 – Grelha para classificação de massas de água superficiais de acordo com as suas características de qualidade para usos múltiplos.

Parâmetro	Unidades	A		B		C		D		E
		Excelente		Boa		Razoável		Má		Muito má
		MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX	-
Arsénio	mg/L	-	0,01	-	0,05	-	-	-	0,1	>0,1
Azoto amoniacal	mg/L NH ₄ ⁺	-	0,5	-	1,5	-	2,5	-	4	>4
Azoto Kjeldahl	mg/L N	-	0,5	-	1	-	2	-	3	>3
Cádmio	mg/L	-	0,001	-	0,005	-	0,005	-	>0,005	
CBO ₅	mg/L O ₂	-	3	-	5	-	8	-	20	>20
CQO	mg/L O ₂	-	10	-	20	-	40	-	80	>80
Chumbo	mg/L	-	0,05	-	-	-	0,1	-	0,1	>0,1
Cianetos	mg/L	-	0,05	-	-	-	0,08	-	0,08	>0,08
Cobre	mg/L	-	0,05	-	0,2	-	0,5	-	1	>1
Coliformes fecais	ufc/100 mL	-	20	-	2000	-	20000	-	>20000	
Coliformes totais	ufc/100 mL	-	50	-	5000	-	50000	-	>50000	
Condutividade	µS/cm, 20°C	-	750	-	1000	-	1500	-	3000	>3000
Crómio	mg/L	-	0,05	-	-	-	0,08	-	0,08	>0,08
Estreptococos fecais	ufc/100 mL	-	20	-	2000	-	20000	-	>20000	
Fenois	mg/L C ₆ H ₅ OH	-	0,001	-	0,005	-	0,01	-	0,1	>0,1
Ferro	mg/L	-	0,5	-	1	-	1,5	-	2	>2
Fosfatos	mg/L P ₂ O ₅	-	0,4	-	0,54	-	0,94	-	1	>1
Fósforo	mg/L	-	0,2	-	0,25	-	0,4	-	0,5	>0,5
Manganês	mg/L	-	0,1	-	0,25	-	0,5	-	1	>1
Mercurio	mg/L	-	0,0005	-	-	-	0,001	-	0,001	>0,001
Nitratos	mg/L	-	5	-	25	-	50	-	80	>80
Oxidabilidade	mg/L	-	3	-	5	-	10	-	25	>25
OD	% saturação	90	-	70	-	50	-	30	-	<30
pH	Escala Sorensen	6,5	8,5	5,5	9	5	10	4,5	11	>11
Selénio	mg/L	-	0,01	-	-	-	0,05	-	0,05	>0,05
SST	mg/L	-	25	-	30	-	40	-	80	>80
Substâncias tensoativas	mg/L, sulfato de lauril e sódio	-	0,2	-	-	-	0,5	-	0,5	>0,5
Zinco	mg/L	-	0,3	-	1	-	3	-	5	>5

- O pH, sendo um parâmetro muito dependente de características geomorfológicas, pode apresentar valores fora deste intervalo sem, contudo, significar alterações de qualidade devidas a poluição.

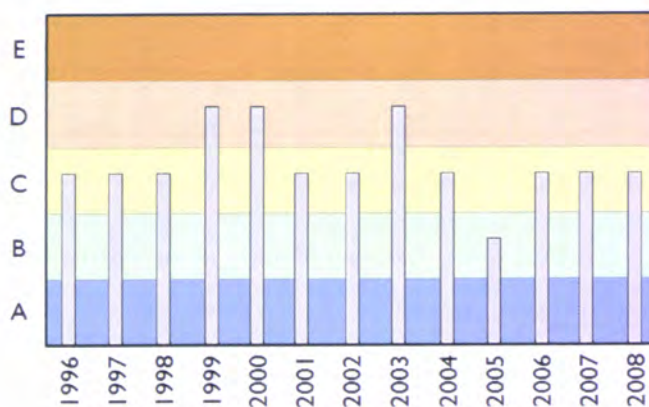


Figura 14 – Classificação da qualidade da água da albufeira do Alvito entre 1996 e 2008.

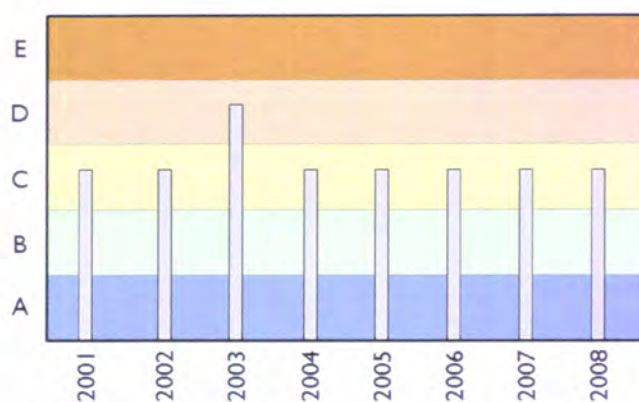


Figura 15 – Classificação da qualidade da água da albufeira do Odivelas entre 2001 e 2008.

2.2.3 QUALIDADE DAS ÁGUAS DOCES SUPERFICIAIS DESTINADAS À PRODUÇÃO DE ÁGUA PARA CONSUMO HUMANO

A qualidade da água de uma albufeira pode, também, ser avaliada de acordo com o critério de classificação das águas brutas destinadas à produção de água para consumo humano (Decreto-Lei nº 236/98, de 1 de Agosto). Este critério classifica as águas superficiais destinadas à produção de água para consumo humano em três classes (A1, A2 e A3), tendo em vista definir esquema de tratamento para as tornar aptas para consumo humano:

- ✓ Classe A1 – águas que necessitam de tratamento físico e desinfecção;
- ✓ Classe A2 – águas que carecem de tratamento físico, químico e desinfecção; e
- ✓ Classe A3 – águas que requerem tratamento físico, químico de afinação e desinfecção.

De acordo com esta metodologia, a classificação da massa de água é obtida parâmetro a parâmetro, tendo em consideração os valores ou intervalos de valores estipulados no anexo I do Decreto-Lei nº 236/98, de 1 de Agosto (Tabela 6), sendo atribuída à massa de água a classificação do parâmetro mais desfavorável.

Tendo em vista a verificação da conformidade, as amostras devem ser colhidas sistematicamente num mesmo ponto, tão próximo quanto possível do local de captação e em intervalos de tempo regulares. A frequência anual mínima de amostragem é função da classe a que pertencem as águas superficiais e do grupo (G1, G2 ou G3) em que o parâmetro se insere, Tabela 7 e Tabela 8.

Tabela 6 – Grelha para classificação de massas de água superficiais destinadas à produção de água para consumo humano (Anexo I do Decreto-Lei nº 236/98, de 1 de Agosto).

Parâmetro	Expressão dos resultados	A1		A2		A3	
		VMR	VMA	VMR	VMA	VMR	VMA
pH, 25°C	Escala de Sorensen	6,5 - 8,5	-	5,5 - 9,0	-	5,5 - 9,0	-
Cor (após filtração simples)	mg/L, escala Pt-CO	10	(O) 20	50	(O) 100	50	(O) 200
Sólidos suspensos totais	mg/L	25	-	-	-	-	-
Temperatura	°C	22	(O) 25	22	(O) 25	22	(O) 25
Condutividade	µS/cm, 20°C	1000	-	1000	-	1000	-
Cheiro	Factor diluição, a 25°C	3	-	10	-	20	-
Nitratos (*)	mg/L	25	(O) 50	-	(O) 50	-	(O) 50
Fluoretos (**)	mg/L	0,7 - 1,0	1,5	0,7 - 1,7	-	0,7 - 1,7	-
Ferro dissolvido (*)	mg/L	0,1	0,3	1	2	1	-
Manganês (*)	mg/L	0,05	-	0,1	-	1	-
Cobre	mg/L	0,02	(O) 0,05	0,05	-	1	-
Zinco	mg/L	0,5	3	1	5	1	5
Boro	mg/L	1	-	1	-	1	-
Arsénio	mg/L	0,01	0,05	-	0,05	0,05	0,1
Cádmio	mg/L	0,001	0,005	0,001	0,005	0,001	0,005
Crómio total	mg/L	-	0,05	-	0,05	-	0,05
Chumbo	mg/L	-	0,05	-	0,05	-	0,05
Selénio	mg/L	-	0,01	-	0,01	-	0,01
Mercúrio	mg/L	0,0005	0,001	0,0005	0,001	0,0005	0,001
Bário	mg/L	-	0,1	-	1	-	1
Cianetos	mg/L	-	0,05	-	0,05	-	0,05

Parâmetro	Expressão dos resultados	A1		A2		A3	
		VMR	VMA	VMR	VMA	VMR	VMA
Sulfatos	mg/L	150	250	150	(O) 250	150	(O) 250
Cloretos	mg/L	200	-	200	-	200	-
Substâncias tensoativas (que reagem com o azul-de-metileno)	mg/L, sulfato de lauril e sódio	0,2	-	0,2	-	0,5	-
Fosfatos (*) (***)	mg/L P ₂ O ₅	0,4	-	0,7	-	0,7	-
Fenóis	mg/L C ₆ H ₅ OH	-	0,001	0,001	0,005	0,01	0,1
Hidrocarbonetos dissolvidos ou emulsionados	mg/L	-	0,05	-	0,2	0,5	1
Hidrocarbonetos aromáticos polinucleares	µg/L	-	0,2	-	0,2	-	1
Pesticidas totais (paratião, hexaclorociclo-hexano, dieldrina e outros)	µg/L	-	1	-	2,5	-	5
CQO (*)	mg/L O ₂	-	-	-	-	30	-
OD (*) (***)	% saturação	70	-	50	-	30	-
CBO ₅ (*)	mg/L O ₂	3	-	5	-	7	-
Azoto Kjeldahl	mg/L N	1	-	2	-	3	-
Azoto amoniacal	mg/L NH ₄ ⁺	0,05	-	1	1,5	2	(O) 4
Substâncias extraíveis com clorofórmio	mg/L	0,1	-	0,2	-	0,5	-
Coliformes totais	ufc/100 mL	50	-	5000	-	50000	-
Coliformes fecais	ufc/100 mL	20	-	2000	-	20000	-
Estreptococos fecais	ufc/100 mL	20	-	1000	-	10000	-
Salmonelas	-	Ausência em 5000 mL	-	Ausência em 1000 mL	-	-	-

(O) Os limites podem ser excedidos em caso de condições geográficas ou meteorológicas excepcionais (n.º 1 do artigo 10.º).

(*) Os limites podem ser excedidos em lagos de pouca profundidade e baixa taxa de renovação.

(**) Os valores indicados constituem os limites das concentrações determinados em função da média anual das temperaturas máximas diárias.

(***) Este parâmetro é incluído para satisfazer as exigências ecológicas de certos meios.

(****) Refere-se a um valor mínimo recomendado.

Tabela 7 – Frequência mínima de amostragem de águas superficiais (Anexo IV do Decreto-Lei n.º 236/98, de 1 de Agosto).

Frequência mínima (número/ano)	Classe da água								
	A1			A2			A3		
	Grupo de parâmetros								
	G1	G2	G3	G1	G2	G3	G1	G2	G3
	4	2	1	8	4	2	12	6	3

A determinação dos parâmetros correspondentes ao grupo G2 implica a determinação dos parâmetros do grupo G1 e, do mesmo modo, a determinação do grupo G3 implica a determinação dos grupos G1 e G2.

Tabela 8 – Divisão dos parâmetros de qualidade de águas superficiais em grupos (Anexo V do Decreto-Lei nº 236/98, de 1 de Agosto).

G1	G2	G3
pH		Fluoretos
Cor		Boro
Sólidos Suspensos Totais		Arsénio
Temperatura	Ferro Dissolvido	Cádmio
Condutividade Eléctrica	Manganês	Crómio Total
Cheiro	Cobre	Chumbo
Nitratos	Zinco	Selénio
Cloretos	Sulfatos	Mercúrio
Fosfatos	Substâncias	Bário
Carência Química de Oxigénio	Tensoactivas	Cianetos
Oxigénio Dissolvido	Fenóis	Hidrocarbonetos Dissolvidos e
Carência Bioquímica de	Azoto Kjeldahl	Emulsionados
Oxigénio	Estreptococos fecais	Hidrocarbonetos Aromáticos
Azoto Amoniacal		Polinucleares
Coliformes totais		Pesticidas Totais
Coliformes fecais		Substâncias Extraíveis com Clorofórmio
		Salmonelas

A aplicação do critério exposto anteriormente à água da albufeira do Alvito durante os anos hidrológicos 1999/2000 a 2007/2008 classifica-a na classe A3. Trata-se, portanto, de uma água bastante poluída que obriga, para além do tratamento físico, a tratamento químico de afinação e a desinfecção. A classificação não sofre alteração no caso de se excluir o parâmetro temperatura. Na verdade, para além deste, os parâmetros CQO, pH, Coliformes Totais e CBO_5 foram os responsáveis pela atribuição da classe A3 à água da albufeira do Alvito.

2.2.4 QUALIDADE MÍNIMA PARA AS ÁGUAS DESTINADAS A REGA

As albufeiras do Alvito e de Odivelas têm, como um dos seus fins, o fornecimento de água para rega. Como tal, devem preencher os critérios estipulados no anexo XVI do Decreto-Lei nº 236/98, de 1 de Agosto (Tabela 9). Estes critérios visam proteger a saúde pública, a qualidade das águas superficiais, as culturas e os solos cuja aptidão para a agricultura pode ser degradada pelo uso de águas de rega de má qualidade.

As águas destinadas a rega são consideradas em conformidade com a norma de qualidade se, para a totalidade das amostras, os valores dos parâmetros respeitarem os valores nela fixados. Para as albufeiras em estudo apresentam-se, na Tabela 10, as situações em que foi ultrapassado o VMR.

Tabela 9 – Grelha para verificação da conformidade das águas superficiais destinadas a rega (Anexo XVI do Decreto-Lei nº 236/98, de 1 de Agosto).

Parâmetros	Unidades	VMR	VMA	Parâmetros	Unidades	VMR	VMA
Alumínio	mg/L	5	20	Lítio	mg/L	2,5	5,8
Arsénio	mg/L	0,1	10	Manganês	mg/L	0,2	10
Bário	mg/L	1	-	Molibdénio	mg/L	0,005	0,05
Berílio	mg/L	0,5	1	Níquel	mg/L	0,5	2
Boro	mg/L	0,3	3,75	Nitratos	mg/L	50	-
Cádmio	mg/L	0,01	0,05	Salinidade	mg/L	640	-
Chumbo	mg/L	5	20	Selénio	mg/L	0,02	0,05
Cloretos	mg/L	70	-	SST	mg/L	60	-
Cobalto	mg/L	0,05	10	Sulfatos	mg/L	575	-
Cobre	mg/L	0,2	5	Vanádio	mg/L	0,1	1
Crómio	mg/L	0,1	20	Zinco	mg/L	2	10
Estanho	mg/L	2	-	pH	Escala de Sorensen	6,5 - 8,4	4,5 - 9,0
Ferro	mg/L	5	-	Coliformes fecais	ufc/100 mL	100	-
Flúor	mg/L	1	15				

A análise da Tabela 10, relativamente à albufeira do Alvito, mostra que apenas em algumas das campanhas de amostragem foram superados os VMR para os parâmetros cloretos, manganês, molibdénio e coliformes fecais salientando-se, no entanto, o facto de não ter sido ultrapassado o VMA. No que respeita à albufeira de Odivelas é possível observar, para o parâmetro pH, a ocorrência de dois desvios pontuais pouco significativos relativamente ao VMA no decorrer do ano de 2001. Os parâmetros SST, molibdénio e coliformes fecais apresentaram, em algumas campanhas de amostragem, desvios ao VMR sem, no entanto, excederem o VMA. Tanto na albufeira do Alvito, como na de Odivelas a água armazenada cumpre as metas estipuladas pela legislação em vigor para águas destinadas a rega.

Tabela 10 – Parâmetros de qualidade da água que superaram o VMR fixado no anexo XVI do Decreto –
– Lei nº 236/98, de 1 de Agosto.

	Parâmetros	Unidade	VMR	VMA	Resultado	Data
ALBUFEIRA DO ALVITO	Cloretos	mg/L	70	-	92	04-11-2003
					214	30-11-2004
					132	07-03-2006
	Manganês	mg/L	0,2	10	0,28	06-05-2003
Molibdénio	mg/L	0,005	0,05	0,03	03-04-2001	
				0,03	07-05-2002	
Coliformes fecais	ufc/100 mL	100	-	220	02-10-2001	
				130	04-06-2002	
				220	06-10-2004	
				330	03-10-2005	
				170	02-11-2005	
				140	24-11-2005	
				110	30-10-2006	
				3300	28-11-2006	
ALBUFEIRA DE ODIVELAS	Molibdénio	mg/L	0,005	0,05	0,03	03-04-2001
	SST	mg/L	60	-	70	07-01-2003
					350	29-01-2008
	pH	Escala de Sorensen	6,5 - 8,4	4,5 - 9,0	9,06 9,09	05-06-2001 03-07-2001
Coliformes fecais	ufc/100 mL	100	-	400	06-03-2001	
				130	31-07-2001	
				180	06-11-2001	
				180	30-07-2002	
				130	04-02-2003	
				350	01-04-2003	
				570	09-09-2003	
				240	03-05-2005	
				140	02-05-2006	
				220	30-10-2006	
				240	02-09-2008	

2.2.5 QUALIDADE MÍNIMA PARA AS ÁGUAS BALNEARES

A Directiva 76/160/CEE, de 8 de Dezembro de 1975 e o Decreto-Lei nº 236/98, de 1 de Agosto definem o que se entende por águas balneares. Estão nesta situação todas as águas interiores (correntes ou paradas), águas de transição (estuarinas) e costeiras que sejam autorizadas para uso de banhos pelas entidades competentes e activamente promovidas a nível local, regional, nacional ou internacional ou, não sendo áreas proibidas, sejam regularmente utilizadas para banhos por um número considerável de banhistas locais e/ou visitantes.

A garantia de segurança da utilização de uma zona balnear passa, cada vez mais, pela qualidade da água. A qualidade das águas balneares é deveras importante, tanto como factor de preservação da saúde como, também, ao nível de qualidade ambiental e de desenvolvimento turístico.

Atendendo ao disposto na legislação comunitária que regulamenta a qualidade das águas para fins balneares (Directiva 76/160/CEE, de 8 de Dezembro de 1975) e à legislação nacional que transpõe esta a mesma directiva (Decreto-Lei n° 236/98, de 1 de Agosto) iniciou-se, em 1993 (ano de entrada em vigor do disposto na directiva), a monitorização das águas balneares nacionais.

A albufeira de Odivelas (Figura 16) foi palco de actividades balneares entre os anos 2003 e 2007, tendo sido alvo de monitorização. Nesse período vários parâmetros superaram o VMR fixado no anexo XV do Decreto-Lei n° 236/98, de 1 de Agosto. No entanto, só em quatro situações os VMA foram superados. O parâmetro pH superou, em três campanhas de amostragem relativas a 2005, o VMA tendo o mesmo sido verificado em relação ao parâmetro Coliformes fecais na campanha de 9 de Setembro de 2003. A partir de 2008 a monitorização foi suspensa uma vez que o Decreto-Lei n° 135/2009, de 3 de Junho, que revogou a Directiva 76/160/CEE, é bastante mais restritivo relativamente à gestão das bacias drenantes para zonas balneares. No caso da bacia hidrográfica da albufeira de Odivelas, a prática de agricultura intensiva e de actividades pecuárias não se afiguram compatíveis com o uso balnear.



Figura 16 – Praia fluvial localizada na albufeira de Odivelas. [SNIRH, 2009].

CAPÍTULO 3 – DESCOBERTA DE CONHECIMENTO EM BASE DE DADOS

Nas duas últimas décadas, no que concerne ao armazenamento de dados, tem-se assistido a um aumento significativo da quantidade de informação armazenada tanto em número de registos como em complexidade. Estima-se que a quantidade de informação duplique a cada 20 meses e que, num futuro próximo, cresça a um ritmo ainda mais elevado [Sousa, 2004].

O extraordinário desenvolvimento das comunicações aliado à imediata disponibilização da informação, bem como o despertar do sentido competitivo entre organizações de todo o mundo são alguns dos factores que justificam a acumulação de dados em bruto. Na verdade, as organizações partem do pressuposto que uma maior quantidade de dados gera melhor informação. No entanto, o facto dos dados, na maioria dos casos, não serem tratados em tempo útil ou a utilização de técnicas de tratamento inadequadas leva a que o resultado obtido nem sempre seja o desejado. Torna-se, pois, vital analisar os dados, extrair conhecimento, generalizar, sublinhar princípios e estruturar a informação. Não basta ter a informação disponível. A obtenção de mais-valias a este nível está fortemente correlacionada com a capacidade de seleccionar a informação certa para o problema específico [Perlovsky, 2001; Turban et al., 2004; Soman et al., 2006].

A análise de dados não é uma discussão recente. Na verdade, desde há muito que ficou claro que o cérebro humano analisa dados, trata informação e adquire conhecimento de modo diferente, através de processos de aprendizagem. Numa perspectiva sistémica a aprendizagem pode ser definida como as alterações do sistema que lhe permitem refazer as mesmas tarefas de uma forma mais eficaz e eficiente no futuro [Adriaans & Zantinge, 1996; Hand et al., 2001; Russell & Norvig, 2003; Costa & Simões, 2008]. A aprendizagem pode envolver duas fases distintas de processamento da informação – a forma indutiva e a forma dedutiva. Na primeira tem-se a aprendizagem baseada em casos, na qual os padrões e as regras são determinados a partir dos dados e das experiências. Na segunda forma, as regras e os padrões são utilizados para criar novos factos.

No que concerne à aprendizagem deve ser realçada a relação entre o modelo de aprendizagem e o seu ambiente. Neste âmbito existem, fundamentalmente, três paradigmas de aprendizagem [Patterson, 1996]:

- ✓ Aprendizagem supervisionada – a aprendizagem é levada a cabo a partir de um conjunto de casos de treino ou exemplos onde, cada um deles, é formado por um vector de entrada e um vector de saída ou de resposta. Deste modo, são fornecidas ao sistema as respostas correctas;
- ✓ Aprendizagem de reforço – nesta técnica não se apresenta a resposta correcta ao sistema. Apenas se dá uma indicação sobre se a resposta apresentada pelo sistema é correcta ou não, sendo esta informação utilizada no melhoramento do seu desempenho; e
- ✓ Aprendizagem não supervisionada – a aprendizagem é feita a partir de descobertas de características nos dados de entrada, procurando regularidades estatísticas ou agrupamentos de padrões nos exemplos de treino. Deste modo, ao sistema não chega qualquer indicação externa sobre a resposta correcta.

3.1 DESCOBERTA DE CONHECIMENTO EM BASE DE DADOS

O desenvolvimento tecnológico tem possibilitado um crescimento galopante no que respeita ao volume de dados armazenado. Como consequência deste aumento efectivo de informação, o seu processamento através dos métodos tradicionais tornou-se cada vez mais difícil e complexo [Michalski et al., 1998; Goebel & Gruenwald, 1999; Santos & Azevedo, 2005; Teixeira, 2005]. No sentido de fazer face a estes problemas emergiu, no início dos anos 90, a área da Descoberta de Conhecimento em Bases de Dados, começando a surgir aplicações vocacionadas para esta tarefa incorporando ferramentas de *Data Mining* [Witten & Frank, 2005; Bramer, 2007; Cios et al., 2007; Taniar, 2008]. Esta área, em progressiva expansão, estuda, investiga e desenvolve processos que possibilitam extrair conhecimento inovador, útil e válido incluído, de forma implícita, nos dados e representá-lo de forma acessível e clara para o utilizador [Apte et al., 2002; Sousa, 2004; Williams & Simoff, 2006].

O termo Descoberta de Conhecimento em Bases de Dados (DCBD) refere-se, portanto, a um processo que envolve a identificação e o reconhecimento de padrões ou tendências numa Base de Dados (BD) de forma autónoma e automática [Fayyad et al., 1996; Thuraisingham, 1999; Apte et al., 2002; Han & Kamber, 2006; Lin et al., 2008].

O processo de DCBD depende de uma nova geração de ferramentas e técnicas de análise de dados e, como se pode observar na Figura 17, envolve diversas etapas que se encontram descritas na Tabela II [Adriaans & Zantinge, 1996; Fayyad et al., 1996; Michalski et al., 1998; Santos & Azevedo, 2005; Teixeira, 2005; Larose, 2006; Soman et al., 2006; Bramer, 2007; Cios et al., 2007; Taniar, 2008].



Figura 17 – Processo de descoberta de conhecimento em bases de dados. Adaptado de Santos e Azevedo [Santos & Azevedo, 2005].

Tabela 11 – Etapas do processo de DCBD.

Etapa do processo de DCBD	Descrição
Seleção	Consiste no estudo/compreensão do domínio de aplicação e na selecção dos dados a analisar. Deste modo, procura-se limitar o espaço de pesquisa, direccionando o foco para subconjuntos de variáveis ou dados, onde será realizada a descoberta de conhecimento.
Pré-Processamento	Comporta a limpeza dos dados (data cleaning) de forma a adequar a BD aos algoritmos que, posteriormente, serão utilizados. Inclui, ainda, a integração de dados heterogéneos e a eliminação de dados incompletos/inconsistentes. São estabelecidas estratégias para eliminar ou minimizar os problemas relacionados com dados omissos/incompletos, com a presença de ruído e de erros. Quando necessário, pode ocorrer, a reconfiguração dos dados para assegurar uma BD uniforme e consistente.
Transformação	Visa armazenar os dados num formato que facilite a sua utilização por parte das ferramentas de <i>Data Mining</i> .
Data Mining	Constitui o coração do processo de DCBD, sendo muitas vezes confundida com ele. Envolve a selecção e utilização dos métodos e técnicas que melhor se enquadrem no cumprimento das metas inicialmente estabelecidas. A referida escolha depende, fundamentalmente, dos objectivos do processo de <i>Data Mining</i> .
Interpretação	Construção de uma representação do conhecimento que seja acessível e clara aos utilizadores. Os padrões identificados pelo sistema são interpretados como conhecimento, o qual pode ser usado para o suporte à tomada de decisão. Este conhecimento traduz-se, por exemplo, em entendimento dos conteúdos da BD ou na explicação de um fenómeno observado.

No início de qualquer projecto de Descoberta de Conhecimento em Bases de Dados é necessário delimitar o domínio sobre o qual se pretende executar o estudo, definir os objectivos a atingir, recolher os dados e/ou varáveis e construir a BD sobre a qual se irá actuar. Um outro aspecto importante a atender prende-se com o estudo do domínio de aplicação. O conhecimento do domínio é o fio condutor no desenvolvimento de qualquer projecto de DCBD [Gopalan & Silvaselvan, 2009].

Antes de se poder dar o processo de DCBD como concluído deve fazer-se a verificação da qualidade da informação obtida. Esta verificação pode ser vista como uma fase de avaliação do desempenho do mecanismo de aprendizagem, sendo determinante para a confiança que se possa depositar nos resultados obtidos.

A Figura 17, que esquematiza o processo de DCBD, pode sugerir a existência de um percurso linear, no entanto, isso nem sempre se verifica podendo, no decurso do projecto de DCBD, ser identificada a necessidade de retorno a fases anteriores. Em geral, nos projectos de *Data Mining* verifica-se um grande dispêndio de tempo nas etapas de preparação dos dados, cerca de 80%, enquanto os restantes 20% são atribuídos à etapa de *Data Mining* [Adriaans & Zantinge, 1996].

3.2 DATA MINING

Consoante o autor, a área de especialização e a abordagem assim se encontram diferentes definições para o termo *Data Mining*. Apresentam-se, de seguida, algumas das definições comumente aceites:

“Data Mining é o processo de encontrar padrões e relações em BD de grande dimensão, previamente desconhecidos e potencialmente interessantes” [Fayyad et al., 1996].

“O termo Data Mining é vulgarmente utilizado como um sinónimo do processo de extracção de informação útil das BD [...] O componente de Data Mining do processo de DCBD está relacionado com os meios algoritmos através dos quais os padrões são extraídos e os dados são enumerados” [Fayyad, 1997].

“Data Mining refere-se à aplicação de métodos de aprendizagem automática, bem como de outros métodos para a enumeração de padrões sobre dados” [Michalski et al., 1998].

“Data Mining é o processo de extrair informação ou conhecimento de conjuntos de dados para os propósitos de tomada de decisão” [Chang et al., 2001].

“O Data Mining procura padrões de interesse numa determinada forma de representação, ou conjunto de representações: classificação, árvores de decisão, regras de indução, regressão, segmentação, e assim por diante” [Sousa, 2004].

“Data Mining refere-se à utilização de uma variedade de técnicas para identificar “nuggets” de informação ou conhecimento em conjuntos de dados, extraindo-os de modo a poderem ser utilizados nas áreas de suporte à decisão, previsão e estimativa. Esses dados são, por norma, volumosos mas de pouca utilidade quando acedidos de forma directa” [Clementine, 2007].

“O objectivo do Data Mining é fazer com que grandes quantidades de dados obtenham sentido, em alguns domínios, na sua maioria sem supervisão” [Cios et al., 2007].

A análise destas definições revela características que são comuns a todas elas, as quais se podem sintetizar na seguinte definição: aplicação de métodos e técnicas em BD, tendo em vista encontrar tendências ou padrões com o intuito de descobrir conhecimento implícito nos dados.

Como foi referido anteriormente, a etapa de *Data Mining* (DM) é apenas uma das várias etapas do processo de DCBD que, de acordo com Berry e Linoff [Berry & Linoff, 2000], é sustentada por três pilares fundamentais, nomeadamente os modelos e técnicas, os dados e a modelação de dados. No que diz respeito aos modelos e técnicas não existe um modelo ou uma técnica adequada a um determinado problema, podendo-se utilizar vários. A escolha está, na maior parte dos casos, relacionada com os objectivos estabelecidos para o processo de DCBD no geral e, para a etapa de DM no particular. Na Tabela 12 apresentam-se e caracterizam-se alguns desses objectivos, divididos em duas tarefas gerais (previsão e descrição). Na tarefa de previsão tem-se por objectivo inferir comportamentos futuros com base nas experiências do passado, ao passo que a tarefa de descrição pretende-se, de uma forma geral, descrever de forma compacta conjuntos de dados ou associações entre variáveis.

Tabela 12 – Caracterização dos objectivos de *Data Mining*. Adaptado de Cortez [Cortez, 2004].

Objectivo		Descrição
Previsão	Classificação	Corresponde à descoberta de uma função que associe um caso a uma das várias classes de classificação possíveis, de forma a classificar um novo objecto de acordo com um padrão de classificação.
	Regressão	Consiste em prever valores futuros ou desconhecidos de uma variável dependente, a partir de exemplos. Corresponde à concepção de um modelo capaz de aprender uma função desconhecida que se aproxime da função dada por um conjunto de características genéricas.
Descrição	Associação ou Dependência	Procura-se encontrar modelos que descrevam dependências significativas entre variáveis, através da identificação de dados fortemente correlacionados. As associações surgem quando várias ocorrências estão ligadas num único evento, podendo surgir a nível estrutural ou a nível quantitativo.
	Sumariação	Pretende-se encontrar uma descrição compacta para subconjuntos de dados através de regras de resumo e de descobertas de relações entre as variáveis. Aplica-se à análise exploratória de dados e à geração de relatórios.
	Segmentação	Consiste em identificar um conjunto finito de categorias ou segmentos para descrever os dados. Os objectos pertencentes ao mesmo segmento são semelhantes entre si e diferentes dos objectos pertencentes a outros segmentos. Pode ser um objectivo intermédio, realizado numa fase inicial, para encontrar grupos homogéneos de dados para posterior aplicação de algoritmos de aprendizagem.
	Visualização	Visa apresentar os resultados através de uma forma visual com o recurso a gráficos ou diagramas. Pretende dar a conhecer informações complexas de uma forma acessível, evidenciando os padrões e as tendências encontradas.

3.3 MODELOS E TÉCNICAS DE *DATA MINING*

Um modelo é definido como a representação matemática de um sistema ou processo com o objectivo de o estudar. Corresponde à estruturação da representação do conhecimento, tendo em vista um determinado objectivo que se pretende alcançar

[Rud, 2001; Santos & Azevedo, 2005; Chakrabarti et al., 2009]. Um modelo pode ser considerado como uma função que atribui a cada exemplo, no domínio definido pelos atributos de entrada, um valor contido no domínio dos atributos de saída. A cada modelo está associado um conjunto de parâmetros que necessitam de ser ajustados ou estimados, a partir de um conjunto de dados, através de um algoritmo de aprendizagem. Após a fase de aprendizagem é possível utilizar o modelo, ou seja, extrapolar novas saídas por introdução de novas entradas [Cortez, 2004; Teixeira, 2005; Poncelet et al., 2008].

Não existe um modelo universal de DM que resolva, de forma eficiente, todos os problemas. Na verdade, a escolha de um determinado algoritmo pode ser considerada uma arte na medida em que é possível utilizar diferentes modelos para a mesma tarefa, com vantagens e desvantagens intrínsecas [Fayyad et al., 1996; Han & Kamber, 2006; Cios et al., 2007; Chakrabarti et al., 2009].

Os modelos e técnicas utilizadas na etapa de DM dependem da natureza do problema e dos objectivos que se pretendem atingir. Na Tabela 13 referem-se algumas das técnicas mais utilizadas na etapa de *Data Mining* consoante os objectivos que se pretendem atingir [Dzeroski, 2006].

Tabela 13 – Tarefas e técnicas usadas no processo de *Data Mining*.

Tarefa		Técnica				
		Árvores de Decisão	Indução de Regras	Redes Neurais Artificiais	Algoritmos Genéticos	Aproximação de Vizinhanças
Previsão	Classificação	x	x	x	x	
	Regressão	x	x	x	x	
Descrição	Associação ou Dependência		x	x		
	Sumariação	x			x	x
	Segmentação	x	x	x	x	x
	Visualização	x				

No presente trabalho, as técnicas utilizadas foram as Árvores de Decisão, a Indução de Regras e as Redes Neurais Artificiais na tarefa de classificação e a Aproximação de Vizinhanças na tarefa de segmentação. Na tentativa de gerar modelos explicativos dos segmentos encontrados, de forma a atribuir um novo caso a um dado segmento, foram novamente utilizadas as Árvores de Decisão e a Indução de Regras.

3.3.1 ÁRVORES DE DECISÃO

As árvores de decisão tiveram a sua origem na área da Aprendizagem Automática (*Automatic Interaction Detection*). Esta análise testa automaticamente todos os valores de um determinado atributo de forma a identificar aqueles que têm uma forte associação com os registos de saída.

A definição mais simples de árvores de decisão é a de uma forma de representação de um conjunto de regras que seguem uma hierarquia de classes ou valores. Expressa uma lógica condicional simples e, do ponto de vista gráfico, são semelhantes a uma árvore.

Uma árvore de decisão é uma representação de um conjunto de regras que permite classificar as instâncias quando percorrida desde o nó da raiz até aos nós terminais ou folhas. Cada nó da árvore especifica um teste para os atributos da instância (variáveis) e, cada ramo descendente desse nó, corresponde a um dos valores possíveis para esse atributo. Uma instância é classificada começando por testar o atributo especificado pelo nó raiz e depois seguindo o ramo correspondente ao valor do atributo. O primeiro nó da árvore de decisão mostra o factor mais correlacionado com os objectos de saída. Os restantes factores são subsequentemente classificados como nós e relacionados com os nós anteriores, possibilitando uma visualização fácil e rápida dos factores que mais se relacionam com os objectos de saída [Quinlan, 1987; Kufirin, 1997; Han & Kamber, 2006; Cios et al., 2007].

Os algoritmos de indução de árvores de decisão utilizam a aprendizagem supervisionada, i.e. são fornecidas ao sistema as respostas correctas, a partir de um conjunto de exemplos onde, cada um deles, é formado por um vector de entrada e um vector de saída ou resposta [Rich & Knight, 1991].

A construção das árvores de decisão é feita a partir dos dados de treino, de uma forma recursiva, subdividindo este conjunto de dados até à obtenção de nós “puros”,

em que cada nó represente apenas uma única classe, ou a satisfação de um critério de paragem [Rich & Knight, 1991]. As árvores geradas têm a seguinte estrutura [Michalski et al., 1998]:

- ✓ Folhas ou nós puros – corresponde às classes/objectos;
- ✓ Nós internos – especificam testes efectuados a um único atributo originando duas ou mais sub-árvores que representam as saídas possíveis;
- ✓ Ramos – correspondem aos possíveis valores dos atributos.

Na construção de árvores de decisão levantam-se dois problemas. O primeiro está relacionado com o atributo a seleccionar para teste num determinado nó e, o segundo, está ligado à paragem da divisão dos exemplos. No que respeita à escolha do atributo, mais informativo, existem várias medidas para avaliar a capacidade de um dado atributo para discriminar as classes. Contudo, todas convergem em dois pontos:

- ✓ Uma divisão que mantém as proporções de classes em todas as partições é inútil;
- ✓ Uma divisão onde em cada partição todos os exemplos são da mesma classe tem utilidade máxima [Sousa, 2004; Teixeira, 2005; Han & Kamber, 2006].

As medidas de partição dividem-se em três tipos:

- ✓ Medida da diferença dada por uma função baseada nas proporções das classes entre o nó corrente e os nós descendentes, valorizando a pureza das partições, e.g., índice de gini, entropia;
- ✓ Medida da diferença dada por uma função baseada nas proporções das classes entre os nós descendentes, que valoriza a disparidade entre as partições; e
- ✓ Medida de independência, que mede o grau de associação entre os atributos e a classe.

No que respeita à paragem da divisão dos exemplos, esta deve acontecer quando se verificar uma destas situações [Sousa, 2004; Teixeira, 2005; Han & Kamber, 2006]:

- ✓ Todos os exemplos pertencem à mesma classe;
- ✓ Todos os exemplos, embora pertencentes a classes diferentes, têm os mesmos valores dos atributos;
- ✓ O número de exemplos a dividir é inferior a um certo limite;

- ✓ O mérito de todos os possíveis testes de partição dos exemplos é muito baixo.

O algoritmo ID3, desenvolvido por Ross Quinlan [Quinlan, 1987], foi um dos algoritmos pioneiros que permitiu o desenvolvimento das árvores de decisão. Desde essa altura têm sido introduzidas melhorias e funcionalidades que resultaram no aparecimento de novas versões/evoluções deste algoritmo como, por exemplo, os algoritmos C4.5 e C5.0.

Uma das limitações exibidas pelo algoritmo ID3 prende-se com a possibilidade de ocorrer sobre-ajustamento, que se manifesta na construção de árvores com elevado desempenho nos dados de treino, mas com um baixo desempenho nos dados de teste.

No algoritmo C4.5 foram introduzidos melhoramentos que permitiram ultrapassar o problema de sobre-ajustamento através da introdução da poda da árvore. A poda da árvore consiste na redução de algumas sub-árvores, transformando-as em folhas, tendo por base a comparação entre os erros nesse nó e a soma dos erros nos nós que dele descendiam. Deste modo consegue-se obter árvores menores, mas com maior capacidade de previsão para novos casos [Quinlan, 1998; Han & Kamber, 2006].

O algoritmo C4.5 permite, ainda, ultrapassar problemas relacionados com valores numéricos, valores omissos, dados com ruído e possibilita a utilização de validação cruzada [Quinlan, 1993].

O algoritmo C5.0 é o sucessor mais recente do algoritmo C4.5, otimizado para lidar com as exigências do mundo real, através do aumento da eficácia ao nível do tempo de processamento e da memória utilizada. Uma das características mais importantes deste algoritmo consiste em gerar vários classificadores, a partir do mesmo conjunto de treino, permitindo depois combiná-los num único classificador final, no qual cada classificador inicial participa com um determinado peso que é ajustado durante a fase de treino. Esta metodologia, designada por “*Boosting*”, permite uma redução dos erros de classificação que, nalguns casos, pode atingir cerca de 40% [Quinlan, 1996; Schapire, 2002].

As árvores de decisão são um dos métodos mais simples e adequados para tratar problemas com muitas dimensões. Trata-se de um método preditivo não linear, rápido e que apresenta uma elevada expressividade na representação do conhecimento

[Quinlan, 1998; Berry & Linoff, 2000; Han & Kamber, 2006; Gopalan & Silvaselvan, 2009]. Na verdade, as árvores de decisão apresentam as seguintes vantagens:

- ✓ Método não-paramétrico – não assume nenhuma distribuição particular para os dados, podendo construir modelos para qualquer função desde que o número de exemplos de treino seja suficiente;
- ✓ Independência da escala das variáveis – transformações monótonas das variáveis não altera a estrutura da árvore de decisão;
- ✓ Facilidade de Interpretação – uma decisão complexa é decomposta numa sucessão de decisões elementares permitindo, de forma rápida, determinar quais os factores que mais influenciam a classificação;
- ✓ Adaptável a problemas de regressão – para além de se poderem aplicar a problemas de classificação (Árvores de Classificação) podem, também, ser utilizadas em problemas de regressão (Árvores de Regressão);
- ✓ Facilidade em lidar com diversos tipos de informação (real, nominal, ordinal);
- ✓ Facilidade em lidar com informação incompleta; e
- ✓ Escolha automática dos atributos mais relevantes.

As principais desvantagens das árvores de decisão prendem-se com o facto de necessitarem de uma grande quantidade de dados para descobrir estruturas complexas, de apresentarem alguma instabilidade, i.e. pequenas perturbações no conjunto de treino podem provocar grandes alterações no modelo e exibirem fronteiras lineares e perpendiculares aos eixos [Berry & Linoff, 2000; Maimon & Rokach, 2005; Taniar, 2008; Gopalan & Silvaselvan, 2009].

3.3.2 INDUÇÃO DE REGRAS

A indução de regras está relacionada com a descoberta de tendências e padrões em grupos de dados, ou seja, aplicação de regras sobre os dados que, posteriormente, são apresentadas sob a forma de lista [Fayyad et al., 1996; Quinlan, 1998; Berson et al., 2000]. Esta técnica consiste na descoberta de regras de previsão, do tipo *SE...ENTÃO*, onde o *SE*, i.e. a “condição” da regra especifica alguns atributos e o *ENTÃO*, i.e. a acção da regra, prevê um valor para um determinado atributo cuja previsão é desejada [Fayyad et al., 1996].

As regras devem, segundo Quinlan [Quinlan, 1993] e Kufirin [Kufirin, 1997], satisfazer três condições:

- ✓ Previsões correctas (se o SE da regra é verdadeiro, o ENTÃO da regra também deve ser verdadeiro);
- ✓ Compreensíveis (devem ser de fácil compreensão); e
- ✓ Úteis para a tomada de decisão (expressar conhecimento que é ou não esperado).

As regras têm dois graus associados, o suporte e a confiança. O suporte está relacionado com o número de casos no qual a regra é encontrada, enquanto a confiança é a probabilidade condicional da regra [Quinlan, 1998].

A indução de regras surge, muitas vezes, associada às árvores de decisão, na medida em que são utilizadas para expressar o conhecimento nelas representado. A regra mais importante corresponde ao primeiro nó da árvore e as seguintes correspondem aos nós subsequentes de acordo com a sua importância [Santos & Azevedo, 2005; Han & Kamber, 2006; Nisbet et al., 2009].

As principais vantagens da técnica de indução de regras prendem-se com a forma directa de lidar com os dados, o desempenho, a facilidade de compreensão, o poder expressivo e, por fim, o facto de serem facilmente inspeccionáveis. As desvantagens relacionam-se com a falta de profundidade e com a dificuldade no manuseamento de informação incompleta ou de valores fora do contexto [Langley & Simon, 1995; Hand et al., 2001; Lin et al., 2008].

3.3.3 REDES NEURONAIS ARTIFICIAIS

As Redes Neurais Artificiais (RNAs), também denominadas por sistemas conexionistas, são modelos simplificados do sistema nervoso central do ser humano. Trata-se de uma estrutura muito interconectada de unidades computacionais paralelas, designadas por neurónios artificiais ou nodos, que possuem capacidade de aprender [Cortez & Neves, 2000; Graupe, 2007]. As RNAs assemelham-se ao comportamento do cérebro em dois aspectos:

- ✓ O conhecimento é adquirido a partir de um ambiente, através de processos de aprendizagem;
- ✓ O conhecimento é armazenado nas conexões (também designadas por ligações) entre nodos.

A maior parte da investigação em RNAs teve, como foi referido, inspiração no sistema nervoso dos seres vivos, em particular o do ser humano. O sistema nervoso dos seres humanos recebe informações dos diferentes órgãos sensoriais e combina-a com a informação já armazenada (obtida por hereditariedade ou por mecanismos de aprendizagem) produzindo uma resposta.

O cérebro humano é uma estrutura extremamente complexa e maciçamente paralela. Os neurónios, apesar de serem cinco a seis vezes mais lentos que uma porta lógica de silício, superam essa lentidão com a utilização de uma estrutura maciçamente paralela [Wu & McLarty, 2000; Galushkin, 2007; Haykin, 2008].

Um neurónio (Figura 18) é constituído por um corpo celular de onde saem vários ramos chamados dendrites, e um ramo mais comprido, isolado, que pode estar ou não ramificado denominado axónio. As dendrites ligam-se a axónios de outros neurónios através de junções especializadas denominadas sinapses. A informação propaga-se entre os neurónios através de sinais electroquímicos, que fazem com que possa haver libertação de mensageiros químicos, os quais podem influenciar o potencial electroquímico do neurónio vizinho [Bose & Liang, 1996; Kartalopoulos, 1996; Arbib, 2003].

Do ponto de vista histórico as RNAs surgiram em 1943, com o trabalho de McCulloch e Pitts [McCulloch & Pitts, 1943]. Neste trabalho, os autores propuseram um modelo simplificado dos neurónios baseado no facto do neurónio, num dado instante, ou estar activo ou inactivo, o que corresponde ao verdadeiro/falso da lógica proposicional ou ao um/zero da álgebra booleana.

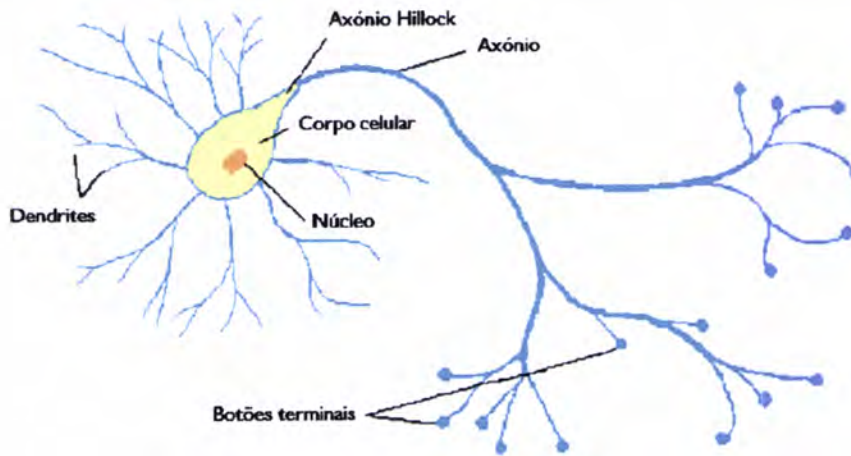


Figura 18 – Representação esquemática de um neurônio.

O referido trabalho despertou o interesse pelas RNAs e atraiu para a área muitos investigadores. Contudo, alguns resultados menos positivos e a falta de algoritmos de aprendizagem levou a que a abordagem conexionista ficasse em segundo plano na década de setenta. No início da década de oitenta, devido a diversas contribuições, houve um ressurgimento do interesse pelas RNAs. Das referidas contribuições destacam-se as redes neuronais recorrentes com conexões simétricas [Hopfield, 1982], as redes não supervisionadas do tipo *Kohonen* [Kohonen, 1982], o aparecimento do algoritmo da retro-propagação [Rumelhart et al., 1986] (que se tornou o algoritmo de treino mais utilizado em redes neuronais unidireccionais), o desenvolvimento de redes denominadas *Radial Basis-Functions* como alternativa às redes neuronais unidireccionais [Broomhead & Lowe, 1988] e o aparecimento de uma classe de redes supervisionadas denominadas Máquinas de Vectors de Suporte (*Support Vector Machines*) para regressão e reconhecimento de padrões [Boser et al., 1992].

Como foi anteriormente referido, o nodo, termo usado para distinguir um neurônio artificial de um neurônio natural, é a unidade de processamento base das RNAs. Existem diversos tipos de nodos, no entanto, todos eles se comportam como comparadores, produzindo uma saída, quando o efeito cumulativo das entradas excede um dado valor limite. Um nodo pode ser representado esquematicamente como se mostra na Figura 19, onde se evidencia a existência de três elementos fundamentais [Rojas, 1996; Cortez, 2002]:

- ✓ Um conjunto de conexões, cada uma com um peso associado (w_{ij}) que representa a força do sinal. No caso de o peso ser positivo tem-se um sinal de excitação e, no caso de ser negativo, tem-se um sinal de inibição. Por vezes existe um neurónio adicional, denominado “bias”, cujo valor é +1 que visa estabelecer as condições operacionais correctas para o nodo;
- ✓ Um integrador, que reduz os n argumentos de entrada, i.e. os estímulos, a um único valor. Nas RNAs utiliza-se a função adição (Σ), pesando todas as entradas numa combinação linear;
- ✓ Uma função de activação (fa), que limita a amplitude do sinal de saída podendo introduzir uma componente não linear.

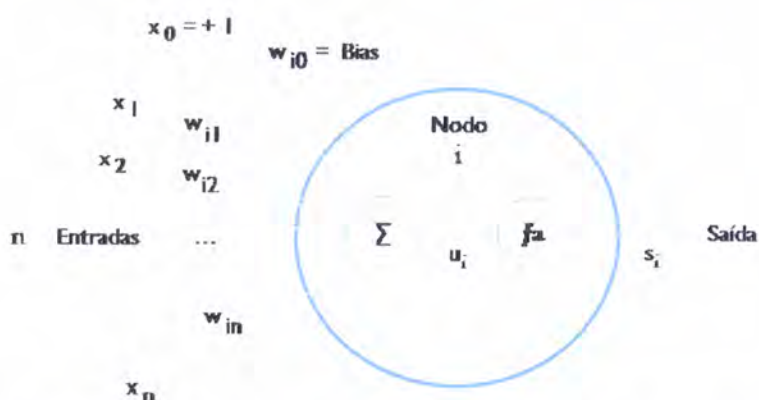


Figura 19 – Estrutura geral de um nodo. Adaptado de Rojas [Rojas, 1996].

Em termos formais um nodo i , com n entradas e uma saída, é descrito pelas seguintes equações:

$$u_i = \sum_{j=1}^n (x_j \times w_{i,j}) + bias \quad (1)$$

$$s_i = fa(u_i) \quad (2)$$

onde u_i denota o ganho do nodo e s_i a saída do nodo [Chester, 1993].

A Tabela 14 mostra algumas das funções de activação mais utilizadas enquanto na Figura 20 tem-se a representação gráfica de algumas dessas funções de activação [Chester, 1993; Hagal et al., 1996; Kartalopoulos, 1996; Graupe, 2007].

Tabela 14 – Algumas funções de activação utilizadas em RNAs.

Nome	Função	Contradomínio
Limiar	$\begin{cases} 1, & u_i \geq 0 \\ 0, & u_i < 0 \end{cases}$	$\{0, 1\}$
Linear	u_i	$] -\infty, +\infty[$
Logística	$\frac{1}{1 + e^{-ku_i}}$	$\{0, 1\}$
Tangente hiperbólica	$\tanh(k u_i)$	$\{-1, 1\}$
Gaussiana	$\frac{-u_i^2}{e^{2k^2}}$	$\{-1, 1\}$
Quadrada	$-\sinh(u_i) u_i^2$	$] -\infty, +\infty[$

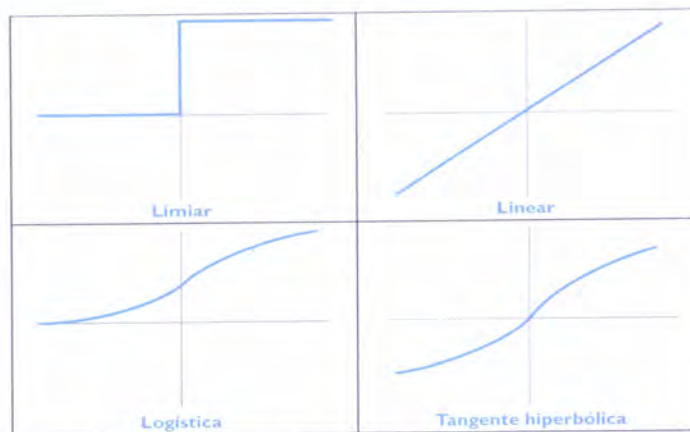


Figura 20 – Alguns exemplos de funções de activação utilizadas em RNAs.

A função limiar, também designada por função *heaviside*, é normalmente utilizada em nodos do tipo *McCulloch & Pitts* [McCulloch & Pitts, 1943], em que a saída toma o valor +1 apenas se o ganho for não-negativo, obedecendo a uma filosofia do tudo ou nada.

As funções de activação logística, tangente hiperbólica, gaussiana e quadrada introduzem uma componente de não-linearidade ao processo computacional. A função logística, também conhecida por função *sigmoid*, cuja forma é aproximada por um “s”, é a função de activação mais utilizada em RNAs. É uma função crescente que exhibe um balanceamento entre um comportamento linear e não linear [Jordan, 1995].

