



UNIVERSIDADE ESTADUAL DA PARAÍBA

CENTRO DE CIÊNCIAS E TECNOLOGIA

DEPARTAMENTO DE ENGENHARIA SANITÁRIA E AMBIENTAL

CLARISSA CÂMARA DE FREITAS

**APLICAÇÃO DE REDES NEURAS ARTIFICIAIS NA
MODELAGEM DE REATORES FOTOQUÍMICOS**

Campina Grande/PB

2012

CLARISSA CÂMARA DE FREITAS

**APLICAÇÃO DE REDES NEURAS ARTIFICIAIS NA
MODELAGEM DE REATORES FOTOQUÍMICOS**

Monografia apresentada ao Curso de Engenharia Sanitária e Ambiental da Universidade Estadual da Paraíba, referente ao Trabalho Acadêmico Orientado, em cumprimento às exigências para obtenção do título de Bacharel em Engenharia Sanitária e Ambiental.

Orientador: Prof. Dr. Fernando Fernandes Vieira

**Campina Grande/PB
2012**

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL – UEPB

F884a Freitas, Clarissa Câmara de.
Aplicação de redes neurais artificiais na modelagem de reatores fotoquímicos [manuscrito] / Clarissa Câmara de Freitas. – 2012.
45 f. : il. color.

Digitado.
Trabalho de Conclusão de Curso (Graduação em Engenharia Sanitária e Ambiental) – Centro de Ciências e Tecnologias, 2012.
“Orientação: Prof. Dr. Fernando Fernandes Vieira, Departamento de Engenharia Sanitária e Ambiental.”

1. Modelagem matemática. 2. Redes neurais artificiais. 3. Processos oxidativos avançados. I. Título.

21. ed. CDD 372.7

CLARISSA CÂMARA DE FREITAS

APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS NA MODELAGEM DE
REACTORES FOTOQUÍMICOS

MONOGRAFIA APROVADA EM: 21 / 06 / 2012 .

NOTA 10 (DEZ)

BANCA EXAMINADORA



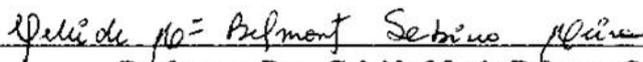
Professor Dr. Fernando Fernandes Vieira

Departamento de Engenharia Sanitária e Ambiental - CCT/UEPB
Orientador



Professor Dr. Rui de Oliveira

Departamento de Engenharia Sanitária e Ambiental - CCT/UEPB
Examinador



Professora Dra. Celeide Maria Belmont Sabino Meira

Departamento de Engenharia Sanitária e Ambiental - CCT/UEPB
Examinadora

*A Deus, cuja força serviu de alicerce em
minha vida, tornando possível superar cada
obstáculo, fazendo deste momento uma
grande conquista.*

AGRADECIMENTOS

A Deus pela grandeza de seu amor, que ao realizar mais um sonho, me mostrou que acreditar e confiar nele é válido sob qualquer circunstância.

Aos meus pais de criação, Miguel e Silene, que na ausência dos pais consanguíneos me concederam amor e confiança para que eu pudesse alcançar pontos extremos.

A Júlia, Joaquim e Marieliza, pela compreensão e carinho nesta importante caminhada.

Aos meus tios, Luiz Alberto e Salete, por terem acreditado nos meus esforços e investido na minha educação, sem os quais não chegaria aqui.

Aos meus mestres, Fernando Fernandes Vieira, Rui de Oliveira e Celeide Maria Belmont Sabino Meira, pela verdadeira fonte de inspiração durante estes cinco anos de caminhada, marcando minha vida profissional e pessoal.

A Emanuel Campos, Thassio Nóbrega, Juscelino Henriques, Igor Ogata e Cayo Farias, pela amizade inigualável, apoio e carinho, cujas habilidades me subsidiaram a experiência atual.

Aos meus companheiros de turma, que trilharam essa longa caminhada sem ceder diante dos desafios.

