

Utilização de Redes Neurais Artificiais na gestão de processos de Digestão Anaeróbia

Hélder Tiago Ferreira da Rocha

Mestrado em Engenharia Civil – Gestão da Construção

Orientador: Professor Jaime Gabriel Silva

JÚRI

Presidente: Professor Doutor Rui de Sousa Camposinhos (ISEP)

Professor Doutor António Jorge Monteiro (IST)

Professor Jaime Gabriel Silva (ISEP)

Engenheiro Milton Andrés Fontes (SIMRIA)

ISEP – Porto, 6 de Setembro de 2013

DEDICATÓRIA

À Joana, pelo apoio e dedicação.

E por nunca me deixar desistir.

AGRADECIMENTOS

Gostaria de expressar o meu agradecimento a todas as pessoas que, de algum modo, contribuíram para a realização deste trabalho:

Ao orientador, Professor Gabriel Silva, por me ter dado a oportunidade de realizar o estágio na SIMRIA, por todo o apoio prestado quer na compreensão do *software* utilizado como nas correções sugeridas no trabalho.

Ao co-orientador, Eng.º Milton Fontes, da SIMRIA, pelo apoio prestado durante o estágio e disponibilidade para esclarecer as dúvidas que lhe fui colocando.

Aos Engenheiros Nuno Brôco e José Gascão, da Direção de Engenharia das Águas de Portugal, pelo apoio prestado na interpretação dos resultados obtidos.

Aos restantes colaboradores da SIMRIA, em particular da ETAR Sul, pela forma cordial com que me receberam e pelo bom ambiente de trabalho criado.

Palavras-chave: Redes Neurais Artificiais, Digestão Anaeróbia, Biogás, Tratamento de Águas Residuais.

RESUMO

Numa Estação de Tratamento de Águas Residuais (ETAR), a otimização do processo de Digestão Anaeróbia (DA) é fundamental para o aumento da produção de biogás, que por sua vez é convertido em energia, essencial para a rentabilidade de exploração de ETAR. No entanto, a complexidade do processo de Digestão Anaeróbia das lamas constitui um obstáculo à sua otimização.

Com este trabalho pretende-se efetuar a análise e tratamento de dados de Digestão Anaeróbia, com recurso a Redes Neurais Artificiais (RNA), contribuindo, desta forma, para a compreensão do processo e do impacto de algumas variáveis na produção de biogás. As Redes Neurais Artificiais são modelos matemáticos computacionais inspirados no funcionamento do cérebro humano, com capacidade para entender relações complexas num determinado conjunto de dados, motivo por que se optou pela sua utilização na procura de soluções que permitem prever o comportamento de uma DA.

Para o desenvolvimento das RNA utilizou-se o programa NeuralToolsTM da PalisadeTM.

Como caso de estudo, a metodologia foi aplicada ao Digestor A da ETAR Sul da SIMRIA, empresa onde teve lugar o estágio curricular que originou o presente trabalho. Nesse contexto, utilizaram-se dados com informação referente aos últimos dois anos de funcionamento do digestor, disponíveis na empresa.

Apesar de se terem verificado certas limitações, na predição em alguns casos particulares, de um modo geral, considera-se que os resultados obtidos permitiram concluir que as redes neuronais modeladas apresentam boa capacidade de generalização na imitação do processo anaeróbio.

Conclui-se, portanto, que o estudo realizado pode constituir um contributo com interesse para a otimização da produção do biogás na DA de ETAR Sul da SIMRIA e que a utilização de RNA poderá ser uma ferramenta a explorar, quer nessa área, quer noutras áreas de gestão de sistemas de saneamento básico.

Keywords: Artificial Neural Networks, Anaerobic Digestion, Biogas, Methane, Wastewaters Treatment.

ABSTRACT

In a wastewater treatment plant (WWTP), the optimization of Anaerobic Digestion (AD) process is fundamental to increase the production of biogas, which in turn is converted into energy, essential to the profitability of the WWTP operation. However, the complexity of sludge anaerobic digestion process is an obstacle to its optimization.

This work tries to carry out the analysis and treatment of Anaerobic Digestion data, using Artificial Neural Networks (ANN), thus contributing to understand the process and the impact of some variables in the biogas production.

ANN are computational mathematical models inspired by the human brain, with the ability to understand complex relationships in a given data set. Therefore, ANN were selected in order to develop a computation solutions allowing to predict the behavior of an AD.

In this work, the program NeuralTools™ from Palisade™ was used to develop de ANN oriented to predict AD outputs.

As a case study, the methodology has been applied to SIMRIA's digester A, at WWTP Sul. SIMRIA is the company where the traineeship that originated this work took place.

In this work, enterprise's data has been used to apply ANN to the AD process, namely information concerning the last two years of the digester operation.

Although some limitations in predicting a few particular cases, in general, the results showed that the modeled neural networks exhibit good generalization capability, in the simulation of the anaerobic process.

Finally, this study can be seen as a contribution to the optimization of biogas production, in AD, and an interesting contribution to biogas production optimization, at SIMRIA's AD in WWTP Sul, as well as concluding that the use of ANN might be a tool to explore, whether in this area or other areas of the management of sanitation systems.

ÍNDICE

| | |
|--|----|
| 1. INTRODUÇÃO | 1 |
| 1.1. Enquadramento Geral..... | 1 |
| 1.2. Breve referência ao processo de tratamento de ETAR..... | 5 |
| 1.4. Descrição Sumária do Estudo | 7 |
| 2. BREVE CARACTERIZAÇÃO DO SECTOR DE SERVIÇO DE ÁGUAS E APRESENTAÇÃO DA EMPRESA..... | 11 |
| 2.1. Sector de Serviço de Águas | 11 |
| 2.1.1. Principais etapas dos serviços de abastecimento de águas e saneamento de águas residuais..... | 12 |
| 2.2. O grupo ADP – Águas de Portugal..... | 14 |
| 2.3. A SIMRIA – Saneamento Integrado dos Municípios da Ria, SA..... | 16 |
| 2.3.1. Apresentação | 16 |
| 2.3.2. Estrutura Organizacional..... | 17 |
| 2.3.3. Relações com <i>Stakeholders</i> | 19 |
| 2.3.4. Principais indicadores da atividade da SIMRIA no triénio 2009/2011..... | 21 |
| 2.3.5. Caracterização dos ecossistemas existentes na área de atuação da SIMRIA | 22 |
| 2.3.6. O Sistema Multimunicipal de Saneamento | 26 |
| 2.3.7. Estação de Tratamento de Águas Residuais Sul (ETAR Sul)..... | 35 |
| 3. PROCESSO DE DIGESTÃO ANAERÓBIA..... | 41 |
| 3.1. Fases do Processo Anaeróbio | 41 |
| 3.2. Fatores que interferem no processo de DA..... | 47 |
| 3.3. Tipos de Reatores Anaeróbios | 54 |
| 3.4. A Digestão Anaeróbia da ETAR Sul da SIMRIA | 60 |
| 4. REDES NEURONAIS ARTIFICIAIS..... | 61 |
| 4.1. Analogia com o modelo biológico | 63 |
| 4.2. Evolução histórica..... | 67 |

| | | |
|-------------|---|------------|
| 4.3. | Princípios de funcionamento e descrição das RNA..... | 69 |
| 4.3.1. | Perceptrão..... | 71 |
| 4.3.2. | Perceptrão Multicamada (PMC)..... | 73 |
| 4.4. | Tipos de Redes Neurais Artificiais..... | 76 |
| 4.4.1. | Redes MLFNN e PNN/GRNN..... | 77 |
| 4.4.2. | Comparação entre redes MLFNN e PNN/GRNN..... | 79 |
| 4.5. | Fases de desenvolvimento de uma RNA..... | 80 |
| 4.5.1. | Preparação dos dados..... | 80 |
| 4.5.2. | Treino da RNA..... | 81 |
| 4.5.3. | Testar a RNA..... | 82 |
| 4.5.4. | Previsão..... | 82 |
| 4.6. | Comparação entre Redes Neurais e Métodos Estatísticos..... | 83 |
| 4.7. | Justificação da aplicação de redes neurais..... | 84 |
| 5. | APRESENTAÇÃO DO CASO DE ESTUDO..... | 87 |
| 5.1. | Enquadramento do estudo..... | 87 |
| 5.2. | Metodologia Proposta..... | 88 |
| 5.2.1. | Considerações Iniciais..... | 88 |
| 5.2.2. | Apresentação do NeuralTools..... | 89 |
| 5.2.3. | Preparação dos dados..... | 90 |
| 5.2.4. | Treino e teste da Rede Neuronal..... | 95 |
| 5.2.5. | Previsão..... | 102 |
| 5.3. | Desenvolvimento do estudo..... | 103 |
| 5.3.1. | Introdução..... | 103 |
| 5.3.2. | Preparação dos dados..... | 103 |
| 5.3.3. | Treino e Teste..... | 108 |
| 5.3.4. | Previsão e Validação..... | 114 |
| 6. | CONSIDERAÇÕES FINAIS..... | 119 |

| | |
|---|-----|
| 6.1. Conclusões..... | 119 |
| 6.2. Sugestões para desenvolvimentos futuros..... | 122 |
| 7. REFERÊNCIAS BIBLIOGRÁFICAS..... | 123 |

ÍNDICE DE FIGURAS

| | |
|---|----|
| Figura 1 – Cadeia de valor do sector de serviços de águas. (Fonte: ERSAR, 2010)..... | 13 |
| Figura 2 – Presença da ADP em Portugal. (Fonte: ADP, 2013) | 15 |
| Figura 3 – Estrutura Acionista da SIMRIA. (Fonte: SIMRIA, 2011a) | 17 |
| Figura 4 – Estrutura Organizacional da SIMRIA. (Fonte: SIMRIA, 2011a) | 18 |
| Figura 5 – Organigrama dos <i>Stakeholders</i> da SIMRIA. (Fonte: SIMRIA, 2011a) | 19 |
| Figura 6 – Vista aérea da Ria de Aveiro na zona da Barra de Aveiro..... | 22 |
| Figura 7 – Vista aérea da Ria de Aveiro próximo do centro da cidade de Aveiro..... | 23 |
| Figura 8 – Imagens de moliceiros. (Fontes: www.net-moliceiro.inovanet.pt e www.geolocation.ws) | 23 |
| Figura 9 – Imagens aéreas da BE/LP. (Fonte: www.prof2000.pt/users/eb23fespan)..... | 24 |
| Figura 10 – Imagens elucidativas do estado da BE/LP. | 25 |
| Figura 11 – Mapa da área de abrangência do Sistema Multimunicipal da SIMRIA..... | 26 |
| Figura 12 – 1ª Fase do Sistema Multimunicipal. (Fonte: SIMRIA, 2013)..... | 28 |
| Figura 13 – 2ª Fase do Sistema Multimunicipal. (Fonte: SIMRIA, 2013)..... | 30 |
| Figura 14 – Imagem da ETAR de Espinho. (Fonte: SIMRIA, 2013)..... | 30 |
| Figura 15 – Cadeia de valor da atividade chave da empresa. (Fonte: SIMRIA, 2011a)..... | 33 |
| Figura 16 – ETAR Sul da SIMRIA. (Fonte: SIMRIA, 2013)..... | 35 |
| Figura 17 – Diagrama Funcional da ETAR. (Fonte: SIMRIA, 2013) | 38 |
| Figura 18 – Representação esquemática das fases do processo de DA. (Fonte: Silveira, 2009) | 42 |
| Figura 19 - Produtos da Acidogénese mediante a pressão parcial de H ₂ . (Fonte: Silveira, 2009) | 44 |
| Figura 20 – Percentagem de NH ₃ e NH ₄ ⁺ em função do pH. (Fonte: Gerardi, 2003)..... | 51 |
| Figura 21 – Representação esquemática dos diferentes tipos de digestores. (Fonte: Silveira, 2009) | 55 |
| Figura 22 – Representação de um neurónio biológico..... | 63 |
| Figura 23 – Representação esquemática de uma sinapse biológica | 64 |
| Figura 24 – Representação simplificada de uma sinapse | 64 |
| Figura 25 – Rede Neuronal Biológica. (Fonte: www.auladeanatomia.com , 2013) | 65 |
| Figura 26 – Representação de duas topologias de redes | 66 |
| Figura 27 – Modelo de um nó de uma RNA. (Fonte: Matos, 2008)..... | 70 |
| Figura 28 – Diferentes representações de redes multicamada..... | 71 |

| | |
|---|-----|
| Figura 29 - Representação esquemática de um perceptrão. (Fonte: Lingireddy e Brion, 2005) | 72 |
| Figura 30 – a) Classes linearmente separáveis; b) Classes não linearmente separáveis. | 72 |
| Figura 31 – Esquema de rede MLFN ou PMC de 3 camadas. (Fonte: Silva, 2013)..... | 73 |
| Figura 32 – Exemplo de uma base de dados para utilizar com o Neural Tools..... | 91 |
| Figura 33 – Caixa de diálogo do Gestor do conjunto de dados..... | 92 |
| Figura 34 – Definição do tipo de variável..... | 93 |
| Figura 35 – Layout do conjunto de dados | 94 |
| Figura 36 – Caixa de diálogo Treinar..... | 96 |
| Figura 37 – Caixa de diálogo Configuração da rede..... | 96 |
| Figura 38 – Caixa de diálogo tempo de execução..... | 97 |
| Figura 39 – Caixa de diálogo Visão Geral de Treino | 98 |
| Figura 40 – Caixa de diálogo “Andamento do Treino” | 99 |
| Figura 41 – Tabelas geradas no relatório do Neural Tools..... | 100 |
| Figura 42 – Exemplos de gráficos gerados no relatório do Neural Tools..... | 101 |

ÍNDICE DE TABELAS

| | |
|---|-----|
| Tabela 1 – Indicadores da atividade da SIMRIA. (Fonte: SIMRIA, 2011a) | 21 |
| Tabela 2 – Zonas de Intervenção e População Equivalente Abrangida pelo sistema..... | 27 |
| Tabela 3 – Valores de Investimento/Financiamento. (Fonte: SIMRIA, 2011b) | 27 |
| Tabela 4 – Descrição do Sistema de Recolha - 1ª Fase. (Fonte: SIMRIA, 2011b) | 29 |
| Tabela 5 – ETARs responsáveis pela componente de tratamento - 1ª Fase. (Fonte: SIMRIA, 2011b)..... | 29 |
| Tabela 6 – Descrição do sistema de rejeição - 1ª Fase. (Fonte: SIMRIA, 2011b)..... | 29 |
| Tabela 7 – Descrição do Sistema de Recolha - 2ª Fase. (Fonte: SIMRIA, 2011b)..... | 31 |
| Tabela 8 – ETARs responsáveis pela componente de tratamento - 2ª Fase. (Fonte: SIMRIA, 2011b)..... | 31 |
| Tabela 9 - Descrição do sistema de rejeição - 2ª Fase. (Fonte: SIMRIA, 2011b) | 31 |
| Tabela 10 – Descrição da expansão da componente de recolha - 3ª fase. (Fonte: SIMRIA, 2011b)..... | 32 |
| Tabela 11 – Infraestruturas ao serviço da SIMRIA..... | 32 |
| Tabela 12 – Valores Limite de Emissão (VLE). (Fonte: SIMRIA, 2011b) | 34 |
| Tabela 13 – Rejeição de efluente (Valores em milhares de m3). (Fonte: SIMRIA, 2011b) | 34 |
| Tabela 14 – Dados de Projeto. (Fonte: SIMRIA, 2013) | 35 |
| Tabela 15 – Caracterização geral de Lamas geradas no tratamento de águas residuais. (Fonte: Di Bernardino, 2001)..... | 40 |
| Tabela 16 – Regimes de temperatura do processo anaeróbio. (Fonte: Metcalf e Eddy, 1991) | 48 |
| Tabela 17 – Diferentes Tipos de Reatores. (Fonte: Mara e Horan, 2003) | 55 |
| Tabela 18 – Variáveis definidas para o treino das RNA..... | 105 |

ÍNDICE DE FÓRMULAS

| | |
|--|-----|
| Fórmula 1 – Tempo de Retenção Hidráulica (TRH)..... | 48 |
| Fórmula 2 – Fórmula de cálculo de um valor de output da rede PMC. (Fonte: Lingireddy e Brion, 2005)..... | 74 |
| Fórmula 3 – Função Sigmoidal. (Fonte: Lingireddy e Brion, 2005) | 74 |
| Fórmula 4 – Eficiência da Digestão | 106 |
| Fórmula 5 – Carga Mássica Volúmica..... | 107 |

LISTA DE ABREVIATURAS

AdP – Águas de Portugal

AGV – Ácidos Gordos Voláteis

BF – Bactérias Fermentativas

BH – Bactérias Hidrolíticas

BHA – Bactérias que realizam a inter-conversão entre o Acetato e o Hidrogénio.

BM – Bactérias Metanogénicas

BMA – Bactérias Metanogénicas que convertem o Acetato.

BMH – Bactérias Metanogénicas que convertem o Hidrogénio.

BRS – Bactérias Redutoras de Sulfatos

CMV – Carga Mássica Volúmica

CQO – Carência Química de Oxigénio

DA – Digestão Anaeróbia

ED – Eficiência da Digestão

EMQ – Erro Médio Quadrático

ERSAR – Entidade Reguladora dos Serviços de Água e Resíduos

ETAR – Estação de Tratamento de Águas Residuais

GRNN – Generalized Regression Neural Network

MLFNN – MultiLayer Feed-forward Neural Network

MS – Matéria Sólida

MV – Matéria Volátil

PEAASAR – Plano Estratégico de Abastecimento de Água e Saneamento de Águas Residuais

PMC – Perceptrão Multicamadas

PNN – Probabilistic Neural Network

RNA – Redes Neuronais Artificiais

SIMRIA – Saneamento Integrado dos Municípios da Ria

TRH – Tempo de Retenção Hidráulica

TRS – Tempo de Retenção de Sólidos

1. INTRODUÇÃO

1.1. Enquadramento Geral

Silveira (2009) refere que a crescente consciencialização das sociedades modernas sobre a importância de se proceder ao tratamento das águas residuais de modo a prevenir a poluição ambiental, assegurar a sustentabilidade do ambiente e evitar as consequências que os resíduos poderão ter na saúde pública, tem resultado no aumento significativo do número de Estações de Tratamento de Águas Residuais (ETARs) e no desenvolvimento de esforços com o objetivo de aumentar a eficiência dos processos de tratamento. De acordo com a ERSAR (2010), para esse efeito, foram implementados na última década, em Portugal, dois planos estratégicos (PEAASAR I e II). O PEAASARII, que se encontra em vigor, define três grandes objetivos estratégicos, traduzidos em objetivos operacionais e respetivas medidas a desenvolver no período de 2007 a 2013, designadamente:

- A universalidade, continuidade e qualidade de serviço;
- A sustentabilidade do sector, implicando a melhoria da produtividade e da eficiência;
- A proteção dos valores ambientais.

No âmbito da universalidade, continuidade e qualidade de serviço foram definidos os seguintes objetivos operacionais:

- Servir cerca de 95% da população do País com sistemas públicos de abastecimento de água e servir cerca de 90% da população total do País com sistemas públicos de saneamento de águas residuais urbanas, sendo que em cada sistema integrado o nível de atendimento desejável deve ser de, pelo menos, 70% da população abrangida;
- Obter níveis adequados de qualidade de serviço, mensuráveis pela conformidade dos indicadores de qualidade do serviço definidos pela entidade reguladora;
- Estabelecer orientações para a definição de tarifas ao consumidor final, evoluindo

tendencialmente para um intervalo razoável e compatível com a capacidade económica das populações.

Em relação à sustentabilidade do sector, os objetivos estabelecidos foram:

- Garantir a recuperação integral dos custos incorridos com os serviços;
- Otimizar a gestão operacional e eliminar custos de ineficiência;
- Contribuir para a dinamização do tecido empresarial privado nacional e local.

Os objetivos relativos à proteção dos valores ambientais são os seguintes:

- Cumprir os objetivos decorrentes do normativo nacional e comunitário;
- Garantir uma abordagem integrada na prevenção e no controlo da poluição provocada pela atividade humana e pelos sectores produtivos
- Aumentar a produtividade e a competitividade do sector através de soluções que promovam a ecoeficiência.

Por outro lado, com a conjuntura económica atual, a par da importância dada às preocupações ambientais e à preservação dos recursos naturais, as empresas deste sector preocupam-se, cada vez mais, em garantir o seu desenvolvimento sustentável, procurando reduzir os custos de exploração através da otimização de processos de tratamento.

Conforme indicado por Gonçalves (2012), desde a primeira grande crise energética na segunda metade do século XX, com o aumento do preço dos combustíveis fósseis e com o aumento das preocupações ambientais, a União Europeia tem vindo a apostar na utilização de combustíveis renováveis de baixo custo, estando entre eles o biogás.

Por este motivo, as energias renováveis têm estado, nos últimos anos, em franca expansão devido à urgência de implementação de soluções alternativas que minimizem a dependência dos países relativamente aos combustíveis fósseis e à urgência em combater o fenómeno do aquecimento global, reduzindo as emissões de gases com efeito de estufa.

Atualmente existe uma preocupação associada à adequabilidade das fontes de combustível, do custo da energia e dos níveis de tratamento crescentes que são exigidos, dos quais resultam consumos energéticos superiores. Consequentemente, o dimensionamento e a operação das ETARs modernas estão, cada vez mais, focalizados na maximização da eficiência do uso da energia elétrica e na redução dos custos de tratamento (Metcalf e Eddy, 1991), pelo menos nos casos em que a dimensão da ETAR torne o investimento inicial recuperável ao longo do tempo.

Como se depreende de Silveira (2009) e Gonçalves (2012), o biogás (biometano) é um gás produzido naturalmente quando a matéria orgânica é decomposta em meio húmido e na ausência de oxigénio, meio anaeróbio, por organismos metabolicamente ativos. O biogás produzido durante o processo de Digestão Anaeróbia (DA) é uma fonte de energia renovável com um potencial significativo. A principal vantagem da produção de biogás, sobre outro tipo de energias renováveis existentes no mercado, relaciona-se com o facto de conter um poder calorífico semelhante ao gás natural a um menor custo, por ser produzido através da matéria biodegradável considerada como resíduo, contribuindo desta forma para a redução da emissão de gases com efeito de estufa para a atmosfera.

O biogás pode ter diversas aplicações, tais como, a combustão direta e a produção de calor, geração de eletricidade, integração na rede de gás natural, utilização como combustível para veículos e, em última análise, em pilhas de combustível. Deste modo, a biomassa apresenta-se como uma fonte de energia renovável bastante viável devido à sua enorme disponibilidade no ambiente terrestre, apresentando um papel crucial na redução da concentração do CO₂ atmosférico.

A energia libertada durante a combustão do gás metano pode substituir a das principais fontes de energia atuais associadas à utilização de combustíveis fósseis, caracterizadas por

um ciclo de carbono com uma escala temporal de milhares de anos. Sendo o metano derivado de biomassa, com um ciclo de carbono com uma escala temporal bastante mais reduzida comparativamente aos combustíveis fósseis, permite que o carbono emitido seja rapidamente fixado com a produção de nova biomassa, acabando por não ficar concentrado na atmosfera sob a forma de CO₂, designando-se, por esta razão, por ciclo neutro de carbono.

Silveira (2009) refere, ainda, que a DA é descrita muitas vezes como um processo instável e a complexidade do processo industrial de produção de biogás constitui um obstáculo à sua otimização. No entanto, os avanços verificados nos últimos 20 anos na compreensão da bioquímica e energética do processo contribuíram para o desenvolvimento do conhecimento das fases mais sensíveis do processo, assim como de estratégias que permitiram aumentar a estabilidade dos reatores anaeróbios.

Apesar dos avanços verificados, o interior de um digestor continua a ser uma caixa negra da qual não se consegue obter informação que permita a monitorização e controlo do processo de digestão. Esta continua a ser a principal dificuldade com que se deparam as entidades gestoras de ETARs. A aplicação de modelos matemáticos como as Redes Neurais Artificiais (RNA) a esta problemática podem prestar um forte contributo para a compreensão do processo de digestão e otimização da produção de biogás.

De acordo com a Palisade (2010), as redes neuronais têm a capacidade de entender relações complexas existentes entre os dados que de outra forma não seria possível perceber. Por imitação das funções do cérebro, conseguem discernir padrões existentes nos dados e, quando são fornecidos dados novos, são capazes de gerar previsões com base nas relações e nos padrões detetados nos dados utilizados no treino.

Embora seja uma ferramenta matemática pouco divulgada, nos últimos anos o recurso a

RNA tem vindo a aumentar, com aplicações de sucesso em diferentes áreas, como por exemplo, na engenharia e ambiente, na modelação de sistemas de abastecimento de água, previsões de precipitação diária, modelação de pH e previsão de salinidade em rios, ou no setor económico-financeiro, na análise e pesquisa de mercado, identificação de fraude com cartões de crédito e avaliação de créditos, entre outras.

1.2. Breve referência ao processo de tratamento de ETAR

O principal objetivo do tratamento dos esgotos domésticos está na correção das suas características de modo a que o efluente tratado possa sofrer uma deposição final de acordo com as normas estabelecidas, de modo a evitar impactes adversos no ecossistema do meio recetor (Van Haandel e Lettinga, 1994).

Durante o processo de tratamento são retiradas lamas que permitem a produção de biogás através de um processo de decomposição, a digestão anaeróbia. O biogás produzido representa uma fonte de energia renovável de grande importância ecológica, podendo ser transformado em energia térmica e/ou elétrica.

Após a digestão, as lamas podem ter como destino final a incineração ou a aplicação em solo agrícola. A reutilização da lama estabilizada como fertilizante agrícola é a opção mais atrativa a nível ambiental, uma vez que possibilita a reciclagem de nutrientes e matéria orgânica.

Apesar dos avanços alcançados nos últimos anos na compreensão e conhecimento do processo de DA terem contribuído para o desenvolvimento de estratégias que permitiram aumentar a estabilidade dos reatores anaeróbios, subsistem ainda muitas dúvidas em relação ao funcionamento dos digestores e sobre qual o contributo das diversas variáveis presentes na produção de biogás. É aqui que ferramentas poderosas como as RNA podem contribuir de forma decisiva para a compreensão e otimização do processo.

1.3. Objetivos

Assim, o principal objetivo na realização deste estudo consiste na aplicação de RNA ao complexo processo de DA, de forma a desenvolver um modelo que permita otimizar a produção de biogás, tirando partido da potencialidade de predição das referidas redes. Para tal, o caminho a seguir no desenvolvimento deste estudo compreende etapas cujos objetivos intermédios podem ser sintetizados nos seguintes aspetos principais:

- A análise da adequação da aplicação de redes neuronais artificiais à modelação do processo de digestão anaeróbia.
- O estudo e perceção das variáveis com maior impacto no funcionamento da digestão anaeróbia, estabelecendo uma hierarquização entre elas.
- Elaboração de um modelo de simulação da produção de biogás, orientado para apoiar as decisões da entidade gestora no processo de digestão anaeróbia.

No que respeita ao primeiro aspeto, pretende-se fornecer, neste documento, uma base para a compreensão e utilização de RNA discutindo-se a diversidade, as propriedades e as aplicações destes modelos e analisando as mais-valias e possibilidades que a sua aplicação poderá trazer em termos da previsão e otimização da produção de biogás.

O segundo aspeto referido tem aplicação mais orientada para o planeamento. Devido à carência de dados, à pouca fiabilidade dos mesmos ou, simplesmente, à carência de ferramentas adequadas ao seu tratamento, desconhece-se, em profundidade, qual a relação e o impacto das variáveis no processo de produção do biogás. Sabe-se, por exemplo, que o controlo da temperatura dentro do digestor é crucial para o bom funcionamento do processo, devendo permanecer constante e, em regime mesófilo, a temperatura de funcionamento ótima encontra-se entre os 30 e os 35°C. Por outro lado a quantidade de matéria orgânica e a concentração de nutrientes que entram no digestor são fundamentais para assegurar o crescimento microbiano. No entanto, devido à quantidade de variáveis

intervenientes, a análise estatística tradicional resulta num grau de complexidade excessivo, sendo ainda que, por vezes, é necessário definir o próprio comportamento do sistema antes do estudo, imputando-se, desde logo, possíveis limitações ao modelo.

Em relação ao último aspeto, e na concretização do principal objetivo do estudo, espera-se conseguir estabelecer gamas de funcionamento ótimo das variáveis definidas como mais importantes no processo de digestão e testar a sua eficácia no funcionamento do digestor, comparando os resultados obtidos pela RNA com a realidade.

A otimização do processo de DA e conseqüente aumento da produção de biogás visa sobretudo aumentar a rentabilidade de exploração da ETAR assegurando a sua sustentabilidade.

Refira-se que o presente trabalho se desenvolveu, no âmbito de um estágio curricular, do Mestrado em Engenharia Civil do Instituto Superior de Engenharia do Porto (ISEP), realizado na empresa de Saneamento Integrado dos Municípios da Ria (SIMRIA).

1.4. Descrição Sumária do Estudo

A SIMRIA assume internamente uma forte preocupação com a otimização do funcionamento da etapa de Digestão Anaeróbia da ETAR Sul e a gestão do biogás gerado, tendo vindo a solicitar assessoria técnica à Direção de Engenharia do grupo Águas de Portugal com a duração de aproximadamente um ano, que decorreu entre Agosto de 2010 e Junho de 2011. Como resultado desse estudo foi identificado um conjunto de medidas, já implementadas, que permitiram uma melhoria da operacionalidade global da linha de lamas. (Gascão e Carvalho, 2011)

No seguimento desta melhoria, considerou-se pertinente realizar o presente estudo da gestão de processo e otimização da produção de biogás com recurso a Redes Neurais

Artificiais. As RNA são um tipo de ferramentas computacionais que operam de forma análoga ao processo biológico de um cérebro que, quando devidamente treinadas, são capazes de entender relações complexas em determinados conjuntos de dados.

A empresa (SIMRIA) dispõe de um *software* (NAVIA) que permite o armazenamento dos dados relacionados com o processo de operação da ETAR recolhidos ao longo do tempo, criando, assim, uma considerável base de dados disponíveis para o treino das redes. A quantidade e qualidade dos dados disponível é um fator muito importante para o sucesso do treino de RNA, sendo que quanto maior for a base de dados melhor será o treino e a capacidade de previsão da rede.

O objeto deste estudo é a etapa de DA, em particular o digestor A, que é parte integrante da linha de lamas da ETAR Sul. O digestor A, após paragem para uma intervenção de reabilitação ao nível da construção civil (sobretudo betão), tem trabalhado em contínuo desde Abril de 2011. O reinício de funcionamento deste digestor coincidiu com a substituição do caudalímetro do biogás, passando de um caudalímetro mecânico para um caudalímetro termomássico, com garantia de maior fiabilidade na leitura dos resultados obtidos. Por este motivo, para o desenvolvimento deste trabalho, os dados recolhidos, para treino e teste das RNA, estão compreendidos entre Abril 2011 e Fevereiro de 2013. Para validação dos resultados foram usados os dados referentes aos meses de Março e Abril de 2013.

A primeira etapa do estudo consistiu no tratamento e manipulação dos dados recolhidos de forma a construir uma base de dados e obter alguns indicadores importantes para o treino das RNA. Em seguida, tentou-se efetuar um saneamento da amostra, composta pela base de dados, através da análise robusta dos dados pelo método de “caixas de bigodes”, sem que, no entanto, os resultados obtidos permitissem excluir um número significativo de casos.

Com uma base de dados definida, analisou-se a adequação da aplicação de RNA à modelação do processo de DA iniciando-se o treino de vários modelos de RNA com diferentes variáveis de entrada de forma a encontrar o modelo que melhor simula o comportamento da digestão anaeróbica verificada no digestor A.

Para o desenvolvimento deste trabalho utilizaram-se os programas Microsoft ExcelTM e Neural ToolsTM do The Decision Tools SuiteTM da PalisadeTM.

O presente trabalho foi, assim, desenvolvido sob a orientação de Silva, na qualidade de docente do Instituto Superior de Engenharia do Porto (ISEP), com acompanhamento do Eng.º Milton Fontes da SIMRIA – responsável pela exploração e manutenção da ETAR Sul da SIMRIA. Além disso, beneficiou ainda do apoio da Direção de Engenharia da AdP, na pessoa do Eng.º Nuno Broco (Diretor) e do Eng.º José Gascão.

A validação dos resultados obtidos implicou uma análise cuidada dos mesmos, contando para isso com o apoio da Direção de Engenharia (DE) da AdP, para averiguar a fiabilidade dos mesmos.

No capítulo 2 efetua-se uma breve descrição do sector de abastecimento de água e tratamento de águas residuais, assim como, do grupo AdP, da SIMRIA e dos ecossistemas protegidos pelo sistema multimunicipal de saneamento da Ria de Aveiro. Procede-se, ainda, à caracterização da ETAR Sul, dos seus efluentes e das fases e processos de tratamento da ETAR.

No capítulo 3 aborda-se, de forma resumida, o processo de degradação anaeróbio, descrevendo as principais fases interdependentes em que pode ser dividido e os fatores ambientais que provocam maior interferência no processo. Neste capítulo também são apresentados os principais tipos de digestores anaeróbios.

O estudo das RNA é desenvolvido no capítulo 4, no qual, após uma breve apresentação

das redes, é estabelecida a sua analogia com sistemas biológicos e procede-se a uma contextualização histórica. De seguida, são descritos os princípios de funcionamento das redes neuronais. Ainda, neste capítulo, são apresentadas as redes neuronais mais comuns e as que são usadas pelo programa NeuralTools, que foi usado no desenvolvimento deste estudo. Serão ainda destacadas diversas questões importantes relacionadas com a modelação das RNA, incluindo a divisão dos dados em conjuntos de treino, teste e validação, os diferentes algoritmos de treino e métodos para validação das RNA treinadas. Por último, efetua-se uma comparação entre as redes neuronais e os métodos estatísticos procedida de uma justificação do método escolhido.

O capítulo 5 será dedicado ao desenvolvimento do estudo, com a descrição do caso de estudo e das técnicas e metodologias utilizadas, e à apresentação dos resultados obtidos.

Finalmente, no capítulo 6 são apresentadas as conclusões do estudo e desenvolvidas recomendações/sugestões para eventuais trabalhos futuros.

2. BREVE CARACTERIZAÇÃO DO SECTOR DE SERVIÇO DE ÁGUAS E APRESENTAÇÃO DA EMPRESA

2.1. Sector de Serviço de Águas

Segundo a ERSAR (2010), o sector de serviços de águas, conforme é usualmente denominado, compreende as atividades de abastecimento de água às populações, urbanas e rurais, e as atividades associadas, como os serviços, o comércio e a pequena indústria inserida na malha urbana. Compreende também a drenagem e o tratamento de águas residuais urbanas, que incluem as águas residuais de origem doméstica, industrial e pluvial. Estas duas atividades têm, tradicionalmente, sido tratadas em conjunto, verificando-se, todavia, que a atividade de abastecimento de água apresenta níveis de atendimento muito superiores aos registados na atividade de saneamento de águas residuais.

O sistema de abastecimento de água é um sistema em alta se é constituído por um conjunto de componentes a montante da rede de distribuição de água, fazendo a ligação do meio hídrico ao sistema em baixa. O sistema de abastecimento de água é um sistema em baixa se é constituído por um conjunto de componentes que ligam o sistema em alta ao utilizador final. O sistema de abastecimento de água presta um serviço em alta e em baixa sempre que vincula o meio hídrico a um utilizador final.

O sistema de drenagem de águas residuais em baixa é um sistema que faz a coleta de águas residuais junto ao produtor rejeitando-as num sistema em alta. O sistema de drenagem de águas residuais em alta é constituído por um conjunto de componentes que permitem a ligação do sistema em baixa ao ponto de rejeição. Se o sistema de drenagem de águas residuais faz afluir as águas residuais urbanas desde o produtor até ao meio recetor, então considera-se que é um sistema em alta e em baixa.

Se para a atividade de abastecimento de água se pode afirmar que o país está já dotado de uma rede que permite abastecer a quase totalidade da população, para a atividade de saneamento de águas residuais a realidade é bem distinta. Existem ainda muitos aglomerados sem drenagem de águas residuais e, para as zonas onde existe drenagem, verificam-se frequentes situações de ausência de tratamento adequado. Isto significa que é expectável que o crescimento que este sector tem vindo a registar se mantenha nos próximos anos, até que toda a rede de infraestruturas em alta e baixa esteja completa.

O nível de cobertura da população com serviços de abastecimento público de água é atualmente, de cerca de 93%, o que corresponde a uma melhoria significativa verificada na última década, em que o valor era de cerca de 80%. Este nível de cobertura encontra-se relativamente próximo da meta de 95% prevista no PEAASAR II.

O nível de cobertura da população com serviços de saneamento de águas residuais urbanas com adequado tratamento é atualmente de cerca de 76%, o que corresponde a uma melhoria significativa quando comparada com o que se passava há cerca de uma década, em que o valor era de 30%. Este nível de cobertura encontra-se ainda distante da meta de 90% prevista no plano estratégico para o sector.

2.1.1. Principais etapas dos serviços de abastecimento de águas e saneamento de águas residuais

As atividades deste sector compreendem um conjunto de etapas distintas e com requisitos muito diferentes para os dois subsectores.

Para o subsector das águas é necessário dispor de origens em quantidade e qualidade suficientes para cobrir as necessidades e de infraestruturas que assegurem a fiabilidade no abastecimento a toda a população e às várias atividades económicas.

Para o subsector das águas residuais é necessário recolher as rejeições provenientes das

fontes tóxicas de origem urbana, assegurar um tratamento adequado e descarregar as águas tratadas em condições que não adulterem a qualidade da água dos meios recetores.

Em termos genéricos pode-se estabelecer a seguinte cadeia de valor para os dois subsectores:

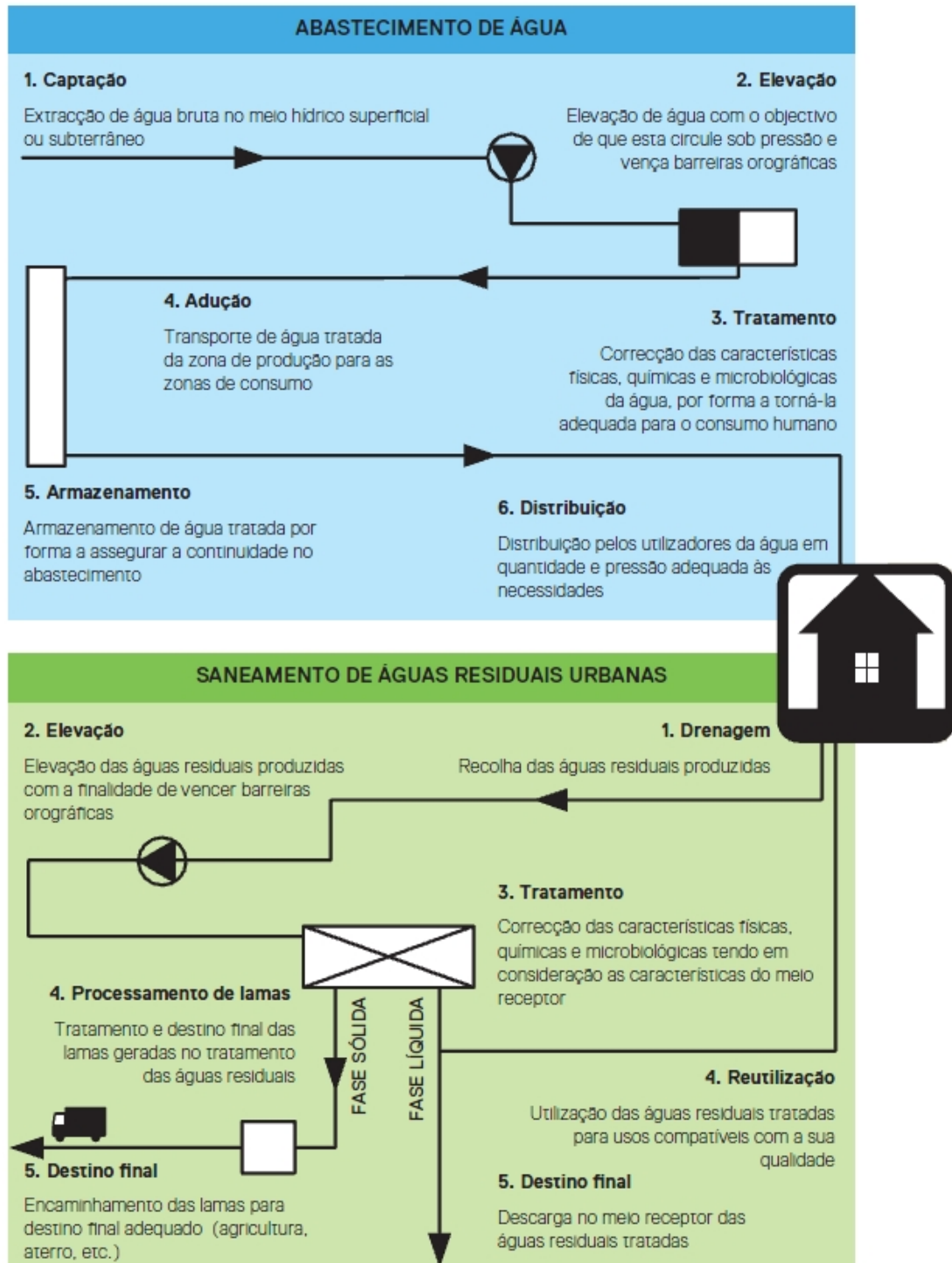


Figura 1 – Cadeia de valor do sector de serviços de águas. (Fonte: ERSAR, 2010)

2.2. O grupo ADP – Águas de Portugal

Conforme referido pela AdP (2013), na última década do século XX, Portugal revelava insuficiências graves nos sistemas de abastecimento de água para consumo público e baixos níveis de atendimento no saneamento de águas residuais. A gestão e exploração destes sistemas eram da responsabilidade exclusiva dos Municípios, predominando os investimentos numa lógica local e não integrada. A única exceção era a EPAL, empresa já centenária, responsável pela distribuição domiciliária de água no Município de Lisboa e pelo fornecimento de água a cerca de 20 Municípios da área da Grande Lisboa.