A forma como os nodos se interligam numa estrutura de rede é denominada de arquitectura ou topologia de rede. Existem inúmeros tipos de arquitectura de RNAs,

cada um com as suas próprias potencialidades. De uma forma geral, caem dentro de duas categorias principais [Haykin, 2008]:

- ✓ Redes Neurais Unidireccionais (RNUs) – são organizadas por camadas, as conexões propagam-se sempre numa só direcção. Na sua forma mais simples são constituídas por uma camada de entrada e por uma camada de saída, no entanto, podem possuir uma ou mais camadas internas, cujos nodos são designados por nodos intermédios (Figura 21a); e
- ✓ Redes Neurais Recorrentes (RNRs) – Está-se perante recorrência quando a saída de um elemento influencia, de algum modo, a entrada desse mesmo elemento criando, assim, um ou mais circuitos fechados (Figura 21b). Ao conter ciclos, as saídas não são função exclusivamente das conexões entre nodos, mas também de uma dimensão temporal. Está-se na presença de um procedimento recursivo, que obedece a uma condição de paragem, com a última iteração a ser dada como a saída para o nodo [Rojas, 1996].



Figura 21 – Arquitectura de uma RNU (a) e de uma RNR (b).

O poder computacional de uma RNA tem por base dois aspectos fundamentais. Por um lado, a topologia que premeia o paralelismo e, por outro, a sua capacidade de aprendizagem e generalização, i.e. consegue responder, de modo adequado, a novas situações com base nas experiências do passado [Groth, 2000; Nisbet et al., 2009]. As RNAs, porém, não conseguem dar resposta a todos os problemas. Não raras vezes necessitam de ser integradas em outros sistemas ou novos paradigmas, tendo em vista

a resolução de problemas [Haykin, 2008]. No entanto, apresentam características únicas, que não se encontram noutros mecanismos ou técnicas para a resolução de problemas [Azoff, 1994; Kartalopoulos, 1996; Patterson, 1996; Haykin, 2008]:

- ✓ Aprendizagem e generalização – conseguem descrever o todo a partir de algumas partes, constituindo-se como formas eficientes de aprendizagem e armazenamento de conhecimento;
- ✓ Processamento maciçamente paralelo – permite que tarefas complexas sejam realizadas num curto espaço de tempo;
- ✓ Não linearidade – consegue dar resposta a problemas reais cuja natureza é não linear;
- ✓ Adaptabilidade – pode adaptar a sua topologia de acordo com mudanças de ambiente;
- ✓ Robustez e degradação suave – permitem processar o ruído ou informação incompleta/contraditória de forma eficiente, assim como são capazes de manter o seu desempenho quando ocorre desactivação de algumas das suas conexões e/ou nodos;
- ✓ Flexibilidade – apresenta um grande domínio de aplicabilidade; e
- ✓ Usabilidade – não necessita de conhecimento explícito acerca da função a ser aprendida.

No que respeita ao treino de redes neuronais, o algoritmo de retro-propagação [Rumelhart et al., 1986] é apontado como o algoritmo de treino mais utilizado em RNUs [Govindaraju & Rao, 2000]. Trata-se de um algoritmo de referência, já que constitui um método eficiente de computação para o treino de RNUs, procurando o mínimo da função de erro no espaço de procura dos pesos, baseando-se em métodos de gradiente descendente. O objectivo do treino é encontrar o conjunto óptimo de valores de pesos que, numa situação ideal, originaria as saídas correctas para uma qualquer entrada.

O algoritmo de retro-propagação envolve essencialmente duas etapas. Na primeira, denominada “propagação frontal”, são fornecidos os vectores de entrada que se propagam entre frente, camada após camada e, com base nas equações (1) e (2), são calculadas as saídas. No final é calculado o erro entre os valores calculados e os

valores reais. Na segunda etapa, denominada “retro-propagação”, o erro calculado anteriormente é propagado para trás, desde a camada de saída até à camada de entrada sendo os pesos das conexões ajustados [Bose & Liang, 1996].

Antes de se iniciar o treino procede-se à escolha dos valores iniciais dos pesos. Em geral, estes devem ser gerados de uma forma aleatória, dentro do intervalo $[-0,7, 0,7]$ [Hastie et al., 2001]. Inicia-se então o treino da rede sendo os pesos iterativamente ajustados. O processo é dado como terminado pela aplicação de critérios de paragem (e.g. quando as alterações nos pesos e no erro forem insignificantes ou até que se atinja um erro mínimo). Uma iteração considera-se terminada quando todos os exemplos de treino tiverem sido considerados.

O algoritmo de retro-propagação é controlado por dois parâmetros, o coeficiente de momento (μ) e taxa de aprendizagem (η), cujos valores se situam no intervalo $[0, 1]$. O coeficiente de momento é utilizado na fase de actualização dos pesos e visa garantir que esta seja feita de um modo consistente. Este parâmetro funciona como uma memória, tendo um efeito estabilizador em situações oscilatórias. A taxa de aprendizagem, por seu lado, controla as mudanças nos pesos. Quando η toma valores pequenos, próximos de zero, as mudanças produzidas são pequenas e podem tornar a aprendizagem lenta. No caso de η ter valores próximos da unidade as variações são demasiado grandes e podem conduzir a instabilidade no treino [Han & Kamber, 2006; Haykin, 2008].

3.3.4 APROXIMAÇÃO DE VIZINHANÇAS

A aproximação de vizinhanças é uma técnica de segmentação baseada no princípio de que registos semelhantes estão próximos uns dos outros, quando analisados numa perspectiva espacial [Perlovsky, 2001; Han & Kamber, 2006; Bramer, 2007; Butenko et al., 2009]. A verificação da localização dos registos, interpretados como pontos no espaço, permite a identificação de regiões, denominadas classes ou segmentos, que definem características comuns para os registos aí incluídos. A segmentação pode ser hierárquica, i.e. as classes estão organizadas numa hierarquia (e.g. dendogramas). Quando, pelo contrário, não existe organização entre as classes a segmentação diz-se não hierárquica.

Neste tipo de abordagem é utilizada a aprendizagem não supervisionada uma vez que não é fornecido ao sistema qualquer indicação sobre as respostas correctas, sendo a avaliação feita pela utilidade dos agrupamentos formados.

As técnicas de segmentação não hierárquica, como a estratégia *k-means*, assentam na construção de partições dos objectos armazenados numa BD, num conjunto de k classes, sendo k um parâmetro de entrada. O processo de obtenção das classes é iniciado com centróides em posições aleatórias, as quais são optimizadas, iterativamente, através da movimentação dos centros. Inicialmente cada objecto é atribuído ao centróide mais próximo sendo, mais comumente, utilizada a distância Euclideana, definida por:

$$d(i, j) = \sqrt{(X_{i1} - X_{j1})^2 + (X_{i2} - X_{j2})^2 + \dots + (X_{in} - X_{jn})^2} \quad (3)$$

onde $i = (X_{i1}, X_{i2}, \dots, X_{in})$ denota o centróide do segmento i e $j = (X_{j1}, X_{j2}, \dots, X_{jn})$ denota um objecto de dimensão n , contido na BD. No passo seguinte o algoritmo calcula o ponto médio de cada segmento, repetindo-se este processo até não haver alterações ou o número de iterações exceda um determinado limite, previamente fixado. De um modo geral, é utilizado, como critério de paragem, a minimização da soma do quadrado dos erros, definida por:

$$E = \sum_{i=1}^k \sum_{p \in C_i} |p - m_i|^2 \quad (4)$$

onde E denota a soma do quadrado dos erros, p denota um objecto da BD e m_i denota o centróide do segmento C_i [Bradley & Fayyad, 1998].

Cada classe é representada através do seu centro de gravidade, i.e. pela localização definida pela média de todos os membros do segmento [Teixeira, 2005; Han & Kamber, 2006].

A principal vantagem da estratégia *k-means* reside na simplicidade e na velocidade de processamento, permitindo ser utilizada em BD de grande dimensionalidade. As principais desvantagens prendem-se com o facto de a solução depender da escolha inicial das posições dos centros e utilizar o conceito de média que, em algumas situações, pode não existir. Nesses casos pode utilizar-se a variante *k-medoid*. Nesta

técnica o centro de gravidade é atribuído ao objecto da classe que está mais próximo do centro [Teixeira, 2005; Han & Kamber, 2006; Chakrabarti et al., 2009; Gopalan & Silvaselvan, 2009].

3.3.5 AVALIAÇÃO DOS MODELOS

Após a construção dos modelos é necessário avaliar o seu desempenho. Existem diversos métodos para estimar a capacidade de generalização de um modelo, como sejam a estatística simples, a validação com divisão da amostra, a validação cruzada e o *Bootstrapping* [Teixeira, 2005; Han & Kamber, 2006; Gopalan & Silvaselvan, 2009].

Um dos métodos mais utilizados para estimar os erros de generalização é a validação com divisão da amostra, que se baseia na divisão dos dados do problema em dois subconjuntos, mutuamente exclusivos, um para gerar o modelo, designado como conjunto de treino, e outro para a sua validação, denominado conjunto de teste. A cardinalidade dos referidos subconjuntos é variável e depende, entre outros factores, da dimensão da base de dados. Este método apresenta como vantagens a sua simplicidade e rapidez e, como desvantagem, a redução efectiva dos dados disponíveis para o treino [Hand et al., 2001; Larose, 2006; Williams & Simoff, 2006; Bramer, 2007].

A validação cruzada é uma variante método de validação com divisão de amostra. Este método permite a utilização de todos os casos disponíveis tanto para construir o modelo como para o validar [Han & Kamber, 2006; Larose, 2006; Williams & Simoff, 2006].

Na validação cruzada k -desdobrável, os dados (P) são divididos em k subconjuntos mutuamente exclusivos (P_1, P_2, \dots, P_k) de cardinalidade semelhante. Os modelos são treinados e testados k vezes (Figura 22). O erro final da generalização é dado pela média dos erros de validação, obtidos durante os k testes. Os valores de k podem variar, embora o valor dez seja o mais utilizado. A validação cruzada é notavelmente superior à validação com divisão da amostra para BD com um número reduzido de registos. A principal desvantagem prende-se com o facto de exigir um grande esforço computacional [Cortez, 2002; Han & Kamber, 2006; Bramer, 2007].

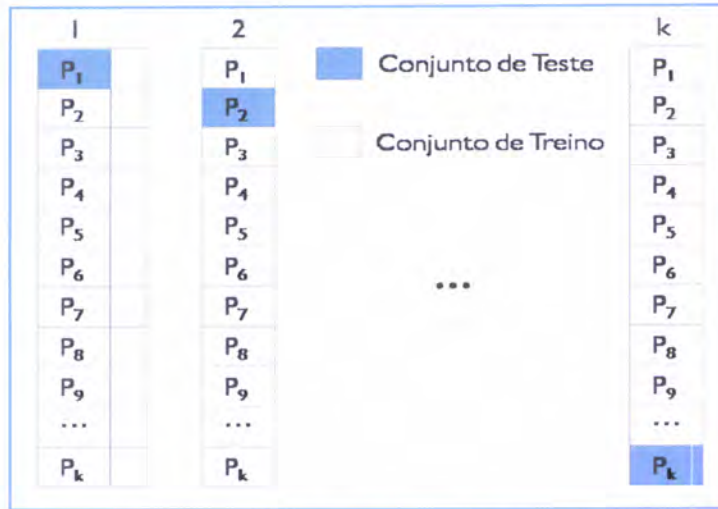


Figura 22 – Validação cruzada com k iterações.

3.3.5.1 AVALIAÇÃO DE MODELOS DE CLASSIFICAÇÃO

Quando se trata de problemas de classificação uma das técnicas de avaliação mais utilizadas é a Matriz de Coincidências [Kohavi & Provost, 1998]. Trata-se de uma matriz $L \times L$ (onde L denota o número de classes possíveis), na qual são mapeados os valores previstos pelo modelo e os valores desejados, sob a forma de tabela de duas entradas. As entradas horizontais referem-se às saídas desejadas, enquanto as saídas devolvidas pelo modelo se encontram nas entradas verticais. As células são preenchidas com os números de instâncias que correspondem ao cruzamento das entradas.

Na Tabela 15 apresenta-se um exemplo de uma matriz de coincidências onde é possível visualizar, de um modo rápido e inequívoco, o desempenho do modelo. Neste caso, relativamente à classe B, o modelo classificou correctamente 46 instâncias e incorrectamente 4. No que concerne à classe A todas as instâncias foram correctamente classificadas.

Tabela 15 – Matriz de coincidências 2x2.

Classe	A	B
A	50	0
B	4	46

3.3.5.2 AVALIAÇÃO DE MODELOS DE REGRESSÃO

Um modelo de regressão tenta estimar o valor associado a cada um dos exemplos. O principal objectivo de um método de regressão é conceber o “melhor” modelo na perspectiva da minimização de uma medida do erro global. Define-se por erro ou resíduo a diferença entre o valor real e o valor previsto, podendo ser calculado um erro global, respeitante a todos os valores previstos, usando as seguintes medidas [Witten & Frank, 2005]:

Erro Absoluto Médio (EAM): $EAM = \frac{\sum_{i=1}^N |e_i|}{N}$

Soma do Quadrado dos Erros (SQE): $SQE = \sum_{i=1}^N e_i^2$

Média do Quadrado dos Erros (MQE): $MQE = \frac{SQE}{N}$

Raiz da Média do Quadrado dos Erros (RMQE): $RMQE = \sqrt{MQE}$

onde $e_i = Y_i - Y'_i$ denota o erro ou resíduo, denotando Y_i e Y'_i , respectivamente o valor real e o valor previsto pelo modelo.

O desempenho de um modelo de regressão será tanto melhor quanto menor for o erro global. As medidas EAM e MQE são duas das métricas mais utilizadas no âmbito da avaliação dos modelos de regressão. Ambas têm as suas origens na estatística, tendo como intuito definir quão melhor é um estimador para um conjunto de valores. Segundo Torgo [Torgo, 1999] estas medidas, quando aplicadas à avaliação de modelos de regressão, servem diferentes propósitos. No caso de ser aceitável cometer alguns erros extremos, desde que o modelo se aproxime, na maior parte das vezes, do valor real, a minimização do EAM é a métrica mais adequada. Quando, pelo contrário, se está perante situações em que é crucial não cometer erros extremos, a MQE deve ser minimizada uma vez que esta métrica, ao utilizar o quadrado do erro amplifica os erros extremos.

CAPÍTULO 4 – PROCEDIMENTO EXPERIMENTAL

A componente experimental deste trabalho divide-se em duas partes. A primeira apresenta um carácter analítico, sendo constituída pela recolha de amostras de água bruta das albufeiras do Alvito e de Odivelas e pela determinação, no laboratório, de alguns parâmetros de qualidade da água. A segunda parte refere-se à construção de uma base de dados de qualidade da água para as albufeiras em causa sendo, posteriormente, utilizada para o desenvolvimento de modelos baseados no processo DCBD. Para o efeito, foram utilizados dados de qualidade da água, hidrométricos e meteorológicos disponibilizados pelo INAG através do portal do SNIRH [SNIRH, 2009]. Na construção da base de dados foram recolhidos os dados referentes ao período compreendido entre Junho de 2001 e Dezembro de 2008.

4.1 CONTROLO DA QUALIDADE DA ÁGUA

O controlo da qualidade da água pode definir-se como o conjunto sistemático de acções de avaliação, realizadas com carácter regular, tendo em vista a manutenção permanente da sua qualidade em conformidade com a norma ou padrão vigente [Alley, 2007].

A qualidade da água dos meios hídricos pode ser afectada por diversos factores quer de origem natural, quer relacionados com a actividade humana. Os equilíbrios naturais que se estabelecem no meio hídrico sofrem frequentemente alterações devidas, principalmente, a pressões externas. Os factores naturais de maior influência na qualidade e na quantidade da água de uma bacia hidrográfica são geológicos, hidrológicos e climáticos [Chin, 2006].

As alterações das características da água decorrentes de actividades antrópicas são, essencialmente, devidas ao lançamento, nos corpos de água, de esgoto bruto ou tratado ineficientemente, à infiltração de materiais poluentes, à utilização de pesticidas na agricultura, à desflorestação e à prática de actividades interditas à área [Tebbutt, 1998; Lehr & Keeley, 2005; Lehr et al., 2005].

Quando substâncias ou impurezas, de origem orgânica ou inorgânica, estão presentes numa água conferem-lhe determinadas propriedades e características que importa conhecer [Bortman et al., 2002]. As características de uma água são avaliadas através de ensaios físico-químicos, biológicos e microbiológicos que, normalmente, estão associados à implementação de um programa de monitorização ambiental. Os resultados analíticos que daí advêm permitem inferir tendências sobre a qualidade da água e perspectivar formas de tratamento ou melhoramento da mesma.

O controlo da qualidade da água fornece um diagnóstico sobre as condições de qualidade da água permitindo, muitas vezes, a detecção de possíveis origens de poluição. A poluição evidenciada em muitas massas de água pode ter uma origem difusa ou pontual. Os resíduos domésticos e industriais estão englobados no grupo das fontes pontuais por se restringirem a pontos de lançamento, o que facilita a colecta e o encaminhamento, tendo em vista um tratamento adequado para posterior lançamento no meio receptor [Harrison, 2001; Miranda & Bernard, 2009].

As fontes de poluição difusas caracterizam-se por apresentarem múltiplos pontos de descarga resultantes do escoamento, em áreas urbanas e/ou agrícolas, ocorrendo durante períodos de chuva, atingindo os poluente concentrações bastante elevadas. A poluição difusa é, muitas vezes, considerada como o principal factor impeditivo do alcance de níveis ambientalmente adequados de qualidade da água. A redução destas fontes requer, geralmente, mudanças nas práticas de uso do solo e programas de educação ambiental eficazes [Peirce et al., 1998; Novotny, 2003; Miranda & Bernard, 2009].

Uma água pode tornar-se inaceitável para determinados usos, muito embora possa ainda satisfazer outros [Conesa-Fernandez, 1997]. Existe, pois, um padrão de qualidade da água consoante a especificidade do uso a que esta se destina. O controlo da qualidade da água destinada a um fim específico visa, principalmente, a protecção da saúde pública uma vez que são inúmeras as doenças de veiculação hídrica causadas, entre outros, por vírus, bactérias, fungos, protozoários [OMS, 2009].

Na avaliação da qualidade da água de uma albufeira, alguns dos principais problemas prendem-se com o facto de as análises serem realizadas sobre amostras pontuais, colhidas em determinados momentos e em certos locais, muitas vezes afastados dos

laboratórios onde se realizam as análises. Estas circunstâncias enfatizam o extremo cuidado que se deve ter na extrapolação dos dados de uma análise pontual, quer para a massa de água no seu todo, quer para o intervalo de tempo entre duas amostragens. Um resultado isolado não deverá ser considerado como se representasse um valor absoluto, mas integrado num conjunto alargado de informações [Mendes & Oliveira, 2004; Zhang, 2007]. As referidas informações não devem ser limitadas à qualidade da água, mas integrar outros dados que permitam contextualizar o resultado (e.g. dados meteorológicos, dados sobre a bacia hidrográfica/albufeira, fontes de poluição pontual, usos do solo). O desejável é dispor de sistemas inteligentes de análises de dados que se integrem em sistemas de apoio ao processo de tomada de decisão.

4.2 AMOSTRAGEM E PRESERVAÇÃO DAS AMOSTRAS

Uma amostra é um conjunto de observações ou porção da população que é estudada de forma a compreender toda a população. De acordo com Zhang [Zhang, 2007], uma amostra representativa é “uma amostra de um universo/todo que seja previsível ter as propriedades médias desse mesmo universo/todo”. O principal objectivo da amostragem é, neste caso, recolher uma amostra de água de volume suficientemente pequeno que possa ser convenientemente transportada e manuseada no laboratório, mas que seja representativa da massa de água. Deste modo, as proporções relativas e as concentrações das espécies que se pretendem dosear não devem variar significativamente até à conclusão das análises [Bartram & Ballance, 1996; Eaton et al., 2005; Zhang, 2007; Yang et al., 2008]. Na verdade, não se pode depositar confiança no resultado da uma análise laboratorial, mesmo que seja seguido um método analítico aconselhado e o equipamento mais adequado, se a amostra não for representativa da água a controlar [Quevauviller et al., 2008].

A preparação do trabalho de campo é um passo essencial para o bom funcionamento do processo de amostragem. Antes da saída para o campo devem ser tidos em consideração aspectos relacionados com o planeamento do trabalho de campo, com a protecção e a segurança dos técnicos de amostragem e com os equipamentos e material a utilizar [Bartram & Ballance, 1996; Weiner, 2000; Popek, 2006]. É de toda a relevância a existência de uma lista de verificação e de preparação do trabalho de

campo onde devem constar, entre outros, o itinerário, um mapa com a localização dos pontos de amostragem e uma lista das amostras a colher em cada estação de amostragem. Deve, ainda, ser elaborada e tida em consideração uma lista de verificação do material necessário à amostragem, devendo ser garantido o transporte de pilhas e/ou baterias extra, dos acessórios necessários para o bom funcionamento dos equipamentos que devem ser, previamente, verificados e/ou calibrados [Weiner, 2000].

Além do exposto no parágrafo anterior, há que ter ainda em consideração algumas medidas básicas de protecção que o operador/técnico de amostragem deve ter em conta aquando da recolha de amostras de água. Tais medidas passam, por exemplo, pelo uso de bata ou capa impermeável, luvas descartáveis, calçado apropriado, desinfecção das mãos e caixa de primeiros socorros.

No presente trabalho, as amostras relativas à água superficial das albufeiras do Alvito e de Odivelas foram colhidas a uma profundidade situada entre os 30 e 50 cm, tendo em vista evitar a recolha de materiais sólidos (folhas, ramos, plásticos) que as pudessem contaminar. Para tal, utilizou-se uma garrafa de van Dorn havendo o cuidado de a lavar duas a três vezes com a água do local. Os recipientes de recolha foram, igualmente, lavados antes de se encherem completamente (no caso de ser necessário, antes do enchimento, adicionaram-se os reagentes de conservação). Em seguida os recipientes foram fechados, acondicionados em mala térmica e transportados a cerca de 4 °C ao abrigo da luz.

Na Figura 23 assinalam-se os locais de amostragem e as estações de monitorização meteorológicas e hidrométricas. Nas estações meteorológicas são medidas as variáveis de estado relativas à fase atmosférica do ciclo hidrológico, tais como a precipitação, a temperatura, a humidade do ar, a velocidade e a direcção do vento, a radiação solar, a insolação e a evapotranspiração. As estações hidrométricas permitem, através da medição dos níveis das albufeiras, quantificar outras variáveis como caudais, áreas inundadas e volumes armazenados.

**Legenda:**

- | | |
|--|--|
| ● Estação de Qualidade da Albufeira de Odivelas (24I/02) | ● Estação de Qualidade da Albufeira de Alvito (24J/02) |
| ● Estação Hidrométrica da Albufeira de Odivelas (24I/01A) | ● Estação Hidrométrica da Albufeira de Alvito (24J/01A) |
| ● Estação Meteorológica da Albufeira de Odivelas (24I/03C) | ● Estação Meteorológica da Albufeira de Alvito (24J/02F) |

Figura 23 – Estações de qualidade, meteorológicas e hidrométricas das albufeiras do Alvito e Odivelas.

Adaptado de SAPO [Sapo, 2009], Google Earth [Google, 2009] e SNIRH [SNIRH, 2009].

A amostragem foi realizada entre Abril e Julho de 2008 e a periodicidade, de acordo com o Decreto-Lei n.º 236/98, de 1 de Agosto, foi mensal. No local, logo após a colheita, foram determinados os parâmetros temperatura, pH, condutividade, oxigénio dissolvido e transparência. No laboratório foram feitas as determinações de Azoto Amoniacal, Azoto Kjeldahl, CBO_5 , Clorofila-a, Coliformes fecais, Coliformes totais, Condutividade, CQO, Fósforo total, Nitratos, Nitritos, Ortofosfatos, Oxidabilidade, OD, pH, SST e Temperatura.

As amostras de água foram recolhidas directamente para frascos adequados e preservadas com os reagentes apropriados para cada tipo de análise. Foram seguidas as recomendações do Standard Methods for the Examination of Water and Wastewater (SMEWW) [Eaton et al., 2005] relativamente ao tipo de recipiente utilizado, o volume recolhido, o modo de preservação da amostra e o intervalo de tempo máximo entre a realização da amostragem e a determinação analítica.

4.3 PROCEDIMENTO ANALÍTICO

Na determinação dos diferentes parâmetros de qualidade da água foram seguidas as técnicas analíticas definidas no Decreto-Lei nº 236/98, de 1 de Agosto. Relativamente aos métodos analíticos utilizados foram seguidos os descritos no SMEWW, nas normas ISO e nas normas NP, de acordo com o apresentado na Tabela 16.

Tabela 16 – Parâmetros determinados e métodos de análise.

Parâmetro	Técnica analítica	Método analítico
Azoto amoniacal	Espectrometria de absorção molecular	SMEWW 4500-NH ₃ F
Azoto Kjeldahl	Espectrometria de absorção molecular	SMEWW 4500- N _{org} B
CBO ₅	Método de Winkler Potenciometria	SMEWW 5210 B SMEWW 4500-O G
Clorofila-a	Espectrometria de absorção molecular	NP 4327:1996
Coliformes fecais	Membrana Filtrante	ISSO 9308:2000
Coliformes totais	Membrana Filtrante	ISSO 9308:2000
Condutividade	Condutimetria	SMEWW 2510 B
CQO	Volumetria	SMEWW 5220 B
Fósforo total	Espectrometria de absorção molecular	SMEWW 4500-P B
Nitratos	Espectrometria de absorção molecular	SMEWW 4500-NO ₃ B
Nitritos	Espectrometria de absorção molecular	SMEWW 4500-NO ₂ B
Ortofósatos	Espectrometria de absorção molecular	SMEWW 4500-P E
Oxidabilidade	Volumetria	NP 731:1969
OD	Método de Winkler Potenciometria	SMEWW 4500-O B SMEWW 4500-O G
pH	Potenciometria	SMEWW 4500-H ⁺ B
SST	Gravimetria	SMEWW 2540 B
Temperatura	Termometria	SMEWW 2550 B

4.4 BASE DE DADOS

A Base de Dados sobre a qual assenta este trabalho foi construída a partir de dados experimentais, por nós obtidos, durante o período compreendido entre Abril a Julho de 2008 e complementada com os dados de qualidade da água, hidrométricos e meteorológicos disponibilizados pelo INAG, através do portal do SNIRH [SNIRH,

2009] ficando coberto o período compreendido entre Junho de 2001 e Dezembro de 2008.

A Base de Dados contém 78 campos e 81 registos. Na Figura 24 apresenta-se um excerto da BD, a qual foi construída utilizando o *Microsoft Excel* e, na Tabela 17, apresentam-se alguns dados estatísticos que permitem caracterizar as diversas variáveis. No Anexo I disponibiliza-se a versão integral da BD utilizada.

1	A	B	C	D	E	F	G	H	I	J	K	L	M	N
	Ano	Mês	Clorofila-a	Coliformes fecais	Coliformes totais	Condutividade	Nitratos	Oxidabilidade	OD	pH	SST	Temperatura	Vol arm	Cota
18	2002	Outubro	4,34	9	270	277	0,31	4,8	145	8,5	3,6	21,6	120884	196,53
19	2002	Novembro	2,27	8	200	274	0,18	4	107	8,5	0,9	20	124105	196,49
20	2002	Dezembro	5,32	5	12	275	0,71	4,4	83,5	8,1	3,1	15	126044	196,7
21	2003	Janeiro	3,38	41	1160	251	1,46	4,8	75	7,9	6,7	13	132500	197,43
22	2003	Fevereiro	8,37	0	7	263	1,46	6,6	84,1	7,8	5,2	11,2	132500	197,63
23	2003	Março	56,37	0	120	263	1,02	7,6	121,8	8,5	11	15,5	132500	197,73
24	2003	Abril	45,1	11	25	268	1,02	5,9	100,8	8,3	7,1	16,3	132500	197,6
25	2003	Mai	8,4	1	900	278	0,93	4,7	95,2	8	25	18,4	132500	197,61
26	2003	Junho	30,25	2	150	272	0,93	4	92,3	8,5	3,9	23	127334	197,5
27	2003	Julho	4,02	90	6400	254	0,09	4,7	80,4	8,3	2,2	24,5	127334	197,59
28	2003	Agosto	2,49	8	60	308	0,89	4,5	112,1	8,5	7,2	27,5	116369	196,23
29	2003	Setembro	1,39	2	790	320	0,93	4,3	91,7	8	8,4	24	112074	195,8
30	2003	Outubro	1,74	60	300	319	1,02	3,9	144,5	8,1	8,7	20,8	113254	195,9
31	2003	Novembro	6,19	50	260	304	0,46	3,8	102,1	8,1	3,1	17,2	118304	196,01
32	2003	Dezembro	7,62	21	200	283	1,59	3,8	116,5	8	2,9	13,8	124109	196,27
33	2004	Janeiro	1,8	0	4	305	1,32	3,4	94,6	8,1	2,7	13	125786	196,79
34	2004	Fevereiro	4,02	6	31	298	1,59	4,5	100,2	8,5	7	13,8	127072	196,92
35	2004	Março	9,51	26	230	300	1,12	4,7	101,4	8,3	5,9	14,2	130779	197,13
36	2004	Abril	11,54	22	72	269	0,48	5,8	105,8	8,5	14	16,8	126500	197,2
37	2004	Mai	1,99	0	70	309	0,23	4	112,4	8,1	2,2	18,3	132500	197,13
38	2004	Junho	1,39	4	15	296	0,19	3,9	121,2	8,4	2	25,1	126044	197,06
39	2004	Julho	1,48	56	1000	303	0,27	4,3	98,2	8,7	2	27,1	121529	196,85
40	2004	Agosto	1,8	30	180	317	0,29	4,2	99,4	8,6	2	27,4	118304	196,64
41	2004	Setembro	1,85	10	120	335	0,09	4,4	86,2	8,5	2	23,9	117040	196,4
42	2004	Outubro	1,29	220	820	340	0,49	3,7	92	8,5	2	22,7	119336	196,26

Figura 24 – Excerto da base de dados utilizada.

Tabela 17 – Caracterização, em termos estatísticos, das variáveis utilizadas.

Albufeira	Parâmetro	Unidade	Mínimo	Máximo	Média	Desvio Padrão	Nº de registos	
Alvito	Qualidade da água	CQO	mg/L O ₂	10	56	19,21	7,42	81
		CBO ₅	mg/L O ₂	2	8	3,08	0,87	81
		Clorofila-a	µg/L	1,27	56,37	7,74	9,96	81
		Coliformes fecais	ufc/100 mL	0	3300	63,69	366,40	81
		Coliformes totais	ufc/100 mL	2	7100	457	1126	81
		Condutividade	µS/cm	227	422	320,56	54,40	81
		Fósforo total	mg/L	0,001	0,29	0,05	0,06	81
		Nitratos	mg/L	0,09	2,97	1,07	0,78	81
		Nitritos	mg/L	0,001	0,16	0,02	0,03	81
		Ortofosfatos	mg/L P ₂ O ₅	0,01	0,29	0,04	0,04	81
		Oxidabilidade	mg/L	2,9	7,6	4,55	0,87	81
		OD	% Saturação	29,0	167	103,02	21,33	81
		pH	Escala de Sorensen	7,8	8,98	8,32	0,26	81
		SST	mg/L	0,6	39	6,13	6,50	81
		Temperatura da água	°C	9,8	27,50	19,13	5,14	81
	Meteorológicos e Hidrométricos	Humidade relativa (*)	%	50,97	90,77	72,31	9,88	81
		Humidade relativa (**)	%	40	93	72,99	12,42	81
		Humidade relativa (***)	%	44,7	91,6	73,18	11,14	81
		Precipitação (*)	mm	0	6,64	1,35	1,33	81
		Precipitação (**)	mm	0	11,98	1,46	2,32	81
		Precipitação (***)	mm	0	14,46	1,61	2,35	81
		Temperatura do ar (*)	°C	7,71	25,05	16,46	5,02	81
		Temperatura do ar (**)	°C	6,36	29,64	16,58	5,54	81
		Temperatura do ar (***)	°C	6,83	27,3	16,52	5,21	81
		Velocidade do vento (*)	m/s	0,92	3,63	1,89	0,50	81
		Velocidade do vento (**)	m/s	0,84	4,22	1,80	0,61	81
		Velocidade do vento (***)	m/s	1,06	3,5	1,84	0,54	81
		Radiação (*)	W/m ²	1542	7742	4564	1932	81
		Radiação (**)	W/m ²	1374	8022	4532	2015	81
		Radiação (***)	W/m ²	1466	7993	4492	2004	81
Volume armazenado	dam ³	48070	132500	105445	26400	81		
Cota	m	173,48	197,73	194,92	3,43	81		

Albufeira	Parâmetro	Unidade	Mínimo	Máximo	Média	Desvio Padrão	Nº de registos	
Odivelas	Qualidade da água	CQO	mg/L O ₂	1	51	22	8,25	81
		CBO ₅	mg/L O ₂	2	5	3	0,49	81
		Clorofila-a	µg/L	1,11	351	15	48	81
		Coliformes Fecais	ufc/100 mL	0	570	43	87	81
		Coliformes Totais	ufc/100 mL	2	19000	663	2561	81
		Condutividade	µS/cm	280	494	392	54	81
		Fósforo	mg/L	0,01	2,919	0,09	0,33	81
		Nitratos	mg/L	0,09	4,6	1,13	0,86	81
		Nitritos	mg/L	0,001	0,22	0,02	0,03	81
		Ortofosfatos	mg/L P ₂ O ₅	0,01	0,18	0,04	0,03	81
		Oxidabilidade	mg/L	2,4	23,2	5,8	2,23	81
		OD	% Saturação	60,5	113,8	81,3	11,7	81
		pH	Escala de Sorensen	7,3	9,09	8,2	0,32	81
		SST	mg/L	1,3	350	11,2	39,1	81
		Temperatura da água	°C	11	29,1	20,1	4,9	81
		Meteorológicos e Hidrométricos	Humidade relativa (*)	%	50,5	90,71	71,06	10,29
	Humidade relativa (**)		%	37,4	96,40	72,10	12,69	81
	Humidade relativa (***)		%	44,5	92,20	72,29	11,49	81
	Precipitação (*)		mm	0	6,92	1,25	1,24	81
	Precipitação (**)		mm	0	8,06	1,26	1,95	81
	Precipitação (***)		mm	0	13,71	1,47	2,17	81
	Temperatura do ar (*)		°C	7,82	25,20	16,74	5,14	81
	Temperatura do ar (**)		°C	6,46	29,94	16,91	5,58	81
	Temperatura do ar (***)		°C	6,93	26,92	16,81	5,24	81
	Velocidade do vento (*)		m/s	0,27	1,26	0,71	0,20	81
	Velocidade do vento (**)		m/s	0,16	1,76	0,69	0,30	81
	Velocidade do vento (***)		m/s	0,2	1,41	0,72	0,26	81
	Radiação (*)		W/m ²	1488	7112	4264	1693	81
	Radiação (**)		W/m ²	1544	7275	4254	1766	81
	Radiação (***)	W/m ²	1482	7271	4207	1758	81	
Volume armazenado	dam ³	33132	95700	65455	18335	81		
Cota	m	93,2	103,2	99,1	2,84	81		

(*) Média relativa ao período entre amostragens; (**) Média relativa aos 5 dias anteriores à amostragem e (***) Média relativa aos 10 dias anteriores à amostragem

CAPÍTULO 5 – APRESENTAÇÃO, INTERPRETAÇÃO E ANÁLISE DE RESULTADOS E MODELOS

A crescente importância que se atribui à DCBD, com recurso a técnicas de *Data Mining* tem, como foi referido anteriormente, dado origem ao desenvolvimento de inúmeras aplicações de software que procuram dar resposta aos problemas e desafios que se têm levantado sendo, actualmente, conhecidas mais de cento e cinquenta ferramentas diferentes [Cruz, 2007]. No presente trabalho procurou-se uma ferramenta não comercial que permitisse, por um lado, implementar Árvores de Decisão e RNAs na tarefa de classificação e, por outro, possibilitasse a segmentação dos dados tendo a escolha recaído na ferramenta WEKA (*Waikato Environment for Knowledge Analysis*).

5.1 BREVE DESCRIÇÃO DA FERRAMENTA WEKA

A ferramenta WEKA permite desenvolver o processo de DCBD, uma vez que disponibiliza diversas formas de preparação de dados, contempla uma série de algoritmos de *Data Mining* e possibilita a utilização de diversos métodos, tendo em vista a validação dos resultados. A ferramenta foi desenvolvida na Universidade de Waikato (Nova Zelândia), faz uso das facilidades disponibilizadas pela programação orientada ao objecto encontrando-se escrito em linguagem *Java*. Está licenciado ao abrigo da *General Public License* sendo, portanto, possível estudar e alterar o respectivo código fonte [Weka, 2004].

O *software* WEKA possui interfaces gráficas amigáveis permitindo ao utilizador escolher, logo no menu inicial, uma de quatro possíveis interfaces possibilitando, deste modo, a sua utilização por utilizadores menos familiarizados com a ferramenta. No que concerne aos dados, estes têm que estar num formato próprio do WEKA (extensão *.arff*), o que pode ser conseguido facilmente utilizando um conversor, disponibilizado pelo *software*, aplicável para ficheiros do tipo “separado por vírgulas” (extensão *.csv*) ou do tipo “separado por tabs” (extensão *.tsv*).

A ferramenta *WEKA* disponibiliza diversos algoritmos de *Data Mining*, entre os quais se salientam as RNAs e as Árvores de Decisão. Relativamente à validação dos resultados permite a utilização de validação com divisão da amostra e de validação cruzada. Além do exposto, permite que o utilizador defina o número de repetições tendo em vista assegurar a significância estatística dos resultados.

5.2 MODELOS DE PREVISÃO DOS PARÂMETROS DA QUALIDADE DA ÁGUA

Como foi referido anteriormente, o processo de DCBD comporta várias etapas, as quais se podem observar na Figura 17. As etapas mencionadas podem ser agrupadas em duas grandes tarefas (Figura 25). A primeira está relacionada com a preparação dos dados no que concerne à sua selecção, pré-processamento e transformação. A segunda tarefa, designada por modelação, inclui a aplicação de ferramentas de DM, a visualização a interpretação e a avaliação dos resultados.

Neste trabalho utilizou-se a ferramenta *Microsoft Excel* na fase inicial de preparação dos dados e a aplicação *WEKA* tanto na preparação dos dados como na fase de modelação.

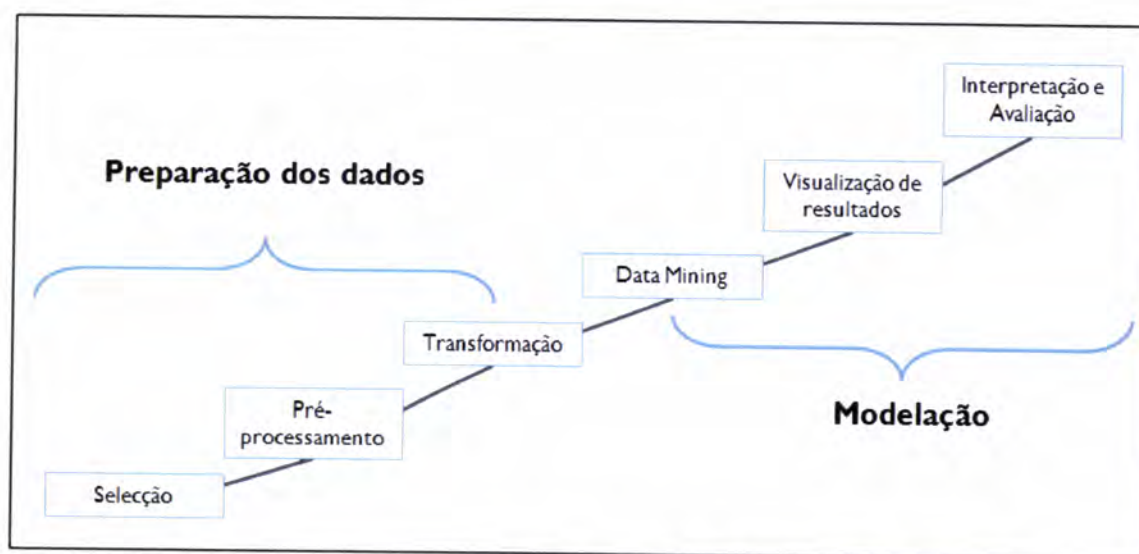


Figura 25 – Etapas de desenvolvimento do processo de DCBD agrupadas em tarefas.

5.2.1 PREPARAÇÃO DOS DADOS

A preparação dos dados inclui, como se pode observar na Figura 18, as etapas de selecção, pré-processamento e transformação dos dados. Na etapa de selecção é realizada a escolha do conjunto de dados e/ou variáveis necessárias ao estudo, bem como a segmentação dos dados, de acordo com determinados critérios. Esta etapa tem como principal objectivo limitar o espaço de pesquisa definindo, assim, o domínio sobre o qual se pretende executar o processo de DCBD.

No presente trabalho, para efeitos de classificação, teve-se em conta o critério proposto pelo INAG, que classifica as massas de água em cinco classes possíveis de acordo com a grelha paramétrica apresentada na Tabela 5.

A qualidade dos dados e a forma como se encontram estruturados pode condicionar o sucesso e a utilidade de uma análise inteligente dos mesmos. Na etapa de pré-processamento são realizados procedimentos que visam melhorar a qualidade dos dados, contribuindo para que a etapa de *Data Mining* seja realizada de forma mais eficiente. Quando se analisa uma BD podem ser identificados diversos tipos de problemas que podem estar relacionados com o tamanho e dimensionamento da BD, a volatilidade dos dados, a ausência de algum atributo relevante, a presença de ruído e com a existência de dados incompletos, omissos ou desconhecidos.

Existem diversas técnicas, com origem na estatística e na matemática, para resolver os problemas relacionados com dados incompletos, omissos ou desconhecidos, tais como os métodos Bayesianos [Gelman et al., 2003] e os métodos de imputação [Little, 1992; Rubin, 1996; Schafer, 1999; Mamedov & Udalov, 2002].

Um dos métodos de imputação mais comuns, em BD com elevado número de registos, consiste na eliminação dos exemplos com dados omissos. No entanto, em algumas situações, este método pode conduzir a alterações significativas na representatividade dos dados relativamente ao universo em estudo e conduzir a amostras muito pequenas.

Na BD inicial existiam registos que apresentavam valores omissos. Contudo, a existência de dados em falta está relacionada com a natureza do problema em estudo, uma vez que, de acordo com o Decreto-Lei nº 238/98, de 1 de Agosto, a periodicidade da amostragem varia consoante o grupo no qual o parâmetro se insere (Tabela 8).

Deste modo, nos meses em que não foi realizada a amostragem relativa aos parâmetros dos grupos G2 e G3 foi-lhes atribuído o valor determinado na campanha imediatamente anterior. Após ter sido feita a atribuição referida anteriormente apenas um reduzido número de registos apresentava valores omissos, tendo sido eliminados da BD.

Após terem sido descartados os registos incompletos e se ter atribuído a cada registo (campanha de amostragem) a respectiva classe de qualidade fez-se a distribuição do número de casos por classe, tendo em vista avaliar a necessidade de adoptar estratégias que possam melhorar o desempenho das ferramentas de *Data Mining*.

Na Figura 26 apresentam-se, respectivamente, as distribuições do número de casos por classe de qualidade da água, para as albufeiras do Alvito e de Odivelas.

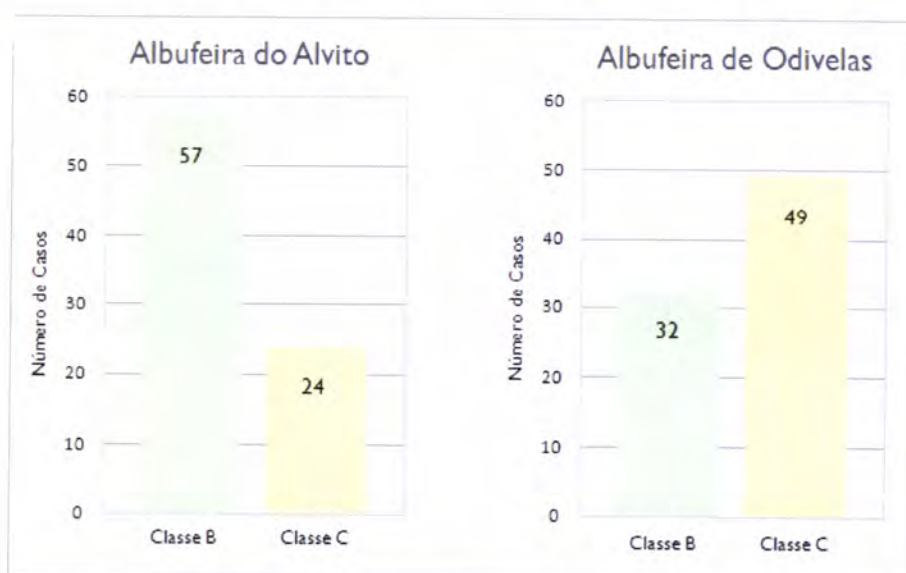


Figura 26 – Distribuição do número de casos por classe de qualidade para as albufeiras estudadas.

A análise das distribuições, apresentadas na Figura 26, mostra a existência de uma desproporção entre o número de casos incluídos em cada uma das classes. Este desequilíbrio, de acordo com Weiss e Provost [Weiss & Provost, 2001], pode originar modelos com pior desempenho na previsão das classes minoritárias, sendo normalmente apontadas duas razões para esta situação. A primeira deriva do facto das regras geradas para a classe minoritária serem baseadas em menos exemplos e, conseqüentemente, mais sobreajustadas. A segunda razão prende-se com o facto de a

classe mais frequentemente prevista ser a maioritária existindo, portanto, uma probabilidade maior de classificar, de forma incorrecta, os exemplos da classe minoritária.

Existem diversas estratégias que visam tornar a distribuição de classes mais equilibrada. Um dos métodos, designado por *under-sampling*, consiste em criar uma amostra mais pequena do conjunto de exemplos da classe maioritária. Em sentido oposto, no método de *over-sampling*, são gerados casos, a partir dos casos iniciais, de forma a aumentar a classe minoritária. No entanto, deve ser salientado que este balanceamento acontece somente para o conjunto de treino mantendo-se, no conjunto de teste, a distribuição original. A principal desvantagem do método de *under-sampling* prende-se com o facto de poder desprezar dados potencialmente úteis. No método de *over-sampling*, por seu lado, as desvantagens estão relacionadas com o aumento da cardinalidade do conjunto de treino e, conseqüentemente, do tempo de computação podendo, ainda, dar origem a problemas sobre-ajustamento [Pal & Jain, 2005].

No presente trabalho, atendendo ao número de exemplos pertencentes a cada uma das classes, para não correr o risco de descartar casos potencialmente relevantes, optou-se pelo método de *over-sampling* à semelhança do que tem sido feito noutros trabalhos [Pinto, 2009].

5.2.2 MODELAÇÃO

Antes de se dar início à fase de construção dos modelos, os objectivos que se pretendem atingir devem estar claramente definidos bem como o conjunto das variáveis de entrada e das variáveis de saída. Neste trabalho procuram-se obter modelos que permitam prever a classe de qualidade da água. Atendendo a que se pretende associar cada registo a uma de várias classes discretas possíveis, está-se perante um problema de classificação, tendo-se recorrido às Árvores de Decisão (AD) e às Redes Neurais Artificiais (RNAs).

A escolha das referidas técnicas prende-se com diversas razões. Por um lado, as AD são umas das técnicas de *Data Mining* mais utilizadas neste tipo de problemas, permitindo uma representação expressiva do conhecimento, de fácil utilização e compreensão. Por outro lado, as RNAs são actualmente consideradas como ferramentas capazes de apreender e representar relações complexas entre variáveis.

Para além das razões expostas anteriormente, é ainda de referir que estas técnicas têm sido aplicadas e originado resultados satisfatórios em estudos de previsão quer de classes de qualidade de água [Vicente, 2004; Santos et al., 2005; Atkins et al., 2007; Pinto et al., 2009; Pinto, 2009] quer de efluentes resultantes da indústria petrolífera [Fernandes, 2009; Fernandes et al., 2009].

No que concerne à construção de modelos utilizando RNAs, de entre as várias arquitecturas de rede que se conhecem, optou-se pela utilização de Redes Neurais Unidireccionais (RNUs) uma vez que esta é umas das arquitecturas mais utilizadas tanto em problemas de classificação como de regressão. Como foi referido anteriormente, esta arquitectura é formada por, pelo menos, três camadas de nodos (a camada de entrada, a camada de saída e uma ou mais camadas intermédias). O número de nodos na camada de entrada corresponde ao número de variáveis independentes sendo a camada de saída constituída por tantos nodos quantas as variáveis dependentes.

Relativamente às camadas intermédias, muito embora tenha sido provado que uma camada intermédia é suficiente para apreender qualquer função contínua [Hornik et al., 1989], outros estudos apontam que, na prática, muitas funções são difíceis de apreender com apenas uma camada intermédia [Cheng & Titterington, 1994; Flood & Kartam, 1994]. Na verdade, não existem regras claras para estabelecer quer o número de camadas intermédias, quer o número de nodos que as constituem. Na maior parte dos casos, o estabelecimento da melhor arquitectura de rede é feito por tentativa e erro, começando-se por uma estrutura inicial que vai sendo selectivamente modificada [Han & Kamber, 2006].

O algoritmo de treino utilizado neste trabalho foi o algoritmo de retro-propagação que, como foi referido anteriormente, tem sido apontado como o algoritmo de treino supervisionado mais utilizado em RNUs. Pelo mesmo motivo escolheu-se, como função de activação, a função logística (Tabela 14). Em todos os casos optou-se por utilizar os parâmetros de configuração sugeridos pelo *software* WEKA (Tabela 18).

Tabela 18 – Valores assumidos por defeito pelo algoritmo de implementação de RNAs.

Parâmetro	Descrição	Valor por Defeito
<i>Auto Build</i>	Constrói as camadas intermédias da rede	<i>True</i>
<i>Decay</i>	Diminui a taxa de aprendizagem (a taxa de aprendizagem de cada iteração é obtida dividindo-a pelo número da iteração)	<i>False</i>
<i>Hidden Layers</i>	Define o número de camadas intermédias	$(attributes + classes) / 2$
<i>Learning Rate</i>	Taxa de aprendizagem	0,3
<i>Momentum</i>	Coefficiente de momento	0,2
<i>Nominal to Binary Filter</i>	Converte os atributos nominais em binários	<i>True</i>
<i>Normalize Attributes</i>	Normaliza os atributos numéricos	<i>True</i>
<i>Normalize Numeric Classes</i>	Normaliza o atributo a prever caso seja numérico	<i>True</i>
<i>Random Seed</i>	Semente para gerar aleatoriamente os pesos iniciais	0
<i>Reset</i>	Reinicia a aprendizagem, com uma taxa de aprendizagem inferior, caso o algoritmo esteja a divergir	<i>True</i>
<i>Training Time</i>	Número de iterações de treino	500
<i>Validation Set Size</i>	Percentagem de dados a serem utilizados para validação	0
<i>Validation Threshold</i>	Número de vezes que o erro pode aumentar nos dados de validação até o treino ser concluído	20

No que respeita à implementação de AD foi utilizado o algoritmo C4.5 que, nesta ferramenta, é designado por J.48 tendo-se, tal como anteriormente, utilizado a parametrização sugerida pelo *software* (Tabela 19). Relativamente aos modelos de segmentação, foi utilizada a estratégia *k-means*, sendo apenas necessário fornecer ao programa o número de segmentos a formar.

Tabela 19 – Valores assumidos por defeito pelo algoritmo de indução de árvores de decisão.

Parâmetro	Descrição	Valor por Defeito
<i>Binary Splits</i>	Divisão binária em atributos nominais	<i>False</i>
<i>Confidence Factor</i>	Factor de confiança utilizado na poda	0,25
<i>Debug</i>	Apresentação de informação adicional	<i>False</i>
<i>Min Num Obj</i>	Número mínimo de instâncias por folha	2
<i>Num Folds</i>	Define os dados utilizados para a poda	3
<i>Seed</i>	Semente para gerar aleatoriamente os índices quando se usa a opção <i>reduced error pruning</i>	1
<i>Unpruned</i>	Impede a poda	<i>False</i>

Tendo em vista assegurar a significância estatística dos resultados obtidos, cada simulação foi executada vinte vezes. Em cada uma delas os registos da BD foram aleatoriamente divididos em dois conjuntos mutuamente exclusivos. O conjunto de treino, composto por 67% dos exemplos e o conjunto de teste, constituído pelos restantes.

5.3 APRESENTAÇÃO, DISCUSSÃO E AVALIAÇÃO DOS MODELOS DE CLASSIFICAÇÃO

Neste trabalho foram realizados modelos de previsão da qualidade da água das albufeiras do Alvito e de Odivelas. Como foi referido anteriormente, trata-se de um problema de classificação, tendo-se recorrido às AD e às RNAs. Em seguida apresentam-se, discutem-se, avaliam-se e comparam-se os modelos obtidos através das referidas técnicas apresentando-se as respectivas matrizes de coincidências.

5.3.1 MODELOS DE PREVISÃO DA QUALIDADE DA ÁGUA – ALBUFEIRA DO

ALVITO

A AD obtida para previsão das classes de qualidade da água da albufeira do Alvito, tendo sido disponibilizados ao algoritmo todos os parâmetros da qualidade da água existentes na BD, encontra-se na Figura 27. A sua análise mostra que a classificação é conseguida utilizando apenas quatro variáveis de entrada (pH, CQO, oxidabilidade e temperatura). O parâmetro pH é o factor que mais influencia a classificação, secundado pelo CQO e pela oxidabilidade, surgindo a temperatura da água no ramo mais afastado da raiz.