A Fúlvia Borges pela importante contribuição em meus estudos através de suas pesquisas, bem como à Universidade de São Paulo.

A todos os professores e funcionários que compõem o Departamento de Engenharia Sanitária e Ambiental, cujas experiências e a transmissão do saber foram marcantes para a construção de minha carreira profissional.

E, por fim, a todos que contribuíram para o êxito na obtenção do título de Bacharel em Engenharia Sanitária e Ambiental.

*“Que os vossos esforços desafiem as
impossibilidades, lembrai-vos de que as
grandes coisas do homem foram
conquistadas do que parecia impossível.”*
Charles Chaplin

RESUMO

As Redes Neurais Artificiais (RNA's) vêm sendo utilizadas como modelos preditivos, possuindo um potencial de exploração no tratamento de dados complexos, com aplicação em mecanismos de reações complexas como os dos reatores fotoquímicos, os quais possuem comportamento não linear. As RNA's possuem ampla capacidade de aprendizagem sendo possível, a partir de dados experimentais, condicionar as respectivas respostas, sem que haja conhecimento prévio de leis físicas ou químicas, não necessitando assim, de descrição matemática do fenômeno. Os Processos Oxidativos Avançados (POA's) se baseiam na geração do radical hidroxila, o qual é um forte agente oxidante, que pode ser empregado na degradação de diversos compostos poluentes, principalmente os recalcitrantes. A maioria dos POA's utilizam uma combinação de catalisadores e uma fonte de radiação eletromagnética para gerar os radicais hidroxila. Devido ao grande número de fatores que podem influenciar os POA's a modelagem fenomenológica torna-se muito complexa, por isso muitas vezes se recorre a procedimentos empíricos para a construção de modelos matemáticos para estes processos. O presente trabalho estudou a aplicação das RNA's na modelagem do processo de degradação do fenol, proveniente de efluentes industriais. Os dados experimentais usados na modelagem foram extraídos de um trabalho desenvolvido na Universidade de São Paulo (BORGES, 2009). Foi estudada a influência da variação do número de neurônios na camada oculta, bem como as funções de ativação intervenientes no processo de transmissão dos dados, o número de dados na camada de entrada e a influência de ruídos externos ao modelo. Os resultados obtidos mostraram que a técnica das RNA's pode ser aplicada demonstrando resultados primorosos para um modelo ajustado em função tangente hiperbólica demonstrando sua melhor aplicação em estudos de predição. Mesmo com um número de dados reduzido, a RNA também demonstrou bom ajuste, de modo que os custos com a etapa experimental podem ser reduzidos. Quanto aos ruídos provocados na rede, ocorreu uma pequena variação de modo que a rede perdeu em ajuste na etapa de validação, entretanto o desempenho da rede se manteve adequada.

PALAVRAS-CHAVE: Modelagem Matemática. Redes Neurais Artificiais. Processos Oxidativos Avançados.

ABSTRACT

Artificial Neural Networks (ANN) have been used as predictive models, having a potential for exploitation in the treatment of complex data, with application in mechanisms of complex reactions such as photochemical reactors, which have nonlinear behavior. The RNA's have ample capacity for learning is possible, from experimental data constrain the answers, without prior knowledge of physical or chemical laws, thus not requiring, the mathematical description of the phenomenon. The advanced oxidation processes (POA's) are based on the generation of OH radical which is a strong oxidizing agent, which can be employed in the degradation of various polluting compounds, especially recalcitrant. Most POA's use a combination of catalysts and a source of electromagnetic radiation to generate hydroxyl radicals. Due to the large number of factors that may influence the POA's phenomenological modeling becomes very complex, so many times by the use of empirical procedures for the construction of mathematical models for these processes. This work studied the application of ANN in modeling the degradation process of phenol from industrial effluents. The experimental data used in modeling were taken from a work at the University of São Paulo (Borges, 2009). The influence of variation in the number of neurons in the hidden layer and the activation functions involved in the process of transmitting data, the number of data in the input layer and the influence of external noise to the model. The results showed that the technique of RNA's can be applied tastefully to demonstrating results in an adjusted model hyperbolic tangent function showing its best application in studies of prediction. Even with a small amount of data, the RNA also showed a good fit, so that the cost of the experimental stage can be reduced. As regards the noise caused in the network, there was a slight variation so that lost in the network validation step adjustment, the performance of the network however is maintained appropriate.