A Águas de Portugal foi constituída em 1993 para ser um dos instrumentos da mudança que iria permitir aumentar, de forma sustentável e em quantidade e qualidade, os níveis de atendimento da população portuguesa nos sectores de abastecimento de água e de saneamento de águas residuais. A criação de Sistemas Multimunicipais de Abastecimento de Água e de Saneamento de Águas Residuais, cuja responsabilidade lhe foi então atribuída, tinha por objetivos sanar as insuficiências dos sistemas de abastecimento de água para consumo público e aumentar os níveis de atendimento da população servida por sistemas de saneamento de águas residuais.

Em Outubro de 2000, em resultado de novas orientações estratégicas e como reconhecimento do mérito e potencialidades da experiência empresarial desenvolvida, foram integradas nas competências do Grupo AdP todas as participações do Estado Português na área do Ambiente, nomeadamente a área de tratamento e valorização de resíduos. Prosseguindo uma função estruturante no sector do Ambiente, o Grupo AdP assume um papel de grande contribuidor para os objetivos operacionais dos planos estratégicos sectoriais do País. Nesta área, o Grupo AdP atua em todas as fases do ciclo urbano da água, integrando a captação, tratamento e a distribuição de água para consumo público e a recolha, o tratamento e a rejeição de águas residuais urbanas e industriais,

incluindo a sua reciclagem e reutilização.

Em parceria com os Municípios é responsável pela gestão e exploração de Sistemas Multimunicipais, cuja prioridade é dotar o País das infraestruturas que permitam nivelar o atendimento das populações e a qualidade de serviço com a média europeia e dos países mais desenvolvidos. Em parceria com o sector privado, gere ainda a concessão de sistemas municipais de abastecimento de água e saneamento de águas residuais.

Tendo por objetivos prestar aos clientes das empresas multimunicipais um serviço de qualidade ao menor custo e contribuir para a dinamização do mercado dos prestadores de serviços, o Grupo promove a contratação externa de serviços de Operação e Manutenção no âmbito dos Sistemas Multimunicipais de Abastecimento de Água e Saneamento de Águas Residuais. Em Portugal, o Grupo AdP participa num conjunto de empresas que, em parceria com os municípios, prestam serviços a cerca de 80% da população portuguesa.

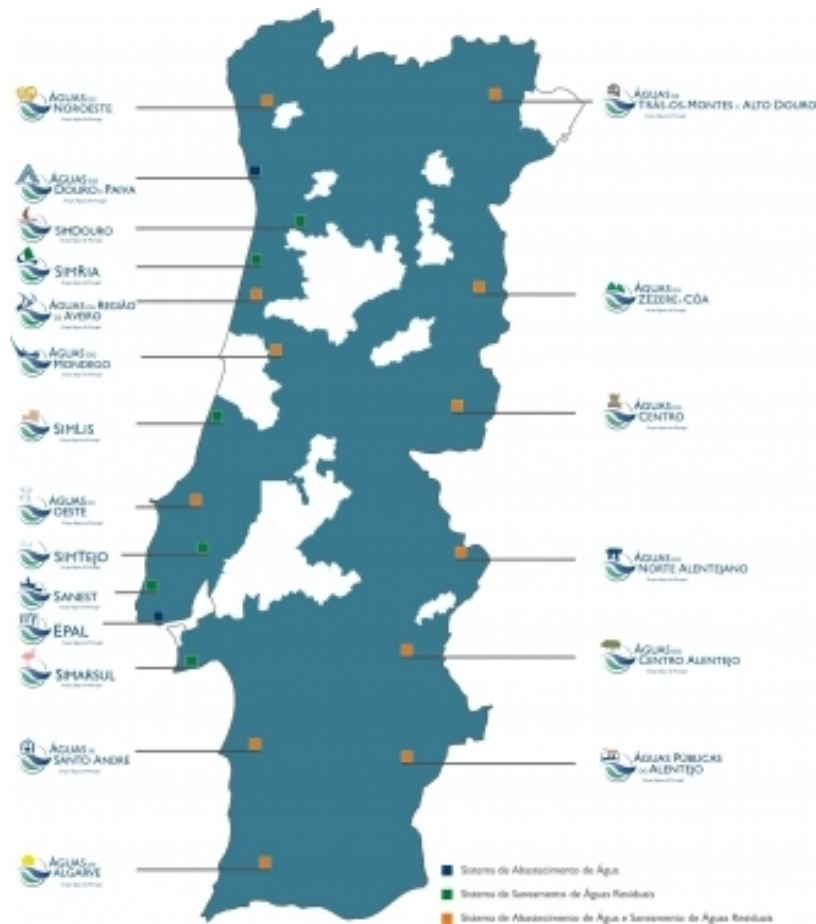


Figura 2 – Presença da ADP em Portugal. (Fonte: ADP, 2013)

2.3. A SIMRIA – Saneamento Integrado dos Municípios da Ria, SA

2.3.1. Apresentação

A SIMRIA – Saneamento Integrado dos Municípios da Ria, SA, é a sociedade anónima responsável pela construção, gestão e exploração do Sistema Multimunicipal de Saneamento da Ria de Aveiro, criado através do Decreto-Lei nº101/97, de 26 de Abril. Em conformidade com o disposto no Decreto-Lei nº 329/2000, de 22 de dezembro, a intervenção da SIMRIA ao nível da requalificação ambiental, foi ampliada aos Municípios de Espinho, Santa Maria da Feira e restante parte de Ovar, no sentido de dar resposta à recuperação do ecossistema da Barrinha de Esmoriz / Lagoa de Paramos.

A SIMRIA é responsável pela recolha, tratamento e rejeição dos efluentes domésticos e industriais dos municípios pertencentes ao Sistema Multimunicipal de Saneamento da Ria de Aveiro, promovendo a requalificação ambiental dos ecossistemas da Ria de Aveiro e da Barrinha de Esmoriz / Lagoa de Paramos num quadro de sustentabilidade económica, financeira, técnica e social. A sua atividade configura, assim, um serviço de interesse geral, com um papel determinante na melhoria da qualidade de vida dos cidadãos e na garantia da coesão económica e social da região, estando sujeita a um conjunto de obrigações de serviço público que importa salientar: universalidade, continuidade, acessibilidade, qualidade do serviço e proteção dos utilizadores.

O Contrato de Concessão, entre o Estado Português e a SIMRIA, foi celebrado a 8 de maio de 2000, por um período de 30 anos. Em virtude do alargamento do Sistema Multimunicipal, com vista à resolução dos problemas de saneamento da Barrinha de Esmoriz, foi celebrado o primeiro aditamento ao Contrato de Concessão a 22 de novembro de 2000. A 19 de janeiro de 2009 foi celebrado o segundo aditamento ao Contrato de Concessão que prolongou o período de concessão de 30 para 50 anos e ampliou a área de

intervenção ao município de Cantanhede, à restante parte do município de Oliveira do Bairro e à bacia da Laje e Cáster no município de Santa Maria da Feira. Após 13 anos de existência a SIMRIA encontra-se numa fase de plena maturidade de exploração das suas infraestruturas.

Detentora de um capital social de 16 712 225 euros, a SIMRIA é uma empresa de direito privado e de capitais maioritariamente públicos, participada pela Águas de Portugal, SGPS, SA, e pelos municípios de Águeda, Albergaria-a-Velha, Aveiro, Cantanhede, Espinho, Estarreja, Ílhavo, Mira, Murtosa, Oliveira do Bairro, Ovar, Santa Maria da Feira e Vagos.

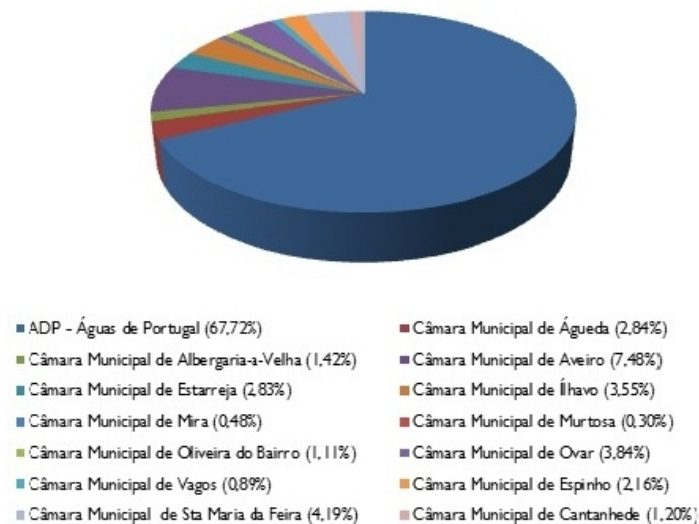


Figura 3 – Estrutura Acionista da SIMRIA. (Fonte: SIMRIA, 2011a)

2.3.2. Estrutura Organizacional

Em termos organizacionais, a SIMRIA encontra-se estruturada segundo duas Direções sob a tutela da Administração, nomeadamente a Direção de Operação e Infraestruturas e a Direção Administrativa e Financeira, as quais são responsáveis pelas áreas funcionais que tutelam.

A reportar diretamente à Administração existem os órgãos de Secretariado da Administração, Sistema de Responsabilidade Empresarial, Comunicação e Imagem,

Departamento de Qualidade do Processo e Assessoria Jurídica. Estes quatro últimos com funções transversais a toda a empresa.

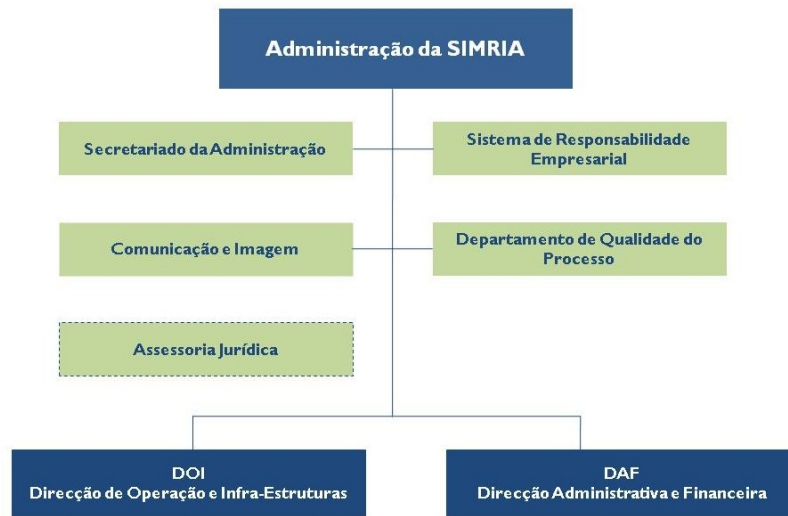


Figura 4 – Estrutura Organizacional da SIMRIA. (Fonte: SIMRIA, 2011a)

Direção de Operação e Infraestruturas

A DOI é responsável por toda a atividade associada à recolha, tratamento e rejeição de efluentes, bem como à manutenção dos equipamentos. Divide-se em três áreas funcionais: Centro Operacional I, responsável pela Operação e Manutenção da ETAR Norte; Centro Operacional II, responsável pela Operação e Manutenção das ETAR Sul e S. Jacinto; Operação e Manutenção de Conduitas, Estações Elevatórias e Pontos de Entrada (OMCEEPE).

É igualmente responsável por toda a atividade associada à conceção e planeamento, através da área funcional Estudos e Planeamento, e por toda a atividade associada à construção e manutenção das infraestruturas, através da área funcional Obras.

Nesta direção está ainda inserida a área funcional Administração de Sistemas, responsável pela informática, telecomunicações, SIG (sistemas de informação geográfica) e automação.

A reportar diretamente à Direção de Operação e Infraestruturas, existe ainda o órgão de *staff* Secretariado.

Direção Administrativa e Financeira

É responsável pela gestão administrativa, económica, financeira e de recursos humanos. É nesta Direção que estão concentrados os serviços de Contabilidade, Compras, Tesouraria, Recursos Humanos, Fundo de Coesão e Gestão de Frota.

2.3.3. Relações com *Stakeholders*

Na qualidade de empresa responsável por um serviço de interesse público, cuja atividade se reflete na requalificação ambiental da região, em prol da melhoria da qualidade de vida das populações, a SIMRIA interage com vários parceiros que, direta ou indiretamente, constituem partes interessadas no desempenho da sua atividade.



Figura 5 – Organigrama dos *Stakeholders* da SIMRIA. (Fonte: SIMRIA, 2011a)

Do Ministério da Agricultura, do Mar, do Ambiente e do Ordenamento do Território, entidade concedente da gestão do Sistema Multimunicipal, emanam um conjunto de orientações, que norteiam a SIMRIA no cumprimento da missão que lhe foi atribuída.

A Águas de Portugal, SGPS, enquanto acionista maioritário detentor de um relevante *know-how* e experiência acumulada no sector, representa um pilar fundamental na gestão estratégica da empresa.

Com um duplo posicionamento, alguns municípios apresentam-se simultaneamente como acionistas e clientes/utilizadores diretos do Sistema Multimunicipal, com quem a SIMRIA estabelece diversas parcerias no sentido de melhorar as taxas de atendimento de saneamento que caracterizam a região. Pelo facto de ter integrado 9 municípios abrangidos pelo Sistema Multimunicipal, a AdRA passou a ser o principal cliente da SIMRIA. Para além da AdRA e dos Municípios, a SIMRIA tem também como cliente direto a Unidade Industrial da Portucel, que apenas utiliza a componente de rejeição do Sistema Multimunicipal, uma vez que possui instalações próprias para o tratamento dos seus efluentes.

As populações, embora indiretamente, beneficiam consideravelmente do sistema multimunicipal e serão a principal parte interessada na qualidade dos serviços da SIMRIA. A empresa tem dado particular importância ao desenvolvimento de diversas ações de comunicação junto das populações no âmbito da Educação Cívica em matéria ambiental. A este nível, refira-se o trabalho de sensibilização realizado, prioritariamente junto de um público infanto-juvenil, na promoção das questões ambientais, em particular dos recursos hídricos.

As relações com as entidades comunitárias sempre se pautaram pelo bom relacionamento entre as partes, traduzido nas diversas candidaturas que suportam a construção e ampliação do Sistema Multimunicipal.

Por fim, importa referir as relações estabelecidas com os fornecedores, as Organizações Não Governamentais (ONG), bem como com as diversas entidades do sector, de onde se destaca a Entidade Reguladora dos Serviços de Águas e Resíduos (ERSAR) e a Administração para os Recursos Hídricos (ARH), com as quais a SIMRIA procura estabelecer relações estáveis e de confiança.

Em suma, em virtude das diversas vertentes que caracterizam a sua atividade, a SIMRIA

possui um conjunto de interfaces, relativamente aos quais procura responder, de forma eficiente, aos diferentes compromissos assumidos, quer sejam técnicos, económicos, sociais ou ambientais, privilegiando um relacionamento sério e virtuosos, suportado em valores de ética e justiça social.

2.3.4. Principais indicadores da atividade da SIMRIA no triénio 2009/2011

Tabela 1 – Indicadores da atividade da SIMRIA. (Fonte: SIMRIA, 2011a)

| | Unidade | 2009 | 2010 | 2011 |
|---|--------------------------------|---------|---------|---------|
| Atividade | | | | |
| Volume de efluente recolhido | 10 ³ m ³ | 32.578 | 37.679 | 37.216 |
| Volume de efluente tratado | 10 ³ m ³ | 22.156 | 26.928 | 26.194 |
| Volume de efluente faturado | 10 ³ m ³ | 32.578 | 36.679 | 37.216 |
| População abrangida | Hab | 468.686 | 468.686 | 468.686 |
| População servida | Hab | 430.355 | 450.941 | 468.686 |
| Municípios abrangidos | Nº | 13 | 13 | 13 |
| Municípios servidos | Nº | 13 | 13 | 13 |
| Infraestruturas | | | | |
| Extensão de rede em funcionamento | km | 266 | 303 | 315 |
| Estações elevatórias e pontos de entrada disponibilizados | Nº | 153 | 167 | 208 |
| Estações elevatórias em funcionamento | Nº | 71 | 74 | 76 |
| ETAR em funcionamento | Nº | 5 | 5 | 5 |
| Emissário submarinos em funcionamento | Nº | 2 | 2 | 2 |
| Economico-Financeiros | | | | |
| Investimento | Milhar € | 9.187 | 13.013 | 6.817 |
| Volume de Negócios | Milhar € | 12.366 | 16.172 | 16.234 |
| Gastos Operacionais | Milhar € | 11.496 | 12.949 | 13.830 |
| Gastos com Pessoal | Milhar € | 2.162 | 2.088 | 1.955 |
| Gastos Financeiros | Milhar € | 3.689 | 3.059 | 4.277 |
| Resultados Operacionais | Milhar € | 2.990 | 5.562 | 5.103 |
| Resultado Líquido do Exercício | Milhar € | 271 | 3.338 | 1.207 |
| Imobilizado Líquido | Milhar € | 209.188 | 218.637 | 207.494 |
| Cash- Flow | Milhar € | 1.931 | 5.823 | 4.444 |
| Ativo Líquido | Milhar € | 245.085 | 255.695 | 240.481 |
| Endividamento | % | 43 | 42 | 45 |
| Recursos Humanos | | | | |
| Colaboradores em 31 de Dezembro | Nº | 83 | 85 | 82 |
| Nº médio de colaboradores | Nº | 81 | 85 | 86 |
| Produtividade | | | | |
| Gastos com pessoal/caudal total | €/m ³ | 0,0664 | 0,0554 | 0,0525 |
| Gastos financeiros/caudal total | €/m ³ | 0,1132 | 0,0812 | 0,1149 |
| Gastos operacionais/caudal total | €/m ³ | 0,3529 | 0,3437 | 0,3716 |

2.3.5. Caracterização dos ecossistemas existentes na área de atuação da SIMRIA

Ao tratar as águas residuais, a SIMRIA concilia as atividades humanas com a preservação dos recursos hídricos, contribuindo assim para o desenvolvimento sustentável da região. A SIMRIA implementou um Programa de Monitorização da Qualidade da Água para garantir que todas as etapas do processo de tratamento são controladas em termos ambientais.

Ria de Aveiro

A Ria de Aveiro faz parte de um ecossistema lagunar com reconhecido valor ambiental, económico e social.

A Ria de Aveiro estende-se, pelo interior, paralelamente ao mar, numa distância de 47 km, no sentido Este-Oeste, desde Ovar até Mira, e com uma largura máxima de 11 km. A Ria é o resultado do recuo do mar, com a formação de cordões litorais que, a partir do séc. XVI, formaram uma laguna que constitui um dos mais importantes e belos acidentes hidrográficos da costa portuguesa. Abarca cerca de 11000 hectares, dos quais 6000 estão permanentemente alagados, e desdobra-se em quatro importantes canais ramificados em esteiros que circundam um sem número de ilhas e ilhotes. Nela desagüam o Vouga, o Antuã e o Boco, tendo como única comunicação com o mar um canal que corta o cordão litoral entre a Barra e S. Jacinto, permitindo o acesso ao Porto de Aveiro, de embarcações de grande calado.



Figura 6 – Vista aérea da Ria de Aveiro na zona da Barra de Aveiro.

(Fonte: <http://portugalfotografiaaerea.blogspot.pt/>)

Rica em peixes e aves aquáticas, possui grandes planos de água, locais de eleição para a prática de todos os desportos náuticos.

Nesta região o ambiente e a cultura estão fortemente entrelaçados e moldaram-se ao longo da História. A pesca, a agricultura, a apanha e transporte do moliço e a produção de sal ainda se praticam recorrendo a técnicas tradicionais. Aqui, o Homem domou os recursos naturais e, fruto dessa relação, foi criado um património cultural único.



Figura 7 – Vista aérea da Ria de Aveiro próximo do centro da cidade de Aveiro.

(Fonte: <http://portugalfotografiaaerea.blogspot.pt/>)

Ainda que tenha vindo a perder, de ano para ano, a importância que já teve na economia aveirense, a produção de sal, utilizando técnicas milenares, é, ainda, uma das atividades tradicionais mais características de Aveiro, havendo, atualmente, dezenas de salinas em laboração.

Muito especialmente no Norte da Ria, os barcos moliceiros, embarcações únicas e de linhas perfeitas, ostentando policromos e ingénuos painéis decorativos continuam a apanhar o moliço, fertilizante de eleição, bem dentro dos mais exigentes e atuais parâmetros ecológicos, que transformou solos estéreis de areia em ubérrimos terrenos agrícolas.



Figura 8 – Imagens de moliceiros. (Fontes: www.net-moliceiro.inovanet.pt e www.geolocation.ws)

Lagoa de Paramos / Barrinha de Esmoriz

A Lagoa de Paramos é uma laguna costeira de média dimensão, comunicando com o Atlântico através de um canal, a Barrinha de Esmoriz. Está situada no litoral Norte de Portugal, entre os concelhos de Ovar e Espinho ocupando uma área de cerca de 396 ha. A Barrinha de Esmoriz/Lagoa de Paramos é a zona húmida mais significativa no litoral Norte de Portugal, entre a ria de Aveiro e o estuário do Rio Minho e é dotada de uma biodiversidade ímpar (270 espécies de avifauna e pelo menos 80 espécies de flora classificadas). É uma área pertencente à Rede Natura 2000 e à REN e classificada como IBA (Important Bird Area), pela Bird Life International.

Esta laguna apresenta uma forma grosseiramente triangular, com cerca de 1500m de comprimento, na direção Norte-Sul, e cerca de 700m de largura, na direção Este-Oeste. Encontra-se afastada do mar cerca de 400m, separada por dunas e pela faixa de areia.

A acumulação de água nesta laguna resulta essencialmente dos volumes drenados por duas ribeiras principais, Ribeira de Paramos que tem a sua foz no lado Norte da Lagoa e Vala de Maceda, que desagua no lado Sul. De um modo geral, trata-se de uma laguna com baixa profundidade, apresentando os seus fundos cobertos por sedimentos móveis (lodo).



Figura 9 – Imagens aéreas da BE/LP. (Fonte: www.prof2000.pt/users/eb23fspan)

A ligação da laguna com o mar era efetuada ocasionalmente por um canal de largura variável com cerca de 600m de comprimento, com tendência a fechar por ação do transporte sólido litoral. A comunicação com o mar verificava-se na altura das marés-vivas

ou quando ocorria precipitação muito elevada. Atualmente não se encontra sob a influência das marés. A comunicação com o mar não é permanente pois a abertura e fecho do canal dependem da ação do mar e/ou da intervenção humana, através de dragagens. Em virtude deste facto, verificam-se flutuações importantes do nível de água, que prejudicam a reprodução das aves que nidificam na barrinha. Quando a comunicação com o mar está interrompida e a Barrinha está cheia, a área do plano de água pode atingir os 56 ha, correspondendo a um volume aproximado de 230.000m³ de água.

Os seus habitats encontram-se ameaçados, devido ao completo abandono a que foi sujeita durante as ultimas décadas. A falta de um plano de gestão, o agravamento da poluição dos canais de água doce, a pressão turística e urbanística, assim como obras mal executadas na tentativa de minimizar os efeitos poluentes, são apenas alguns dos fatores que contribuíram para o agravamento da situação. Sendo um espaço a preservar mas que se encontra gravemente ameaçado



Figura 10 – Imagens elucidativas do estado da BE/LP.

(Fonte: <http://cusquicesdeesmoriz.blogspot.pt/>)

Desde a construção da ETAR de Esmoriz (já desativada) e da ETAR de Espinho verificou-se uma significativa melhoria da qualidade da água da vala de Maceda. Apesar de algumas melhorias, a situação na Ribeira de Paramos (lado Norte) mantém-se preocupante, continuando esta ribeira muito poluída, devido aos efluentes industriais descarregados, sem tratamento, a montante da ribeira, daí resultando a péssima qualidade da água na Barrinha de Esmoriz.

2.3.6. O Sistema Multimunicipal de Saneamento

A conceção do Sistema Multimunicipal de Saneamento da Ria de Aveiro, nomeadamente o dimensionamento e localização das suas infraestruturas foi objeto de diversos estudos, em articulação com as entidades competentes e municípios, no sentido de assegurar a otimização técnica e económica da sua construção bem como da sua exploração.

O Sistema Multimunicipal de Saneamento da Ria de Aveiro apresenta uma solução para a coleta, tratamento e destino final dos efluentes da Ria de Aveiro e da Barrinha de Esmoriz/Lagoa de Paramos, cujo equilíbrio tem vindo a ser ameaçado pelos problemas de poluição devido às descargas de águas residuais domésticas e, sobretudo, industriais.

A construção do sistema municipal, que foi desenvolvida em 3 fases, conforme representado na figura seguinte, encontra-se concluída no que se refere aos investimentos previstos no Contrato de Concessão.

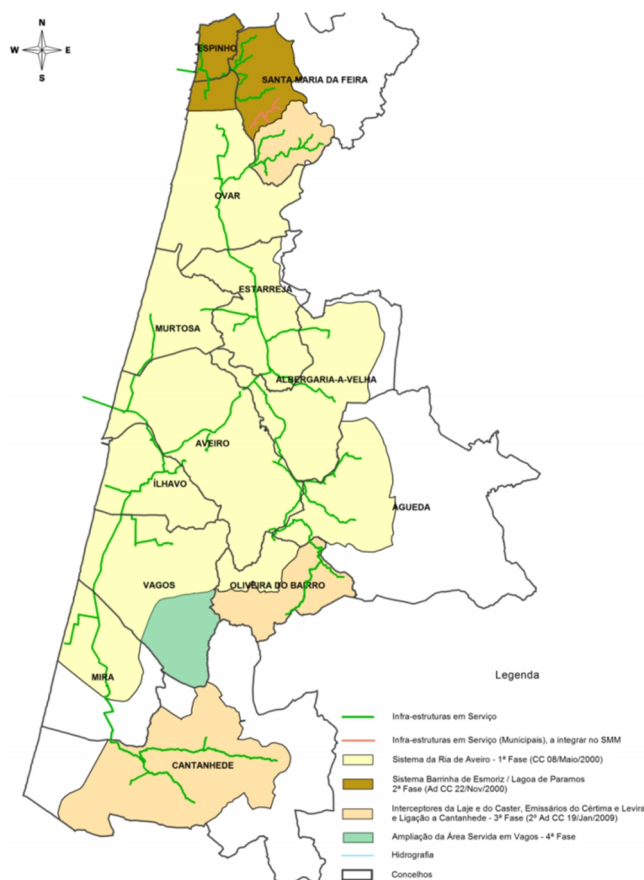


Figura 11 – Mapa da área de abrangência do Sistema Multimunicipal da SIMRIA.

(Fonte: SIMRIA, 2011b)

Tabela 2 – Zonas de Intervenção e População Equivalente Abrangida pelo sistema.

(Fonte: SIMRIA, 2011b)

| | Zonas de Intervenção | População Equivalente |
|----------------|----------------------|-----------------------|
| 1ª Fase | Águeda | 800.000 hab |
| | Albergaria-a-Velha | |
| | Aveiro | |
| | Estarreja | |
| | Ílhavo | |
| | Mira | |
| | Murtosa | |
| | Oliveira do Bairro | |
| | Ovar | |
| | Vagos | |
| 2ª Fase | Espinho | 220.000 hab |
| | Ovar | |
| | Santa Maria da Feira | |
| 3ª Fase | Cantanhede | 121.520 hab |
| | Oliveira do Bairro | |
| | Santa Maria da Feira | |

Investimento/financiamento

O investimento total previsto para cumprimento da missão da SIMRIA ascende a 234,5 milhões de euros, estando já realizados mais de 90% deste valor.

A Comunidade Europeia também contribui na cobertura financeira deste investimento através do Fundo de Coesão e do POVT.

Tabela 3 – Valores de Investimento/Financiamento. (Fonte: SIMRIA, 2011b)

| | | Investimento (milhões de €) | Comparticipação do Fundo de Coesão |
|----------------|---------------------|--------------------------------|---------------------------------------|
| 1ª Fase | Ria de Aveiro | 154 | 85% |
| 2ª Fase | Barrinha de Esmoriz | 23 | 59% |
| | Lagoa de Paramos | | |
| 3ª Fase | Cértima e Levira | 12,4 | 48% |
| | Laje e Cáster | | |
| | Cantanhede | | |

Para além da participação comunitária do Fundo de Coesão, o projeto também é financiado por capitais próprios e capitais alheios.

Infraestruturas

1ª Fase



Figura 12 – 1ª Fase do Sistema Multimunicipal. (Fonte: SIMRIA, 2013)

O Sistema Multimunicipal foi concebido inicialmente para servir os municípios de Águeda, Albergaria-a-Velha, Aveiro, Estarreja, Ílhavo, Mira, Murtosa, Oliveira do Bairro, Ovar e

Vagos, abrangendo um total de 800 mil habitantes-equivalentes, em três grandes componentes:

I – Componente de Recolha de Águas Residuais (Subsistemas Norte, Sul e Torreira / S. Jacinto)

Tabela 4 – Descrição do Sistema de Recolha - 1ª Fase. (Fonte: SIMRIA, 2011b)

| | Extensão (km) | Estações Elevatórias (nº) | Pontos de Entrada (nº) |
|-------------------------------|--------------------------|--------------------------------------|-----------------------------------|
| Intercetor Norte | 62,8 | 15 | 21 |
| Intercetor Sul | 37,2 | 8 | 8 |
| Intercetor Vouga | 45,2 | 16 | 25 |
| Emissário Aveiro / Ílhavo | 5,3 | 4 | 3 |
| Emissário Barra / Costa Nova | 3,9 | 2 | 2 |
| Emissário Gafanha da Nazaré | 3,7 | 1 | 1 |
| Emissário da Torreira | 10,7 | 5 | 5 |
| Conduta Elevatória S. Jacinto | 1,6 | 1 | 1 |
| Emissários Complementares | 4,9 | 2 | 2 |

II – Componente de Tratamento de Águas Residuais

Tabela 5 – ETARs responsáveis pela componente de tratamento - 1ª Fase. (Fonte: SIMRIA, 2011b)

| | População Abrangida HP* (hab-eq) | Caudal Médio Diário HP* (hab-eq) |
|-----------------|---|---|
| ETAR Norte | 272.000 | 48.705 |
| ETAR Sul | 159.700 | 39.278 |
| ETAR S. Jacinto | 9.377 | 1.065 |

*Horizonte de Projeto

III – Componente de Rejeição (Intercetor Geral e Exutor Submarino de São Jacinto)

Tabela 6 – Descrição do sistema de rejeição - 1ª Fase. (Fonte: SIMRIA, 2011b)

| | Extensão (km) | Estações Elevatórias (nº) |
|----------|----------------------|----------------------------------|
| Geral | 24,7 | 3 |
| Exutor I | 3,4 | -- |

A 1ª fase do Sistema Multimunicipal está totalmente disponibilizada aos clientes municipais e Portucel, encontrando-se em curso o Plano de Manutenção de Infraestruturas associado a esta fase.

2ª Fase

No ano 2000 o Sistema foi alargado aos municípios de Espinho e de Santa Maria da Feira (parcialmente) e reforçado em Ovar, de forma a integrar o projeto de despoluição da Barrinha de Esmoriz / Lagoa de Paramos.

O subsistema da Barrinha de Esmoriz / Lagoa de Paramos envolve a ligação à ETAR de Espinho dos subsistemas de Espinho, Ovar (Esmoriz e Cortegaça) e de Santa Maria da Feira (bacia que drena para a Barrinha de Esmoriz / Lagoa de Paramos), sendo que os efluentes tratados na ETAR de Espinho são posteriormente encaminhados para o Exutor Submarino de Espinho através do qual são rejeitados no mar.

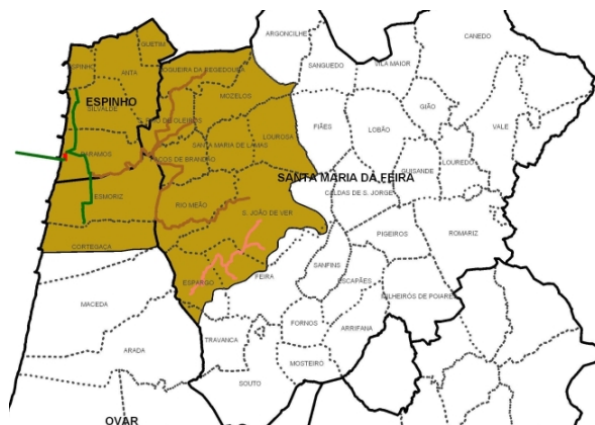


Figura 13 – 2ª Fase do Sistema Multimunicipal. (Fonte: SIMRIA, 2013)

Para o efeito, a ETAR de Espinho foi remodelada e ampliada para receber um caudal de 30.890 m³/dia no horizonte do projeto. O Sistema da Barrinha de Esmoriz / Lagoa de Paramos implicou ainda a execução de cerca de 29 km de Condutas e 5 Estações Elevatórias.



Figura 14 – Imagem da ETAR de Espinho. (Fonte: SIMRIA, 2013)

Com o alargamento do Sistema à Barrinha de Esmoriz / Lagoa de Paramos, o Sistema Multimunicipal de Saneamento da Ria de Aveiro, inicialmente dimensionado para uma população total de oitocentos mil habitantes-equivalentes, passou a servir um universo populacional que ultrapassa um milhão de habitantes-equivalentes, num horizonte de 30 anos.

I – Componente de Recolha de Águas Residuais (Subsistema Barrinha de Esmoriz / Lagoa de Paramos)

Tabela 7 – Descrição do Sistema de Recolha - 2ª Fase. (Fonte: SIMRIA, 2011b)

| | Extensão (km) | Estações Elevatórias (nº) | Pontos de Entrada (nº) |
|--|--------------------------|--------------------------------------|-----------------------------------|
| Intercetor de Esmoriz/Cortegaça | 6,0 | 1 | 1 |
| Emissário de Espinho | 3,0 | 0 | 2 |
| Intercetores de Rio Maior/Silvalde/Beire | 20,0 | 4 | 36 |

II – Componente de Tratamento de Águas Residuais

Tabela 8 – ETARs responsáveis pela componente de tratamento - 2ª Fase. (Fonte: SIMRIA, 2011b)

| | População Abrangida HP* (hab-eq) | Caudal Médio Diário HP* (hab-eq) |
|-----------------|---|---|
| ETAR de Espinho | 194232 | 30890 |
| ETAR da Remolha | 8874 | 1040 |

*Horizonte de Projeto

III – Componente de Rejeição (Exutor Submarino de Espinho)

Tabela 9 - Descrição do sistema de rejeição - 2ª Fase. (Fonte: SIMRIA, 2011b)

| | Extensão (km) | Estações Elevatórias (nº) |
|-----------|----------------------|----------------------------------|
| Exutor II | 2,0 | -- |

A 2ª fase do Sistema Multimunicipal está totalmente disponibilizada aos clientes municipais (Ovar, Espinho e Santa Maria da Feira), encontrando-se em curso o Plano de Manutenção de Infraestruturas associado a esta fase.

3ª Fase

A 3ª fase corresponde a todos os investimentos de expansão da rede a novas áreas territoriais e/ou municípios. Refira-se nomeadamente o prolongamento do Intercetor Vouga para recolher os efluentes gerados na totalidade do concelho de Oliveira do Bairro, o prolongamento do Intercetor Sul de forma a abranger o concelho de Cantanhede e o prolongamento do Intercetor Norte ao longo das bacias das ribeiras do Cáster e da Laje no concelho de Santa Maria da Feira.

I – Componente de Recolha de Águas Residuais

Tabela 10 – Descrição da expansão da componente de recolha - 3ª fase. (Fonte: SIMRIA, 2011b)

| | Extensão (km) | Estações Elevatórias (nº) | Pontos de Entrada (nº) |
|--------------------------------|--------------------------|--------------------------------------|-----------------------------------|
| Emissários do Cértima e Levira | 15,5 | 5 | 11 |
| Emissário de Cantanhede | 16,0 | 4 | 15 |
| Emissários de Laje e Cáster | 24,0 | 3 | 35 |

Com a conclusão da terceira fase, estão em funcionamento as seguintes infraestruturas:

Tabela 11 – Infraestruturas ao serviço da SIMRIA

ETAR's

- Norte
- Sul
- Espinho
- S. Jacinto
- Remolha
- Santa Catarina
- Salgueiro
- Ouca

Condutas

- 310 km

Estações Elevatórias e Pontos de Entrega

- 76 Estações Elevatórias
- 150 Pontos de Entrada Isolados

Emissários Submarinos

- 2

Desde o início da sua atividade, a SIMRIA tem vindo a aumentar e melhorar a sua rede de infraestruturas. Com os investimentos realizados e a qualidade do serviço da SIMRIA, evita-se que sejam descarregados na ria cerca de 3 milhões de m³ de efluente, garantindo o tratamento adequado antes da sua rejeição no mar.

Operação

O ciclo que caracteriza a atividade da SIMRIA tem início na identificação das necessidades de saneamento em alta da região que, depois de suportadas num contrato de concessão e no recurso à comparticipação financeira comunitária, desencadeiam uma sucessão de etapas através das quais a SIMRIA dá cumprimento à missão a que se propôs.

Na Operação do Sistema Multimunicipal, a SIMRIA recorre a todo o *know-how* e experiência acumulada no sentido, não só, de criar valor intrínseco no desenvolvimento da sua atividade, mas sobretudo de obter um desempenho que seja devidamente valorizado pelas diferentes interfaces com que se relaciona.

Seguidamente encontra-se ilustrada a **cadeia de valor** associada à atividade chave da empresa:

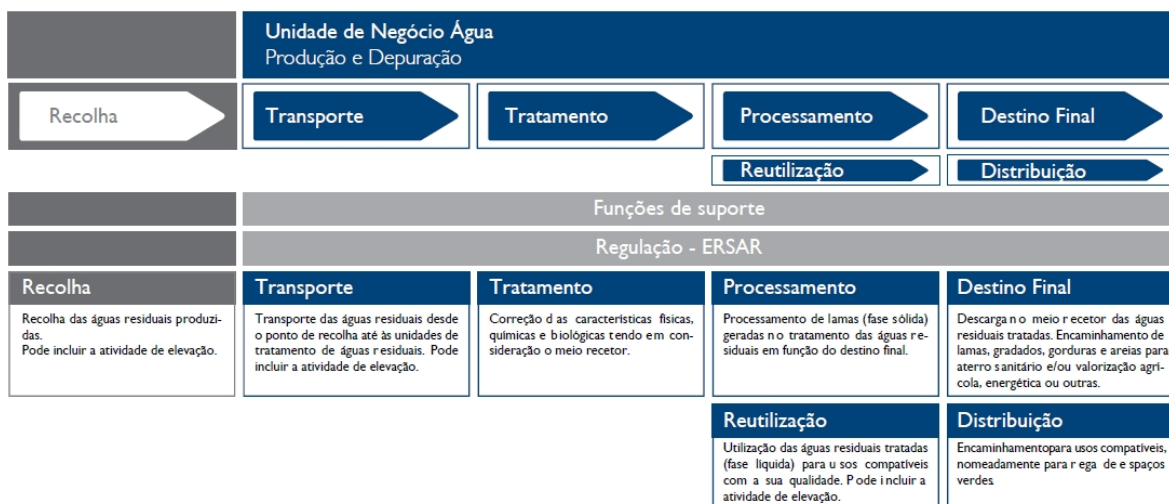


Figura 15 – Cadeia de valor da atividade chave da empresa. (Fonte: SIMRIA, 2011a)

Recolha, tratamento e rejeição

As equipas dos Centros Operacionais (CO), mantêm a seu cargo a operação e manutenção da ETAR Norte (COI) e da ETAR Sul (COII). Este facto permite reforçar ainda mais o domínio das equipas da SIMRIA, sobre a vertente operacional do processo de tratamento, o que se traduz numa melhoria da qualidade do serviço prestado. A exploração da ETAR de Espinho permanece a cargo da LUSÁGUA - Gestão de Águas, SA, ao abrigo do contrato existente.

A realização de análises semanais e mensais ao efluente tratado e rejeitado nos Exutores Submarinos de São Jacinto e Espinho, permitem atestar o cumprimento dos parâmetros constantes nas respetivas licenças de descarga. São ainda realizadas análises às águas balneares que permitem aferir a qualidade das mesmas.

Tabela 12 – Valores Limite de Emissão (VLE). (Fonte: SIMRIA, 2011b)

| | Ph | SST | CQO | CBO ₅ |
|--------------|-------|-----|-----|------------------|
| ETAR Norte | 6 - 9 | 60 | 125 | 25 |
| ETAR Sul | 6 - 9 | 60 | 15 | 25 |
| ETAR Espinho | 6 - 9 | 35 | 125 | 25 |

Tabela 13 – Rejeição de efluente (Valores em milhares de m³). (Fonte: SIMRIA, 2011b)

| | Exutor Submarino S. Jacinto | Exutor Submarino Espinho | ETAR da Remolha | ETARs de Vagos |
|------|--------------------------------|-----------------------------|--------------------|-------------------|
| 2003 | 22284 | 1769 | -- | -- |
| 2004 | 23620 | 2035 | -- | -- |
| 2005 | 22955 | 2767 | -- | -- |
| 2006 | 28687 | 3135 | -- | -- |
| 2007 | 28022 | 2843 | -- | -- |
| 2008 | 27710 | 2984 | -- | -- |
| 2009 | 28208 | 4369 | -- | -- |
| 2010 | 31404 | 5849 | 426 | -- |
| 2011 | 31476 | 5355 | 385 | -- |
| 2012 | 28964 | 4718 | 318 | 42 |

2.3.7. Estação de Tratamento de Águas Residuais Sul (ETAR Sul)



Legenda:

1. Edifício de Pré-Tratamento
2. Decantação Primária
3. Tratamento Biológico
4. Decantação Secundária
5. Estação Elevatória IS8
6. Espessamento de lamas
7. Digestão Primária
8. Digestão Secundária
9. Edifício de Digestão, Desidratação e PT
10. Silo de lamas
11. Gasómetro
12. Edifício de Exploração e Oficinas
13. Portaria

Figura 16 – ETAR Sul da SIMRIA. (Fonte: SIMRIA, 2013)

Caracterização da ETAR

Características técnicas:

A ETAR Sul localiza-se na freguesia da Gafanha da Encarnação, concelho de Ílhavo, encontrando-se em funcionamento desde Agosto de 2002.