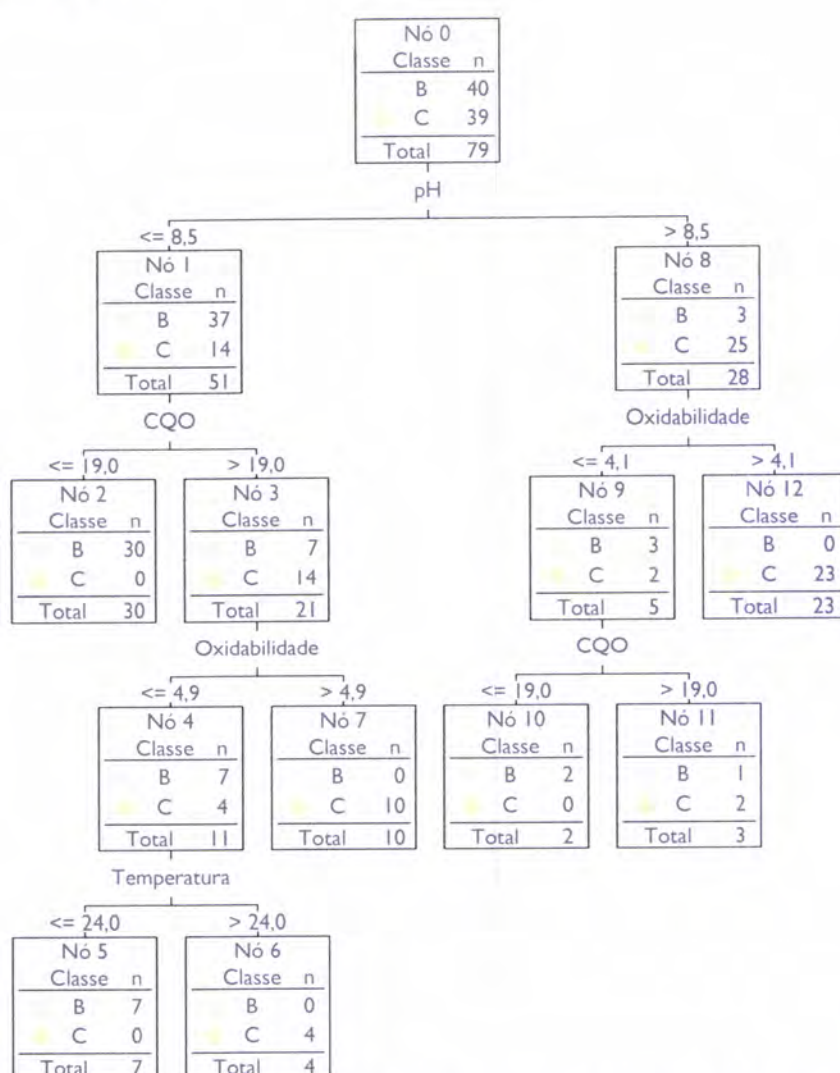


Figura 27 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito.

A informação contida nas AD pode ser traduzida em termos de um conjunto de regras. Na Tabela 20 apresentam-se as regras obtidas a partir do modelo apresentado na Figura 27 destacando-se, em cada uma delas, a confiança e o suporte.

Tabela 20 – Conjunto de regras relativo ao modelo apresentado na Figura 27.

Regras para a Classe B	Regras para a Classe C
<p>Regra 1 Se pH \leq 8,5 e CQO \leq 19,0 Então \rightarrow B (30; 1,0)</p>	<p>Regra 1 Se pH \leq 8,5 e CQO $>$ 19,0 e Oxidabilidade \leq 4,9 e Temperatura $>$ 24,0 Então \rightarrow C (4; 1,0)</p>
<p>Regra 2 Se pH \leq 8,5 e CQO $>$ 19,0 e Oxidabilidade \leq 4,9 e Temperatura \leq 24,0 Então \rightarrow B (7; 1,0)</p>	<p>Regra 2 Se pH \leq 8,5 e CQO $>$ 19,0 e Oxidabilidade $>$ 4,9 Então \rightarrow C (10; 1,0)</p>
<p>Regra 3 Se pH $>$ 8,5 e Oxidabilidade \leq 4,1 e CQO \leq 19,0 Então \rightarrow B (2; 1,0)</p>	<p>Regra 3 Se pH $>$ 8,5 e Oxidabilidade \leq 4,1 e CQO $>$ 19,0 Então \rightarrow C (3; 0,67)</p>
	<p>Regra 4 Se pH $>$ 8,5 e Oxidabilidade $>$ 4,1 Então \rightarrow C (23; 1,0)</p>

A regra 1 (Se pH \leq 8,5 e CQO \leq 19,0 Então \rightarrow Classe B) é, das regras geradas para a classe B, a que exhibe suporte mais elevado, 30. No que diz respeito à classe C, é a regra 4 (Se pH $>$ 8,5 e oxidabilidade $>$ 4,1 Então \rightarrow classe C), que apresenta suporte mais elevado, 23, destacando-se, em ambas as regras, o facto da confiança ser um.

A acuidade do modelo, medida em termos de percentagem de acertos relativamente ao número de casos apresentados, é de 98,7% (78 casos classificados correctamente em 79) para o conjunto de treino e de 97,0% (32 casos classificados correctamente em 33) para o conjunto de teste, Tabela 21. O facto dos acertos relativos em cada uma

das classes serem semelhantes, quer para o conjunto de treino quer para o conjunto de teste, mostra que o modelo tem boa capacidade de generalização, não revelando indícios de sobre-ajustamento.

Tabela 21 – Matriz de coincidências relativa ao modelo apresentado na Figura 27.

Classe	Conjunto de Treino		Conjunto de Teste	
	B	C	B	C
B	39	1	17	1
C	0	39	0	15

No que respeita às RNAs optou-se, como foi referido anteriormente, pela utilização de RNUs. Foram construídas e avaliadas diversas topologias de rede, tendo o número óptimo de camadas e de nodos sido determinado por tentativa e erro. A RNU seleccionada para prever as classes de qualidade apresenta a topologia 3-9-3-2, i.e. uma camada de entrada com três nodos, duas camadas intermédias com, respectivamente, nove e três nodos e uma camada de saída com dois nodos, correspondentes às classes B e C, Figura 28.

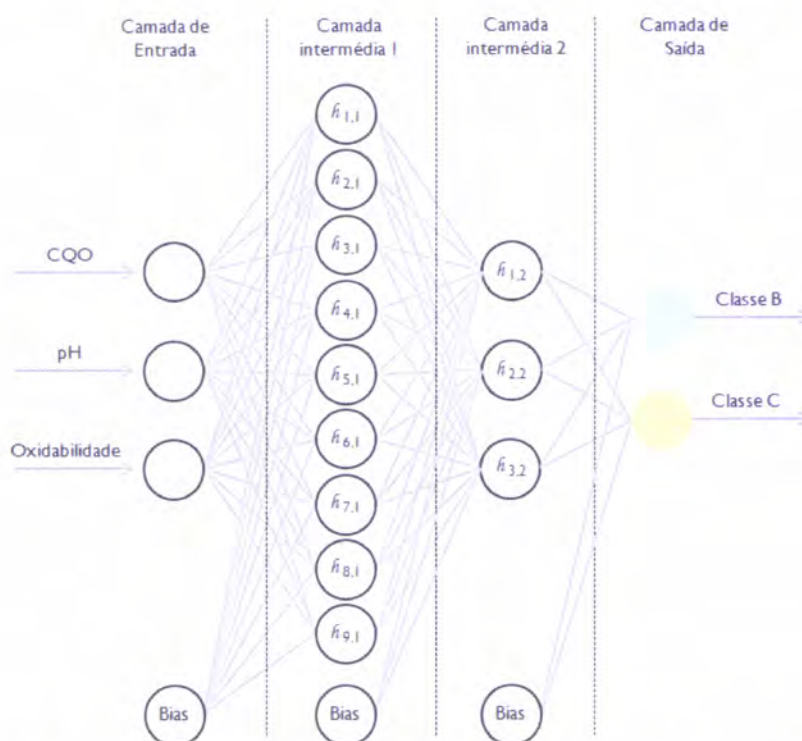


Figura 28 – RNA obtida para previsão da qualidade da água da albufeira do Alvito.

O modelo obtido faz uso, como variáveis de entrada, de apenas três parâmetros (pH, oxidabilidade e CQO) apesar de estarem disponíveis todos os parâmetros de qualidade da água existentes na BD. Os pesos associados às conexões entre nodos de camadas adjacentes encontram-se no Anexo II, nas Tabelas II.1, II.2 e II.3.

A acuidade global do modelo, medida em termos de percentagem de acertos relativamente ao número de casos apresentados é de 97,4% (75 casos classificados correctamente em 77) para o conjunto de treino e de 97,1% (34 casos classificados correctamente em 35) para o conjunto de teste, Tabela 22. A semelhança entre a percentagem de respostas correctas relativamente a cada uma das classes, quer para o conjunto de treino quer para o conjunto de teste, mostra que o modelo tem boa capacidade de generalização, não revelando indícios de sobre-ajustamento.

Tabela 22 – Matriz de coincidências relativa ao modelo apresentado na Figura 28.

Classe	Conjunto de Treino		Conjunto de Teste	
	B	C	B	C
B	36	2	18	1
C	0	39	0	16

A análise conjunta da Tabela 21 e Tabela 22 mostra que os modelos construídos com recurso a AD e a RNAs apresentam desempenhos similares. No que diz respeito às variáveis de entrada, há a salientar o facto de ambos os modelos fazerem uso dos parâmetros pH, CQO e oxidabilidade tendo, no entanto, o modelo construído com recurso a AD utilizado o parâmetro temperatura, muito embora no nível mais afastado da raiz.

Na comparação dos dois modelos, um aspecto que tem que ser destacado prende-se com o facto de os erros cometidos apenas ocorrerem na previsão da classe B o que, do ponto de vista da utilização da água, não trás grandes problemas uma vez que esta tem qualidade superior à que é prevista. Na verdade, erros na previsão da classe C serão sempre de evitar, uma vez que uma água desta classe tem um uso mais limitado que uma água da classe B.

Tendo em vista, por um lado, avaliar a importância das variáveis pH, CQO e oxidabilidade na previsão das classes de qualidade e, por outro lado, quais os

parâmetros que poderiam ser utilizadas para sua substituição, foram gerados novos modelos tendo-se, numa primeira fase, retirado sucessivamente uma dessas variáveis. Em seguida, as referidas variáveis foram retiradas duas a duas e, finalmente, foram retiradas em simultâneo. Os modelos obtidos, as regras geradas e os pesos associados às conexões entre os nodos das camadas adjacentes encontram-se no Anexo III. Na Tabela 23 comparam-se as AD e as RNAs obtidas em cada uma das referidas situações apresentando-se a topologia da rede/profundidade da AD, os parâmetros utilizados como entrada, as matrizes de coincidências e a acuidade dos modelos, medida em termos de percentagem de acertos relativamente ao número de casos apresentados.

A análise da Tabela 23 mostra que no caso de se retirar apenas um dos parâmetros utilizados como entradas no modelo inicial (pH, CQO ou oxidabilidade) conseguem-se, de um modo geral, modelos de previsão de maior complexidade, traduzida na utilização de um maior número de variáveis de entrada, topologias com maior número de nodos ou AD com maior profundidade, i.e. com maior número de níveis. A análise do desempenho destes modelos mostra que têm menor acuidade, na ordem de 95% de classificações correctas para o conjunto de treino, independentemente da técnica de *Data Mining* utilizada (AD ou RNAs). No que respeita ao conjunto de teste, a acuidade obtida é próxima ou, em alguns casos, superior à obtida para o respectivo conjunto de treino, não revelando os modelos indícios de sobre-ajustamento. A única excepção verifica-se para o modelo 3, em que se verifica uma diminuição acentuada da acuidade. Um outro aspecto a salientar prende-se com o facto de todos os modelos obtidos, excepto o anteriormente referido, não cometerem erros na previsão da classe C o que, como foi referido, é relevante do ponto de vista da utilização da água.

Quando são retirados dois parâmetros em simultâneo a acuidade dos modelos, no que concerne ao conjunto de treino diminui sendo, no entanto, superior a 90%. Relativamente ao conjunto de teste a acuidade obtida foi semelhante à obtida para o respectivo conjunto de treino tendo, em dois casos, sido ligeiramente superior (modelos 11 e 14). Tal como anteriormente, ocorre uma excepção (modelo 9), havendo uma diminuição relativamente acentuada (92,4% para o conjunto de treino contra 85,3% para o conjunto de teste). Contudo, apesar de exibirem um desempenho aceitável, todos eles, à excepção do modelo 10, cometem erros de previsão para a classe C o que pode, de alguma forma por em causa a sua utilização. Por fim, quando

se retiram simultaneamente o pH, a CQO e a oxidabilidade os modelos exibem acuidade inferior e, além disso, apresentam falhas na previsão da classe C.

As RNAs parecem ser mais adequadas para a previsão das classes de qualidade da água da albufeira do Alvito, uma vez que não cometem falhas relativamente à classe C, quando se retira apenas um dos três parâmetros em estudo. Quando são retirados, simultaneamente, dois dos parâmetros, desde que se permita o acesso à oxidabilidade, também se verifica 100% de acerto relativamente à classe C.

No que concerne às AD, são cometidos erros na previsão da classe C tanto no caso de se retirar apenas o parâmetro pH como no caso de se retirarem simultaneamente dois ou mais parâmetros. Um outro aspecto que se destaca nestes modelos está relacionado com o facto de, na ausência do pH, ser escolhido para o nó raiz a CQO e, no caso de ambos estarem ausentes, a escolha recair sobre a oxidabilidade.

Tabela 23 – Síntese dos diversos modelos obtidos para previsão da qualidade da água da albufeira do Alvito.

Parâmetro Retirado	Árvores de Decisão											Redes Neurais Artificiais												
	Modelo	Profundidade	Entradas	Matrizes de Coincidências						Acuidade		Notas	Modelo	Topologia	Entradas	Matrizes de Coincidências						Acuidade		Notas
				Conjunto de Treino			Conjunto de Teste			Conjunto de Treino	Conjunto de Teste					Conjunto de Treino			Conjunto de Teste			Conjunto de Treino	Conjunto de Teste	
Nenhum	1	4	Oxidabilidade pH CQO Temperatura	Classe	B	C	Classe	B	C	98.7%	97.0%	Figura 27	2	3-9-3-2	CQO Oxidabilidade pH	Classe	B	C	Classe	B	C	97.4%	97.1%	Figura 28
				B	39	1	B	17	1							B	18	1						
				C	0	39	C	0	15							C	0	16						
pH	3	5	CQO OD Temperatura CBO ₅ SST Condutividade	Classe	B	C	Classe	B	C	94.8%	88.6%	Anexo III.1	4	5-27-8-2	CQO OD Temperatura Oxidabilidade CBO ₅	Classe	B	C	Classe	B	C	94.7%	94.6%	Anexo III.1
				B	38	3	B	16	2							B	20	2						
				C	1	35	C	2	15							C	0	15						
CQO	5	4	pH Oxidabilidade OD Condutividade	Classe	B	C	Classe	B	C	94.8%	93.8%	Anexo III.2	6	4-29-17-2	pH Oxidabilidade Volume Arm. Condutividade	Classe	B	C	Classe	B	C	96.2%	94.1%	Anexo III.2
				B	34	4	B	17	2							B	11	2						
				C	0	39	C	0	13							C	0	21						
Oxidabilidade	7	4	CQO pH Cota Temperatura Condutividade	Classe	B	C	Classe	B	C	96.1%	97.0%	Anexo III.3	8	5-4-4-2	CQO pH Volume Arm. Temperatura Condutividade	Classe	B	C	Classe	B	C	96.3%	97.1%	Anexo III.3
				B	37	3	B	16	1							B	14	1						
				C	0	37	C	0	16							C	0	19						
pH e CQO	9	4	Oxidabilidade Condutividade CBO ₅ Temperatura	Classe	B	C	Classe	B	C	92.4%	85.3%	Anexo III.4	10	3-4-3-2	Oxidabilidade Volume Arm. Temperatura	Classe	B	C	Classe	B	C	94.9%	93.9%	Anexo III.4
				B	35	5	B	14	3							B	16	2						
				C	1	38	C	2	15							C	0	15						
pH e Oxidabilidade	11	5	CQO Temperatura SST OD	Classe	B	C	Classe	B	C	90%	91.2%	Anexo III.5	12	6-10-6-2	CQO OD Volume Arm. Temperatura Condutividade CBO ₅	Classe	B	C	Classe	B	C	92.5%	91.2%	Anexo III.5
				B	30	7	B	18	2							B	15	1						
				C	1	42	C	1	13							C	2	16						
CQO e Oxidabilidade	13	5	pH CBO ₅ Cota Temperatura Condutividade	Classe	B	C	Classe	B	C	93.5%	91.2%	Anexo III.6	14	3-6-4-2	pH Volume Arm. CBO ₅	Classe	B	C	Classe	B	C	92.2%	93.9%	Anexo III.6
				B	34	4	B	17	2							B	19	1						
				C	1	38	C	1	14							C	1	12						
pH, CQO e Oxidabilidade	15	5	Cota Temperatura OD CBO ₅	Classe	B	C	Classe	B	C	87.2%	79.4%	Anexo III.7	16	3-5-2-2	Temperatura CBO ₅ Condutividade	Classe	B	C	Classe	B	C	85.2%	88.6%	Anexo III.7
				B	35	6	B	11	5							B	14	1						
				C	4	33	C	2	16							C	3	17						

Os modelos apresentados anteriormente fazem uso de parâmetros como a oxidabilidade e a CQO que, devido à sua natureza, impedem a aferição da qualidade da água em tempo real. Deste modo, procuraram-se obter novos modelos tendo em vista contornar esse inconveniente, tendo-se recorrido aos parâmetros meteorológicos e hidrométricos. Os modelos obtidos, as regras geradas e os pesos associados às conexões entre os nodos das camadas adjacentes encontram-se no Anexo IV, enquanto na Tabela 24 se compara a AD e a RNA obtida, apresentando-se a topologia da rede/profundidade da AD, os parâmetros utilizados como entrada, as matrizes de coincidências e a acuidade dos modelos, medida em termos de percentagem de acertos relativamente ao número de casos apresentados. A análise da Tabela 24 mostra que ambos os modelos usam, como entradas, as mesmas cinco variáveis (T_{10} , PP_{am} , volume armazenado na albufeira, HR_{am} e PP_5), apresentando acuidades relativamente elevadas (cerca de 80% e 90%, respectivamente para a AD e RNA). Contudo, verificam-se alguns erros de classificação no respeito à classe C. Na tentativa de resolver esta situação disponibilizou-se o parâmetro pH. A escolha deste parâmetro prende-se com duas razões fundamentais. Por um lado, como se mostrou anteriormente, influencia o desempenho dos modelos de classificação e, por outro, é um parâmetro que pode ser facilmente determinado no local, em contínuo, não inviabilizando a aferição da qualidade da água em tempo real.

Tabela 24 – Síntese dos modelos de previsão da qualidade da água da albufeira do Alvito utilizando os parâmetros hidrométricos e meteorológicos.

Parâmetros	Modelo	Profundidade/ /Topologia	Entradas	Matrizes de Coincidências						Acuidade	
				Conjunto de Treino			Conjunto de Teste			Conjunto de Treino	Conjunto de Teste
Hidrométricos e Meteorológicos	Árvore de Decisão	5	T_{10} PP_{am} Volume Arm. HR_{am} PP_5	Classe	B	C	Classe	B	C	81,4%	80,6%
				B	29	6	B	18	4		
				C	7	28	C	2	7		
	RNA	5-20-9-2	Volume Arm. PP_5 T_{10} HR_{am} PP_{am}	Classe	B	C	Classe	B	C	92,1%	93,9%
				B	34	3	B	20	0		
				C	2	24	C	2	11		

A AD obtida para previsão das classes de qualidade da água da albufeira do Alvito utilizando, para além do pH, os parâmetros hidrométricos e meteorológicos encontra-se na Figura 29. A sua análise mostra que a classificação é conseguida utilizando seis variáveis de entrada (pH, volume armazenado, PP_5 , T_{10} , T_{am} , PP_{am} e HR_{am}).

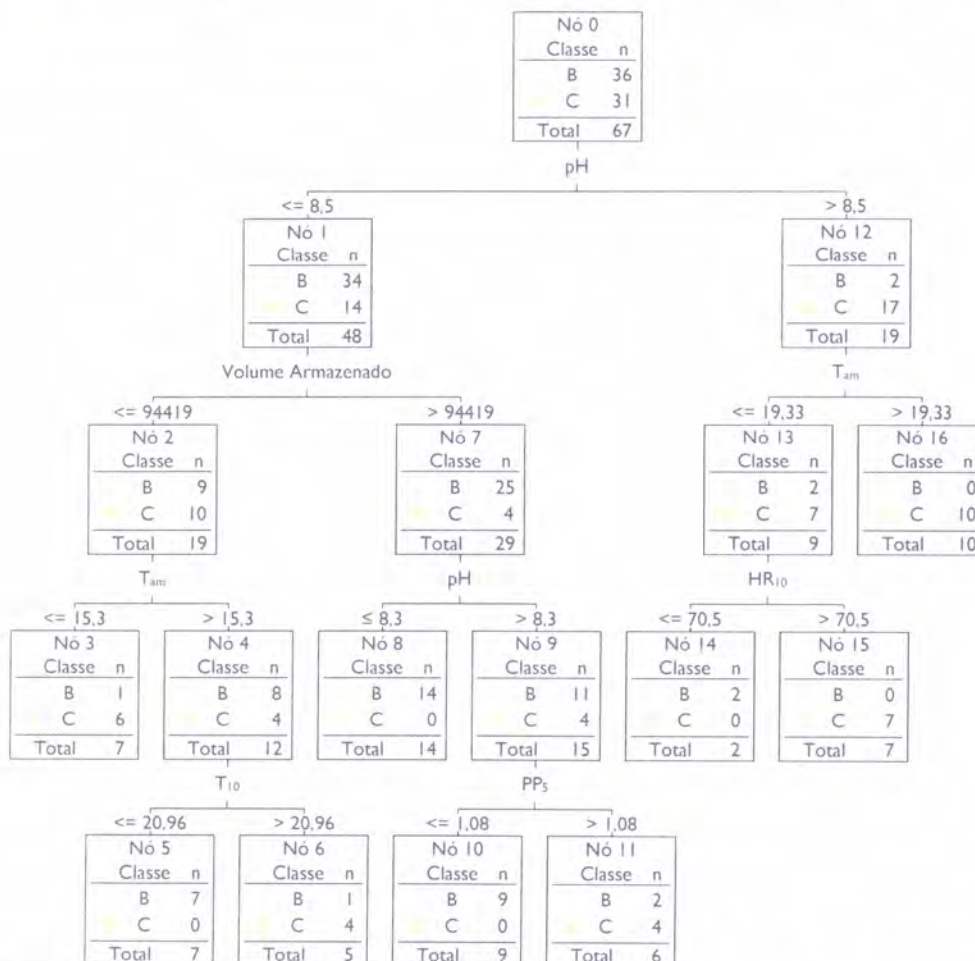


Figura 29 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito utilizando o pH e os parâmetros hidrométricos e meteorológicos.

A acuidade do modelo, medida em termos de percentagem de acertos relativamente ao número de casos apresentados, é de 94,0% (63 casos classificados correctamente em 67) para o conjunto de treino e de 93,8% (30 casos classificados correctamente em 32) para o conjunto de teste, Tabela 25. O facto dos acertos relativos em cada uma das classes serem semelhantes, quer para o conjunto de treino quer para o conjunto de teste, mostra que o modelo tem boa capacidade de generalização, não revelando indícios de sobre-ajustamento.

Tabela 25 – Matriz de coincidências relativa ao modelo apresentado na Figura 29.

Classe	Conjunto de Treino		Conjunto de Teste	
	B	C	B	C
B	32	4	20	2
C	0	31	0	10

No Anexo V apresenta-se o conjunto de regras relativas a este modelo. As regras com suporte mais elevado, para ambas as classes de qualidade, apresentam confiança um e são, respectivamente a regra 2 (Se $pH \leq 8,5$ e volume armazenado > 94419 e $pH \leq 8,3$ Então \rightarrow Classe B), cujo suporte é 14 e a regra 5 (Se $pH > 8,5$ e $T_{am} > 19,33$ Então \rightarrow Classe C), cujo o suporte é 10.

No que respeita às RNAs, a RNU seleccionada para prever as classes de qualidade apresenta a topologia 5-6-2-2, i.e. uma camada de entrada com cinco nodos, duas camadas intermédias com, respectivamente, seis e dois nodos e uma camada de saída com dois nodos, correspondentes às classes B e C, Figura 30. O modelo obtido faz uso, como variáveis de entrada, de cinco parâmetros (pH , volume armazenado HR_{am} , T_{am} e PP_{am}). Os pesos associados às conexões entre nodos de camadas adjacentes encontram-se no Anexo V, nas Tabelas V.2, V.3 e V.4.

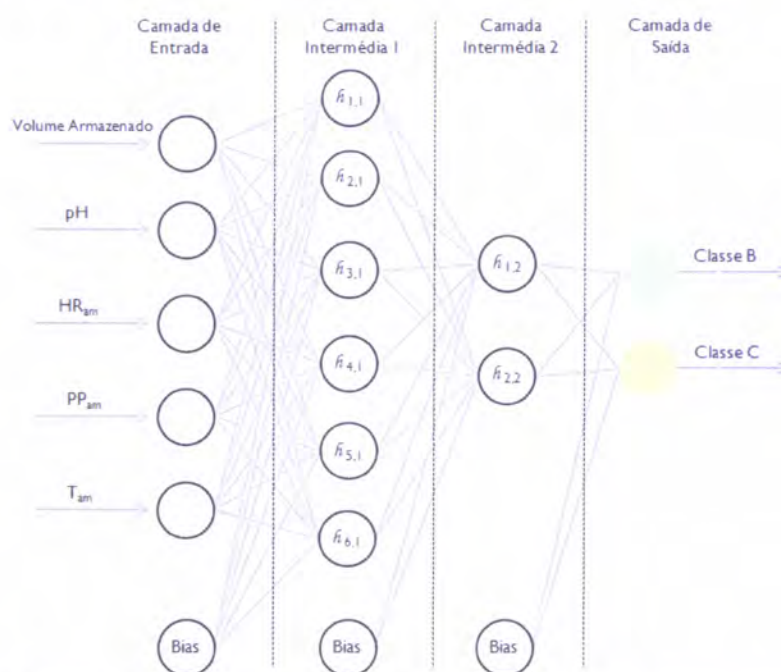


Figura 30 – RNA obtida para previsão da qualidade da água da albufeira do Alvito utilizando o pH e os parâmetros hidrométicos e meteorológicos.

A acuidade global do modelo, medida em termos de percentagem de acertos relativamente ao número de casos apresentados é de 94,4% (68 casos classificados correctamente em 72) para o conjunto de treino e de 94,3% (33 casos classificados correctamente em 35) para o conjunto de teste, Tabela 26. A semelhança entre a percentagem de respostas correctas relativamente a cada uma das classes, quer para o conjunto de treino quer para o conjunto de teste, mostra que o modelo tem boa capacidade de generalização, não revelando indícios de sobre-ajustamento.

Tabela 26 – Matriz de coincidências relativa ao modelo apresentado na Figura 30.

Classe	Conjunto de Treino		Conjunto de Teste	
	B	C	B	C
B	29	4	21	2
C	0	39	0	12

A análise conjunta da Tabela 25 e da Tabela 26 mostra que os modelos construídos com recurso a AD e a RNAs apresentam desempenhos similares. No que diz respeito às variáveis de entrada há a salientar, por um lado, o facto de ambos os modelos fazerem uso dos parâmetros pH e volume armazenado e, por outro, o facto de a RNA utilizar dados meteorológicos relativos ao período entre amostragens.

5.3.2 MODELOS DE PREVISÃO DA QUALIDADE DA ÁGUA – ALBUFEIRA DE ODIVELAS

A AD obtida para previsão das classes de qualidade da água da albufeira de Odivelas, tendo sido disponibilizados ao algoritmo todos os parâmetros da qualidade da água existentes na BD, encontra-se na Figura 31. A sua análise mostra que a classificação é conseguida utilizando seis variáveis de entrada (CQO, SST, Oxidabilidade, Condutividade, OD e Temperatura). O parâmetro CQO é o factor que mais influencia a classificação, secundado pelos SST e pela oxidabilidade, surgindo condutividade, o oxigénio dissolvido e a temperatura da água nos níveis mais afastados da raiz.

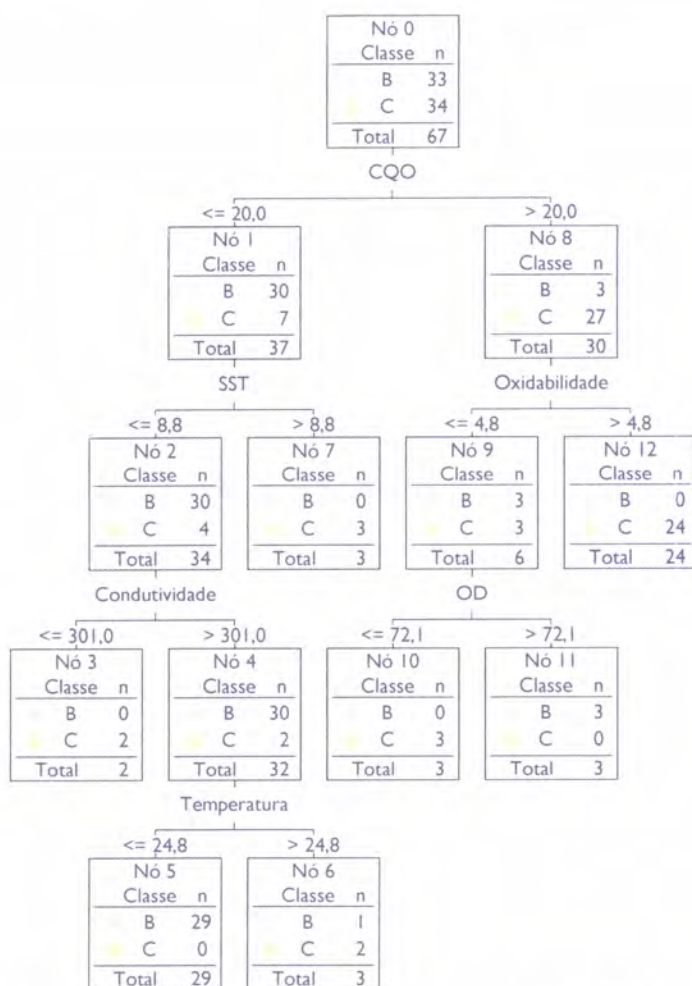


Figura 31 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas.

A acuidade do modelo, medida em termos de percentagem de acertos relativamente ao número de casos apresentados, é de 98,5% (66 casos classificados correctamente em 67) para o conjunto de treino e de 96,9% (31 casos classificados correctamente em 32) para o conjunto de teste, Tabela 27. O facto dos acertos relativos em cada uma das classes serem semelhantes, quer para o conjunto de treino quer para o conjunto de teste, mostra que o modelo tem boa capacidade de generalização, não revelando indícios de sobre-ajustamento.

Tabela 27 – Matriz de coincidência relativa ao modelo apresentado na Figura 31.

Classe	Conjunto de Treino		Conjunto de Teste	
	B	C	B	C
B	32	1	19	1
C	0	34	0	12

Na Tabela 28 apresentam-se as regras obtidas a partir deste modelo destacando-se, em cada uma delas, a confiança e o suporte.

Tabela 28 – Conjunto de regras relativo ao modelo apresentado na Figura 31.

Regras para a Classe B	Regras para a Classe C
<p>Regra 1 Se CQO \leq 20,0 e SST \leq 8,8 e Condutividade $>$ 301,0 e Temperatura \leq 24,8 Então \rightarrow B (29; 1,0)</p>	<p>Regra 1 Se CQO \leq 20,0 e SST \leq 8,8 e Condutividade \leq 301,0 Então \rightarrow C (2; 1,0)</p>
<p>Regra 2 Se CQO $>$ 20,0 e Oxidabilidade \leq 4,8 e Oxigénio Dissolvido $>$ 72,1 Então \rightarrow B (3; 1,0)</p>	<p>Regra 2 Se CQO \leq 20,0 e SST \leq 8,8 e Condutividade $>$ 301,0 e Temperatura $>$ 24,8 Então \rightarrow C (3; 0,67)</p>
	<p>Regra 3 Se CQO \leq 20,0 e SST $>$ 8,8 Então \rightarrow C (3; 1,0)</p>
	<p>Regra 4 Se CQO $>$ 20,0 e Oxidabilidade \leq 4,8 e Oxigénio Dissolvido \leq 72,1 Então \rightarrow C (3; 1,0)</p>
	<p>Regra 5 Se CQO $>$ 20,0 e Oxidabilidade $>$ 4,8 Então \rightarrow C (24; 1,0)</p>

As regras com suporte mais elevado, para ambas as classes de qualidade, apresentam confiança um e são, respectivamente, a regra 1 (Se CQO \leq 20,0 e SST \leq 8,8 e condutividade $>$ 301,0 e temperatura \leq 24,8 Então \rightarrow Classe B) cujo suporte é 29 e a regra 5 (Se CQO $>$ 20,0 e oxidabilidade $>$ 4,8 Então \rightarrow classe C) cujo suporte é 24.

No que respeita às RNAs optou-se, tal como anteriormente, pela utilização de RNUs. Foram construídas e avaliadas diversas topologias de rede, tendo o número óptimo de camadas e de nodos sido determinado por tentativa e erro. A RNU seleccionada para prever as classes de qualidade apresenta a topologia 4-14-5-2, i.e. uma camada de entrada com três nodos, duas camadas intermédias com, respectivamente, catorze e cinco nodos e uma camada de saída com dois nodos, correspondentes às classes B e C, Figura 32. O modelo obtido faz uso, como variáveis de entrada, de apenas quatro parâmetros (CQO, oxidabilidade, OD e SST), apesar de estarem disponíveis todos os parâmetros de qualidade da água existentes na BD. Os pesos associados às conexões entre nodos de camadas adjacentes encontram-se no Anexo VI, nas Tabelas VI.1, VI.2 e VI.3.

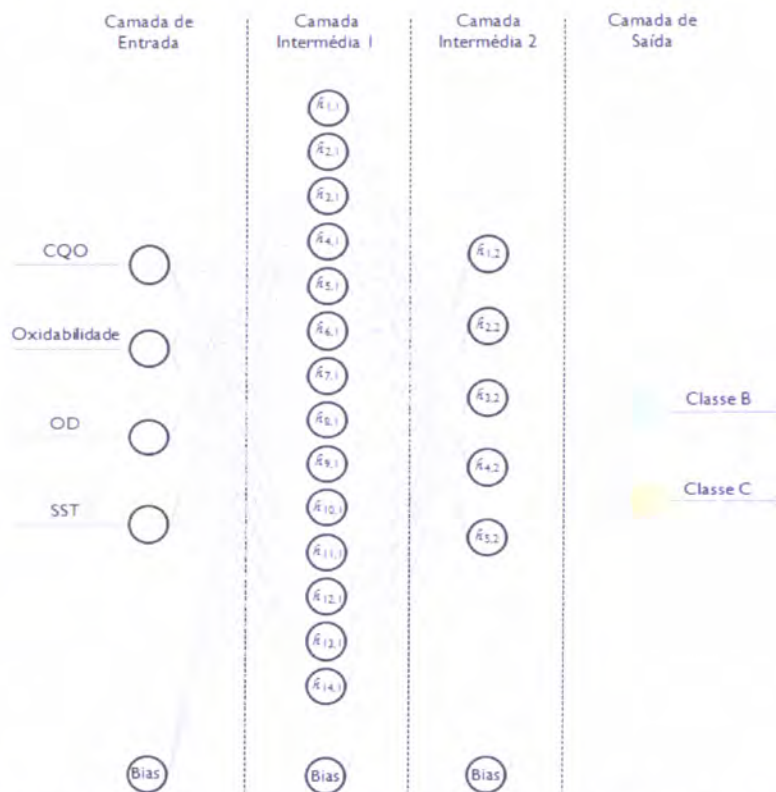


Figura 32 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas.

A acuidade global do modelo, medida em termos de percentagem de acertos relativamente ao número de casos apresentados é de 98,5% (66 casos classificados correctamente em 67) para o conjunto de treino e de 97,0% (32 casos classificados correctamente em 33) para o conjunto de teste, Tabela 29. A semelhança entre a percentagem de respostas correctas relativamente a cada uma das classes, quer para o

conjunto de treino quer para o conjunto de teste, mostra que o modelo tem boa capacidade de generalização, não revelando indícios de sobre-ajustamento.

Tabela 29 – Matriz de coincidências relativa ao modelo apresentado na Figura 32.

Classe	Conjunto de Treino		Conjunto de Teste	
	B	C	B	C
B	34	1	19	1
C	0	32	0	13

A análise conjunta da Tabela 27 e da Tabela 29 mostra que os modelos construídos com recurso a AD e a RNAs apresentam desempenhos similares. No que diz respeito às variáveis de entrada há a salientar o facto de ambos os modelos fazerem uso dos parâmetros CQO, SST, oxidabilidade e OD tendo, no entanto, o modelo construído com recurso a AD utilizado, ainda, os parâmetros condutividade e temperatura nos níveis mais afastados da raiz.

Na comparação dos dois modelos destaca-se, tal como no caso anterior, o facto de os erros cometidos apenas ocorrerem na previsão da classe B o que, do ponto de vista da utilização da água, não trás grandes desvantagens.

Tendo em vista, por um lado, avaliar a importância das variáveis CQO, SST, oxidabilidade e OD na previsão das classes de qualidade e, por outro lado, quais os parâmetros que poderiam ser utilizadas em sua substituição, foram gerados novos modelos tendo-se, numa primeira fase, retirado sucessivamente uma dessas variáveis. Em seguida, foram retiradas, em simultâneo a oxidabilidade e a CQO, depois foram removidas três variáveis tendo permanecido a oxidabilidade ou a CQO e, finalmente, foram retiradas as quatro em simultâneo. Os modelos obtidos, as regras geradas e os pesos associados às conexões entre os nodos das camadas adjacentes encontram-se no Anexo VII. Na Tabela 30 comparam-se as AD e as RNAs obtidas em cada uma das referidas situações apresentando-se a topologia da rede/profundidade da AD, os parâmetros utilizados como entrada, as matrizes de coincidências e a acuidade dos modelos, medida em termos de percentagem de acertos relativamente ao número de casos apresentados.

Tabela 23 – Síntese dos diversos modelos obtidos para previsão da qualidade da água da albufeira do Alvito.

Parâmetro Retido	Árvores de Decisão									Redes Neuronais Artificiais														
	Modelo	Profundidade	Entradas	Matrizes de Coincidências			Acuidade		Notas	Modelo	Topologia	Entradas	Matrizes de Coincidências			Acuidade		Notas						
				Conjunto de Treino	Conjunto de Teste		Conjunto de Treino	Conjunto de Teste					Conjunto de Treino	Conjunto de Teste		Conjunto de Treino	Conjunto de Teste							
Nenhum	1	4	Oxidabilidade pH CQO Temperatura	Classe	B	C	Classe	B	C	98.7%	97.0%	Figura 27	2	3-9-3-2	CQO Oxidabilidade pH	Classe	B	C	Classe	B	C	97.4%	97.1%	Figura 28
pH	3	5	CQO OD Temperatura CBO ₅ SST Condutividade	Classe	B	C	Classe	B	C	94.8%	88.6%	Anexo III.1	4	5-27-8-2	CQO OD Temperatura Oxidabilidade CBO ₅	Classe	B	C	Classe	B	C	94.7%	94.6%	Anexo III.1
CQO	5	4	pH Oxidabilidade OD Condutividade	Classe	B	C	Classe	B	C	94.8%	93.8%	Anexo III.2	6	4-29-17-2	pH Oxidabilidade Volume Arm. Condutividade	Classe	B	C	Classe	B	C	96.2%	94.1%	Anexo III.2
Oxidabilidade	7	4	CQO pH Cota Temperatura Condutividade	Classe	B	C	Classe	B	C	96.1%	97.0%	Anexo III.3	8	5-4-4-2	CQO pH Volume Arm. Temperatura Condutividade	Classe	B	C	Classe	B	C	96.3%	97.1%	Anexo III.3
pH e CQO	9	4	Oxidabilidade Condutividade CBO ₅ Temperatura	Classe	B	C	Classe	B	C	92.4%	85.3%	Anexo III.4	10	3-4-3-2	Oxidabilidade Volume Arm. Temperatura	Classe	B	C	Classe	B	C	94.9%	93.9%	Anexo III.4
pH e Oxidabilidade	11	5	CQO Temperatura SST OD	Classe	B	C	Classe	B	C	90%	91.2%	Anexo III.5	12	6-10-6-2	CQO OD Volume Arm. Temperatura Condutividade CBO ₅	Classe	B	C	Classe	B	C	92.5%	91.2%	Anexo III.5
CQO e Oxidabilidade	13	5	pH CBO ₅ Cota Temperatura Condutividade	Classe	B	C	Classe	B	C	93.5%	91.2%	Anexo III.6	14	3-6-4-2	pH Volume Arm. CBO ₅	Classe	B	C	Classe	B	C	92.2%	93.9%	Anexo III.6
pH, CQO e Oxidabilidade	15	5	Cota Temperatura OD CBO ₅	Classe	B	C	Classe	B	C	87.2%	79.4%	Anexo III.7	16	3-5-2-2	Temperatura CBO ₅ Condutividade	Classe	B	C	Classe	B	C	85.2%	88.6%	Anexo III.7

Tabela 30 – Síntese dos diversos modelos obtidos para previsão da qualidade da água da albufeira de Odivelas.

Parâmetro Retornado	Árvores de Decisão										Redes Neurais Artificiais																																					
	Modelo	Profundidade	Entradas	Matrizes de Coincidências						Acuidade		Nota	Modelo	Topologia	Entradas	Matrizes de Coincidências						Acuidade		Nota																								
				Conjunto de Treino			Conjunto de Teste			Conjunto de Treino	Conjunto de Teste					Conjunto de Treino			Conjunto de Teste			Conjunto de Treino	Conjunto de Teste																									
Nenhum	17	4	CQO Oxidabilidade Temperatura Condutividade SST OD	Classe	B	C	Classe	B	C				18	4-14-5-2	CQO Oxidabilidade OD SST	Classe	B	C	Classe	B	C				18	4-14-5-2	CQO Oxidabilidade OD SST	Classe	B	C	Classe	B	C	98.5%	96.9%	Figura 31	18	4-14-5-2	CQO Oxidabilidade OD SST	Classe	B	C	Classe	B	C	98.5%	97.0%	Figura 32
CQO	19	4	Temperatura Oxidabilidade SST OD	Classe	B	C	Classe	B	C				20	3-4-5-2	Oxidabilidade OD Temperatura	Classe	B	C	Classe	B	C				20	3-4-5-2	Oxidabilidade OD Temperatura	Classe	B	C	Classe	B	C	91.2%	90.6%	Anexo VII.1	20	3-4-5-2	Oxidabilidade OD Temperatura	Classe	B	C	Classe	B	C	92.4%	90.0%	Anexo VII.1
Oxidabilidade	21	2	CQO SST pH	Classe	B	C	Classe	B	C				22	5-30-13-2	CQO SST Temperatura OD pH	Classe	B	C	Classe	B	C				22	5-30-13-2	CQO SST Temperatura OD pH	Classe	B	C	Classe	B	C	91.0%	90.0%	Anexo VII.2	22	5-30-13-2	CQO SST Temperatura OD pH	Classe	B	C	Classe	B	C	91.0%	90.3%	Anexo VII.2
SST	23	3	CQO Oxidabilidade	Classe	B	C	Classe	B	C				24	6-9-8-2	CQO Oxidabilidade Temperatura pH Cota OD	Classe	B	C	Classe	B	C				24	6-9-8-2	CQO Oxidabilidade Temperatura pH Cota OD	Classe	B	C	Classe	B	C	95.5%	93.3%	Anexo VII.3	24	6-9-8-2	CQO Oxidabilidade Temperatura pH Cota OD	Classe	B	C	Classe	B	C	95.6%	93.6%	Anexo VII.3
OD	25	5	CQO Oxidabilidade pH SST Temperatura	Classe	B	C	Classe	B	C				26	5-10-4-2	CQO Oxidabilidade Temperatura SST Volume Arm.	Classe	B	C	Classe	B	C				26	5-10-4-2	CQO Oxidabilidade Temperatura SST Volume Arm.	Classe	B	C	Classe	B	C	96.9%	93.5%	Anexo VII.4	26	5-10-4-2	CQO Oxidabilidade Temperatura SST Volume Arm.	Classe	B	C	Classe	B	C	93.4%	91.4%	Anexo VII.4
CQO e Oxidabilidade	27	4	SST OD Temperatura Volume Arm.	Classe	B	C	Classe	B	C				28	6-6-3-2	OD Temperatura Volume Arm. pH Cota SST	Classe	B	C	Classe	B	C				28	6-6-3-2	OD Temperatura Volume Arm. pH Cota SST	Classe	B	C	Classe	B	C	89.2%	87.1%	Anexo VII.5	28	6-6-3-2	OD Temperatura Volume Arm. pH Cota SST	Classe	B	C	Classe	B	C	85.9%	86.2%	Anexo VII.5
SST, OD e Oxidabilidade	29	4	CQO Cota pH	Classe	B	C	Classe	B	C				30	5-5-1-2	CQO Cota Volume Arm. pH Temperatura	Classe	B	C	Classe	B	C				30	5-5-1-2	CQO Cota Volume Arm. pH Temperatura	Classe	B	C	Classe	B	C	91.0%	90.3%	Anexo VII.6	30	5-5-1-2	CQO Cota Volume Arm. pH Temperatura	Classe	B	C	Classe	B	C	90.8%	90.3%	Anexo VII.6
SST, OD e CQO	31	5	Oxidabilidade Temperatura pH Cota Volume Arm.	Classe	B	C	Classe	B	C				32	5-5-6-2	Oxidabilidade Temperatura pH Cota Volume Arm.	Classe	B	C	Classe	B	C				32	5-5-6-2	Oxidabilidade Temperatura pH Cota Volume Arm.	Classe	B	C	Classe	B	C	91.0%	90.6%	Anexo VII.7	32	5-5-6-2	Oxidabilidade Temperatura pH Cota Volume Arm.	Classe	B	C	Classe	B	C	90.8%	89.7%	Anexo VII.7
CQO, Oxidabilidade, SST e OD	33	2	Temperatura Volume Arm.	Classe	B	C	Classe	B	C				34	4-3-2-2	Temperatura pH Cota Volume Arm.	Classe	B	C	Classe	B	C				34	4-3-2-2	Temperatura pH Cota Volume Arm.	Classe	B	C	Classe	B	C	78.8%	76.7%	Anexo VII.8	34	4-3-2-2	Temperatura pH Cota Volume Arm.	Classe	B	C	Classe	B	C	70.2%	70.0%	Anexo VII.8

A análise da Tabela 30 mostra que no caso de se retirar apenas a CQO ou a oxidabilidade conseguem-se modelos de previsão com desempenho inferior, na ordem de 90% de classificações correctas, tanto para o conjunto de treino como para o conjunto de teste, independentemente da técnica de *Data Mining* utilizada (AD ou RNAs).

No caso de se retirar os SST ou o OD a acuidade dos modelos é bastante próxima da obtida para os modelos iniciais. Deste modo, independentemente do algoritmo utilizado, os parâmetros CQO e oxidabilidade são aqueles que parecem mais influenciar a classificação. Outro aspecto que ressalta da análise da Tabela 30 relaciona-se com o facto de o número de acertos relativos à classe B permanecer praticamente inalterado e, portanto, a perda de acuidade se situar ao nível da classe C. A única excepção ocorre quando é retirada a oxidabilidade, em que se verifica diminuição da acuidade tanto na classe B como na classe C (modelos 21 e 22). Quando se retira, em simultâneo, a CQO e a oxidabilidade (modelos 27 e 28) a acuidade dos modelos diminui bastante, o que vem confirmar os resultados anteriores que apontam estas variáveis como as que mais influenciam a classificação.

Retirando, simultaneamente, os SST, o OD e a CQO ou a oxidabilidade (modelos 29, 30, 31 e 32) o desempenho é semelhante ao obtido retirando apenas a CQO (modelos 19 e 20) ou a oxidabilidade (modelos 21 e 22). Do exposto resulta que os SST e o OD não são muito relevantes, do ponto de vista da acuidade, desde que uma das variáveis de entrada seja a CQO ou a oxidabilidade. Quando se retira as quatro variáveis em conjunto os modelos obtidos (modelo 33 e 34) apresentam, como seria de esperar, desempenhos fracos, principalmente na previsão da classe C.

A discussão dos resultados apresentados na Tabela 30 deixa transparecer que os parâmetros que mais influenciam a capacidade preditiva dos modelos são a CQO e a oxidabilidade, independentemente de se tratar de AD ou RNAs. Deste modo, a aferição da qualidade da água em tempo real fica comprometida em virtude de não poderem ser determinados no local, necessitando das etapas de recolha, preservação e transporte das amostras até ao laboratório.

Tal como no caso da albufeira do Alvito, procuraram obter-se novos modelos utilizando, como variáveis de entrada, os parâmetros meteorológicos e hidrométricos.

Os modelos obtidos, as regras geradas e os pesos associados às conexões entre os nodos das camadas adjacentes encontram-se no Anexo VIII, enquanto na Tabela 31 se compara a AD e a RNA obtida, apresentando-se a topologia da rede/profundidade da AD, os parâmetros utilizados como entrada, as matrizes de coincidências e a acuidade dos modelos, medida em termos de percentagem de acertos relativamente ao número de casos apresentados.

Tabela 31 – Síntese dos modelos de previsão da qualidade da água da albufeira de Odivelas utilizando os parâmetros hidrométricos e meteorológicos.

Parâmetros	Modelo	Profundidade/ /Topologia	Entradas	Matrizes de Coincidências						Acuidade	
				Conjunto de Treino			Conjunto de Teste			Conjunto de Treino	Conjunto de Teste
Hidrométricos e Meteorológicos	Árvore de Decisão	5	T _{am} Volume Arm. VV ₅ Rad ₅ HR ₅	Classe	B	C	Classe	B	C	85,9%	86,4%
				B	20	5	B	9	1		
				C	3	29	C	2	10		
	RNA	4-4-2-2	T ₁₀ HR _{am} Rad ₁₀ Rad ₅	Classe	B	C	Classe	B	C	86,2%	87,0%
				B	18	6	B	10	1		
				C	2	32	C	2	10		

A análise da Tabela 31 mostra que a AD usa, como variáveis de entradas, a T_{am}, o volume armazenado, a VV₅, a Rad₅ e a HR₅. Em termos de desempenho apresenta uma acuidade de 85,9 % para o conjunto de treino e de 86,4% para o conjunto de teste, verificando-se, no entanto, alguns erros de classificação no respeito à classe C. No que concerne à utilização de RNAs, as variáveis de entrada do modelo são a T₁₀, HR_{am}, a Rad₁₀ e a Rad₅. Em termos de desempenho, apresenta acuidade semelhante à AD (86,2% para o conjunto de treino e 87,0 % para o conjunto de teste) verificando-se, também, alguns erros de classificação no respeito à classe C.

Na tentativa de resolver esta situação disponibilizaram-se os parâmetros OD e pH atendendo ao facto de poderem ser facilmente determinado no local, em contínuo, não inviabilizando a aferição da qualidade da água em tempo real e terem sido utilizados, como parâmetros de entrada, em alguns dos modelos apresentados na Tabela 31.

A AD obtida para previsão das classes de qualidade da água da albufeira de Odivelas utilizando, para além do OD e do pH, os parâmetros hidrométricos e meteorológicos encontra-se na Figura 33. A sua análise mostra que a classificação é conseguida utilizando oito variáveis de entrada (pH, VV_5 , PP_5 , PP_{am} , HR_{10} , VV_{10} , T_5 e HR_5), destacando-se o facto de serem preferencialmente utilizados os dados meteorológicos referentes aos cinco e dez dias anteriores à amostragem.