KEYWORDS: Mathematical Modeling. Artificial Neural Networks. Advanced Oxidation Processes.

LISTA DE ILUSTRAÇÕES

Figura 3.1 - Neurônio Biológico.....	18
Figura 3.2 - Modelo não-linear de um neurônio.....	19
Figura 4. 3 - Esquema do Plano Experimental.....	27
Figura 4.4 -Topologia da rede: 5 neurônios na camada oculta (a); 10 neurônios na camada oculta (b); 15 neurônios na camada oculta (c) e 20 neurônios na camada oculta (d).....	29
Figura 5.5 - Gráfico de resultados experimentais versus calculados em função do tempo. (a) 80 mM de peróxido de hidrogênio; 25 g/L de cloreto de sódio e 0,2 mM de Ferro II. (b) 80 mM de peróxido de hidrogênio; 50 g/L de cloreto de sódio e 0,2 mM de Ferro II.....	38
Figura 5.6 - Comportamento da rede com 10 neurônios e funções de ativação e junção aditiva tangente hiperbólica.....	39
Figura 5.7 - Performance da rede, com melhor validação na época 6.....	40
Figura 5.8 - Comportamento da rede com dados contaminados com 1% de erro.....	41
Figura 5.9 - Performance da rede com dados contaminados.....	42

LISTA DE TABELAS

Tabela 5.1 - Valores de coeficiente de determinação em função da arquitetura para função tangente hiperbólica nas camadas oculta e de saída.	33
Tabela 5.2 - Valores de coeficiente de determinação em função da arquitetura para função tangente hiperbólica na camada oculta e sigmoideal na camada de saída.	34
Tabela 5.3 - Valores de coeficiente de determinação em função da arquitetura para função tangente hiperbólica na camada oculta e linear na camada de saída.....	34
Tabela 5. 4 - Valores de coeficiente de determinação em função da arquitetura para função sigmoideal na camada oculta e camada de saída.	35
Tabela 5. 5 - Valores de coeficiente de determinação em função da arquitetura para função linear na camada oculta e camada de saída.	35
Tabela 5.6 - Valores de coeficiente de determinação em função da arquitetura para função junção aditiva linear e função de ativação tangente hiperbólica. ..	36
Tabela 5.7 - Valores de coeficiente de determinação em função da arquitetura para função junção aditiva sigmoideal e função de ativação linear.	36
Tabela 5.8 - Valores de coeficiente de determinação em função da arquitetura para função junção aditiva sigmoideal e função de ativação tangente hiperbólica.	37
Tabela 5.9 - Valores referentes a variação do número de dados	40

LISTA DE ABREVIATURAS E SIGLAS

Net	- Função de junção aditiva
x	- Sinais de entradas
w	- Pesos das conexões sinápticas
$v(n)$	- Combinação linear associado a cada entrada
b	- Bias
$e(n)$	- Erro associado a cada entrada
$d(n)$	- Vetor de sinais de entrada desejados
m	- Variação de numero de entradas
i	- Posição do neurônio
j	- Posição dos pesos em relação a posição do neurônio
n	- Neurônio
E	- Erro global instantâneo
N	- Número de iterações
ϕ	- Função de ativação
δ	- Gradientes locais
Δw_{ij}	- Ajuste de pesos
Δb_j	- Ajuste do <i>bias</i>
η	- Taxa de aprendizado
ε	- Erro global médio
Σ	- Somatório