Atualmente recebe e trata os efluentes provenientes de Aveiro (parte), Ílhavo, Mira, Vagos e Cantanhede.

Tabela 14 – Dados de Projeto. (Fonte: SIMRIA, 2013)

| Dados de Base | | |
|------------------------------------|----------|----------|
| | Ano 2018 | Ano 2038 |
| População Equivalente (hab) | 159.700 | 182.000 |
| Caudal médio (m ³ /dia) | 39.278 | 46.635 |
| Caudal de ponta (l/h) | 860 | 1.020 |

O montante de investimento na construção da ETAR foi de 13,2 milhões de euros, com participação em 85% pelo Fundo de Coesão.

Características Principais:

- Receção de efluentes provenientes do Intercetor Sul, do Emissário Barra/Costa Nova e do Emissário Aveiro/Ílhavo.

- Pré-tratamento:
 - 2 tamisadores (5 mm) + 1 grelha manual (20 mm)
 - 2 desarenadores /desengorduradores (2x308 m³)
- Medição de caudal através de medidor eletromagnético.
- Decantação Primária:
 - 2 decantadores circulares ($\emptyset=31,5$ m e $V=2750$ m³)
- Tratamento Biológico:
 - 2 tanques de arejamento (2x3882 m³)
 - Lamas ativadas de média carga e baixa idade
 - 4 turbinas de arejamento superficial em cada tanque
- Decantação secundária:
 - 2 decantadores circulares de sucção ($\emptyset=38$ m e $V=4150$ m³)
- Espessamento de lamas:
 - 2 espessadores para as lamas primárias (2x276 m³)
 - 1 flotor para as lamas secundárias (250 m³)
 - Edifício de pressurização e tanque de mistura de lamas (15 m³)
- Digestão anaeróbia:
 - 2 digestores primários (2x3000 m³)
 - 1 digestor secundário (1370 m³)
- Aquecimento das lamas:
 - 2 caldeiras com potência unitária de 350.000 Kcal/h, alimentadas a biogás e gás natural
 - 2 permutadores de calor
- Desidratação de lamas:
 - 2 centrifugas

- Armazenamento de lamas:
 - Silo com 100 m³ de capacidade
- Desodorização por via química, no edifício de pré-tratamento, na sala de desidratação das lamas e nos espessadores.
- Potência instalada: 1000KVA
- Destino final do efluente tratado – Oceano Atlântico, após descarga através do exutor submarino de S. Jacinto
- Captação própria de água
- Ligação à rede municipal de abastecimento de água e ligação à rede de distribuição de gás natural.
- Moto-gerador (725 Kw)

Caracterização do efluente

A água residual doméstica pode ser dividida em 2 categorias distintas que se distinguem pela sua origem:

- **água “cinzenta”** – produzida nos banhos/duches, cozinha, lavagem de roupa;
- **água “negra”** – resultante da utilização dos sanitários.

Estes 2 tipos de água residual são misturadas e transportadas ao longo do sistema de esgoto para a ETAR. Durante o percurso, esta água residual pode ainda ser diluída com águas de escorrência das chuvas e misturada com outro tipo de efluentes provenientes da atividade comercial e industrial que também são descarregados no sistema de esgoto público.

Fases do tratamento

O processo de tratamento do efluente recebido pela ETAR desenvolve-se por duas linhas distintas, a linha de água e a linha de lamas (fase líquida e fase sólida), conforme se pode verificar no diagrama seguinte.

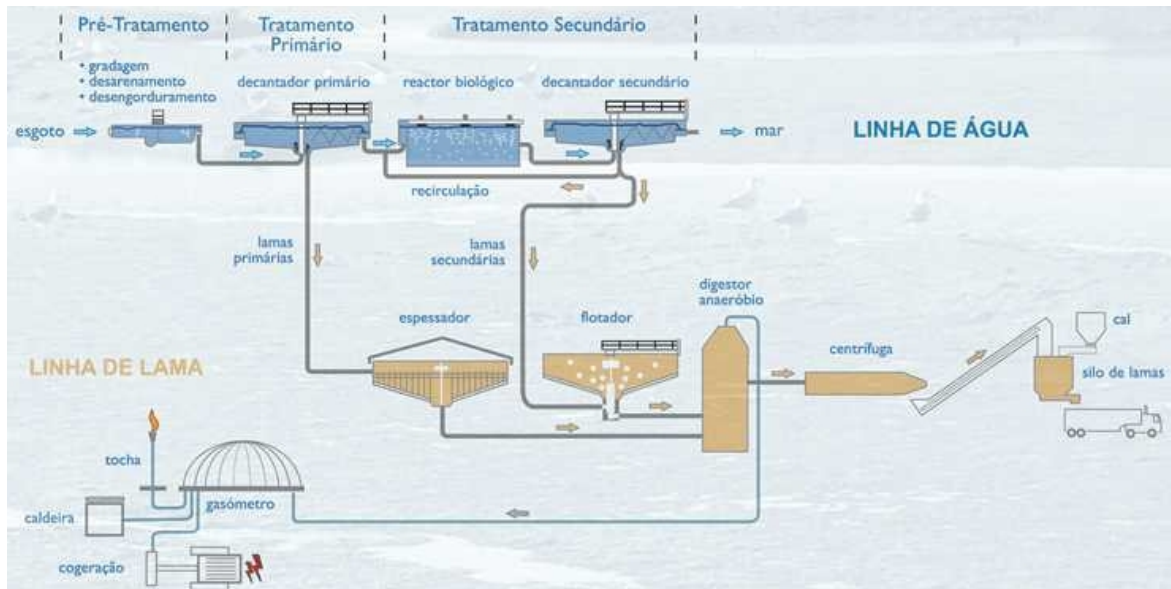


Figura 17 – Diagrama Funcional da ETAR. (Fonte: SIMRIA, 2013)

Linha de água

Na linha de água procede-se à correção das características físicas, químicas e microbiológicas e posterior descarga no meio recetor das águas residuais tratadas, tendo em consideração as características do meio recetor.

Numa primeira fase do tratamento do esgoto, **Pré-tratamento**, são realizadas operações preliminares que visam remover os sólidos grosseiros, areias e gorduras contidos no esgoto, sendo posteriormente realizado um **Tratamento Primário** de sedimentação efetuado em decantadores. A sedimentação permite uma remoção de sólidos suspensos (SS) superior a 90% (Di Bernardino, 2001). Esta operação permite atingir um nível de depuração do esgoto de 30-35% em termos de CBO_5 , reduzindo significativamente a carga poluente do esgoto. Após o tratamento primário, a matéria poluente remanescente é constituída

essencialmente por coloides.

No **tratamento biológico secundário**, a matéria orgânica do efluente resultante do tratamento primário é degradada por intermédio de microrganismos aeróbios que são introduzidos em reatores biológicos sujeitos a arejamento. No efluente do **decantador secundário**, o esgoto contém uma quantidade elevada de biomassa que se desenvolve à custa da degradação de parte significativa do material orgânico e dos nutrientes presentes no esgoto. Estes microrganismos vão sedimentar no fundo do decantador sendo removidos por decantação. Após o **tratamento secundário**, as águas residuais tratadas apresentam um nível reduzido de poluição por matéria orgânica, podendo na maioria dos casos, serem descarregadas no ambiente recetor, que neste caso é o Oceano Atlântico.

O processo completo de tratamento permite atingir, na sua globalidade, uma eficiência de tratamento superior a 90% em termos de CBO_5 , produzindo lamas primárias e lamas secundárias. Os tratamentos primário e secundário possibilitam, deste modo, uma redução significativa do volume de lamas comparativamente ao volume de esgoto inicial. No entanto estas lamas concentram ainda a carga orgânica poluente retirada ao esgoto, apresentando um estado elevado de putrefação, viscosidade e mau cheiro (Di Bernardino, 2001).

Linha de Lamas

Na linha de lamas realiza-se o tratamento das lamas geradas no tratamento das águas residuais e, após o tratamento, o encaminhamento das lamas para destino final adequado (agricultura, aterro, etc.). Nas ETAR's são produzidas quantidades significativas de lamas:

Tabela 15 – Caracterização geral de Lamas geradas no tratamento de águas residuais. (Fonte: Di Berardino, 2001)

| | % de sólidos na Lama | Resíduo Seco (g/hab.dia) | Volume (l/hab.dia) |
|--------------------------------|-----------------------------|---------------------------------|---------------------------|
| Lamas primárias não espessadas | 0,2 – 2,0 | 55 | 1,1 |
| Lamas primárias espessadas | 4,0 – 10 | -- | 0,2 – 0,5 |
| Lamas de decantação secundária | 0,5 – 1,5 | 35 | 0,7 – 1,1 |
| Lamas Mistas espessadas | 5 - 10 | 86 | 0,9 – 1,2 |

As lamas primárias, que correspondem a cerca de 2,5 - 3,5 % do volume de esgoto e resultam do processo de sedimentação dos SS, são recolhidas do decantador primário e encaminhadas para os espessadores. As lamas secundárias são removidas do decantador secundário e são reencaminhadas para o reator biológico (recirculação de lamas) ou encaminhadas para o flotador. As lamas espessadas no espessador e no flotador são recolhidas e misturadas numa câmara de lamas mistas antes de serem enviadas para o digestor anaeróbio. Os tanques de espessamento como os espessadores e o flotador permitem aumentar a concentração das lamas, possibilitando uma redução do seu volume e do caudal a enviar para a digestão.

No digestor anaeróbio processa-se a digestão anaeróbia das lamas da qual resulta a produção de biogás rico em metano.

As lamas digeridas passam por mais um processo de aumento da sua concentração através do espessamento por centrifugação antes de serem armazenadas num silo para posterior encaminhamento para destino final (aterro ou agricultura).

O biogás produzido é armazenado num gasómetro que abastece as caldeiras para o aquecimento das lamas e o moto-gerador no qual se processa a co-geração de eletricidade.

As águas residuais municipais e as lamas de esgoto apresentam-se como um substrato com características adequadas para aplicação na área biotecnológica, uma vez que apresentam uma composição muito particular que contém os compostos necessários para o crescimento microbiano.

3. PROCESSO DE DIGESTÃO ANAERÓBIA

Conforme referido em diversa bibliografia, nomeadamente em Metcalf e Eddy (1991) a Digestão Anaeróbia (DA) é um processo que ocorre naturalmente no ambiente, assim como em digestores em condições controladas, mediado por uma comunidade complexa de microrganismos que promovem a decomposição e degradação da matéria orgânica, na ausência de oxigénio, permitindo, deste modo, a estabilização bioquímica.

Na DA podem ser utilizados diferentes tipos de substratos, nomeadamente, resíduos: como lamas resultantes do tratamento de águas residuais (domésticas ou industriais), resíduos resultantes da atividade agropecuária (estruume de vacas, porcos e aves), fração orgânica dos resíduos sólidos urbanos e efluentes resultantes da atividade industrial (indústria de bebidas, produtos alimentares, amido, açúcar, processamento de papel, matadouros, químicos, leite, cosméticos, entre outros). Os produtos finais da DA são o biogás, constituído essencialmente por metano (CH_4) e dióxido de carbono (CO_2), e o produto digerido, constituído por matéria orgânica e biomassa microbiana. O biogás produzido possui um poder calorífico de aproximadamente 21-25 Mj/m^3 , cerca de 30 a 40 % inferior ao do gás natural (Appels et al., 2008).

O funcionamento do processo está dependente da ação cooperativa e sequencial de diversas bactérias de diferentes grupos tróficos que apresentam relações de sintrofia (benefício mútuo), cooperando entre si de forma a aproveitar, de modo eficiente, quantidades muito reduzidas de energia que normalmente se verificam no processo de conversão de matéria orgânica em metano (Schink, 1997).

3.1. Fases do Processo Anaeróbio

Conforme referido em Silveira (2009), as vias metabólicas de degradação dos compostos durante o processo de DA são complexas e não completamente conhecidas, processando-se

segundo uma sequência diversa de reações. Resumidamente, o processo anaeróbio pode ser dividido em quatro fases interdependentes (podendo identificar-se uma 5ª fase que ocorre em simultâneo com as restantes quatro, que é a Sulfidogénese):

- **Hidrólise** extracelular da matéria complexa
- **Acidogénese** dos açúcares e aminoácidos em ácidos gordos voláteis (AGV) e álcoois;
- **Acetogénese** sintrófica, onde os álcoois e AGV são convertidos a acetato, hidrogénio ou formato. Esta fase ocorre em simbiose com as bactérias metanogénicas (BM).
- **Metanogénese**, onde o acetato, hidrogénio e formato são convertidos em metano.

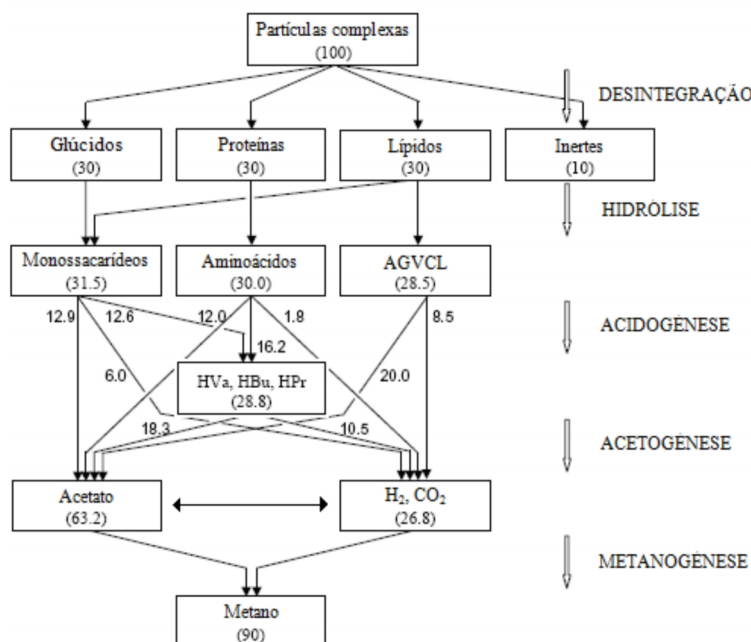


Figura 18 – Representação esquemática das fases do processo de DA. (Fonte: Silveira, 2009)

Hidrólise

Gerardi (2003) refere que o primeiro passo da DA é a hidrólise, onde enzimas hidrolíticas extra-celulares degradam compostos de elevado peso molecular (glúcidos, proteínas e lípidos) em compostos monoméricos e oligoméricos solúveis.

As bactérias hidrolíticas (BH) excretam enzimas que hidrolisam o substrato particulado em moléculas mais reduzidas, permitindo que estas atravessem as membranas celulares,

podendo ser utilizadas pelas bactérias fermentativas (BF) durante o processo subsequente de acidogénese.

Segundo Angelidaki e Sanders (2004), os glúcidos são hidrolisados a monossacáridos, as proteínas a aminoácidos e os lípidos a ácidos gordos voláteis de cadeia longa (AGVCL) e glicerol. O bom funcionamento do processo de hidrólise está dependente do contato eficiente entre a biomassa e o substrato. A eficiência da atividade enzimática está dependente, de um modo geral, de vários fatores como composição e concentração do substrato, produção e estabilidade das enzimas, disponibilidade de substrato, pH, temperatura do líquido em digestão e concentração de AGV.

Acidogénese

Em Schink (1997) é descrito que durante a acidogénese, ou fermentação ácida, os produtos resultantes da fase de hidrólise são degradados pelas BF, constituídas por uma série de estirpes bacterianas anaeróbias obrigatórias e facultativas. Estas bactérias possuem um metabolismo variado, apresentando diferentes vias de conversão donde resultam produtos fermentativos diversos, essencialmente AGV (propionato, butirato e valerato), succinato, lactato e álcoois.

A quantidade e o tipo de produtos obtidos na fase de acidogénese dependem fundamentalmente da pressão parcial de hidrogénio (H_2) no digestor, que está associada à atividade de bactérias utilizadoras de H_2 , como por exemplo, algumas espécies de BM e bactérias redutoras de sulfatos (BRS) (McInerney, 1999).

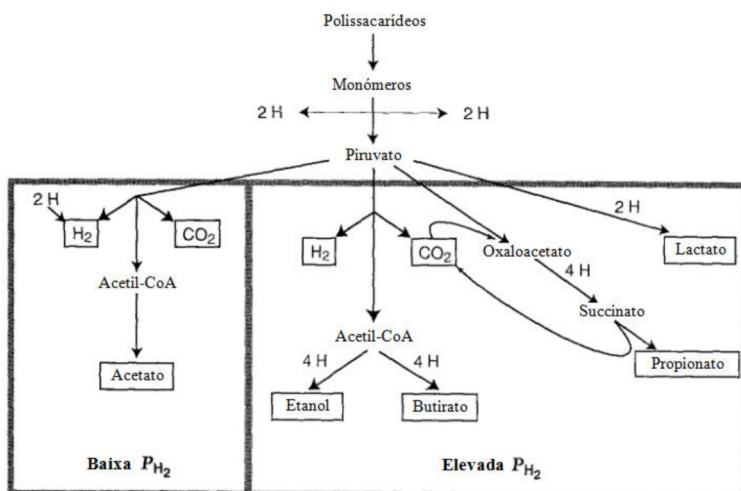


Figura 19 - Produtos da Acidogênese mediante a pressão parcial de H₂. (Fonte: Silveira, 2009)

Hwang et al. (2004) referem que nos digestores anaeróbicos em boas condições operacionais, a concentração de H₂ é normalmente muito baixa devido à enorme capacidade de utilização de H₂ por parte das bactérias utilizadoras de hidrogénio, seguindo principalmente a via metabólica que resulta na produção de acetato. A via metabólica de produção de produtos intermédios (AGV, álcoois e lactato) vai assumindo preponderância à medida que se verifica um aumento da pressão parcial de H₂ derivado, por exemplo, a uma redução do valor do pH (<6). Uma redução do valor do pH pode levar à inibição do processo fermentativo, se atingir um valor de pH igual ou inferior a 4.

O efeito de concentrações crescentes de substrato tem um efeito análogo ao do abaixamento do pH.

Acetogénese

De acordo com McInerney (1999), na acetogénese as bactérias acetogénicas (BA) procedem à degradação dos compostos formados durante a fase anterior (acidogénese), como os AGV, álcoois, alguns aminoácidos e compostos aromáticos. Os produtos resultantes da degradação destes compostos são H₂, formato, CO₂ e acetato, que vão servir de substrato para as BM.

Schink (1997) também refere que mesmo em condições ideais (baixas concentrações de acetato e de H_2), a quantidade de energia disponível para suportar a formação de ATP e o crescimento microbiano é muito baixa. Esta é provavelmente a razão que explica as baixas taxas de crescimento verificadas nos consórcios sintróficos. A temperatura é outro fator que influencia a termodinâmica das reações de acetogênese. A formação de H_2 a partir da oxidação de ácidos orgânicos é energeticamente mais favorável a temperaturas mais elevadas, enquanto o consumo de H_2 pelas BM é menos favorável a temperaturas superiores. No entanto, a velocidade de degradação dos ácidos orgânicos aumenta normalmente com temperaturas mais elevadas devido ao aumento da atividade bacteriana, da sua taxa de crescimento e da degradação do substrato.

Metanogénese

Em conformidade com Silveira (2009) as BM constituem um grupo diverso de microrganismos que obtêm energia para o crescimento a partir de reações que conduzem à produção de metano. Durante a metanogénese, o acetato, H_2 , CO_2 , formato, metanol e metilamina são convertidos em CH_4 e CO_2 pelas BM. As BM estão dependentes do bom funcionamento de todas as fases anteriores para o fornecimento dos substratos adequados para a sua atividade.

As BM dividem-se em 3 grupos tróficos que apresentam vias metabólicas diferentes:

- as BMA, que convertem o acetato em metano;
- as BMH, que convertem o H_2 e CO_2 em metano;
- as BHA, que realizam a inter-conversão entre o acetato, H_2 e CO_2 .

Para Fey e Conrad (2000) cerca de 60-90% do metano produzido em digestores anaeróbios resulta da degradação do acetato pelas BMA. A incapacidade das BMA em degradar o acetato pode resultar numa descida do pH e na paragem de todo o processo anaeróbio. Em

condições de temperatura elevada ($>30^{\circ}\text{C}$), as BHA ganham competitividade sobre as BMA na oxidação do acetato em H_2 e CO_2 . A atividade das BHA está dependente da manutenção de uma baixa concentração de acetato e da baixa pressão de H_2 .

A metanogénese é afetada pelas condições de operação do reator como a temperatura, caudal de alimentação, carga orgânica na alimentação e composição da alimentação, afetando o equilíbrio entre BMA, BMH e BHA.

Sulfidogénese

De acordo com Lens et al. (1998) para efluentes com concentrações elevadas de sulfato (SO_4^{2-}) ou sulfito (SO_3^{2-}) ou outras formas oxidadas de enxofre pode ocorrer outro processo em simultâneo com estas quatro fases de DA que pode ser designado por Sulfidogénese, no qual as bactérias redutoras de sulfato (BRS) reduzem estes compostos convertendo-os em sulfuretos.

A produção de sulfuretos durante a sulfidogénese pode apresentar alguns problemas técnicos no funcionamento do processo anaeróbio, como por exemplo:

- A presença de concentrações elevadas de sulfuretos apresentar efeitos tóxicos na atividade das BM, BA e BRS;
- Parte da matéria orgânica da água residual será utilizada pelas BRS, reduzindo, consequentemente, o potencial de produção de metano por parte das BM.
- Para além disso a qualidade do biogás produzido fica reduzida devido ao aumento da fração de H_2S resultante da produção de sulfuretos na fase líquida. A utilização do biogás fica, deste modo, dependente da remoção da fração de H_2S .
- Os sulfuretos, em concentrações elevadas, apresentam um odor desagradável e provocam problemas de corrosão em condutas, motores e queimadores. A minimização destes efeitos requer custos acrescidos de manutenção e de investimento.

- Parte dos sulfuretos estão presentes no efluente do reator anaeróbio, provocando uma diminuição da eficiência do tratamento, uma vez que os sulfuretos contribuem para o CQO da água residual. Um tratamento suplementar subsequente torna-se um requisito obrigatório de modo a remover os sulfuretos.

No entanto, a ocorrência do processo de sulfidogênese apresenta algumas vantagens como a remoção dos compostos oxidados de enxofre (sulfato, sulfito e tiosulfato) do efluente e a remoção de metais pesados por precipitação com os sulfuretos, reduzindo o seu potencial tóxico.

3.2. Fatores que interferem no processo de DA

Metcalf e Eddy (1991) referem que em todos os processos de tratamento biológicos de águas residuais, a remoção eficaz de poluentes e contaminantes depende, não só do potencial metabólico dos microrganismos, mas também da existência de condições ambientais adequadas que permitam a sua atividade.

Temperatura

É referido em Angelidaki e Sanders (2004) que o controlo da temperatura é crucial para o bom funcionamento do processo, uma vez que este parâmetro influencia a atividade biológica dos microrganismos e, conseqüentemente, a velocidade a que ocorre o processo de digestão e de biodegradação dos compostos.

A manutenção de uma temperatura constante no digestor é crucial para o bom funcionamento do processo anaeróbio, uma vez que flutuações pontuais e/ou frequentes na temperatura do digestor podem provocar efeitos severos nas bactérias, particularmente nas metanogénicas. Para além disso, a temperatura tem um efeito significativo na pressão parcial de H_2 , influenciando a cinética do metabolismo sintrófico.

Existem diferentes gamas de temperatura nas quais o processo de DA pode ocorrer, sendo conhecidos três regimes de temperaturas que podem ser aplicados, conforme exemplificado na seguinte tabela:

Tabela 16 – Regimes de temperatura do processo anaeróbio. (Fonte: Metcalf e Eddy, 1991)

| Regime | Temperatura, °C | |
|------------|-----------------|-----------------|
| | Gama | Intervalo Ótimo |
| Psicrófilo | -10 – 30 | 12 – 18 |
| Mesófilo | 20 – 50 | 25 – 40 |
| Termófilo | 35 – 75 | 55 – 65 |

Tempo de Retenção Hidráulica (TRH)

De acordo com Metcalf e Eddy (1991) o TRH é um parâmetro fundamental para o dimensionamento de um sistema de tratamento anaeróbio e exprime o tempo médio que um determinado volume de lamas permanece no digestor anaeróbio. Para um reator com um volume V (m^3), sujeito a uma alimentação com um caudal Q (m^3/dia), o TRH é obtido pela seguinte expressão:

$$TRH = \frac{V}{Q} \text{ (dias)}$$

Fórmula 1 – Tempo de Retenção Hidráulica (TRH)

O TRH deve ser sempre superior ao tempo de duplicação das bactérias de crescimento lento, nomeadamente, das BM, de forma a garantir a degradação da matéria orgânica com a permanência de um quantitativo suficiente e equilibrado de microrganismos no digestor. A maximização do TRH é preferível para uma maior estabilidade do processo e a minimização da produção de lamas.

PH e alcalinidade

Em Hwang et al. (2004) é descrito o efeito do pH na atividade enzimática dos microrganismos. Cada enzima apresenta atividade para uma determinada gama de pH, exibindo atividade máxima a um valor de pH ótimo. Cada grupo de microrganismos possui uma gama diferente de pH ótimo. As BM atuam num intervalo estreito de pH (entre 5,5-8,5), com um pH ótimo entre 6,5 e 8. As BF atuam numa gama de pH mais ampla, suportando valores entre 4 e 8,5.

O valor do pH afeta o equilíbrio ácido-base dos compostos no digestor. Para valores de pH baixos, os AGV livres podem causar inibição por ácidos fracos, enquanto para valores de pH elevados, a amônia livre pode causar inibição por bases fracas. Em digestores anaeróbios de cultura mista, a gama ótima de pH situa-se entre 6,6 e 7,4.

A capacidade tampão do líquido em digestão é um parâmetro importante para a estabilidade do processo. A capacidade tampão pode ser definida como a resistência de uma solução a mudanças de pH.

Agitação

Em conformidade com Appels et al. (2008) a agitação adequada do digestor é fundamental para o funcionamento ótimo do processo de DA. A agitação tem a função de melhorar o contato entre o substrato do afluente com a biomassa no interior do digestor e possibilita condições uniformes de temperatura e de concentração de substrato. Para além disso, a agitação evita a formação de camadas de espuma superficiais e a deposição de lama no fundo do digestor.

Os métodos de agitação auxiliares utilizados vulgarmente são a recirculação exterior das lamas, agitação mecânica interna ou injeção interna do biogás produzido. (Igoni et al., 2008).

Nutrientes (razão carbono/azoto)

Conforme referido em Mara e Horan (2003), a concentração de nutrientes no digestor tem uma importância crucial no desempenho do processo, por serem essenciais para assegurar o crescimento microbiano e a síntese de enzimas e cofatores indispensáveis para as reações bioquímicas e metabólicas. Os nutrientes necessários em maior quantidade (macronutrientes), durante o processo anaeróbio, são o azoto, o fósforo e o enxofre. As concentrações de carbono e azoto determinam, muitas vezes, o desempenho do processo de DA, uma vez que um destes elementos constitui normalmente um fator limitante. No processo de DA o carbono constitui a fonte de energia para os microrganismos enquanto o azoto estimula o crescimento microbiano. Se a concentração de azoto for limitante, as populações microbianas permanecem em número reduzido, demorando mais tempo a decomposição do carbono disponível.

As bactérias utilizam o carbono disponível no meio cerca de 30 a 35 vezes mais rapidamente comparativamente à taxa com que convertem o azoto. Por esta razão, para a operação ótima do digestor, a razão carbono/azoto deve ser da ordem de 30:1 (Igoni et al., 2008)

Toxicidade e inibição

De acordo com Mara e Horan (2003) existem diversos compostos que potencialmente podem ser inibitórios ao processo anaeróbio, como alguns compostos que fazem parte do substrato no afluente, ou subprodutos resultantes do metabolismo das bactérias. Estes compostos podem abrandar o processo de degradação (toxicidade) ou causar a paragem do processo (inibição). Alguns compostos responsáveis por estes fenómenos são a amónia, os sulfuretos, os AGV, os metais pesados e os metais alcalinos e alcalino-terrosos.

Amónia

Em Gerardi (2003) é indicado que a amónia é um composto libertado durante a hidrólise através da degradação biológica da matéria azotada que se encontra presente essencialmente na forma de proteínas e ureia. Apesar de ser essencial para o crescimento bacteriano, a amónia pode inibir o processo de DA quando presente em concentrações elevadas. Os valores da concentração de amónia a partir dos quais se verifica inibição do processo variam, essencialmente, em função dos tipos de substrato e das condições ambientais (temperatura e pH).

O ião amónio (NH_4^+) e a amónia livre (NH_3) são as duas formas predominantes de azoto inorgânico presente no líquido em digestão. As quantidades de formas reduzidas de nitrogénio – NH_4^+ e NH_3 – em digestores anaeróbios estão dependentes do valor de pH. Um aumento de pH resulta no aumento da produção de amónia (NH_3), enquanto a diminuição de pH resulta na produção de mais iões de amoníaco (NH_4^+). A toxicidade devido à amónia pode ser evitada se o valor do pH se mantiver num intervalo ótimo entre 6,8 e 7,2.

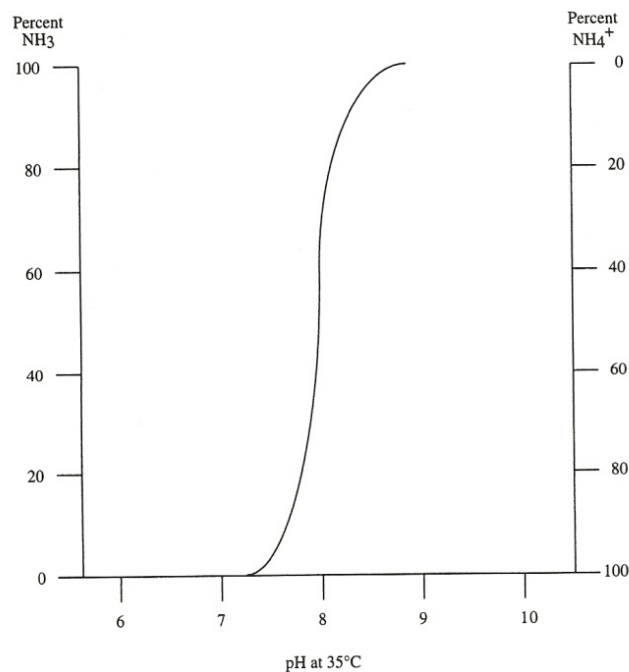


Figura 20 – Percentagem de NH_3 e NH_4^+ em função do pH. (Fonte: Gerardi, 2003)

Sulfuretos

Silveira (2009) refere que em termos energéticos, o sulfureto é a forma mais estável do enxofre em ambiente anaeróbio. O sulfureto é altamente reativo, corrosivo e tóxico para microrganismos, plantas, animais e para o Homem. Consequentemente, os problemas associados ao tratamento de águas residuais com uma concentração elevada em sulfatos derivam essencialmente da produção de sulfuretos pelas BRS. A inibição causada pela redução do sulfato a sulfuretos ocorre em 2 fases:

- Inibição primária – devido à competição das BRS pelo mesmo substrato orgânico e inorgânico, acabando por suprimir a produção de metano
- Inibição secundária – resultante da toxicidade dos sulfuretos produzidos para diversos grupos de bactérias

O resultado da competição entre as BRS e os outros microrganismos anaeróbios determina a concentração de sulfuretos no reator. Os sulfuretos são tóxicos tanto para as BM como para as BRS.

Os sulfuretos produzidos em reatores anaeróbios encontram-se distribuídos em várias formas distintas: S_2^- , HS^- e H_2S em solução, H_2S no biogás e ainda sulfuretos metálicos insolúveis de acordo com o equilíbrio químico e físico.

Ácidos Gordos Voláteis (AGV)

Conforme indicado em Gerardi (2003), os AGV são os principais produtos intermédios formados durante o processo de DA. Para valores baixos de pH, os AGV tornam-se tóxicos devido ao aumento da fração não dissociada. Para além do pH, a concentração inibitória de AGV depende também do poder tampão no reator.

Concentrações elevadas de AGV no líquido em digestão são resultado de perturbações que se verificam no processo, nomeadamente, variações na temperatura, sobrecarga orgânica

ou presença de compostos tóxicos. Nestas situações, as BM são incapazes de remover o hidrogénio e os AGV produzidos pelas BF e BA. Devido a esta situação, ocorre a acumulação de ácidos no digestor que leva a uma descida do pH a níveis que podem ser inibitórios para as BM. A acumulação de ácidos no digestor pode resultar na paragem de todo o processo.

Metais pesados

Segundo Zayed e Winter (2000), os metais pesados podem encontrar-se nos efluentes industriais e, eventualmente, em quantidades significativas. Os principais metais pesados cuja presença no esgoto é mais comum, podendo tornar-se problemática a partir de determinadas concentrações, são o crómio, o ferro, o cobalto, o cobre, o zinco, o cádmio e o níquel. Uma característica particular dos metais pesados reside no facto de, ao contrário das outras substâncias tóxicas, estes não serem biodegradáveis, podendo acumular-se até atingirem concentrações potencialmente tóxicas. O efeito tóxico derivado da presença de metais pesados é atribuído à disrupção do funcionamento e estrutura das enzimas.

O efeito de estimulação ou inibição dos microrganismos anaeróbios é determinado pela concentração total de metais, as formas químicas em que se encontram, pH e potencial redox. A solubilidade dos metais aumenta para valores de pH mais reduzidos. A toxicidade causada por metais pesados está mais correlacionada com a concentração iónica na forma livre do metal do que com a sua concentração total. De um modo geral as BF e BA são mais resistentes à toxicidade causada por metais pesados do que as BM. A quantidade de sólidos presentes no digestor é um fator que tem muita influência na toxicidade relacionada com metais pesados. Quanto maior o teor em sólidos menor a susceptibilidade do sistema entrar em inibição devido à concentração de metais pesados.

Metais alcalinos e alcalino-terrosos

De acordo com Silveira (2009) a adição excessiva de sais ao digestor, com o objetivo de controlar o valor de pH, pode resultar na inibição do processo anaeróbio. Os metais alcalinos e alcalino-terrosos são benéficos ao processo anaeróbio, estimulando as bactérias. No entanto, quando presentes em concentrações excessivas tornam-se inibitórios. A concentração em nutrientes, o período de aclimação e a ocorrência de sinergismo e de antagonismo são os fatores que influenciam a inibição provocada por iões. A utilização de magnésio, ou o efeito combinado dos iões cálcio e potássio, pode ser muito eficaz na redução da toxicidade provocada pelo sódio, enquanto o efeito combinado do cálcio e do sódio pode ser eficaz na redução da toxicidade provocada pelo potássio.

3.3. Tipos de Reatores Anaeróbios

Os reatores anaeróbios convencionais podem operar em modo descontínuo, contínuo ou semi-contínuo.

No processo descontínuo, a matéria orgânica é introduzida na sua totalidade no interior do reator, ficando retida durante um determinado período de tempo até ao final do processo de degradação. No processo contínuo a alimentação é introduzida de forma contínua enquanto na semi-contínua é introduzida de modo intermitente. A operação em modo contínuo e semi-contínuo são preferíveis por permitirem o crescimento constante dos microorganismos através do controlo do caudal de alimentação.

A escolha do tipo de reator é determinada pelas características do resíduo. O quadro seguinte evidencia as principais diferenças operacionais dos vários reatores que podem ser utilizados num processo de DA.

Tabela 17 – Diferentes Tipos de Reactores. (Fonte: Mara e Horan, 2003)

| Tipo de reactor | Alimentação (kg CQO/m ³ .dia) | TRH (h) | Remoção CQO (%) |
|--|---|------------|--------------------|
| Reactor de mistura completa | 1-5 | 240-360 | 60-80 |
| Reactor anaeróbio de contacto | 1-6 | 24-120 | 70-95 |
| Reactor anaeróbio descontínuo sequencial | 1-10 | 6-24 | 75-90 |
| Reactor de filtro anaeróbio | 2-15 | 10-85 | 80-95 |
| Reactor de leito fluidizado | 2-50 | 1-4 | 80-90 |
| Reactor UASB | 2-30 | 2-72 | 80-95 |
| Reactor com anteparas | 3-35 | 75-95 | |
| Digestão anaeróbia em 2 fases | 5-30 | 20-150 | 70-85 |

De referir que, normalmente, apenas se opta pela inclusão de DA em ETARs que abrangem populações iguais ou superiores a 50.000 hab.eq, devido ao elevado custo de exploração deste processo. Em seguida, encontram-se esquematizadas as principais configurações de digestores anaeróbios e descritas as suas principais características:

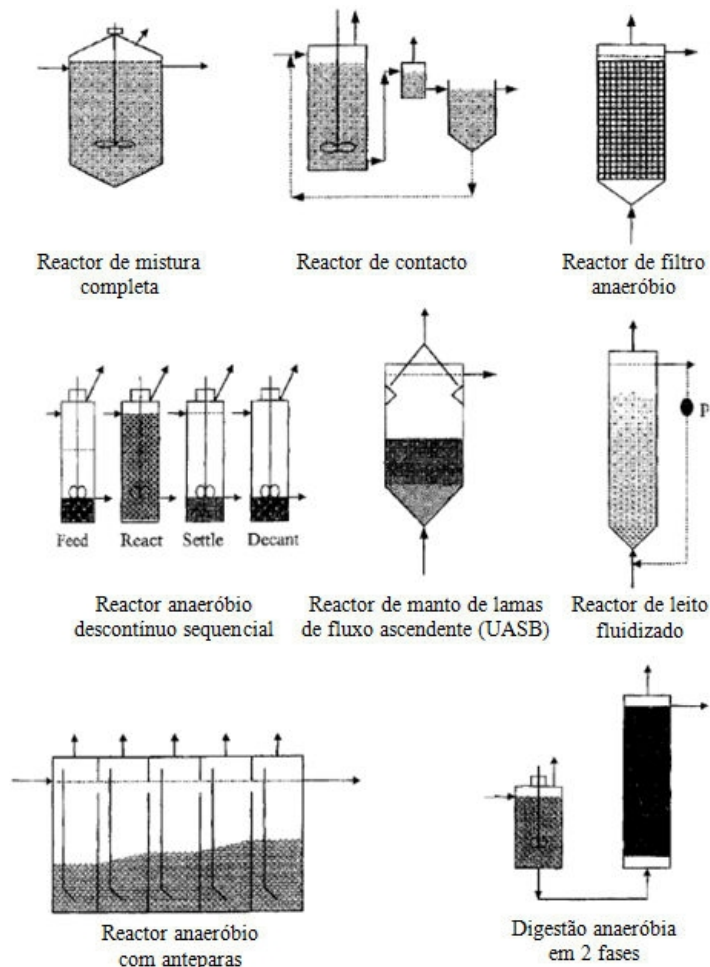


Figura 21 – Representação esquemática dos diferentes tipos de digestores. (Fonte: Silveira, 2009)

Em Mara e Horan (2003) são descritos os diversos tipos de reatores existentes, mencionados em seguida:

Reatores de mistura completa

A alimentação deste tipo de reatores pode ser realizada de modo contínuo ou intermitente e a lama no seu interior está sujeita a agitação.

Nestes reatores, a biomassa encontra-se suspensa no líquido em digestão e será removida em conjunto com o efluente, pelo que o TRS é igual ao TRH. Como tal, a sua operação está dependente do crescimento contínuo de nova biomassa, de modo a substituir a biomassa que se perde pelo efluente. O crescimento das bactérias anaeróbias é bastante lento, particularmente algumas estirpes de BM que possuem um tempo de duplicação de 5 a 10 dias. Por este facto, o TRH deste tipo de digestores nunca deve ser inferior a 10 dias, de forma a assegurar um tratamento eficaz e um quantitativo populacional adequado no interior do digestor.

Reator de contato

O fator diferenciador deste tipo de reator reside na aplicação de um segundo órgão onde a lama é decantada, para posteriormente ser reenviada para o primeiro reator, permitindo um maior período de contato da biomassa com o substrato.

Devido à recirculação da lama o TRS é superior ao TRH, contribuindo para uma maior eficiência da digestão comparativamente ao digestor completamente agitado. A agitação adequada destes digestores é essencial para que a transferência de massa entre a alimentação e a biomassa ativa decorra de modo eficaz.

A maior desvantagem deste tipo de digestores resulta do facto da sua operacionalidade estar dependente do bom funcionamento do reator de sedimentação, que por vezes está

condicionado devido ao crescimento de bactérias filamentosas e à produção de biogás que dificultam o processo de sedimentação.

Reator de filtro anaeróbio

O filtro anaeróbio é um processo onde é utilizado um filme biológico fixo que providencia uma superfície de aderência que suporta os microrganismos. O filme biológico formado pelas bactérias é sustentado pela assimilação dos nutrientes e dos poluentes da água residual que circula em fluxo ascendente. Quando este filme se desprende pode sair com o efluente ou sedimentar no fundo do reator. Os filtros anaeróbios foram o primeiro sistema implementado que eliminou a necessidade de se proceder à separação de sólidos e reciclagem, possibilitando TRS bastante superiores ao TRH.

A principal desvantagem deste sistema reside no facto dos custos associados a esta tecnologia serem consideravelmente elevados.

Reator anaeróbio descontínuo sequencial

Este tipo de digester opera segundo um ciclo sequencial constituído por 4 fases: alimentação; reação; sedimentação; decantação. Todas estas fases ocorrem num único digester. A alimentação e decantação processam-se em modo descontínuo, enquanto o crescimento se processa num sistema em suspensão. No entanto, uma parte significativa do tempo despendido neste ciclo está associada à sedimentação da biomassa, pelo que o volume do reator tem que ser superior, quando comparado com o dos reatores que operam continuamente. Esta desvantagem é compensada pela sua enorme simplicidade, não sendo necessária uma fase subsequente de sedimentação nem de reciclagem da biomassa.