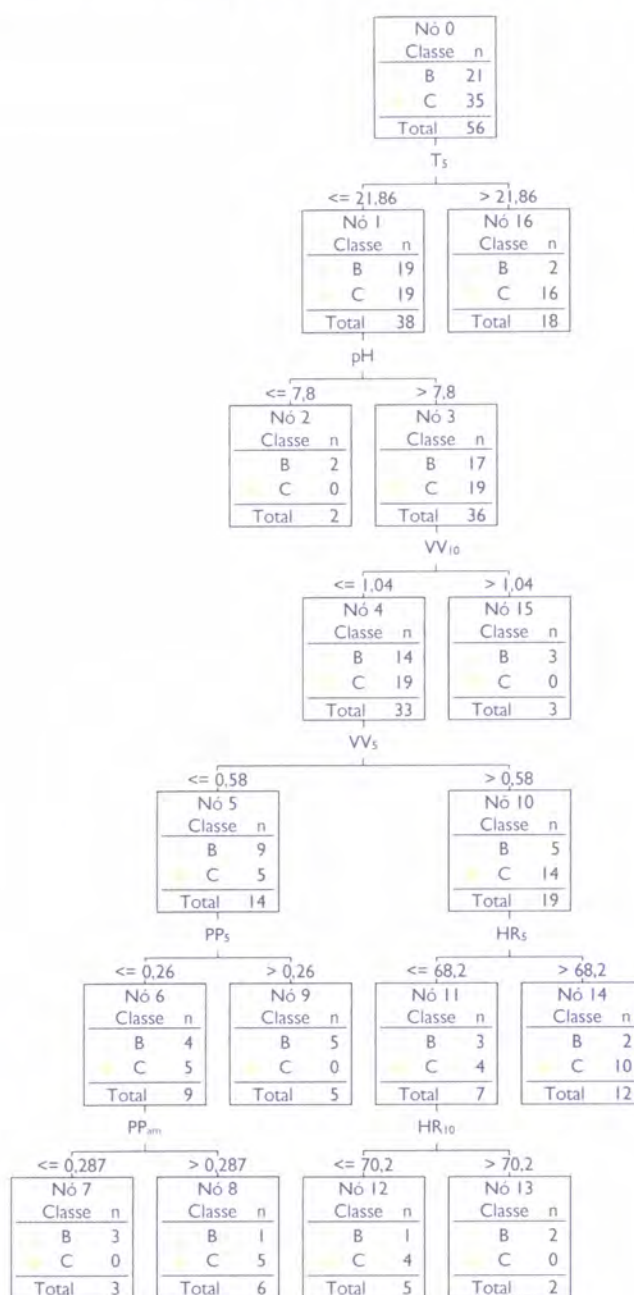


Figura 33 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas utilizando o pH, o OD e os parâmetros hidrométricos e meteorológicos.

A acuidade do modelo, medida em termos de percentagem de acertos relativamente ao número de casos apresentados, é de 89,3% (50 casos classificados correctamente em 56) para o conjunto de treino e de 88,0% (22 casos classificados correctamente em 25) para o conjunto de teste, Tabela 32. O facto dos acertos relativos em cada uma das classes serem semelhantes, quer para o conjunto de treino quer para o conjunto de teste, mostra que o modelo tem boa capacidade de generalização, não revelando indícios de sobre-ajustamento.

Tabela 32 – Matriz de coincidências relativa ao modelo apresentado na Figura 33.

Classe	Conjunto de Treino		Conjunto de Teste	
	B	C	B	C
B	15	6	9	3
C	0	35	0	13

No Anexo IX apresenta-se o conjunto de regras relativas a este modelo, destacando-se a regra 3 (Se $T5 \leq 21,86$ e $pH > 7,8$ e $VV10 \leq 1,04$ e $VV5 \leq 0,58$ e $PP5 > 0,26$ Então \rightarrow Classe B) que é, das regras geradas para a classe B, a que exhibe suporte mais elevado, 5. No que diz respeito à classe C, é a regra 4 (Se $T5 > 21,86$ Então \rightarrow Classe C), que apresenta suporte mais elevado, 18.

Relativamente às RNAs, a RNU seleccionada para prever as classes de qualidade apresenta a topologia 7-4-2-2, i.e. uma camada de entrada com sete nodos, duas camadas intermédias com, respectivamente, quatro e dois nodos e uma camada de saída com dois nodos, correspondentes às classes B e C, Figura 34. O modelo obtido faz uso, como variáveis de entrada, de sete parâmetros (HR_{am} , VV_5 , T_{am} , PP_{am} , Rad_{30} , volume armazenado e pH). Os pesos associados às conexões entre nodos de camadas adjacentes encontram-se no Anexo IX, nas Tabelas IX.2, IX.3 e IX.4.

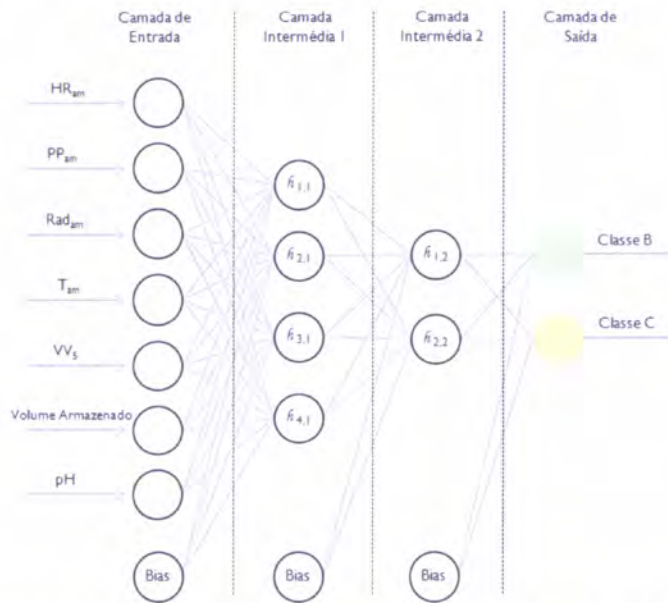


Figura 34 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas utilizando o pH, o OD e os parâmetros hidrométicos e meteorológicos.

A acuidade global do modelo, medida em termos de percentagem de acertos relativamente ao número de casos apresentados é de 91,1% (51 casos classificados correctamente em 56) para o conjunto de treino e de 91,7% (22 casos classificados correctamente em 24) para o conjunto de teste, Tabela 33. A semelhança entre a percentagem de respostas correctas relativamente a cada uma das classes, quer para o conjunto de treino quer para o conjunto de teste, mostra que o modelo tem boa capacidade de generalização, não revelando indícios de sobre-ajustamento.

Tabela 33 – Matriz de coincidências relativa ao modelo apresentado na Figura 34.

Classe	Conjunto de Treino		Conjunto de Teste	
	B	C	B	C
B	16	5	9	2
C	0	35	0	13

A análise conjunta da Tabela 32 e da Tabela 33 mostra que os modelos construídos com recurso a AD e a RNAs apresentam desempenhos similares, embora a acuidade do primeiro seja ligeiramente inferior. No que diz respeito às variáveis de entrada há a salientar, por um lado, o facto de ambos os modelos fazerem uso dos parâmetros pH, VV_5 e PP_{2m} e, por outro, o facto de a RNA utilizar, preferencialmente, dados

meteorológicos relativos ao período entre amostragens enquanto a árvore de decisão, como foi assinalado anteriormente, utilizar preferencialmente os dados meteorológicos referentes aos cinco e dez dias anteriores à amostragem.

Um outro aspecto que deve ser destacado prende-se com o facto de ambos os algoritmos não terem seleccionado o parâmetro OD. Este facto está em concordância com os resultados apresentados e discutidos anteriormente e que apontam para a pouca relevância deste parâmetro no que concerne à classificação da massa de água.

5.4 APRESENTAÇÃO, DISCUSSÃO E AVALIAÇÃO DOS MODELOS DE SEGMENTAÇÃO

Tendo em vista procurar tendências, regularidades ou padrões nos dados físico-químicos de qualidade da água referentes às albufeiras em estudo utilizaram-se técnicas de aprendizagem não supervisionada, tendo sido construídos modelos de segmentação utilizando a estratégia *k-means*. As variáveis de entrada utilizadas foram os parâmetros físico-químicos seleccionados pelos modelos apresentados na secção anterior, nomeadamente CQO, condutividade, oxidabilidade, OD, pH, SST e temperatura da água.

Nesta técnica, como foi referido anteriormente, é necessário fornecer o parâmetro *k*, que corresponde ao número de segmentos a formar, tendo-se utilizado valores compreendidos entre dois e cinco. Quando se utilizou valores de *k* superiores a três alguns segmentos eram constituídos por um número de exemplos muito reduzido não tendo, por este motivo sido considerados. Este facto pode ser devido ao número de casos existentes na BD não ser muito elevado.

Na fase de construção de modelos explicativos da segmentação verificou-se que os modelos de três segmentos conduziram a regras cujo suporte era baixo e, por essa razão, também não foram considerados. Em seguida apresentam-se os modelos de segmentação para $k = 2$, caracterizando-se os segmentos formados através dos seus centros de gravidade e procurando-se estabelecer associações entre os segmentos e o mês em que foi feita a amostragem.

Tendo em vista atribuir um novo caso a um determinado segmento geraram-se modelos explicativos da segmentação, tendo-se optado por utilizar AD. Os motivos subjacentes a essa escolha prendem-se com o seu elevado poder expressivo e com a facilidade de interpretação e de utilização.

5.4.1 MODELOS DE SEGMENTAÇÃO – ALBUFEIRA DO ALVITO

A aplicação da estratégia *k-means* aos parâmetros físico-químicos anteriormente referidos originou dois segmentos com, respectivamente, 36 e 45 registos. Os centros de gravidade de cada um dos segmentos encontram-se na Tabela 34. As variáveis cujo índice de importância é próximo da unidade encontram-se destacadas a negrito. O referido índice é calculado como um menos o valor de significância do teste *t-Student*. Neste caso, a hipótese nula considera que as médias das variáveis são as mesmas para todos os segmentos. Deste modo, se a variável influenciar a definição do segmento a hipótese nula é rejeitada, o valor da significância do teste *t-Student* é próximo de zero e a importância é próxima da unidade.

Tabela 34 – Centros de gravidade dos segmentos obtidos com $k = 2$.

Variável	Segmento 1	Segmento 2
Condutividade	306,639 ± 47,974	331,689 ± 57,139
CQO	20,656 ± 8,478	18,044 ± 6,301
OD	100,886 ± 16,672	104,729 ± 24,477
Oxidabilidade	4,889 ± 1,003	4,282 ± 0,64
pH	8,148 ± 0,19	8,453 ± 0,231
SST	5,572 ± 4,327	6,567 ± 7,845
Temperatura	14,389 ± 2,763	22,92 ± 2,994

A análise da Tabela 34 mostra que a oxidabilidade é a variável mais importante na definição do segmento 1, enquanto que para o segmento 2 as variáveis mais importantes são o pH e a temperatura. Na verdade o segmento 2 concentra as amostras cuja temperatura é mais elevada e portanto deve estar relacionada com os meses mais quentes do ano.

Na perspectiva de testar a hipótese anteriormente formulada construiu-se o gráfico apresentado na Figura 35 em que a espessura das linhas denota a maior ou menor associação entre as variáveis.

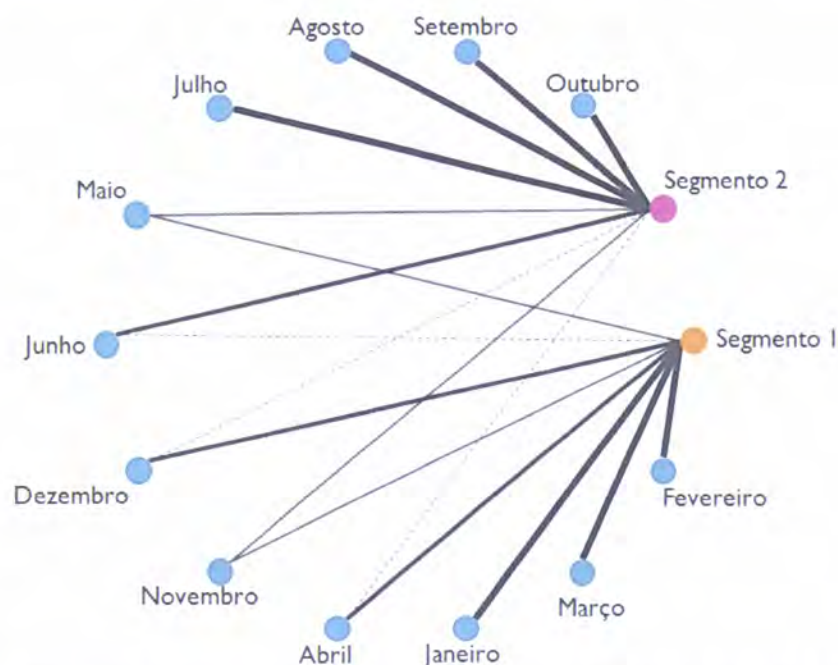


Figura 35 – Associações entre os segmentos e os meses do ano em que foi realizada a amostragem.

A análise da Figura 35 mostra que o segmento 1 é formado pelas amostras colhidas nos meses de Janeiro, Fevereiro e Março e ainda por grande parte das amostras colhidas em Dezembro e Abril. O segmento 2, por seu lado, é constituído pelas amostras colhidas nos meses de Julho, Agosto, Setembro e Outubro e grande parte das amostras referentes a Junho. Relativamente às amostras colhidas nos meses de Maio e Novembro dividem-se pelos dois segmentos.

Do exposto ressalta que os parâmetros físico-químicos estão, de alguma forma, relacionados com as condições meteorológicas. Deste modo, na tentativa de se gerarem modelos explicativos dos segmentos formados, foram seguidas duas estratégias. Na estratégia 1 as árvores de decisão são construídas utilizando, como variáveis de entradas, os parâmetros físico-químicos que estiveram na base do modelo de segmentação. Na estratégia 2 os parâmetros físico-químicos são substituídos pelos parâmetros meteorológicos e hidrométricos. Na Figura 36 apresentam-se as árvores de decisão obtidas para cada uma das situações encontrando-se, na Tabela 36, as regras geradas para cada modelo e na Tabela 35 as respectivas matrizes de coincidência.

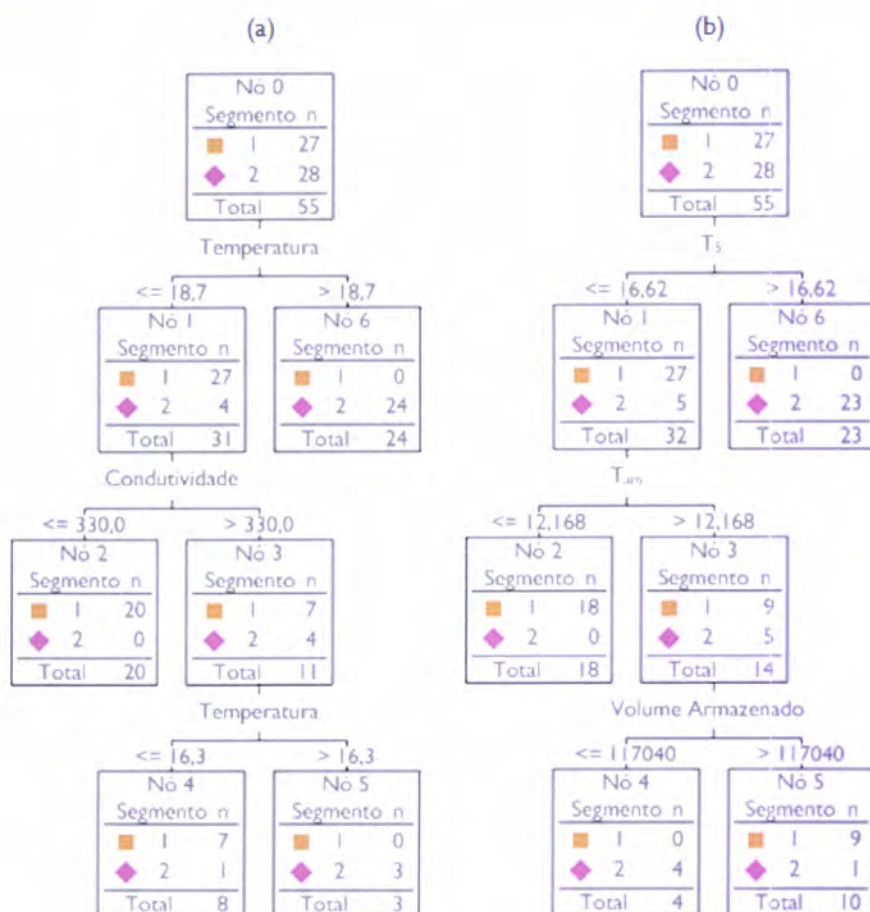


Figura 36 – Árvore de decisão explicativa do modelo de segmentação apresentado na Tabela 36, seguindo a estratégia 1 (a) e a estratégia 2 (b).

A análise da Tabela 35 mostra que a acuidade dos modelos, medida em termos de percentagem de respostas correcta, é bastante elevada quer para o conjunto de treino quer para o conjunto de teste e independente da estratégia escolhida. As regras obtidas (Tabela 36) têm suporte elevado e, para a maior parte delas, a confiança é um.

Tabela 35 – Matrizes de coincidências relativas aos modelos apresentados na Figura 36.

	Segmento	Matriz de Coincidências				Acuidade	
		Conjunto de Treino		Conjunto de Teste		Conjunto de Treino	Conjunto de Teste
		1	2	1	2	1	2
Estratégia 1	1	27	0	8	1	98,2%	96,2%
	2	1	27	0	17		
Estratégia 2	1	27	0	13	0	98,2%	96,6%
	2	1	27	1	15		

Tabela 36 – Conjunto de regras relativo aos modelos apresentados na Figura 36.

	Regras para o Segmento 1	Regras para o Segmento 2
Estratégia 1	Regra 1 Se Temperatura \leq 18,7 e Condutividade \leq 330,0 Então \rightarrow Segmento 1 (20; 1,0)	Regra 1 Se Temperatura \leq 18,7 e Condutividade $>$ 330,0 e Temperatura $>$ 16,3 Então \rightarrow Segmento 2 (3; 1,0)
	Regra 2 Se Temperatura \leq 18,7 e Condutividade $>$ 330,0 e Temperatura \leq 16,3 Então \rightarrow Segmento 1 (8; 0,88)	Regra 2 Se Temperatura $>$ 18,7 Então \rightarrow Segmento 2 (24; 1,0)
Estratégia 2	Regra 1 Se $T_5 \leq$ 16,62 e $T_{am} \leq$ 12,168 Então \rightarrow Segmento 1 (18; 1,0)	Regra 1 Se $T_5 \leq$ 16,62 e $T_{am} >$ 12,168 e Volume Armazenado \leq 117040 Então \rightarrow Segmento 2 (4; 1,0)
	Regra 2 Se $T_5 \leq$ 16,62 e $T_{am} >$ 12,168 e Volume Armazenado $>$ 117040 Então \rightarrow Segmento 1 (10; 0,9)	Regra 2 Se $T_5 >$ 16,62 Então \rightarrow Segmento 2 (23; 1,0)

5.4.2 MODELOS DE SEGMENTAÇÃO – ALBUFEIRA DE ODIVELAS

A aplicação da estratégia k-means aos parâmetros físico-químicos referentes à albufeira de Odivelas originou dois segmentos com, respectivamente, 52 e 28 registos. Os centros de gravidade de cada um dos segmentos formados encontram-se na Tabela 37, realçando-se as variáveis cujo índice de importância é próximo da unidade.

A análise da Tabela 37 mostra que, tal como anteriormente, um dos segmentos agrupa as amostras cuja temperatura da água é mais elevada (segmento 1) sendo as variáveis mais importantes na sua definição a temperatura e o pH. No que concerne ao segmento 2 a variável mais importante é a oxidabilidade e agrupa as amostras cuja temperatura é mais baixa.

Tabela 37 – Centros de gravidade dos segmentos obtidos com $k = 2$.

Variável	Segmento 1	Segmento 2
Condutividade	395,769 ± 52,244	383,107 ± 56,582
CQO	23,471 ± 8,588	20,271 ± 6,227
OD	83,779 ± 12,012	76,85 ± 10,11
Oxidabilidade	5,466 ± 0,984	5,714 ± 1,189
pH	8,357 ± 0,262	7,954 ± 0,175
SST	6,563 ± 5,855	7,729 ± 12,473
Temperatura	22,946 ± 3,127	14,989 ± 2,674

Na Figura 37 encontram-se representadas as associações entre os segmentos formados e os meses do ano em que foi realizada a amostragem. O segmento 2 inclui as amostras colhidas em Dezembro e Fevereiro e grande parte das referentes ao mês de Janeiro. O segmento 1, por seu lado, inclui as amostras colhidas em Julho, Agosto, Setembro e Outubro e grande parte das recolhidas em Abril e Junho. Relativamente às amostras referentes aos meses de Março, Maio e Novembro dividem-se pelos dois segmentos.

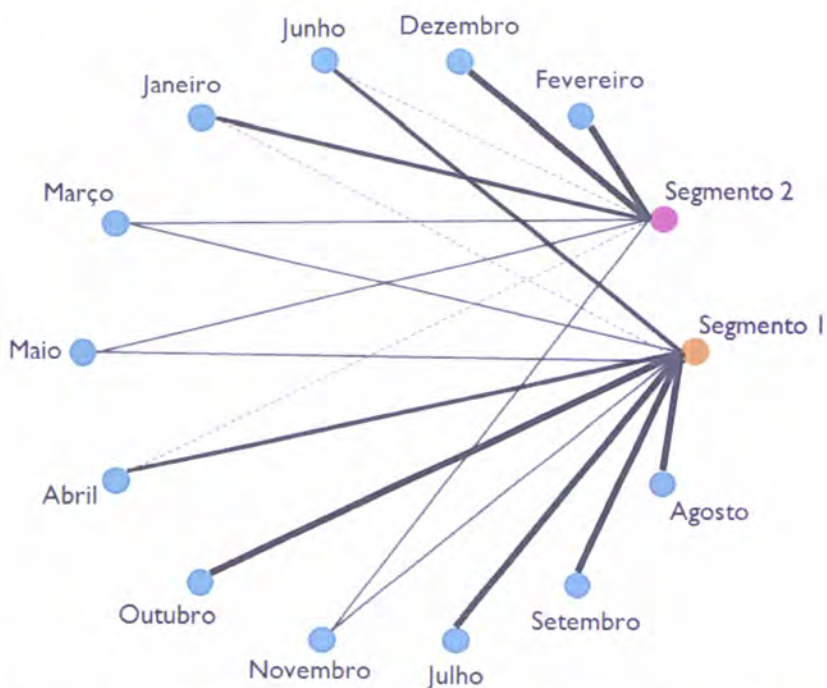


Figura 37 – Associações entre os segmentos e os meses do ano em que foi realizada a amostragem.

Tal como anteriormente, foi possível explicar a segmentação quer recorrendo às variáveis utilizadas na sua construção (Figura 38 a) quer utilizando as variáveis

hidrométricas e meteorológicas (Figura 38 b). As regras geradas, tal como no caso da albufeira do Alvito, apresentam suporte elevado e confiança próxima da unidade (Tabela 38).



Figura 38 – Árvore de decisão explicativa do modelo de segmentação apresentado na Tabela 38 seguindo a estratégia 1 (a) e a estratégia 2 (b).

No que concerne à avaliação dos modelos (Tabela 39), verifica-se que ambos apresentam acuidade elevada, superior a 98% e 96%, respectivamente, para o conjunto de treino e para o conjunto de teste, independentemente da estratégia escolhida. O facto dos acertos relativos em cada um dos segmentos ser semelhante mostra que o modelo não revela indícios de sobre-ajustamento.

Tabela 38 – Conjunto de regras relativo aos modelos apresentados na Figura 38.

		Regras para o Segmento 1		Regras para o Segmento 2	
Estratégia 1	Regra 1 Se Temperatura $\leq 20,0$ e pH $\leq 8,4$ e OD $> 83,9$ e pH $> 8,03$ Então \rightarrow Segmento 1 (3; 1,0)	Regra 3 Se Temperatura > 20 Então \rightarrow Segmento 1 (29; 1,0)	Regra 1 Se Temperatura $\leq 20,0$ e pH $\leq 8,4$ e OD $\leq 83,9$ Então \rightarrow Segmento 2 (28; 1,0)		
	Regra 2 Se Temperatura $\leq 20,0$ e pH $> 8,4$ Então \rightarrow Segmento 1 (4; 1,0)		Regra 2 Se Temperatura $\leq 20,0$ e pH $\leq 8,4$ e OD $> 83,9$ e pH $\leq 8,03$ Então \rightarrow Segmento 2 (7; 0,86)		
Estratégia 2	Regra 1 Se $T_{am} \leq 17,984$ e $Rad_5 > 2736,2$ e $HR_{10} > 68,5$ e $PP_5 \leq 3,98$ e $T_{10} > 10,26$ Então \rightarrow Segmento 1 (9; 1,0)		Regra 1 Se $T_{am} \leq 17,984$ e $Rad_5 \leq 2736,2$ Então \rightarrow Segmento 2 (26; 1,0)	Regra 3 Se $T_{am} \leq 17,984$ e $Rad_5 > 2736,2$ e $HR_{10} > 68,5$ e $PP_5 \leq 3,98$ e $T_{10} \leq 10,26$ Então \rightarrow Segmento 2 (2; 1,0)	
	Regra 2 Se $T_{am} > 17,984$ Então \rightarrow Segmento 1 (26; 1,0)		Regra 2 Se $T_{am} \leq 17,984$ e $Rad_5 > 2736,2$ e $HR_{10} \leq 68,5$ Então \rightarrow Segmento 2 (6; 0,83)	Regra 4 Se $T_{am} \leq 17,984$ e $Rad_5 > 2736,2$ e $HR_{10} > 68,5$ e $PP_5 > 3,98$ Então \rightarrow Segmento 2 (3; 1,0)	

Tabela 39 – Matrizes de coincidências relativas aos modelos apresentados na Figura 38.

	Segmento	Matriz de Coincidências				Acuidade	
		Conjunto de Treino		Conjunto de Teste		Conjunto de Treino	Conjunto de Teste
		1	2	1	2	1	2
Estratégia 1	1	36	1	15	1	98,6%	96,9%
	2	0	34	0	16		
Estratégia 2	1	35	1	16	0	98,6%	97,0%
	2	0	36	1	16		

CAPÍTULO 6 – CONCLUSÕES FINAIS E PERSPECTIVAS FUTURAS

Após a conclusão do trabalho, é chegado o momento de olhar para trás, analisar o percurso feito e perspectivar o trabalho futuro.

O presente trabalho dá uma contribuição para a resolução de problemas relacionados com a qualidade da água, tendo-se recorrido a técnicas e metodologias da Área Científica da Inteligência Artificial e da Descoberta de Conhecimento em Bases de Dados utilizando, como ferramentas, a Indução de Regras, as Árvores de Decisão e as Redes Neurais Artificiais. Desenvolveram-se modelos de previsão da qualidade da água com base nos parâmetros físico-químicos, para as albufeiras em estudo (Alvito e Odivelas), que se caracterizam por apresentarem desempenhos elevados, superiores a 95%, com a particularidade de não cometerem erros de classificação para a classe C. Outro aspecto a destacar prende-se com o facto de as variáveis mais importantes serem a oxidabilidade e a CQO independentemente da albufeira e da técnica de DM utilizada.

Atendendo a que os modelos referidos no parágrafo anterior utilizam variáveis que não permitem a aferição da qualidade da água em tempo real, procurou-se obter modelos de previsão utilizando parâmetros meteorológicos e hidrométricos. A acuidade dos modelos obtidos foi razoável, superior a 80%, com o inconveniente de cometeram algumas falhas de classificação para a classe C. Na tentativa de encontrar solução para este problema disponibilizou-se aos algoritmos, para além dos parâmetros meteorológicos e hidrométricos, o pH. Nestas circunstâncias, conseguiram-se obter modelos que não cometem erros de classificação na classe C e que exibem acuidades próximas de 95%, no caso da albufeira do Alvito e, no caso da albufeira de Odivelas, próximas de 90%.

Do exposto no parágrafo anterior ressalta que os modelos desenvolvidos permitem aferir a qualidade da água em tempo real e potenciar a previsão, dando resposta ao principal desafio lançado constituindo-se, assim, como exemplo da aplicação das ferramentas e técnicas de *Data Mining* à resolução de problemas do foro ambiental.

Tendo em vista compreender a dinâmica das albufeiras procurou-se extrair conhecimento dos dados físico-químicos, através de métodos que utilizassem a aprendizagem não-supervisionada, tendo recorrido a modelos de segmentação utilizando a estratégia *k-means*.

Os modelos de segmentação obtidos para ambas as albufeiras são semelhantes e mostram dois comportamentos distintos, em função da época do ano, ocorrendo a transição entre as referidas dinâmicas em épocas do ano bem definidas, correspondendo ao mês de Novembro e aos meses referentes à Primavera.

O segmento respeitante ao período quente é definido pelas variáveis temperatura e pH, enquanto que o segmento respeitante à época fria é definido pela oxidabilidade. Estes resultados, tal como os obtidos na tarefa de classificação, realçam a importância do pH e da oxidabilidade no comportamento das albufeiras e na previsão da qualidade da água. Além do exposto, os resultados obtidos parecem ainda mostrar que os parâmetros físico-químicos estão, de alguma forma relacionados com as condições meteorológicas, permitindo a sua conjugação contribuir para o muito que há a fazer para diagnosticar, requalificar e preservar as águas superficiais.

Este trabalho abre portas para que no futuro se procure dar resposta a, pelo menos, dois desafios. O primeiro passa por complementar a base de dados com os volumes libertados pela albufeira do Alvito e caracterizar o do troço da Ribeira de Odivelas que faz a ligação entre as duas albufeiras, para tentar compreender o modo como a água armazenada na Barragem de Odivelas é influenciada pela água vinda da albufeira do Alvito e quais as razões subjacentes ao facto de a água da albufeira de Odivelas apresentar, de um modo geral, qualidade inferior quando comparada com a armazenada na albufeira do Alvito.

O segundo, passa por alargar a área de estudo, aplicando-a também à albufeira de Alqueva fazendo, dessa forma a previsão não só do sistema Alvito-Odivelas, mas também do conjunto Alqueva-Alvito-Odivelas.

REFERÊNCIAS BIBLIOGRÁFICAS

A

[Adriaans & Zantinge, 1996] Adriaans, P. & Zantinge, D., Data Mining. Addison-Wesley, Edimburgo, 1996.

[Alley, 2007] Alley, E. R., Water Quality Control Handbook. 2ª Ed., McGraw-Hill, New York, U.S.A., 2007.

[Ambiente 1997] Decreto - Lei n.º 152/1997, de 19 de Junho. 1997.

[Apte et al., 2002] Apte, C., Liu, B., Pednault, E. & Smith, P., Business applications of Data Mining. *Communications of the ACM*, 45: 49-53, 2002.

[Arbib, 2003] Arbib, A., The Handbook of Brain Theory and Neural Networks. 2ª Ed, CIP, London, 2003.

[Atkins et al., 2007] Atkins, J. P., Burdon, D. & Allen, J. H., An application of contingent valuation and decision tree analysis to water quality improvements. *Marine Pollution Bulletin*, 55: 591-602, 2007.

[Azoff, 1994] Azoff, E., Neural Networks Time Series Forecasting of Financial Markets. John Wiley & Sons, USA, 1994.

B

[Bartram & Ballance, 1996] Bartram, J. & Ballance, R., Water Quality Monitoring - a practical guide to the design and implementation of freshwater quality studies and monitoring programmes. UNEP/WHO, 1996.

[Benndorf & Recknagel, 1982] Benndorf, J. & Recknagel, F., Problems of application of ecological model {SALMO} to lakes and reservoirs having various trophic states. *Ecological Modelling*, 17: 139-145, 1982.

[Berry & Linoff, 2000] Berry, M. & Linoff, G., Mastering Data Mining: The Art and Science of Customer Relationships Management. John Wiley & Sons, New York, U.S.A., 2000.

[Berson et al., 2000] Berson, A., Smith, S. & Thearling, K., Building Data Mining Applications for CRM. Mc-Graw Hill, USA, 2000.

[Bobbin & Recknagel, 2001] Bobbin, J. & Recknagel, F., Inducing explanatory rules for the prediction of algal blooms by genetic algorithms. *Environment International*, 27: 237-242, 2001.

[Bortman et al., 2002] Bortman, M., Brimblecombe, P., Cunningham, M. A., Cunningham, W. P. & Freedman, W. Eds., Environmental Encyclopedia. 3ª Ed., Thomson Gale, Farmington Hills, U.S.A., 2002.

[Bose & Liang, 1996] Bose, N. & Liang, P., Neural Networks Fundamentals with Graphs, Algorithms and Applications. McGraw-Hill, USA, 1996.

[Boser et al., 1992] Boser, B., Guyon, I. & Vapnik, V., A training algorithm for optimal margin classifiers. *Proceedings of 5th Annual Workshop on Computational Learning Theory*, Pittsburgh, U.S.A., 144-152, ACM, 1992.

[Bradley & Fayyad, 1998] Bradley, P. S. & Fayyad, U. M., Refining Initial Points for K-Means Clustering. Em J. Shavlik Eds., *Proceedings of 15th International Conference on Machine Learning (ICML98)*, 91- 99, Morgan Kaufmann: San Francisco, 1998.

[Bramer, 2007] Bramer, M., Principles of Data Mining. Springer-Verlag, Londres, U.K., 2007.

[Broomhead & Lowe, 1988] Broomhead, D. & Lowe, D., Multivariable functional interpolation and adaptative networks. *Complex Systems*, 2: 321-355, 1988.

[Butenko et al., 2009] Butenko, S., Chaovalitwongse, W. A. & Pardalos, P. M., Clustering Challenges in Biological Networks. World Scientific, New Jersey, 2009.

C

[Carvalhosa, 1983] Carvalhosa, A., Esquema geológico do Maciço de Évora. Em Eds., Vol. pp. 201-208, Serviço Geológico Portugal, 1983.

[Carvalhosa, 1985] Carvalhosa, A., Sobre os gneisses leptiníticos da faixa de Escoural-Alvito (Alentejo) – Nota preliminar. Em Eds., Vol. pp. 213-218, Bol. Soc. Geol Portugal, 1985.

[Chakrabarti et al, 2009] Chakrabarti, Cox, Frank, Güting, Han, Jiang, Kamber, Lighthstone, Nadeau, Neapolitan, Pyle, Refaat, Schneider, Teorey & Witten, Data Mining: Know it all. Elsevier, 2009.

[Chang et al., 2001] Chang, G., Healey, M., McHugh, J. & Wang, J., Mining the Word Wide Web - An Information Search Approach. Kluwer Academic Publishers, USA, 2001.

[Chau, 2006] Chau, K.-W., A review on integration of artificial intelligence into water quality modelling. *Marine Pollution Bulletin*, 52: 726-733, 2006.

[Chaves & Kojiri, 2007] Chaves, P. & Kojiri, T., Deriving reservoir operational strategies considering water quantity and quality objectives by stochastic fuzzy neural networks. *Advances in Water Resource*, 30: 1329-1341, 2007.

[Chen & Ji, 2007] Chen, G. Q. & Ji, X., Chemical exergy based evaluation of water quality. *Ecological Modelling*, 200: 259-268, 2007.

[Chen, 2003] Chen, L., A study of applying genetic programming to reservoir trophic state evaluation using remote sensor data. *International Journal of Remote Sensing*, 24: 2265-2275, 2003.

[Cheng & Titterington, 1994] Cheng, B. & Titterington, D., Neural networks: a review from a statistical perspective. *Statistical Science* 9, 1994.

[Chester, 1993] Chester, M., Neural Network - a tutorial. PRT Prentice-Hall, USA, 1993.

[Chin, 2006] Chin, D. A., Water-Quality Engineering in Natural Systems. John Wiley & Sons, New Jersey, USA, 2006.

[Cios et al., 2007] Cios, K. J., Pedrycz, W., Swiniarski, R. W. & Kurgan, L. A., Data Mining a Knowledge Discovery Approach. Springer, New York, U.S.A., 2007.

[Clementine, 2007] Clementine, Clementine SPSS – User Guide version 11.1, Integral Solution limited. 2007.

[Conesa-Fernandez, 1997] Conesa-Fernandez, V. V., Guia metodologica para la evaluación del impacto ambiental. 3ª ed., Ediciones Mundi-Prensa, Madrid, Espanha, 1997.

[Cortez, 2002] Cortez, P., Modelos Inspirados na Natureza para a Previsão de Séries Temporais, Tese de Doutoramento, Departamento de Informática, Universidade do Minho, Braga, Portugal, 2002.

[Cortez, 2004] Cortez, P., Aprendizagem e Avaliação de Modelos – Apontamentos Pedagógicos. Departamento de Sistemas de Informação, Universidade do Minho, Guimarães, Portugal. 2004.

[Cortez & Neves, 2000] Cortez, P. & Neves, J., Redes Neurais Artificiais. Universidade do Minho, Braga, 2000.

[Costa & Simões, 2008] Costa, E. & Simões, A., Inteligência artificial - fundamentos e aplicações. 1ª Ed., FCA, 2008.

[Cruz, 2007] Cruz, A., Data Mining via Redes Neurais Artificiais e Máquinas de Vectores de Suporte, Universidade de Minho, 2007.

[Cunha et al., 1980] Cunha, L. V., Gonçalves, A. S., Figueiredo, V. A. & Lino, M., A Gestão da Água. Princípios fundamentais e sua aplicação em Portugal. Fundação Calouste Gulbenkian, Lisboa, Portugal, 1980.

D

[Dejak et al., 1987b] Dejak, C., Mazzei Lalatta, I., Messina, E. & Pecenik, G., A two-dimensional diffusion model of the Venice lagoon and relative open boundary conditions. *Ecol. Modelling*, 37: 21-45, 1987a.

[Dejak et al., 1987c] Dejak, C., Mazzei Lalatta, I., Messina, E. & Pecenik, G., An advection-diffusion pollution model of the lagoon of Venice. *Ecol. Modelling*, 37: 47-57, 1987b.

[Dejak et al., 1987d] Dejak, C., Mazzei Lalatta, I., Messina, E. & Pecenik, G., Steady-state achievement by introduction of true tidal velocities in a pollution model of the Venice lagoon. *Ecol. Modelling*, 37: 59-79, 1987c.

[Dejak et al., 1987e] Dejak, C., Mazzei Lalatta, I., Molin, M. & Pecenik, G., Tidal three-dimensional diffusion in a model of the lagoon of Venice and reliability conditions for its numerical integration. *Ecol. Modelling*, 37: 81-101, 1987d.

[Dejak et al., 1987a] Dejak, C., Mazzei Lalatta, I., Meregalli, L. & Pecenik, G., Development of a mathematical eutrophication model of the lagoon of Venice. *Ecol. Modelling*, 37: 1-20, 1987e.

[Dejak et al., 1989] Dejak, C., Franco, D., Pasters, R. & Pecenik, G., A steady state achieving 3D eutrophication-diffusion submodel. *Environmental Software*, 4: 94-101, 1989.

[DRAOTA, 1999] Identificação e Caracterização das Fontes Poluidoras de Carácter Pontual Localizadas em Bacias Hidrográficas das Albufeiras Destinadas à Produção de Água e em Bacias Drenantes para Zonas Balneares. Relatório da Direcção Regional do Ambiente e Ordenamento do Território – Alentejo, Direcção Regional do Ambiente e Ordenamento do Território. 1999.

[Dzeroski, 2002] Dzeroski, S., Environmental Sciences. Em W. Klösgen & J. M. Zytkow Eds., *Handbook of Data Mining and Knowledge Discovery*, Vol. pp. 817-830, Oxford University Press, Oxford, U.K., 2002.

[Dzeroski, 2006] Dzeroski, S., Environmental Applications of Data Mining. Em Australian Academy of Science Eds., *Proceedings of Elizabeth and Frederick White Conference - Mastering the Data Explosion in the Earth and Environmental Sciences*, Shine Dome, Canberra, Austrália. 2006.

E

[Eaton et al., 2005] Eaton, A., Clesceri, L., Rice, E. & Greenberg, A. Eds., *Standard Methods for the Examination of Water and Wastewater*. 21^a Ed., American Public Health Association, USA, 2005.

[Elshorbagy & Ormsbee, 2006] Elshorbagy, A. & Ormsbee, L., Object-oriented modeling approach to surface water quality management. *Environmental Modelling & Software*, 21: 689-698, 2006.

[Esteves, 1988] Esteves, F. A., *Fundamentos de Limnologia*. 2^a Edição, Interciências/Finep, São Paulo, 1988.

[Esteves & Barbosa, 1986] Esteves, F. A. & Barbosa, F., *Eutrofização artificial, a doença dos lagos*. Ciência hoje, Rio de Janeiro, 1986.

[Eurostat, 2009] <http://epp.eurostat.ec.europa.eu> (Visitado em 06-Out de 2009).

[Eurostat, 2009a] <http://epp.eurostat.ec.europa.eu> (Visitado em 6 Out de 2009).

F

[FAO, 2006] FAO, World reference base for soil resources 2006 - A framework for international classification, correlation and communication. Publishing Management Service, Roma, 2006.

[Fayyad, 1997] Fayyad, U., Mining Databases: Towards Algorithms for Knowledge Discovery. *IEEE Computer Society Technical Committee on Data Engineering*, 21: 39-48, 1997.

[Fayyad et al., 1996] Fayyad, U., Piatetsky-Shapiro, G., Smith, P. & Uthurusamy, R., Advances in Knowledge Discovery and Data Mining. AAAI Press/MIT Press, Massachusetts, USA, 1996.

[Fernandes, 2009] Fernandes, A. T., Previsão da Classificação da Qualidade dos Efluentes da Indústria Petrolífera com Recurso a Sistemas Inteligentes, Dissertação de Mestrado, Universidade de Évora, Évora, Portugal, 2009.

[Fernandes et al., 2009] Fernandes, A. V., Vicente, H. & Neves, J., Solving Challenging Problems in the Oil Industry Using Artificial Intelligence Based Tools. Em Diganta B. Das, V. Nassehi & L. Deka Eds., *Proceedings of Industrial Simulation Conference' 2009 - ISC' 2009*, 325-331, Eurosis - ETI Publication, Loughborough, UK. 2009.

[Flood & Kartam, 1994] Flood, I. & Kartam, N., Neural network in civil engineering: I. Principles and understanding. *Journal of Computational in Civil Engineering*, 8, 131-148, 1994.

[Fonseca, 1995] Fonseca, P., Estudo de sutura varisca no SW ibérico nas regiões de Serpa-Beja-Torrão e Alvito-Viana do Alentejo, Dep. Geol. Fac. Ciências, Univ. Lisboa, 1995.

[Fonseca et al., 1999] Plano de Bacia Hidrográfica do Rio Sado, Anexo Temático 8 - Usos e Ocupações do Domínio Hídrico, Parte I - Utilizações do Domínio Hídrico. Ministério do Ambiente e do Ordenamento do Território, 1999a.

[Fonseca et al., 1999] Plano de Bacia Hidrográfica do Rio Sado, Anexo Temático I - Análise biofísica, Parte I - Análise geomorfológica. Ministério do Ambiente e do Ordenamento do Território, 1999b.

G

[Galushkin, 2007] Galushkin, A. I., *Neural Networks Theory*. Springer, New York, 2007.

[Gelman et al., 2003] Gelman, A., Carlin, J. B., Stern, H. S. & Rubin, D. B., *Bayesian Data Analysis*. 2ª Ed., Chapman & Hall / CRC, Boca Raton, U.S.A., 2003.

[Goebel & Gruenwald, 1999] Goebel, M. & Gruenwald, L., A survey of Data Mining and knowledge discovery software tools. *ACM SIGKDD Explorations*, 1: 20-33, 1999.

[Gomes, 2000] Gomes, E. M. C., *Metamorfismo de rochas carbonatadas siliciosas da região de Alvito (Alentejo, Sul de Portugal)* Universidade de Coimbra, 2000.

[Google Earth, 2009] (Visitado em 12 de Julho de 2009).

[Gopalan & Silvaselvan, 2009] Gopalan, N. P. & Silvaselvan, B., *Data Mining: Techniques and Trends*. PHI, 2009.

[Govindaraju & Rao, 2000] Govindaraju, R. S. & Rao, A. R., *Artificial Neural Networks in Hydrology* Kluwer Academic Publishers, Dordrecht, Holanda, 2000.

[Graupe, 2007] Graupe, D., *Principles of artificial neural networks*. 2nd Edition, World Scientific, 2007.

[Groth, 2000] Groth, R., *Data Mining: Building Competitive Advantage*. Prentice Hall PTR, USA, 2000.

H

[Hagal et al., 1996] Hagal, M., Demuth, H. & Beale, M., *Neural Network Design*. PWS Publishing Company, USA, 1996.

[Han & Kamber, 2006] Han, J. & Kamber, M., *Data Mining: Concepts and Techniques*. 2ª Ed., Morgan Kauffmann Publishers, San Francisco, U.S.A., 2006.

[Hand et. al, 2001] Hand, D., Mannila, H. & Smith, P., Principles of Data Mining. Massachusetts Institute of Technology, London, 2001.

[Harrison, 2001] Harrison, R., Pollution: Causes, Effects and Control. Fourth, Royal Society of Chemistry, 2001.

[Hastie et al., 2001] Hastie, T., Tibshirani, R. & Friedman, J., The elements of statistical learning: data Mining, inference, and prediction. Springer-Verlag, USA, 2001.

[Hasting et al., 1993] Hasting, A., Hom, C. L., Ellner, S., Turchin, P. & Godfray, H., Chaos in ecology: is mother nature a strange attractor? *Ann. Rev. Ecol. Syst.*, 34: 1-33, 1993.

[Haykin, 2008] Haykin, S., Neural Networks and learning machines. 3^a Ed, Prentice Hall, 2008.

[Hopfield, 1982] Hopfield, J., Neural networks and physical systems with emergent collective computational abilities. *The National Academy of Science*, 79: 2554-2558, 1982.

[Hornik et al., 1989] Hornik, K., Stinchcombe, M. & White, H., Multilayer feed-forward networks are universal approximators. *Neural Networks*. 1989.

J

[Jeong et al., 2003] Jeong, K.-S., Kim, D.-K., Whigham, P. & Joo, G.-J., Modelling *Microcystis aeruginosa* bloom dynamics in the Nakdong River by means of evolutionary computation and statistical approach. *Ecological Modelling*, 161: 67-78, 2003.

[Jeong et al., 2001] Jeong, K.-S., Joo, G.-J., Kim, H.-W., Ha, K. & Recknagel, F., Prediction and elucidation of phytoplankton dynamics in the Nakdong River (Korea) by means of a recurrent artificial neural network. *Ecological Modelling*, 146: 115-129, 2001.

[Jordan, 1995] Jordan, M., Why the logistic function? a tutorial discussion on probabilities and neural networks. MIT. USA, 1995.

[Jørgensen, 1995] Jørgensen, S. E., State of the art of ecological modelling in limnology. *Ecological Modelling*, 78: 101-115, 1995.

[Jørgensen, 1999] Jørgensen, S. E., State-of-art of ecological modelling with emphasis on development of structural dynamic models. *Ecological Modelling*, 120: 75-96, 1999.

[Jørgensen & Bendoricchio, 2001] Jørgensen, S. E. & Bendoricchio, G., *Fundamentals of Ecological Modelling*. 3ª Ed, Elsevier, Amesterdão, 2001.

K

[Karmakar & Mujumdar, 2007] Karmakar, S. & Mujumdar, P. P., A two-phase grey fuzzy optimization approach for water quality management of a river system. *Advances in Water Resources*, 30: 1218-1235, 2007.

[Kartalopoulos, 1996] Kartalopoulos, V., *Understanding Neural Network and Fuzzy Logic - Basic Concepts and Applications*. IEEE Press, USA, 1996.

[Karul et al., 2000] Karul, C., Soyupak, S., Çilesiz, A., Akbay, N. & Germen, E., Case studies on the use of neural networks in eutrophication modelling. *Ecological Modelling*, 134: 145-152, 2000.

[Kmet & Straškraba, 1989] Kmet, T. & Straskraba, M., Global behaviour of a generalized aquatic ecosystem model. *Ecol. Modelling*, 45: 95-110, 1989.

[Kock, 1999] Kock, N., *Process Improvement and Organizational Learning: the role of collaboration technologies*. Hershey: Idea Group Publishing, 1999.

[Kohavi & Provost, 1998] Kohavi, R. & Provost, F., Glossary of Terms. *Machine Learning*, 30: 271-274, 1998.

[Kohonen, 1982] Kohonen, T., Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43: 59-69, 1982.

[Kufrin, 1997] Kufrin, R., Generating C4.5 Productions Rules in Parallel. *Proceedings of 4th National Conference on Artificial Intelligence (AAAI-97)*, Rhode Island, U.S.A., pp. 565-570, AAAI Press, 1997.

[Kuo et al., 2006] Kuo, J.-T., Wang, Y.-Y. & Lung, W.-S., A hybrid neural-genetic algorithm for reservoir water quality management. *Water Research*, 40: 1367-1376, 2006.

[Kuo et al., 2007] Kuo, J.-T., Hsieh, M.-H., Lung, W.-S. & She, N., Using artificial neural network for reservoir eutrophication prediction. *Ecological Modelling*, 200: 171-177, 2007.

L

[Lampert, 2007] Lampert, W. & Sommer, U., Limnoecology, the ecology of lakes and streams. 2ª edição, Oxford, 2007.

[Langley & Simon, 1995] Langley, P. & Simon, H., Applications of Machine Learning and Rule Induction. *Communications of the ACM*, 38: 54-64, 1995.

[Larose, 2006] Larose, D. T., Data Mining Methods and Models. John Wiley & Sons, Hoboken, New Jersey, U.S.A., 2006.

[Lee et al., 2003] Lee, J. H., Huang, Y., Dickman, M. & Jayawardena, A. W., Neural network modelling of coastal algal blooms. *Ecological Modelling*, 159: 179–210, 2003.

[Lehr & Keeley, 2005] Lehr, J. H. & Keeley, J. Eds., Water Encyclopedia: Surface and Agricultural Water. John Wiley & Sons, Hoboken, New Jersey, USA, 2005.

[Lehr et al., 2005] Lehr, J. H., Keeley, J. & Lehr, J. Eds., Water Encyclopedia: Water Quality and Resource Development. 1ª, John Wiley & Sons, Hoboken, New Jersey, USA, 2005.

[Lek & Guégan, 1999] Lek, S. & Guégan, J. F., Artificial neural network as a tool in ecological modelling, an introduction. *Ecological Modelling*, 120: 65-73, 1999.

[Lin et. al, 2008] Lin, T. Y., Xie, Y., Wasilewska, A. & Liau, C., Data Mining: Foundations and Practice. Springer, 2008.

[Little, 1992] Little, R., Regression with missing X's: a review. *Journal of the American Statistical Association*, 87: 1227-1237, 1992.

[Liu et al., 2007] Liu, S., Butler, D., Brazier, R., Heathwaite, L. & Khu, S.-T., Using genetic algorithms to calibrate a water quality model. *Science of The Total Environment*, 374: 260-272, 2007.

[Lotka, 1924] Lotka, A. J., Contribution to the energetics of evolution. *Proc. Natl. Acad. Sci.*, 8: 147-150, 1924.

M

[Maimon & Rokach, 2005] Maimon, O. & Rokach, L., Decomposition Methodology for Knowledge Discovery and Data Mining: Theory and Applications. World Scientific, 2005.

[Mamedov & Udalov, 2002] Mamedov, A. & Udalov, S., A computer tool to develop individual based models for simulation of population interactions. *Ecological Modelling*, 147: 53-68, 2002.

[Matoso et al., 2004] Matoso, A., Rasga, M., Santana, M. & Murteira, M., Principais albufeiras do Alentejo monitorizadas. Comissão de Coordenação da Região Alentejo, Direcção de Serviços de Monitorização Ambiental, Alentejo, Portugal. 2004.

[McCulloch & Pitts, 1943] McCulloch, W. & Pitts, W., A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5: 115-133, 1943.

[Mendes & Oliveira, 2004] Mendes, B. & Oliveira, F. J., Qualidade da Água para Consumo Humano. Lidel - Edições Técnicas, Lisboa, Portugal, 2004.

[Metcalf & Eddy, 2003] Metcalf & Eddy, Wastewater engineering, Treatment Disposal Reuse. Fourth edition, Mackgraw-Hill Inc, New York, 2003.

[Michalski et al., 1998] Michalski, R., Bratko, I. & Miroslav, K., Machine Learning and Data Mining Methods and Applications. John Willey & Sons, Londres, U.K., 1998.

[Miranda & Bernard, 2009] Miranda, F. R. & Bernard, L. M., Lake pollution research progress. Nova Science Publisher, inc, New York, 2009.

[Mota, 2006] Mota, S., Introdução à engenharia ambiental. 4ª Edição, Abes, Rio de Janeiro, 2006.

N

[Neves et al., 1998] Neves, J., Machado, J., Cortez, P., Arteiro, J. & Vicente, H., A distributed Web-based water quality diagnosis and simulation in lakes and reservoirs. Em A. Bargiela & E. Kerckhoffs Eds., Simulation Technology: Science and Art, *Proceedings of ESS'98 - 10th European Simulation Symposium and Exhibition*, 296-301, SCS, Nottingham, U.K. 1998.

[Nisbet et. al, 2009] Nisbet, R., Elder, J. & Miner, G., Handbook of Statistical Analysis & Data Mining Applications. Elsevier, 2009.

[Nonaka & Takeuchi, 1997] Nonaka, I. & Takeuchi, H., Criação de conhecimento na empresa. Como as empresas japonesas geram a dinâmica da inovação. Rio de Janeiro: Campus, 1997.

[Novotny, 2003] Novotny, V., Water Quality – Diffuse pollution and watershed management. John Wiley & Sons, Inc, , New York, 2003.

O

[OCDE, 2003] OCDE, Improving Water Management : Recent OECD Experience. IWA publishing, 2003.

[OCDE, 2006] OCDE, Improving Water Management: Recent OECD Experience, Policy Brief. Paris, 2006.

[OMS, 2009] <http://www.who.int/en/> (Visitado em 13 de Abril de 2009).

P

[Pal & Jain, 2005] Pal, N. R. & Jain, L. Eds., Advanced Techniques in Knowledge Discovery and Data Mining. Springer-Verlag, Londres, U.K., 2005.

[Patterson, 1996] Patterson, D., Artificial Neural Networks - Theory and Application. Prentice Hall, New York, U.S.A., 1996.

[Peirce et. al,1998] Peirce, J. J., Weiner, R. E. & Vesilind, E. A., Environmental Pollution and control. 4ª Ed, Butterworth-Heinemann, Oxford, 1998.