SUMÁRIO

1.0 INTRODUÇÃO	13
2.0 OBJETIVOS	15
3.0 REVISÃO DE LITERATURA.....	16
3.1 Redes Neurais Artificiais	16
3.2 Neurônios Biológicos	16
3.3 Neurônio Artificial	18
3.4 Histórico.....	19
3.5 Redes Neurais Artificiais e o processo de Aprendizagem	20
3.5.1 <i>Perceptrons de Múltiplas Camadas (MLP's)</i>	21
3.6 Processos Oxidativos Avançados (POA's)	22
3.2.1 <i>Fotólise de H₂O₂</i>	24
3.2.2 <i>Fotocatálise com TiO₂</i>	24
3.2.3 <i>Processos oxidativos em meio salino</i>	25
3.2.4 <i>Reações Fenton e foto-Fenton</i>	25
4.0 MATERIAIS E MÉTODOS.....	26
4.1 Equipamento	26
4.2 Modelagem Matemática.....	27
4.2.1 <i>Coleta e seleção dos dados</i>	27
4.2.2 <i>Configuração da rede</i>	28
4.2.3 <i>Treinamento</i>	30
4.2.4 <i>Teste</i>	30
4.2.5 <i>O Algoritmo</i>	30
4.2.6 <i>A propagação do sinal de erro</i>	31
5.0 RESULTADOS.....	33
6.0 CONCLUSÕES	43
REFERÊNCIAS	44

1.0 INTRODUÇÃO

A modelagem matemática é uma ferramenta que propicia a representação de um cenário físico através de metodologias e formulações matemáticas. Portanto, em sistemas ambientais, a modelagem constitui um instrumento notável, já que a mesma é capaz de representar de modo eficaz as atribuições do meio. Sua importância está fundamentada na redução de custos com plantas experimentais, sendo possível caracterizar e simular cenários sem haver a necessidade de sua reprodução imediata no meio físico.

O uso da modelagem matemática subsidia a produção do conhecimento de diversos cenários através da matemática, proporcionando a simulação dos sistemas reais, a fim de se avaliar o comportamento dos mesmos. Neste contexto, a modelagem possui recursos capazes de representar cenários complexos e, através de ferramentas computacionais, resolverem os problemas que envolvem as variáveis presentes no meio, de modo que se torna viável a aplicação de técnicas sem a necessidade de interferir no processo em escala real, possibilitando um envolvimento menor de riscos e custos inerentes a tais projetos.

Para tanto, são diversos os modelos disponíveis para as mais diversas áreas de conhecimento, tais como engenharia, física, química, biologia, economia, entre outras. Cada modelo possui o grau de complexidade envolvido para cada situação inerente. Entretanto, mais especificamente para questões relacionadas ao saneamento e tratamento de efluentes industriais, existem variáveis constituintes para tais processos que se tornam sobretudo difíceis para serem argumentados através de valores em modelos matemáticos clássicos, exigindo um tratamento mais complexo para solução destes problemas.

Em contrapartida a esta problemática, inserem-se num contexto mais amplo da modelagem computacional as Redes Neurais Artificiais (RNA's). Porém, as RNA's fazem parte de um teor abrangente que culminou em uma linha de pesquisa: a Inteligência Artificial. Estudos realizados por pesquisadores de diversas áreas do conhecimento, a partir da década de 70, período coincidente com a primeira visualização individual de um neurônio, tinham como meta a tentativa de reproduzir a complexa máquina cerebral. Estes estudos desenvolveram modelos empíricos complexos para a simulação de cenários que possuem variáveis extremamente

múltiplas, de difícil representação nos modelos clássicos, caracterizados pela interconexão das diversas informações inerentes ao modelo.

A partir de então, foi possível traçar uma linha tênue que liga os modelos de RNA's com a necessidade de se estabelecer sistemas de tratamento de efluentes em processos produtivos que constituem contaminantes de difícil remoção. Em um contexto mais específico, para o tratamento de diversos tipos de efluentes industriais faz-se necessária a aplicação de técnicas de tratamento integradas, dada a difícil remoção de poluentes contidos nestes efluentes.