Reator de manto de lamas de fluxo ascendente (UASB)

É a tecnologia usada com maior frequência para o tratamento de águas residuais domésticas e industriais (Van Haandel e Lettinga, 1994).

Nos sistemas UASB, o resíduo líquido apresenta fluxo ascendente através de um manto espesso de lama granular anaeróbia. O substrato entra pela parte inferior do reator entrando em contato com o leito de lamas sedimentadas. A parte superior do reator possui um sistema de separação dos estados sólido, líquido e gasoso. No interior do reator existem zonas separadas específicas para as diferentes fases da digestão, sedimentação e separação gás/líquido. A mistura do conteúdo do reator é proporcionada pela produção de metano no interior do manto e pelo fluxo hidráulico. (McInerney, 1999).

A lama granular contém concentrações elevadas de bactérias imobilizadas que apresentam propriedades que propiciam a colonização. O elevado potencial de sedimentação dos grânulos de lama permite uma retenção superior de biomassa no interior do reator comparativamente a outros sistemas de tratamento. Isto permite que este sistema atinja uma taxa de alimentação orgânica e concentrações de biomassa bastante elevadas sem que sejam necessários materiais de suporte permitindo uma redução de custos significativa.

Reator de leito fludizado

Este reator surgiu da tentativa de solucionar as dificuldades associadas à separação da biomassa em reatores de mistura completa ou de contato. Para solucionar este problema, pretendia-se acumular por aderência a maior quantidade de biomassa ativa possível e os sólidos suspensos mais finos.

A operação é realizada em fluxo ascendente de modo a permitir a fluidização do meio. A biomassa adere à superfície de partículas de reduzidas dimensões que são mantidas em suspensão pelo fluxo ascendente. Com este tipo de reator atingiu-se uma maior atividade

específica da biomassa por volume do reator, maximizando a área superficial disponível para a aderência das bactérias, minimizando o volume ocupado pelo meio.

O maior risco durante a operação do reator de leito fluidizado está na perda de partículas de biomassa do interior do reator derivado a alterações bruscas na densidade das partículas, no caudal ou na produção de biogás.

Reator anaeróbio com anteparas

Este tipo de reator é constituído por uma série de anteparas verticais alternadas por entre as quais ocorre o fluxo de água residual. É designado também por reator de fluxo-pistão.

As principais vantagens associadas à utilização deste tipo de reatores são: a elevada estabilidade e fiabilidade devido à aderência dos sólidos no meio filtrante, funcionando até a própria biomassa como meio de suporte para a aderência de microrganismos; a disposição das anteparas permite reduzir a perda de biomassa pelo efluente e recuperar rapidamente de choques hidráulicos ou choques de carga orgânica; não necessita de equipamento extra para separação de lamas e biogás; possibilita o tratamento de praticamente todo o tipo de águas residuais, de baixa a elevada carga orgânica.

Digestão anaeróbia em duas fases

A aplicação deste processo implica a utilização de dois reatores separadamente, um para a hidrólise e acidogénese e outro para a acetogénese e metanogénese. Estes reatores encontram-se dispostos em série, permitindo a otimização independente de cada fase do processo de digestão. Os benefícios estão associados ao facto dos microrganismos acidogénicos e metanogénicos possuírem diferentes requisitos nutricionais, características fisiológicas, pH ótimo e cinética de crescimento.

As principais vantagens associadas a este processo são o maior controlo dos processos, a

possibilidade de eliminação da biomassa acidogénea em excesso evitando deste modo a inibição da metanogénese, a composição do biogás com maior percentagem de metano. Como principais desvantagens temos a disrupção das relações sintróficas, a elevada acumulação de lamas no primeiro reator e uma operação mais complexa.

3.4. A Digestão Anaeróbia da ETAR Sul da SIMRIA

Os digestores da ETAR Sul da SIMRIA, objeto deste estudo, são reatores de mistura completa, com alimentação contínua, agitação por injeção interna do biogás produzido e recirculação exterior das lamas para aquecimento. A utilização deste tipo de digestores requer volumes elevados, devido ao elevado TRH necessário para o correto funcionamento do digestor. Os digestores concebidos para esta ETAR possuem um volume de 3000 m³, operando com TRH superiores a 15 dias. Este tipo de reatores, conforme já referido atrás, é particularmente interessante para efluentes com elevado teor de sólidos e de matéria orgânica, como é o caso das lamas domésticas, por permitirem a reposição de microrganismos no interior do digestor. No entanto, a perda de biomassa durante a sua operação impossibilita a sua utilização para a maioria dos efluentes industriais, sendo que na ETAR Sul o volume deste tipo de efluente não é significativo.

Em relação à temperatura, este digestor opera em regime mesófilo. Da experiência na operação deste digestor, ao longo do tempo, resulta que, a temperatura de funcionamento ótima, deverá encontrar-se entre os 30 e os 35°C, enquadrando-se no intervalo ótimo (mais alargado) definido por Metcalf e Eddy (1991).

4. REDES NEURONAIAS ARTIFICIAIS

De acordo com diversos autores, entre os quais Silva (2013), as Redes Neurais Artificiais (RNA), ou simplesmente Redes Neurais, são modelos matemáticos ou ferramentas computacionais que operam de forma análoga aos processos biológicos de um cérebro e possuem capacidade para entender relações complexas existentes num determinado conjunto de dados.

Em Moreira (1997), é indicada a seguinte comparação: o sistema nervoso central dos animais recebe informação do exterior, armazena e processa essa informação e transmite uma resposta/reação para o exterior. A observação do desempenho que este apresenta tem revelado uma extraordinária capacidade para executar rápida e eficientemente tarefas de grande complexidade, tais como, o processamento em paralelo da informação, a memória associativa e a capacidade para classificar e generalizar conceitos. Estes factos têm servido de motivação quer para o estudo detalhado da constituição do cérebro quer para a sua mimetização através da conceção de sistemas com as capacidades atrás referidas. Por outro lado, esses sistemas têm conhecido diversas designações, tais como, redes neuronais, modelos conexionistas, processadores distribuídos em paralelo, neurocomputadores, sistemas inteligentes naturais ou máquinas de algoritmos de aprendizagem (Klerfors, 1998).

Em conformidade com Lingireddy e Brion (2005), na tentativa de compreender o funcionamento do cérebro humano, matemáticos e cientistas desenvolveram modelos matemáticos capazes de imitar a capacidade de generalização e reconhecimento de padrões do cérebro humano. Estes modelos matemáticos, quando devidamente treinados, são capazes de compreender padrões subjacentes a um determinado conjunto de dados de aprendizagem e podem generalizar e fornecer informação fiável e precisa em relação aos dados que não foram vistos anteriormente pelo modelo. Este processo é semelhante à

maneira como o cérebro humano assimila e regista determinados padrões na sua memória e reconhece novos padrões que são semelhantes aos padrões anteriormente observados.

A utilidade dos modelos de redes neuronais reside no facto de poderem ser usadas para inferir uma função não-linear apenas a partir de observações. Isto é particularmente útil em aplicações nas quais a complexidade dos dados ou das tarefas a serem processados torna a modelagem de tal função impraticável.

De acordo com Matos (2008) citando Haykin (1999) as RNA podem ser definidas da seguinte forma:

“Uma rede neuronal artificial é um processador massivamente distribuído em paralelo constituído por unidades de processamento simples, que tem a propensão natural para armazenar conhecimento experimental e torná-lo disponível para ser utilizado. Assemelha-se ao cérebro em dois aspetos:

- 1. O conhecimento é adquirido pela rede a partir do meio envolvente através de um processo de aprendizagem;*
- 2. As intensidades das ligações entre neurónios, conhecidas como pesos sinápticos, são usadas para armazenar o conhecimento adquirido.”*

4.1. Analogia com o modelo biológico

O córtex cerebral é constituído por unidades celulares independentes designadas por neurónios que podem comunicar entre si através das chamadas ligações sinápticas ou sinapses (Moreira, 1997). Todas as sensações, pensamentos e sentimentos são resultado de sinais que atravessam e são processadas pelos neurónios (Lingireddy e Brion, 2005).

O neurónio é a unidade base do sistema nervoso. Os neurónios diferem segundo as suas funções e localização, contudo, podemos afirmar que o neurónio típico apresenta três componentes: o corpo celular ou soma, que inclui o núcleo celular, as dendrites e o axónio (Soares, 2013).

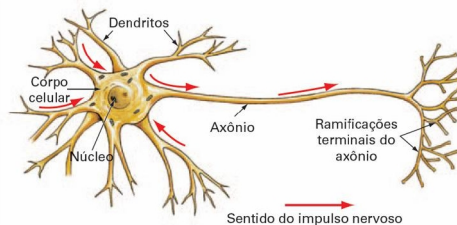


Figura 22 – Representação de um neurónio biológico.

Para Brondino (1999) os neurónios apresentam uma estrutura simples e cada um permite efetuar três funções básicas: entrada, processamento e saída de sinais nervosos. As dendrites são conexões através das quais os neurónios recebem os estímulos nervosos de outros neurónios ou órgãos sensoriais que depois são processados no núcleo do neurónio. Estes impulsos passam através do corpo da célula e são transmitidos por um único axónio para outro neurónio ou outros órgãos. A região na qual o sinal passa do axónio de um neurónio (pré-sináptico) para a dendrite de outro neurónio (pós-sináptico) é designada por sinapse. O tamanho de uma sinapse altera-se em função da frequência com que os sinais passam através dessa área. O tamanho da sinapse controla a intensidade do sinal transmitido aos neurónios adjacentes. As sinapses não são ligações físicas, mas sim ligações químicas temporárias que acontecem por intermédio de substâncias denominadas neurotransmissores.

A figura apresenta o diagrama da ligação entre dois neurónios:

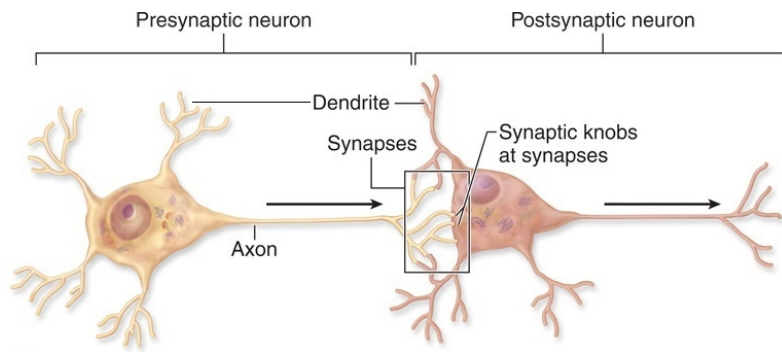


Figura 23 – Representação esquemática de uma sinapse biológica

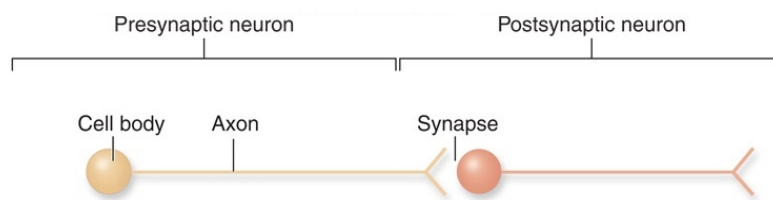


Figura 24 – Representação simplificada de uma sinapse

A comunicação é realizada unidireccionalmente por sinais nervosos elétricos (predominantemente no interior de cada neurónio) e químicos (nas regiões terminais das ligações sinápticas) de uma forma descontínua, isto é, por impulsos. Estes últimos são desencadeados em cada neurónio sempre que um certo potencial de ativação é ultrapassado em resultado da recepção de um ou mais sinais nervosos nas sinapses. De referir que os sinais nervosos são, em geral, amplificados (ou pesados) de forma diferenciada ao atravessar as diferentes sinapses de um neurónio (Moreira, 1997).

“Quando os neurónios disparam, propaga-se uma corrente elétrica a partir do corpo celular e ao longo do axónio. Quando esta corrente atinge a sinapse, desencadeia a libertação de substâncias químicas conhecidas como neurotransmissores. Num neurónio excitatório, a interação cooperativa de outros neurónios cujas sinapses são adjacentes determina se o neurónio seguinte disparará ou não, isto é, se produzirá o seu próprio potencial de ativação que conduzirá à libertação dos neurotransmissores e assim sucessivamente. (...)

Em média cada neurónio forma cerca de mil sinapses. Se pensarmos que existem mais de

10 bilhões de neurónios e mais de 10 triliões de sinapses, podemos concluir que cada neurónio tende a relacionar-se com uns tantos outros, mas nunca com todos ou com quase todos.” (António Damásio, 2013)

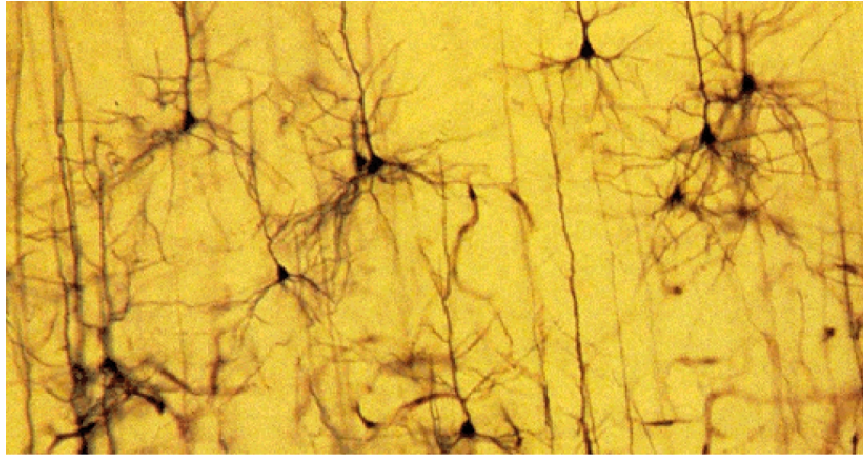


Figura 25 – Rede Neuronal Biológica. (Fonte: www.auladeanatomia.com, 2013)

Segundo Matos (2008), quando em 1943 McCulloch e Pitts conceberam o seu modelo computacional de um neurónio biológico e lançaram as fundações das RNA, pretendiam desenvolver um sistema que permitisse imitar o funcionamento do cérebro humano. O desenvolvimento destes sistemas foi acompanhado pelo progresso do conhecimento na área da neurologia, o que contribuiu para a criação de modelos com características semelhantes aos sistemas biológicos.

Assim como o sistema nervoso, a estrutura de uma rede neuronal é composta por unidades conectadas designadas como nós ou neurónios, em que cada um recebe valores de *inputs*, efetua cálculos relativamente simples com os mesmos e gera um *output*. Então, o valor de *output* de um neurónio é transmitido como um dos *inputs* de outro neurónio ou como *output* final de todo o sistema. No seu conjunto, uma RNA é um sistema que recebe *inputs*, computa-os e devolve um ou mais *outputs*. (Palisade, 2010)

A arquitetura típica das redes neuronais apresenta diferentes camadas, nomeadamente, a camada de entrada ou de *inputs*, a camada de saída ou de *output* e, pelo menos, uma camada intermédia ou oculta.

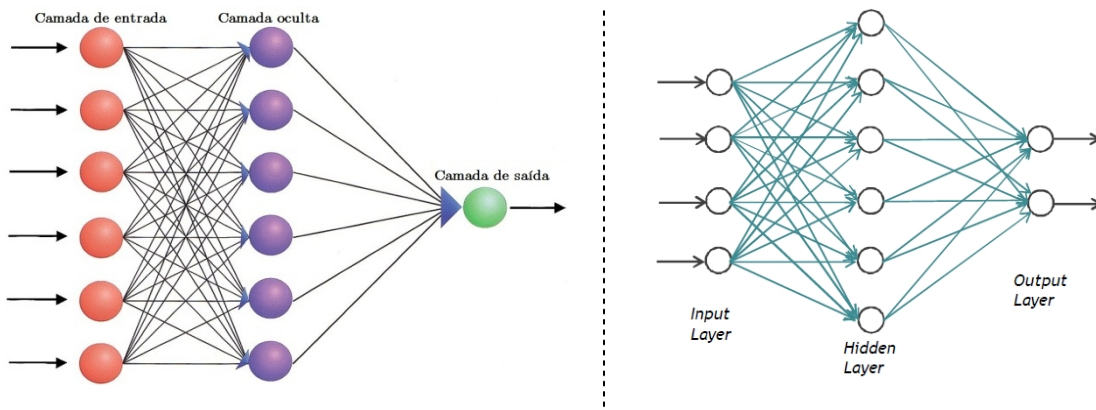


Figura 26 – Representação de duas topologias de redes

Da mesma forma que a aprendizagem no cérebro humano ocorre mediante a apresentação repetitiva de dados, que permitem perceber padrões complexos associados a diversos fenômenos, uma RNA também é treinada através da apresentação repetitiva de dados, ajustando a intensidade das ligações sinápticas dos neurónios das diferentes camadas. A intensidade da ligação sináptica é referida como um peso da conexão, que é simplesmente um valor numérico. Começando com um conjunto de pesos de conexão determinados aleatoriamente, o processo de treino ajusta os pesos de forma a que as previsões do modelo se aproximem dos valores reais.

O cérebro humano tem demonstrado que não existe praticamente nada que seja impossível de aprender! Contudo, a aprendizagem ocorre em neurónios simples, mas bastante interligados, e nas respectivas sinapses. De modo análogo, a aprendizagem nas RNAs ocorre em nós simples, mas altamente interligados, que representam os neurónios e as conexões associadas representam as sinapses. As funções simples, mas altamente interligadas, e os respetivos pesos das ligações permitem às RNAs modelar diversos fenômenos complexos sem introduzir erros desnecessários. (Lingireddy e Brion, 2005)

Os neurónios funcionam em grupos denominados por redes, sendo que cada rede pode ser formada por milhares de neurónios interconectados. Portanto, o cérebro pode ser visto como um conjunto de redes neuronais. (Brondino, 1999)

As RNA têm verificado uma evolução impressionante nos últimos anos com aplicações de

sucesso nas mais diversas áreas. No entanto, as formas matemáticas das redes neuronais são muito mais simples do que as suas homólogas biológicas. Além disso, nem todas as redes neuronais artificiais são modelos de redes neuronais biológicas. O sistema nervoso é uma rede de complexidade admirável, incomparavelmente superior às RNA modernas sob todos os pontos de vista (Matos, 2008).

4.2. Evolução histórica

Como resulta de Moreira (1997), a busca pelo desenvolvimento de ferramentas que imitam o funcionamento de um cérebro humano teve início em 1943 por McCulloch e Pitts com o desenvolvimento do primeiro modelo lógico-matemático de um neurónio artificial, como modelo simplificado de neurónios biológicos. Contudo, Liginreddy e Brion (2005) recordam que este modelo carecia da capacidade de aprendizagem.

De acordo com os mesmos autores, a primeira e melhor regra de aprendizagem foi apresentada por Donald Hebb ficando conhecida como a regra de Hebb ou regra *hebbiana* de aprendizagem. Esta regra é descrita no seu livro “A Organização do Comportamento”, em 1949, e segundo a mesma, se um neurónio recebe uma entrada de outro neurónio e se ambos estão altamente ativos (matematicamente têm o mesmo sinal), o peso sináptico entre os neurónios deve ser reforçado, ou seja, a conexão entre dois neurónios simultaneamente ativos é reforçada pela ativação repetida dessa sinapse aumentando a sua eficiência. Em 1956, Taylor desenvolveu uma rede de memória associativa usando a regra de Hebb.

Matos (2008), refere que na sequência da crescente atividade científica em torno desta área de estudo, em 1958 Rosenblatt concebeu o Perceptrão, criando um algoritmo iterativo de aprendizagem que resulta num filtro com capacidade de adaptação. O Perceptrão é a forma mais simples de uma rede neuronal utilizada para a classificação de padrões

considerados linearmente separáveis, ou seja, padrões que se encontrem em lados opostos de um hiperplano. Basicamente, este modelo consiste num único neurónio com pesos sinápticos e constantes (*bias*) ajustáveis. De acordo com o estabelecido no *Teorema de Convergência do Perceptrão*, este modelo tem a capacidade de encontrar sempre solução, num número de passos finito, desde que os dados correspondam a um conjunto linearmente separável. Na altura este teorema provocou algum entusiasmo entre os investigadores desta área.

Conforme referido por Liginreddy e Brion (2005), pouco tempo depois, em 1960, Widrow e Hoff desenvolveram um algoritmo de treino baseado no erro quadrático médio (EQM) que implementaram na ADALINE (ADAPtative LINear ELEment). Dois anos depois, era proposta por Widrow a MADALINE (Multiple ADALINE), uma das primeiras redes neuronais com vários componentes adaptáveis que tinha a capacidade de aprender. Ao longo dos anos 60 muitos investigadores reportaram descobertas e aplicações bem-sucedidas na área da neurocomputação.

Em conformidade com Brondino (1999), na sequência destes sucessos o progresso das RNA parecia bem encaminhado e o seu campo de aplicação virtualmente ilimitado. No entanto, em 1969, Minsky e Papert provaram que um perceptrão de camada única não consegue executar algumas funções lógicas simples tais como a disjunção exclusiva (ou-exclusivo), demonstrando matematicamente que há limites fundamentais nas operações que perceptrões de camada única podem executar. Embora esta seja apenas uma limitação para perceptrões de camada única, não existia nenhum algoritmo conhecido para perceptrões interligados em múltiplas camadas.

Como mencionado em Liginreddy e Brion (2005), os argumentos convincentes de Minsky e Papert que permitiram constatar severas limitações das redes neuronais, tiveram um efeito desencorajador, levando a que a maioria dos fundos de investigação destinados a esta área

de estudo fossem desviados para outros campos de investigação e a maioria dos investigadores abandonasse este domínio das redes neuronais por mais de uma década. Apenas alguns continuaram a desenvolver estudos neste domínio e as soluções tardaram em aparecer.

De acordo com Matos (2008), apesar de tudo, durante este período, foram desenvolvidos vários trabalhos que contribuíram para algumas descobertas importantes feitas no início dos anos oitenta e que permitiram o renascer das RNA. Uma das descobertas mais importantes aconteceu em 1986 quando Rumelhart, Hinton e Williams publicaram o algoritmo de retropropagação do erro (Back-Propagation) que permitiu ultrapassar as limitações assinaladas por Minsky e é certamente o algoritmo de aprendizagem mais divulgado. A autoria deste algoritmo deve ser partilhada com Werbos (1974), Bryson e Ho (1969), Parker (1985) e LeCun (1985) (Haykin, 1999). A renovação do interesse pelas RNA na década de oitenta originou grandes progressos neste campo, surgindo novas arquiteturas de redes como os mapas auto-organizáveis (SOM) de Kohonen, as redes de base radial (RBR) de Broomhead e Lowe, as redes de Hopfield e as máquinas de apoio vetorial de Vapnik já no início dos anos noventa.

Conforme refere Lobo (2007), foi, de resto, ao longo da década de 90 que se consolidou o recurso às redes neuronais, com um uso generalizado nas mais diversas áreas. Para além de uma utilização isolada desta ferramenta, tem sido muito difundida a aplicação de redes neuronais em conjunto com outros métodos estatísticos ou probabilísticos, complementando-se, de forma a otimizar o funcionamento das RNA.

4.3. Princípios de funcionamento e descrição das RNA

Uma rede neuronal é, por definição, um sistema de elementos de processamento simples, neurónios ou nós, que estão ligados a uma rede por um conjunto de pesos sinápticos. A

rede é definida pela sua arquitetura, pela magnitude dos pesos e pelo modo de operação dos elementos de processamento.

O neurónio é um elemento de processamento que recebe um número de *inputs*, pesa-os, soma-os, adiciona constantes e utiliza o resultado como o argumento para uma função de transferência, que resulta no *output* do neurónio. Os neurónios estão interligados por canais de comunicação (conexões) que transportam dados numéricos e operam apenas com os seus dados locais e com os valores de entrada que recebem através dessas conexões. Todos os nós da rede produzem uma soma ponderada dos seus *inputs*, transformam esse valor através da função de transferência e transmitem um valor de *output*.

De acordo com Matos (2008) o nó é a unidade base de uma RNA. Como tal, é necessário conhecê-lo para compreender o funcionamento da totalidade da rede. Na figura seguinte, sintetiza-se o funcionamento de um nó genérico.

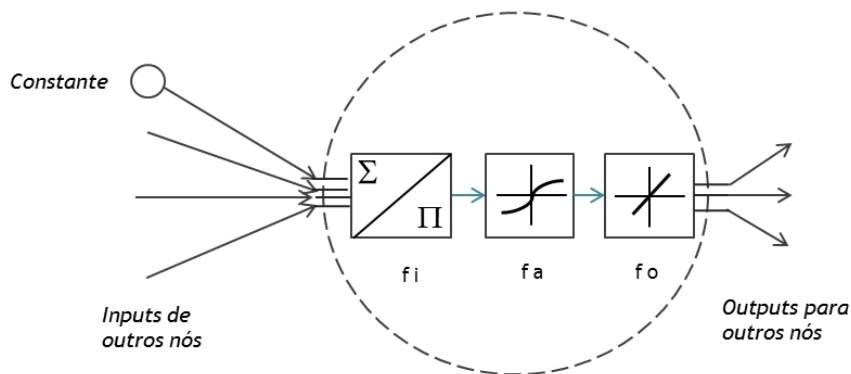


Figura 27 – Modelo de um nó de uma RNA. (Fonte: Matos, 2008)

O nó apresentado recebe ligações de três nós a montante e está ligado a três nós para jusante. As funções de transformação que podem ocorrer no interior do nó são as seguintes:

- Função de Input (fi) – normalmente, corresponde ao somatório dos *inputs* num único valor real, mas outras conceções podem ser aplicadas (produto).
- Função de ativação (fa) – esta função faz corresponder o resultado da função de *input* a um nível de ativação do nó, variando normalmente entre -1 e 1 ou entre 0 e 1.

- Função de output (f_o) – aplica-se esta função para produzir um sinal a partir do nível de ativação. Na maioria dos casos esta função não é aplicada, pois, geralmente, transmite-se simplesmente o nível de ativação.

Conforme referido por Silva (2013), existem vários tipos de RNA, diferindo na estrutura, no género de computações realizadas dentro dos neurónios ou até nos algoritmos de treino. No entanto, as redes mais comuns são constituídas por neurónios ordenados em camadas, fazendo com que cada neurónio receba como *inputs* apenas os outputs dos neurónios das camadas anteriores ou inputs externos, isto é, não existem ligações entre neurónios da mesma camada. As redes neuronais conseguem modelar funções de, praticamente, qualquer complexidade, estando o número de camadas e de nós em cada camada associado à complexidade da função.

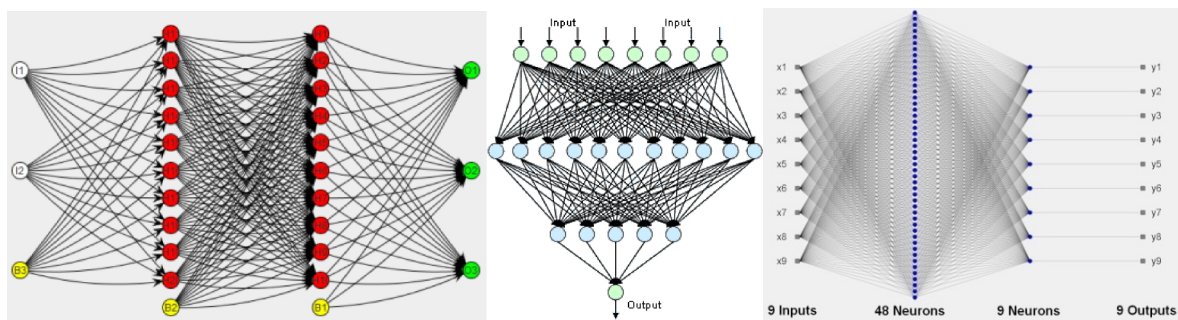


Figura 28 – Diferentes representações de redes multicamada.

O conhecimento destas redes é adquirido a partir de condições reais (conjunto de dados), através de um processo de aprendizagem (algoritmo de treino). Ou seja, para determinar os valores dos pesos sinápticos é necessário um conjunto de dados que permitam aferir a relação entre os valores de entrada da rede e os valores de saída.

4.3.1. Perceptrão

De acordo com Haykin (1999), o Perceptrão proposto por Rosenblatt (1958) – o primeiro algoritmo com aprendizagem supervisionada – é a forma mais simples de uma rede

neuronal utilizada para a classificação de padrões considerados linearmente separáveis, ou seja, padrões que se encontrem em lados opostos de um hiperplano. Basicamente, este modelo consiste num único neurónio com pesos sinápticos e constantes (*bias*) ajustáveis. O esquema de nó utilizado por Rosenblatt para desenvolver o seu perceptrão teve como base o modelo de neurónio artificial concebido por McCulloch e Pitts (1943).

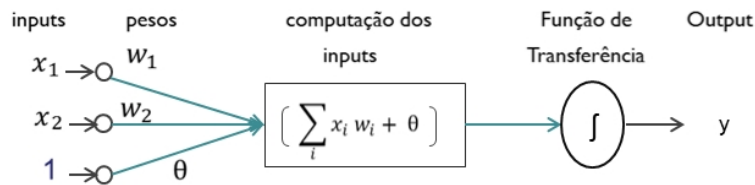


Figura 29 - Representação esquemática de um perceptrão. (Fonte: Lingireddy e Brion, 2005)

Com efeito, como refere Haykin (1999), Rosenblatt provou que se os padrões (vectores) usados para treinar o *perceptrão* forem extraídos de duas classes linearmente separáveis, então o algoritmo de treino converge e coloca a superfície de decisão na forma de um hiperplano entre as duas classes. A prova da convergência do algoritmo ficou conhecida como o *Teorema de Convergência do Perceptrão*.

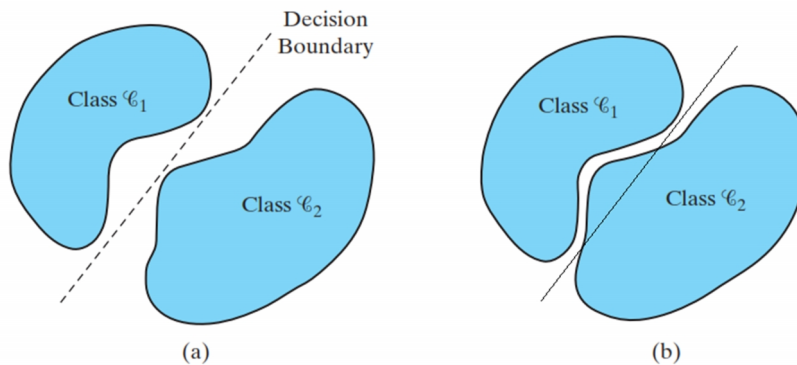


Figura 30 – a) Classes linearmente separáveis; b) Classes não linearmente separáveis. (Fonte: Haykin, 1999)

Segundo Lingireddy e Brion (2005) o perceptrão concebido por Rosenblatt é um perceptrão de camada única. Conforme comprovado por Minsky e Papert, perceptrões de camada única apenas conseguem classificar padrões linearmente separáveis. Como a maioria dos problemas de classificação não possuem classes linearmente separáveis, esta

limitação impôs severas limitações à aplicação do perceptron. Mais tarde, Rumelhart, Hinton e Williams (1986) apresentaram uma solução para este problema com a publicação do algoritmo de retro-propagação do erro.

Em conformidade com Matos (2008) estudos posteriores provaram que aumentando o número de nós, aumenta também o número de classes que o perceptron permite separar. O perceptron foi paradigmático e é interessante na introdução a redes mais complexas, em particular ao PMC que é nele inspirado.

4.3.2. Perceptron Multicamada (PMC)

Silva (2013) refere que as redes mais comuns, que são constituídas por camadas, conhecem diferentes designações. Na terminologia anglo-saxónica são conhecidas como Feed-forward Networks ou Perceptron Networks, traduzindo para português, podem ser designadas por Redes Multicamadas ou Perceptron Multicamada, por seguirem o princípio do Perceptron, aplicando-o a várias camadas. A arquitetura típica destas redes apresenta uma camada de entrada (input), uma camada de saída (output) e, pelo menos, uma camada intermédia ou oculta (hidden layer). Os nós da camada de entrada recebem os valores das variáveis de entrada do modelo e circulam pela rede até atingir a camada de saída produzindo os valores de saída. Em seguida é apresentada uma arquitetura exemplificativa de uma rede multicamada.

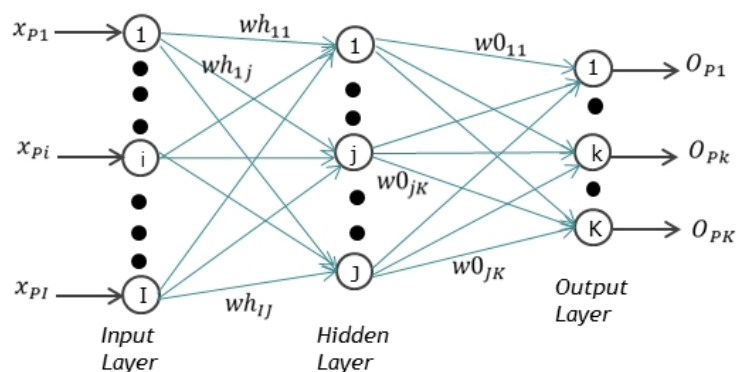


Figura 31 – Esquema de rede MLFN ou PMC de 3 camadas. (Fonte: Silva, 2013)

Matematicamente, uma RNA de 3 camadas com I nós de entrada, J nós ocultos numa única camada oculta e K nós de saída, pode ser expressa da seguinte forma:

$$O_{Pk} = f_1 \left(\sum_{j=1}^J w^o_{jk} \cdot f_2 \left(\sum_{i=1}^I w^h_{ij} \cdot x_{pi} + b^j_1 \right) + b^k_2 \right)$$

Fórmula 2 – Fórmula de cálculo de um valor de output da rede PMC. (Fonte: Lingireddy e Brion, 2005)

Em que:

O_{Pk} – representa o valor de saída do k^o nó da camada de saída para o vetor p;

x_{pi} – é o valor de entrada no i^o nó da camada de entrada do vetor p;

w^o_{jk} – é o peso da conexão entre o nó j da camada oculta e o nó k da camada de saída;

w^h_{ij} – é o peso da conexão entre o nó i da camada de entrada e o nó j da camada oculta;

b^j_1 e b^k_2 – são termos contantes (bias);

f_1 e f_2 – representam funções de ativação.

É necessária uma função de ativação adequada para as unidades escondidas, para se introduzir não-linearidade na rede, o que proporciona a capacidade de capturar relações não lineares entre os valores de entrada e de saída. Para as unidades de saída, a seleção da função de ativação é baseada na distribuição dos valores de referência (reais).

Silva (2013) menciona que Jordan (1995) considera uma boa escolha, para a função de ativação, a função logística sigmoideal, pois é a função de ativação mais comum e produz resultados dentro da gama de 0 a 1. Esta função tem a seguinte forma:

$$f(x) = \frac{1}{1 + e^{-x}}$$

Fórmula 3 – Função Sigmoideal. (Fonte: Lingireddy e Brion, 2005)

Estas redes neuronais, normalmente, são treinadas usando um algoritmo de retro-propagação do erro (back-propagation).

De acordo com a Palisade (2010) as RNAs são normalmente utilizadas em dois tipos de problemas:

- **Problemas de Classificação** – nos quais se tenta determinar a que tipo de categoria pertence um item desconhecido.

Exemplos: diagnósticos médicos, reconhecimento de padrões e previsão da capacidade de pagamento de créditos bancários.

- **Problemas numéricos** – situações em que é necessário prever um resultado numérico específico.

Exemplos: previsões de preços de ações, previsão do nível de vendas num período futuro.

Quando as redes neuronais são usadas para prever valores numéricos, geralmente, produzem apenas um output. Isso acontece porque as redes de output único são mais fiáveis do que as que produzem múltiplos outputs e praticamente qualquer problema de previsão poder ser processado por redes de output único. Em situações em que seja necessário prever duas variáveis, em vez de construir uma única rede para esse efeito, é preferível criar uma rede para prever cada uma das variáveis que se pretendem prever.

Por outro lado, as redes neuronais usadas para previsão de classificação ou categoria têm múltiplos outputs. Supondo, por exemplo, que queremos prever uma classificação de “Bom”, “Satisfatório” ou “Mau”. Neste caso, a rede terá três outputs numéricos. O output mais alto indicará a categoria selecionada pela rede.

As redes neuronais podem ser aplicadas em problemas de regressão, classificação e compactação de dados, como também em situações onde existem interações não-lineares entre as variáveis dependentes e as independentes.

Santos et al. (2005) referem que a capacidade de aprender através de exemplos e de generalizar a informação aprendida representam, sem dúvida, atributos importantes para a

escolha de uma solução neuronal de problemas diversos. A generalização – associada à capacidade da rede aprender através de um conjunto de exemplos representativos do problema que se pretende estudar e, posteriormente, fornecer respostas coerentes para dados não apresentados anteriormente – é uma demonstração de que a capacidade das RNA vai muito além de mapear relações de entrada e saída. As RNA são capazes de extrair informações não apresentadas de forma explícita através dos exemplos.

De acordo com Lingireddy et al. (2005) a ideia base na modelação de RNA é identificar relações implícitas entre os valores de entrada e de saída. Para isso, é necessário que os valores de entrada da rede neuronal contenham informação pertinente para o propósito da rede, de modo a que exista uma função matemática que os relacione com valores de saída, com o desejado grau de precisão. Uma RNA não conseguirá aprender relações que não existam.

4.4. Tipos de Redes Neurais Artificiais

Existem diferentes tipologias de redes neuronais que se podem encontrar na literatura. As redes neuronais podem ser classificadas de diversas formas, em relação a:

- Arquitetura – redes planares ou constituídas por camadas;
- Processo de aprendizagem – com supervisão ou sem supervisão;
- Dinâmica – transferência de informação sequencial ou simultânea.

Para além das redes mais comuns, as redes multicamadas – como o PMC ou MLFNN e as PNN/GRNN – Matos (2008) refere outros modelos de redes neuronais, como as Máquinas de Boltzmann ou as redes de Hopfield, que constituíram marcos importantes na área das RNA e são exemplos de conceitos inteiramente novos. Lobo (2007) também refere outras redes como as redes de funções de base radial (RBF), os mapas auto-organizados (SOM) e as máquinas vetoriais de suporte (SVM), entre outras.

O software NeuralTools da Palisade, que irá ser utilizado no presente trabalho, utiliza redes neurais Multicamada Feed-Foward (MLFNN) ou Percepção Multicamada (MLP), redes de Regressão Generalizada (GRNN) – para previsões numéricas – e rede Probabilística (PNN) – para previsões de categoria. Como este caso de estudo trata de previsões numéricas, foram utilizadas as redes MLF e GRNN e apenas essas serão descritas em seguida.

4.4.1. Redes MLFNN e PNN/GRNN

De acordo com a Palisade (2010), existem vários tipos de redes neurais, que se diferenciam pela estrutura, pelos tipos de cálculos efetuados dentro dos neurônios e pelos algoritmos de treino.

MLFNN

Uma rede neuronal multicamadas é tipicamente composta por camadas alinhadas de neurónios ou nós. Neste tipo de rede, os dados de entrada são apresentados na primeira camada a que chamamos camada de entrada ou *input*. Esta camada distribui as informações de entrada para as camadas seguintes, as camadas ocultas ou intermediárias. A última camada é camada de saída (output), na qual a solução do problema é obtida. As camadas de entrada e de saída poderão ser separadas por uma ou mais camadas ocultas. Na grande maioria das aplicações, considera-se apenas uma camada oculta.

Além disso, os neurónios de uma camada estão conectados apenas aos neurónios da camada imediatamente posterior, não existindo conexões entre neurónios da mesma camada e a comunicação é efetuada unidireccionalmente. De salientar que cada neurónio está conectado com todos os neurónios da camada seguinte.

O número de nós na camada de entrada corresponde ao número de variáveis que serão

usadas para alimentar a rede neuronal e os nós assumem o valor dos dados de entrada.

Normalmente são escolhidas as variáveis mais relevantes para o problema em estudo.

Fatores que determinam o comportamento da rede:

- Topologia – o número de camadas ocultas e o número de nós das camadas;
- “Pesos” das conexões – parâmetro atribuído a cada conexão – e termos constantes – parâmetro atribuído a cada neurónio;
- Função de ativação/transferência – usada para converter os *inputs* de cada neurónio no output correspondente.

GRNN

As redes PNN e GRNN funcionam de forma semelhante, distinguindo-se pelo tipo de problema a que podem ser aplicadas. Se a variável dependente for categórica usa-se a rede PNN, mas se for numérica recorre-se a uma rede GRNN (como neste caso de estudo).

As redes GRNN, apresentam sempre o mesmo número de camadas – uma de *input*, uma de *output* e duas ocultas. Na camada de *input* cada nó representa uma variável de entrada. A primeira camada oculta – camada padrão – contém um nó para cada caso de treino. A segunda camada oculta – camada de soma – é constituída por nós de numeradores e denominadores e o seu tamanho depende de diversos fatores relacionados com os dados de treino. Cada neurónio da camada de padrões calcula a sua distância em relação ao caso apresentado. Os valores passados aos nós de numeradores e denominadores são funções da distância e do valor dependente. Os nós da camada de soma efetuam a soma dos seus *inputs*, enquanto o nó de *output* os divide para gerar a previsão.

Quando um caso é apresentado à rede, ela calcula o valor dependente previsto usando os valores dependentes de cada caso de treino, sendo que os casos de treino mais próximos contribuem de forma mais significativa para o valor de *output*.

De acordo com a Palisade (2010) quanto mais próximo um caso conhecido está de um desconhecido, mais peso é dado ao estimar o valor dependente desconhecido. A GRNN baseia-se nestes conceitos intuitivos.