[Perlovsky, 2001] Perlovsky, L. I., Neural Networks and Intellect: Using Model-Based Concepts. OXFORD UNIVERSITY, Oxford, 2001.

[Piçarra, 2000] Piçarra, J. M., Estudo estratigráfico do sector Estremoz-Barrancos, Zona de Ossa-Morena, Portugal, Vol. I – Litoestratigrafia do intervalo Câmbrico médio – Devónico inferior. 95 p & Vol. II – Bioestratigrafia do intervalo ordovícico – devónico inferior, Tese de Doutoramento, Dep. Geociências, Universidade de Évora, 2000.

[Pinto et al., 2009] Pinto, A., Fernandes, A. V., Vicente, H. & Neves, J., Optimizing Water Treatment Systems Using Artificial Intelligence Based Tools. Em C. A. Brebbia e V. Popov Eds., Water Resource Management V, WIT Transactions on Ecology and the Environmental, Vol. 125, pp. 185-194, WIT Press, , Southampton, UK, 2009.

[Pinto, 2009] Pinto, A. M., Aplicação de Sistemas Inteligentes à Optimização de um Sistema de Tratamento de Águas, Dissertação de Mestrado, Universidade de Évora, Évora, Portugal, 2009.

[Poncelet et al., 2008] Poncelet, P., Teisseire, M. & Masseglia, F., Data Mining Patterns: New Methods and Applications. Information Science Reference, New York, 2008.

[Popek, 2006] Popek, E., Sampling and analysis of environment chemical pollutants. AP, 2006.

Q

[Quevauviller et al., 2008] Quevauviller, P., Borchers, U., Thompson, C. & Simonart, T. Eds., The Water Framework Directive Ecological and Chemical Status Monitoring. John Wiley & Sons, Chichester, U. K., 2008.

[Quinlan, 1993] Quinlan, J., C4.5 Programs for Machine Learning. Morgan Kaufmann Publishers Inc, USA, 1993.

[Quinlan, 1987] Quinlan, J. R., Simplifying decision trees. *International Journal of Man-Machine Studies*, 27: 221-234, 1987.

[Quinlan, 1996] Quinlan, J. R., Bagging, boosting and C4.5. Em Eds., *Proceedings of AAAI'96 National Conference on Artificial Intelligence*, 725-730, AAAI Press, Menlo Park, USA. 1996.

[Quinlan, 1998] Quinlan, R., Programs for Machine Learning. Morgan Kaufmann Publishers, USA, 1998.

R

[Ray et al., 2001] Ray, S., Berec, L., Straškraba, M. & Jørgensen, S. E., Optimization of exergy and implications of body sizes of Phytoplankton and Zooplankton in an aquatic ecosystem model. *Ecological Modelling*, 140: 219-234, 2001.

[Recknagel, 2003] Recknagel, F., Simulation of aquatic food web and species interactions by adaptive agents embodied with evolutionary computation: a conceptual framework. *Ecological Modelling*, 170: 291-302, 2003.

[Recknagel & Benndorf, 1982] Recknagel, F. & Benndorf, J., Validation of the ecological simulation model {SALMO}. *Int. Rev.*, 67: 113-125, 1982.

[Rich & Knight, 1991] Rich, E. & Knight, K., Artificial Intelligence. 2ª Ed., McGraw-Hill Inc., New York, U.S.A., 1991.

[Riley & Stefan, 1988] Riley, M. J. & Stefan, H. G., MINLAKE: a dynamic lake water quality simulation model. *Ecological Modelling*, 43: 155-182, 1988.

[Rodrigues, 1992] Rodrigues, A., Modelação Matemática da Qualidade de Água em Albufeiras, Universidade Nova de Lisboa, Portugal, 1992.

[Rodrigues et al., 2001] Rodrigues, R., Quadrado, F. L. & Lopes, A. R., A qualidade da água em Portugal. *SNIRH*, 2001.

[Rojas, 1996] Rojas, R., Neural Networks - A Systematic Introduction. Springer-Verlag, Berlin, Alemanha, 1996.

[Rosas, 2003] Rosas, F., Estudo tectónico do sector de Viana do Alentejo – Alvito: Evolução geodinâmica e modelação analógica de estruturas em afloramentos chave, Dep. Geol. Fac. Ciências Univ. Lisboa. Tese de Doutoramento em Geologia na especialidade de geodinâmica interna, 2003.

[Rubin, 1996] Rubin, D. B., Multiple Imputation after 18 years. *Journal of the American Statistical Association*, 91: 473-489, 1996.

[Rud, 2001] Rud, O., Data Mining Cookbook John Wiley & Sons, USA, 2001.

[Rumelhart et al., 1986] Rumelhart, D., Hinton, G. & Williams, R., Learning Internal Representation by Error Propagation. Em David E. Rumelhart & James L. McClelland

Eds., *Parallel Distributed Processing*, Vol. 1: Foundations, pp. 318-362, MIT Press, Massachusetts, U.S.A., 1986.

[Russell & Norvig, 2003] Russell, S. J. & Norvig, P., *Artificial Intelligence: A Modern Approach 2ª Ed.*, Prentice Hall, New Jersey, U.S.A., 2003.

S

[Salençon & Thébault, 1996] Salençon, M.-J. & Thébault, J.-M., Simulation model of a mesotrophic reservoir (Lac de Pareloup, France): MELODIA, an ecosystem reservoir management model. *Ecological Modelling*, 84: 163-187, 1996.

[Salomonsen & Jensen, 1996] Salomonsen, J. & Jensen, J., Use of lake model to examine exergy response to changes in phytoplankton growth parameters and species composition. *Ecological Modelling*, 87: 41-49, 1996.

[Santos, 2001] Santos, A. R., *Gestão do conhecimento: Uma experiência para o sucesso empresarial*. Editora Champagnat, 2001.

[Santos & Azevedo, 2005] Santos, M. & Azevedo, C., *Data mining - descoberta de conhecimento em base de dados*. FCA, 2005.

[Santos et al., 2005] Santos, M. F., Cortez, P., Quintela, H., Neves, J., Vicente, H. & Arteiro, J., *Ecological Mining - A Case Study on Dam Water Quality*. Em A. Zanasi, C.A. Brebbia & N.F.F. Ebecken Eds., *Data Mining VI - Data Mining, Text Mining and their Business Applications*, WIT Transactions of Information and Communication Technologies, Vol. 35, pp. 523-531, WIT Press, Southampton, UK, 2005.

[Sapo, 2009] <http://mapas.sapo.pt/> (Visitado em 27 de Novembro de 2009).

[Scardi, 2001] Scardi, M., Advances in neural network modeling of phytoplankton primary production. *Ecological Modelling*, 146: 33-45, 2001.

[Schafer, 1999] Schafer, J., *Analysis of Incomplete Multivariate Data*. Chapman & Hall / CRC, Boca Raton, U.S.A., 1999.

[Schapire, 2002] Schapire, E. R., The boosting approach to machine learning: An overview. *Proceedings of MSRI Workshop on Nonlinear Estimation and Classification*, 2002.

- [Scheer et al., 1998] Scheer, W., Allweyer, T. & Ege, C., Business Process Reengineering and Knowledge Management. Actas do ISI, 1998.
- [Scholten et al., 2005] Scholten, M., Foekema, E., Dokkum, H., Kaag, N. & Jak, M., Eutrophication management and ecotoxicology. Springer, 2005.
- [Schwarzenbach et al., 2003] Schwarzenbach, R. P., Gschwend, P. M. & Imboden, D. M., Environmental Organic Chemistry. 2ª Ed., John Wiley & Sons, Hoboken, U.S.A., 2003.
- [Silva, 2002] Silva, M. P., Anuário Estatístico da Região Alentejo 2001. Instituto Nacional de Estatística, Direcção Regional do Alentejo, 2002.
- [SNIRH, 2009] <http://snirh.pt/> (Visitado em 20 de Agosto de 2009).
- [Soman et al., 2006] Soman, K. P., Diwakar, S. & Ajay, V., Data Mining: Theory and Practice. Prentice-Hall, 2006.
- [Sousa, 2004] Sousa, C., Data Mining: Metodologias, Tecnologias, Modelos e Aplicações, Dissertação de Mestrado, Universidade do Minho, 2004.
- [Spellman, 2008] Spellman, F. R., The Science of Water - Concepts and Applications. 2ª Ed., CRC Press, Boca Raton, U.S.A., 2008.
- [Stewart & Howell, 2003] Stewart, B. & Howell, T., Encyclopedia of water science. Dekker, 2003.
- [Straskraba, 1994] Straskraba, M., Ecotechnology models for reservoir water quality management. *Ecological Modelling*, 74: 1-38, 1994.
- [Straskraba, 2001] Straskraba, M., Natural control mechanisms in models of aquatic ecosystems. *Ecological Modelling*, 140: 195-205, 2001.
- [Streeter & Phelps, 1925] Streeter, H. & Phelps, E., A study of pollution and the natural purification of the Ohio River. Public Health Bulletin nº 146, USA Public Health Service, 1925.
- [Suárez, 1999] Suárez, I., Mastering chaos in ecology. *Ecol. Modelling*, 117: 305-314, 1999.
- [Sveiby, 1998] Sveiby, K. E., A nova riqueza das organizações. Campus, Rio de Janeiro, 1998.

T

[Taniar, 2008] Taniar, D. Eds., Data Mining and Knowledge Discovery Technologies. IGI Publishing, Hershey, New York, U.S.A., 2008.

[Tebbutt, 1998] Tebbutt, T. H., Principles of Water Quality Control. 5^a Ed., Butterworth-Heinemann, Oxford, U.K., 1998.

[Teixeira, 2005] Teixeira, H. Q., Sistemas de Conhecimento Baseados em Data Mining: Aplicação à Análise da Estabilidade de Estruturas Metálicas, Dissertação de Mestrado, Universidade do Minho, Braga, Portugal, 2005.

[Terra, 2005] Terra, J. C. C., Gestão do conhecimento: o grande desafio empresarial. 5^a edição, Elsevier, Rio de Janeiro, 2005.

[Thébault & Saleçon, 1993] Thébault, J.-M. & Saleçon, M.-J., Simulation model of a mesotrophic reservoir (Lac de Pareloup, France): biological model. *Ecological Modelling*, 65: 1-30, 1993.

[Thuraisingham, 1999] Thuraisingham, B., Data Mining Technologies, Techniques, Tools and Trends. CRC Press LLC, USA, 1999.

[Torgo, 1999] Torgo, L., Árvores de Regressão – Métodos e Aplicações. LIACC, FEP Porto, 1999.

[Turban et al., 2004] Turban, E., Aronson, J. E. & Liang, T.-P., Decision Support Systems and Intelligent Systems. 7^a Edição, Prentice Hall, New Jersey, USA, 2004.

[Tyagi et al., 2008] Tyagi, P., Chandramouli, V., Lingireddy, S. & Buddhi, D., Relative Performance of Artificial Neural Networks and Regression Models in Predicting Missing Water Quality Data. *Environmental Engineering Science*, 25: 657-668, 2008.

V

[Varis et al., 2008] Varis, O., Tortajada, C. & Biswas, A. K., Water Resources Development and Management: Management of Transboundary Rivers and Lakes. Springer, 2008.

[Vayenas & Pavlou, 2001] Vayenas, D. & Pavlou, S., Chaotic dynamics of a microbial system of coupled food chains. *Ecological Modelling*, 136: 285-295, 2001.

- [Veolia, 2009] <http://www.aguasdevalongo.net> (Visitado em 29 de Setembro de 2009).
- [Vicente, 2004] Vicente, H., Especificação e Prototipação de Sistemas de Gestão e Controlo da Qualidade da Água de Albufeiras, Departamento de Química, Universidade de Évora, Portugal, 2004.
- [Volterra, 1926] Volterra, V., Fluctuations in the abundance of a species considered mathematically. *Nature*, 188: 558-560, 1926.

W

- [Weiner, 2000] Weiner, E. R., Applications of Environmental Chemistry: a practical guide for Environmental professionals. Lewis Publishers, New York, 2000.
- [Weiss & Provost, 2001] Weiss, G. & Provost, F., The effect of class distribution on classifier learning: an empirical study. Em Technical Report ML-TR-44 Eds., *Proceedings of Department of Computer Science, Rutgers University*, 2001.
- [Weka, 2004] <http://www.cs.waikato.ac.nz/ml/weka/> (Visitado em 18 de Novembro de 2009).
- [Wetzel, 1993] Wetzel, R., Limnologia. Fundação Calouste Gulbenkian, Lisboa, 1993.
- [Wilhelm & Bruggemann, 2000] Wilhelm, T. & Bruggemann, R., Goal functions for the development of natural systems. *Ecological Modelling*, 132: 231-246, 2000.
- [Williams & Simoff, 2006] Williams, G. J. & Simoff, S. J., Data Mining: Theory, Methodology, Techniques, and Applications Springer, 2006.
- [Wilson & Recknagel, 2001] Wilson, H. & Recknagel, F., Towards a generic artificial neural networks model for dynamic predictions of algal abundance in freshwater lakes. *Ecological Modelling*, 146: 69-84, 2001.
- [Witten & Frank, 2005] Witten, I. H. & Frank, E., Data mining - Practical Machine Learning Tools and Techniques. 2ª Ed, Elsevier, 2005.
- [Wu & McLarty, 2000] Wu, C. H. & McLarty, J. W., Neural Networks and Genome Informatics. Elsevier, Maryland, USA, 2000.

Y

[Yang et al., 2008] Yang, W., Nan, J. & Sun, D., An online water quality monitoring and management system developed for the Liming River basin in Daqing, China. *Journal of Environmental Management*, 88: 318–325, 2008.

Z

[Zhang, 2007] Zhang, C., *Fundamentals of Environmental Sampling and Analysis*. John Wiley & Sons, Inc., Hoboken, New Jersey, USA, 2007.

ANEXO I

Neste anexo apresenta-se a BD utilizada neste trabalho, referente às albufeiras do Alvito e de Odivelas, que inclui dados de qualidade da água, hidrométricos e meteorológicos.

Tabela I.1 – Dados da qualidade da água da albufeira do Alvito.

Data	CQO	CBO ₅	Clorofila-a	Coliformes fecais	Coliformes totais	Condutividade	Fosforo total	Nitratos	Nitritos	Ortofosfatos	Oxidabilidade	OD	pH	SST	Temperatura
	mg/L	mg/L	µg/L	ufc/100 mL	ufc/100 mL	uS/cm	mg/L	mg/L	mg/L	mg/L P ₂ O ₅	mg/L	%	E. Sorensen	mg/L	°C
05-Jun-2001	10,5	2	3,7	95	1000	227	0,013	0,44	0,019	0,006	3,9	97	8,63	0,6	24,9
03-Jul-2001	16,3	2	6,2	20	310	240	0,01	0,09	0,001	0,021	4,2	98	8,98	3,3	25,6
01-Ago-2001	20,9	2	10,8	1	300	245	0,16	0,53	0,004	0,205	3,6	126	8,66	9,2	25,3
04-Set-2001	12	2	6,5	0	1600	232	0,043	0,09	0,004	0,034	4,6	84	8,46	4	25,5
02-Out-2001	14,3	2	9,2	220	7100	251	0,018	0,22	0,002	0,019	4,3	87,8	8,66	3,2	21,5
06-Nov-2001	18,8	2	39,1	50	1600	242	0,021	0,35	0,109	0,023	5,1	104,2	7,97	4,6	18,6
04-Dez-2001	18,8	2,56	31,1	5	10	247	0,043	0,71	0,013	0,094	6	103	8,17	3,2	12,6
08-Jan-2002	15	3	19,2	22	120	245	0,032	1,02	0,021	0,036	5,2	91	8,4	3,8	10,5
05-Fev-2002	16	3	17,8	11	250	248	0,022	0,75	0,02	0,03	4,6	93	8	6	12,8
05-Mar-2002	14	4	6,7	3	3	248	0,023	0,89	0,006	0,021	4,9	99	7,9	4,6	12,6
02-Abr-2002	21	3	6,1	0	90	245	0,018	0,58	0,009	0,021	5,8	105	8,4	1,6	18,7
07-Mai-2002	17	3	3,0	9	40	254	0,02	0,31	0,006	0,01	4,3	89	8,5	2,7	17
04-Jun-2002	18	3	1,4	130	1100	257	0,018	0,27	0,005	0,027	4	110	8,4	3,2	20,4
02-Jul-2002	16	3	1,3	21	400	263	0,019	0,13	0,003	0,014	4,7	135	8,4	0,6	24,2
01-Ago-2002	13	3	2,9	6	190	271	0,04	0,13	0,001	0,067	4,9	167	8	1,4	25,6
03-Set-2002	25	3	1,7	11	230	281	0,017	0,18	0,003	0,035	4,5	147	8,2	1,9	23,5
01-Out-2002	11	3	4,3	9	270	277	0,015	0,31	0,002	0,022	4,8	145	8,5	3,6	21,6
01-Nov-2002	13	3	2,3	8	200	274	0,015	0,18	0,001	0,033	4	107	8,5	0,9	20
01-Dez-2002	10	3	5,3	5	12	275	0,022	0,71	0,044	0,035	4,4	83,5	8,1	3,1	15
07-Jan-2003	15	3	3,4	41	1160	251	0,04	1,46	0,036	0,033	4,8	75	7,9	6,7	13
04-Fev-2003	16	3	8,4	0	7	263	0,018	1,46	0,01	0,037	6,6	84,1	7,8	5,2	11,2
05-Mar-2003	31	8	56,4	0	120	263	0,063	1,02	0,007	0,034	7,6	121,8	8,5	11	15,5
01-Abr-2003	15	5	45,1	11	25	268	0,038	1,02	0,004	0,047	5,9	100,8	8,3	7,1	16,3
06-Mai-2003	34	3	8,4	1	900	278	0,064	0,93	0,001	0,015	4,7	95,2	8	25	18,4
03-Jun-2003	16	3	30,3	2	150	272	0,022	0,93	0,002	0,013	4	92,3	8,5	3,9	23
01-Jul-2003	10	3	4,0	90	6400	254	0,158	0,09	0,001	0,029	4,7	80,4	8,3	2,2	24,5
01-Ago-2003	14	3	2,5	8	60	308	0,23	0,89	0,001	0,294	4,5	112,1	8,5	7,2	27,5
09-Set-2003	10	3	1,4	2	790	320	0,019	0,93	0,02	0,01	4,3	91,7	8	8,4	24
07-Out-2003	23	3	1,7	60	300	319	0,037	1,02	0,009	0,023	3,9	144,5	8,1	8,7	20,8
04-Nov-2003	16	3	6,2	50	260	304	0,021	0,46	0,158	0,05	3,8	102,1	8,1	3,1	17,2
02-Dez-2003	20	3	7,6	21	200	283	0,034	1,59	0,031	0,037	3,8	116,5	8	2,9	13,8
13-Jan-2004	18	3	1,8	0	4	305	0,03	1,32	0,013	0,064	3,4	94,6	8,1	2,7	13
10-Fev-2004	19	3	4,0	6	31	298	0,026	1,59	0,012	0,023	4,5	100,2	8,5	7	13,8
09-Mar-2004	22	3	9,5	26	230	300	0,032	1,12	0,007	0,023	4,7	101,4	8,3	5,9	14,2
06-Abr-2004	56	8	11,5	22	72	269	0,001	0,48	0,009	0,023	5,8	105,8	8,5	14	16,8
04-Mai-2004	18	3	2,0	0	70	309	0,012	0,23	0,007	0,015	4	112,4	8,1	2,2	18,3
01-Jun-2004	15	3	1,4	4	15	296	0,013	0,19	0,001	0,023	3,9	121,2	8,4	2	25,1
01-Jul-2004	17	3	1,5	56	1000	303	0,063	0,27	0,001	0,055	4,3	98,2	8,7	2	27,1
01-Ago-2004	16	3	1,8	30	180	317	0,014	0,29	0,001	0,023	4,2	99,4	8,6	2	27,4
07-Set-2004	17	3	1,9	10	120	335	0,023	0,09	0,001	0,023	4,4	86,2	8,5	2	23,9
06-Out-2004	14	3	1,3	220	820	340	0,016	0,49	0,001	0,025	3,7	92	8,5	2	22,7
02-Nov-2004	14	3	4,1	3	250	330	0,026	0,44	0,016	0,023	3,6	94	8,1	3,8	18,2
01-Dez-2004	18	3	3,9	6	11	328	0,016	2,97	0,101	0,023	3,7	131,9	8,3	2	11,7
11-Jan-2005	16	3	6,8	9	19	324	0,018	0,8	0,016	0,032	3,6	100,2	8,3	2,4	10,5
09-Fev-2005	18	3	3,4	2	2	327	0,019	0,8	0,014	0,038	3,7	138,6	8,1	2	9,8
08-Mar-2005	16	3	2,5	1	3	329	0,023	0,84	0,011	0,041	4,4	142,7	8,2	5	10,7
05-Abr-2005	23	3	2,5	0	4	334	0,281	0,66	0,007	0,028	4,3	130,7	8,6	5	17,6
03-Mai-2005	15	3	1,4	12	25	338	0,012	0,58	0,008	0,013	3,6	130,9	8,6	5	17,1
01-Jun-2005	20	3	2,4	2	28	348	0,018	0,58	0,001	0,025	4,6	116,5	8,6	5	24
01-Jul-2005	17	3	2,2	11	110	359	0,016	0,14	0,004	0,025	4,1	97,8	8,2	5	24,7
01-Jun-2006	15	3	3,3	0	21	355	0,021	0,27	0,015	0,023	3,3	103	8,8	5	23
01-Jul-2006	23	3	6,5	30	110	364	0,076	0,22	0,003	0,064	5,1	92	8,6	21	24,8
01-Ago-2006	21	3	1,6	6	140	375	0,289	0,09	0,002	0,032	4,4	95	8,7	11	26
05-Set-2006	19	3	11,5	70	1100	388	0,165	0,09	0,001	0,023	4,6	92	8,7	14	25,8
02-Out-2006	16	3	5,8	10	670	383	0,051	0,09	0,003	0,053	3,4	104	8,4	5	22,7
01-Nov-2006	16	3	9,7	110	380	378	0,033	2	0,044	0,023	2,9	92	8,1	6,6	20,3
01-Dez-2006	20	3	23,2	3300	3300	343	0,06	2	0,059	0,035	5,6	82	8	13	16,2
02-Jan-2007	15	3	5,6	0	110	347	0,048	2	0,025	0,028	4,7	117	8	5	13,7
01-Fev-2007	14	3	5,5	7	150	348	0,04	2	0,027	0,023	4,7	68	7,9	5	10,3
01-Mar-2007	17	3	8,5	2	1100	337	0,041	2	0,019	0,023	4,3	108	8,1	5	15,3
01-Abr-2007	18	3	5,1	20	20	337	0,014	2	0,016	0,023	4,1	95	8,2	5	14
02-Mai-2007	16	3	3,3	8	35	341	0,076	2	0,019	0,078	5,1	97	8,5	5	18,5
01-Jun-2007	12	3	5,6	22	200	352	0,019	2	0,003	0,023	5,1	112	8,6	5	20,1
01-Jul-2007	23	3	4,6	80	220	358	0,02	2	0,004	0,023	4,5	112	8,7	5	23,7
01-Ago-2007	18	3	5,5	7	40	370	0,016	2	0,002	0,023	4,1	93,7	8,6	5	20
04-Set-2007	13	3	3,5	30	60	385	0,022	2	0,004	0,023	4,3	54,3	8,6	5	24,7
01-Out-2007	25	3	2,8	18	80	379	0,014	2	0,004	0,023	4,4	102	8,2	5	21,2
01-Nov-2007	11	3	2,0	1	3	385	0,252	2	0,004	0,082	7	87,1	8,2	5	19,9
01-Dez-2007	25	3	1,6	9	11	377	0,02	2	0,021	0,032	5,2	96,9	8,1	5	14,3
02-Jan-2008	32	3	7,8	21	41	381	0,06	2	0,088	0,023	6,4	100	8,1	5	12,3
01-Fev-2008	28	3	29,1	2	6	385	0,034	2	0,013	0,023	6,6	103,4	8	5	10,9
01-Mar-2008	29	3	6,5	4	60	383	0,044	2	0,014	0,044	5,9	118	8	5	15,8
01-Abr-2008	20	3	3,5	3	54	391	0,105	2	0,016	0,131	5,2	76,7	8,3	5	16,1
01-Mai-2008	32	3	7,9	5	110	378	0,131	2	0,004	0,032	4,4	127,3	8,6	39	20,3
01-Jun-2008	33	3	5,6	2	8	377	0,025	2	0,002	0,023	4,4	81,9	8,1	5	18,9
01-Jul-2008	35	3	4,1	7	160	399	0,031	2	0,002	0,023	4,3	82,7	8,4	5,6	26,2
01-Ago-2008	39	3	3,8	1	130	406	0,035	2	0,002	0,023	4,5	105	8,6	7,1	27
02-Set-2008	26	3	3,9	37	220	412	0,01	2	0,002	0,023	4,4	29	8,1	5	24,5
01-Out-2008	20	3	8,9	0	180	422	0,081	2	0,021	0,023	3,9	99	8,1	5	22,2
01-Nov-2008	18	3	3,7	12	120	421	0,03	2	0,043	0,023	3,5	95	8,1	38	18
01-Dez-2008	17	3	7,2	3	12	414	0,03	2	0,004	0,023	3,5	141	8,3	5,9	15

Tabela I.2 – Dados meteorológicos e hidrométricos referentes à albufeira do Alvito.

Data	Vol Arm. dam ³	Cota m	HR _{am} %	HR _s %	HR ₁₀ %	PP _{am} mm	PP _s mm	PP ₁₀ mm	T _{am} °C	T _s °C	T ₁₀ °C	VV _{am} m/s	VV _s m/s	VV ₁₀ m/s	Rad _{am} W/m ²	Rad _s W/m ²	Rad ₁₀ W/m ²
05-Jun-2001	130740	197,30	69,10	54,00	54,20	0,64	0,00	0,00	18,06	22,80	23,61	2,08	2,34	2,16	6482	8022	7993
03-Jul-2001	132500	197,50	57,37	55,00	59,00	0,70	0,00	0,00	21,83	24,60	22,91	2,88	2,42	2,62	7742	7880	7715
01-Ago-2001	115982	196,07	60,65	64,40	60,10	0,02	0,00	0,00	22,00	22,44	22,69	3,04	2,18	2,52	7147	6549	7052
04-Set-2001	115724	195,99	62,13	60,60	66,60	0,03	0,00	0,03	23,16	23,24	22,43	2,43	2,06	2,13	6594	6155	5926
02-Out-2001	120239	196,20	68,23	80,80	76,90	2,09	9,64	5,50	20,68	19,06	18,51	2,14	2,20	2,31	4856	3658	3923
06-Nov-2001	123464	196,28	80,44	80,20	75,20	5,02	11,98	5,99	17,65	16,62	18,45	2,77	3,60	3,26	2893	2529	2469
04-Dez-2001	127334	196,63	72,80	84,00	78,00	2,50	0,06	1,10	10,89	9,36	10,02	3,33	1,52	1,76	2696	2060	2298
08-Jan-2002	130440	196,97	80,65	82,20	82,10	1,86	0,02	0,07	9,64	9,20	10,96	3,63	3,36	3,50	1742	2273	1659
05-Fev-2002	126500	197,22	84,72	86,17	86,00	0,70	3,94	2,18	10,54	11,44	11,45	2,16	2,20	1,93	2329	2413	2582
05-Mar-2002	132500	197,50	79,00	66,50	74,45	0,93	3,84	1,92	11,11	9,80	10,63	1,80	1,70	1,45	4758	4945	4897
02-Abr-2002	132500	197,50	73,29	77,20	74,40	3,15	1,12	0,56	16,89	14,58	16,43	1,74	1,64	1,50	4912	4325	4875
07-Mai-2002	132500	197,52	72,28	69,33	70,18	0,98	0,62	0,32	14,36	13,94	15,21	3,00	4,22	3,38	5710	6108	6459
04-Jun-2002	132500	197,41	74,86	72,00	65,50	0,51	0,00	0,00	16,86	19,42	18,50	1,95	1,44	1,99	5911	6168	6500
02-Jul-2002	127334	197,27	59,88	64,20	61,40	0,13	0,00	0,00	21,69	20,44	20,96	1,99	1,64	1,96	7623	7771	7815
01-Ago-2002	122174	197,13	59,56	52,60	58,40	0,02	0,00	0,00	22,08	25,42	23,51	2,12	1,80	1,80	7509	7323	7336
03-Set-2002	120884	196,57	63,39	69,83	67,91	2,42	0,46	0,93	22,17	20,00	21,18	1,78	1,30	1,54	6410	5267	6040
01-Out-2002	120884	196,53	80,08	81,80	82,00	2,48	0,38	1,10	19,33	19,74	19,49	1,82	1,10	1,32	4057	3114	3515
01-Nov-2002	124105	196,49	83,75	86,50	85,45	1,38	0,14	1,55	18,15	15,40	17,73	1,54	1,16	2,15	2609	2254	2297
01-Dez-2002	126044	196,70	86,13	90,00	90,10	3,04	6,62	5,59	12,79	11,02	10,68	1,35	1,42	1,44	2014	1631	1653
07-Jan-2003	132500	197,43	88,97	89,80	89,50	3,03	1,14	1,22	12,03	11,36	11,34	1,71	1,16	1,13	1542	1374	1491
04-Fev-2003	132500	197,63	78,07	73,60	70,80	1,49	0,00	0,09	9,19	8,18	9,89	2,05	2,20	2,56	2212	2889	2856
05-Mar-2003	132500	197,73	84,90	88,40	87,60	2,56	0,16	2,56	10,26	12,98	11,66	1,75	0,84	1,50	2675	3189	2724
01-Abr-2003	132500	197,60	79,19	86,40	83,90	1,33	4,46	3,14	13,36	13,40	13,63	1,93	1,88	2,34	3637	3644	3437
06-Mai-2003	132500	197,61	78,00	71,80	74,70	3,58	2,44	1,44	14,17	15,12	15,09	2,06	2,08	1,97	4813	6148	5735
03-Jun-2003	127334	197,50	62,10	66,40	58,00	0,44	0,28	0,14	19,18	20,42	20,12	2,15	1,90	2,30	7031	6011	6804
01-Jul-2003	127334	197,59	63,40	71,20	69,00	0,08	0,46	0,23	21,93	19,76	20,43	2,01	2,08	1,94	6989	7082	7101
01-Ago-2003	116369	196,23	63,30	63,00	62,40	0,04	0,00	0,00	21,58	23,42	22,37	2,07	2,40	2,44	7112	7366	7556
09-Set-2003	112074	195,80	64,97	68,00	71,00	0,00	0,00	0,00	22,76	20,28	20,70	1,79	2,06	1,77	6158	5841	5922
07-Out-2003	113254	195,90	63,17	66,40	74,30	2,04	2,14	6,12	21,63	18,60	19,05	1,99	2,38	2,18	4751	4253	3701
04-Nov-2003	118304	196,01	77,97	81,80	82,80	3,03	5,84	5,18	16,02	13,06	13,95	1,76	1,54	1,96	3130	2302	2405
02-Dez-2003	124109	196,27	83,70	88,40	87,70	2,44	1,98	3,06	12,97	10,32	10,56	1,72	1,16	1,56	1935	1547	1466
13-Jan-2004	125786	196,79	86,45	87,00	85,30	0,58	1,58	0,89	10,93	12,68	11,35	1,60	1,54	1,40	1757	1753	1938
10-Fev-2004	127072	196,92	89,52	92,00	87,70	1,24	0,10	0,73	11,08	12,06	12,28	1,41	1,14	1,17	1921	2741	2061
09-Mar-2004	130779	197,13	81,60	81,80	76,80	1,77	0,50	0,34	9,66	12,12	9,90	1,55	1,26	1,33	2897	3809	3591
06-Abr-2004	126500	197,20	78,52	78,00	82,10	2,08	1,86	3,91	12,15	13,28	11,77	1,94	1,42	1,64	4122	4897	3760
04-Mai-2004	132500	197,13	69,13	74,20	69,00	0,40	1,52	0,76	14,01	12,38	15,20	1,79	1,82	1,67	5455	4410	5427
01-Jun-2004	126044	197,06	71,90	69,60	74,64	1,39	0,00	2,56	16,21	19,74	18,35	1,73	1,80	1,48	5613	7335	6294
01-Jul-2004	121529	196,85	58,83	45,60	58,90	0,00	0,00	0,00	23,77	27,86	24,70	1,67	1,84	1,85	7666	7846	7705
01-Ago-2004	118304	196,64	50,97	40,00	51,20	0,01	0,00	0,00	24,66	29,64	26,25	1,94	1,90	1,84	7476	6875	7238
07-Set-2004	117040	196,40	65,71	77,00	70,00	0,56	1,08	0,55	22,28	19,56	20,66	1,60	1,04	1,45	5591	4611	5340
06-Out-2004	119336	196,26	58,50	48,40	51,10	0,00	0,00	0,00	22,25	23,24	22,88	1,34	1,14	1,10	4973	4652	4812
02-Nov-2004	118304	196,37	75,13	85,20	86,40	3,94	3,14	7,76	17,09	12,68	14,12	1,47	0,86	1,14	3083	2774	2571
01-Dez-2004	118304	196,34	77,03	84,20	82,70	0,94	1,76	0,88	12,73	13,16	12,59	1,55	1,60	1,34	2423	1708	2012
11-Jan-2005	118046	196,30	79,55	82,80	85,60	0,17	0,00	0,05	9,01	7,18	7,51	2,18	1,36	1,51	1939	2228	2135
09-Fev-2005	118046	196,25	75,43	81,20	74,30	0,06	0,34	0,17	7,71	6,76	6,83	1,85	1,58	1,84	2585	2611	2770
08-Mar-2005	116756	196,20	66,07	55,00	61,60	0,36	0,00	0,88	7,93	8,14	7,31	2,56	3,10	3,22	3479	4309	3422
05-Abr-2005	115982	196,20	74,16	75,00	79,10	1,14	3,02	2,24	13,66	15,58	14,74	2,15	2,26	2,10	3883	4289	3956
03-Mai-2005	114080	196,08	70,44	62,60	64,00	0,10	0,00	0,01	15,30	18,20	17,14	2,15	1,74	1,87	5681	6081	6193
01-Jun-2005	113844	195,96	63,61	68,00	60,40	1,05	3,12	1,56	18,25	19,92	19,55	1,95	1,88	1,95	6412	5947	6568
01-Jul-2005	103814	195,79	58,68	67,20	56,10	1,10	0,00	0,00	23,03	21,04	24,24	1,78	1,88	1,75	6785	7004	6719
01-Jun-2006	75772	194,67	61,77	52,60	55,70	0,00	0,02	0,01	18,88	23,74	20,56	1,79	1,54	1,88	6650	6638	6866
01-Jul-2006	73992	194,30	64,13	61,80	69,20	0,10	0,00	0,20	22,01	22,34	21,28	1,14	1,68	1,38	6353	7300	6610
01-Ago-2006	71945	173,48	58,20	58,80	58,80	0,02	0,00	0,05	24,04	23,66	24,66	1,20	1,50	1,24	7038	7599	7111
05-Set-2006	70492	191,83	53,61	43,60	44,70	0,11	0,00	0,00	25,05	27,24	27,30	1,33	1,04	1,30	6363	6161	6275
02-Out-2006	78175	191,77	66,60	86,40	82,90	1,97	0,90	1,04	21,72	19,50	19,33	1,27	0,96	1,11	4615	3363	3787
01-Nov-2006	83847	192,65	80,20	82,40	87,40	6,64	4,68	14,46	18,73	18,30	18,35	1,37	1,38	1,50	2926	2493	2168
01-Dez-2006	91608	193,55	84,90	84,20	87,30	4,00	6,58	3,74	9,17	14,00	14,16	0,92	1,50	1,06	2187	1955	1985
02-Jan-2007	90949	194,04	82,45	89,60	84,00	0,90	0,06	0,03	9,64	11,66	10,06	1,63	1,46	1,63	1923	2092	2083
01-Fev-2007	97519	194,08	89,87	80,20	83,70	0,54	1,66	1,39	8,90	6,36	7,02	1,44	2,50	1,93	2038	1843	2096
01-Mar-2007	97519	194,63	89,81	90,00	87,30	3,31	1,52	2,52	10,54	11,98	10,87	1,74	1,38	1,64	2359	3002	2797
01-Abr-2007	97746	194,65	75,41	78,60	72,20	0,90	2,34	1,17	12,02	10,60	10,72	2,05	1,60	2,50	4216	4613	4659
02-Mai-2007	97292	194,69	78,17	77,00	76,40	2,60	3,60	2,67	14,06	13,30	14,92	1,69	2,18	1,93	4953	5644	5343
01-Jun-2007	97519	194,63	72,10	78,20	77,30	1,01	0,12	0,46	16,71	14,96	15,88	1,67	1,66	1,57	5928	5524	5069
01-Jul-2007	95018	194,52	72,10	66,20	76,80	0,63	0,16	1,69	18,87	19,12	18,59	1,51	1,82	1,35	6367	7121	5737
01-Ago-2007	94223	194,31	60,43	63,20	63,70	0,06	0,20	0,10	21,78	19,26	20,52	1,96	2,56	2,07	7609	7502	7377
04-Set-2007	94419	194,04	60,39	58,60	66,70	0,21	0,00	0,56	22,49	23,54	22,32	1,60	1,06	1,12	6313		

Tabela I.3 – Dados da qualidade da água da albufeira de Odivelas.

Data	CQO	CBO ₅	Clorofila-a	Coliformes fecais	Coliformes totais	Condutividade	Fosforo total	Nitratos	Nitritos	Ortofosfatos	Oxidabilidade	OD	pH	SST	Temperatura
	mg/L	mg/L	µg/L	ufc/100 mL	ufc/100 mL	µS/cm	mg/L	mg/L	mg/L	mg/L P ₂ O ₅	mg/L	%	E. Sorensen	mg/L	°C
05-Jun-2001	22,9	3,58	10,7	40	13000	280	0,02	1,28	0,033	0,013	6,1	104,8	9,06	1,3	24,8
03-Jul-2001	26,8	5,07	15,9	0	100	295	0,051	0,13	0,003	0,008	6,6	85,3	9,09	7,2	25,5
01-Ago-2001	22,4	2	10,9	130	500	304	0,015	0,53	0,004	0,044	5,42	88,2	8,41	4,2	25,3
04-Set-2001	16	2	4,6	72	2700	308	0,013	0,4	0,005	0,011	5,1	83,4	8,26	4	25,3
02-Out-2001	12,4	3,66	9,6	5	230	309	0,012	0,38	0,003	0,022	5,2	75,8	8,23	5,6	22,3
06-Nov-2001	12,9	2,38	16,9	180	370	301	0,012	0,53	0,004	0,009	5,9	68,5	8,03	4,8	19,2
04-Dez-2001	13,7	2,05	13,6	19	19	299	0,06	0,84	0,02	0,058	7,9	60,7	7,97	4,4	13
08-Jan-2002	24	5	27,2	8	120	301	0,093	0,49	0,019	0,065	7,6	79,3	8,1	5,6	12,3
05-Fev-2002	15	3	5,0	4	130	306	0,022	0,71	0,018	0,03	5,9	82,1	7,9	3,4	12,7
05-Mar-2002	16	3	2,3	2	5	310	0,014	0,44	0,019	0,019	8,2	89,7	7,9	2	14,2
02-Abr-2002	42	3	39,4	45	110	307	0,07	0,44	0,006	0,025	9,3	99,4	8,6	20	19
07-Mai-2002	22	3	1,6	12	3115	324	0,021	0,31	0,007	0,013	4,7	78,9	8,3	5,6	16,6
04-Jun-2002	17	5	10,5	11	40	323	0,039	0,27	0,006	0,02	6,7	90	8,2	5	21,6
02-Jul-2002	21	3	1,3	1	5	334	0,036	0,22	0,006	0,008	5,9	62	8,2	2,1	26
01-Ago-2002	13	4	2,0	180	240	343	0,032	0,13	0,003	0,04	5,9	70	8,3	4	25,1
03-Set-2002	10	3	1,9	3	17	350	0,026	0,22	0,001	0,024	5,5	100	8	1,3	23
01-Out-2002	16	3	6,0	13	70	343	0,033	0,18	0,001	0,053	5,1	99,9	8,4	4,2	21,9
01-Nov-2002	24	3	5,9	41	240	346	0,058	0,75	0,021	0,128	7,8	87,7	8,1	5,6	21,3
01-Dez-2002	13	3	3,8	66	600	344	0,057	0,58	0,007	0,05	5,2	69,1	7,7	13	14,8
07-Jan-2003	19	3	3,8	35	120	341	0,059	1,02	0,016	0,07	5,4	83,6	7,9	70	12
04-Fev-2003	11	4	3,6	130	250	347	0,038	1,28	0,023	0,045	7,6	89,9	7,8	7,1	12
05-Mar-2003	27	3	12,2	4	170	344	0,041	1,42	0,019	0,02	6,6	98,5	8,5	2,9	15,6
01-Abr-2003	24	3	9,8	350	410	360	0,031	1,33	0,011	0,046	5,8	79,7	8,2	2,2	16,6
06-Mai-2003	51	3	5,0	14	14	358	0,027	1,24	0,011	0,036	5,7	103	8,1	4,1	17,5
03-Jun-2003	50	3	3,4	19	110	365	0,023	0,97	0,007	0,033	5,6	61,9	8,6	2,6	20,9
01-Jul-2003	34	3	19,3	1	20	378	0,026	1,06	0,002	0,026	6,2	65,4	8,3	5,8	24
01-Ago-2003	23	3	4,3	0	180	386	0,161	1,11	0,002	0,087	6,2	94	8,7	6,1	26,6
09-Set-2003	13	3	2,1	570	19000	393	0,036	0,93	0,002	0,018	5,8	89,4	7,9	32	22,8
07-Out-2003	16	3	1,7	4	40	389	0,039	1,11	0,004	0,023	5,6	87,3	8,1	5,2	21,9
04-Nov-2003	21	3	4,0	90	1100	383	0,043	0,66	0,009	0,036	4,9	92,8	7,8	6,1	16,8
02-Dez-2003	20	3	1,1	49	150	372	0,061	1,2	0,221	0,073	6,8	77,3	7,8	7,7	14,8
13-Jan-2004	19	3	1,3	54	110	393	0,081	1,48	0,019	0,074	4,2	65,6	8,1	4,7	11,3
10-Fev-2004	16	3	1,7	11	17	376	0,019	1,68	0,021	0,043	4,8	69	7,6	8,7	13,6
09-Mar-2004	18	3	1,3	2	2	386	0,024	1,48	0,016	0,023	5	92,1	8	2,9	15,3
06-Abr-2004	18	4	2,1	3	40	385	0,01	0,68	0,018	0,023	5,7	85,6	8,5	2,4	18,5
04-Mai-2004	21	3	1,4	12	40	395	0,027	0,62	0,017	0,035	5,8	91,4	8	2,4	18,7
01-Jun-2004	20	3	2,4	8	24	382	0,089	0,21	0,005	0,023	4,8	78,6	8,5	2,4	23,4
01-Jul-2004	21	3	1,6	10	80	391	0,048	0,34	0,004	0,023	2,4	65,2	8,7	4,2	26
01-Ago-2004	20	3	1,8	6	120	401	0,023	0,45	0,003	0,035	6,4	82,1	8,6	7,6	25
07-Set-2004	21	3	1,9	1	13	425	0,029	0,15	0,001	0,023	5,8	80,4	8,1	2	24,2
06-Out-2004	21	3	1,9	3	30	432	0,022	0,49	0,002	0,023	5,2	74	8,1	3,3	22
02-Nov-2004	17	3	1,7	1	23	424	0,025	0,49	0,024	0,023	4,9	64,7	7,8	3,5	18,6
01-Dez-2004	23	3	2,2	21	60	421	0,035	0,53	0,037	0,043	4,3	63,1	8	3,6	14,6
11-Jan-2005	20	3	1,9	15	330	422	0,026	0,62	0,028	0,032	4,8	74,7	8,1	2,6	12,6
09-Fev-2005	19	3	2,1	5	130	424	0,03	0,62	0,017	0,032	4,7	80,5	8,2	2	11,8
08-Mar-2005	19	3	2,0	3	300	426	0,027	0,66	0,017	0,057	5,3	92,7	8	5	11
05-Abr-2005	24	3	4,3	0	1500	430	0,482	0,66	0,012	0,032	5,4	92,8	8,6	5	16,8
03-Mai-2005	20	3	2,1	240	250	437	0,016	0,58	0,014	0,014	4,6	81,9	8,1	5	20,9
01-Jun-2005	24	3	3,0	24	49	447	0,06	0,58	0,004	0,072	5,4	79,2	8,5	5	25
01-Jul-2005	19	3	2,2	34	80	460	0,02	0,16	0,001	0,023	4,4	80,3	8,4	7,1	26
01-Jun-2006	22	3	10,3	22	120	437	0,037	0,09	0,026	0,023	5,4	70	8,3	6,4	25
01-Jul-2006	23	3	6,5	0	400	425	0,056	0,09	0,005	0,03	5,2	87	8,5	14	24
01-Ago-2006	22	3	5,4	0	30	438	0,277	0,09	0,003	0,031	5,5	89,7	8,6	5	27
05-Set-2006	19	3	1,8	3	16	453	0,105	0,09	0,001	0,039	4,3	83,9	8,6	7,7	29,1
02-Out-2006	16	3	3,7	72	72	453	0,054	0,09	0,001	0,085	4,7	76,2	8,4	7,9	23
01-Nov-2006	20	3	82,6	220	2100	443	0,059	2	0,069	0,056	4,9	87	8,2	7,1	21
01-Dez-2006	22	3	7,4	38	150	362	0,11	2	0,048	0,182	7,7	60,5	7,7	6,9	20
02-Jan-2007	23	3	12,6	18	46	374	0,092	2	0,032	0,133	6,1	97	8,2	8,8	17,4
01-Fev-2007	21	3	13,5	15	39	383	0,072	2,8	0,015	0,031	6,4	78,8	8,2	8,5	14,2
01-Mar-2007	19	3	6,5	21	21	381	0,059	2,2	0,025	0,102	5,9	83,2	8,2	5	20,3
01-Abr-2007	18	3	3,7	10	10	390	0,025	2	0,035	0,023	5,9	113,8	8,3	5	18,3
02-Mai-2007	21	3	6,9	19	160	396	0,032	2	0,042	0,023	5,1	65,8	8,5	5	19,3
01-Jun-2007	20	3	9,6	47	190	401	0,048	2	0,011	0,023	5,1	95,2	8,6	8,1	23,5
01-Jul-2007	24	3	4,1	27	390	413	0,085	2	0,003	0,023	4,3	75,7	8,6	31	24,9
01-Ago-2007	18	3	7,9	1	24	427	0,026	2	0,002	0,023	4,1	68,9	8,5	5	24,6
04-Set-2007	21	3	4,8	28	110	469	0,027	2	0,003	0,032	4,6	72,1	8,3	5,4	26
01-Out-2007	25	3	4,3	17	90	442	0,027	2	0,002	0,023	5,9	68,1	8,2	5	22,1
01-Nov-2007	22	3	15,0	0	29	445	0,127	2	0,008	0,023	5,6	83,5	8,3	5	22,7
01-Dez-2007	27	3	3,4	27	660	494	0,056	2,7	0,021	0,023	6,4	66,1	7,9	8,6	17,5
02-Jan-2008	39	3	10,6	32	140	480	0,039	2,4	0,025	0,034	5,9	71	7,9	8,3	14,6
01-Fev-2008	1	2,3	350,7	5	40	455	2,919	2	0,029	0,023	23,2	79	7,3	350	12,6
01-Mar-2008	41	3	23,3	9	33	455	0,48	4,6	0,03	0,023	5,4	89,6	7,9	9,4	18,6
01-Abr-2008	20	3	5,1	4	23	430	0,163	2	0,02	0,134	5,6	87,9	8,2	6,5	20
01-Mai-2008	33	3	254,0	21	52	426	0,088	2	0,025	0,023	4,6	75,1	8,2	5	22
01-Jun-2008	37	3	2,4	31	50	420	0,042	2	0,027	0,047	4,6	76,5	7,7	5	18,9
01-Jul-2008	31	3	1,6	7	1700	444	0,024	2	0,005	0,023	5,1	86,6	8,1	5	27,8
01-Ago-2008	38	3	3,0	1	10	456	0,035	2	0,002	0,023	4,9	97,8	8,6	5,9	28,9
02-Set-2008	33	3	3,3	240	410	468	0,01	2	0,004	0,023	5,1	90,3	7,8	5,9	25
01-Out-2008	26	3	14,0	1	110	469	0,012	2	0,003	0,023	4,5	66	8,3	5	24,5
01-Nov-2008	20	3	4,1	4	120	468	0,02	2	0,023	0,023	4,8	73	8,1	5,1	18
01-Dez-2008	18	3	1,8	10	14	465	0,04	2	0,065	0,023	4,5	80,5	8	6,7	14

Tabela I.4 – Dados meteorológicos e hidrométricos referentes à albufeira de Odivelas.