Frente a tal cenário, os reatores fotoquímicos surgem como uma técnica de remoção eficaz para diversas substâncias poluentes que são de difícil remoção. O funcionamento destes apresenta-se como uma alternativa eficaz para o tratamento de água contaminada por poluentes como o fenol, o qual, sendo abundante nos efluentes industriais, mais especificamente em indústrias petroquímicas e químicas, é de toxicidade elevada e também se apresenta como inibidor ao tratamento biológico. Portanto, os reatores fotoquímicos são técnicas que se baseiam na geração de radicais fortemente oxidantes e pouco seletivos, de modo que promovam a degradação de uma grande variedade de poluentes. Deste modo, Processos Oxidativos Avançados (POA's) entram no contexto de reatores fotoquímicos, gerando radicais hidroxilas para a oxidação de poluentes sob diversas formas de reações químicas utilizando diversos oxidantes e indicados para compostos altamente tóxicos e recalcitrantes, promovendo a transformação dos mesmos em dióxido de carbono e água ou subprodutos com toxicidade atenuada.

Porém, a alta concentração de sais presentes nos efluentes industriais inibe a completa atuação dos Processos Oxidativos Avançados, fazendo com que as moléculas não degradadas permaneçam no efluente tratado. Desta forma, a integração com outros processos de remoção é útil para a completa eficiência de tratamento. O presente trabalho pretende avaliar mediante a simulação de RNA's, as interações pertinentes a reatores fotoquímicos pelas variáveis que permeiam o processo de tratamento.

O trabalho ora proposto representa uma contribuição importante às pesquisas em andamento, valendo-se de resultados já obtidos pela experiência de pesquisas da equipe da Universidade de São Paulo que atingem estudos avançados em experimentos e modelagem matemática de processos de remoção de fenol em efluentes industriais por diversas técnicas de tratamento.

2.0 OBJETIVOS

2.1 Objetivo Geral

Estudar a aplicabilidade da modelagem de reatores fotoquímicos, utilizando as Redes Neurais Artificiais, usando dados experimentais do tratamento de uma solução contendo fenol.

2.2 Objetivos Específicos

- Verificar a influência do número de neurônios na camada oculta;
- Avaliar a qualidade do ajuste de acordo com as funções de transferência;
- Analisar o desempenho da rede em consonância com a contaminação dos dados;
- Estudar o comportamento da rede em função do número de experimentos usados no treinamento, validação e teste.

3.0 REVISÃO DE LITERATURA

3.1 Redes Neurais Artificiais

As Redes Neurais Artificiais são um conjunto de procedimentos que buscam simular o comportamento do cérebro humano de modo que se possam representar cenários complexos de aprendizagem e caracterização de padrões tal qual o mesmo numa ótica de neurocomputação o faz.

Trata-se de um estudo que transporta através de modelos matemáticos, o funcionamento do cérebro humano. As Redes Neurais Artificiais (RNA's) interpretam o cérebro humano como um sistema de processamento de informações, admitindo que ele corresponda ao computador não-linear e altamente complexo (HAYKIN, 2001).

Através desta capacidade de organizar seus constituintes estruturais, os neurônios, torna-se possível a realização de certos processamentos. Mediante esta capacidade, surge a necessidade de se desenvolver modelos que possuam uma aproximação do funcionamento neurológico humano.

As Redes Neurais tornam-se dispositivos importantes quando as classes de problemas são de complexidade algorítmica extrema, de modo que sua resolução convencional é inviável. Além disso, as RNA's são tolerantes a ruídos e falhas, o que advém da capacidade de generalização da rede, suportando as imperfeições do ambiente mantendo os resultados desejados.