A função de distância calculada nos neurónios da camada padrão usa fatores de suavização. O treino de uma rede GRNN consiste em otimizar os fatores de suavização para minimizar o erro no conjunto de treino. A medida de erro, usada durante o treino, para avaliar diferentes conjuntos de fatores de suavização é o erro médio quadrático (EMQ).

A principal vantagem das redes PNN e GRNN é que, ao contrário das redes MLFNN, elas não precisam de nenhuma configuração. Ao mesmo tempo, o grau de exatidão das previsões, de modo geral, é comparável aos das redes MLFNN.

4.4.2. Comparação entre redes MLFNN e PNN/GRNN

Cada uma destas redes oferece vantagens e desvantagens, conforme descrevemos a seguir:

Vantagens das redes GRNN/PNN:

- Treino rápido;
- Não requerem especificação da topologia (número de camadas ocultas e nós);
- Além de efetuar classificação, as redes PN retornam as probabilidades de o caso se enquadrar em diferentes categorias dependentes.

Vantagens das redes MLFNN:

- Apresentam uma estrutura mais pequena, por isso as previsões são mais rápidas;
- São mais confiáveis fora dos intervalos dos dados de treino (por exemplo, quando o valor de uma variável independente não pertence à faixa de valores usados nos dados de treino). No entanto, a previsão com redes MLFNN fora do intervalo dos dados de treino é muito arriscada.

- São capazes de efetuar generalizações a partir de conjuntos de treino muito pequenos.

4.5. Fases de desenvolvimento de uma RNA

A implementação de uma rede neuronal compreende quatro fases distintas que são a preparação dos dados, o treino da rede, o teste e a previsão. Essas mesmas fases são descritas em seguida:

4.5.1. Preparação dos dados

Conforme mencionado por Lingireddy et al. (2005), uma aplicação bem-sucedida de um modelo de RNA requer uma preparação adequada dos dados de input, assim como, uma sólida compreensão do processo a ser modelado é essencial para a correta seleção das variáveis de entrada. Esta compreensão fundamental não só ajudará na escolha de dados de *input* apropriados, mas também ajudará a evitar a escolha de dados com erros ou dados menos fiáveis, que poderão confundir o processo de treino.

Segundo o mesmo autor, o objetivo da preparação dos dados é definir um conjunto de dados apropriados a serem utilizados na RNA, escolher as variáveis de entrada e identificar as variáveis a prever. Encontrar variáveis de entrada para a modelação de RNA e recolher dados suficientes para o treino da rede, geralmente, é mais demorado e exige maior esforço do que treinar a própria rede. Frequentemente, os dados disponíveis são subdivididos em dois ou três conjuntos. O conjunto com a maior quantidade de dados (entre 70 a 80%) é usado para o treino, outro conjunto é usado para o teste e, se necessário, poderá existir um terceiro conjunto para validação. O conjunto de treino é composto por casos conhecidos que serão usados no processo de aprendizagem da rede, isto é, para ajustar os parâmetros ou pesos da rede. O conjunto de teste é utilizado para avaliar a capacidade de generalização da rede. O conjunto de validação permite confirmar

resultados considerados aceitáveis na fase de teste, atestando a capacidade de generalização da rede.

A quantidade de dados disponíveis é um fator muito importante para o sucesso de uma rede neuronal. Normalmente, quanto maior for a quantidade de casos melhor será o treino da rede ou, pelo menos, com um maior número de casos será mais fácil para a rede perceber as relações existentes entre os dados. Com um número insuficiente de exemplos, tanto o desempenho da rede como a sua capacidade de generalização serão pobres. No entanto, não importa apenas a quantidade de dados, a qualidade dos mesmos também tem de ser garantida.

4.5.2. Treino da RNA

Em consonância com o que foi apresentado por Silva (2013), o processo de treino de uma rede neuronal é efetuado a partir de um conjunto de dados dos quais conhecemos os valores de output. Este é o processo de aprendizagem que permite afinar os parâmetros de computação da rede. O objetivo é fazer com que a rede retorne valores de *output* aproximados dos reais, face a determinados valores de entrada. O processo é orientado pelos dados fornecidos para treino e pelo algoritmo de treino. Este último seleciona vários conjuntos de parâmetros e avalia cada um aplicando a RNA a cada caso de treino (uma linha de informação), verificando a qualidade das respostas da rede, comparando-as com os valores reais. Cada conjunto de parâmetros designa-se por tentativa (*trial*). O algoritmo escolhe novos conjuntos de parâmetros tendo em consideração os resultados de anteriores tentativas. Não são disponibilizados todos os dados conhecidos para o processo de treino da rede, guardando-se parte para o teste e, se necessário, para a validação.

4.5.3. Testar a RNA

De acordo com a Palisade (2010), durante esta fase, a rede neuronal que foi previamente treinada é testada para verificar o seu desempenho, ou seja, a qualidade com que consegue prever valores de casos conhecidos, mas que não lhe foram fornecidos na fase de treino. A informação usada na fase de teste é, geralmente, um subconjunto de dados históricos, cujos *outputs* são conhecidos.

A fase de teste permite medir estatisticamente o desempenho da rede, nomeadamente, a percentagem de respostas conhecidas que foram previstas corretamente. Também é determinado o erro médio quadrático com base no qual é escolhida a melhor rede, sendo assim, a rede que apresentar menor erro médio quadrático na fase de teste é considerada a rede com maior garantia de sucesso na previsão de novos valores desconhecidos.

4.5.4. Previsão

Em conformidade com a Palisade (2010) nesta fase uma rede neuronal já treinada e testada é usada para prever valores de output desconhecidos. Este é o principal objetivo das RNA, ou seja, as redes neuronais são treinadas e testadas para adquirirem a capacidade de generalização que lhes permite efetuar previsões, com um determinado grau de confiança, quando inserimos novos dados.

O processo de treino e teste são interativos e, por vezes, longos. Normalmente, o treino pode ser efetuado várias vezes, com configurações diferentes, até gerar uma rede neuronal que apresente os melhores resultados ao ser testada. Depois de se obter a melhor rede, esta pode ser usada para efetuar previsões rapidamente.

4.6. Comparação entre Redes Neuronais e Métodos Estatísticos

As RNA são usadas para aproximar funções, assim como as regressões lineares, e usam-se para classificação, da mesma forma que análises discriminantes e regressões logísticas. A vantagem das redes neuronais prende-se com o facto de serem capazes de modelar funções extremamente complexas. As técnicas de otimização de modelos lineares já eram bem conhecidas antes de serem inventadas as RNA, em meados do século XX. O desenvolvimento de algoritmos eficazes para treinar redes neuronais ainda demorou algum tempo. No entanto, atualmente, existe uma grande variedade de algoritmos sofisticados para o treino de RNA, o que as torna uma alternativa mais interessante do que os métodos convencionais. (Palisade, 2010)

O poder de RNAs na aproximação de relações funcionais complexas elimina a necessidade de se conhecer a forma como essas relações representam o fenómeno a ser modelado. As ferramentas de modelação tradicionais requerem a especificação inicial da relação funcional (linear, polinomial, exponencial ou alguma outra função complexa). A maioria da modelação de dados ambientais é altamente complexa e a capacidade dos modelos de RNAs em compreender relações complexas com apenas algumas alterações simples (por exemplo, alterar o número de nós da camada oculta) é certamente uma vantagem incomparável. (Linigireddy e Brion, 2005)

4.7. Justificação da aplicação de redes neuronais

A DA é descrita muitas vezes como um processo instável e a complexidade do processo industrial de produção de biogás constitui um obstáculo à sua otimização. Apesar dos avanços verificados nos últimos 20 anos na compreensão da bioquímica e energética do processo que contribuíram para o desenvolvimento de estratégias que permitiram aumentar a estabilidade dos reatores anaeróbios, continua a ser um fenómeno complexo em que intervêm muitas variáveis e cujas inter-relações não são totalmente conhecidas. O interior de um digestor anaeróbio funciona como uma caixa negra da qual não se consegue obter informação, em tempo útil, que permita a monitorização e controlo do processo de digestão. Atualmente, o controlo do processo é efetuado apenas com base na experiência acumulada no passado e nos conceitos teóricos do processo anaeróbio conhecidos. O resultado de qualquer opção ou decisão não é conhecido no imediato, o que torna complexa a monitorização do processo. Esta continua a ser a principal dificuldade com que se deparam as entidades gestoras das ETARs.

Espera-se, no entanto, que a aplicação de modelos matemáticos, como as Redes Neurais Artificiais (RNA), a esta problemática, possa prestar um forte contributo para a compreensão do processo de digestão e otimização da produção de biogás, constituindo-se como um sistema de apoio à decisão na gestão do processo anaeróbio.

Conforme já foi mencionado neste trabalho, as redes neuronais têm capacidade de entender relações complexas existentes entre os dados que de outra forma não seria possível perceber. Por imitação das funções do cérebro, conseguem discernir padrões existentes nos dados e, quando são fornecidos dados novos, são capazes de gerar previsões com base nas relações e nos padrões detetados nos dados utilizados no treino. Esta metodologia possui a vantagem de detectar implicitamente qualquer relação não-linear entre uma variável resposta e as variáveis explicativas.

Devido ao facto de não haver necessidade de independência e normalidade das variáveis em estudo, bem como a sua grande capacidade de aprendizagem a partir do ambiente, a aplicação de redes neuronais artificiais na análise do processo de digestão anaeróbia pode ser muito interessante. Além do mais, de acordo com Santos et al. (2005), o processamento neuronal é capaz de extrair relações das variáveis de entrada diretamente sobre os espaços de dimensão elevada que tipicamente as caracterizam, tornando tal processamento uma ferramenta valiosa em problemas complexos de reconhecimento de padrões. Por outro lado, as redes neuronais podem trabalhar em conjunto com outras técnicas de processamento, permitindo que se utilize o conhecimento acumulado numa determinada área de aplicação. Assim, os dados podem ser pré-processados, identificando-se a informação relevante à tarefa de processamento de interesse, e a rede neuronal irá operar sobre esta informação qualificada, ao invés de trabalhar com os dados brutos. Desta maneira, evitam-se modelos neuronais de alta complexidade, que normalmente são pouco práticos.

As RNA têm vindo a ser desenvolvidas e aplicadas nos mais diversos campos da ciência para responder a desafios tão distintos como o processamento de linguagem e imagens, reconhecimento de padrões, classificação de dados, otimização, planeamento, monitorização e controlo. Assim como, na área da engenharia e ambiente, na modelação de sistemas de abastecimento de água, previsões de precipitação diária, modelação de pH e previsão de salinidade em rios, ou no setor económico-financeiro, na análise e pesquisa de mercado, identificação de fraude com cartões de crédito e avaliação de créditos, e até na medicina, para apoio ao diagnóstico médico.

Para Matos (2008) as razões da crescente utilização das RNA em todos estes campos prendem-se, sem dúvida, com os benefícios que lhes estão associados:

- Aprendizagem – uma RNA pode ser iniciada “sem conhecimento” e treinada com um

dado conjunto de exemplos (inputs e outputs). Através do treino, as intensidades das ligações entre os nós alteram-se de tal forma que a rede aprende a produzir o resultado desejado para os dados introduzidos;

- Generalização – se um novo conjunto de dados, diferentes dos exemplos conhecidos, for introduzido na rede, esta produz o melhor resultado possível com base nos exemplos usados.
- Paralelismo massivo potencial – no decurso do processamento da informação, vários neurónios “disparam” simultaneamente.
- Robustez – Se alguns neurónios falharem, o sistema pode ainda ter um bom desempenho.
- Correspondência Parcial – Na maior parte dos casos uma correspondência total aos casos de exemplo não é necessária.

Segundo Liginreddy et al. (2005) as RNA apresentam uma vantagem substancial em relação aos métodos estatísticos que reside no facto de não ser necessário conhecer qual a relação funcional entre as variáveis que regem o fenómeno a modelar.

5. APRESENTAÇÃO DO CASO DE ESTUDO

5.1. Enquadramento do estudo

Com a conjuntura económica atual, a par da importância dada às preocupações ambientais e à preservação dos recursos naturais, as empresas do sector do tratamento de águas residuais preocupam-se, cada vez mais, em garantir o seu desenvolvimento sustentável, procurando reduzir os custos de exploração através da otimização de processos de tratamento.

A SIMRIA não é exceção, pelo que tem demonstrado uma forte preocupação com a otimização do funcionamento da etapa de Digestão Anaeróbia da ETAR Sul e a gestão do biogás gerado, tendo solicitado, antes da realização deste estágio, uma assessoria técnica à Direção de Engenharia do grupo Águas de Portugal com a duração de aproximadamente um ano, que decorreu entre Agosto de 2010 e Junho de 2011. Como resultado desse estudo foram identificadas um conjunto de medidas a serem implementadas que permitiram uma melhoria da operacionalidade global da linha de lamas.

Neste contexto, realizou-se o presente estudo de análise e tratamento de dados de Digestão Anaeróbia, da ETAR Sul da SIMRIA, para gestão de processo com recurso a redes neurais. Este trabalho, cujo objetivo passa por relacionar algumas variáveis explicativas com a produção de biogás, pode auxiliar a tomada de decisões no processo de produção, com vista à otimização da produção de biogás. As conclusões e considerações que resultaram da assessoria técnica prestada pela AdP, foram o ponto de partida para o desenvolvimento deste estudo.

5.2. Metodologia Proposta

Para o desenvolvimento deste trabalho utilizaram-se os programas Microsoft ExcelTM e Neural ToolsTM do The Decision Tools SuiteTM da PalisadeTM.

5.2.1. Considerações Iniciais

Relativamente à aplicação informática utilizada no presente trabalho, importa referir as questões seguintes, dado que há no mercado outras aplicações informáticas destinadas a esta área. Da leitura de Matos (2008), em particular, conclui-se ser possível recorrer à aplicação MatlabTM (com Neural Network ToolboxTM e Genetic Algorithms ToolboxTM), que segundo o orientador se enquadra num conjunto de ferramentas muito divulgadas em diversas escolas de ensino superior, nomeadamente em Portugal. O recurso ao Neural ToolsTM, um dos aplicativos do pacote Decision Tools SuiteTM, da Palisade (que inclui também uma ferramenta de otimização baseada em algoritmos genéticos – o EvolverTM) foi decidida em conjunto com o orientador, pelo facto dessas ferramentas estarem disponíveis no Departamento de Engenharia Civil do ISEP, tendo também a Simria adquirido, no contexto do estágio em que este trabalho se realizou, uma versão para estudante, destinada a esse fim. Vale a pena referir, ainda, que em Schuyler (2001) são analisados, comparativamente, alguns programas mais divulgados da área da análise de decisão, onde se insere o pacote Decision Tools SuiteTM, sendo de admitir, pelo que esse autor descreve, que a PalisadeTM e a Crystal BallTM são dois dos fornecedores mais credíveis, nessa área de aplicações informáticas.

Como habitual nestas questões de aplicativos informáticos, torna-se evidente que há fornecedores que apresentam uma elevada qualidade para aplicações específicas, sendo mais fortes numas áreas, mas menos noutras, o que parece não ser o caso dos referidos acima. Entretanto, convém registar que, à data da publicação do livro acima referido, a

Palisade ainda não tinha integrado o Neural Tools™ no pacote Decision Tools Suite™, segundo informação do orientador desta dissertação. Ainda assim, avaliando os dados referidos, considerou-se que a adoção do Neural Tools™ para suporte informático a este trabalho era uma solução adequada, representando um bom compromisso de uma aplicação “amigável”, pela sua orientação para a comercialização no mercado informático generalizado, e tecnicamente robusta, dada a credibilidade do fabricante que a fornece – a Palisade – há muitos anos no mercado e com uma posição dominante a defender, como se depreende de Silva (2008). Considerou-se, pois, que para uma abordagem pragmática com recurso a ferramentas deste tipo e para os modelos que uma empresa da área do saneamento básico terá interesse em desenvolver, a aplicação informática escolhida é claramente adequada aos objetivos pretendidos.

5.2.2. Apresentação do NeuralTools

Na parte inicial deste estágio foi realizado um Workshop, na SIMRIA, pelo Eng.º Gabriel Silva, orientador desta dissertação, para introdução dos conceitos base das redes neuronais e apresentação do Neural Tools aos colaboradores da empresa e de outras empresas do grupo AdP, com recurso a alguns exemplos práticos que acompanham o programa.

De acordo com a Palisade (2010), o Neural Tools é uma ferramenta informática poderosa, flexível e fácil de usar que permite efetuar previsões através de recursos avançados de redes neuronais num ambiente conhecido, o Microsoft Excel. Os procedimentos do Neural Tools – como a definição de conjunto de dados, o treino e teste de redes neuronais e a previsão de valores usando redes treinadas – podem ser executados com os dados contidos no Excel, e os relatórios e tabelas de análise também são criadas no Excel.

Este *software* pode ser aplicado a problemas numéricos ou de classificação, permitindo qualquer combinação de variáveis numéricas ou categóricas para variáveis dependentes ou

independentes e as relações entre elas podem ser lineares ou não lineares.

Os conjuntos de dados a usar nas RNA tendem a ser grandes, de modo que, normalmente, é possível dividi-los em três grupos, um para o treino da rede, um para o teste (quando os valores da variável dependente são conhecidos) e outro para a previsão (quando os valores da variável dependente não são conhecidos).

O Neural Tools permite desenvolver as redes neuronais nas quatro etapas já referidas neste trabalho (preparação de dados, treino, teste e previsão da rede). O desenvolvimento destas etapas no NeuralTools e os comandos associados serão descritos nos pontos seguintes.

5.2.3. Preparação dos dados

Antes de se criar qualquer rede neuronal, deve-se examinar os dados cuidadosamente e determinar o que se pretende alcançar.

O Neural Tools trabalha sobre folhas de cálculo do Excel, permitindo estruturar os dados por casos e variáveis. Cada coluna da folha de cálculo representa uma variável, com o nome da mesma a ser colocado na primeira linha da tabela do conjunto de dados, e cada linha representa um caso com diferentes valores das variáveis escolhidas. Cada caso contém um conjunto de valores de variáveis independentes e um valor conhecido ou não para a variável dependente ou de output.

| | A | B | C | D | E | F | G | H | I | J |
|----|----|-------------|------------|------------------------|--------------------------|------|---------------------------|---|---|---|
| 1 | | | Lamas | Digestor A | | | Biogás | | | |
| 2 | # | Data | SL4 %MV | Q_{ent} (m^3) | Temp. ($^{\circ}C$) | CMV | $Q_{Biogás}$ (m^3) | | | |
| 3 | 1 | 17-Abr-2011 | | 97 | 34,4 | | | | | |
| 4 | 2 | 18-Abr-2011 | 1,7 | 109 | 33,1 | 0,62 | | | | |
| 5 | 3 | 19-Abr-2011 | 1,5 | 138 | 33,6 | 0,69 | | | | |
| 6 | 4 | 20-Abr-2011 | | 203 | 33,3 | | 1547 | | | |
| 7 | 5 | 21-Abr-2011 | | 130 | 32,8 | | 2683 | | | |
| 8 | 6 | 22-Abr-2011 | | 167 | 33,3 | | 2584 | | | |
| 9 | 7 | 23-Abr-2011 | | 168 | 33,9 | | 2819 | | | |
| 10 | 8 | 24-Abr-2011 | | 168 | 34,6 | | 2755 | | | |
| 11 | 9 | 25-Abr-2011 | | 169 | 35,1 | | 2983 | | | |
| 12 | 10 | 26-Abr-2011 | 1,8 | 152 | 35,5 | 0,91 | 1287 | | | |
| 13 | 11 | 27-Abr-2011 | | 169 | 35,5 | | 1177 | | | |
| 14 | 12 | 28-Abr-2011 | 1,1 | 167 | 35,2 | 0,61 | 1006 | | | |
| 15 | 13 | 29-Abr-2011 | | 219 | 35,3 | | 1103 | | | |
| 16 | 14 | 30-Abr-2011 | | 322 | 30,9 | | 1067 | | | |
| 17 | 15 | 1-Mai-2011 | | 353 | 35,6 | | 1172 | | | |
| 18 | 16 | 2-Mai-2011 | | 196 | 36,0 | | 1463 | | | |
| 19 | 17 | 3-Mai-2011 | 2,0 | 200 | 36,3 | 1,33 | 1537 | | | |
| 20 | 18 | 4-Mai-2011 | | 258 | 36,6 | | 1292 | | | |
| 21 | 19 | 5-Mai-2011 | 2,1 | 244 | 35,7 | 1,71 | 1153 | | | |
| 22 | 20 | 6-Mai-2011 | | 249 | 35,5 | | 1011 | | | |
| 23 | 21 | 7-Mai-2011 | | 271 | 35,9 | | 915 | | | |

Figura 32 – Exemplo de uma base de dados para utilizar com o Neural Tools

Para a preparação dos dados o NeuralTools disponibiliza o Gestor do Conjunto de Dados. Este comando permite definir conjuntos de dados e editar ou excluir variáveis que podem ser usadas para o treino, teste e previsão das redes neuronais.

A caixa de diálogo do Gestor do Conjunto de Dados permite adicionar ou remover conjuntos de dados, atribuir nomes aos mesmos e especificar o *layout*. Para definir um conjunto de dados tem que se seleccionar qualquer célula dentro do conjunto de dados e clicar no botão “Gerenciador de Conjunto de Dados” no separador NeuralTools. Isso faz aparecer a seguinte caixa de diálogo:

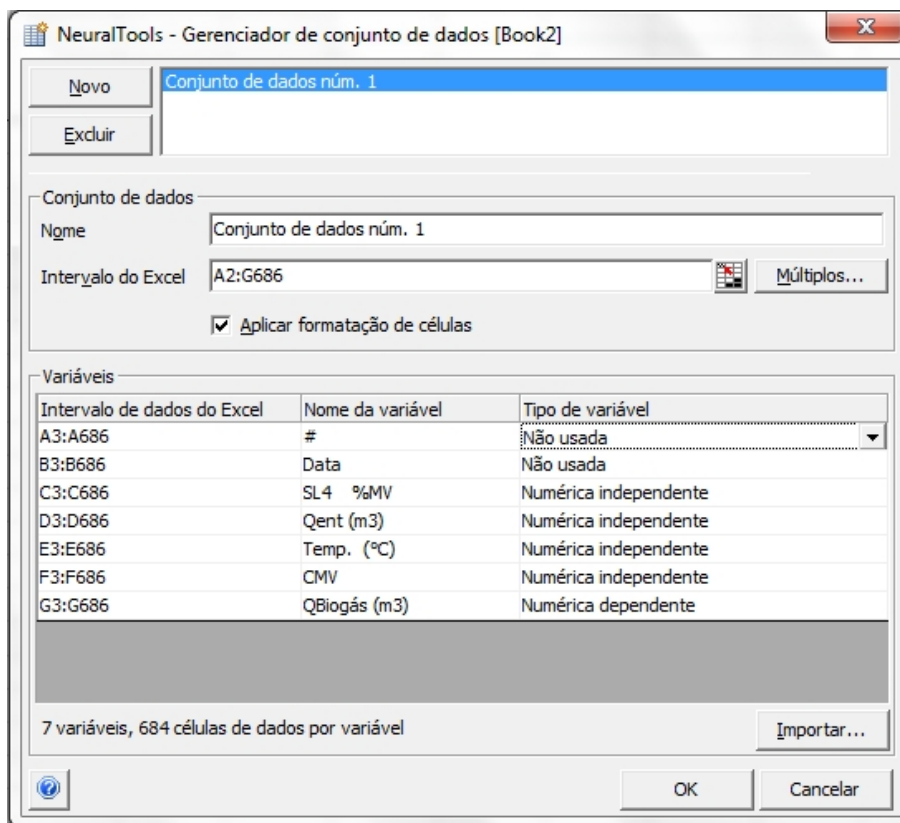


Figura 33 – Caixa de diálogo do Gestor do conjunto de dados

As opções na caixa de diálogo do Gestor do conjunto de dados são:

- Novo e Excluir – para adicionar um novo conjunto de dados ou excluir um conjunto existente.
- Nome – permite especificar o nome do conjunto de dados.
- Intervalo do Excel – especifica o intervalo do Excel correspondente ao conjunto de dados. Se forem selecionados vários intervalos de células de um conjunto de dados, este item será antecedido pelo rótulo Múltiplos.
- Aplicar formatação de células – acrescenta uma grade e cores que identificam os conjuntos de dados.
- Múltiplos – o botão “Múltiplos” permite especificar intervalos de células individuais que formam um conjunto de dados, com múltiplos intervalos de células.
- Importar – permite copiar neste conjunto de dados os tipos de variáveis de um outro conjunto de dados ou de uma rede neural treinada.

Na grade da caixa de diálogo do Gestor do Conjunto de Dados são apresentados: o Intervalo de dados do Excel, o Nome da variável e o Tipo de variável. Clicando no tipo de variável aparece uma lista que permite identificar o tipo de variável.

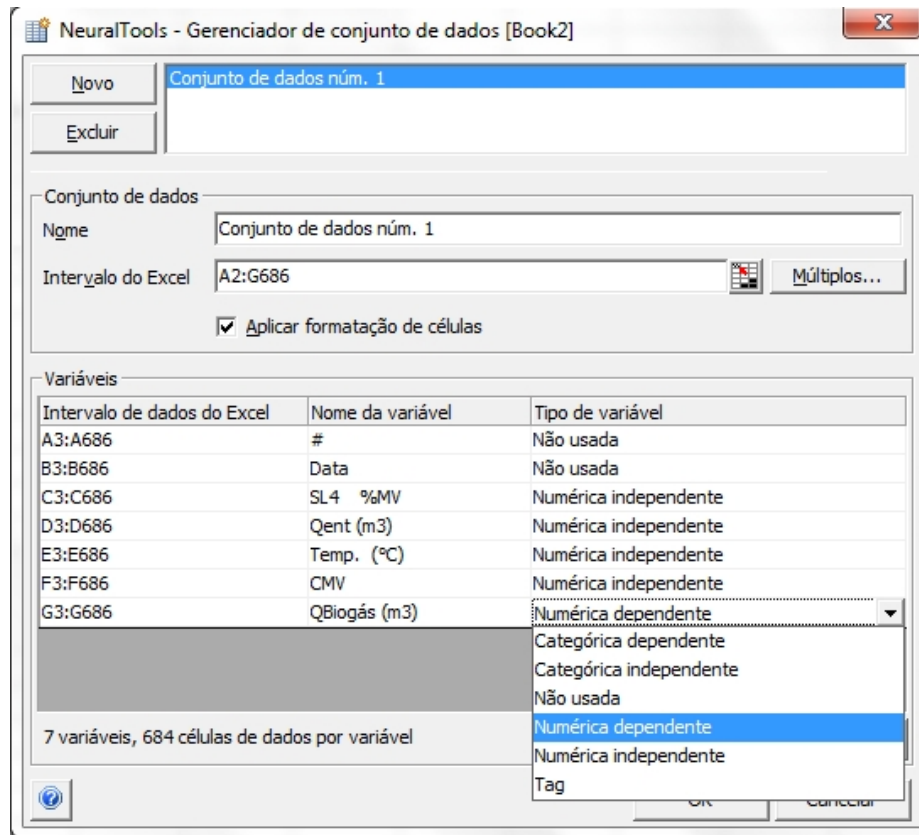


Figura 34 – Definição do tipo de variável

As opções de tipo de variável são:

- Categorica dependente – é uma variável dependente ou de output cujos valores possíveis são tomados de um conjunto de possíveis categorias.
- Categorica independente – é uma variável independente cujos valores possíveis são tomados de um conjunto de possíveis categorias.
- Numérica dependente – é uma variável dependente ou de output cujos valores possíveis são numéricos.
- Numérica independente – é uma variável independente cujos valores possíveis são numéricos.

- Tag – é uma variável que assume possíveis valores de “treino”, “teste” ou “previsão”.

Esse tipo de variável é escolhido pelo utilizador para identificar casos a serem usados especificamente para treino, teste e previsão.

- Não usada – é uma variável que não será usada na rede neural.

Após a definição do conjunto de dados, a folha de cálculo adquire a seguinte formatação:

| | | Lamas | | Digestor A | | Biogás | |
|----|-------------|---------|------------------------------------|------------|------|---------------------------------------|--|
| # | Data | SL4 %MV | Q _{ent} (m ³) | Temp. (°C) | CMV | Q _{biogás} (m ³) | |
| 1 | 17-Abr-2011 | | 97 | 34,4 | | | |
| 2 | 18-Abr-2011 | 1,7 | 109 | 33,1 | 0,62 | | |
| 3 | 19-Abr-2011 | 1,5 | 138 | 33,6 | 0,69 | | |
| 4 | 20-Abr-2011 | | 203 | 33,3 | | 1547 | |
| 5 | 21-Abr-2011 | | 130 | 32,8 | | 2683 | |
| 6 | 22-Abr-2011 | | 167 | 33,3 | | 2584 | |
| 7 | 23-Abr-2011 | | 168 | 33,9 | | 2819 | |
| 8 | 24-Abr-2011 | | 168 | 34,6 | | 2755 | |
| 9 | 25-Abr-2011 | | 169 | 35,1 | | 2983 | |
| 10 | 26-Abr-2011 | 1,8 | 152 | 35,5 | 0,91 | 1287 | |
| 11 | 27-Abr-2011 | | 169 | 35,5 | | 1177 | |
| 12 | 28-Abr-2011 | 1,1 | 167 | 35,2 | 0,61 | 1006 | |
| 13 | 29-Abr-2011 | | 219 | 35,3 | | 1103 | |
| 14 | 30-Abr-2011 | | 322 | 30,9 | | 1067 | |
| 15 | 1-Mai-2011 | | 353 | 35,6 | | 1172 | |
| 16 | 2-Mai-2011 | | 196 | 36,0 | | 1463 | |
| 17 | 3-Mai-2011 | 2,0 | 200 | 36,3 | 1,33 | 1537 | |
| 18 | 4-Mai-2011 | | 258 | 36,6 | | 1292 | |
| 19 | 5-Mai-2011 | 2,1 | 244 | 35,7 | 1,71 | 1153 | |
| 20 | 6-Mai-2011 | | 249 | 35,5 | | 1011 | |
| 21 | 7-Mai-2011 | | 271 | 35,9 | | 915 | |

Figura 35 – Layout do conjunto de dados

Apenas como curiosidade, referir que o NeuralTools tem capacidade para processar na mesma sessão:

- Até 256 conjuntos de dados contidos na mesma pasta de trabalho.
- Até 16384 variáveis para cada conjunto de dados
- O número de casos de dados por variável e de casos por conjunto de dados limita-se somente à memória disponível no Excel (cerca de 16.277.216 casos)

5.2.4. Treino e teste da Rede Neuronal

Depois de definido um conjunto de dados procede-se ao treino e teste da rede neuronal.

O treino permite treinar uma rede neuronal a partir de um conjunto de dados com valores de output conhecidos. O processo de teste, que é realizado após o treino, permite medir o desempenho da rede treinada, na previsão dos valores de output. Os processos de treino e teste são interativos e, às vezes, longos. Depois de se obter a melhor rede, esta pode ser usada para efetuar previsões de valores de output desconhecidos.

O Neural Tools permite treinar, testar e prever os valores de output em simultâneo, ou seja, numa única etapa, para isso precisamos de ter uma base de dados com valores de output conhecidos para o treino e teste e valores de output desconhecidos para a previsão.

Existem três botões no grupo redes neuronais do separador NeuralTools que permitem escolher as definições da rede neuronal. Basicamente, este grupo permite definir os dados que serão utilizados no treino e no teste de uma rede e configurar o algoritmo de treino a usar e os critérios do tempo de execução desse algoritmo.

Treinar

Ao clicar no botão Treinar aparece uma caixa de diálogo de treino, com o separador Treinar visível.

Nesta caixa podem-se seleccionar as definições para o treino da rede neuronal. Pode-se escolher a opção “testar automaticamente em casos aleatoriamente seleccionados”, definindo a percentagem de casos destinados para o teste. É possível seleccionar a opção “prever automaticamente valores dependentes ausentes”, usando esses dados para previsão e, em algumas versões, está disponível a previsão instantânea, que permite mudar os valores das variáveis independentes alterando o resultado da previsão instantaneamente.

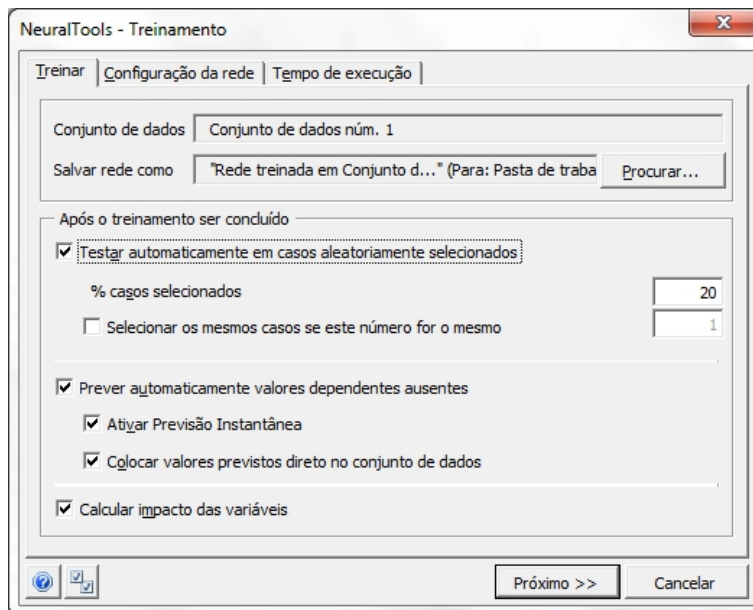


Figura 36 – Caixa de diálogo Treinar

Permite, ainda, seleccionar a opção “calcular impacto das variáveis” que possibilita a análise do impacto que cada variável de *input* teve no treino da rede neuronal.

Configurações de rede

Ao clicar no separador Configuração da Rede aparece a seguinte caixa de diálogo.

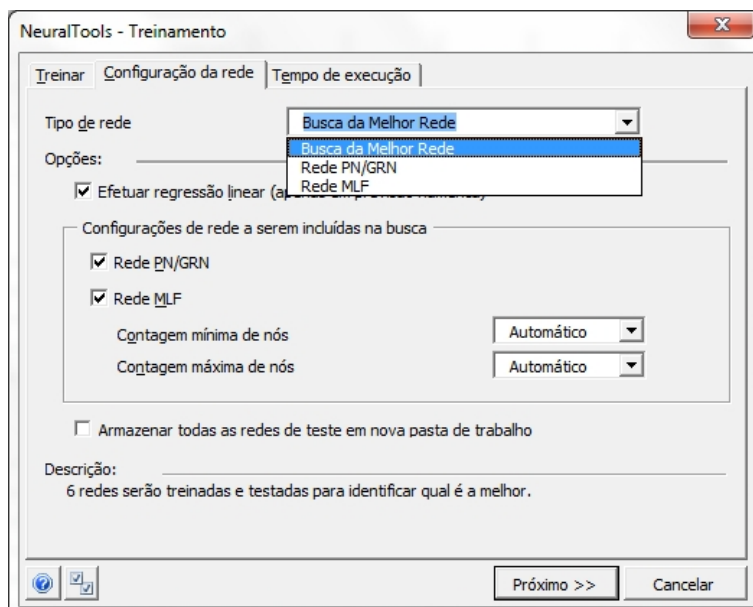


Figura 37 – Caixa de diálogo Configuração da rede

Existem diversas opções para estabelecer o tipo de rede que será gerado pelo NeuralTools, dependendo do problema a modelar. As RNA permitem resolver problemas numéricos ou de classificação. A lista de opções oferece três possibilidades de escolha para o treino da rede. Para a previsão de classificação/categoria existem dois tipos de redes disponíveis: Redes Neurais Probabilísticas (PNN) e redes Multi-Layer Feedforward (MLF). A previsão numérica pode ser realizada usando-se redes MLF assim como Redes Neurais de Regressão Generalizada (GRNN), que são semelhantes às redes PNN.

O NeuralTools possibilita ainda um recurso de busca da Melhor Rede. Quando essa opção é selecionada, o NeuralTools treina e testa uma variedade de configurações de redes neurais até gerar a que produz as melhores previsões para os seus dados. A melhor configuração é determinada com base nos resultados obtidos no teste. Quando se está a começar é aconselhável usar a opção busca da melhor rede.

Tempo de execução

Se clicar no separador Tempo de Execução aparece a seguinte caixa de diálogo:

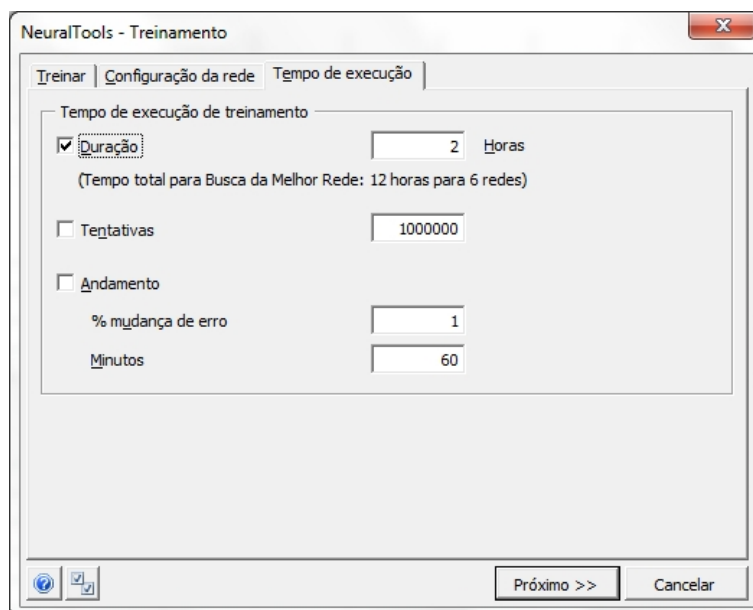


Figura 38 – Caixa de diálogo tempo de execução

Esta caixa permite configurar a duração do treino. Para isso são disponibilizados três critérios de paragem (duração, nº de tentativas e andamento – função da % mínima de mudança de erro num determinado período de tempo), sendo obrigatório definir pelo menos um, a duração. É muito difícil saber quanto tempo será necessário para treinar devidamente uma rede neuronal. Inicialmente poderá ser preferível aceitar os padrões existentes por definição. Com a aquisição de alguma experiência no treino de redes semelhantes é aconselhável definir critérios de tempo mais curtos desde que não interfira com o desempenho da rede.

Em qualquer das caixas de diálogo anteriores pode-se clicar no botão “Próximo”, assim que se tiverem escolhido todas as definições do algoritmo de treino, e aparecerá a seguinte caixa de diálogo.

Visão geral de treino

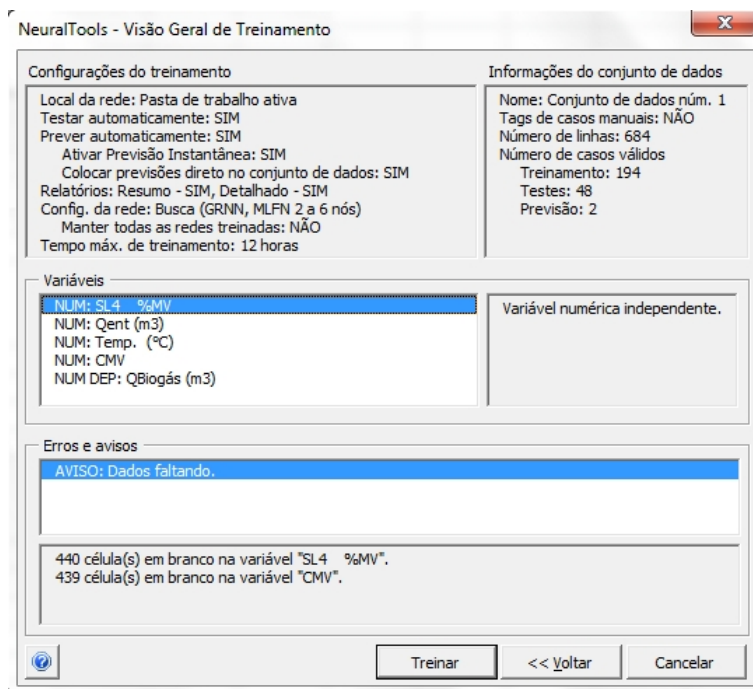


Figura 39 – Caixa de diálogo Visão Geral de Treino

Esta caixa de diálogo apresenta um resumo do que será realizado durante o treino da rede, de acordo com as opções escolhidas anteriormente. Como o treino é o processo mais demorado na modelação da rede neural, convém examinar as definições do treino antes de prosseguir. O NeuralTools tenta verificar se existem alguns problemas com os dados. Caso exista alguma anomalia é transmitindo um aviso na caixa “erros e avisos” para que se possa corrigir, se necessário, antes de iniciar o treino. Em seguida, pode-se clicar no botão treinar para começar o treino.

Processo de treino

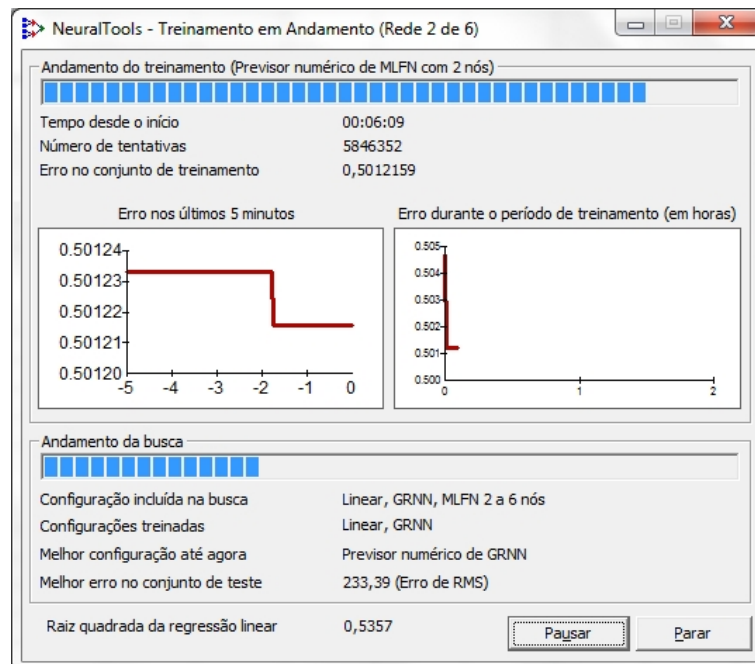


Figura 40 – Caixa de diálogo “Andamento do Treino”

Enquanto o NeuralTools efetua o treino da rede neural, é possível visualizar o andamento do mesmo, através dos gráficos que são atualizados em tempo real. O treino é interrompido quando é atingido algum dos critérios de paragem definidos anteriormente, como, por exemplo, a duração máxima do treino.