Data	Vol Arm. dam ³	Cota m	HR _{am} %	HR ₅ %	HR ₁₀ %	PP _{am} mm	PP ₅ mm	PP ₁₀ mm	T _{am} °C	T ₅ °C	T ₁₀ °C	VV _{am} m/s	VV ₅ m/s	VV ₁₀ m/s	Rad _{am} W/m ²	Rad ₅ W/m ²	Rad ₁₀ W/m ²
05-Jun-2001	85444	102,58	66,65	56,80	51,40	0,36	0,00	0,00	19,13	22,94	24,84	0,55	0,76	0,58	5992	7275	7271
03-Jul-2001	77641	102,58	57,97	55,80	59,90	0,79	0,00	0,00	22,39	25,00	23,29	0,83	0,90	0,84	7112	7153	7086
01-Ago-2001	79810	100,97	59,97	63,60	60,10	0,02	0,00	0,00	22,64	22,10	23,24	1,04	0,72	0,87	6544	6146	6503
04-Set-2001	76081	101,13	61,61	60,20	66,20	0,02	0,00	0,00	23,50	23,66	22,90	0,76	0,76	0,69	6002	5655	5388
02-Out-2001	77313	100,86	68,50	79,80	77,70	1,38	4,72	3,39	20,92	19,62	18,73	0,59	0,60	0,55	3916	3207	3646
06-Nov-2001	81408	101,04	79,71	78,40	74,30	3,60	5,00	2,50	17,98	16,86	18,79	0,66	0,80	0,83	2595	2154	2069
04-Dez-2001	83090	101,51	74,47	85,20	82,30	1,37	0,06	0,97	11,01	9,26	9,85	0,60	0,18	0,21	2306	1735	1954
08-Jan-2002	86709	101,88	80,48	80,20	82,30	2,61	0,08	1,77	9,96	9,90	11,34	0,80	0,66	0,87	1488	2008	1482
05-Fev-2002	88037	102,17	84,35	86,80	85,10	1,85	3,18	1,82	10,93	11,62	11,87	0,46	0,48	0,42	2003	2241	2337
05-Mar-2002	91423	102,21	82,33	78,25	75,33	4,38	6,43	4,88	9,97	9,27	9,04	0,54	0,55	0,63	2473	2604	2828
02-Abr-2002	94096	102,55	78,58	77,80	74,70	2,44	0,26	0,15	13,86	14,88	16,53	0,56	0,46	0,40	3868	4083	4621
07-Mai-2002	90176	102,75	70,23	63,40	66,60	1,37	0,32	0,18	15,14	15,14	15,39	0,59	0,58	0,65	5602	5588	6003
04-Jun-2002	83678	102,25	70,38	74,00	75,75	0,30	0,00	0,00	16,08	17,00	18,40	0,46	0,55	0,45	5174	5892	6145
02-Jul-2002	75014	101,54	62,33	63,40	62,40	0,10	0,00	0,01	20,60	20,84	20,90	0,84	0,78	0,90	6317	6789	6538
01-Ago-2002	70745	100,59	55,53	54,20	59,20	0,03	0,00	0,04	23,41	25,46	17,64	0,96	0,96	0,83	6549	6671	5937
03-Set-2002	68275	99,96	62,23	61,60	62,60	0,02	0,00	0,00	22,35	21,86	21,41	0,82	0,50	0,53	5955	5442	5176
01-Out-2002	67502	99,77	78,40	81,40	82,20	2,15	0,30	0,66	19,98	21,04	20,08	0,68	0,34	0,45	3982	3368	3385
01-Nov-2002	68345	99,68	81,93	83,20	79,80	1,65	0,14	2,42	18,32	17,22	17,71	0,52	0,20	0,73	3009	3431	3335
01-Dez-2002	73208	99,60	85,84	88,20	89,70	2,79	8,06	5,53	13,95	11,06	11,28	0,48	0,70	0,66	2300	1837	1634
07-Jan-2003	82080	100,47	89,45	90,00	89,40	2,91	2,86	1,85	12,42	11,84	11,83	0,78	0,56	0,60	1530	1544	1568
04-Fev-2003	89285	101,50	77,39	74,80	71,10	1,78	0,02	0,01	8,90	8,18	10,20	0,56	0,56	0,45	2435	3392	3291
05-Mar-2003	94809	102,53	85,18	86,20	86,90	2,40	0,12	2,34	10,63	13,66	12,17	0,43	0,16	0,45	2717	3262	2809
01-Abr-2003	95700	102,99	78,97	86,80	84,50	1,31	4,66	3,37	13,68	13,78	14,11	0,52	0,60	0,69	3735	3906	3453
06-Mai-2003	93027	103,17	78,20	72,00	74,80	3,03	2,50	1,66	14,69	15,64	15,62	0,67	0,64	0,60	4605	5783	5513
03-Jun-2003	92671	102,54	61,48	66,40	56,00	0,45	0,24	0,12	19,69	21,02	20,95	0,66	0,72	0,76	6683	5664	6457
01-Jul-2003	78298	101,65	61,27	68,20	66,20	0,03	0,16	0,08	22,65	20,74	21,31	0,84	0,94	0,87	6469	6337	6501
01-Ago-2003	78134	101,05	62,37	61,60	61,10	0,10	0,00	0,00	22,21	23,98	23,06	0,99	1,14	1,20	6517	6856	7016
09-Set-2003	78052	101,23	64,97	67,40	70,20	0,01	0,00	0,00	23,08	20,78	20,96	0,86	1,08	0,87	5558	5635	5275
07-Out-2003	77970	100,97	60,27	64,40	73,30	2,32	1,90	6,96	22,21	19,18	19,55	0,77	0,68	0,79	4595	4243	3472
04-Nov-2003	80062	101,02	77,55	83,40	83,80	2,84	5,82	4,91	16,28	13,16	14,15	0,64	0,74	0,96	3160	2349	2477
02-Dez-2003	80463	101,25	84,57	89,40	89,50	2,10	0,54	1,00	13,12	10,54	10,73	0,57	0,36	0,45	1960	1570	1483
13-Jan-2004	84688	101,70	87,48	96,40	90,90	0,71	0,72	0,56	10,95	12,78	11,38	0,39	0,20	0,24	1840	1549	1896
10-Fev-2004	87503	101,85	90,71	92,00	92,20	1,43	0,24	1,11	11,38	11,94	13,01	0,35	0,32	0,40	1880	2531	2300
09-Mar-2004	88216	102,10	81,62	81,40	76,20	1,94	0,42	0,27	9,81	12,42	10,26	0,51	0,36	0,39	2915	3633	3582
06-Abr-2004	86622	102,18	79,10	76,60	81,40	1,85	1,92	3,40	12,51	14,12	12,44	0,61	0,44	0,59	4226	5442	4045
04-Mai-2004	83426	101,96	67,57	72,00	66,80	0,39	1,80	0,90	14,57	12,94	15,30	0,59	0,84	0,65	5237	3802	4962
01-Jun-2004	76574	101,59	70,81	68,00	72,60	0,21	0,00	0,22	16,78	20,46	18,86	0,77	1,02	0,81	5039	6805	5967
01-Jul-2004	67994	100,80	58,58	47,80	59,10	0,01	0,00	0,04	23,89	27,80	25,07	0,93	0,90	1,04	6911	7190	7012
01-Ago-2004	61974	99,87	50,47	37,40	49,60	0,00	0,00	0,00	24,98	29,94	26,65	1,10	1,00	1,02	6800	6309	6596
07-Set-2004	58168	98,67	65,19	72,20	69,50	0,34	0,38	0,29	22,67	20,00	20,99	0,86	0,56	0,89	5187	3979	4853
06-Out-2004	57337	98,20	57,67	48,00	52,20	0,02	0,00	0,01	22,19	22,32	22,18	0,67	0,42	0,47	4875	4486	4674
02-Nov-2004	56853	98,15	74,00	85,60	86,40	3,10	2,66	5,89	9,25	12,94	14,47	0,75	0,50	0,62	3041	2736	2492
01-Dez-2004	56991	98,09	76,58	83,20	81,00	0,96	2,06	1,04	12,57	13,30	12,64	0,43	0,70	0,46	2503	1579	1965
11-Jan-2005	55809	98,03	78,32	78,40	82,70	0,15	0,06	0,11	8,92	7,26	7,47	0,55	0,22	0,23	2060	2470	2159
09-Fev-2005	54183	97,82	70,94	77,20	68,50	0,10	0,26	0,18	7,82	7,22	7,31	0,39	0,36	0,39	2700	2722	2990
08-Mar-2005	53702	97,59	61,18	52,00	58,80	0,86	0,02	1,56	8,39	8,56	7,76	0,73	1,02	1,03	3584	4651	3580
05-Abr-2005	52738	97,49	72,20	74,00	77,30	1,17	2,42	2,24	14,27	16,20	15,48	0,77	0,86	0,85	3782	3882	3841
03-Mai-2005	48302	96,86	66,85	66,60	70,40	0,06	0,02	0,02	14,96	16,42	15,52	0,90	0,84	0,87	5245	5546	5417
01-Jun-2005	44180	96,08	61,83	66,80	58,90	1,11	1,78	0,94	18,87	20,40	20,05	0,98	0,94	0,96	5992	5281	6034
01-Jul-2005	39624	94,85	57,77	66,40	55,00	0,93	0,00	0,00	23,31	21,60	24,63	0,91	1,00	0,91	6199	6298	6328
01-Jun-2006	55447	95,23	59,10	45,60	50,80	0,01	0,02	0,02	19,33	25,02	21,52	0,78	0,78	0,91	6301	6222	6550
01-Jul-2006	49507	97,87	61,35	59,80	67,30	0,90	0,00	1,51	22,61	22,68	21,68	0,89	1,26	1,14	5752	6459	5999
01-Ago-2006	40621	96,41	56,20	58,80	57,80	0,08	0,00	0,19	24,35	23,92	25,03	0,90	1,20	1,02	6210	6756	6298
05-Set-2006	37104	94,86	52,97	44,60	44,50	0,14	0,00	0,00	25,20	26,70	26,92	0,87	0,66	0,76	5804	5653	5798
02-Out-2006	45190	94,20	66,23	85,20	81,70	1,37	0,54	0,72	21,89	19,86	19,83	0,80	0,46	0,61	4450	3302	3658
01-Nov-2006	53039	95,99	78,43	78,80	85,00	6,92	3,70	13,71	19,12	18,86	18,81	0,96	1,02	1,21	2885	2633	2172
01-Dez-2006	59828	97,80	85,35	83,60	87,30	4,07	4,56	2,54	16,23	14,50	14,49	0,59	1,28	0,83	2037	1635	1644
02-Jan-2007	60244	98,52	82,26	88,60	82,40	0,81	0,12	0,07	9,55	11,98	10,17	0,38	0,34	0,39	1919	1982	1974
01-Fev-2007	64554	98,57	88,90	79,00	84,30	0,78	1,78	2,03	8,88	6,46	6,93	0,27	0,72	0,54	1912	1980	2094
01-Mar-2007	64484	99,19	88,90	89,80	86,80	2,07	0,62	1,36	10,89	12,28	11,20	0,65	0,60	0,75	2335	2979	2907
01-Abr-2007	64203	99,20	72,75	78,00	70,00	0,46	0,62	1,36	12,53	10,50	11,19	0,69	0,48	0,89	4070	4135	4385
02-Mai-2007	62043	98,84	76,13	74,80	74,30	2,56	3,98	2,92	14,59	13,84	15,48	0,72	1,10	0,92	4899	5445	5224
01-Jun-2007	57752	98,22	69,83	76,40	74,80	1,53	0,60	2,13	17,25	15,66	16,44	0,82	1,02	0,98	5678	4923	4783
01-Jul-2007	49507	98,05	69,88	86,60	78,10	0,84	2,44	1,95	19,61	20,08	19,56	0,92	1,03	0,96	5604	4809	4276
01-Ago-2007	42737	97,18	56,90	60,20	61,40	0,03	0,12	0,06	22,97	20,04	21,24	1,26	1,76	1,41	6294	6384	6221
04-Set-2007	39120	95,32	57,68	56,00	64,60	0,29	0,00	0,69	22,98	24,06	22,69	0,97	0,68	0,66	5657	5610	5328
01-Out-2007	37272	94,26	70,33	67,80	67,												

ANEXO II

Neste anexo apresentam-se os pesos associados às conexões entre os nodos das camadas adjacentes para a RNA apresentada na Figura 28. Os nodos das camadas intermédias são designados genericamente por h_{ij} , denotando o nodo i da camada intermédia j .

Tabela II.1 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura 28.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$	$h_{5,1}$	$h_{6,1}$	$h_{7,1}$	$h_{8,1}$	$h_{9,1}$
CQO	-21,20	-18,42	11,16	14,81	12,44	2,42	14,43	11,58	18,46
Oxidabilidade	-7,41	-5,84	8,88	-0,18	-5,90	6,58	8,27	19,95	2,06
pH	27,41	32,81	14,68	-1,17	10,09	18,32	1,87	2,93	1,68
Bias	2,84	-18,06	-13,68	-2,07	1,05	-12,96	-8,52	-7,67	-5,40

Tabela II.2 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura 28.

	$h_{1,2}$	$h_{2,2}$	$h_{3,2}$
$h_{1,1}$	5,71	16,47	5,35
$h_{2,1}$	-14,54	-1,53	-0,80
$h_{3,1}$	-1,74	-7,48	-4,25
$h_{4,1}$	4,81	-4,96	-1,34
$h_{5,1}$	-1,05	-7,99	0,61
$h_{6,1}$	-7,78	-4,92	-1,22
$h_{7,1}$	5,37	-6,20	-6,53
$h_{8,1}$	4,64	1,52	-12,81
$h_{9,1}$	5,35	-6,72	-3,40
Bias	2,43	8,49	4,08

Tabela II.3 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura 28.

	Classe B	Classe C
$h_{1,2}$	6,94	-7,00
$h_{2,2}$	6,46	-6,50
$h_{3,2}$	7,17	-7,25
Bias	-9,79	9,86

ANEXO III

Neste anexo apresentam-se os modelos relativos à previsão da qualidade da água da albufeira do Alvito, obtidos com recurso a AD e RNAs e com restrição de alguns dos parâmetros de entrada. Apresentam-se, ainda, o conjunto de regras geradas destacando-se, no final, o suporte e a confiança da regra (AD) e os pesos associados às conexões entre os nodos das camadas adjacentes (RNAs). Os nodos das camadas intermédias são designados genericamente por h_j , denotando o nodo i da camada intermédia j .

III.1 – Modelos excluindo o parâmetro pH

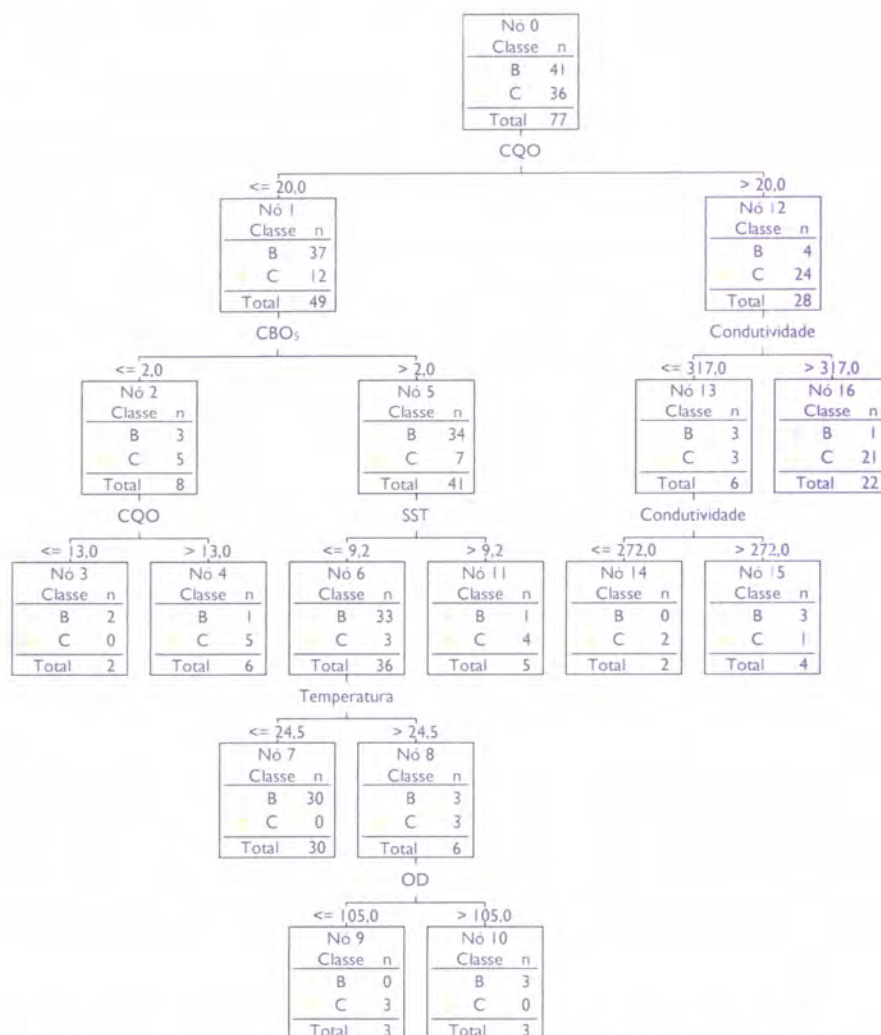


Figura III.1 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro pH.

Tabela III.1 – Conjunto de regras relativo ao modelo apresentado na Figura III.1.

Regras para a Classe B	Regras para a Classe C
<p>Regra 1</p> <p>Se CQO \leq 20,0 e CBO₅ \leq 2,0 e CQO \leq 13,0</p> <p>Então → B (2; 1,0)</p>	<p>Regra 1</p> <p>Se CQO \leq 20,0 e CBO₅ \leq 2,0 e CQO $>$ 13,0</p> <p>Então → C (6; 0,83)</p>
<p>Regra 2</p> <p>Se CQO \leq 20,0 e CBO₅ $>$ 2,0 e SST \leq 9,2 e Temperatura \leq 24,5</p> <p>Então → B (30; 1,0)</p>	<p>Regra 2</p> <p>Se CQO \leq 20,0 e CBO₅ $>$ 2,0 e SST \leq 9,2 e Temperatura $>$ 24,5 e Oxigénio Dissolvido \leq 105,0</p> <p>Então → C (3; 1,0)</p>
<p>Regra 3</p> <p>Se CQO \leq 20,0 e CBO₅ $>$ 2,0 e SST \leq 9,2 e Temperatura $>$ 24,5 e Oxigénio Dissolvido $>$ 105,0</p> <p>Então → B (3; 1,0)</p>	<p>Regra 3</p> <p>Se CQO \leq 20,0 e CBO₅ $>$ 2,0 e SST $>$ 9,2</p> <p>Então → C (5; 0,80)</p>
<p>Regra 4</p> <p>Se CQO $>$ 20,0 e Condutividade \leq 317,0 e Condutividade $>$ 272,0</p> <p>Então → B (4; 0,75)</p>	<p>Regra 4</p> <p>Se CQO $>$ 20,0 e Condutividade \leq 317,0 e Condutividade \leq 272,0</p> <p>Então → C (2; 1,0)</p>
	<p>Regra 5</p> <p>Se CQO $>$ 20,0 e Condutividade $>$ 317,0</p> <p>Então → C (22; 0,95)</p>

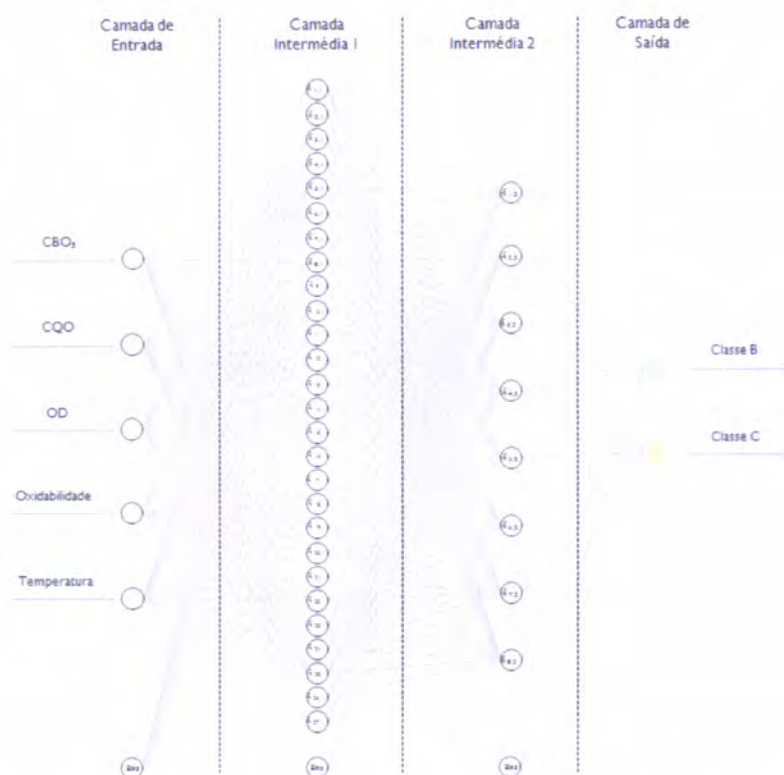


Figura III.2 – RNA obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro pH.

Tabela III.2 – Pesos associados às conexões entre a camada de entrada e a camada intermédia I para o modelo apresentado na Figura III.2.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$	$h_{5,1}$	$h_{6,1}$	$h_{7,1}$	$h_{8,1}$	$h_{9,1}$
CBO ₅	-5,66	0,35	-4,46	-12,39	-4,29	-5,58	8,37	-2,91	-3,12
CQO	6,04	-0,18	3,35	16,97	2,66	8,10	-5,24	0,32	1,81
OD	-3,05	-0,11	-1,93	-8,85	-3,52	-9,11	2,21	-2,64	-1,62
Oxidabilidade	0,97	-0,29	0,98	4,01	-3,70	4,99	-0,09	-1,06	1,19
Temperatura	2,35	-1,33	4,42	3,88	8,91	4,03	-3,88	6,54	3,93
Bias	-1,20	-0,69	-2,65	-0,86	-2,82	1,16	0,98	-2,25	-2,62

	$h_{10,1}$	$h_{11,1}$	$h_{12,1}$	$h_{13,1}$	$h_{14,1}$	$h_{15,1}$	$h_{16,1}$	$h_{17,1}$	$h_{18,1}$
CBO ₅	-3,37	-0,48	-0,29	1,04	1,38	-5,71	3,67	0,28	2,09
CQO	2,81	0,59	2,25	-0,85	-2,86	5,13	-4,13	-134,00	-2,09
OD	-2,94	-0,96	2,31	-1,26	-0,52	-3,35	5,98	0,93	1,36
Oxidabilidade	0,63	-0,41	6,19	-1,57	-0,04	1,41	-0,39	1,26	1,13
Temperatura	3,96	-1,29	-3,80	0,37	3,64	2,79	-1,15	2,25	1,33
Bias	-2,17	-0,04	-2,26	-0,45	-0,91	-0,53	-2,83	-1,94	-1,49

	$h_{10,1}$	$h_{20,1}$	$h_{21,1}$	$h_{22,1}$	$h_{23,1}$	$h_{24,1}$	$h_{25,1}$	$h_{26,1}$	$h_{27,1}$
CBO ₅	3,43	5,46	6,04	-3,10	2,00	5,65	-4,70	-7,53	0,89
CQO	-6,98	-8,98	-10,49	7,94	-4,07	-9,34	5,18	6,49	-1,96
OD	2,26	0,92	7,94	-5,42	-0,63	8,42	-2,68	-4,95	-1,15
Oxidabilidade	0,57	-2,39	-6,11	6,57	-4,72	-5,57	2,06	1,66	-1,28
Temperatura	-2,08	0,95	-5,83	5,34	4,72	-2,58	0,83	4,28	-1,06
Bias	-0,46	1,53	1,12	-1,92	0,27	-0,79	-0,18	-1,65	0,48

Tabela III.3 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura III.2.

	$h_{1,2}$	$h_{2,2}$	$h_{3,2}$	$h_{4,2}$	$h_{5,2}$	$h_{6,2}$	$h_{7,2}$	$h_{8,2}$
$h_{1,1}$	-1,15	-1,94	1,58	-2,18	-1,38	-0,32	-2,82	-0,07
$h_{2,1}$	0,27	0,79	0,36	0,06	0,12	0,89	0,41	0,55
$h_{3,1}$	-1,34	-1,73	1,64	-1,99	-1,55	-2,39	-2,72	-1,22
$h_{4,1}$	0,70	-5,87	4,92	-5,21	-4,71	2,34	-6,60	1,27
$h_{5,1}$	-7,02	1,66	4,18	-1,23	-1,26	-3,22	-2,01	-1,48
$h_{6,1}$	-2,15	-3,51	3,68	-3,73	-3,53	3,87	-3,55	1,85
$h_{7,1}$	2,35	2,96	0,63	1,59	0,94	1,64	4,41	0,48
$h_{8,1}$	-3,26	-1,31	2,49	-1,09	-1,09	-4,21	-0,89	-2,70
$h_{9,1}$	-1,39	-0,74	1,26	-1,78	-1,89	-2,21	-1,42	-1,57
$h_{10,1}$	-1,64	-1,06	2,09	-2,36	-2,02	-1,92	-1,88	-1,53
$h_{11,1}$	-0,08	-0,03	1,44	-0,25	-0,18	1,12	0,19	0,57
$h_{12,1}$	3,81	0,09	-2,23	-3,89	-2,74	2,02	0,24	1,06
$h_{13,1}$	-0,73	0,99	1,05	0,51	0,73	-0,03	1,09	-0,45
$h_{14,1}$	-0,71	1,76	1,18	-0,66	-0,24	-2,45	2,08	-1,54
$h_{15,1}$	-1,65	-1,97	1,93	-1,94	-2,03	0,34	-3,14	0,21
$h_{16,1}$	2,29	2,09	-1,71	0,08	0,04	-1,30	2,65	-1,18
$h_{17,1}$	0,31	0,75	0,26	-1,41	-0,79	-2,05	0,70	-0,87
$h_{18,1}$	1,32	1,38	0,26	-1,18	-0,62	-1,50	1,34	-0,83
$h_{19,1}$	1,05	2,95	-0,22	0,84	0,14	0,12	3,24	0,28
$h_{20,1}$	-1,54	3,54	1,14	1,72	1,62	-1,61	4,74	-0,32
$h_{21,1}$	0,42	2,63	-1,68	5,72	4,18	-1,90	3,89	-1,66
$h_{22,1}$	1,09	-0,38	2,36	-6,22	-5,02	1,16	-0,67	1,21
$h_{23,1}$	-2,34	3,09	5,79	1,44	1,26	-3,07	3,62	-1,66
$h_{24,1}$	0,89	4,63	-2,46	2,79	2,82	-3,35	5,30	-1,76
$h_{25,1}$	-0,62	-2,16	1,33	-2,31	-2,01	0,69	-2,71	0,78
$h_{26,1}$	-2,09	-3,09	2,91	-2,99	-2,04	-0,54	-3,14	-0,05
$h_{27,1}$	-0,74	0,98	0,51	1,81	0,75	0,93	1,58	0,68
Bias	0,64	2,36	2,93	-0,21	-0,46	1,04	3,03	0,52

Tabela III.4 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura III.2.

	Classe B	Classe C
$h_{1,2}$	-7,56	7,76
$h_{2,2}$	3,16	-3,50
$h_{3,2}$	3,66	-3,79
$h_{4,2}$	5,44	-5,02
$h_{5,2}$	4,21	-4,18
$h_{6,2}$	4,78	-5,03
$h_{7,2}$	4,60	-5,17
$h_{8,2}$	1,91	-2,75
Bias	-7,44	8,58

III.2 – Modelos excluindo o parâmetro CQO

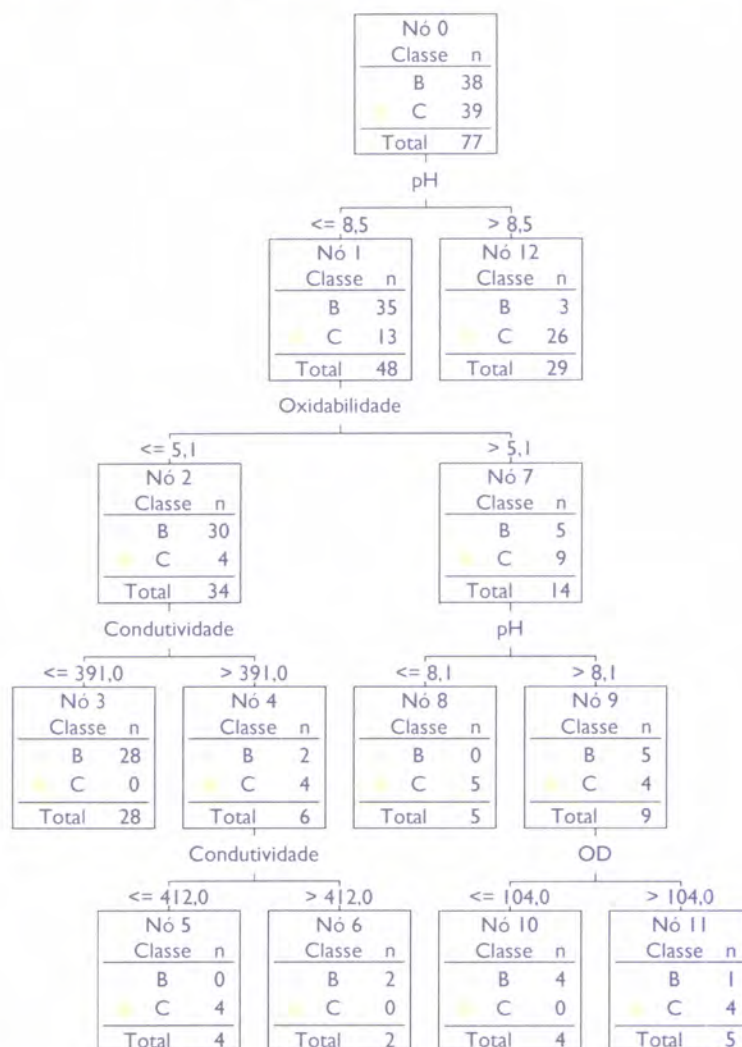


Figura III.3 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro CQO.

Tabela III.5 – Conjunto de regras relativo ao modelo apresentado na Figura III.3.

Regras para a Classe B	Regras para a Classe C	
<p>Regra 1</p> <p>Se pH ≤ 8,5 e Oxidabilidade ≤ 5,1 e Condutividade ≤ 391,0</p> <p>Então → B (28; 1,0)</p>	<p>Regra 1</p> <p>Se pH ≤ 8,5 e Oxidabilidade ≤ 5,1 e Condutividade > 391,0 e Condutividade ≤ 412,0</p> <p>Então → C (4; 1,0)</p>	<p>Regra 3</p> <p>Se pH ≤ 8,5 e Oxidabilidade > 5,1 e pH > 8,1 e OD > 104,0</p> <p>Então → C (5; 0,80)</p>
<p>Regra 2</p> <p>Se pH ≤ 8,5 e Oxidabilidade ≤ 5,1 e Condutividade > 391,0 e Condutividade > 412,0</p> <p>Então → B (2; 1,0)</p>	<p>Regra 2</p> <p>Se pH ≤ 8,5 e Oxidabilidade > 5,1 e pH ≤ 8,1</p> <p>Então → C (5; 1,0)</p>	<p>Regra 4</p> <p>Se pH > 8,5</p> <p>Então → C (29; 0,90)</p>
<p>Regra 3</p> <p>Se pH ≤ 8,5 e Oxidabilidade > 5,1 e pH > 8,1 e OD ≤ 104,0</p> <p>Então → B (4; 1,0)</p>		

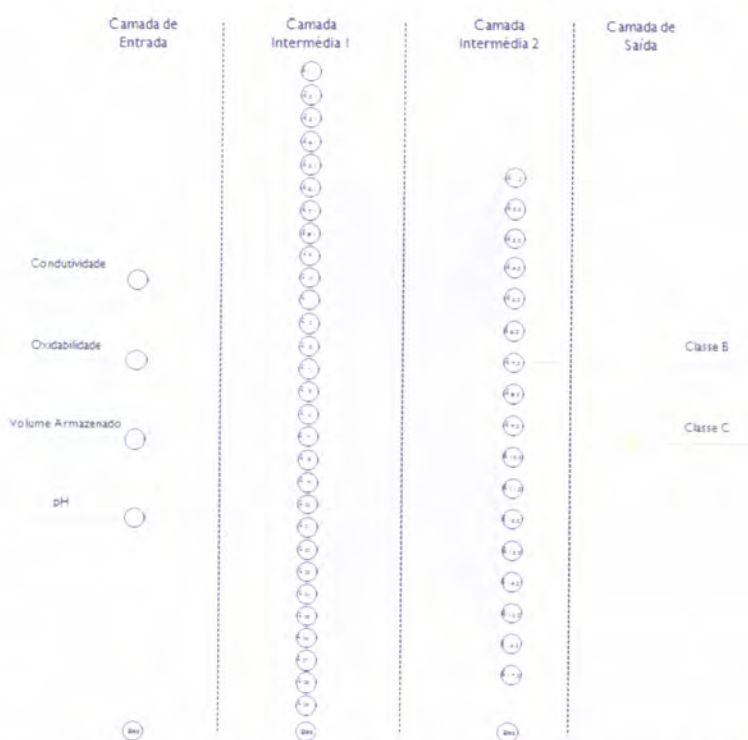


Figura III.4 – RNA obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro CQO.

Tabela III.6 – Pesos associados às conexões entre a camada de entrada e a camada intermédia I para o modelo apresentado na Figura III.4.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$	$h_{5,1}$	$h_{6,1}$	$h_{7,1}$	$h_{8,1}$	$h_{9,1}$	$h_{10,1}$
Condutividade	-1,65	1,37	-0,22	0,68	-0,82	2,50	-4,92	-0,32	-0,77	-1,70
Oxidabilidade	3,08	-5,62	0,05	-0,68	-6,71	1,98	-0,59	-1,03	3,52	-0,51
Volume Arm.	-2,59	3,93	-0,56	0,51	4,13	1,96	-1,78	0,27	-0,68	-0,95
pH	9,87	-13,05	-1,93	-4,31	-7,78	-5,58	6,51	-2,73	0,40	1,72
Bias	-4,11	5,84	-0,29	0,67	3,50	-2,07	-2,79	0,28	-2,30	-1,35

	$h_{11,1}$	$h_{12,1}$	$h_{13,1}$	$h_{14,1}$	$h_{15,1}$	$h_{16,1}$	$h_{17,1}$	$h_{18,1}$	$h_{19,1}$	$h_{20,1}$
Condutividade	0,60	3,49	-1,44	-0,89	1,04	9,22	0,21	-1,25	1,50	-0,35
Oxidabilidade	0,97	9,52	-5,89	0,11	0,12	11,82	-1,39	-12,05	1,57	-10,38
Volume Arm.	-0,25	-0,11	2,19	-0,68	2,23	-12,14	0,98	0,48	-0,76	-2,40
pH	-0,11	-2,69	-5,96	-0,71	-6,86	-2,36	-4,99	-6,23	-0,44	-4,65
Bias	-2,50	-5,31	3,72	-0,64	0,62	-2,46	1,39	6,88	-2,58	3,94

	$h_{21,1}$	$h_{22,1}$	$h_{23,1}$	$h_{24,1}$	$h_{25,1}$	$h_{26,1}$	$h_{27,1}$	$h_{28,1}$	$h_{29,1}$
Condutividade	-0,69	-1,07	-0,55	-0,33	-0,16	-0,55	-13,08	1,31	-2,07
Oxidabilidade	-1,96	-2,20	-1,49	-3,26	-1,61	-2,83	-3,64	-5,93	-10,55
Volume Arm.	0,02	-4,05	0,03	0,39	0,98	1,21	7,54	2,23	7,45
pH	-2,38	7,94	-2,02	-3,16	-3,02	-4,37	13,40	-9,77	-9,71
Bias	0,46	-1,49	-0,08	0,89	0,07	1,78	-5,12	4,99	3,01

Tabela III.7 – Pesos associados às conexões entre a camada intermédia I e a camada intermédia 2 para o modelo apresentado na Figura III.4.

	$h_{1,2}$	$h_{2,2}$	$h_{3,2}$	$h_{4,2}$	$h_{5,2}$	$h_{6,2}$	$h_{7,2}$	$h_{8,2}$	$h_{9,2}$
$h_{1,1}$	3,10	2,71	-1,73	0,17	-0,59	-0,94	-0,30	0,06	3,38
$h_{2,1}$	-5,78	-4,44	1,28	-1,36	-0,46	-0,31	-1,27	-2,13	-4,57
$h_{3,1}$	-1,42	0,36	0,15	-0,41	-0,48	-0,57	-0,11	-0,03	-0,96
$h_{4,1}$	-2,37	-0,99	0,01	-0,10	0,00	-0,41	-0,49	-0,16	-1,79
$h_{5,1}$	-3,50	-2,08	2,29	-0,70	-0,17	-0,41	-0,82	-2,50	-2,32
$h_{6,1}$	-0,38	-1,14	-0,85	0,08	-1,11	-0,15	-0,69	1,91	-4,16
$h_{7,1}$	0,09	-0,85	-1,04	-1,01	0,74	0,19	0,00	1,16	2,84
$h_{8,1}$	-1,73	-0,51	-0,02	-0,04	-0,33	0,08	0,02	-0,78	-1,27
$h_{9,1}$	0,54	0,56	-0,89	0,32	-0,41	-0,42	0,14	1,06	-0,23
$h_{10,1}$	0,27	0,07	-0,31	0,09	0,28	-0,68	-0,45	-0,24	0,98
$h_{11,1}$	1,11	0,90	-0,79	-0,54	-1,01	-0,68	-0,28	0,37	-0,82
$h_{12,1}$	0,41	3,79	-2,89	1,32	-0,71	-0,34	-0,27	5,52	-3,77
$h_{13,1}$	-2,58	-1,95	1,86	-0,70	-0,83	-0,13	-0,48	-2,59	-1,28
$h_{14,1}$	-0,79	0,66	-0,03	-0,64	0,06	-0,42	-0,22	-0,41	-0,28

	$h_{1,2}$	$h_{2,3}$	$h_{3,3}$	$h_{4,2}$	$h_{5,2}$	$h_{6,2}$	$h_{7,2}$	$h_{8,2}$	$h_{9,2}$
$h_{15,1}$	-1,21	-1,93	-0,17	-0,88	-0,45	-0,87	-0,81	0,53	-3,93
$h_{16,1}$	1,39	-0,12	-3,52	-0,62	0,22	-0,04	-0,14	9,86	4,43
$h_{17,1}$	-1,65	-1,32	-0,18	-0,62	-0,70	0,15	-0,17	-0,82	-2,21
$h_{18,1}$	-5,24	-2,51	2,70	-1,02	-0,41	-0,14	-0,94	-5,34	-0,45
$h_{19,1}$	0,50	1,25	-0,53	0,51	-0,01	-0,23	0,17	1,53	-1,11
$h_{20,1}$	-5,55	-2,17	2,18	-0,31	-0,26	0,16	-0,08	-3,66	0,74
$h_{21,1}$	-1,24	-0,46	0,71	-0,77	-0,48	-0,36	-0,37	-1,37	-0,49
$h_{22,1}$	1,02	0,98	-0,49	-0,19	0,49	-0,57	0,79	-0,32	3,75
$h_{23,1}$	-1,07	-0,67	0,56	-0,12	-0,03	-0,04	-0,09	-1,07	-0,73
$h_{24,1}$	-1,35	-1,01	1,00	-0,36	-0,36	-0,76	-0,14	-1,69	-1,01
$h_{25,1}$	-0,70	-0,73	-0,16	-0,51	-0,28	-0,79	-0,70	-1,14	-1,24
$h_{26,1}$	-2,50	-1,68	0,67	-0,15	-0,59	0,21	-0,72	-1,38	-1,50
$h_{27,1}$	1,20	-1,53	-1,90	-1,60	0,47	-0,09	0,19	5,94	4,34
$h_{28,1}$	-5,87	-2,95	1,43	-1,24	-0,59	0,04	-0,41	-2,97	-2,08
$h_{29,1}$	-5,99	-2,96	3,83	-0,61	-0,35	-0,33	-0,49	-3,52	-2,57
Bías	-0,05	1,76	-2,09	0,23	-0,70	-0,74	-1,22	-3,05	-0,49

	$h_{10,2}$	$h_{11,2}$	$h_{12,2}$	$h_{13,2}$	$h_{14,1}$	$h_{15,2}$	$h_{16,2}$	$h_{17,2}$
$h_{1,1}$	-0,24	-0,19	-0,39	0,69	0,63	2,84	-0,39	-0,64
$h_{2,1}$	-0,26	-0,11	-1,01	-2,71	-1,39	-5,50	-0,74	-0,46
$h_{3,1}$	-0,10	0,27	-0,24	0,35	-0,57	-0,32	-0,16	-0,45
$h_{4,1}$	-0,59	-0,19	-0,18	-0,39	-1,15	-1,75	0,02	-0,78
$h_{5,1}$	-0,64	-0,60	-0,77	-1,62	-0,36	-2,03	-0,58	-0,42
$h_{6,1}$	-0,95	-0,50	-1,01	-0,92	-1,86	-1,05	-0,89	-0,41
$h_{7,1}$	0,52	-0,74	0,63	-0,25	0,94	-1,14	0,38	-0,24
$h_{8,1}$	-0,27	-0,21	0,09	-0,83	-0,26	-1,13	-0,50	0,04
$h_{9,1}$	-0,85	-0,59	-0,26	0,67	-0,57	-0,31	-0,09	-0,64
$h_{10,1}$	0,00	0,15	-0,65	0,55	0,41	-0,29	-0,05	0,01
$h_{11,1}$	-0,69	-0,54	-0,65	-0,19	-0,22	0,75	-0,72	-0,66
$h_{12,1}$	-0,65	-0,61	-0,21	2,32	-1,60	0,85	-0,82	-0,84
$h_{13,1}$	-0,43	-0,47	0,08	-0,85	-1,12	-3,05	-0,37	-0,44
$h_{14,1}$	0,07	-0,69	0,66	0,35	-0,46	-0,34	-0,23	-0,33
$h_{15,1}$	-0,78	-0,21	-0,70	-1,44	-1,23	-2,63	-0,54	-0,79
$h_{16,1}$	0,27	-0,09	-0,03	1,89	1,27	-3,91	0,21	0,29
$h_{17,1}$	-0,04	0,05	-0,45	-1,31	-1,01	-1,92	-0,47	-0,16
$h_{18,1}$	-0,84	-0,07	-0,69	-2,39	-0,46	-2,82	-0,13	-0,37
$h_{19,1}$	-0,71	-0,43	0,18	-0,08	-0,64	1,38	-0,67	-0,31
$h_{20,1}$	0,22	-0,13	-0,80	-1,17	-0,11	-3,12	-0,72	-0,19
$h_{21,1}$	-0,13	-0,46	-0,32	0,01	-0,43	-0,69	-0,56	-0,31
$h_{22,1}$	-0,34	-0,15	0,44	0,84	2,52	1,15	-0,04	0,86
$h_{23,1}$	0,27	-0,06	-0,48	-0,51	-0,47	-0,44	-0,38	-0,13
$h_{24,1}$	-0,05	-0,63	-0,53	-0,26	-0,09	-0,99	0,09	-0,44

	$h_{10,2}$	$h_{11,2}$	$h_{12,2}$	$h_{13,2}$	$h_{14,2}$	$h_{15,2}$	$h_{16,2}$	$h_{17,2}$
$h_{25,1}$	-0,75	-0,75	-0,19	-0,37	-1,31	-0,68	-0,57	-0,69
$h_{26,1}$	-0,28	-0,08	-0,78	-1,05	-0,53	-1,23	-0,15	-0,18
$h_{27,1}$	-0,05	-0,95	0,11	-0,41	1,69	-4,28	-0,19	1,01
$h_{28,1}$	0,21	-0,45	-0,94	-2,81	-1,27	-3,05	-0,24	-0,68
$h_{29,1}$	0,13	-0,03	-0,73	-1,25	-0,78	-3,57	-0,58	-0,49
Bias	-0,77	-0,66	-0,76	0,29	-0,26	-0,68	-0,54	-0,72

Tabela III.8 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura III.4.

	Classe B	Classe C
$h_{1,2}$	-5,03	5,08
$h_{2,2}$	4,83	-4,85
$h_{3,2}$	2,85	-2,71
$h_{4,2}$	1,42	-1,44
$h_{5,2}$	-0,83	0,51
$h_{6,2}$	-0,24	-0,06
$h_{7,2}$	0,05	0,29
$h_{8,2}$	-5,50	5,66
$h_{9,2}$	-5,65	5,82
$h_{10,2}$	-0,16	0,59
$h_{11,2}$	0,02	-0,52
$h_{12,2}$	0,19	0,57
$h_{13,2}$	2,44	-2,58
$h_{14,2}$	-2,18	1,72
$h_{15,2}$	5,64	-5,43
$h_{16,2}$	0,04	0,57
$h_{17,2}$	-0,59	0,72
Bias	1,57	-1,64

III.3 – Modelos excluindo o parâmetro Oxidabilidade

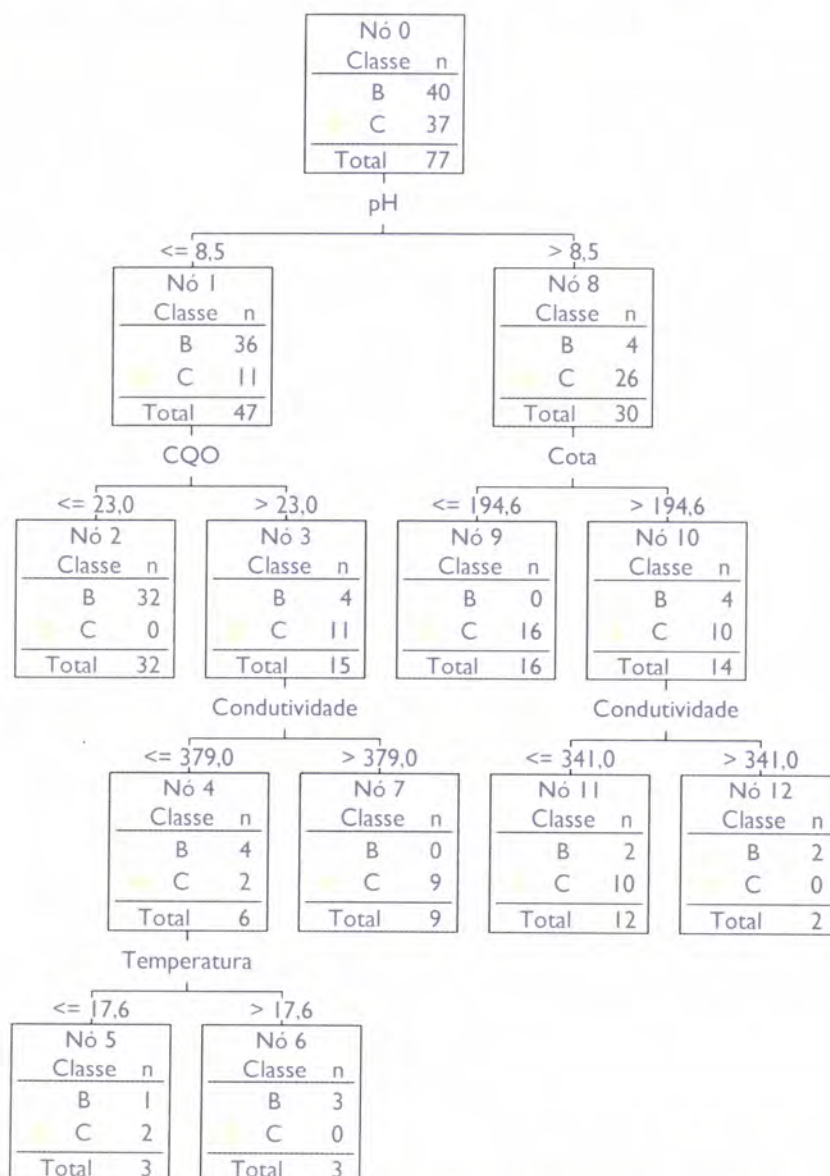


Figura III.5 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro oxidabilidade.

Tabela III.9 – Conjunto de regras relativo ao modelo apresentado na Figura III.5.

Regras para a Classe B	Regras para a Classe C	
<p>Regra 1</p> <p>Se pH ≤ 8,5 e CQO ≤ 23,0</p> <p>Então → B (32; 1,0)</p>	<p>Regra 1</p> <p>Se pH ≤ 8,5 e CQO > 23,0 e Condutividade ≤ 379,0 e Temperatura ≤ 17,6</p> <p>Então → C (3; 0,67)</p>	<p>Regra 4</p> <p>Se pH > 8,5 e Cota > 194,6 e Condutividade ≤ 341,0</p> <p>Então → C (12; 0,83)</p>

Regras para a Classe B	Regras para a Classe C	
Regra 2 Se pH \leq 8,5 e CQO $>$ 23,0 e Condutividade \leq 379,0 e Temperatura $>$ 17,6 Então \rightarrow B (3; 1,0)	Regra 2 Se pH \leq 8,5 e CQO $>$ 23,0 e Condutividade $>$ 379,0 Então \rightarrow C (9; 1,0)	
Regra 3 Se pH $>$ 8,5 e Cota $>$ 194,6 e Condutividade $>$ 341,0 Então \rightarrow B (2; 1,0)	Regra 3 Se pH $>$ 8,5 e Cota \leq 194,6 Então \rightarrow C (16; 1,0)	

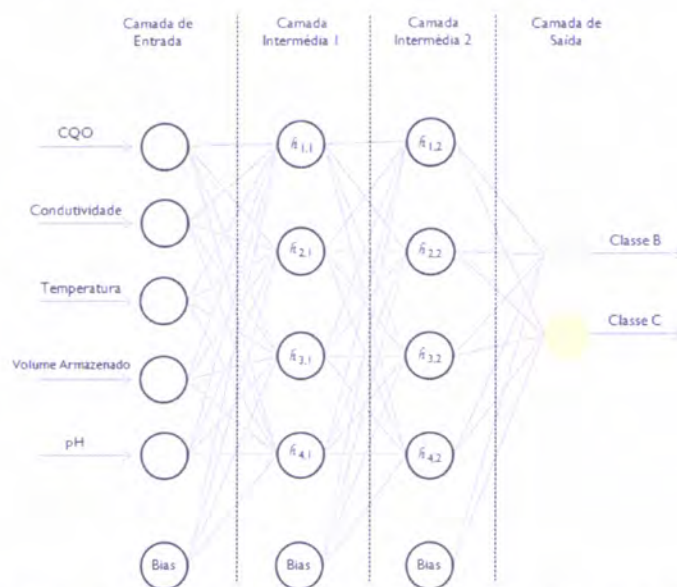


Figura III.6 – RNA obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro oxidabilidade.

Tabela III.10 – Pesos associados às conexões entre a camada de entrada e a camada intermédia I referente ao modelo apresentado na Figura III.6.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$
CQO	-19,79	-14,85	6,71	6,78
Condutividade	6,28	9,37	5,69	0,29
Temperatura	-4,70	2,21	-5,36	2,91
Volume Arm.	1,01	0,97	-10,79	-7,27
pH	3,31	-9,29	-1,68	19,09
Bias	-0,12	4,85	5,17	-9,26

Tabela III.11 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 referente ao modelo apresentado na Figura III.6.

	$h_{1,2}$	$h_{2,2}$	$h_{3,2}$	$h_{4,2}$
$h_{1,1}$	4,12	-9,99	-1,95	0,00
$h_{2,1}$	-1,08	-3,72	-5,51	1,48
$h_{3,1}$	-9,24	5,19	8,01	5,89
$h_{4,1}$	5,80	1,29	7,28	2,19
Bias	-3,84	2,91	2,51	4,17

Tabela III.12 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída referente ao modelo apresentado na Figura III.6.

	Classe B	Classe C
$h_{1,2}$	-7,01	7,04
$h_{2,2}$	-7,95	7,91
$h_{3,2}$	-4,99	4,88
$h_{4,2}$	-3,12	3,28
Bias	12,26	-12,36

III.4 – Modelos excluindo os parâmetros pH e CQO

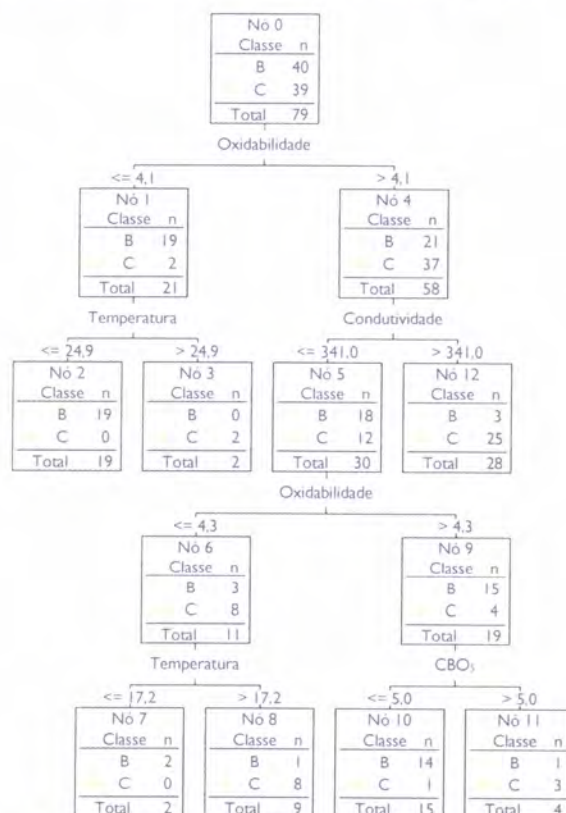


Figura III.7 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito, excluindo os parâmetros pH e CQO.

Tabela III.13 – Conjunto de regras relativo ao modelo apresentado na Figura III.7.

Regras para a Classe B	Regras para a Classe C	
Regra 1 Se Oxidabilidade $\leq 4,1$ e Temperatura $\leq 24,9$ Então $\rightarrow B$ (19; 1,0)	Regra 1 Se Oxidabilidade $\leq 4,1$ e Temperatura $> 24,9$ Então $\rightarrow C$ (2; 1,0)	Regra 4 Se Oxidabilidade $> 4,1$ e Condutividade $> 341,0$ Então $\rightarrow C$ (28; 0,89)
Regra 2 Se Oxidabilidade $> 4,1$ T Condutividade $\leq 341,0$ e Oxidabilidade $\leq 4,3$ e Temperatura $\leq 17,2$ Então $\rightarrow B$ (2; 1,0)	Regra 2 Se Oxidabilidade $> 4,1$ e Condutividade $\leq 341,0$ e Oxidabilidade $\leq 4,3$ e Temperatura $> 17,2$ Então $\rightarrow C$ (9; 0,89)	
Regra 3 Se Oxidabilidade $> 4,1$ e Condutividade $\leq 341,0$ e Oxidabilidade $> 4,3$ e $CBO_5 \leq 5,0$ Então $\rightarrow B$ (15; 0,93)	Regra 3 Se Oxidabilidade $> 4,1$ e Condutividade $\leq 341,0$ e Oxidabilidade $> 4,3$ e $CBO_5 > 5,0$ Então $\rightarrow C$ (4; 0,75)	

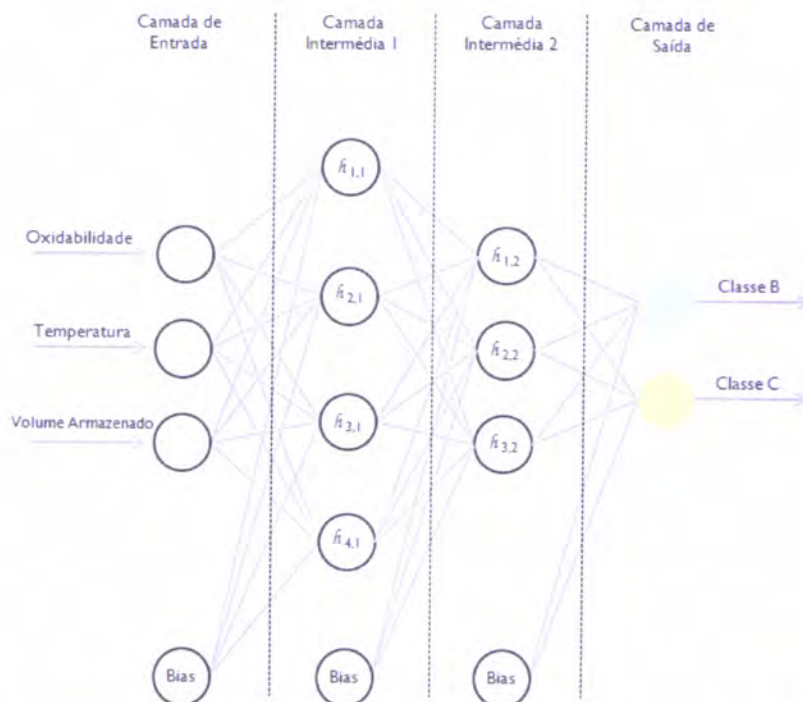


Figura III.8 – RNA obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro pH e CQO.

Tabela III.14 – Pesos associados às conexões entre a camada de entrada e a camada intermédia I referente ao modelo apresentado na Figura III.8.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$
Oxidabilidade	-23,57	-14,07	20,19	9,09
Temperatura	2,87	-11,04	5,95	1,96
Volume Arm.	9,69	17,66	-1,10	-19,69
Bias	-3,16	-5,38	-10,16	8,81

Tabela III.15 – Pesos associados às conexões entre a camada intermédia I e a camada intermédia 2 referente ao modelo apresentado na Figura III.8.

	$h_{1,2}$	$h_{2,2}$	$h_{3,2}$
$h_{1,1}$	-9,81	8,13	7,27
$h_{2,1}$	-6,64	-0,50	10,07
$h_{3,1}$	9,30	-8,69	-5,72
$h_{4,1}$	4,21	-11,33	-3,24
Bias	-7,13	0,68	-8,29

Tabela III.16 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída referente ao modelo apresentado na Figura III.8.