3.2 Neurônios Biológicos

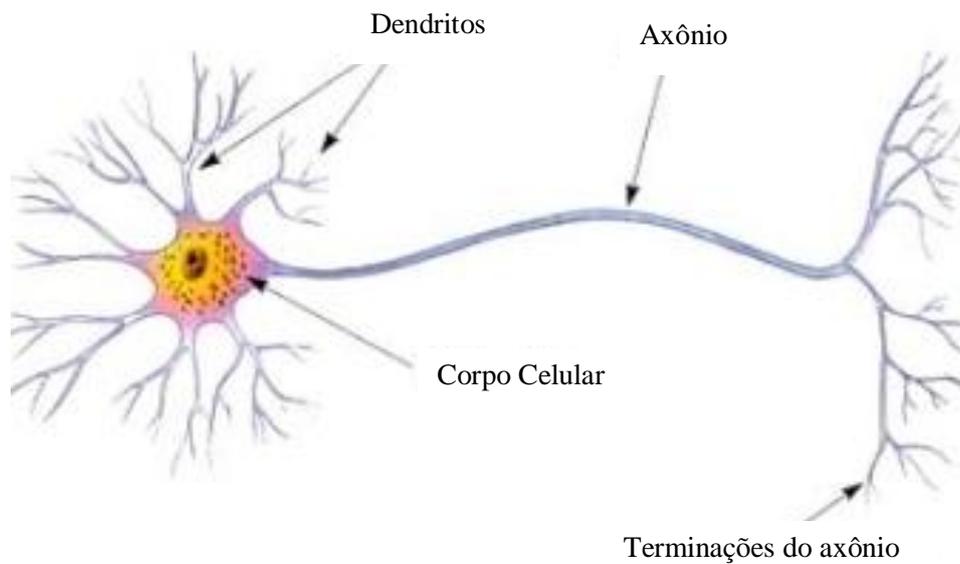
Pode-se dizer, portanto, que os primeiros esforços em se modelar matematicamente o cérebro humano surgiram logo após a primeira visualização do neurônio fisiológico por Camillo Golgi, em 1875. Golgi descobriu, ao acaso, que ao se colorir um conjunto de células neurais é possível a visualização individual. Golgi foi o contemporâneo a realizar estudos desta natureza, numa era em que as estruturas nervosas eram completamente desconhecidas (BARRETO, 1996).

Mas os estudos de Santiago Ramón y Cajal, em 1891, apresentaram dois resultados importantes para a evolução das pesquisas neurológicas. Primeiro Cajal adotou a noção de sistema nervoso, postulando sobre a comunicação entre as células

pela sinapse (BARRETO, 1996). A segunda foi de que a interconexão entre neurônios seria altamente específica e estruturada. Descrevendo diversas estruturas cerebrais, ele fez com que as pesquisas se voltassem não apenas para a estrutura dos neurônios, mas também para o desconhecido campo das muitas estruturas neuronais (BARRETO, 2002).

A partir de então, pode-se avaliar que o neurônio típico é composto por um corpo celular ou soma, um axônio tubular e várias ramificações arbóreas conhecidas como dendritos. Esta célula nervosa faz parte de um sistema que controla as ações rápidas do corpo, com tempo de resposta igualmente rápido. Existem dois tipos principais de células no cérebro: os neurônios e a glia. É comum atribuir aos neurônios as principais funções cerebrais, já que a função da glia ainda está por ser entendida. Existem aproximadamente 10^{11} neurônios no cérebro humano, com um número de sinapses variando entre 100 e 100.000, sendo tanto excitatórias quanto inibitórias. As sinapses são definidas como a conexão entre o fim do bulbo e o dendrito do próximo neurônio, a Figura 3.1 mostra a ilustração de um neurônio, onde pode-se perceber estruturas que compõem os dendritos chamados de espinhas dendritais. Tais estruturas formam, segundo Haykin (1994), microcircuitos neurais, que se referem a agrupamentos de sinapses organizadas em padrões de conectividade para produzir uma ação de interesse.

As sinapses se propagam graças a um potencial de ação existente dentro e fora das células, constituído de íons de Na^+ e K^+ , que provocam um ponto de equilíbrio de -85 milivolts, sendo o interior da célula negativo com relação ao exterior.