Relatório de treino

Os relatórios de treino gerados automaticamente no final do treino, permitem analisar os resultados da rede. Indicam qual o desempenho da rede treinada, fornecendo dados estatísticos como a % de más previsões, valores de diferentes tipos de medição do erro.

NeuralTools: Treinamento, Teste Automático e Previsão Automática de rede neural

Executado por: Geral
 Data: terça-feira, 9 de Julho de 2013 02:57:13
 Conjunto de dados: Conjunto de dados núm. 1
 Rede: Rede treinada em Conjunto de dados núm. 1

| Resumo | |
|--|---|
| Informações da rede | |
| Nome | Rede treinada em Conjunto de dados núm. 1 |
| Configuração incluída na busca | GRNN, MLFN 2 a 6 nós |
| Melhor configuração | Previsor numérico de MLFN (4 nós) |
| Local | Esta pasta de trabalho |
| Variáveis categóricas independentes | 0 |
| Variáveis numéricas independentes | 4 (SL4 %MV, Qent (m3), Temp. (°C), CMV) |
| Variável dependente | Var. numérica (QBiogás (m3)) |
| Treinamento | |
| Número de casos | 194 |
| Tempo de treinamento | 00:08:20 |
| Número de tentativas | 4942080 |
| Motivo da parada | Parada efetuada pelo usuário |
| % previsões inadequadas (15% tolerância) | 26,2887% |
| Erro médio quadrático | 172,10 |
| Erro médio absoluto | 125,09 |
| Desvio padrão de erro absoluto | 118,21 |
| Teste | |
| Número de casos | 48 |
| % previsões inadequadas (15% tolerância) | 43,7500% |
| Erro médio quadrático | 226,37 |
| Erro médio absoluto | 168,03 |
| Desvio padrão de erro absoluto | 151,69 |
| Previsão | |
| Número de casos | 2 |
| Previsão Instantânea ativada | SIM |
| Conjunto de dados | |
| Nome | Conjunto de dados núm. 1 |
| Número de linhas | 684 |
| Tags de casos manuais | NÃO |
| Análise de impacto de variáveis | |
| SL4 %MV | 31,8034% |
| Qent (m3) | 24,9736% |
| CMV | 22,7757% |
| Temp. (°C) | 20,4472% |

| Busca da Melhor Rede | | | |
|----------------------|-------------|----------------------|---------------------------------|
| | Erro de RMS | Tempo de treinamento | Motivo da parada do treinamento |
| Previsor linear | 266,53 | 00:00:00 | Parada automática |
| GRNN | 233,39 | 00:00:00 | Parada automática |
| MLFN 2 nós | 265,09 | 00:07:17 | Parada efetuada pelo usuário |
| MLFN 3 nós | 255,72 | 00:04:44 | Parada efetuada pelo usuário |
| MLFN 4 nós | 226,37 | 00:08:20 | Parada efetuada pelo usuário |
| MLFN 5 nós | 307,73 | 00:04:34 | Parada efetuada pelo usuário |
| MLFN 6 nós | 237,43 | 00:10:29 | Parada efetuada pelo usuário |

| Previsor linear comparado a rede neural | | |
|---|-----------------|-------------|
| | Previsor linear | Rede neural |
| R-quadrada (Treinamento) | 0,5357 | - |
| Erro médio quadrático (Treinamento) | 275,56 | 172,10 |
| Erro médio quadrático (Teste) | 266,53 | 226,37 |

Figura 41 – Tabelas geradas no relatório do Neural Tools

Neste relatório podem ser disponibilizados também gráficos e histogramas das fases de treino e teste. Quando se seleciona a opção “calcular impacto das variáveis” surge um gráfico de barras com os valores da percentagem de impacto de cada variável.

A partir dos dados do relatório e dos gráficos é possível proceder à interpretação dos resultados.

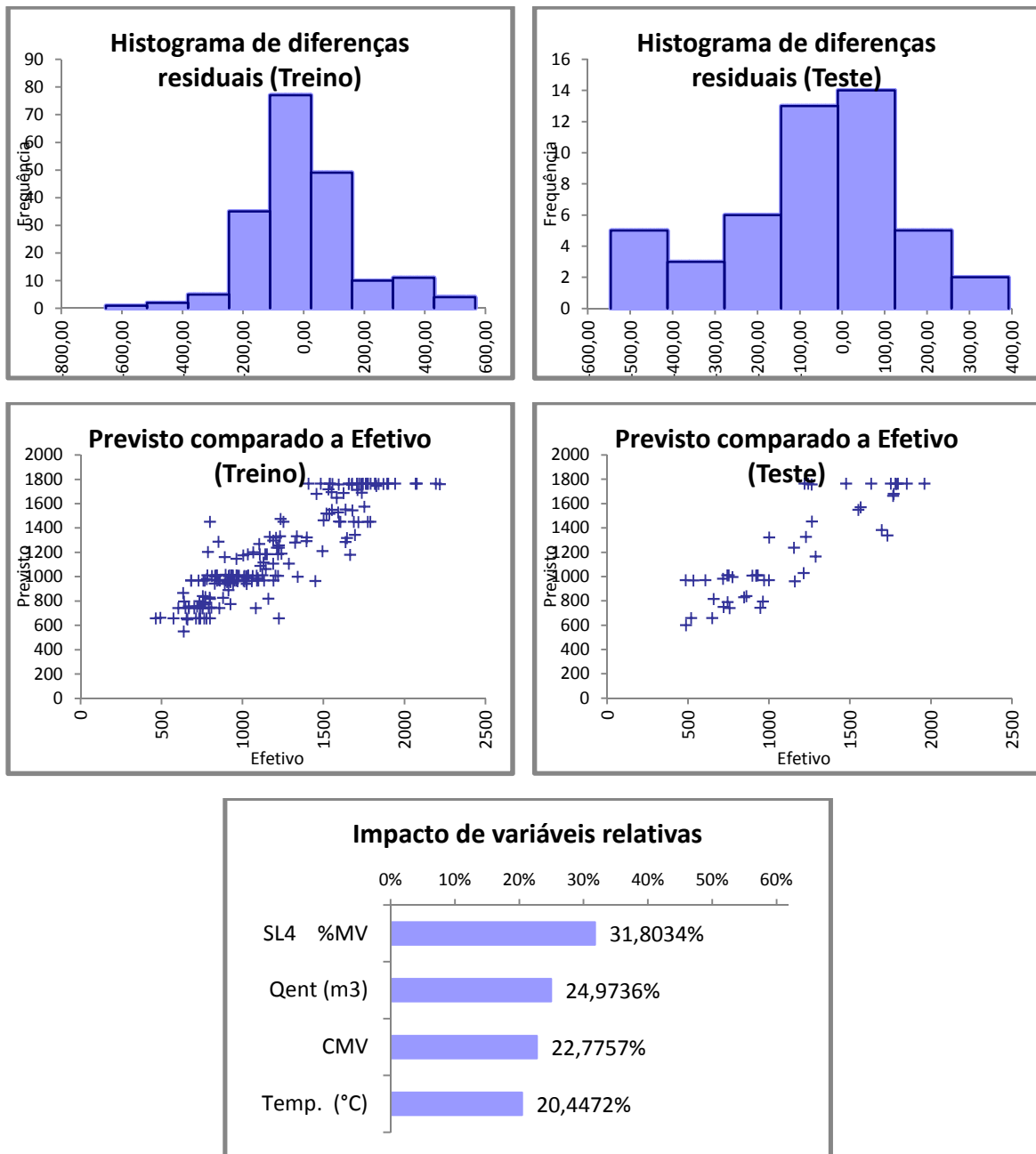


Figura 42 – Exemplos de gráficos gerados no relatório do Neural Tools

5.2.5. Previsão

Uma rede neural treinada é usada para prever valores de output desconhecidos. Depois de treinada e testada, a rede pode ser usada para prever outputs de dados de novos casos.

O NeuralTools disponibiliza dois métodos de previsão:

1) Método de previsão para valores desconhecidos da variável dependente de um conjunto de dados, através do comando de previsão do NeuralTools

2) Previsão instantânea (somente nas versões Industrial e Estudante), em que os valores das variáveis independentes de determinado caso podem ser modificados e o NeuralTools calcula automaticamente o valor de output previsto.

A caixa de diálogo Previsão ajuda a definir o processo de previsão de valores para um determinado conjunto de dados. Só é possível fazer a previsão para casos em que faltam os valores de output. Os valores de output previstos aparecem ao lado dos casos para os quais foi efetuada a previsão.

Previsão Instantânea

A previsão instantânea permite modificar os dados de *input* e observar as alterações que essas mudanças provocam nas previsões.

Quando a Previsão Instantânea está ativada, o NeuralTools acrescenta automaticamente uma fórmula de Excel na célula que contém o valor previsto. Portanto, se os valores das variáveis independentes de um caso específico forem mudados, o valor previsto será automaticamente recalculado, por essa fórmula. Ao usar a Previsão Instantânea, basta digitar os dados referentes aos novos casos diretamente no Excel e gerar uma nova previsão automaticamente, sem passar pela caixa de diálogo Previsão.

5.3. Desenvolvimento do estudo

5.3.1. Introdução

Antes, ainda, de estar definida a base de dados, foram realizados alguns testes iniciais com um reduzido conjunto de casos recolhidos sem qualquer preparação ou tratamento inicial, de forma a estabelecer um primeiro contacto com este software e para adquirir alguma sensibilidade no treino das RNA. Estes primeiros testes foram realizados com o acompanhamento do orientador desta dissertação.

Uma vez que o software de RNA (NeuralTools), utilizado neste estudo, apenas estava disponível no ISEP e isso implicaria deslocações constantes entre o estabelecimento de ensino no Porto e a ETAR Sul em Ílhavo, decidiu-se proceder à aquisição deste software na versão estudante para o desenvolvimento das redes no local de estágio.

5.3.2. Preparação dos dados

No decorrer do processo de operação da ETAR são recolhidos diversos dados, através de um plano de amostragem, relacionados com o funcionamento da ETAR. A empresa (SIMRIA) implementou um *software* (NAVIA) que permite o armazenamento dos dados recolhidos ao longo do tempo, criando, assim, uma considerável base de dados disponíveis para o treino das RNA. A quantidade e qualidade dos dados disponíveis é um fator muito importante para o sucesso do treino de RNA. Parte-se do princípio que, quanto maior for a base de dados, melhor será o treino e a capacidade de previsão da rede ou, pelo menos, espera-se que, com um maior número de casos, seja mais fácil para a rede perceber as relações existentes entre os dados. Com um número insuficiente de exemplos, tanto o desempenho da rede como a sua capacidade de generalização serão pobres.

Existem diferentes tipos de dados recolhidos e com periodicidades diferentes. Alguns dados

são recolhidos *in loco* e outros são obtidos através de análises laboratoriais às amostras de efluente e lamas recolhidas pelos operadores. Os resultados dessas análises são carregados no mesmo programa. Este *software* permite consultar o histórico dos dados recolhidos ao longo dos últimos anos.

Na definição das variáveis e dados a recolher foram considerados alguns critérios importantes como a disponibilidade e fiabilidade dos dados e o impacto esperado das variáveis no processo de digestão. Na prática, a seleção dos *inputs* de uma rede neuronal deve ser baseada nas seguintes considerações:

- a) Conhecimento prévio sobre o processo a modelar
- b) Disponibilidade e qualidade dos dados necessários para o conjunto de treino.

Sendo assim, as primeiras semanas de estágio na ETAR Sul foram dedicadas à compreensão do processo de digestão anaeróbia. Para isso, foi fundamental o apoio do co-orientador de estágio no esclarecimento das muitas dúvidas que foram sendo colocadas e na sugestão de alguma bibliografia para consulta.

Conforme mencionado, anteriormente, no presente trabalho, Lingireddy e Brion (2005) defendem que uma aplicação bem-sucedida de um modelo de RNA requer uma preparação adequada dos dados de input, assim como, uma sólida compreensão do processo a ser modelado é essencial para a correta seleção das variáveis de entrada. Esta compreensão fundamental não só ajudará na escolha de dados de *input* apropriados, mas também ajudará a evitar a escolha de dados com erros ou dados menos fiáveis, que poderão confundir o processo de treino.

Portanto, com o apoio do co-orientador, foi definido um conjunto de dados representativos de toda a gama de funcionamento do processo de digestão. Todos os parâmetros que foram recolhidos e utilizados como entradas na modelação das redes neuronais de previsão da produção de biogás estão apresentados na seguinte tabela.

Tabela 18 – Variáveis definidas para o treino das RNA

| Nome da variável ¹ | Periodicidade de amostragem | Tipo de variável |
|--------------------------------------|--|------------------------|
| % MS SL4 (entrada do digestor) | 2/semana (até Janeiro de 2012) 3/semana (a partir de Fevereiro de 2013) | Resultado Laboratorial |
| % MV SL4 (entrada do digestor) | 2/semana (até Janeiro de 2012) 3/semana (a partir de Fevereiro de 2013) | Resultado Laboratorial |
| MV/MS SL4 (entrada do digestor) | 2/semana (até Janeiro de 2012) 3/semana (a partir de Fevereiro de 2013) | Calculada |
| % MS SL5 (recirculação de lamas) | 2/semana | Resultado Laboratorial |
| % MV SL5 (recirculação de lamas) | 2/semana | Resultado Laboratorial |
| MV/MS SL5 (recirculação de lamas) | 2/semana | Calculada |
| Eficiência da Digestão | 2/semana | Calculada |
| pH | 2/semana | Resultado Laboratorial |
| Alcalinidade | 2/semana | Resultado Laboratorial |
| AGV | 2/semana | Resultado Laboratorial |
| AGV/Alcalinidade | 2/semana | Calculada |
| Q _{ent} | Diária | Medida |
| TRH | Diária | Calculada |
| Temperatura | Média diária de 3 medições | Medida |
| CMV | Diária | Calculada |
| Q _{biogás} | Diária | Medida |

Os parâmetros utilizados serão descritos em seguida:

- %MS – representa a percentagem de matéria sólida presente nas lamas.
- %MV – representa o conteúdo em matéria orgânica da lama. Fração passível de ser digerida.
- MV/MS – permite verificar qual o potencial do resíduo para a digestão.
- Eficiência da Digestão (ED) – é um indicador da percentagem de matéria orgânica que é digerida no digestor. É calculada pela seguinte equação:

¹ Os códigos SL4 e SL5 presentes em algumas variáveis referem-se à localização da recolha das amostras:

SL4 – entrada do digestor

SL5 – ponto na recirculação de lamas

$$ED = 1 - \frac{\%MV (SL5)}{\%MV (SL4)}$$

Fórmula 4 – Eficiência da Digestão

- pH – representa o valor do pH das lamas completamente misturadas dentro do digestor.
- Alcalinidade – representa a capacidade de resistência a alterações de pH causadas pela adição de ácidos (acidificação). A alcalinidade resulta da presença de hidróxidos, carbonatos e bicarbonatos e é expressa em termos de carbonato de cálcio. A concentração da alcalinidade em águas residuais é importante em sistemas de remoção biológica de nutrientes, como é o caso da DA. A estabilidade do digestor é reforçada por uma elevada concentração da alcalinidade.
- Acidos Gordos Voláteis (AGV) – são os principais produtos intermédios formados durante o processo de DA. A presença de altas concentrações de ácidos voláteis provoca a diminuição da concentração da alcalinidade e uma descida do pH, podendo tornar o meio tóxico para algumas bactérias presentes no processo de digestão.
- AGV/Alcalinidade – conforme mencionado anteriormente os AGV provocam a diminuição da alcalinidade, por isso, o quociente entre os valores de AGV e a alcalinidade é importante para monitorizar o processo, não devendo ser superior a 0,2.
- Caudal de entrada no digestor (Q_{ent}) – representa o volume de lamas que entram no digestor ao longo de um dia. Por isso é apresentado em m³/dia. Está dependente da quantidade e características de efluente que é encaminhado para a ETAR.
- Tempo de Retenção Hidráulica (TRH) – o TRH é um parâmetro fundamental para o dimensionamento de um sistema de tratamento anaeróbio e exprime o tempo médio que um determinado volume de lamas permanece no digestor anaeróbio. Para um reator com um volume V (m³), sujeito a uma alimentação com um caudal Q (m³/dia),

o TRH é obtido pela expressão representada na fórmula 1 deste trabalho.

- Temperatura – refere-se à temperatura das lamas dentro do digestor. É usado o valor médio de três medições de temperatura efetuadas ao longo do dia. Conforme já foi referido no presente trabalho o controlo e a manutenção de uma temperatura constante no digestor é crucial para o bom funcionamento do processo anaeróbio, uma vez que flutuações pontuais e/ou frequentes da temperatura podem provocar efeitos severos. Neste caso de estudo o digestor opera em regime mesófilo num intervalo de temperaturas entre 30 a 35 °C. A operação fora de um intervalo ótimo de temperatura (acima ou baixo) pode provocar inibição do processo de digestão. Na modelação das redes neuronais foram experimentadas algumas médias móveis da temperatura para diferentes durações. No entanto não se verificaram alterações significativas nos resultados obtidos. Sendo a temperatura influenciada pelas diversas reações e outros fatores que ocorrem nos digestores, concluiu-se que o valor médio da temperatura diária representa, por si só, uma média dos diversos parâmetros que interferem no processo de digestão, não sendo necessária a utilização de médias móveis desta variável, conforme se veio a verificar. Este raciocínio aplica-se a todas as variáveis referentes ao interior do digestor.

- Carga Mássica Volúmica (CMV) – a CMV é um parâmetro fundamental para a monitorização do processo de digestão anaeróbia e relaciona a quantidade de matéria orgânica alimentada num dia com o volume do digestor, pela seguinte expressão:

$$CMV = \frac{Q_{ent} \cdot \%MV \cdot 10}{V}$$

Fórmula 5 – Carga Mássica Volúmica

- Qbiogás – este parâmetro representa o caudal diário de produção do biogás que é produzido pela DA. Sendo apresentado em m³/dia.

Como o principal objetivo da modelação do processo de digestão anaeróbia através de redes neuronais é a otimização da produção de biogás, o $Q_{\text{biogás}}$ foi definido como output das RNA e as restantes variáveis representam os inputs.

Entretanto, foi definido como objeto do estudo do processo de DA o digestor A da ETAR Sul, pelos seguintes motivos:

1. Este digestor, após paragem para uma intervenção de reabilitação ao nível da construção civil (sobretudo betão), tem trabalhado em contínuo desde Abril de 2011.
2. O reinício de funcionamento deste digestor coincidiu com a substituição do caudalímetro de medição do biogás, passando de um caudalímetro mecânico para um caudalímetro termomássico, com garantia de maior fiabilidade na leitura dos resultados obtidos.
3. No início deste estágio o outro digestor disponível na ETAR Sul (digestor B) encontrava-se em fase de desativação.

Pelo que foi referido anteriormente, os dados recolhidos para o treino e teste das redes neuronais a serem modeladas, referem-se apenas ao Digestor A e estão compreendidos entre Abril de 2011 e Fevereiro de 2013. Desta forma foi possível elaborar a base de dados que consta no Anexo I. Durante a elaboração desta base de dados foram analisados os valores das variáveis, caso a caso, para tentar despistar possíveis erros de medição ou digitação de valores.

5.3.3. Treino e Teste

Após a preparação dos dados e com uma base de dados definida, estavam reunidas as condições para iniciar a fase de treino e teste na modelação das redes neuronais.

Antes de se iniciar a fase de treino devem ser consideradas algumas limitações das RNA. No conjunto de variáveis de *input* não se devem apresentar variáveis que estejam

diretamente relacionadas, que partilhem algum dos parâmetro que as definem e por isso variem da mesma forma ou sejam inversamente proporcionais, como por exemplo, o Q_{ent} e o TRH. Estas variáveis são inversamente proporcionais e o TRH varia em função do Q_{ent} . A colocação destas duas variáveis no mesmo conjunto de dados pode levar a rede a assumir esta relação como muito importante e atribuir-lhes maior importância do que às restantes variáveis, desvirtuando o resultado final.

Apesar de, na diversa bibliografia sobre RNA, estas duas etapas serem descritas em separado, o NeuralTools permite que estas aconteçam em simultâneo, da perspectiva do utilizador. Isso é bastante útil, no uso da ferramenta, por tornar desnecessária mais uma intervenção do utilizador, após o treino, embora internamente a ferramenta desenvolva o treino primeiro e o teste depois. Para beneficiar dessa solução, quando se procede à definição de um conjunto de dados, é possível escolher as percentagens ou quantidade de casos do conjunto de dados que serão reservadas para o treino e para o teste da rede neuronal. Com base nos valores definidos pelo utilizador, o Neural Tools escolhe aleatoriamente os casos que usa para treino e os que usa para teste. No modo de procura da melhor rede é escolhida, como melhor rede, aquela que apresentar menor erro médio quadrático (EMQ) na fase de teste.

Sendo assim, daqui em diante, sempre que forem referidos os termos treino ou teste de uma rede, referem-se ao treino e teste em simultâneo.

Primeiro conjunto de redes

Para a modelação das primeiras redes, foram definidos conjuntos de dados com um grande número de variáveis de entrada, escolhidas de forma mais ou menos aleatória, para perceber o comportamento do programa, a capacidade das RNA e quais as variáveis com maior impacto no treino e nos resultados de previsão. Para isso, o Neural Tools

disponibiliza no final de cada treino um relatório completo no qual são indicados os valores das diferentes medições de erro para as redes treinadas, compara-os com a regressão linear (quando solicitado) e analisa o impacto de cada variável na modelação da rede.

Estas redes foram treinadas no modo de busca da melhor rede. Esta opção do Neural Tools permite treinar e testar uma variedade de possíveis configurações até identificar a que apresenta o melhor desempenho. Previsões numéricas podem ser realizadas usando redes Multi-Layer Feedforward (MLF) ou Redes Neurais de Regressão Generalizada (GRNN). Por definição, em cada busca da melhor rede foram treinadas redes GRNN e MLF de 2 a 6 nós numa camada oculta, comparando os resultados com uma regressão linear. A principal vantagem das redes GRNN reside na sua rapidez de processamento, quase instantânea, quando comparadas com as MLF, que são muito mais lentas, podendo demorar muitos minutos, ou até horas, para encontrar a melhor solução.

Ao longo das diversas tentativas realizadas foram-se excluindo as variáveis com menor impacto nos resultados, diminuindo a quantidade de *inputs* das redes e alternando algumas das variáveis no conjunto de dados. Foi possível constatar que, geralmente, são alcançados melhores resultados nas redes com menor número de *inputs*. A análise dos resultados obtidos permitiu identificar um conjunto de variáveis com maior impacto no treino das redes neuronais que alcançaram melhores resultados na modelação da produção de biogás. Lingireddy e Brion (2005) referem que, inicialmente, com o intuito de proporcionar um modelo robusto, podem ser usadas tantas variáveis quanto possível, que podem ou não ter um impacto significativo sobre os resultados. Isto é possível porque as redes neuronais, normalmente, geram um modelo adequado mesmo quando são fornecidos alguns *inputs* pouco fiáveis. No entanto, a fim de se encontrar um modelo parcimonioso, o número de variáveis utilizadas na modelação da rede deve ser otimizado. Essa otimização pode ser efetuada da seguinte forma:

1. Modelar uma rede neuronal usando todas as variáveis de entrada possíveis;
2. Eliminar, progressivamente, algumas variáveis de entrada, de acordo com o seu impacto no resultado final (removendo as que apresentarem pouco contributo, ou mesmo nenhum);
3. Verificar o efeito da redução do número de variáveis na modelação da rede;
4. Remodelar a rede neuronal, com um reduzido número de variáveis, experimentando diferentes combinações, até se obter o melhor resultado possível.

A seleção das variáveis de entrada é um ponto muito importante para a modelação de uma rede neuronal com sucesso.

Em algumas das redes experimentadas optou-se pela normalização dos dados de entrada convertendo os valores iniciais de forma a aproximar as diferentes grandezas, para evitar a atribuição de maior importância a variáveis com valores de maior grandeza. Esta preocupação é refletida em alguma da bibliografia estudada. No entanto, esta normalização não teve o efeito esperado, não se verificando melhorias significativas. Pois, conforme mencionado no manual do programa, o Neural Tools tem um algoritmo incorporado que faz precisamente essa tarefa, procede à normalização dos dados, de forma a evitar este efeito de atribuição de maior importância a variáveis com valores mais altos.

A melhor rede encontrada neste primeiro processo foi a rede designada por “Teste RNA12”, constituída por 5 variáveis de entrada – temperatura, %MV (em SL4 e SL5), TRH e CMV – e como variável de saída o $Q_{\text{biogás}}$.

Na análise dos relatórios gerados pelo Neural Tools para estas redes foi possível detetar um valor de erro bastante desviado dos restantes, o que permitiu detetar um erro grosseiro num dos casos da base de dados. Após a retificação deste erro procedeu-se, novamente, ao treino de um novo conjunto de redes com base nos melhores resultados obtidos anteriormente.

Pode ser consultada, no AnexoII, uma tabela com os principais resultados deste conjunto inicial de redes treinadas. Um exemplo dos relatórios gerados pelo NeuralTools para cada rede será inserido no AnexoIII.

Segundo conjunto de redes

Os resultados obtidos anteriormente permitiram definir um conjunto mais reduzido de variáveis com maiores garantias de sucesso na modelação do processo de digestão anaeróbia da ETAR Sul. Então, decidiu-se testar algumas redes com diferentes combinações dessas variáveis. Entre as variáveis consideradas mais representativas do processo de digestão encontram-se: o Q_{ent} ; a %MV e %MS em SL4 e SL5; a CMV; o TRH e a Temperatura. Nos testes deste segundo conjunto de redes optou-se por experimentar algumas alterações, nomeadamente, a exclusão dos casos em que a Eficiência da Digestão (ED) era igual ou inferior a zero e a introdução de diferentes médias móveis da temperatura sem que, conforme referido anteriormente na descrição das variáveis, se tenham verificado alterações significativas nos resultados obtidos, devido ao motivo já indicado. No entanto, as diferentes redes experimentadas obtiveram resultados mais constantes e ligeiramente melhores do que as anteriores. Foi, inclusivamente, encontrada uma rede com menor erro do que a melhor rede encontrada anteriormente, a rede “Treino RNA12”, que em relação à anterior apenas difere numa variável, trocando o TRH pelo Q_{ent} de lamas no digestor.

O facto de se excluírem as variáveis em que a ED era neutra ou negativa aliado à escolha de variáveis presentes nos dados em menor número – como é o caso da %MV, que apenas é determinada duas a três vezes por semana – representa um fator menos favorável para o treino das redes, por diminuir em demasia a quantidade de casos disponíveis no conjunto de dados. Sabendo que a capacidade de generalização das RNA tende a aumentar com o

tamanho do conjunto de dados, ao usarem-se conjuntos de dados com menor número de casos, o desempenho de rede poderá ser afetado.

Foi elaborada uma tabela com os resultados destas redes que consta no AnexoIV.

Terceiro conjunto de redes

Analisando os resultados obtidos, optou-se por reduzir ainda mais o número de *inputs* a considerar no treino das próximas redes, alternando diferentes combinações. Estas redes foram treinadas novamente no modo de busca da melhor rede. Neste conjunto de tentativas foram obtidos resultados bastante animadores, nomeadamente, em três redes (“TrTsRNA3”, “TrTsRNA4” e “TrTsRNA9”), sendo uma delas constituída por 4 inputs e outras duas apenas por 3 inputs. Pode-se ainda concluir que as variáveis com maior impacto na modelação do processo anaeróbio são a Temperatura das lamas, a %MV na entrada do digestor, a CMV, e o Q_{ent} ou o TRH. Os resultados obtidos por estas redes podem ser consultados no AnexoV, inseridos numa tabela simplificada para facilitar a análise dos mesmos. Nessa tabela, assim como nas anteriores, podem-se consultar os valores do EMQ e a percentagem de más previsões em cada fase, as variáveis de entrada para cada rede e o impacto de cada uma na modelação da rede, entre outras informações.

Última fase – treino das melhores redes

Na expectativa de conseguir melhorar os resultados obtidos por estas três redes, foram testadas, consecutivamente, para cada uma delas, diversas redes com a mesma topologia e as mesmas variáveis, de acordo com os resultados obtidos na procura da melhor rede. Durante o treino destas redes foi sendo alterada a percentagem de tolerância para más previsões, para verificar se teria algum impacto no resultado final. Constatou-se que esta alteração não produz qualquer efeito nos resultados do EMQ obtido pelas redes. Apenas

faz diminuir ou aumentar a percentagem de más previsões conforme se aumente ou diminua a tolerância.

De facto, foi possível melhorar os resultados obtidos até então, encontrando-se a melhor rede desenvolvida durante este estudo, a rede “BNTrTs RNA3.16”, com as variáveis – Temperatura, Qent, %MV (SL4) e CMV.

As tabelas com os resultados obtidos para as três topologias de rede encontram-se no AnexoVI, assim como, a tabela dos dados usados e o relatório gerado na melhor rede (BNTrTs RNA3.16).

5.3.4. Previsão e Validação

Para verificar a capacidade de generalização e o desempenho das melhores redes neuronais encontradas para modelar o processo de digestão anaeróbia, foram definidos três novos conjuntos de dados (dois para previsão e um para validação) que as redes não conheceram durante a fase de treino, para efetuar previsões de produção de biogás, ou seja, prever o $Q_{\text{biogás}}$ produzido associado a estes *inputs* sem conhecer os valores de *output*.

Previsão

Com os dois primeiros conjuntos de dados elaborados para previsão da produção de biogás não se pretendia efetuar qualquer comparação com outros valores conhecidos, pretendeu-se apenas analisar o comportamento da rede na previsão de produção de biogás, fazendo variar de forma constante e alternadamente o valor de cada uma das variáveis de entrada, para que fosse possível verificar o desempenho da rede na previsão de valores de output, ao longo de toda a gama de valores do processo anaeróbio do digester.

Estes dois conjuntos não foram baseados em dados reais. Foram definidos valores dentro da gama de funcionamento do digester, uma vez que as RNA não têm capacidade para

extrapolar. O estudo destas previsões foi efetuado apenas para a melhor rede encontrada. Os valores das variáveis foram organizados em tabelas e elaboraram-se diversos gráficos com os resultados obtidos pela melhor rede encontrada, que permitiram analisar e avaliar o desempenho da rede de forma mais intuitiva. Um bom exemplo disso, são os gráficos tridimensionais elaborados para os valores do $Q_{\text{biogás}}$ previstos para as diferentes temperaturas, variando num eixo a %MV e no outro o TRH (que varia de forma inversamente proporcional ao Q_{ent}). Este e outros gráficos interessantes para a análise dos resultados podem ser consultados no AnexoVII.

Da análise destes gráficos resulta a conclusão de que as redes conseguiram aproximar-se do comportamento do processo de digestão anaeróbia, sendo possível definir gamas de valores de funcionamento ótimo.

Validação

Para efetuar uma comparação entre os valores previstos pelas redes e os valores reais, foi criado um conjunto de dados de validação. Este conjunto de dados é composto pelos valores de funcionamento do digester, obtidos nos meses de Março e Abril de 2013. O conjunto de validação permite avaliar a capacidade de generalização da rede.

Comparando os resultados previstos pelas redes com os valores reais neste conjunto, verifica-se que as redes neuronais sentem alguma dificuldade em se aproximar de alguns casos particulares, mas revelam uma tendência relativamente próxima da realidade. De qualquer forma, nesses casos em que a diferença de resultados é maior, face aos valores de *input* disponíveis para treino e análise das redes, associados a esses resultados, o comportamento da rede neuronal está mais próximo do que seria espectável teoricamente, do que os valores realmente obtidos no processo anaeróbio.

Identificando e analisando um destes casos particulares, verificamos que quando a %MV

(SL4), que representa a porção de matéria orgânica que está a entrar no digestor, aumenta e as restantes variáveis não sofrem alterações significativas, seria de prever que a produção de biogás aumentasse perante estas condições, conforme previsto pela rede, mas na realidade não foi isso que aconteceu no processo de DA.

Estas diferenças podem dever-se a diversos fatores, entre eles, erros induzidos durante a recolha de amostras ou valores recolhidos que podem não ser representativas da realidade do processo, assim como, a necessidade de incluir outras variáveis que não foram consideradas na modelação da rede, mas que poderão representar um contributo a considerar no processo de digestão.

A base de dados criada para a validação das redes, os valores obtidos nas previsões das redes e os gráficos de comparação entre estes valores e a realidade estão contidos no AnexoVIII.

Para se perceber a capacidade de generalização efetiva das melhores redes, foi elaborada uma base de dados com as médias mensais das variáveis de entrada, usadas em cada rede, com o intuito de diluir alterações significativas, mas pontuais, nos dados.

O resultado obtido não deixa de ser surpreendente, pela aproximação verificada entre os valores previstos e os valores reais, atestando a capacidade de generalização das redes. A tabela e o gráfico obtido com esses valores estão inseridos no AnexoIX.

Por último, resolveu-se efetuar outra experiência. Foram recolhidos alguns dados da digestão anaeróbia de outra ETAR da mesma empresa, a ETAR Norte, para previsão na rede desenvolvida para a ETAR Sul. A ETAR Norte apresenta características, quer de efluente, como dos próprios digestores, completamente diferentes da ETAR Sul. Com esta tentativa procurou-se testar a possibilidade de aplicar a mesma rede a outros processos de DA. Como seria de esperar, os resultados obtidos permitiram concluir que não é possível aplicar uma rede modelada para um determinado processo a outros processos com

características diferentes, conforme se pode verificar no gráfico apresentado no AnexoX. Confirma-se, assim, a incapacidade de extrapolação das RNA. Para se modelar outro processo tem que se percorrer o mesmo caminho desenvolvido neste trabalho. Preparar os dados, treinar e testar várias redes até se obter a melhor solução possível, efetuar previsões e validar o modelo obtido.

A validação de todos os resultados obtidos implicou uma análise cuidada dos mesmos, contando com o apoio de colegas da Direção de Engenharia da AdP, além do orientador e do co-orientador do trabalho, com vista a averiguar a fiabilidade desses resultados.

6. CONSIDERAÇÕES FINAIS

6.1. Conclusões

A otimização do processo de Digestão Anaeróbia para produção de biogás é essencial para a rentabilidade de exploração de uma estação de tratamento de águas residuais, devido aos elevados consumos energéticos associados às operações de tratamento. O recurso a uma fonte de energia renovável e com baixo custo associado, como o biogás, em detrimento de energias fósseis ou, até mesmo, de outras energias renováveis, apresenta vantagens sob o ponto de vista económico, ecológico e ambiental que não podem ser menosprezadas pelas entidades gestoras de ETAR.

Compreendendo-se a complexidade do processo anaeróbio – do qual não se conhecem as relações existentes entre as diversas variáveis intervenientes, assim como, qual o contributo efetivo de cada uma na produção de biogás – as redes neuronais apresentam-se como uma ferramenta com enorme potencial para a modelação deste processo, devido à capacidade para entender relações complexas existentes num determinado conjunto de dados. Por imitação das funções do cérebro, as RNA conseguem discernir padrões existentes nos dados e, quando são fornecidos dados novos, são capazes de gerar previsões com base nas relações e nos padrões apreendidos durante o treino e teste da rede.

Neste trabalho, explorou-se a aplicação dessas redes à modelação do processo de digestão anaeróbia, através do *software* Neural Tools™ da Palisade. Esta aplicação informática permite desenvolver redes neuronais de forma rápida e bastante intuitiva, num ambiente de trabalho bem conhecido, o Microsoft Excel™.

As redes neuronais desenvolvidas neste estudo – do tipo MLF e GRNN – produziram resultados que se podem considerar bastante interessantes, tendo-se encontrado redes com bom desempenho na modelação do processo de digestão. Analisando os resultados obtidos

pelas redes neuronais, foi possível encontrar um conjunto de variáveis com maior influência no processo de digestão.

A melhor rede obtida foi treinada com as seguintes variáveis de input: %MV, CMV, Temperatura e Caudal de entrada de lamas no digestor (Q_{ent}). As previsões apresentadas pela rede permitem perceber gamas de funcionamento ótimo para diferentes variáveis. No entanto, nem todas as variáveis são passíveis de serem controladas durante a operação do digestor, por representarem características intrínsecas do efluente. Os gráficos elaborados com os resultados das previsões permitem determinar os valores de funcionamento ótimo para as variáveis que é possível controlar, em função das variáveis mencionadas anteriormente.

Em relação aos dados utilizados para validação das redes, alguns valores de previsão ficaram afastados dos valores reais. Mas uma análise pormenorizada dos dados, nesses casos particulares, permitiu perceber que o comportamento da rede coaduna-se com o esperado teoricamente, só que tal reação não veio a acontecer na realidade, por motivos que provavelmente escapam à normalidade do funcionamento de uma DA e cuja ocorrência será, em princípio, pouco habitual. Conforme se pode verificar num caso em que a %MV (SL4) – que representa a porção de matéria orgânica que está a entrar no digestor – aumenta e as restantes variáveis não sofrem alterações significativas, seria de prever que a produção de biogás aumentasse perante estas condições. Esta reação foi prevista pela rede, mas na realidade não foi isso que aconteceu no processo de DA. A razão para isto acontecer pode estar relacionada com erros induzidos durante a recolha de amostras, que podem não ser representativas da realidade do processo. De referir ainda, que não foi possível considerar todas as condicionantes do processo de digestão, como a presença de algumas substâncias tóxicas, por exemplo, a presença de sulfuretos.

No entanto, na validação com valores de médias mensais, das variáveis de entrada, é

possível concluir que as redes apresentam uma muito boa capacidade de generalização, com os valores previstos a ficarem muito próximos dos valores reais.

Em síntese, acredita-se que a utilização de redes neurais artificiais para a modelação do processo de digestão anaeróbia constitui uma temática promissora, podendo configurar-se como uma ferramenta a explorar nas diversas áreas da gestão de sistemas de saneamento básico.

O facto deste estudo ter sido desenvolvido no âmbito de um estágio curricular, representa uma mais-valia, por permitir a integração num ambiente de trabalho, ficando a conhecer a organização e o funcionamento da empresa, nomeadamente, através do acompanhamento das atividades desenvolvidas em alguns departamentos. O acompanhamento da operação da ETAR, permitiu acompanhar e observar o funcionamento das diferentes fases de tratamento de águas residuais, facilitando a compreensão dos diversos fatores inerentes ao processo.

De salientar o agradável ambiente de trabalho e a cordialidade de todos os colaboradores, nos diferentes locais de trabalho, o que facilitou a integração na empresa.

Por outro lado, o presente trabalho, constituiu, por si só, um desafio aliciante e suplementar, por se enquadrar numa temática, até aqui, desconhecida para o autor, por não ter sido abordada na formação anterior, nem ser habitualmente abordada, porventura, na maioria dos cursos de Engenharia Civil.

De facto, as redes neurais artificiais, possivelmente ainda pouco divulgadas, apresentam-se como uma ferramenta com muita potencialidade, com aplicações de sucesso em diferentes áreas, como a engenharia e ambiente, mas também no setor económico-financeiro ou, até, no apoio a diagnóstico médico, conforme se deduz da bibliografia consultada.

6.2. Sugestões para desenvolvimentos futuros

Conforme referido no desenvolvimento do estudo, o facto de algumas variáveis que apresentam impacto significativo no processo de aprendizagem das redes, existirem em menor número, condicionou, de sobremaneira, a quantidade de dados disponíveis para treinar as redes neuronais. Com o intuito de aumentar o número de dados disponíveis para futuros desenvolvimentos destas redes, sugere-se que, caso seja viável e comportável para a empresa, sejam recolhidas e analisadas uma maior quantidade de amostras destas variáveis. Considera-se que um conjunto de dados com maior número de casos poderá contribuir para a obtenção de melhores resultados na imitação do processo de digestão anaeróbia.

Em diversos autores, entre os quais Matos (2008), é referida a complementaridade das redes neuronais com outras aplicações, nomeadamente, na preparação dos dados a serem usados no treino das redes. Ora, aplicações que permitam efetuar a análise e normalização dos dados, poderão contribuir para a melhoria dos resultados obtidos, permitindo identificar e excluir dados menos fiáveis e passíveis de introduzirem erro no treino das redes.

Conforme resulta das observações aqui mencionadas, este trabalho não se encontra fechado, podendo ser desenvolvido no futuro, com possibilidade de melhorar o trabalho aqui apresentado. Espera-se que este trabalho constitua um ponto de partida para a compreensão do complexo processo de digestão anaeróbia e contribua para a otimização do processo.

7. REFERÊNCIAS BIBLIOGRÁFICAS

Angelidaki I. e Sanders W., 2004. Assessment of the anaerobic biodegradability of macropollutants. *Reviews in Environmental Science and Bio/Technology*

Appels L., Baeyens J., Degève J. e Dewil R., 2008. Principles and potential of the anaerobic digestion of waste-activated sludge. *Progress in Energy and Combustion Science*.

Brondino N.C.M., 1999. Estudo da influência da acessibilidade no valor de lotes urbanos através do uso de Redes Neurais. Universidade de São Paulo. (Dissertação de Douturamento)

Di Bernardino S., 2006. Desenvolvimento de um processo tecnológico inovador para o tratamento dos efluentes do processamento de sementes de alfarroba, Tese de dissertação de Doutoramento, FCT-UNL, Lisboa.