	Classe B	Classe C
$h_{1,2}$	-8,32	8,31
$h_{2,2}$	-9,51	9,49
$h_{3,2}$	8,43	-8,42
Bias	4,94	-4,93

III.5 – Modelos excluindo os parâmetros pH e Oxidabilidade

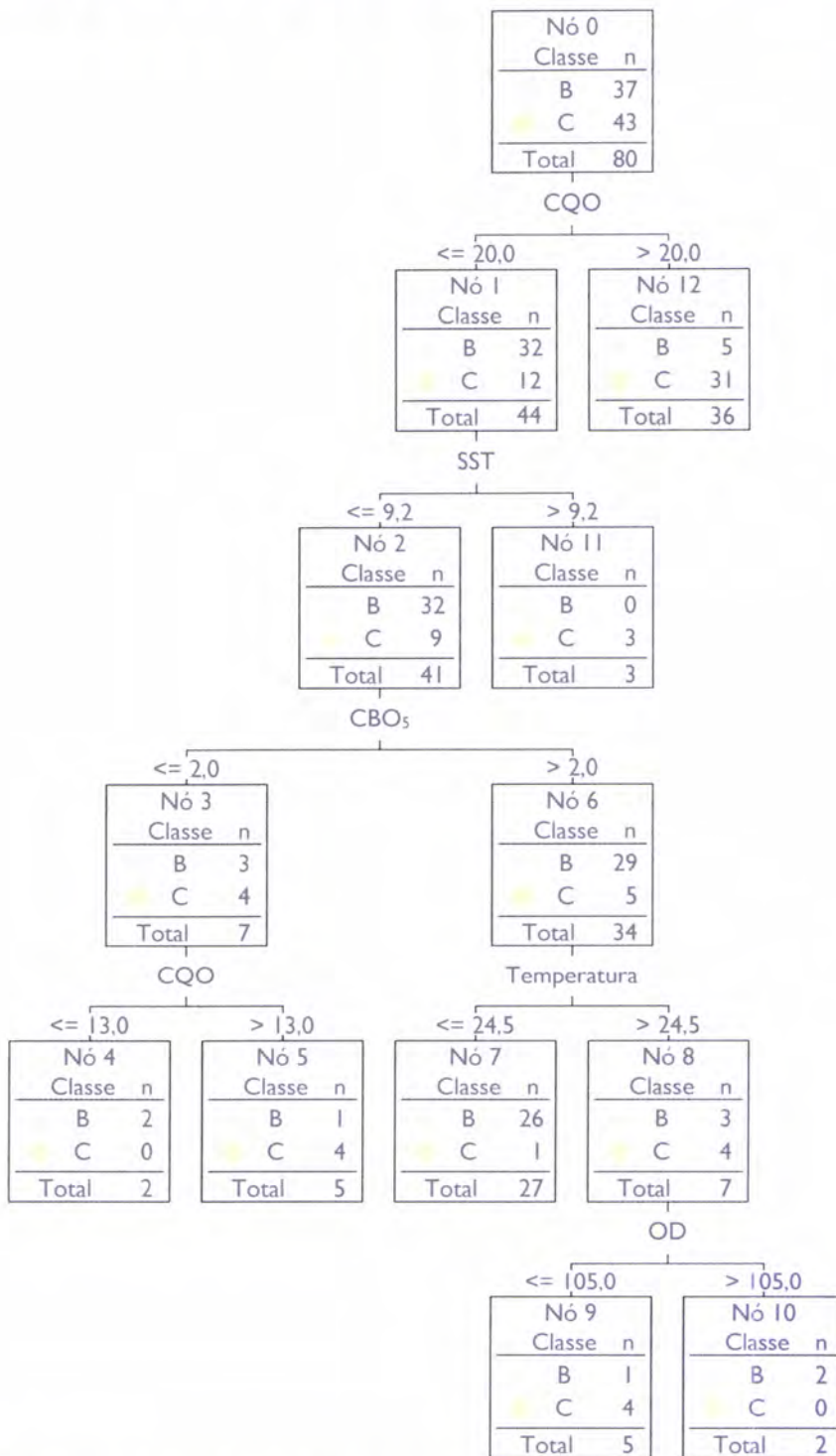


Figura III.9 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito, excluindo os parâmetros pH e oxidabilidade.

Tabela III.17 – conjunto de regras relativo ao modelo apresentado na Figura III.9.

Regras para a Classe B		Regras para a Classe C	
Regra 1 Se CQO \leq 20,0 e SST \leq 9,2 e CBO ₅ \leq 2,0 e CQO \leq 13,0 Então \rightarrow B (2; 1,0)	Regra 3 Se CQO \leq 20,0 e SST \leq 9,2 e CBO ₅ $>$ 2,0 e Temperatura $>$ 24,5 e OD $>$ 105,0 Então \rightarrow B (2; 1,0)	Regra 1 Se CQO \leq 20,0 e SST \leq 9,2 e CBO ₅ \leq 2,0 e CQO $>$ 13,0 Então \rightarrow C (5; 0,80)	Regra 3 Se CQO \leq 20,0 e SST $>$ 9,2 Então \rightarrow C (3; 1,0)
Regra 2 Se CQO \leq 20,0 e SST \leq 9,2 e CBO ₅ $>$ 2,0 e Temperatura \leq 24,5 Então \rightarrow B (27; 0,96)	Regra 2 Se CQO \leq 20,0 e SST \leq 9,2 e CBO ₅ $>$ 2,0 e Temperatura $>$ 24,5 e OD \leq 105,0 Então \rightarrow C (5; 0,80)	Regra 4 Se CQO $>$ 20,0 Então \rightarrow C (36; 0,86)	

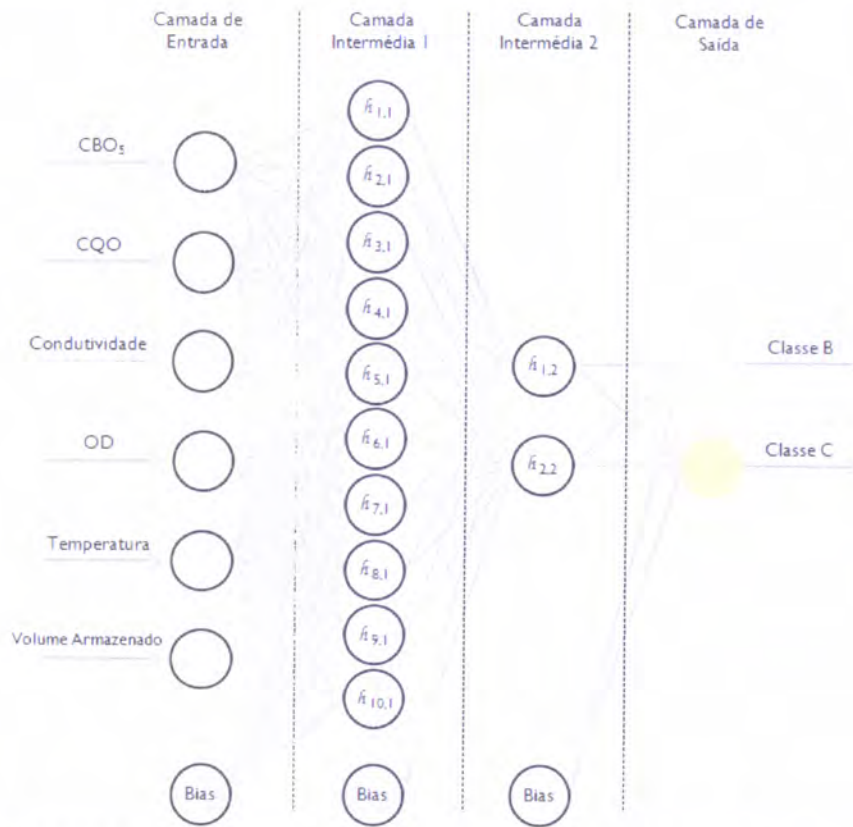


Figura III.10 – RNA obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro pH e oxidabilidade.

Tabela III.18 – Pesos associados às conexões entre a camada de entrada e a camada intermédia I referente ao modelo apresentado na Figura III.10.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$	$h_{5,1}$	$h_{6,1}$	$h_{7,1}$	$h_{8,1}$	$h_{9,1}$	$h_{10,1}$
CBO ₅	6,52	-14,79	-5,23	4,45	2,39	3,10	2,67	1,56	2002,00	2,74
CQO	9,14	6,53	16,88	17,06	7,66	5,02	-9,12	-7,16	24,77	-12,79
Condutividade	-5,83	-717,00	-2,99	-6,10	-1,07	-1,09	0,91	0,11	-5,37	0,49
OD	-5,99	-5,52	-10,36	-10,99	-5,12	-3,94	5,00	4,52	-13,51	7,43
Temperatura	-1,10	21,95	4,09	0,22	-6,21	-5,94	-3,37	-1,50	3,11	-3,34
Volume Arm.	-2,53	-10,95	-2,40	-3,20	-3,16	-2,89	0,88	-0,12	-3,59	-0,58
Bias	3,37	-2,43	3,14	4,66	3,08	2,86	-0,28	-0,41	4,25	-0,23

Tabela III.19 – Pesos associados às conexões entre a camada intermédia I e a camada intermédia 2 referente ao modelo apresentado na Figura III.10.

	$h_{1,2}$	$h_{2,2}$	$h_{3,2}$	$h_{4,2}$	$h_{5,2}$	$h_{6,2}$
$h_{1,1}$	-3,48	-0,83	-3,46	-2,00	-1,55	-1,42
$h_{2,1}$	-3,46	-2,78	-4,34	-4,15	-3,61	-3,76
$h_{3,1}$	-3,25	1,74	-4,60	-3,96	2,49	1,79
$h_{4,1}$	-3,25	-0,89	-4,78	-3,58	-1,26	-0,69
$h_{5,1}$	-3,16	-1,32	-3,72	-2,54	-1,96	-1,70
$h_{6,1}$	-2,51	-1,57	-2,51	-1,75	-2,26	-1,43
$h_{7,1}$	-0,92	-1,85	3,06	3,28	-2,42	-2,11
$h_{8,1}$	2,52	-1,41	2,44	2,29	-1,90	-1,36
$h_{9,1}$	1,79	1,61	-6,82	-6,87	1,72	1,91
$h_{10,1}$	-4,33	-2,92	3,86	3,54	-2,40	2,80
Bias	3,28	0,30	4,45	3,59	0,45	0,31

Tabela III.20 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída referente ao modelo apresentado na Figura III.10.

	Classe B	Classe C
$h_{1,2}$	2,59	-2,96
$h_{2,2}$	4,89	-3,96
$h_{3,2}$	5,05	-4,89
$h_{4,2}$	3,49	-3,08
$h_{5,2}$	5,08	-5,75
$h_{6,2}$	5,04	-4,73
Bias	-8,91	8,69

Regras para a Classe B	Regras para a Classe C
<p>Regra 2</p> <p>Se $\text{pH} \leq 8,5$ e $\text{Cota} \leq 193,9$ e $\text{Condutividade} > 412,0$</p> <p>Então $\rightarrow B$ (2; 1,0)</p>	<p>Regra 2</p> <p>Se $\text{pH} \leq 8,5$ e $\text{Cota} \leq 193,9$ e $\text{Condutividade} \leq 412,0$ e $\text{Temperatura} > 17,2$ e $\text{Condutividade} > 388,0$</p> <p>Então $\rightarrow C$ (4; 1,0)</p>
<p>Regra 3</p> <p>Se $\text{pH} \leq 8,5$ e $\text{Cota} > 193,9$ e $\text{CBO}_5 \leq 5,0$</p> <p>Então $\rightarrow B$ (30; 0,97)</p>	<p>Regra 3</p> <p>Se $\text{pH} \leq 8,5$ e $\text{Cota} > 193,9$ e $\text{CBO}_5 > 5,0$</p> <p>Então $\rightarrow C$ (3; 1,0)</p>
	<p>Regra 4</p> <p>Se $\text{pH} > 8,5$</p> <p>Então $\rightarrow C$ (27; 0,85)</p>

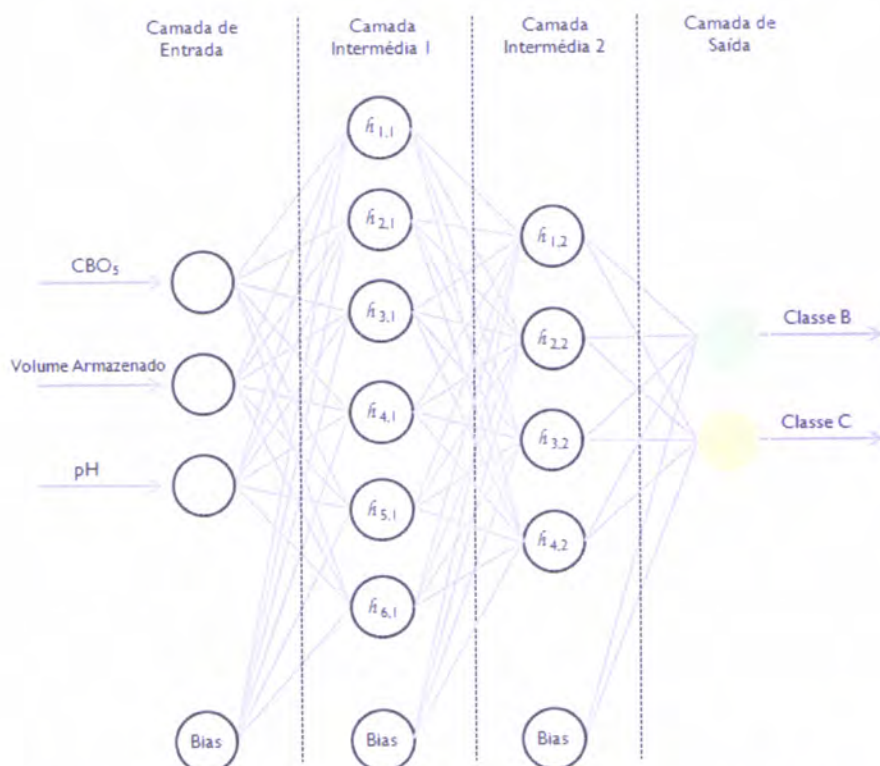


Figura III.12 – RNA obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro CQO e oxidabilidade.

Tabela III.22 – Pesos associados às conexões entre a camada de entrada e a camada intermédia I referente ao modelo apresentado na Figura III.12.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$	$h_{5,1}$	$h_{6,1}$
CBO_5	-5,06	0,26	19,09	-8,73	5,62	2,99
Volume Arm.	22,01	-17,97	-19,10	-5,53	-18,69	-9,29
pH	-17,35	16,65	-10,02	35,34	3,37	22,71
Bias	-4,45	2,07	9,43	-16,80	4,97	-4,40

Tabela III.23 – Pesos associados às conexões entre a camada intermédia I e a camada intermédia 2 referente ao modelo apresentado na Figura III.12.

	$h_{1,2}$	$h_{2,2}$	$h_{3,2}$	$h_{4,2}$
$h_{1,1}$	-2,47	-11,35	5,74	14,08
$h_{2,1}$	-0,29	8,60	-5,30	-4,81
$h_{3,1}$	-16,78	2,90	-1,40	-5,26
$h_{4,1}$	-15,60	1,23	-8,78	-5,20
$h_{5,1}$	7,94	7,07	-2,51	-5,72
$h_{6,1}$	8,83	3,24	-7,49	-3,77
Bias	5,45	-0,14	4,13	5,77

Tabela III.24 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída referente ao modelo apresentado na Figura III.12.

	Classe B	Classe C
$h_{1,2}$	8,05	-7,82
$h_{2,2}$	7,94	-7,61
$h_{3,2}$	9,09	-8,72
$h_{4,2}$	4,29	-4,10
Bias	-14,06	13,51

III.7 – Modelos excluindo os parâmetros pH, CQO e oxidabilidade

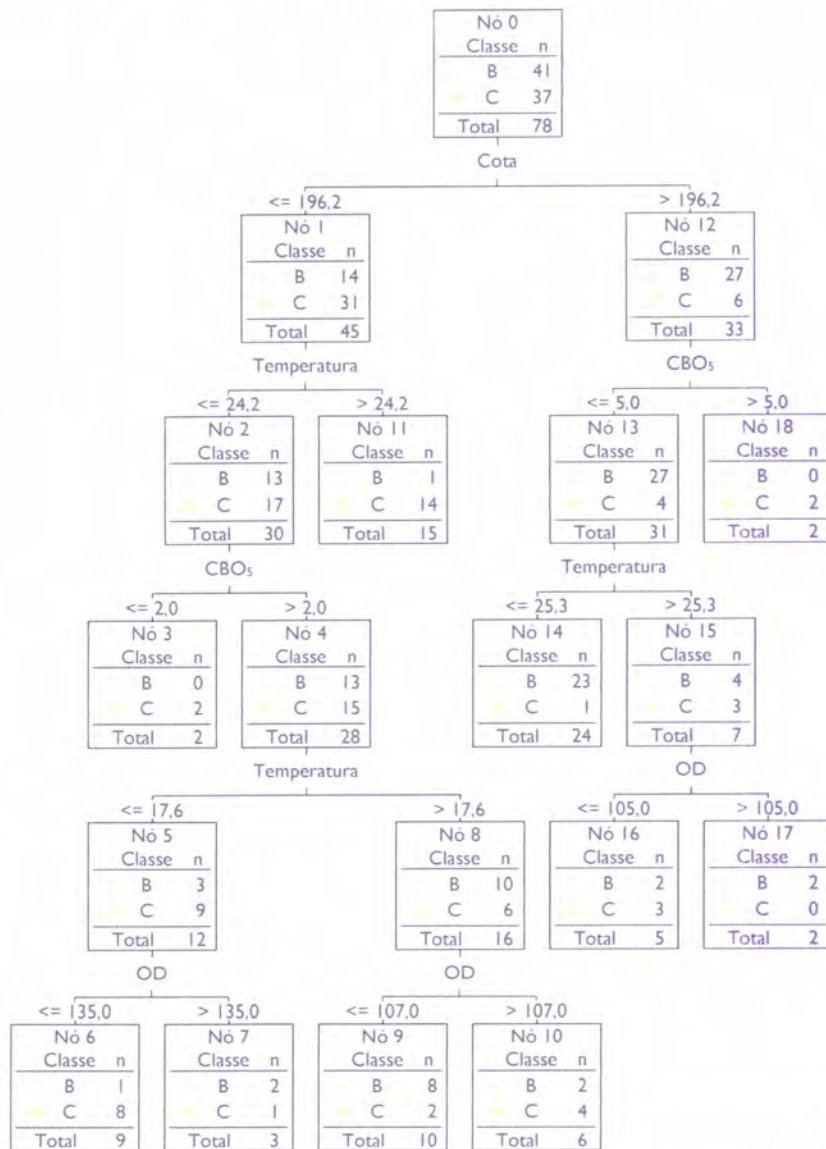


Figura III.13 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito, excluindo os parâmetros pH, CQO e oxidabilidade.

Tabela III.25 – Conjunto de regras relativo ao modelo apresentado na Figura III.13.

Regras para a Classe B	Regras para a Classe C	
<p>Regra 1</p> <p>Se Cota ≤ 196,2 e Temperatura ≤ 24,2 e CBO₅ > 2,0 e Temperatura ≤ 17,6 e OD > 135,0</p> <p>Então → B (3; 0,67)</p>	<p>Regra 1</p> <p>Se Cota ≤ 196,2 e Temperatura ≤ 24,2 e CBO₅ ≤ 2,0</p> <p>Então → C (2; 1,0)</p>	<p>Regra 5</p> <p>Se Cota > 196,2 e CBO₅ ≤ 5,0 e Temperatura > 25,3 e OD ≤ 105,0</p> <p>Então → C (5; 0,60)</p>

Regras para a Classe B	Regras para a Classe C	
<p>Regra 2 Se Cota \leq 196,2 e Temperatura \leq 24,2 e CBO₅ > 2,0 e Temperatura > 17,6 e OD \leq 107,0 Então → B (10; 0,80)</p>	<p>Regra 2 Se Cota \leq 196,2 e Temperatura \leq 24,2 e CBO₅ > 2,0 e Temperatura \leq 17,6 e OD \leq 135,0 Então → C (9; 0,89)</p>	<p>Regra 6 Se Cota > 196,2 e CBO₅ > 5,0 Então → C (2; 1,0)</p>
<p>Regra 3 Se Cota > 196,2 e CBO₅ \leq 5,0 e Temperatura \leq 25,3 Então → B (24; 0,96)</p>	<p>Regra 3 Se Cota \leq 196,2 e Temperatura \leq 24,2 e CBO₅ > 2,0 e Temperatura > 17,6 e OD > 107,0 Então → C (6; 0,67)</p>	
<p>Regra 4 Se Cota > 196,2 e CBO₅ \leq 5,0 e Temperatura > 25,3 e OD > 105,0 Então → B (2; 1,0)</p>	<p>Regra 4 Se Cota \leq 196,2 e Temperatura > 24,2 Então → C (15; 0,93)</p>	

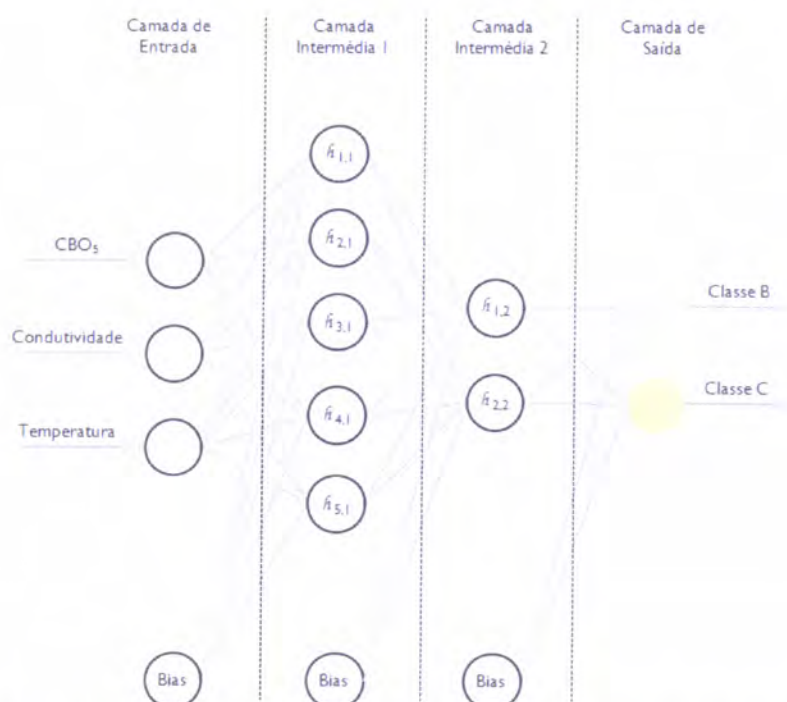


Figura III.14 – RNA obtida para previsão da qualidade da água da albufeira do Alvito, excluindo o parâmetro pH, CQO e oxidabilidade.

Tabela III.26 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 referente ao modelo apresentado na Figura III.14.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$	$h_{5,1}$
CBO_S	-16,26	-2,49	-17,53	-9,62	-6,99
Condutividade	-2,92	-12,77	-23,53	-19,71	-5,98
Temperatura	-23,65	-11,82	-4,68	-0,64	-17,13
Bias	25,17	15,57	15,57	13,61	8,41

Tabela III.27 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 referente ao modelo apresentado na Figura III.14.

	$h_{1,2}$	$h_{2,2}$
$h_{1,1}$	-9,18	-14,76
$h_{2,1}$	-3,25	7,19
$h_{3,1}$	-14,97	2,76
$h_{4,1}$	16,05	-4,94
$h_{5,1}$	13,28	-1,96
Bias	-1,08	8,31

Tabela III.28 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída referente ao modelo apresentado na Figura III.14.

	Classe B	Classe C
$h_{1,2}$	4,83	-4,75
$h_{2,2}$	-8,54	8,44
Bias	0,61	-0,60

ANEXO IV

Neste anexo apresentam-se os modelos relativos à previsão da qualidade da água da albufeira do Alvito, tendo em consideração a utilização de parâmetros meteorológicos e hidrométricos, obtidos com recurso a Árvores de Decisão e RNAs (Tabela 24). Apresentam-se, ainda, o conjunto de regras geradas destacando-se, no final, o suporte e a confiança da regra (Árvores de Decisão) e os pesos associados às conexões entre os nodos das camadas adjacentes (RNAs). Os nodos das camadas intermédias são designados genericamente por h_{ij} , denotando o nodo i da camada intermédia j .

IV.1 – Modelos gerados com os parâmetros meteorológicos e hidrométricos

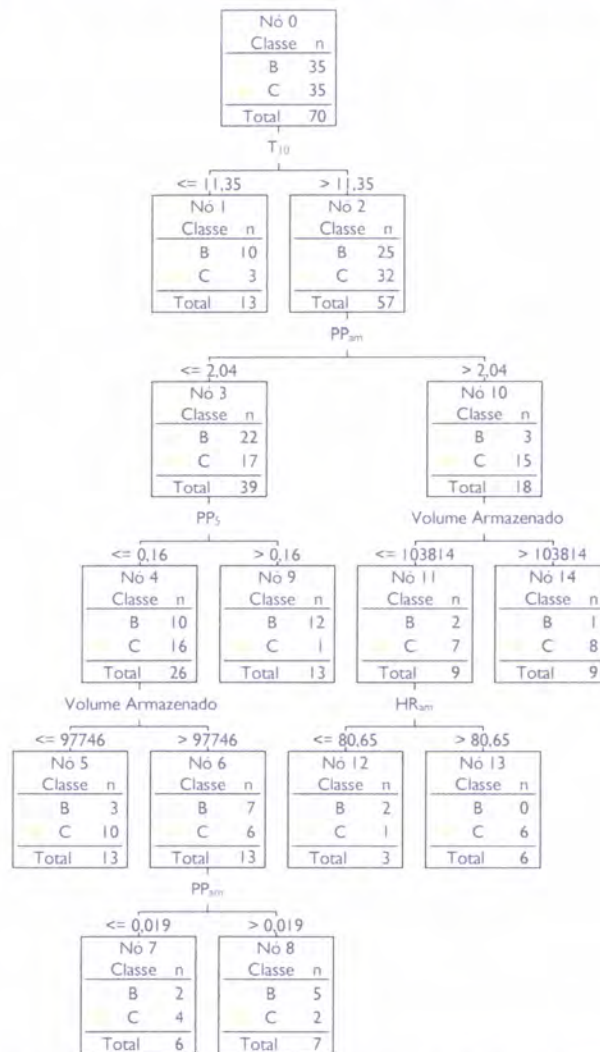


Figura IV.1 – Árvore de decisão obtida para previsão da qualidade da água da albufeira do Alvito, tendo em consideração a utilização de parâmetros meteorológicos e hidrométricos.

Tabela IV.1 – Conjunto de regras relativo ao modelo apresentado na Figura IV.1.

Regras para a Classe B	Regras para a Classe C
<p>Regra 1</p> <p>Se $T_{10} \leq 11,35$</p> <p>Então $\rightarrow B$ (13; 0,77)</p>	<p>Regra 1</p> <p>Se $T_{10} > 11,35$ e $PP_{am} \leq 2,04$ e $PP_5 \leq 0,16$ e Volume Armazenado ≤ 97746</p> <p>Então $\rightarrow C$ (13; 0,77)</p>
<p>Regra 2</p> <p>Se $T_{10} > 11,35$ e $PP_{am} \leq 2,04$ e $PP_5 \leq 0,16$ e Volume Armazenado > 97746 e $PP_{am} > 0,019$</p> <p>Então $\rightarrow B$ (7; 0,71)</p>	<p>Regra 2</p> <p>Se $T_{10} > 11,35$ e $PP_{am} \leq 2,04$ e $PP_5 \leq 0,16$ e Volume Armazenado > 97746 e $PP_{am} \leq 0,019$</p> <p>Então $\rightarrow C$ (6; 0,67)</p>
<p>Regra 3</p> <p>Se $T_{10} > 11,35$ e $PP_{am} \leq 2,04$ e $PP_5 > 0,16$</p> <p>Então $\rightarrow B$ (13; 0,92)</p>	<p>Regra 3</p> <p>Se $T_{10} > 11,35$ e $PP_{am} > 2,04$ e Volume Armazenado ≤ 103814 e $HR_{am} > 80,65$</p> <p>Então $\rightarrow C$ (6; 1,0)</p>
<p>Regra 4</p> <p>Se $T_{10} > 11,35$ e $PP_{am} > 2,04$ e Volume Armazenado ≤ 103814 e $HR_{am} \leq 80,65$</p> <p>Então $\rightarrow B$ (3; 0,67)</p>	<p>Regra 4</p> <p>Se $T_{10} > 11,35$ e $PP_{am} > 2,04$ e Volume Armazenado > 103814</p> <p>Então $\rightarrow C$ (9; 0,89)</p>

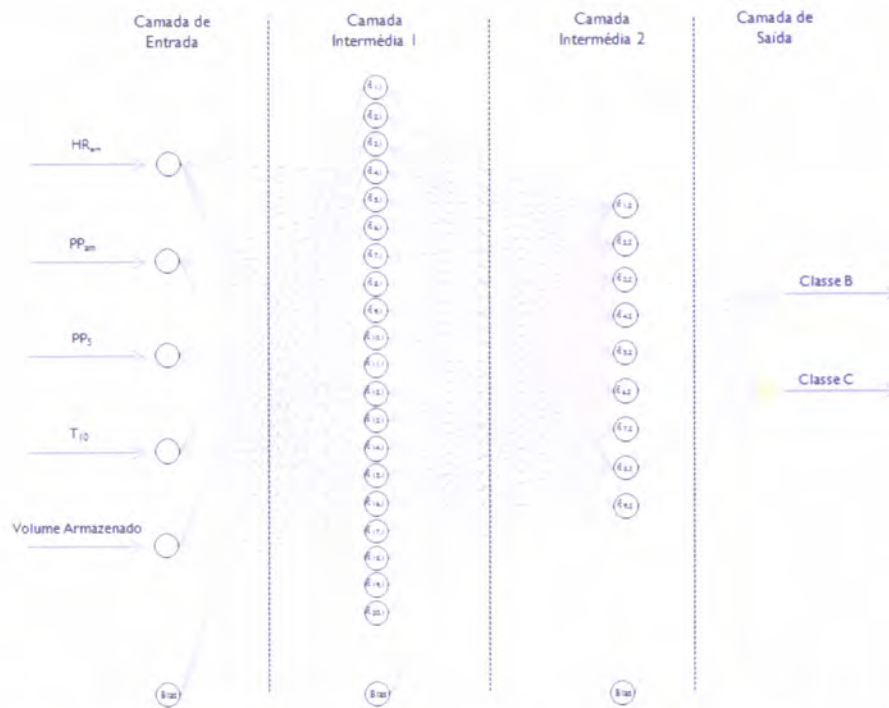


Figura IV.2 – RNA obtida para previsão da qualidade da água da albufeira do Alvito, tendo em consideração a utilização de parâmetros meteorológicos e hidrométricos.

Tabela IV.2 – Pesos associados às conexões entre a camada de entrada e a camada intermédia I para o modelo apresentado na Figura IV.2.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$	$h_{5,1}$	$h_{6,1}$	$h_{7,1}$	$h_{8,1}$	$h_{9,1}$	$h_{10,1}$
HR_{3m}	3,84	-2,76	5,66	-2,43	-1,33	1,46	-6,74	1,96	.0,95	2,6
PP_{3m}	18,17	-3,73	5,69	-5,32	-0,91	-1,67	2,32	8,85	-1,09	-6,09
PP_5	-3,98	0,09	10,53	0,56	1,79	-5,96	-5,13	11,09	1,88	-2,27
T_{10}	-4,93	4,99	-8,51	-2,29	-1,63	-3,76	4,01	1,08	-1,35	6,82
Volume Arm.	-1,55	-5,66	4,56	-5,61	-3,29	13,39	-14,52	-9,77	-3,5	-3,45
Bias	-7,19	-0,82	-5,97	4,12	1,74	-7,12	8,46	-3,42	1,58	-3,63

	$h_{11,1}$	$h_{12,1}$	$h_{13,1}$	$h_{14,1}$	$h_{15,1}$	$h_{16,1}$	$h_{17,1}$	$h_{18,1}$	$h_{19,1}$	$h_{20,1}$
HR_{3m}	3,13	-1,06	4,04	9,45	0,96	-1,59	4,67	-0,68	-4,23	-1,17
PP_{3m}	-8,38	-1,79	-8,15	-7,98	-4,41	-1,17	-3,4	3,01	-12,71	-2,15
PP_5	-8,39	1,36	-4,2	-15,56	-7,67	1,13	18,09	-3,06	-3,85	4,22
T_{10}	-17,41	-1,62	9,81	-2,38	1,82	-1,53	-7	-4,07	15,12	-1,44
Volume Arm.	23,13	-2,9	-5,41	-8,84	-7,79	-2,82	-3,12	5,87	-5,68	-4,43
Bias	-6,19	1,06	-4,53	5	2,8	1,34	-3,32	-6,51	-7,19	2,39

Tabela IV.3 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura IV.2.

	$h_{1,2}$	$h_{2,2}$	$h_{3,2}$	$h_{4,2}$	$h_{5,2}$	$h_{6,2}$	$h_{7,2}$	$h_{8,2}$	$h_{9,2}$
$h_{1,1}$	-5,22	3,82	-1,29	-2,37	-1,04	2,39	2,35	-2,92	-6,75
$h_{2,1}$	-1,32	-3,82	-0,39	-1,09	-0,74	-1,00	-1,31	-0,88	-0,84
$h_{3,1}$	5,18	8,55	-1,26	-1,30	-1,09	3,18	3,85	-1,19	5,91
$h_{4,1}$	-0,82	-3,93	1,28	2,79	0,93	-1,99	-2,39	1,17	-1,06
$h_{5,1}$	0,60	-1,80	0,45	1,75	0,67	-1,39	-1,00	0,25	1,34
$h_{6,1}$	-2,45	4,11	-2,13	-3,22	-1,47	0,99	1,14	-2,65	-2,95
$h_{7,1}$	-1,75	-5,46	2,10	3,08	2,16	-3,68	-4,08	2,44	-2,14
$h_{8,1}$	-2,16	-2,66	-4,23	-5,12	-3,16	-1,62	-1,69	-5,11	-3,82
$h_{9,1}$	0,29	-2,44	0,065	1,63	0,39	-0,68	-0,86	0,15	1,38
$h_{10,1}$	-1,39	-1,24	-1,06	-1,90	-1,26	-1,11	-0,78	-2,07	-2,12
$h_{11,1}$	2,38	8,53	-4,98	-9,24	-5,30	3,43	4,05	-8,63	2,38
$h_{12,1}$	0,67	-2,19	0,44	1,57	0,87	-0,78	-1,25	0,98	0,08
$h_{13,1}$	-0,96	-1,77	-2,57	-3,08	-1,47	-1,75	-1,53	-2,54	-1,73
$h_{14,1}$	-4,52	-5,35	-2,10	-3,22	-1,70	-2,01	-2,21	-1,19	-5,07
$h_{15,1}$	-2,00	-5,35	0,01	-0,66	-0,74	-2,23	-2,04	-0,45	-2,01
$h_{16,1}$	0,03	-2,02	0,61	1,71	0,67	-0,78	-1,19	0,27	0,70
$h_{17,1}$	-6,79	-8,43	-0,29	-1,35	-0,64	-3,48	-3,06	-0,43	-7,74
$h_{18,1}$	-2,22	2,01	0,27	0,17	0,17	1,78	1,02	0,07	-1,85
$h_{19,1}$	-0,6	-3,06	-3,01	-5,12	-2,99	-1,57	-2,09	-3,99	-0,87
$h_{20,1}$	-0,23	-3,99	1,13	1,78	0,83	-1,43	-1,52	0,34	0,42
Bias	-2,06	-0,55	1,41	2,85	0,95	0,06	-0,12	2,61	-2,14

Tabela IV.4 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura IV.2.

	Classe B	Classe C
$h_{1,2}$	-5,13	4,78
$h_{2,2}$	7,08	-6,98
$h_{3,2}$	2,44	-3,09
$h_{4,2}$	4,49	-4,58
$h_{5,2}$	2,72	-2,39
$h_{6,2}$	1,94	-1,86
$h_{7,2}$	2,34	-2,07
$h_{8,2}$	4,04	-3,57
$h_{9,2}$	-6,52	6,11
Bias	-6,07	5,97

ANEXO V

Apresentam-se o conjunto de regras geradas para o modelo apresentado na Figura 29, destacando-se, no final, o suporte e a confiança da regra (Árvores de Decisão). Apresentam-se, ainda, os pesos associados às conexões entre os nodos das camadas adjacentes (RNAs), referente ao modelo apresentado na Figura 30. Os nodos das camadas intermédias são designados genericamente por h_{ij} , denotando o nodo i da camada intermédia j .

Tabela V.1 – Conjunto de regras relativo ao modelo apresentado na Figura 29.

Regras para a Classe B	Regras para a Classe C	
<p>Regra 1 Se pH \leq 8,5 e Volume Armazenado \leq 94419 e $T_{am} >$ 15,3 e $T_{10} \leq$ 20,96 Então \rightarrow B (7; 1,0)</p>	<p>Regra 1 Se pH \leq 8,5 e Volume Armazenado \leq 94419 e $T_{am} \leq$ 15,3 Então \rightarrow C (7; 0,86)</p>	<p>Regra 5 Se pH $>$ 8,5 e $T_{am} >$ 19,33 Então \rightarrow C (10; 1,0)</p>
<p>Regra 2 Se pH \leq 8,5 e Volume Armazenado $>$ 94419 e pH \leq 8,3 Então \rightarrow B (14; 1,0)</p>	<p>Regra 2 Se pH \leq 8,5 e Volume Armazenado \leq 94419 e $T_{am} >$ 15,3 e $T_{10} >$ 20,96 Então \rightarrow C (5; 0,80)</p>	
<p>Regra 3 Se pH \leq 8,5 e Volume Armazenado $>$ 94419 e pH $>$ 8,3 e $PP_5 \leq$ 1,08 Então \rightarrow B (9; 1,0)</p>	<p>Regra 3 Se pH \leq 8,5 e Volume Armazenado $>$ 94419 e pH $>$ 8,3 e $PP_5 >$ 1,08 Então \rightarrow C (6; 0,67)</p>	
<p>Regra 4 Se pH $>$ 8,5 e $T_{am} \leq$ 19,33 e $HR_{10} \leq$ 70,5 Então \rightarrow B (2; 1,0)</p>	<p>Regra 4 Se pH $>$ 8,5 e $T_{am} \leq$ 19,33 e $HR_{10} >$ 70,5 Então \rightarrow C (7; 1,0)</p>	

Tabela V.2 – Pesos associados às conexões entre a camada de entrada e a camada intermédia I para o modelo apresentado na Figura 30.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$	$h_{5,1}$	$h_{6,1}$
HR_{arm}	2,13	6,18	7,11	8,28	-10,71	5,62
PP_{arm}	14,57	-5,73	0,1	9,08	4,24	12,51
T_{arm}	-3,09	3,54	-10,04	-0,02	3,77	-0,41
Volume Arm.	-17,36	10,79	7,98	-7,48	-7,99	-3,47
pH	10,67	-9,83	-9,67	-3,63	15,44	2,28
Bias	5,69	-8,08	-0,62	-1,57	-4,12	4,59

Tabela V.3 – Pesos associados às conexões entre a camada intermédia I e a camada intermédia 2 para o modelo apresentado na Figura 30.

	$h_{1,2}$	$h_{2,2}$
$h_{1,1}$	-10,65	7,78
$h_{2,1}$	7,98	-7,21
$h_{3,1}$	5,18	10,15
$h_{4,1}$	-0,47	-10,14
$h_{5,1}$	-5,09	11,77
Bias	0,73	-4,96

Tabela V.4 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura 30.

	Classe B	Classe C
$h_{1,2}$	8,09	-8,14
$h_{2,2}$	-6,67	6,84
Bias	4,7	-4,68

ANEXO VI

Neste anexo apresentam-se os pesos associados às conexões entre os nodos das camadas adjacentes para a RNA apresentada na Figura 32. Os nodos das camadas intermédias são designados genericamente por h_{ij} , denotando o nodo i da camada intermédia j .

Tabela VI.1 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura 32.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$	$h_{5,1}$	$h_{6,1}$	$h_{7,1}$
CQO	13,57	-6,51	25,96	-16,23	-0,85	17,16	-3,58
OD	4,16	-5,08	1,09	3,44	-6,97	-6,19	-0,47
Oxidabilidade	2,74	-0,25	0,95	-4,77	-1,10	0,74	-1,95
SST	5,37	-1,60	10,47	-8,43	-0,35	8,08	-2,59
Bias	-6,24	1,97	-8,65	4,45	1,29	-7,89	-0,67

	$h_{8,1}$	$h_{9,1}$	$h_{10,1}$	$h_{11,1}$	$h_{12,1}$	$h_{13,1}$	$h_{14,1}$
CQO	-1,25	-12,16	-12,83	-20,69	-20,79	-15,38	1,92
OD	-12,11	-0,87	4,69	9,74	2,23	1,91	-21,03
Oxidabilidade	-4,68	-2,86	-5,06	-7,97	-2,82	-4,66	-4,33
SST	-2,53	-6,45	-6,21	-5,80	-9,96	-8,24	2,86
Bias	9,26	3,76	2,54	5,72	6,91	4,46	5,64

Tabela VI.2 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura 32.

	$h_{1,2}$	$h_{2,2}$	$h_{3,2}$	$h_{4,2}$	$h_{5,2}$
$h_{1,1}$	5,01	3,87	-5,35	-3,43	-4,89
$h_{2,1}$	-3,15	-2,87	-1,16	3,24	2,51
$h_{3,1}$	7,53	10,49	-4,30	-1,81	-10,13
$h_{4,1}$	-2,71	-6,14	2,29	-1,54	2,08
$h_{5,1}$	-3,14	-0,48	-1,36	3,85	1,19
$h_{6,1}$	-4,29	7,35	-5,38	-3,87	-4,98
$h_{7,1}$	-1,83	-1,40	0,91	0,67	-0,52
$h_{8,1}$	-9,27	-6,19	-4,71	-3,27	-4,19
$h_{9,1}$	-4,29	-4,85	2,51	-3,26	1,89
$h_{10,1}$	-1,72	-5,28	3,19	1,41	1,30

	$h_{1,2}$	$h_{2,2}$	$h_{3,2}$	$h_{4,2}$	$h_{5,2}$
$h_{1,1}$	-1,83	-7,43	3,23	-1,40	-4,53
$h_{12,1}$	-2,79	-7,61	3,88	-7,05	4,45
$h_{13,1}$	-4,12	-5,47	2,83	-1,96	2,35
$h_{14,1}$	-12,39	-2,12	-5,51	-0,74	2,74
Bias	3,49	4,33	-0,56	0,75	1,06

Tabela VI.3 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura 32.

	Classe B	Classe C
$h_{1,2}$	-7,21	7,26
$h_{2,2}$	-4,69	4,72
$h_{3,2}$	4,59	-4,53
$h_{4,2}$	-4,74	4,77
$h_{5,2}$	-6,25	6,29
Bias	6,23	-6,26

ANEXO VII

Neste anexo apresentam-se os modelos relativos à previsão da qualidade da água da albufeira de Odivelas, obtidos com recurso a Árvores de Decisão e RNAs e com restrição de alguns dos parâmetros de entrada. Apresentam-se, ainda, o conjunto de regras geradas destacando-se, no final, o suporte e a confiança da regra (Árvores de Decisão) e os pesos associados às conexões entre os nodos das camadas adjacentes (RNAs). Os nodos das camadas intermédias são designados genericamente por h_{ij} , denotando o nodo i da camada intermédia j .

VII.1 – Modelos excluindo o parâmetro CQO

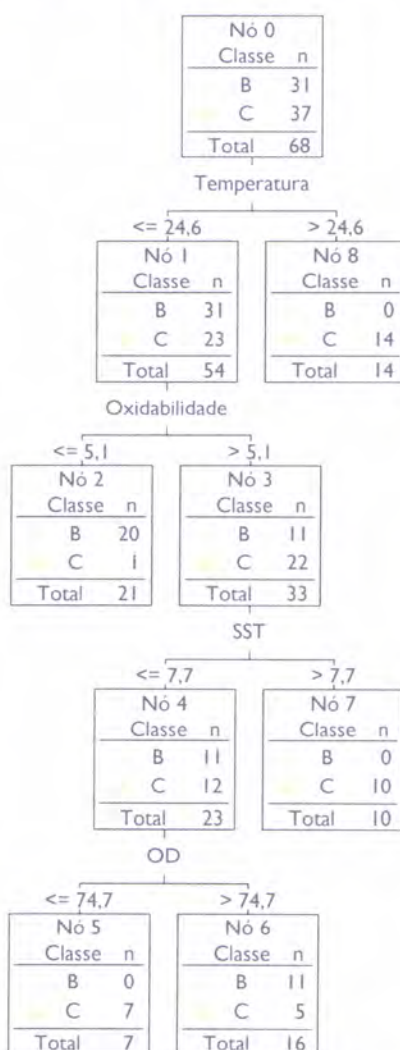


Figura VII.1 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo o parâmetro CQO.

Tabela VII.1 – Conjunto de regras relativo ao modelo apresentado na Figura VII.1.

Regras para a Classe B	Regras para a Classe C	
<p>Regra 1</p> <p>Se Temperatura $\leq 24,6$ e Oxidabilidade $\leq 5,1$</p> <p>Então $\rightarrow B$ (21; 0,95)</p>	<p>Regra 1</p> <p>Se Temperatura $\leq 24,6$ e Oxidabilidade $> 5,1$ e SST $\leq 7,7$ e Oxigénio Dissolvido $\leq 74,7$</p> <p>Então $\rightarrow C$ (7; 1,0)</p>	<p>Regra 3</p> <p>Se Temperatura $> 24,6$</p> <p>Então $\rightarrow C$ (14; 1,0)</p>
<p>Regra 2</p> <p>Se Temperatura $\leq 24,6$ e Oxidabilidade $> 5,1$ e SST $\leq 7,7$ e Oxigénio Dissolvido $> 74,7$</p> <p>Então $\rightarrow B$ (16; 0,69)</p>	<p>Regra 2</p> <p>Se Temperatura $\leq 24,6$ e Oxidabilidade $> 5,1$ e SST $> 7,7$</p> <p>Então $\rightarrow C$ (10; 1,0)</p>	

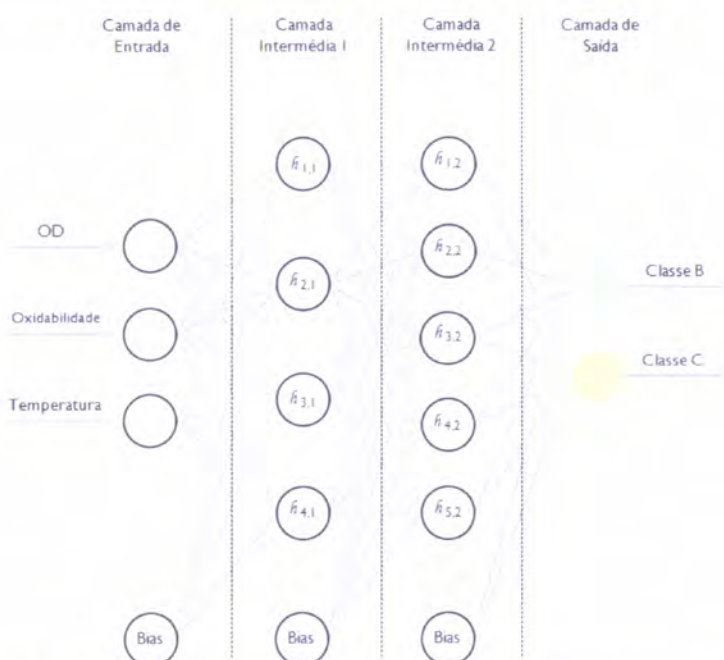


Figura VII.2 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo o parâmetro CQO.

Tabela VII.2 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura VII.2.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$
OD	20,08	-24,78	-13,54	-18,12
Oxidabilidade	-33,05	-14,71	11,31	2,29
Temperatura	-7,75	-11,06	32,04	-8,37
Bias	2,93	20,69	-12,54	18,52

Tabela VII.3 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura VII.2.

	$h_{1,2}$	$h_{2,2}$	$h_{3,2}$	$h_{4,2}$	$h_{5,2}$
$h_{1,1}$	-9,59	2,92	0,66	-2,65	-12,40
$h_{2,1}$	-8,07	6,31	-0,34	2,34	-12,10
$h_{3,1}$	-9,67	-1,33	3,38	2,71	22,85
$h_{4,1}$	8,38	7,88	-1,14	2,78	-5,79
Bias	4,75	6,51	8,92	2,71	-1,13

Tabela VII.4 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura VII.2.

	Classe B	Classe C
$h_{1,2}$	-7,97	8,07
$h_{2,2}$	-11,29	11,24
$h_{3,2}$	12,02	-12,07
$h_{4,2}$	-5,50	5,67
$h_{5,2}$	-4,64	4,65
Bias	7,18	-7,25

VII.2 – Modelos excluindo o parâmetro oxidabilidade

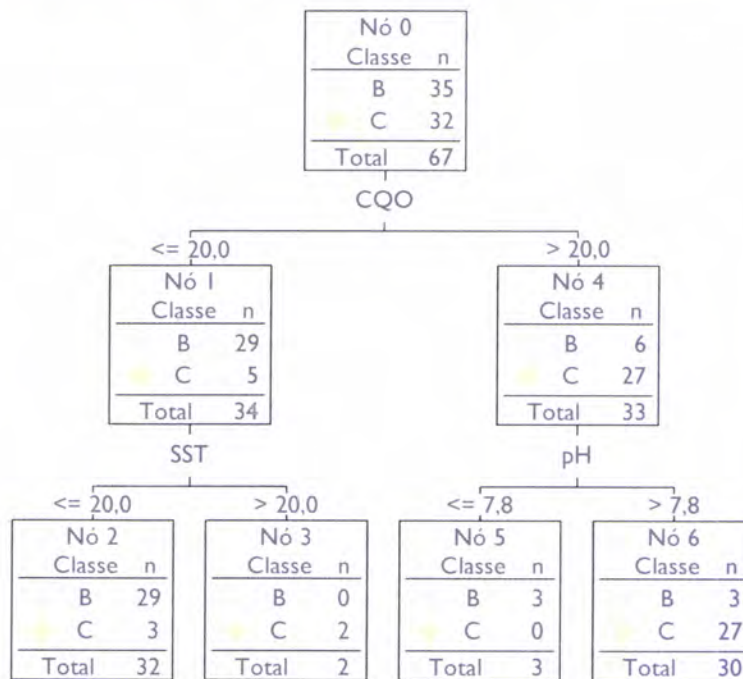


Figura VII.3 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo o parâmetro oxidabilidade.

Tabela VII.5 – Conjunto de regras relativo ao modelo apresentado na Figura VII.3.

Regras para a Classe B	Regras para a Classe C
<p>Regra 1 Se CQO \leq 20,0 e SST \leq 20,0 Então \rightarrow B (32; 0,91)</p>	<p>Regra 1 Se CQO \leq 20,0 e SST $>$ 20,0 Então \rightarrow C (2; 1,0)</p>
<p>Regra 2 Se CQO $>$ 20,0 e pH \leq 7,8 Então \rightarrow B (3; 1,0)</p>	<p>Regra 2 Se CQO $>$ 20,0 e pH $>$ 7,8 Então \rightarrow C (30; 0,90)</p>

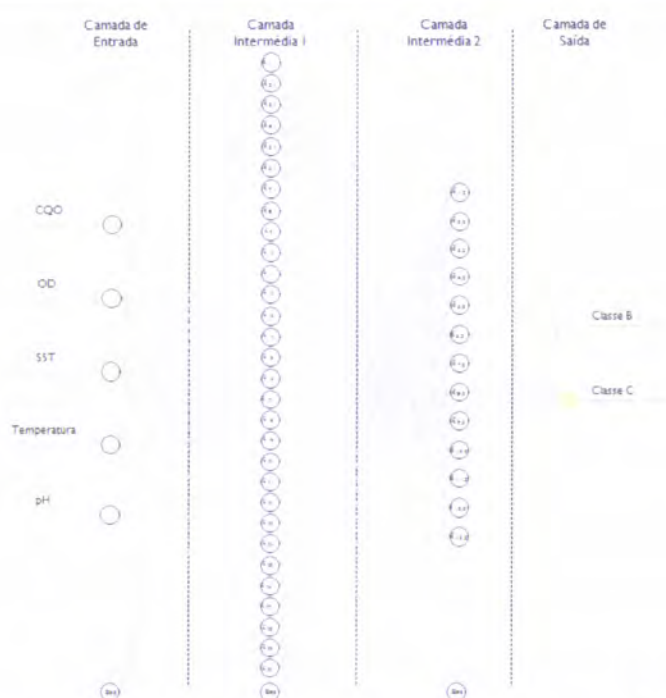


Figura VII.4 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo o parâmetro oxidabilidade.