Figura 3.1 - Neurônio Biológico

Fonte: *Adaptado de* Baughman (1995).

3.3 Neurônio Artificial

Apesar de haver grandes esforços em se modelar matematicamente o cérebro humano, tudo o que se conseguiu, até os dias atuais, foi uma aproximação elementar. Como base destas informações, temos a representação de um neurônio segundo Haykin (1994) na Figura 3.2, onde as sinapses ou elos de conexão são caracterizados individualmente por um peso, ou força própria, inspirados nas forças sinápticas. Desta forma temos o chamado *net* do neurônio, que é definida como a representação matemática do neurônio biológico, dado pela Equação 1.

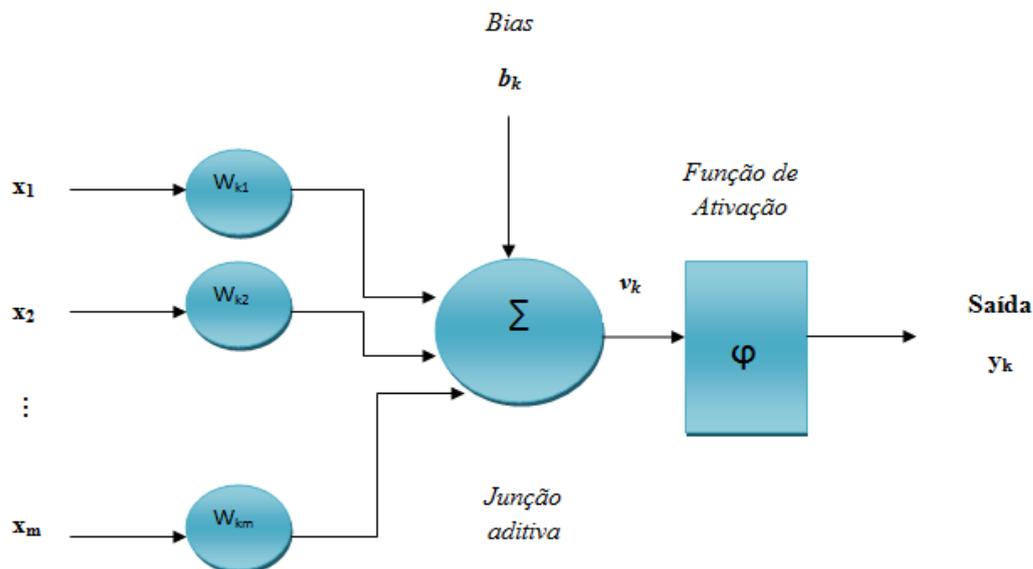
$$Net = \sum_1^m x_j w_{ij} \quad (1)$$

Onde w_{ij} representa o peso da conexão sináptica do neurônio i para o neurônio j . Os sinais de entrada x_m transmitidos são ponderados pelos seus respectivos pesos w_{km} e pelo bias b_k .

Ocorre também a presença de um somador \sum , que soma os sinais de entrada, ponderados pelos respectivos pesos. A este conjunto de operações de entrada denomina-se combinador linear.

Há ainda a presença de uma função de ativação φ que restringe o intervalo do valor de saída y_k a um valor finito.

Figura 3.2 - Modelo não-linear de um neurônio



Fonte: *Adaptado de Haykin (1994)*

3.4 Histórico

Com o propósito de reproduzir o funcionamento do cérebro humano de modo que se faça emergir comportamento inteligente, McCulloch e Pitts (1943) conceberam o primeiro modelo de neurônio artificial, baseados em uma modelagem matemática que buscou reproduzir o neurônio biológico. Tal acontecimento foi considerado o marco inicial dos novos estudos em Redes Neurais, e o marco final da chamada *Época Antiga* das mesmas, caracterizada pelas conquistas da psicologia e neurofisiologia, apesar daquela se sobressair em termos de estudos avançados.