ERSAR, 2010. Relatório Anual do Sector de Águas e Resíduos em Portugal. Volume 1 – Caracterização Geral do Sector. Entidade Reguladora dos Serviços de Águas e Resíduos. Lisboa

Fey A. e Conrad R., 2000. Effect of temperature on carbon and electron flow and on the archaeal community in methanogenic rice field soil. *Applied and Environmental Microbiology*

Gascão J. e Carvalho M., 2011. Assessoria Técnica de Apoio à Optimização do Funcionamento da Digestão Anaeróbia da ETAR Sul da SIMRIA. Águas de Portugal.

Gerardi M.H., 2003. The Microbiology of Anaerobic Digesters. John Wiley & Sons, Inc., Hoboken, New Jersey.

Gonçalves C.D.C., 2012. Modelação do Processo de Digestão Anaeróbia da FORSU à escala industrial. Instituto Superior Técnico da Universidade Técnica de Lisboa. (Dissertação de Mestrado)

Haykin S., 1999. Neural Networks – A Comprehensive Foundation. Pearson Education. Singapura

Hwang M. H., Jang N. J., Hyum S. H. e Kim I. S., 2004. Anaerobic bio-hydrogen production from ethanol fermentation: the role of pH. Journal of Biotechnology

Igoni A.H., Ayotamuno M.J., Eze C.L., Ogaji S.O.T. e Probert S.D., 2008. Designs of anaerobic digesters for producing biogas from municipal solid-waste. Applied Energy.

Klerfors D., 1998. Artificial Neural Networks. School of Business and Administration. Saint Louis University.

Lingireddy S. e Brion G.M., 2005. Artificial Neural Networks in Water Supply Engineering. ASCE. Reston, Virginia.

Lens P.N.L., Visser A., Janssen A.J.H., Pol L.W.H. e Lettinga G., 1998. Biotechnological treatment of sulfate-rich wastewaters. *Critical Reviews in Environmental Science and Technology*

Lobo V., 2007. Sistemas de Apoio à Decisão – Redes Neurais (MLP). EN/ISEGI. Volume 1.1.

Mara D. e Horan N., 2003. *The handbook of water and wastewater microbiology*. Academic Press, University of Leeds, Reino Unido.

Matos J.P.G.S.C., 2008. Utilização de Redes Neurais Artificiais para a Modelação da Degradação de Sistemas de Distribuição de Água. Instituto Superior Técnico da Universidade Técnica de Lisboa. (Dissertação de Mestrado)

McInerney M. J., 1999. Anaerobic metabolism and its regulation. In: Rehm H.J. e Reed G. (Eds.), *Biotechnology. Environmental Processes I – Wastewater Treatment*, Wiley – VCH. Volume 11a.

Metcalf & Eddy, 1991. *Wastewater Engineering: Treatment, Disposal and Reuse*. McGraw Hill Book Co., 3ª edição, Edição Internacional, Singapura

Moreira M.A., 1997. *Introdução às Redes Neurais*.

Palisade, 2010. *Neural Tools – Add-In de Redes Neural para o Microsoft Excel. Manual do Utilizador. Versão 5.7*. Palisade Corporation. Ithaca. N.Y.

Santos A.M., Seixas J.M., Pereira B.B., Medronho R.A., 2005. Usando Redes Neurais Artificiais e Regressão Logística Predição da Hepatite A. Ver Bras Epidemiol.

Schink B., 1997. Energetics of syntrophic cooperation in methanogenic degradation. Microbiology and Molecular Biology Reviews.

Schuyler J., 2001. Risk and Decision Analysis in Projects, Second Edition, Project Management Institute, Inc., Pennsylvania.

Silva J.G., 2008. Análise de decisão usando métodos estocásticos – casos de aplicação à gestão na área do ambiente, GESCON 2008. Fórum Internacional de Gestão da Construção. F.E.U.P., Porto, Portugal.

Silva J.G., 2013. Workshop sobre utilização de Redes Neurais Artificiais em empresa da área do Saneamento Básico.

Silveira G.S., 2009. Estudo de co-digestão anaeróbia de lamas domésticas com o efluente da indústria da fermentação da levedura do pão da empresa Mauri Fermentos. Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa. (Dissertação de Mestrado)

SIMRIA, 2011a. Relatório e Contas 2011. SIMRIA – Sistema Integrado de Municípios da Ria, SA. Aveiro.

SIMRIA, 2011b. Relatório de Sustentabilidade 2011. SIMRIA – Sistema Integrado de Municípios da Ria, SA. Aveiro.

Van Haandel, A.C. e Lettinga, G., 1994. Anaerobic sewage treatment – A practical guide for regions with a hot climate. John Wiley & Sons Limitada, Reino Unido.

Zayed G. e Winter J., 2000. Inhibition of methane production from whey by heavy metals - protective effect of sulfide. Applied Microbiology Biotechnology.

ANEXOS

Nota: Os anexos completos, em ficheiros Excel, estão gravados no CD que acompanha este trabalho.

Nestes anexos apenas constam imagens representativas desses ficheiros.

ANEXO I

Bases de Datos

| # | Data | Lamas | | | | | | | | | | | Digestor A | | | | Biogás |
|----|-------------|------------|------------|--------------|-------------|-------------|--------------|------------------------|------------|----------------------|-------------|----------------|---------------------------------------|---------------|---------------|------|--|
| | | SL4 %MS | SL4 %MV | SL4 MV/MS | SL5A %MS | SL5A %MV | SL5 MV/MS | Eficiência Digestão | SL5A pH | SL5A Alcalinidade | SL5A AGV | AGV/A Ical. | Q _{ent} (m ³) | TRH (dias) | Temp. (°C) | CMV | Q _{biogás} (m ³) |
| 1 | 17-Abr-2011 | | | | | | | | | | | | 97 | 30,9 | 34,4 | | |
| 2 | 18-Abr-2011 | 2,3 | 1,7 | 0,74 | 1,5 | 0,9 | 0,60 | 0,47 | 7,6 | 3020 | 250 | 0,08 | 109 | 27,5 | 33,1 | 0,62 | |
| 3 | 19-Abr-2011 | 2,2 | 1,5 | 0,68 | 1,8 | 1,1 | 0,61 | 0,27 | 7,7 | 2570 | 340 | 0,13 | 138 | 21,7 | 33,6 | 0,69 | |
| 4 | 20-Abr-2011 | | | | | | | | | | | | 203 | 14,8 | 33,3 | 0,00 | 1547 |
| 5 | 21-Abr-2011 | | | | | | | | | | | | 130 | 23,1 | 32,8 | 0,00 | 2683 |
| 6 | 22-Abr-2011 | | | | | | | | | | | | 167 | 18,0 | 33,3 | 0,00 | 2584 |
| 7 | 23-Abr-2011 | | | | | | | | | | | | 168 | 17,9 | 33,9 | 0,00 | 2819 |
| 8 | 24-Abr-2011 | | | | | | | | | | | | 168 | 17,9 | 34,6 | 0,00 | 2755 |
| 9 | 25-Abr-2011 | | | | | | | | | | | | 169 | 17,8 | 35,1 | 0,00 | 2983 |
| 10 | 26-Abr-2011 | 2,4 | 1,8 | 0,75 | 1,8 | 1,1 | 0,61 | 0,39 | 7,5 | 2400 | 610 | 0,25 | 152 | 19,7 | 35,5 | 0,91 | 1287 |
| 11 | 27-Abr-2011 | | | | | | | | | | | | 169 | 17,8 | 35,5 | 0,00 | 1177 |
| 12 | 28-Abr-2011 | 1,8 | 1,1 | 0,61 | 2,3 | 1,7 | 0,74 | -0,55 | 7,3 | 2460 | 490 | 0,20 | 167 | 18,0 | 35,2 | 0,61 | 1006 |
| 13 | 29-Abr-2011 | | | | | | | | | | | | 219 | 13,7 | 35,3 | 0,00 | 1103 |
| 14 | 30-Abr-2011 | | | | | | | | | | | | 322 | 9,3 | 30,9 | 0,00 | 1067 |
| 15 | 1-Mai-2011 | | | | | | | | | | | | 353 | 8,5 | 35,6 | 0,00 | 1172 |
| 16 | 2-Mai-2011 | | | | | | | | | | | | 196 | 15,3 | 36,0 | 0,00 | 1463 |
| 17 | 3-Mai-2011 | 2,8 | 2,0 | 0,71 | 1,8 | 1,1 | 0,61 | 0,45 | 7,0 | 2600 | 290 | 0,11 | 200 | 15,0 | 36,3 | 1,33 | 1537 |
| 18 | 4-Mai-2011 | | | | | | | | | | | | 258 | 11,6 | 36,6 | 0,00 | 1292 |
| 19 | 5-Mai-2011 | 2,6 | 2,1 | 0,81 | 1,7 | 1,0 | 0,59 | 0,52 | 7,4 | 2700 | 370 | 0,14 | 244 | 12,3 | 35,7 | 1,71 | 1153 |
| 20 | 6-Mai-2011 | | | | | | | | | | | | 249 | 12,0 | 35,5 | 0,00 | 1011 |
| 21 | 7-Mai-2011 | | | | | | | | | | | | 271 | 11,1 | 35,9 | 0,00 | 915 |

Neste anexo apresentam-se apenas as primeiras 21 linhas da base de dados elaborada para treinar as redes neurais. A base de dados completa poderá ser consultada no ficheiro Anexo I, assim como, as restantes bases de dados de previsão e validação.

ANEXO II

Tabela de Resultados do primeiro conjunto de redes

| Código RNA | Variáveis de entrada | Impacto variáveis | Nº dados | Esquema | Variáveis de saída | Resultados da RNA | | Observações |
|------------------------|------------------------|-------------------|----------|---------|--------------------|------------------------------------|------------|--|
| Teste RNA1 | 8 SL4 %MS | 29,26% | 242 | | 5 - Q biogás | Melhor rede | GRNN | RNA treinada por Hélder Rocha e Eng Milton |
| | 10 SL4 MV/MS | 20,46% | | | | %más previsões (treino-20% toler.) | 7,73% | |
| | 2 TRH | 23,60% | | | | % más previsões (teste-20% toler.) | 20,83% | |
| | 3 Temperatura digestor | 9,90% | | | | Erro médio quadrático (treino) | 114,71 | |
| | 4 CMV | 16,77% | | | | Erro médio quadrático (teste) | 226,75 | |
| Teste RNA2 | 2 TRH | 17,48% | 191 | | 5 - Q biogás | Melhor rede | MLFN 2 nós | RNA treinada por Hélder Rocha e Eng Milton |
| | 3 Temperatura digestor | 7,75% | | | | %más previsões (treino-20% toler.) | 19,61% | |
| | 4 CMV | 9,06% | | | | % más previsões (teste-20% toler.) | 34,21% | |
| | 8 SL4 %MS | 9,79% | | | | Erro médio quadrático (treino) | 173,84 | |
| | 10 SL4 MV/MS | 15,99% | | | | Erro médio quadrático (teste) | 280,01 | |
| | 11 SL5 %MS | 8,77% | | | | | | |
| | 13 SL5 MV/MS | 8,07% | | | | | | |
| | 15 SL5 pH | 3,04% | | | | | | |
| | 17 SL5 AGV | 9,58% | | | | | | |
| | 18 AGV/Alcalinade | 10,48% | | | | | | |
| Teste RNA3 | 9 SL4 %MV | 31,74% | 191 | | 5 - Q biogás | Melhor rede | MLFN 4 nós | RNA treinada por Hélder Rocha e Eng Milton |
| | 2 TRH (dias) | 19,31% | | | | %más previsões (treino-20% toler.) | 16,99% | |
| | 10 SL4 MV/MS | 14,59% | | | | % más previsões (teste-20% toler.) | 26,32% | |
| | 18 AGV/Alcal. | 14,29% | | | | Erro médio quadrático (treino) | 157,63 | |
| | 4 CMV | 12,36% | | | | Erro médio quadrático (teste) | 254,04 | |
| 3 Temperatura digestor | 7,71% | | | | | | | |
| Teste RNA4 | 2 TRH | 17,39% | 191 | | 5 - Q biogás | Melhor rede | MLFN 2 nós | RNA treinada por Hélder Rocha e Eng Milton |
| | 3 Temperatura digestor | 4,50% | | | | %más previsões (treino-15% toler.) | 31,17% | |
| | 4 CMV | 14,71% | | | | % más previsões (teste-15% toler.) | 41,03% | |
| | 8 SL4 %MS | 13,78% | | | | Erro médio quadrático (treino) | 192 | |
| | 10 SL4 MV/MS | 13,03% | | | | Erro médio quadrático (teste) | 271,57 | |
| | 11 SL5 %MS | 5,31% | | | | | | |
| | 13 SL5 MV/MS | 7,19% | | | | | | |
| | 15 SL5 pH | 6,03% | | | | | | |
| | 17 SL5 AGV | 9,35% | | | | | | |
| 18 AGV/Alcalinade | 8,72% | | | | | | | |
| Teste RNA5 | 2 TRH (dias) | 31,78% | 242 | | 5 - Q biogás | Melhor rede | GRNN | RNA treinada por Hélder Rocha e Eng Milton |
| | 10 SL4 MV/MS | 21,19% | | | | %más previsões (treino-15% toler.) | 4,64% | |
| | 4 CMV | 29,17% | | | | % más previsões (teste-15% toler.) | 33,33% | |
| | 3 Temperatura digestor | 17,87% | | | | Erro médio quadrático (treino) | 74,82 | |
| | | | | | | Erro médio quadrático (teste) | 233,77 | |

Neste anexo apresentam-se apenas os resultados das primeiras redes treinadas no primeiro conjunto de redes. A tabela completa poderá ser consultada no ficheiro Anexo II.

ANEXO III

Exemplo de um relatório gerado pelo Neural Tools

NeuralTools: Treinamento, Teste Automático e Previsão Automática de rede neural

Executado por: Geral

Data: sexta-feira, 15 de Março de 2013 16:55:07

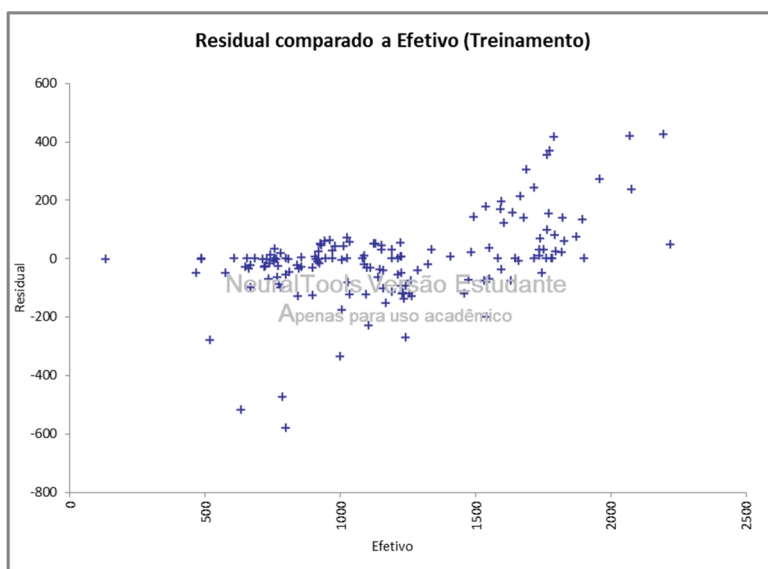
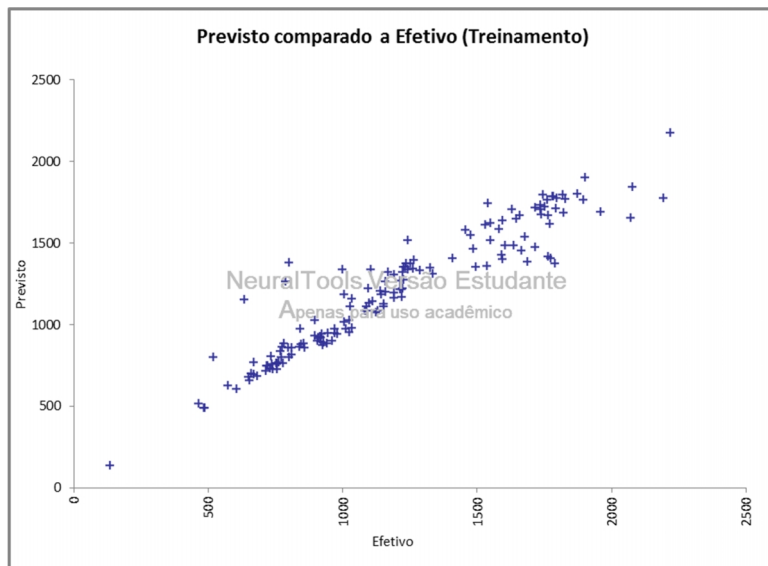
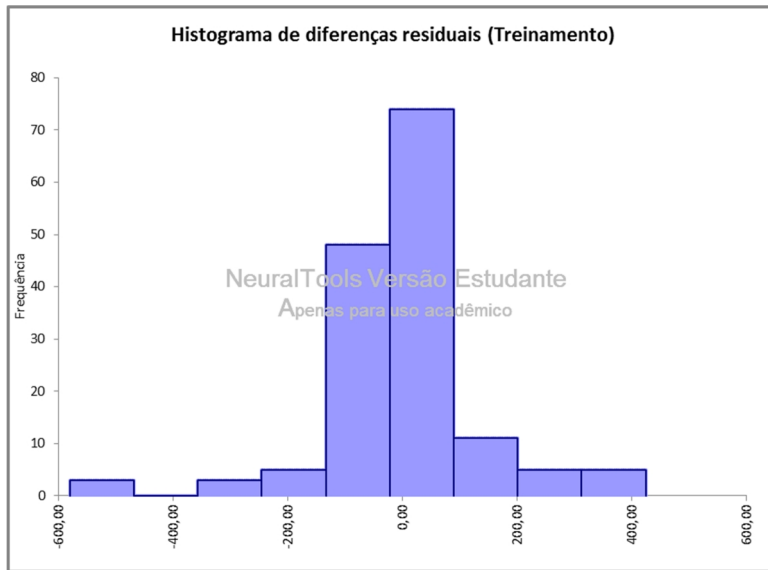
Conjunto de dados: Conjunto de dados núm. 10

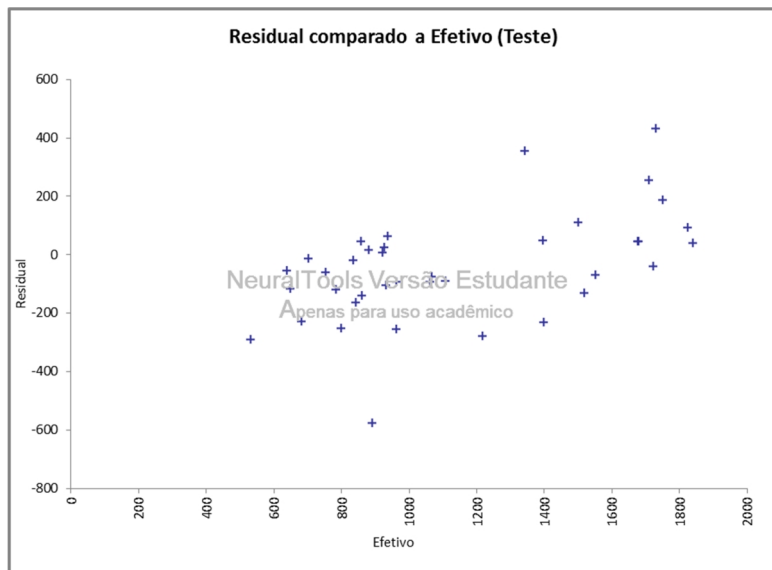
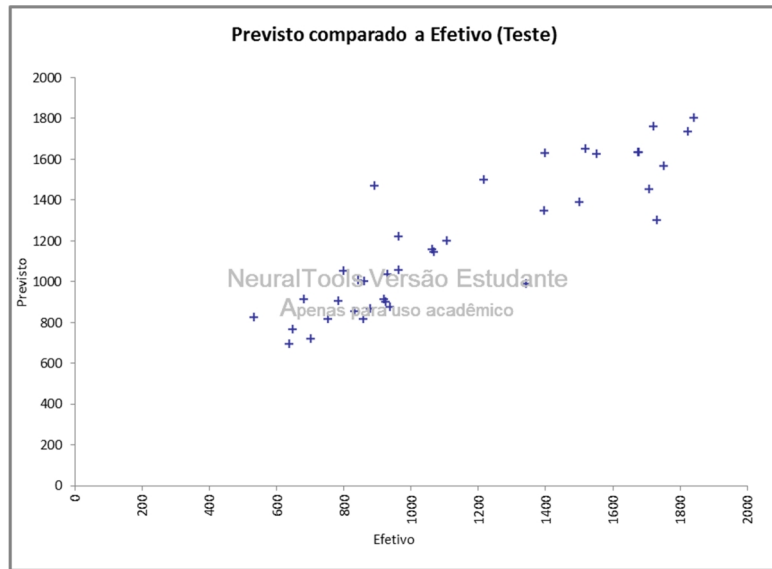
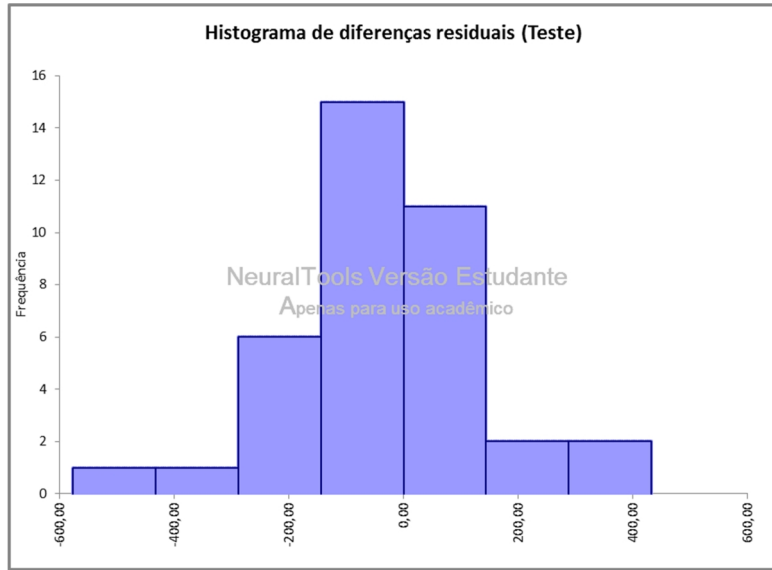
Rede: Rede treinada em Conjunto de dados núm. 10

| | |
|--|--|
| Resumo | |
| <i>Informações da rede</i> | |
| Nome | Rede treinada em Conjunto de dados núm. 10 |
| Configuração incluída na busca | GRNN, MLFN 2 a 6 nós |
| Melhor configuração | Previsor numérico de GRNN |
| Local | Esta pasta de trabalho |
| Variáveis categóricas independentes | 0 |
| Variáveis numéricas independentes | 5 (SL4 %MV, SL5A %MV, TRH (dias), Temp. (°C), CMV) |
| Variável dependente | Var. numérica (QBiogás (m3)) |
| <i>Treinamento</i> | |
| Número de casos | 154 |
| Tempo de treinamento | 00:00:00 |
| Número de tentativas | 83 |
| Motivo da parada | Parada automática |
| % previsões inadequadas (15% tolerância) | 9,7403% |
| Erro médio quadrático | 139,21 |
| Erro médio absoluto | 84,92 |
| Desvio padrão de erro absoluto | 110,31 |
| <i>Teste</i> | |
| Número de casos | 38 |
| % previsões inadequadas (15% tolerância) | 34,2105% |
| Erro médio quadrático | 186,79 |
| Erro médio absoluto | 138,77 |
| Desvio padrão de erro absoluto | 125,04 |
| <i>Previsão</i> | |
| Número de casos | 2 |
| Previsão Instantânea ativada | SIM |
| <i>Conjunto de dados</i> | |
| Nome | Conjunto de dados núm. 10 |
| Número de linhas | 684 |
| Tags de casos manuais | NÃO |
| <i>Análise de impacto de variáveis</i> | |
| Temp. (°C) | 32,7852% |
| SL4 %MV | 25,4566% |
| TRH (dias) | 23,0128% |
| SL5A %MV | 16,4466% |
| CMV | 2,2987% |

| Busca da Melhor Rede | | | |
|----------------------|-------------|----------------------|---------------------------------|
| | Erro de RMS | Tempo de treinamento | Motivo da parada do treinamento |
| Previsor linear | 224,72 | 00:00:00 | Parada automática |
| GRNN | 186,79 | 00:00:00 | Parada automática |
| MLFN 2 nós | 224,22 | 00:39:53 | Parada efetuada pelo usuário |
| MLFN 3 nós | 188,31 | 00:22:05 | Parada efetuada pelo usuário |
| MLFN 4 nós | 223,82 | 00:10:19 | Parada efetuada pelo usuário |
| MLFN 5 nós | 260,44 | 00:10:27 | Parada efetuada pelo usuário |
| MLFN 6 nós | 231,85 | 00:17:36 | Parada efetuada pelo usuário |

| Previsor linear comparado a rede neural | | |
|---|-----------------|-------------|
| | Previsor linear | Rede neural |
| R-quadrada (Treinamento) | 0,5347 | - |
| Erro médio quadrático (Treinamento) | 290,04 | 139,21 |
| Erro médio quadrático (Teste) | 224,72 | 186,79 |





ANEXO IV

Tabela de Resultados do segundo conjunto de redes

| Código RNA | Variáveis de entrada | Impacto variáveis | Nº dados | Esquema | Variáveis de saída | Resultados da RNA | | Observações |
|-------------|----------------------|---------------------------|----------|---------|--------------------|------------------------------------|------------|---|
| Treino RNA1 | 3 | Temperatura digestor (°C) | 29,19% | | 5 - Q biogás | Melhor rede | GRNN | Nesta rede foram excluídos os dados com eficiência de digestão negativa. (<0) |
| | 9 | SL4 %MV | 26,78% | | | %más previsões (treino-15% toler.) | 4,76% | |
| | 2 | TRH (dias) | 24,64% | | | % más previsões (teste-15% toler.) | 27,03% | |
| | 12 | SL5 %MV | 16,51% | | | Erro médio quadrático (treino) | 96,77 | |
| | 4 | CMV | 2,88% | | | Erro médio quadrático (teste) | 192,84 | |
| Treino RNA2 | 4 | CMV | 33,55% | | 5 - Q biogás | Melhor rede | MLFN 2 nós | Foi considerada a média móvel (7 dias) do TRH. |
| | 9 | SL4 %MV | 32,88% | | | %más previsões (treino-15% toler.) | 34,42% | |
| | 3 | Temperatura digestor (°C) | 13,86% | | | % más previsões (teste-15% toler.) | 47,37% | |
| | 12 | SL5 %MV | 11,32% | | | Erro médio quadrático (treino) | 214,48 | |
| | 24 | Média móvel TRH 7 dias | 8,39% | | | Erro médio quadrático (teste) | 269,98 | |
| Treino RNA3 | 9 | SL4 %MV | 37,78% | | 5 - Q biogás | Melhor rede | GRNN | Nesta rede foram excluídos os dados com eficiência de digestão < ou = a 0. |
| | 2 | TRH (dias) | 26,94% | | | %más previsões (treino-15% toler.) | 18,75% | |
| | 3 | Temperatura digestor (°C) | 22,23% | | | % más previsões (teste-15% toler.) | 50,00% | |
| | 12 | SL5 %MV | 13,04% | | | Erro médio quadrático (treino) | 151,97 | |
| | | | | | | Erro médio quadrático (teste) | 340,88 | |

(...)

| | | | | | | | | |
|--------------|----|---------------------------|--------|--|--------------|------------------------------------|--------------|--|
| Treino RNA10 | 9 | SL4 %MV | 32,20% | | 5 - Q biogás | Melhor rede | GRNN | Nesta rede foram excluídos os dados com eficiência de digestão < ou = a 0. |
| | 2 | TRH (dias) | 29,64% | | | %más previsões (treino-15% toler.) | 13,29% | |
| | 26 | Média móvel temp. 18 dias | 22,16% | | | % más previsões (teste-15% toler.) | 36,11% | |
| | 12 | SL5 %MV | 15,41% | | | Erro médio quadrático (treino) | 156,39 | |
| | 4 | CMV | 0,59% | | | Erro médio quadrático (teste) | 219,89 | |
| Treino RNA11 | 3 | Temperatura digestor (°C) | 45,04% | | 5 - Q biogás | Melhor rede | MLFN (6 nós) | Nesta rede foram excluídos os dados com eficiência de digestão < ou = a 0. |
| | 9 | SL4 %MV | 32,65% | | | %más previsões (treino-15% toler.) | 30,41% | |
| | 1 | Q entrada digestor (m3) | 22,31% | | | % más previsões (teste-15% toler.) | 41,67% | |
| | | | | | | Erro médio quadrático (treino) | 177,10 | |
| | | | | | | Erro médio quadrático (teste) | 215,19 | |
| Treino RNA12 | 4 | CMV | 33,40% | | 5 - Q biogás | Melhor rede | MLFN (3 nós) | Nesta rede foram excluídos os dados com eficiência de digestão < ou = a 0. |
| | 9 | SL4 %MV | 22,53% | | | %más previsões (treino-15% toler.) | 30,52% | |
| | 12 | SL5 %MV | 17,57% | | | % más previsões (teste-15% toler.) | 34,21% | |
| | 1 | Q entrada digestor (m3) | 14,97% | | | Erro médio quadrático (treino) | 199,24 | |
| | 3 | Temperatura digestor (°C) | 11,53% | | | Erro médio quadrático (teste) | 177,27 | |

Neste anexo apresentam-se alguns resultados das redes treinadas no segundo conjunto de redes. A tabela completa poderá ser consultada no ficheiro Anexo IV.

ANEXO V

Tabela de Resultados do terceiro conjunto de redes

| Código RNA | Variáveis de entrada | Impacto variáveis | Nº dados | Esquema | Variáveis de saída | Resultados da RNA | | Observações |
|------------|-----------------------------|-------------------|----------|---------|--------------------|-------------------------------------|--------|--|
| TrTs RNA1 | 1 Q entrada digestor (m3) | 24,91% | 180 | | 5 - Q biogás | Melhor rede | GRNN | Nesta rede foram excluídos os dados com eficiência de digestão < ou = a 0. |
| | 4 CMV | 24,90% | | | | % más previsões (treino-15% toler.) | 8,33% | |
| | 9 SL4 %MV | 17,84% | | | | % más previsões (teste-15% toler.) | 36,11% | |
| | 3 Temperatura digestor (°C) | 17,22% | | | | Erro médio quadrático (treino) | 126,07 | |
| | 12 SL5 %MV | 15,13% | | | | Erro médio quadrático (teste) | 262,84 | |
| TrTs RNA2 | 3 Temperatura digestor (°C) | 33,97% | 180 | | 5 - Q biogás | Melhor rede | GRNN | Nesta rede foram excluídos os dados com eficiência de digestão < ou = a 0. |
| | 9 SL4 %MV | 26,56% | | | | % más previsões (treino-15% toler.) | 12,50% | |
| | 2 TRH (dias) | 22,09% | | | | % más previsões (teste-15% toler.) | 41,67% | |
| | 4 CMV | 8,85% | | | | Erro médio quadrático (treino) | 141,01 | |
| | 12 SL5 %MV | 8,53% | | | | Erro médio quadrático (teste) | 246,46 | |
| TrTs RNA3 | 3 Temperatura digestor (°C) | 28,61% | 242 | | 5 - Q biogás | Melhor rede | GRNN | |
| | 9 SL4 %MV | 26,94% | | | | % más previsões (treino-15% toler.) | 17,01% | |
| | 4 CMV | 22,39% | | | | % más previsões (teste-15% toler.) | 27,08% | |
| | 1 Q entrada digestor (m3) | 22,06% | | | | Erro médio quadrático (treino) | 143,19 | |
| | | | | | | Erro médio quadrático (teste) | 179,77 | |
| TrTs RNA4 | 3 Temperatura digestor (°C) | 33,69% | 242 | | 5 - Q biogás | Melhor rede | GRNN | |
| | 1 Q entrada digestor (m3) | 33,69% | | | | % más previsões (treino-15% toler.) | 16,49% | |
| | 9 SL4 %MV | 32,62% | | | | % más previsões (teste-15% toler.) | 33,33% | |
| | | | | | | Erro médio quadrático (treino) | 139,23 | |
| | | | | | | Erro médio quadrático (teste) | 182,71 | |
| TrTs RNA5 | 9 SL4 %MV | 32,05% | 242 | | 5 - Q biogás | Melhor rede | GRNN | |
| | 3 Temperatura digestor (°C) | 29,04% | | | | % más previsões (treino-15% toler.) | 10,31% | |
| | 1 Q entrada digestor (m3) | 28,10% | | | | % más previsões (teste-15% toler.) | 33,33% | |
| | 8 SL4 %MS | 10,82% | | | | Erro médio quadrático (treino) | 116,17 | |
| | | | | | | Erro médio quadrático (teste) | 213,81 | |

(...)

| | | | | | | | | |
|------------|-----------------------------|--------|-----|--|--------------|-------------------------------------|--------|--|
| TrTs RNA9 | 9 SL4 %MV | 36,11% | 242 | | 5 - Q biogás | Melhor rede | GRNN | |
| | 3 Temperatura digestor (°C) | 33,17% | | | | % más previsões (treino-15% toler.) | 12,37% | |
| | 2 TRH (dias) | 30,72% | | | | % más previsões (teste-15% toler.) | 31,25% | |
| | | | | | | Erro médio quadrático (treino) | 141,72 | |
| | | | | | | Erro médio quadrático (teste) | 198,34 | |
| TrTs RNA10 | 3 Temperatura digestor (°C) | 35,94% | 243 | | 5 - Q biogás | Melhor rede | GRNN | |
| | 4 CMV | 32,52% | | | | % más previsões (treino-15% toler.) | 15,98% | |
| | 2 TRH (dias) | 31,54% | | | | % más previsões (teste-15% toler.) | 28,57% | |
| | | | | | | Erro médio quadrático (treino) | 139,64 | |
| | | | | | | Erro médio quadrático (teste) | 208,80 | |

Neste anexo apresentam-se alguns resultados das redes treinadas no terceiro conjunto de redes. A tabela completa poderá ser consultada no ficheiro Anexo V.

ANEXO VI

Tabelas de resultados das 3 melhores redes

| Código RNA | Variáveis de entrada | Impacto variáveis | Nº dados | Esquema | Variáveis de saída | Resultados da RNA | | |
|---------------|----------------------|---------------------------|----------|---------|--------------------|-------------------|------------------------------------|--------|
| TrTs RNA3 | 3 | Temperatura digestor (°C) | 28,61% | 242 | | 5 - Q biogás | Melhor rede | GRNN |
| | 9 | SL4 %MV | 26,94% | | | | %más previsões (treino-15% toler.) | 17,01% |
| | 4 | CMV | 22,39% | | | | % más previsões (teste-15% toler.) | 27,08% |
| | 1 | Q entrada digestor (m3) | 22,06% | | | | Erro médio quadrático (treino) | 143,19 |
| | | | | | | | Erro médio quadrático (teste) | 179,77 |
| BNTrTs RNA3.1 | 3 | Temperatura digestor (°C) | 32,82% | 242 | | 5 - Q biogás | Rede | GRNN |
| | 4 | CMV | 25,37% | | | | %más previsões (treino-15% toler.) | 15,46% |
| | 9 | SL4 %MV | 21,19% | | | | % más previsões (teste-15% toler.) | 35,42% |
| | 1 | Q entrada digestor (m3) | 20,62% | | | | Erro médio quadrático (treino) | 130,86 |
| | | | | | | | Erro médio quadrático (teste) | 258,69 |
| BNTrTs RNA3.2 | 4 | CMV | 30,68% | 242 | | 5 - Q biogás | Rede | GRNN |
| | 1 | Q entrada digestor (m3) | 25,30% | | | | %más previsões (treino-15% toler.) | 17,53% |
| | 3 | Temperatura digestor (°C) | 24,16% | | | | % más previsões (teste-15% toler.) | 35,42% |
| | 9 | SL4 %MV | 19,87% | | | | Erro médio quadrático (treino) | 137,58 |
| | | | | | | | Erro médio quadrático (teste) | 199,66 |

(...)

| | | | | | | | | |
|----------------|---|---------------------------|--------|-----|--|--------------|------------------------------------|--------|
| BNTrTs RNA3.16 | 3 | Temperatura digestor (°C) | 38,30% | 242 | | 5 - Q biogás | Rede | GRNN |
| | 1 | Q entrada digestor (m3) | 29,85% | | | | %más previsões (treino-20% toler.) | 14,43% |
| | 9 | SL4 %MV | 26,83% | | | | % más previsões (teste-20% toler.) | 12,50% |
| | 4 | CMV | 5,02% | | | | Erro médio quadrático (treino) | 157,73 |
| | | | | | | | Erro médio quadrático (teste) | 139,66 |
| BNTrTs RNA3.17 | 3 | Temperatura digestor (°C) | 38,10% | 242 | | 5 - Q biogás | Rede | GRNN |
| | 9 | SL4 %MV | 30,82% | | | | %más previsões (treino-20% toler.) | 9,28% |
| | 1 | Q entrada digestor (m3) | 28,60% | | | | % más previsões (teste-20% toler.) | 18,75% |
| | 4 | CMV | 2,49% | | | | Erro médio quadrático (treino) | 134,99 |
| | | | | | | | Erro médio quadrático (teste) | 172,98 |

Neste anexo apresentam-se apenas alguns resultados de uma das 3 melhores redes encontradas, a “BNTrTs RNA3”. As tabelas completas das 3 melhores redes poderão ser consultadas no ficheiro Anexo VI.

| # | Data | SL4 %MV | Q _{ent} (m ³) | Temp. (°C) | CMV | Q _{total} (m ³) | Tag usada | Previsão | Adequada/Inadequada | Residual |
|----|-------------|------------|---------------------------------------|---------------|------|---|-----------|----------|---------------------|----------|
| 1 | 17-Abr-2011 | | 97 | 34,4 | | | | | | |
| 2 | 18-Abr-2011 | 1,7 | 109 | 33,1 | 0,62 | 656 | prever | #NOME? | | |
| 3 | 19-Abr-2011 | 1,5 | 138 | 33,6 | 0,69 | 688 | prever | #NOME? | | |
| 4 | 20-Abr-2011 | | 203 | 33,3 | | 1547 | | | | |
| 5 | 21-Abr-2011 | | 130 | 32,8 | | 2683 | | | | |
| 6 | 22-Abr-2011 | | 167 | 33,3 | | 2584 | | | | |
| 7 | 23-Abr-2011 | | 168 | 33,9 | | 2819 | | | | |
| 8 | 24-Abr-2011 | | 168 | 34,6 | | 2755 | | | | |
| 9 | 25-Abr-2011 | | 169 | 35,1 | | 2983 | | | | |
| 10 | 26-Abr-2011 | 1,8 | 152 | 35,5 | 0,91 | 1287 | testar | 1347 | Adequada | -60 |
| 11 | 27-Abr-2011 | | 169 | 35,5 | | 1177 | | | | |
| 12 | 28-Abr-2011 | 1,1 | 167 | 35,2 | 0,61 | 1006 | testar | 1070 | Adequada | -64 |
| 13 | 29-Abr-2011 | | 219 | 35,3 | | 1103 | | | | |
| 14 | 30-Abr-2011 | | 322 | 30,9 | | 1067 | | | | |
| 15 | 1-Mai-2011 | | 353 | 35,6 | | 1172 | | | | |
| 16 | 2-Mai-2011 | | 196 | 36,0 | | 1463 | | | | |
| 17 | 3-Mai-2011 | 2,0 | 200 | 36,3 | 1,33 | 1537 | treinar | | | |
| 18 | 4-Mai-2011 | | 258 | 36,6 | | 1292 | | | | |
| 19 | 5-Mai-2011 | 2,1 | 244 | 35,7 | 1,71 | 1153 | testar | 867 | Inadequada | 286 |
| 20 | 6-Mai-2011 | | 249 | 35,5 | | 1011 | | | | |
| 21 | 7-Mai-2011 | | 271 | 35,9 | | 915 | | | | |
| 22 | 8-Mai-2011 | | 217 | 35,5 | | 923 | | | | |
| 23 | 9-Mai-2011 | | 236 | 36,0 | | 912 | | | | |
| 24 | 10-Mai-2011 | 1,8 | 287 | 35,9 | 1,72 | 768 | treinar | | | |
| 25 | 11-Mai-2011 | | 262 | 35,9 | | 729 | | | | |
| 26 | 12-Mai-2011 | | 137 | 35,4 | | 805 | | | | |
| 27 | 13-Mai-2011 | 1,7 | 250 | 36,0 | 1,42 | 844 | treinar | | | |
| 28 | 14-Mai-2011 | | 327 | 36,1 | | 845 | | | | |
| 29 | 15-Mai-2011 | | 297 | 35,5 | | 946 | | | | |
| 30 | 16-Mai-2011 | | 275 | 35,5 | | 910 | | | | |
| 31 | 17-Mai-2011 | 1,8 | 220 | 35,2 | 1,32 | 897 | treinar | | | |
| 32 | 18-Mai-2011 | 1,7 | 173 | 35,6 | 0,98 | 1034 | testar | 1314 | Inadequada | -280 |
| 33 | 19-Mai-2011 | | 168 | 35,7 | | 994 | | | | |
| 34 | 20-Mai-2011 | | 182 | 36,0 | | 1125 | | | | |
| 35 | 21-Mai-2011 | | 192 | 35,7 | | 1151 | | | | |
| 36 | 22-Mai-2011 | | 191 | 35,6 | | 1219 | | | | |
| 37 | 23-Mai-2011 | 0,7 | 135 | 36,5 | 0,32 | 1083 | treinar | | | |
| 38 | 24-Mai-2011 | | 191 | 33,1 | | 1247 | | | | |
| 39 | 25-Mai-2011 | 2,1 | 193 | 36,1 | 1,35 | 1262 | testar | 1428 | Adequada | -166 |
| 40 | 26-Mai-2011 | | 190 | 36,0 | | 1087 | | | | |
| 41 | 27-Mai-2011 | | 245 | 35,9 | | 939 | | | | |
| 42 | 28-Mai-2011 | | 191 | 35,5 | | 1210 | | | | |
| 43 | 29-Mai-2011 | | 194 | 35,7 | | 1247 | | | | |
| 44 | 30-Mai-2011 | 1,9 | 180 | 36,0 | 1,14 | 1337 | treinar | | | |
| 45 | 31-Mai-2011 | | 193 | 33,6 | | 1272 | | | | |
| 46 | 1-Jun-2011 | 2,2 | 140 | 35,6 | 1,03 | 1001 | treinar | | | |
| 47 | 2-Jun-2011 | | 187 | 35,4 | | 1043 | | | | |
| 48 | 3-Jun-2011 | | 197 | 35,8 | | 1224 | | | | |
| 49 | 4-Jun-2011 | | 147 | 35,7 | | 924 | | | | |
| 50 | 5-Jun-2011 | | 237 | 35,6 | | 1495 | | | | |
| 51 | 6-Jun-2011 | 1,1 | 187 | 35,7 | 0,69 | 1131 | treinar | | | |

(...)