Tabela VII.6 – Pesos associados às conexões entre a camada de entrada e a camada intermédia I para o modelo apresentado na Figura VII.4.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$	$h_{5,1}$	$h_{6,1}$	$h_{7,1}$	$h_{8,1}$	$h_{9,1}$	$h_{10,1}$
CQO	2,60	3,03	9,24	-2,61	7,75	2,13	7,19	-2,79	0,84	1,97
OD	-0,80	-3,66	-1,92	-0,58	2,83	-2,84	-2,83	-0,63	-2,75	-1,28
SST	1,27	2,24	4,03	-1,94	3,68	2,17	4,23	-1,44	0,98	1,55
Temperatura	0,30	6,31	-0,99	-0,99	-5,04	2,83	3,24	-0,49	1,37	1,33
pH	-0,38	-0,56	-2,74	-0,22	-1,19	-0,39	-0,54	-0,56	-0,30	0,41
Bias	-2,22	-3,85	-2,59	0,24	-3,46	-2,53	-4,36	0,21	-1,19	-2,85

	$h_{11,1}$	$h_{12,1}$	$h_{13,1}$	$h_{14,1}$	$h_{15,1}$	$h_{16,1}$	$h_{17,1}$	$h_{18,1}$	$h_{19,1}$	$h_{20,1}$
CQO	-5,19	-3,65	-2,78	-15,29	-7,16	16,13	1,55	-16,25	-12,39	-9,32
OD	-0,23	-0,74	-1,02	-0,26	-1,30	0,99	0,15	-0,21	-0,24	-0,52
SST	-2,12	-2,16	-0,32	-3,12	-1,81	6,53	0,37	-6,45	-2,29	-3,85
Temperatura	-0,41	-1,03	0,09	0,44	-0,19	-0,13	-2,32	1,18	0,40	-0,67
pH	-0,54	-0,02	-0,24	-1,12	-0,11	1,11	-0,75	-0,59	-0,72	-0,39
Bias	1,04	0,69	-0,32	4,48	1,98	-8,19	-0,50	6,62	3,23	3,05

	$h_{21,1}$	$h_{22,1}$	$h_{23,1}$	$h_{24,1}$	$h_{25,1}$	$h_{26,1}$	$h_{27,1}$	$h_{28,1}$	$h_{29,1}$	$h_{30,1}$
CQO	3,89	20,85	18,68	8,49	-14,64	8,16	-0,04	5,85	9,34	-3,79
OD	-0,91	0,69	1,35	-1,22	-0,27	-0,99	-0,69	-2,20	-5,98	-0,34
SST	2,26	7,05	3,95	3,38	-3,56	3,51	-0,56	3,50	3,17	-2,56
Temperatura	-0,21	0,19	-1,01	2,14	0,10	0,33	-0,62	2,72	4,89	-0,81
pH	-0,32	1,29	0,91	0,71	-1,11	0,06	-0,74	0,64	-1,31	-0,25
Bias	-2,14	-9,49	-6,97	-5,04	4,56	-3,22	-0,93	-4,26	-4,58	0,56

Tabela VII.7 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura VII.4.

	$h_{1,2}$	$h_{2,2}$	$h_{3,2}$	$h_{4,2}$	$h_{5,2}$	$h_{6,2}$	$h_{7,2}$	$h_{8,2}$	$h_{9,2}$	$h_{10,2}$	$h_{11,2}$	$h_{12,2}$	$h_{13,2}$
$h_{1,1}$	-0,79	-0,36	-0,62	-1,51	-0,68	-0,38	-0,28	0,02	-0,41	-1,19	-0,38	-0,92	-0,62
$h_{2,1}$	-0,56	0,93	-0,43	-2,39	-1,01	-2,51	-2,99	-1,16	-0,49	0,61	-1,21	-2,41	-0,34
$h_{3,1}$	-0,84	-2,55	-0,87	-2,91	-0,65	-1,37	-1,86	-1,07	-1,12	-2,17	-0,95	-1,99	-0,71
$h_{4,1}$	-0,05	-0,39	-0,49	-0,31	-0,35	1,35	1,88	-0,01	-0,16	-0,25	0,07	1,18	0,14
$h_{5,1}$	-0,25	-4,27	-0,09	-1,84	-0,73	-2,25	-2,25	-0,06	-0,72	-2,79	-0,24	-1,74	-0,31
$h_{6,1}$	-0,94	1,01	-0,87	-1,89	-0,89	-1,14	-1,09	-0,72	-0,27	-0,23	-1,07	-0,72	-0,55
$h_{7,1}$	-1,02	-0,17	-1,05	-2,76	-0,82	-2,14	-2,16	-0,33	-0,59	-1,12	-1,62	-2,06	-1,22
$h_{8,1}$	0,14	0,20	-0,28	-0,21	0,18	1,66	-1,09	-0,42	0,14	-0,17	-0,49	1,48	-0,01
$h_{9,1}$	-0,69	0,07	-0,85	-2,28	-0,71	-0,57	-0,36	-0,39	-0,44	0,54	-0,39	-0,04	-0,49
$h_{10,1}$	-0,12	0,27	-0,35	-1,29	-1,13	-0,66	-0,72	-0,35	-0,58	-0,48	-0,68	-0,57	-0,49
$h_{11,1}$	0,50	0,51	0,01	0,91	0,28	2,19	2,05	0,32	0,14	0,76	0,46	1,71	-0,28
$h_{12,1}$	0,42	-0,01	-0,21	-0,36	-0,11	2,14	1,56	0,47	0,10	-0,37	-0,07	1,72	0,09
$h_{13,1}$	-0,73	-0,57	-0,65	-0,46	-0,12	1,32	1,02	-0,49	-0,71	0,64	-0,49	1,18	-0,21
$h_{14,1}$	-0,19	2,99	-0,22	2,30	0,29	3,74	5,08	0,00	0,43	3,06	0,30	2,62	0,07
$h_{15,1}$	-0,09	1,38	-0,29	-0,72	-0,13	2,06	2,82	-0,47	0,04	1,11	-0,49	1,80	0,13
$h_{16,1}$	-1,26	-4,52	-0,68	-3,81	-1,15	-2,56	-3,57	-1,21	-0,53	-3,40	-0,68	-2,28	-1,03
$h_{17,1}$	-0,32	-2,08	-0,63	-0,90	-0,21	0,32	-0,49	-0,79	-0,22	-1,68	-0,53	0,29	-0,28
$h_{18,1}$	0,19	1,57	0,61	1,38	0,45	4,51	4,98	0,13	0,59	1,50	0,61	3,35	0,67
$h_{19,1}$	0,41	2,63	0,49	1,34	0,54	3,47	3,62	0,02	0,09	1,67	0,06	2,36	-0,21
$h_{20,1}$	0,42	1,06	0,05	0,84	0,18	2,98	3,83	0,22	-0,11	-1,07	0,11	2,23	-0,21

	$h_{1,2}$	$h_{2,2}$	$h_{3,2}$	$h_{4,2}$	$h_{5,2}$	$h_{6,2}$	$h_{7,2}$	$h_{8,2}$	$h_{9,2}$	$h_{10,2}$	$h_{11,2}$	$h_{12,2}$	$h_{13,2}$
$h_{21,1}$	-0,66	-1,19	-0,59	-1,83	-0,83	-0,89	-0,68	-0,48	-0,79	-1,18	-1,23	-0,29	-0,29
$h_{22,1}$	-0,91	-5,29	-1,02	-4,78	-1,43	-4,19	-4,37	-0,76	-1,29	-4,67	-1,35	-2,88	-0,64
$h_{23,1}$	-0,63	-5,76	-0,69	-3,42	-0,77	-3,00	-3,16	-0,68	-1,29	-4,89	-0,54	-2,68	-1,01
$h_{24,1}$	-0,95	-1,62	-1,09	-2,69	-1,29	-1,55	-2,21	-0,43	-0,74	-1,06	-0,43	-1,37	-0,56
$h_{25,1}$	0,21	2,65	0,18	1,95	-0,01	3,60	5,11	0,01	-0,01	2,60	0,57	3,31	0,28
$h_{26,1}$	-0,46	-1,92	-0,72	-2,88	-1,26	-0,69	-1,89	-0,86	-1,04	-1,73	-0,73	-0,52	-1,22
$h_{27,1}$	0,14	-0,67	-0,63	-1,01	0,09	0,74	0,89	-0,09	-0,87	-0,55	-0,84	0,10	-0,23
$h_{28,1}$	-0,58	-0,23	-0,82	-3,19	-0,46	-1,33	-2,19	-1,00	-1,08	-0,99	-0,52	-1,45	-0,85
$h_{29,1}$	-0,94	-1,49	-0,30	-3,56	-0,59	-3,93	-4,22	-0,34	-1,22	-1,55	-0,35	-2,74	-0,39
$h_{30,1}$	-0,19	-0,02	0,31	0,56	0,11	1,85	1,52	-0,17	-0,19	0,05	0,17	1,93	0,06
Bias	-0,72	-1,29	-0,18	-1,37	-0,56	4,20	5,49	-0,23	-0,41	-0,72	-0,62	3,74	-0,74

Tabela VII.8 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura VII.4.

	Classe B	Classe C
$h_{1,2}$	0,88	-0,95
$h_{2,2}$	-5,74	5,80
$h_{3,2}$	0,72	-0,79
$h_{4,2}$	5,24	-5,58
$h_{5,2}$	1,48	-0,77
$h_{6,2}$	3,13	-2,25
$h_{7,2}$	3,05	-3,66
$h_{8,2}$	0,55	-0,81
$h_{9,2}$	0,78	-0,93
$h_{10,2}$	-4,67	4,54
$h_{11,2}$	1,48	-0,76
$h_{12,2}$	1,76	-2,10
$h_{13,2}$	0,81	-0,73
Bias	-2,11	2,11

VII.3 – Modelos excluindo o parâmetro SST

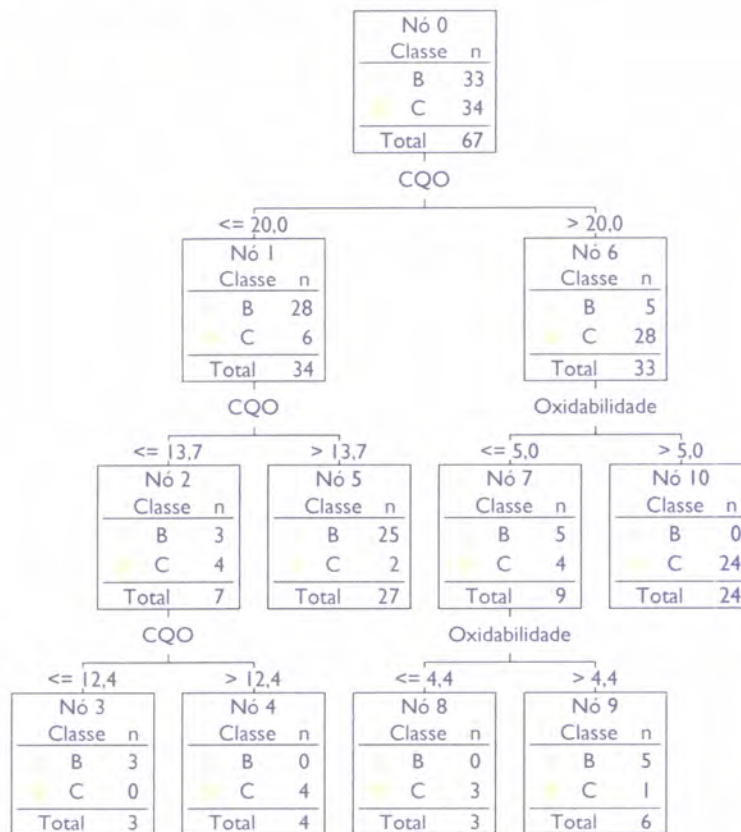


Figura VII.5 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo o parâmetro SST.

Tabela VII.9 – Conjunto de regras relativo ao modelo apresentado na Figura VII.5.

Regras para a Classe B		Regras para a Classe C	
<p>Regra 1 Se CQO ≤ 20 e CQO ≤ 13,7 e CQO ≤ 12,4</p> <p>Então → B (3; 1,0)</p>	<p>Regra 3 Se CQO > 20 e Oxidabilidade ≤ 5,0 e Oxidabilidade > 4,4</p> <p>Então → B (6; 0,83)</p>	<p>Regra 1 Se CQO ≤ 20 e CQO ≤ 13,7 e CQO > 12,4</p> <p>Então → C (4; 1,0)</p>	<p>Regra 3 Se CQO > 20 e Oxidabilidade > 5,0</p> <p>Então → C (24; 1,0)</p>
<p>Regra 2 Se CQO ≤ 20 e CQO > 13,7</p> <p>Então → B (27; 0,93)</p>		<p>Regra 2 Se CQO > 20 e Oxidabilidade ≤ 5,0 e Oxidabilidade ≤ 4,4</p> <p>Então → C (3; 1,0)</p>	

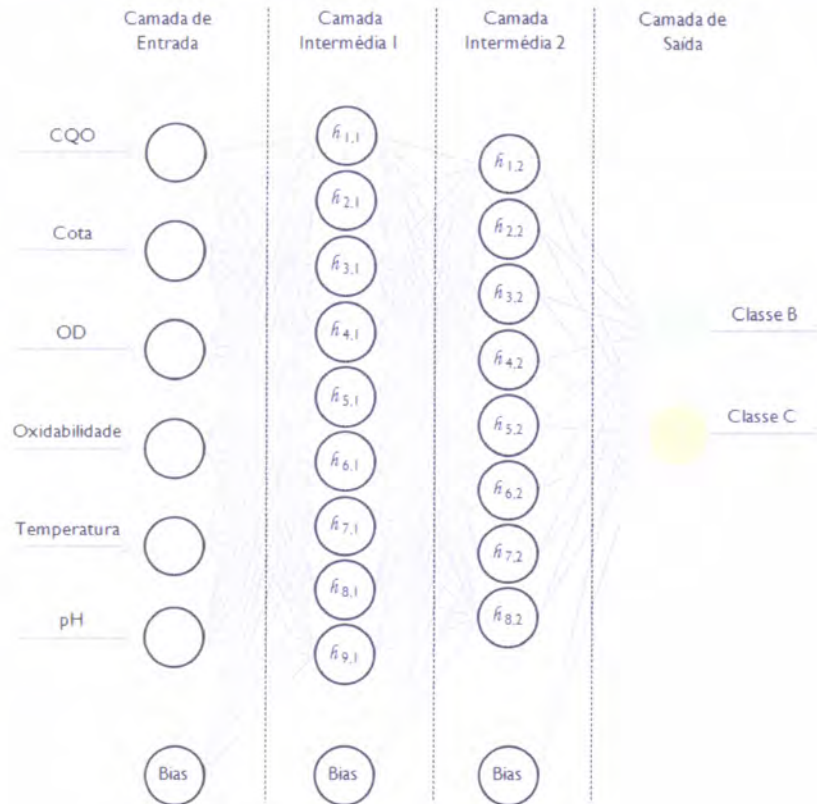


Figura VII.6 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo o parâmetro SST.

Tabela VII.10 – Pesos associados às conexões entre a camada de entrada e a camada intermédia I para o modelo apresentado na Figura VII.6.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$	$h_{5,1}$	$h_{6,1}$	$h_{7,1}$	$h_{8,1}$	$h_{9,1}$
CQO	5,32	12,13	6,66	3,66	-12,59	5,06	-12,37	-3,77	4,11
Cota	-10,02	-5,18	-3,23	-4,13	-16,36	-5,95	-8,52	-4,85	6,81
OD	-0,06	3,44	4,65	-1,47	2,69	-5,92	-13,47	0,51	-8,2
Oxidabilidade	1,69	13,86	8,46	8,36	-7,17	11,42	9,15	-3,72	-8,38
Temperatura	-15,15	0,63	4,71	5,96	0,63	0,36	11,45	-0,37	-0,48
pH	6,03	11,7	5,36	1,19	-1,68	4,51	-7,78	0,79	-6,57
Bias	5,43	-14,24	-12,09	-6,19	13,15	-4,89	11,35	1,92	-1,61

Tabela VII.11 – Pesos associados às conexões entre a camada intermédia I e a camada intermédia 2 para o modelo apresentado na Figura VII.6.

	$h_{1,2}$	$h_{2,2}$	$h_{3,2}$	$h_{4,2}$	$h_{5,2}$	$h_{6,2}$	$h_{7,2}$	$h_{8,2}$
$h_{1,1}$	-1,2	-2,41	-1,92	-4,37	-1,7	-2,27	-1,59	-1,57
$h_{2,1}$	-2,997	-2,9	-2,25	-6,73	-3,38	-1,57	-2,69	-3,48
$h_{3,1}$	-1,21	-1,19	-1,66	-7,08	-1,89	-1,01	-2,52	-2,37

	$h_{1,2}$	$h_{2,2}$	$h_{3,2}$	$h_{4,2}$	$h_{5,2}$	$h_{6,2}$	$h_{7,2}$	$h_{8,2}$
$h_{4,1}$	-1,66	-0,23	-1,18	-3,72	-1,21	-0,56	-1,13	-0,57
$h_{5,1}$	5,3	0,27	2,73	7,09	3,36	4,84	4,1	7,54
$h_{6,1}$	-1,48	-0,97	-0,58	-7,89	-0,83	0,07	-1,43	0,37
$h_{7,1}$	-6,28	-3,04	-1,52	-1,96	-4,65	-1,62	-1,91	-0,9
$h_{8,1}$	0,67	-0,17	-0,09	3,91	0,66	1,89	1,03	1,9
$h_{9,1}$	1,06	2,31	2,18	5,6	2,36	-0,2	1,31	-1,32
Bias	1,87	0,34	0,41	3,58	0,48	0,97	0,63	-0,86

Tabela VII.12 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura VII.6.

	Classe B	Classe C
$h_{1,2}$	3,91	-2,95
$h_{2,2}$	2,88	-2,44
$h_{3,2}$	2,27	-1,73
$h_{4,2}$	6,28	-5,68
$h_{5,2}$	2,46	-2,65
$h_{6,2}$	2,41	-2,28
$h_{7,2}$	1,68	-1,64
$h_{8,2}$	3,61	-3,36
Bias	-12,46	11,08

Tabela VII.13 – Conjunto de regras relativo ao modelo apresentado na Figura VII.7.

Regras para a Classe B	Regras para a Classe C	
Regra 1 Se CQO \leq 20,0 e pH \leq 8,5 e SST \leq 9,4 e CQO \leq 13,7 e CQO \leq 12,4 Então \rightarrow B (2; 1,0)	Regra 1 Se CQO \leq 20,0 e pH \leq 8,5 e SST \leq 9,4 e CQO \leq 13,7 e CQO $>$ 12,4 Então \rightarrow C (3; 0,67)	Regra 4 Se CQO $>$ 20,0 e Oxidabilidade \leq 4,9 e Temperatura $>$ 23,0 Então \rightarrow C (3; 1,0)
Regra 2 Se CQO \leq 20,0 e pH \leq 8,5 e SST \leq 9,4 e CQO $>$ 13,7 Então \rightarrow B (25; 1,0)	Regra 2 Se CQO \leq 20,0 e pH \leq 8,5 e SST $>$ 9,4 Então \rightarrow C (2; 1,0)	Regra 5 Se CQO $>$ 20,0 e Oxidabilidade $>$ 4,9 Então \rightarrow C (23; 1,0)
Regra 3 Se CQO $>$ 20,0 e Oxidabilidade \leq 4,9 e Temperatura \leq 23,0 Então \rightarrow B (5; 0,80)	Regra 3 Se CQO \leq 20,0 e pH $>$ 8,5 Então \rightarrow C (2; 1,0)	

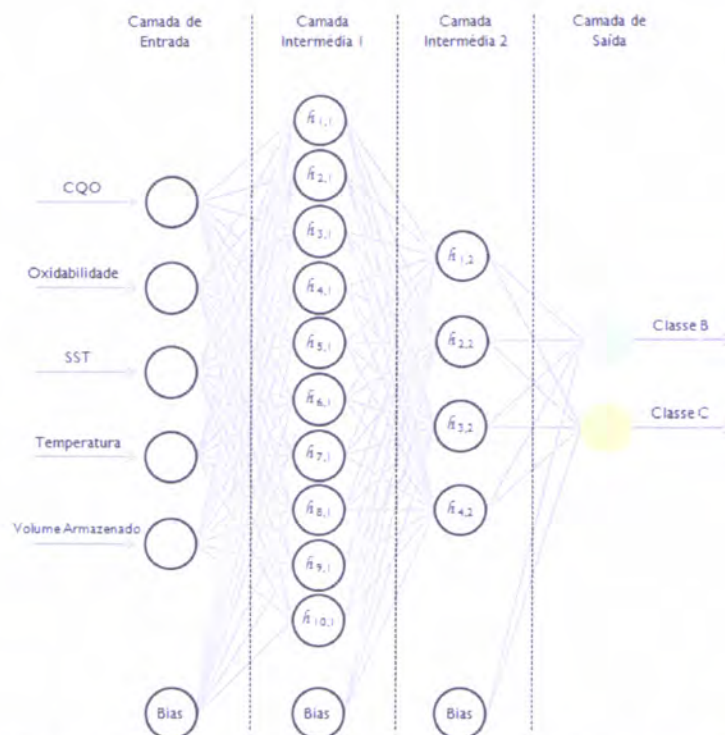


Figura VII.8 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo o parâmetro oxigénio dissolvido.

Tabela VII.14 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura VII.8.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$	$h_{5,1}$	$h_{6,1}$	$h_{7,1}$	$h_{8,1}$	$h_{9,1}$	$h_{10,1}$
CQO	-11,09	4,34	21,54	-26,77	-12,21	7,72	-16,16	-13,7	24,69	8,26
Oxidabilidade	-3,69	2,58	8,05	-4,08	-3,49	-8,84	-3,62	2,27	5,11	3,88
SST	-1,26	-0,17	7,83	-12,92	-7,43	4,97	-10,73	-2,83	6,37	7,36
Temperatura	-8,48	20,43	2,17	1,75	1,67	-11,52	0,85	1,76	2,61	-12,78
Volume Arm.	-1,72	-5,25	1,42	1,09	8,49	14,67	8,49	-4,47	3,11	12,42
Bias	13,19	-14,25	-9,76	9,75	-1,06	-6,82	1,07	6,52	-11,62	2,79

Tabela VII.15 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura VII.8.

	$h_{1,2}$	$h_{2,2}$	$h_{3,2}$	$h_{4,2}$
$h_{1,1}$	5,21	3,14	-8,65	-0,39
$h_{2,1}$	-7,62	1,62	6,09	-6,08
$h_{3,1}$	1,74	-0,74	8,23	-1,59
$h_{4,1}$	3,08	5,98	-7,92	2,76
$h_{5,1}$	-2,03	-0,61	-8,52	1,38
$h_{6,1}$	-9,55	-15,88	3,59	2,78
$h_{7,1}$	-3,24	-0,24	-7,72	1,77
$h_{8,1}$	0,71	4,84	-5,81	-3,96
$h_{9,1}$	1,82	-0,57	8,09	-1,62
$h_{10,1}$	-5,22	-4,29	3,27	-7,83
Bias	3,63	5,09	1,38	1,34

Tabela VII.16 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura VII.8.

	Classe B	Classe C
$h_{1,2}$	5,41	-5,57
$h_{2,2}$	-5,46	5,79
$h_{3,2}$	-6,48	6,76
$h_{4,2}$	4,12	-4,45
Bias	3,34	-3,57

VII.5 – Modelos excluindo os parâmetros CQO e oxidabilidade

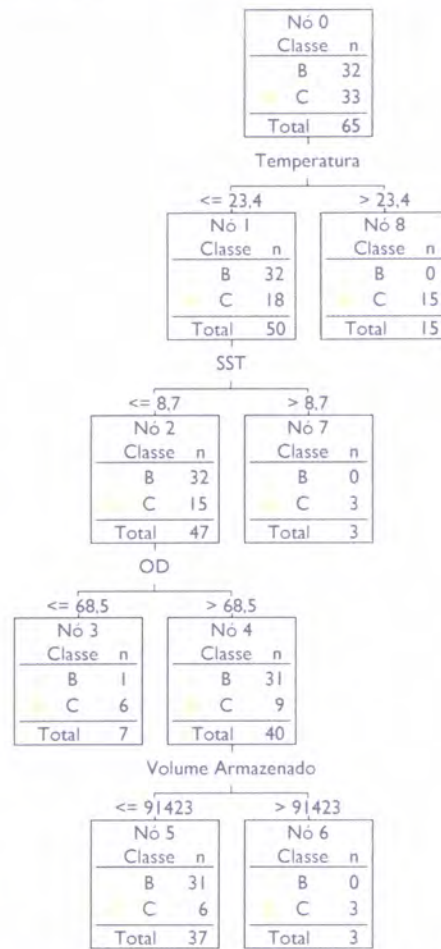


Figura VII.9 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo os parâmetros CQO e oxidabilidade.

Tabela VII.17 – Conjunto de regras relativo ao modelo apresentado na Figura VII.9.

Regras para a Classe B	Regras para a Classe C	
<p>Regra 1</p> <p>Se Temperatura $\leq 23,4$ e SST $\leq 8,7$ e OD $> 68,5$ e Volume Armazenado ≤ 91423</p> <p>Então \rightarrow B (37; 0,84)</p>	<p>Regra 1</p> <p>Se Temperatura $\leq 23,4$ e SST $\leq 8,7$ e OD $\leq 68,5$</p> <p>Então \rightarrow C (7; 0,86)</p>	<p>Regra 3</p> <p>Se Temperatura $\leq 23,4$ e SST $> 8,7$</p> <p>Então \rightarrow C (3; 1,0)</p>
	<p>Regra 2</p> <p>Se Temperatura $\leq 23,4$ e SST $\leq 8,7$ e OD $> 68,5$ e Volume Armazenado > 91423</p> <p>Então \rightarrow C (3; 1,0)</p>	<p>Regra 4</p> <p>Se Temperatura $> 23,4$</p> <p>Então \rightarrow C (15; 1,0)</p>

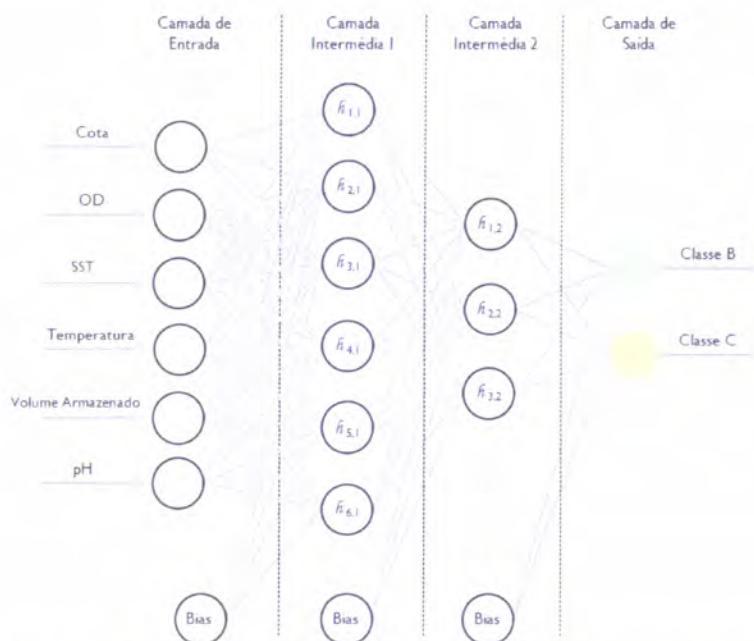


Figura VII.10 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo os parâmetros CQO e oxidabilidade.

Tabela VII.18 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura VII.10.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$	$h_{5,1}$	$h_{6,1}$
Cota	-1,40	3,24	-2,92	6,31	1,57	1,32
OD	10,27	1,71	-1,12	-6,22	-23,72	-11,06
SST	-7,82	-0,49	14,61	1,31	1,59	1,87
Temperatura	-7,85	22,49	1,68	8,49	-4,43	10,85
Volume Arm.	0,18	-6,34	6,17	-1,94	-1,36	-10,09
pH	-4,07	-3,39	13,68	0,51	3,39	7,78
Bias	1,43	-10,34	-6,67	-3,66	6,59	3,73

Tabela VII.19 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura VII.10.

	$h_{1,2}$	$h_{2,2}$	$h_{3,2}$
$h_{1,1}$	-5,22	6,05	7,15
$h_{2,1}$	6,95	0,36	-11,19
$h_{3,1}$	-6,89	-7,34	3,59
$h_{4,1}$	4,31	-1,46	-6,05
$h_{5,1}$	6,11	-2,79	-12,42
$h_{6,1}$	6,32	-3,61	-3,19
Bias	1,01	0,66	11,01

Tabela VII.20 – pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura VII.10.

	Classe B	Classe C
$h_{1,2}$	8,78	-8,73
$h_{2,2}$	7,16	-7,17
$h_{3,2}$	7,65	-7,61
Bias	-10,26	10,21

VII.6 – Modelos excluindo os parâmetros SST, OD e oxidabilidade

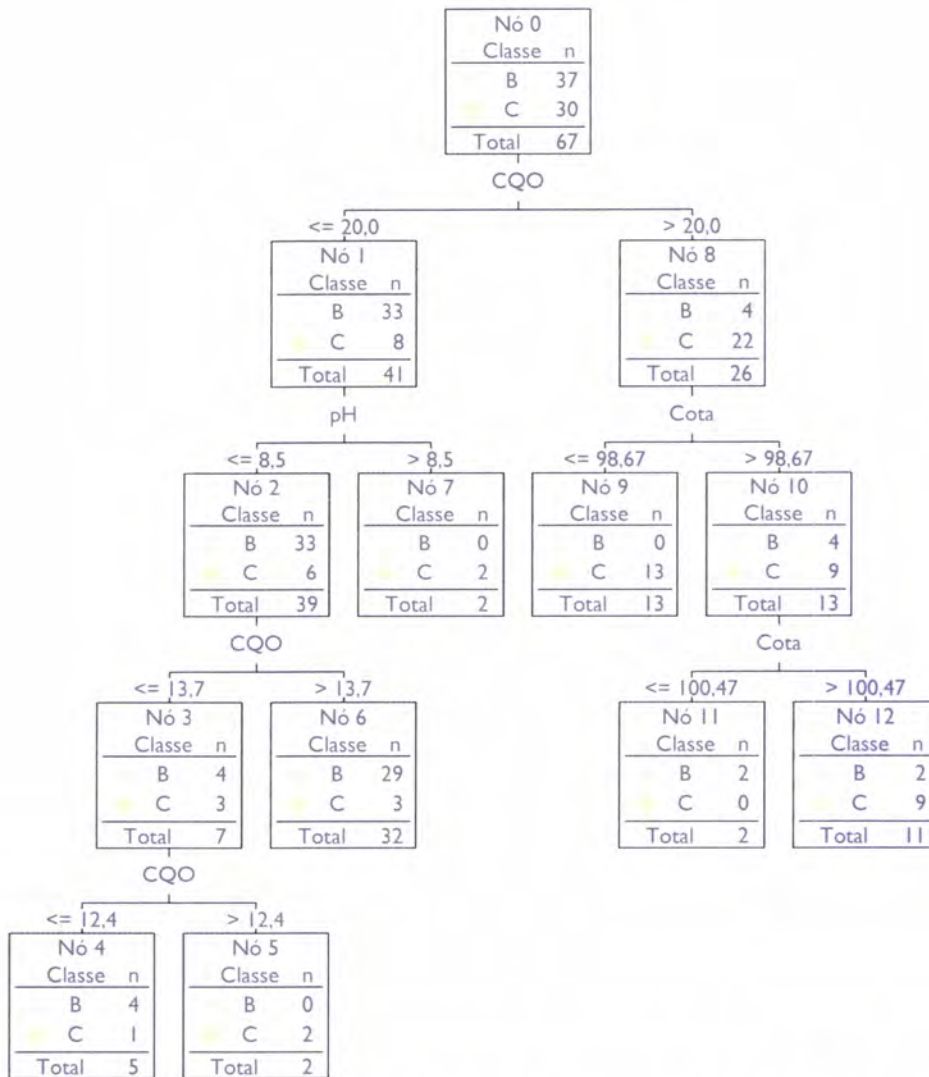


Figura VII.11 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo os parâmetros SST, OD e oxidabilidade.

Tabela VII.21 – Conjunto de regras relativo ao modelo apresentado na Figura VII.11.

Regras para a Classe B		Regras para a Classe C	
Regra 1 Se CQO ≤ 20 e pH ≤ 8,5 e CQO ≤ 13,7 e CQO ≤ 12,4 Então → B (5; 0,80)	Regra 3 Se CQO > 20 e Cota > 98,67 e Cota ≤ 100,47 Então → B (2; 1,0)	Regra 1 Se CQO ≤ 20 e pH ≤ 8,5 e CQO ≤ 13,7 e CQO > 12,4 Então → C (2; 1,0)	Regra 3 Se CQO > 20 e Cota ≤ 98,67 Então → C (13; 1,0)
Regra 2 Se CQO ≤ 20 e pH ≤ 8,5 e CQO > 13,7 Então → B (32; 0,91)		Regra 2 Se CQO ≤ 20 e pH > 8,5 Então → C (2; 1,0)	Regra 4 Se CQO > 20 e Cota > 98,67 e Cota > 100,47 Então → C (11; 0,82)



Figura VII.12 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo os parâmetros SST, OD e oxidabilidade.

Tabela VII.22 – Pesos associados às conexões entre a camada de entrada e a camada intermédia I para o modelo apresentado na Figura VII.12.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$	$h_{5,1}$
CQO	-27,67	-30,54	-25,64	-11,16	-30,81
Cota	4,93	-4,87	6,59	-0,14	-7,45
Temperatura	-4,63	-1,23	-6,68	0,52	-4,57
Volume Arm.	6,66	1,62	10,35	4,67	1,48
pH	-1,97	2,95	2,38	1,78	-3,6
Bias	8,34	12,47	11,45	12,39	18,32

Tabela VII.23 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura VII.12.

	$h_{1,2}$
$h_{1,1}$	-9,17
$h_{2,1}$	-9,11
$h_{3,1}$	-6,98
$h_{4,1}$	7,87
$h_{5,1}$	-12,1
Bias	7,41

Tabela VII.24 – pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura VII.12.

	Classe B	Classe C
$h_{1,2}$	-6,29	6,03
Bias	1,85	-1,82

VII.7 – Modelos excluindo os parâmetros SST, OD e CQO

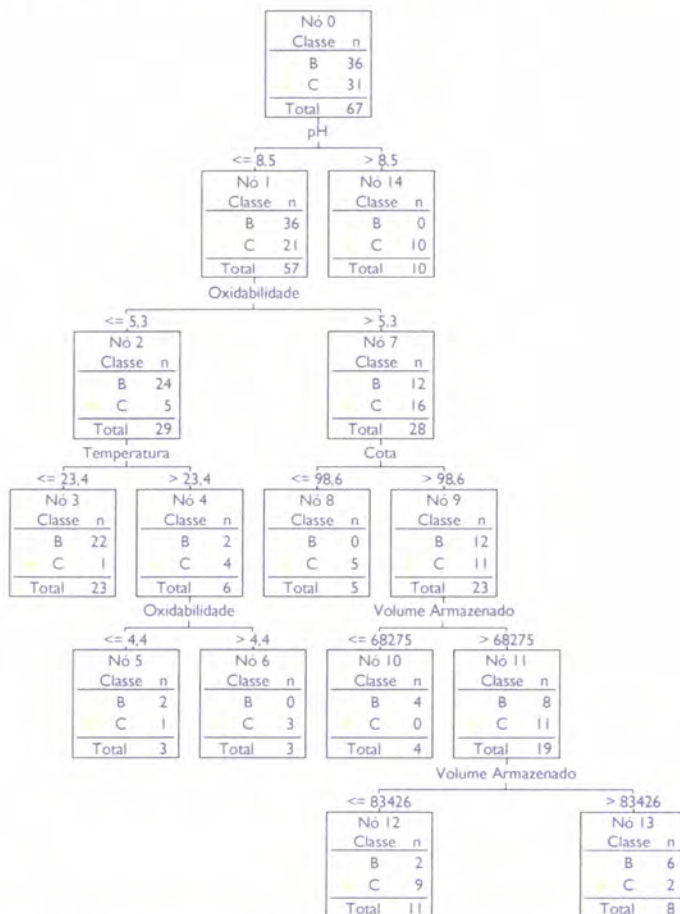


Figura VII.13 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo os parâmetros SST, OD e CQO.

Tabela VII.25 – conjunto de regras relativo ao modelo apresentado na Figura VII.13.

Regras para a Classe B	Regras para a Classe C
<p>Regra 1 Se pH \leq 8,5 e Oxidabilidade \leq 5,3 e Temperatura \leq 23,4 Então \rightarrow B (23; 0,96)</p>	<p>Regra 1 Se pH \leq 8,5 e Oxidabilidade \leq 5,3 e Temperatura $>$ 23,4 e Oxidabilidade $>$ 4,4 Então \rightarrow C (3; 1,0)</p>
<p>Regra 2 Se pH \leq 8,5 e Oxidabilidade \leq 5,3 e Temperatura $>$ 23,4 e Oxidabilidade \leq 4,4 Então \rightarrow B (3; 0,67)</p>	<p>Regra 2 Se pH \leq 8,5 e Oxidabilidade $>$ 5,3 e Cota \leq 98,6 Então \rightarrow C (5; 1,0)</p>
<p>Regra 3 Se pH \leq 8,5 e Oxidabilidade $>$ 5,3 e Cota $>$ 98,6 e Volume Armazenado \leq 68275 Então \rightarrow B (4; 1,0)</p>	<p>Regra 3 Se pH \leq 8,5 e Oxidabilidade $>$ 5,3 e Cota $>$ 98,6 e Volume Armazenado $>$ 68275 e Volume Armazenado \leq 83426 Então \rightarrow C (11; 0,82)</p>
<p>Regra 4 Se pH \leq 8,5 e Oxidabilidade $>$ 5,3 e Cota $>$ 98,6 e Volume Armazenado $>$ 68275 e Volume Armazenado $>$ 83426 Então \rightarrow B (8; 0,75)</p>	<p>Regra 4 Se pH $>$ 8,5 Então \rightarrow C (10; 1,0)</p>

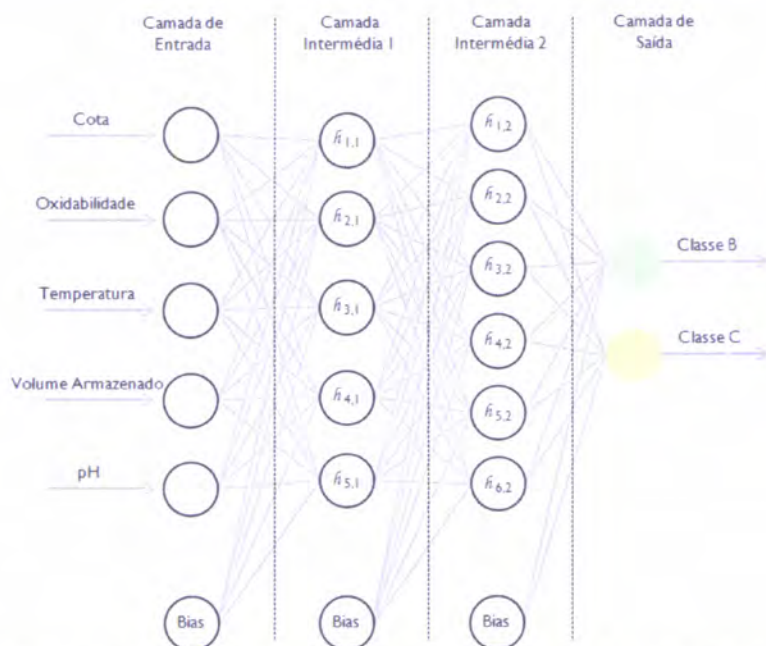


Figura VII.14 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo os parâmetros SST, OD e CQO.

Tabela VII.26 – pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura VII.14.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$	$h_{5,1}$
Cota	-2,42	3,44	-9,45	7,58	-3,2
Oxidabilidade	4,91	-10,5	9,16	2,87	18,5
Temperatura	1,74	-18,58	-37,16	-9,21	36,09
Volume Arm.	2,33	-5,52	-10,33	24,42	-1,78
pH	32,23	8,67	24,64	-12,26	-23,96
Bias	-24,19	0,89	9,22	5,63	-11,54

Tabela VII.27 – pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura VII.14.

	$h_{1,2}$	$h_{2,2}$	$h_{3,2}$	$h_{4,2}$	$h_{5,2}$	$h_{6,2}$
$h_{1,1}$	-7,46	-7,84	-2,93	-0,48	-9,65	-0,91
$h_{2,1}$	5,78	-0,79	4,24	4,16	-0,83	5,07
$h_{3,1}$	2,67	-8,7	-7,68	0,99	-11,17	2,8
$h_{4,1}$	-13,15	0,34	8,34	-5,51	-1,33	-3,37
$h_{5,1}$	-2,08	-10,63	-4,93	0,93	-13,49	-3,29
Bias	7,32	5,21	-2,96	3,1	7,66	-1,39

Tabela VII.28 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura VII.14.

	Classe B	Classe C
$h_{1,2}$	6,2	-6,01
$h_{2,2}$	3,18	-3,29
$h_{3,2}$	1,68	-1,49
$h_{4,2}$	2,14	-2,14
$h_{5,2}$	6,16	-6,04
$h_{6,2}$	4,53	-4,48
Bias	-9,72	9,52

VII.8 – Modelos excluindo os parâmetros SST, OD, oxidabilidade e CQO

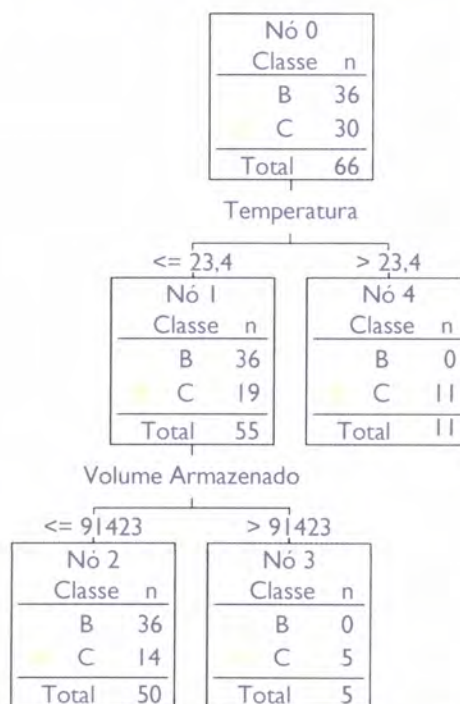


Figura VII.15 – Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo os parâmetros SST, OD, oxidabilidade e CQO.

Tabela VII.29 – Conjunto de regras relativo ao modelo apresentado na Figura VII.15.

Regras para a Classe B	Regras para a Classe C	
Regra 1 Se Temperatura ≤ 23,4 e Volume Armazenado ≤ 91423 Então → B (50; 0,72)	Regra 1 Se Temperatura ≤ 23,4 e Volume Armazenado > 91423 Então → C (5; 1,0)	Regra 2 Se Temperatura > 23,4 Então → C (11; 1,0)

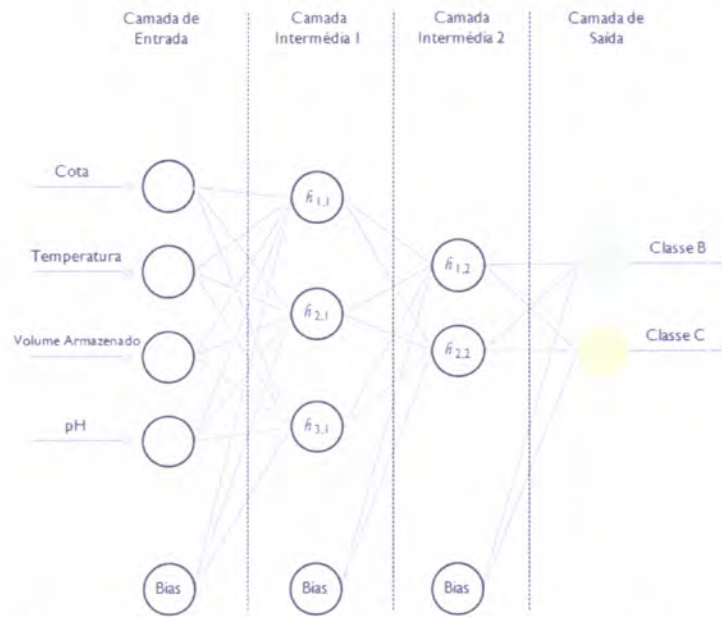


Figura VII.16 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas, excluindo os parâmetros SST, OD, oxidabilidade e CQO.

Tabela VII.30 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura VII.16.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$
Cota	4,07	3,83	3,79
Temperatura	19,11	-26,19	-9,21
Volume Arm.	-5,96	1,54	-3,49
pH	-3,61	-4,73	-4,52
Bias	-8,37	14,89	3,13

Tabela VII.31 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura VII.16.

	$h_{1,2}$	$h_{2,2}$
$h_{1,1}$	12,19	7,18
$h_{2,1}$	-4,78	-8,98
$h_{3,1}$	4,3	-4,14
Bias	1,08	-0,83

Tabela VII.32 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura VII.16.

	Classe B	Classe C
$h_{1,2}$	2,57	-2,57
$h_{2,2}$	-5,17	5,17
Bias	0,03	-0,03

ANEXO VIII

Neste anexo apresentam-se os modelos relativos à previsão da qualidade da água da albufeira de Odivelas, tendo em consideração a utilização de parâmetros meteorológicos e hidrométricos, obtidos com recurso a Árvores de Decisão e RNAs (Figura 31). Apresentam-se, ainda, o conjunto de regras geradas destacando-se, no final, o suporte e a confiança da regra (Árvores de Decisão) e os pesos associados às conexões entre os nodos das camadas adjacentes (RNAs). Os nodos das camadas intermédias são designados genericamente por h_{ij} , denotando o nodo i da camada intermédia j .

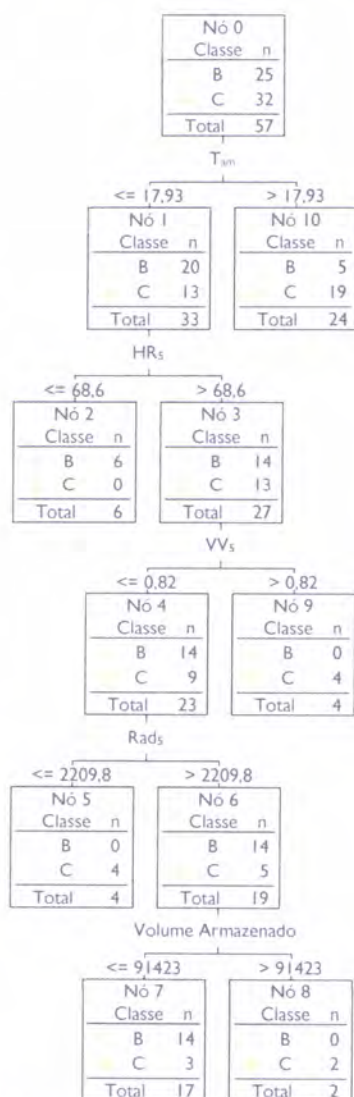


Figura VIII.1 - Árvore de decisão obtida para previsão da qualidade da água da albufeira de Odivelas, tendo em consideração a utilização de parâmetros meteorológicos e hidrométricos.

Tabela VIII.1 – Conjunto de regras relativo ao modelo apresentado na Figura VIII.1.

Regras para a Classe B	Regras para a Classe C	
<p>Regra 1</p> <p>Se $T_{am} \leq 17,93$ e $HR_5 \leq 68,6$</p> <p>Então → B (6; 1,0)</p>	<p>Regra 1</p> <p>Se $T_{am} \leq 17,93$ e $HR_5 > 68,6$ e $VV_5 \leq 0,82$ e $Rad_5 \leq 2209,8$</p> <p>Então → C (4; 1,0)</p>	<p>Regra 3</p> <p>Se $T_{am} \leq 17,93$ e $HR_5 > 68,6$ e $VV_5 > 0,82$</p> <p>Então → C (4; 1,0)</p>
<p>Regra 2</p> <p>Se $T_{am} \leq 17,93$ e $HR_5 > 68,6$ e $VV_5 \leq 0,82$ e $Rad_5 > 2209,8$ e Volume Armazenado ≤ 91423</p> <p>Então → B (17; 0,82)</p>	<p>Regra 2</p> <p>Se $T_{am} \leq 17,93$ e $HR_5 > 68,6$ e $VV_5 \leq 0,82$ e $Rad_5 > 2209,8$ e Volume Armazenado > 91423</p> <p>Então → C (2; 1,0)</p>	<p>Regra 4</p> <p>Se $T_{am} > 17,93$</p> <p>Então → C (24; 0,79)</p>



Figura VIII.2 – RNA obtida para previsão da qualidade da água da albufeira de Odivelas, tendo em consideração a utilização de parâmetros meteorológicos e hidrométricos.

Tabela VIII.2 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura VIII.2.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$
HR_{am}	-0,63	1,28	9,16	6,61
Rad_{10}	4,94	-12,66	1,87	8,09
Rad_5	11,82	-12,01	-10,65	9,03
T_{10}	-1,33	-1,53	16,14	5,32
Bias	7,42	1,44	-6,42	-12,45

Tabela VIII.3 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo na Figura VIII.2.

	$h_{1,2}$	$h_{2,2}$
$h_{1,1}$	-3,57	6,41
$h_{2,1}$	11,22	-6,78
$h_{3,1}$	0,32	-10,19
$h_{4,1}$	7,64	-6,84
Bias	-0,45	3,41

Tabela VIII.4 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura VIII.2.

	Classe B	Classe C
$h_{1,2}$	-3,15	3,15
$h_{2,2}$	4,73	-4,72
Bias	0,69	-0,69

ANEXO IX

Apresentam-se o conjunto de regras geradas para o modelo apresentado na Figura 33, destacando-se, no final, o suporte e a confiança da regra (Árvores de Decisão). Apresentam-se, ainda, os pesos associados às conexões entre os nodos das camadas adjacentes (RNAs), referente ao modelo apresentado na Figura 34. Os nodos das camadas intermédias são designados genericamente por h_{ij} , denotando o nodo i da camada intermédia j .

Tabela IX.1 – Conjunto de regras relativo ao modelo apresentado na Figura 33.

Regras para a Classe B		Regras para a Classe C	
Regra 1 Se $T_5 \leq 21,86$ e $pH \leq 7,8$ Então $\rightarrow B$ (2; 1,0)	Regra 4 Se $T_5 \leq 21,86$ e $pH > 7,8$ e $VV_{10} \leq 1,04$ e $VV_5 > 0,58$ e $HR_5 \leq 68,2$ e $HR_{10} > 70,2$ Então $\rightarrow B$ (2; 1,0)	Regra 1 Se $T_5 \leq 21,86$ e $pH > 7,8$ e $VV_{10} \leq 1,04$ e $VV_5 \leq 0,58$ e $PP_5 \leq 0,26$ e $PP_{am} > 0,287$ Então $\rightarrow C$ (6; 0,83)	Regra 4 Se $T_5 > 21,86$ Então $\rightarrow C$ (18; 0,89)
Regra 2 Se $T_5 \leq 21,86$ e $pH > 7,8$ e $VV_{10} \leq 1,04$ e $VV_5 \leq 0,58$ e $PP_5 \leq 0,26$ e $PP_{am} \leq 0,287$ Então $\rightarrow B$ (3; 1,0)	Regra 5 Se $T_5 \leq 21,86$ e $pH > 7,8$ e $VV_{10} > 1,04$ Então $\rightarrow B$ (3; 1,0)	Regra 2 Se $T_5 \leq 21,86$ e $pH > 7,8$ e $VV_{10} \leq 1,04$ e $VV_5 > 0,58$ e $HR_5 \leq 68,2$ e $HR_{10} \leq 70,2$ Então $\rightarrow C$ (5; 0,80)	
Regra 3 Se $T_5 \leq 21,86$ e $pH > 7,8$ e $VV_{10} \leq 1,04$ e $VV_5 \leq 0,58$ e $PP_5 > 0,26$ Então $\rightarrow B$ (5; 1,0)		Regra 3 Se $T_5 \leq 21,86$ e $pH > 7,8$ e $VV_{10} \leq 1,04$ e $VV_5 > 0,58$ e $HR_5 > 68,2$ Então $\rightarrow C$ (12; 0,83)	

af

Tabela IX.2 – Pesos associados às conexões entre a camada de entrada e a camada intermédia 1 para o modelo apresentado na Figura 34.

	$h_{1,1}$	$h_{2,1}$	$h_{3,1}$	$h_{4,1}$
HR_{am}	0,82	-21,37	7,06	-8,74
PP_{am}	-1,21	15,69	11,67	-13,05
Rad_{am}	2,11	-13,38	-17,52	-2,24
T_{20}	1,12	23,11	1,73	-6,37
VV_5	-0,11	5,09	-2,66	28,59
Volume Arm.	2,02	12,02	6,01	-14,45
pH	7,06	8,72	-9,23	5,56
Bias	2,45	-3,72	-1,5	9,28

Tabela IX.3 – Pesos associados às conexões entre a camada intermédia 1 e a camada intermédia 2 para o modelo apresentado na Figura 34.

	$h_{1,2}$	$h_{2,2}$
$h_{1,1}$	5,92	2,78
$h_{2,1}$	-10,72	4,91
$h_{3,1}$	-10,85	-0,53
$h_{4,1}$	-12,31	6,64
Bias	8,04	-2,02

Tabela IX.4 – Pesos associados às conexões entre a camada intermédia 2 e a camada de saída para o modelo apresentado na Figura 34.

	Classe B	Classe C
$h_{1,2}$	5,96	-5,77
$h_{2,2}$	-4,3	3,36
Bias	2,25	-1,39