Ficheiro Excel usado no treino da melhor rede (BNTrTs RNA3.16)

Relatório de treino e teste gerado pela melhor rede encontrada (BNTrTs RNA3.16)

NeuralTools: Treinamento, Teste Automático e Previsão Automática de rede neural

Executado por: Geral

Data: quinta-feira, 4 de Abril de 2013 10:52:26

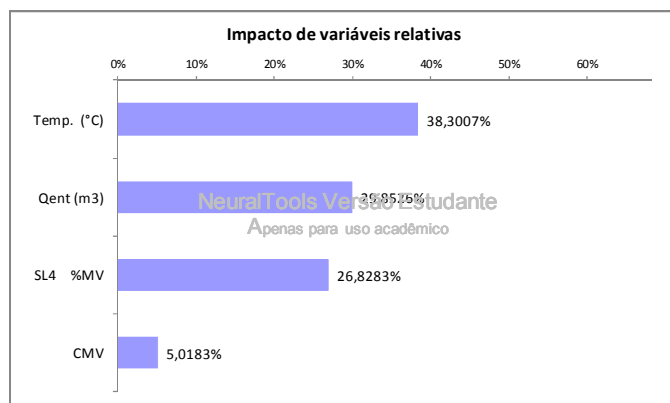
Conjunto de dados: Conjunto de dados núm. 16

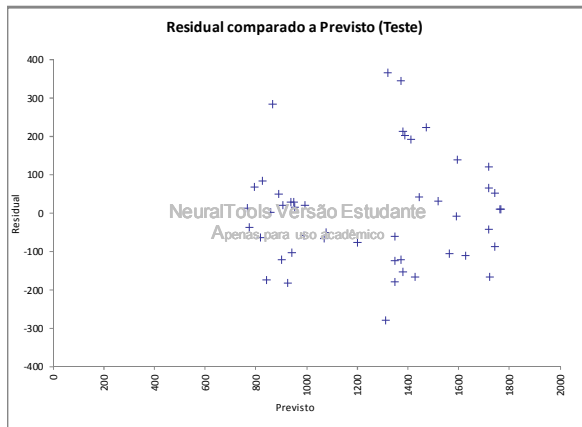
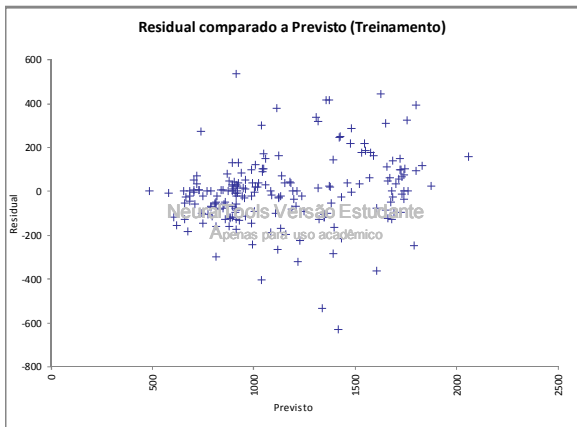
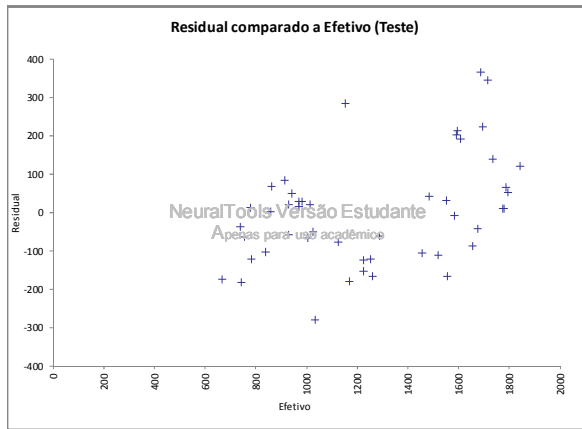
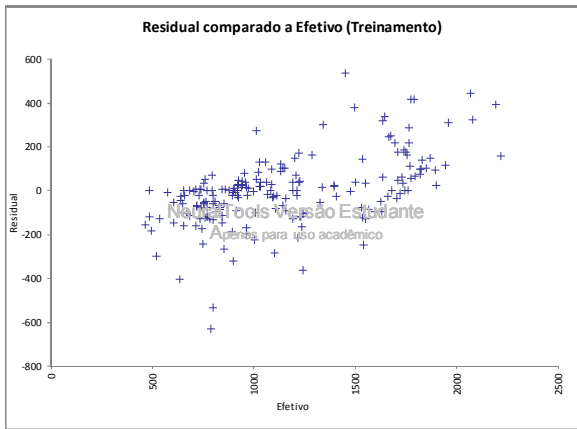
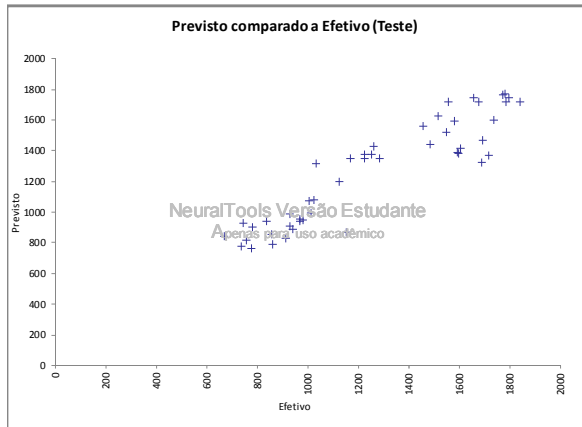
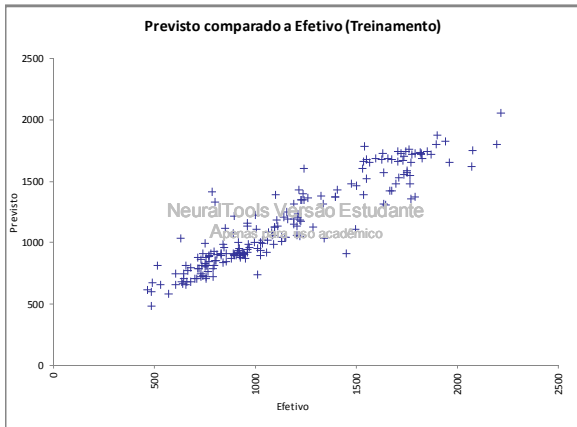
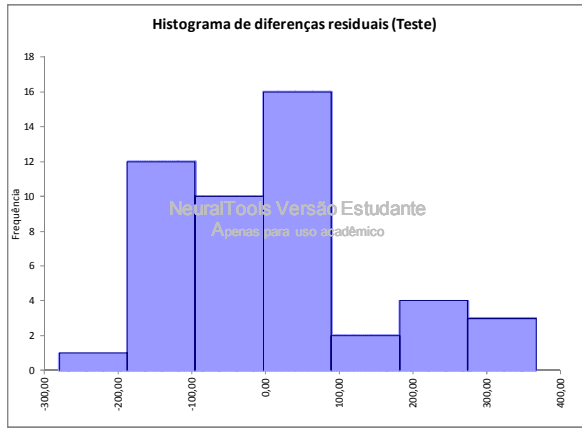
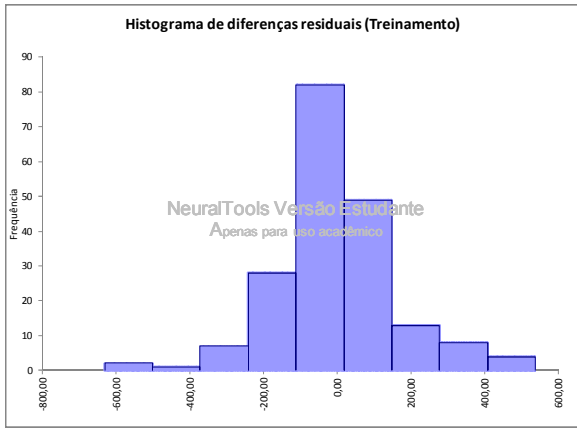
Rede: Rede treinada em Conjunto de dados núm. 16

| Resumo | |
|--|--|
| Informações da rede | |
| Nome | Rede treinada em Conjunto de dados núm. 16 |
| Configuração | Previsor numérico de GRNN |
| Local | Esta pasta de trabalho |
| Variáveis categóricas independentes | 0 |
| Variáveis numéricas independentes | 4 (SL4 %MV, Qent (m3), Temp. (°C), CMV) |
| Variável dependente | Var. numérica (QBiogás (m3)) |
| Treinamento | |
| Número de casos | 194 |
| Tempo de treinamento | 00:00:00 |
| Número de tentativas | 69 |
| Motivo da parada | Parada automática |
| % previsões inadequadas (20% tolerância) | 14,4330% |
| Erro médio quadrático | 157,73 |
| Erro médio absoluto | 108,92 |
| Desvio padrão de erro absoluto | 114,09 |
| Teste | |
| Número de casos | 48 |
| % previsões inadequadas (20% tolerância) | 12,5000% |
| Erro médio quadrático | 139,66 |
| Erro médio absoluto | 107,84 |
| Desvio padrão de erro absoluto | 88,75 |
| Previsão | |
| Número de casos | 2 |
| Previsão Instantânea ativada | SIM |
| Conjunto de dados | |
| Nome | Conjunto de dados núm. 16 |
| Número de linhas | 684 |
| Tags de casos manuais | NÃO |
| Análise de impacto de variáveis | |
| Temp. (°C) | 38,3007% |
| Qent (m3) | 29,8526% |
| SL4 %MV | 26,8283% |
| CMV | 5,0183% |

Previsor linear comparado a rede neural

| | Previsor linear | Rede neural |
|-------------------------------------|-----------------|-------------|
| R-quadrada (Treinamento) | 0,5310 | -- |
| Erro médio quadrático (Treinamento) | 286,38 | 157,73 |
| Erro médio quadrático (Teste) | 213,14 | 139,66 |

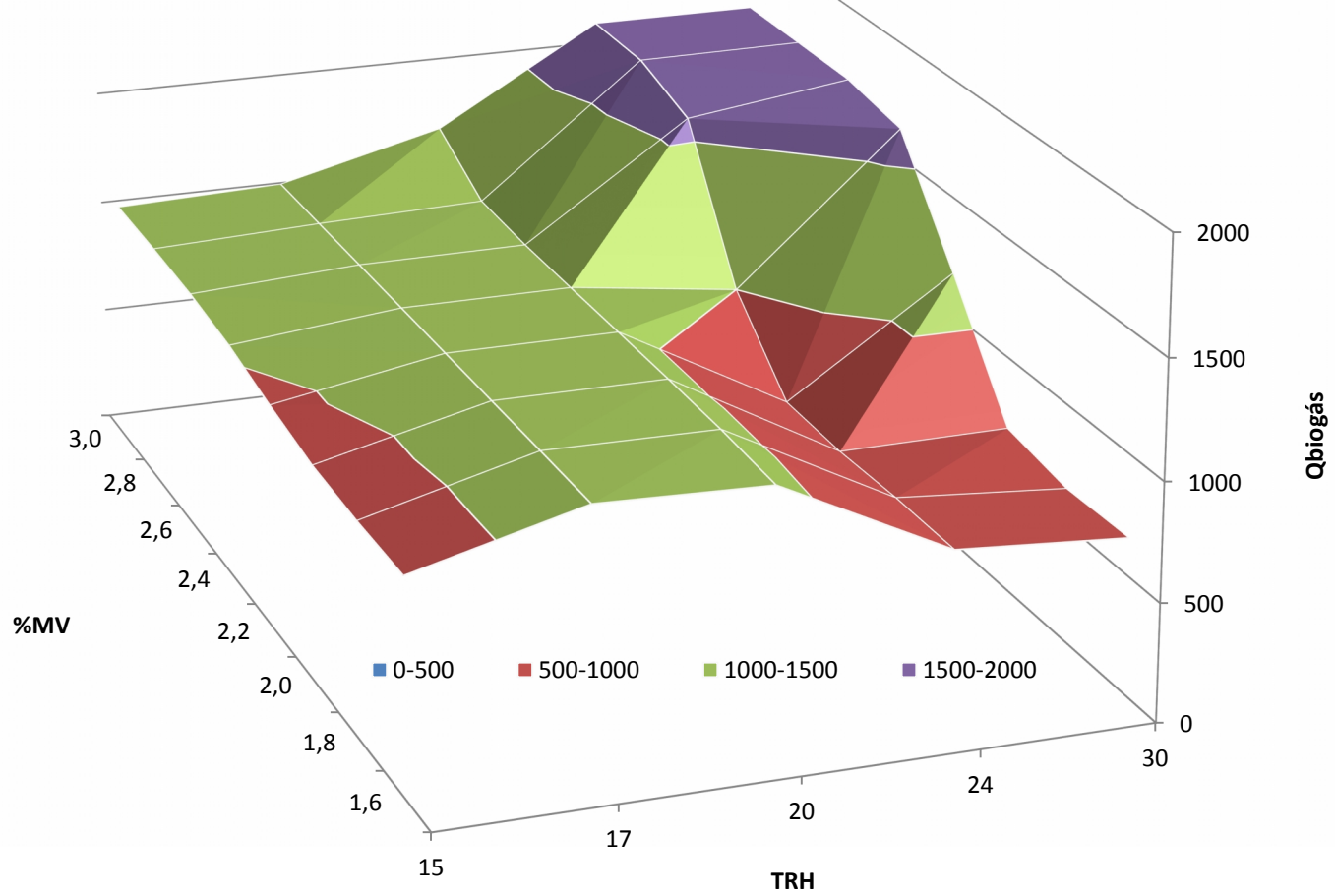




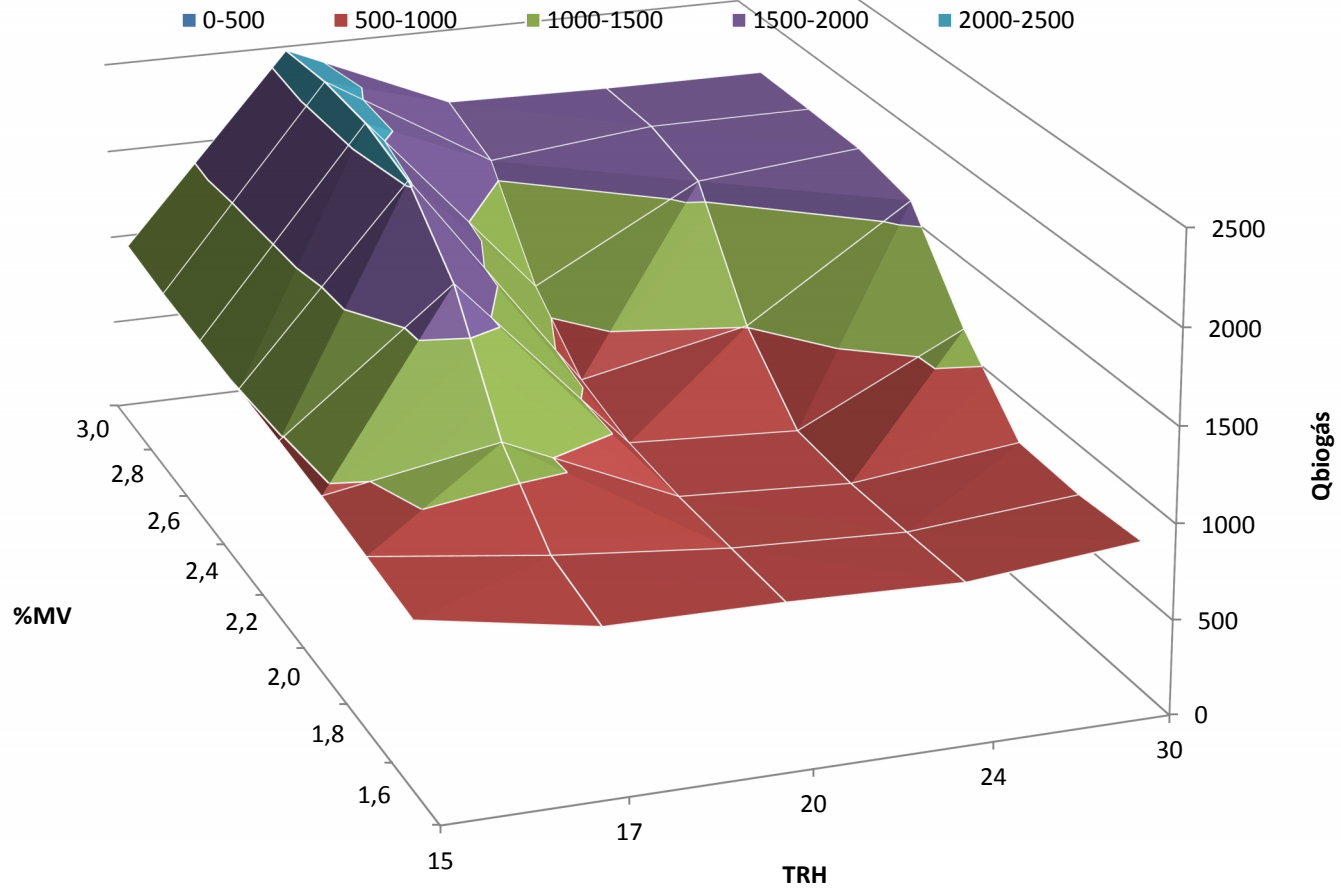
ANEXO VII

Gráficos de interpretação dos dados de previsão

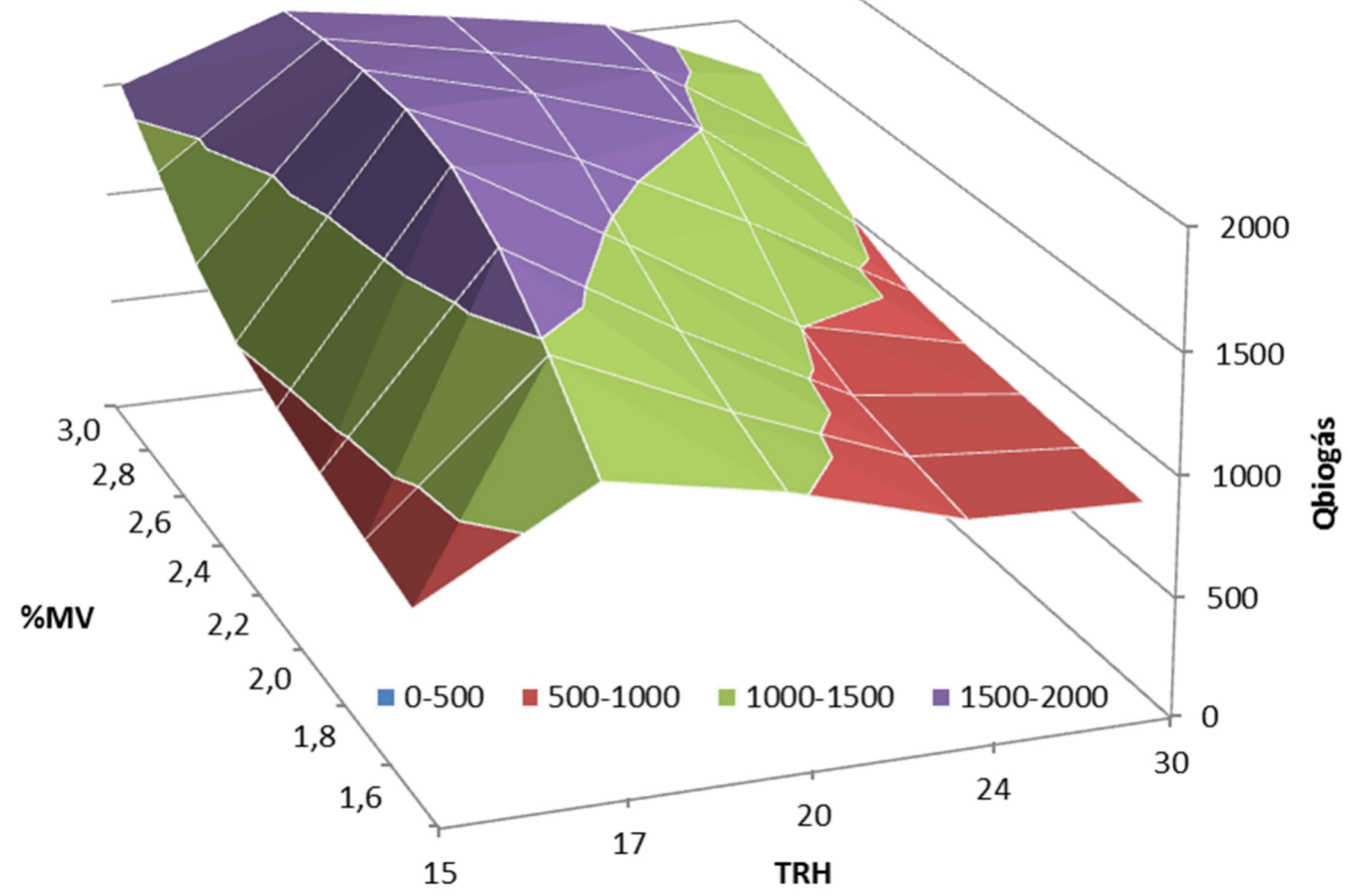
Temperatura 33 oC

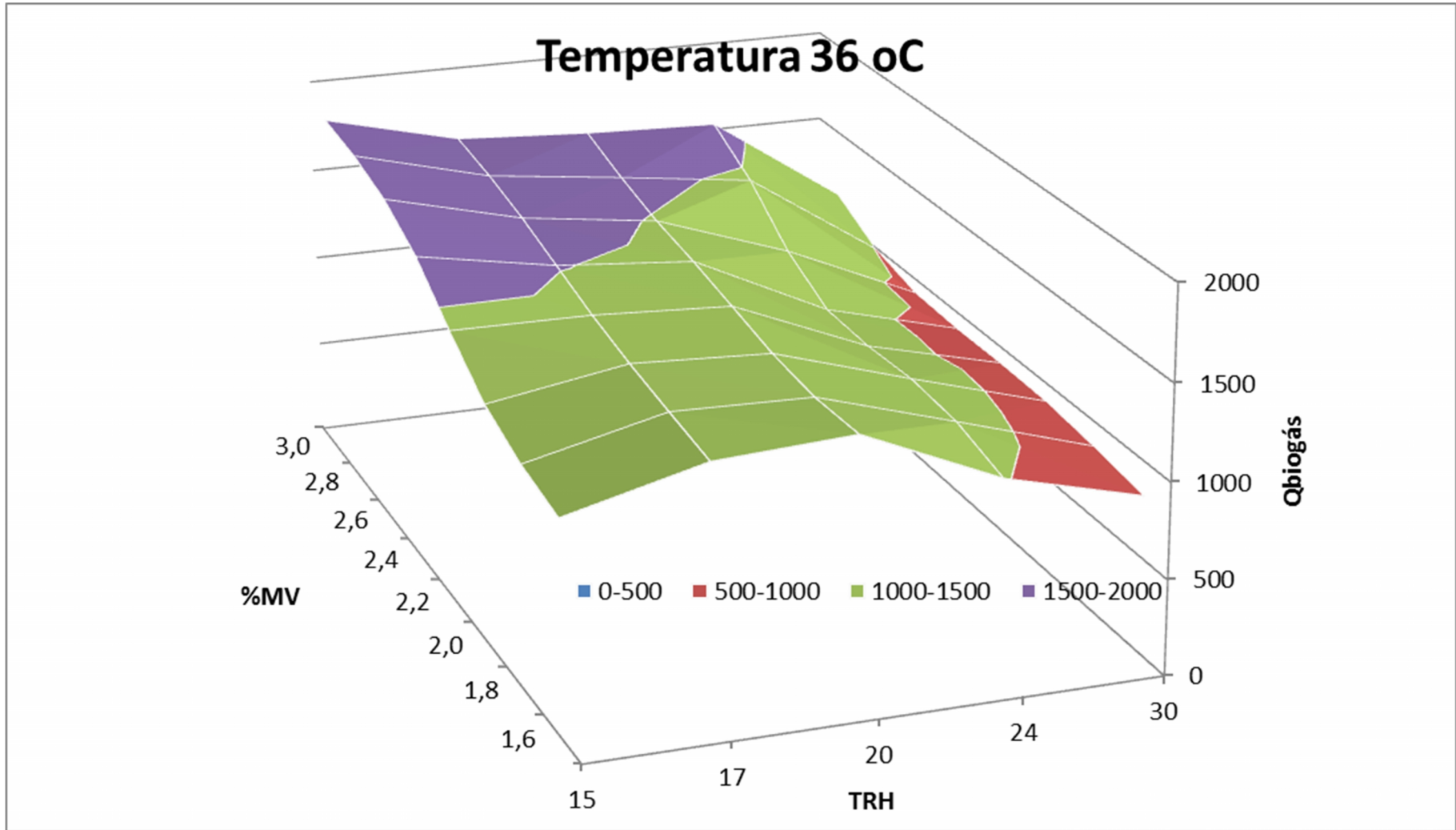


Temperatura 34 oC



Temperatura 35 oC

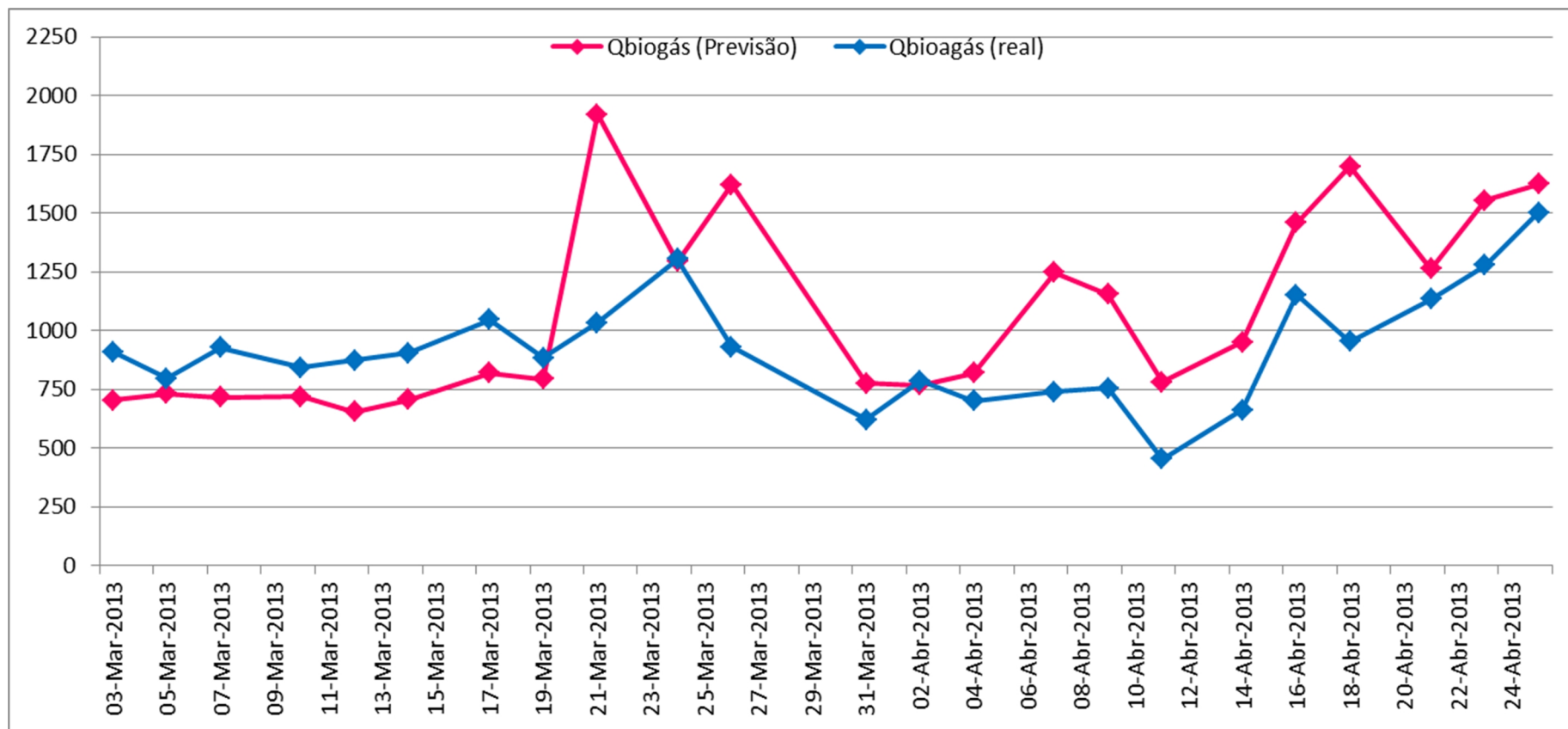




Para além dos gráficos com os resultados das previsões da melhor rede, aqui apresentados, podem ainda ser consultados mais gráficos e dados referentes a essas mesmas previsões no ficheiro Anexo VII.

ANEXO VIII

Gráficos dos dados de validação



Neste anexo apresenta-se apenas o gráfico com os resultados de validação da melhor rede treinada. Os gráficos com os dados de validação, das restantes redes, poderão ser consultados no ficheiro Anexo VIII.

ANEXO IX

Gráfico das previsões com médias mensais

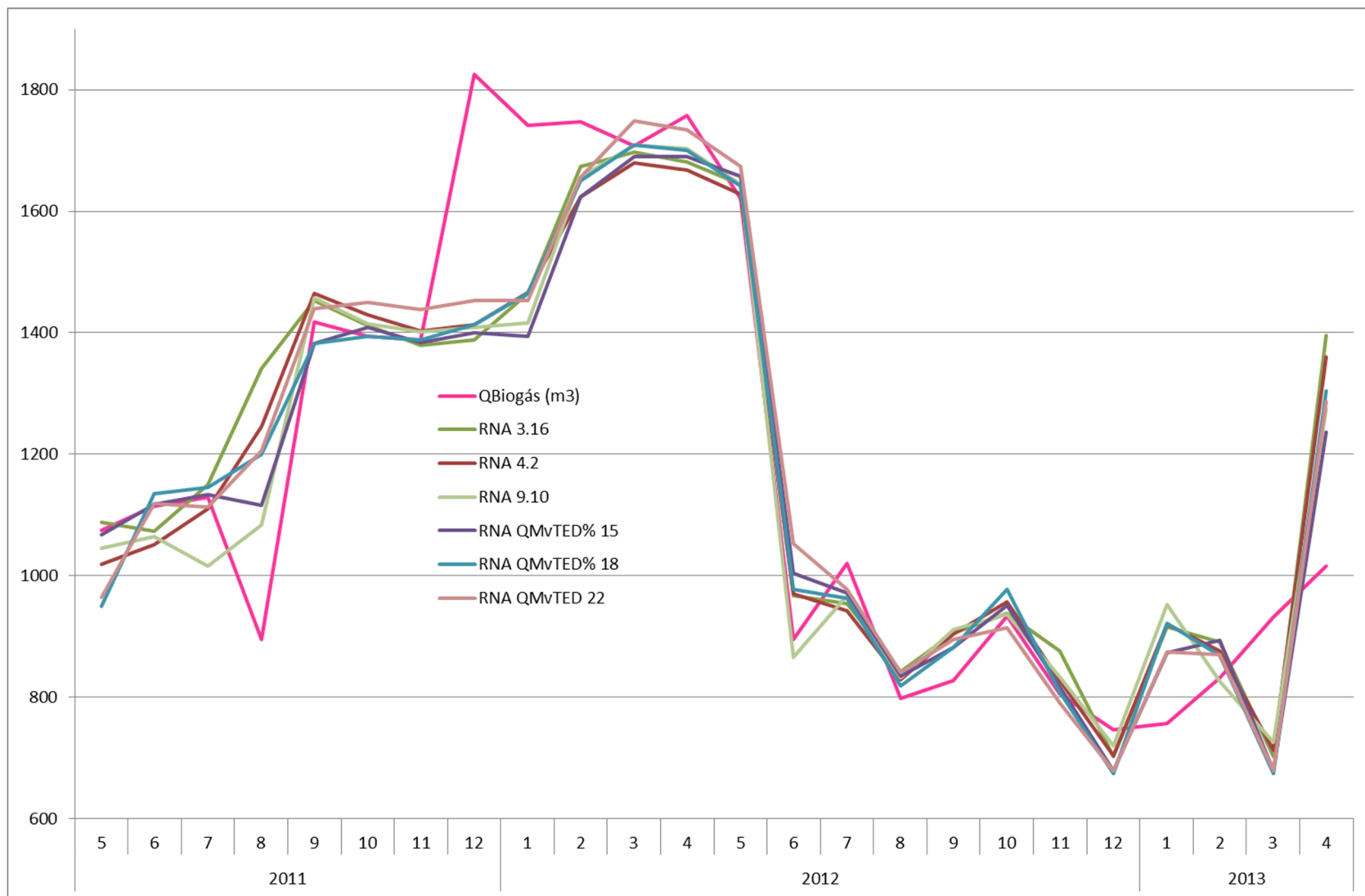


Gráfico com os resultados de previsões de várias redes, com base nas médias mensais das variáveis.

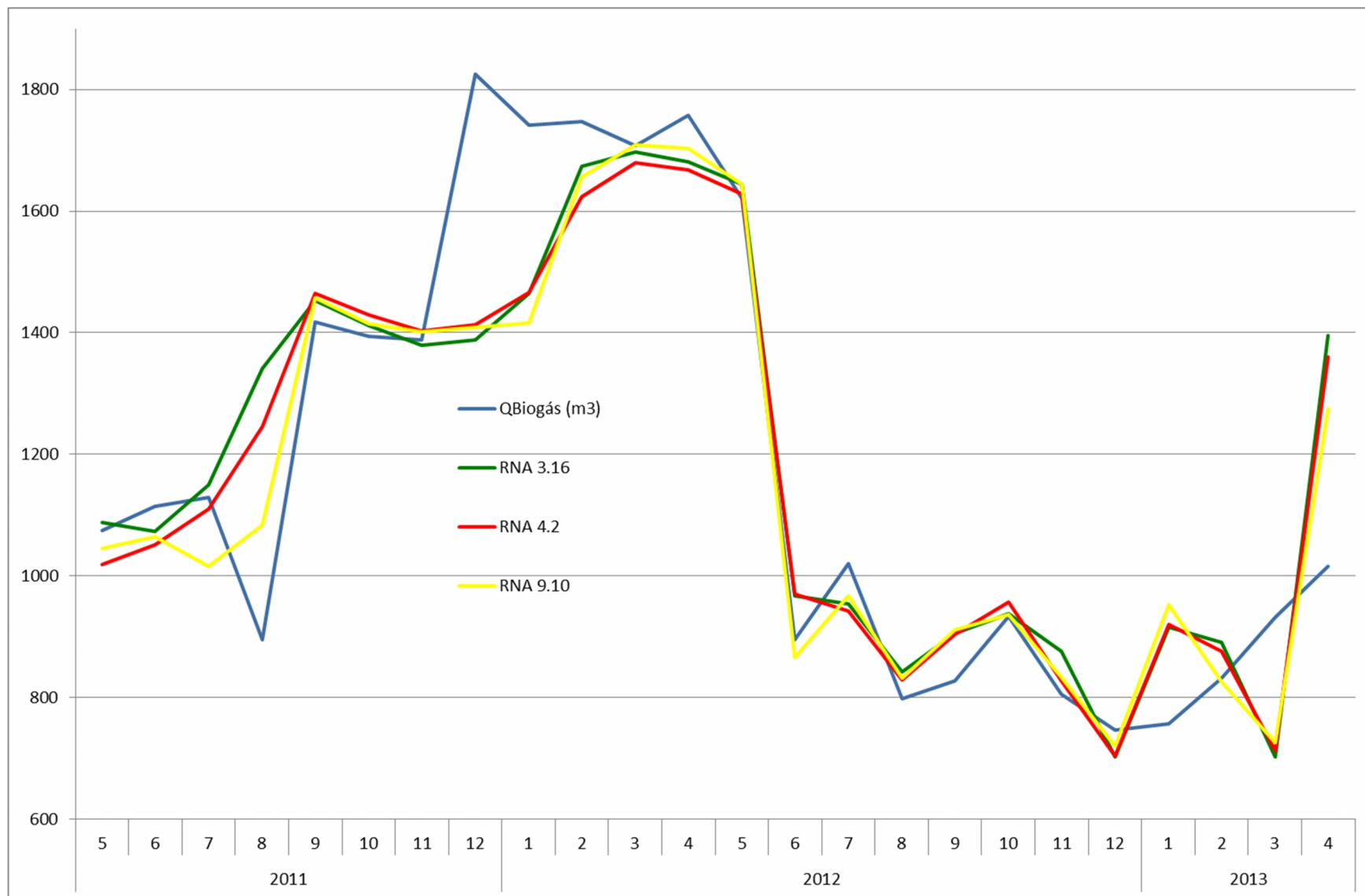


Gráfico com os resultados de previsões das 3 melhores redes, com base nas médias mensais das variáveis.

ANEXO X

Gráfico das previsões com médias mensais para a
ETAR Norte

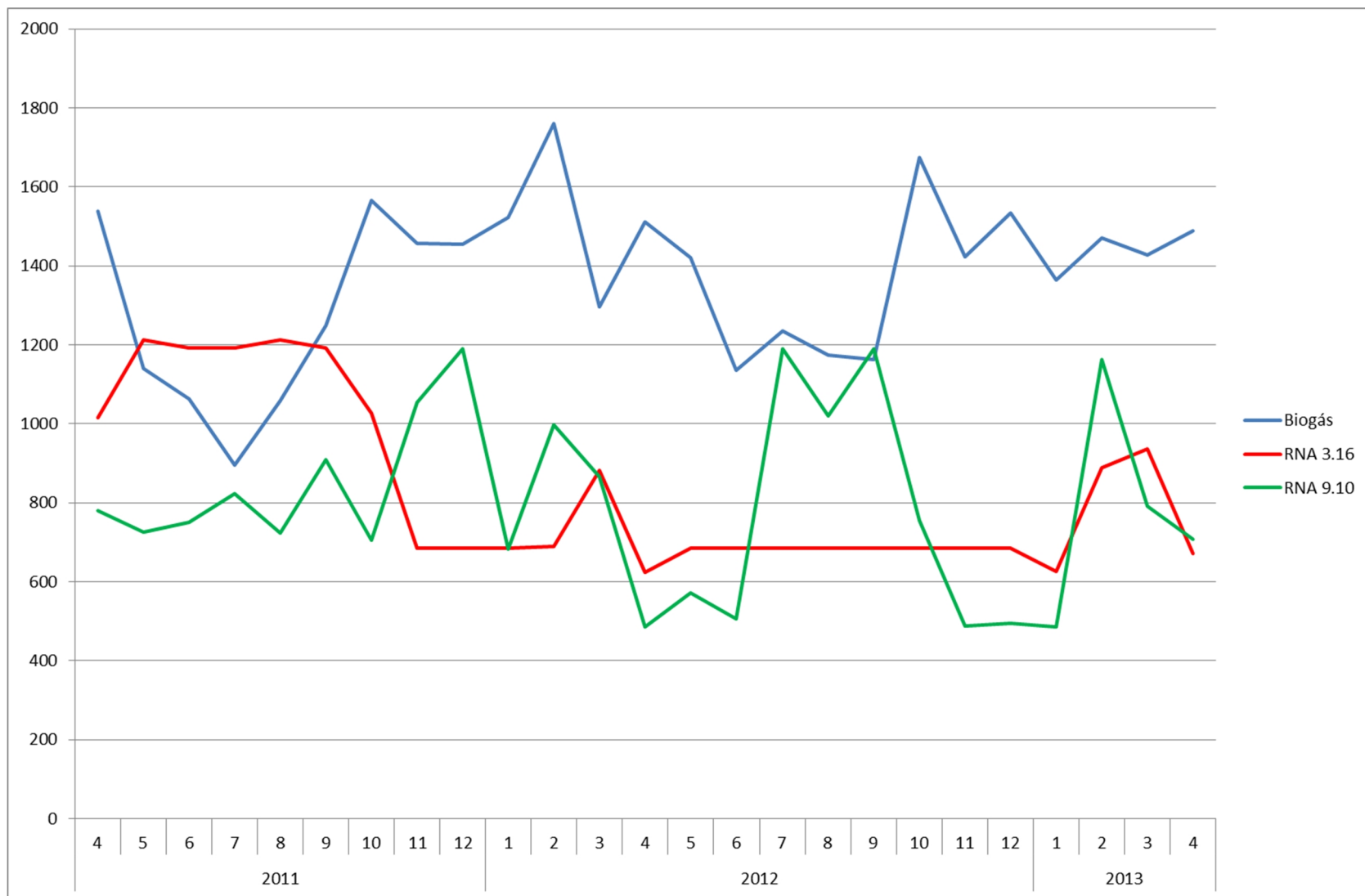


Gráfico com os resultados de previsões efetuadas com os valores de outra ETAR (ETAR Norte), com base nas médias mensais das variáveis.