Foi no encontro de *Darth-mouth College*, em 1956, que concretizou-se tais esforços para estudar as Inteligências Artificiais, motivando a posterior publicação do livro sob o título de *Automata Studies* agregando em seu conteúdo o primeiro artigo que tratava de redes neurais como uma arquitetura computacional. Nasceram, então, concomitantemente os paradigmas simbólicos e conexionistas de inteligência artificial.

Nos paradigmas de Inteligência Artificial Simbólicos, a inteligência global é simulada sem levar em consideração os mecanismos responsáveis por este comportamento. Já nos paradigmas de Inteligência Artificial Conexionista, a reprodução

do comportamento inteligente se baseia no funcionamento do cérebro humano mediante simulação das mesmas conexões fisiológicas. Tal fundamento faz emergir a era Romântica da abordagem conexionista, tendo seu princípio com o trabalho de Donald Hebb, com seu livro *Organization of Behavior*, sugerindo através do mesmo a intenção de se inspirar na natureza humana para reproduzir comportamentos inteligentes. A ideia de que qualquer problema poderia ser facilmente resolvido pelas redes neurais ganhou ainda mais intensidade com o livro *Perceptron* de Frank Rosenblatt, cujo título é definido como uma rede neural de duas camadas de neurônios capaz de aprender de acordo com a lei de Hebb.

Em 1969, Minsky e Papert declaram em seu livro *Perceptrons* que a rede proposta por Rosenblatt não era capaz de distinguir padrões linearmente separáveis (MINSKY & PAPERT, 1969).

Tal publicação fez surgir a chamada era das Trevas, que durou até 1981, contudo pesquisas continuaram sendo desenvolvidas mas de modo mais discreto e tímido.

Porém, foi com o algoritmo de treinamento de Paul Werbos criado em 1974, popularizado em 1986, que o algoritmo de *Backpropagation* que marcou o renascimento da Inteligência Artificial Conexionista, também caracterizada pela popularização dos Sistemas especialistas e pela Primeira Conferência Internacional em Redes Neurais, marcando o final do período.

Desde 1987 até os dias atuais temos o período conhecido como Era Contemporânea, onde logo no início do período Gallant (1988) publica seu célebre artigo sobre sistemas especialistas conexionistas. Este foi o ponto de partida para a união das duas abordagens de inteligência artificial.

3.5 Redes Neurais Artificiais e o processo de Aprendizagem

O processo de conhecimento nas redes neurais artificiais ocorre pela ponderação dos pesos da conexão entre os neurônios. Sendo assim, encontrar a melhor solução para um determinado problema seria encontrar o melhor ajuste entre estes pesos, e, por conseguinte, a melhor topologia da rede.

Uma abordagem geral do que vem a ser o processo de aprendizagem, é a que foi adotada por Braga et al. (2000), onde a aprendizagem é o processo de apresentação

dos parâmetros de uma rede neural, os quais são ajustados de acordo com os estímulos do ambiente ao seu entorno.

A forma de aprendizado recebe duas classificações quanto ao modo de operação: supervisionado ou não supervisionado. O aprendizado supervisionado implica na presença de um professor que confere o quanto a rede está próxima de uma solução confiável, usado para previsão de resultados, como o caso da previsão de enchentes em um sistema hidrelétrico. Entretanto, o aprendizado não supervisionado não possui conhecimento *a priori* das saídas da rede, funcionando como meio de distinguir padrões e classes diferentes. Tal procedimento se adequa no reconhecimento de padrões, como a vinculação de dados pessoais no instante em que se passa o cartão de crédito num caixa eletrônico.

Entre os tantos algoritmos existentes no estudo de RNA's, o algoritmo de *backpropagation* constitui o algoritmo de treinamento mais utilizado, consistindo de um aprendizado supervisionado que visa achar iterativamente a mínima diferença entre as saídas desejadas e as saídas obtidas, executando, então, a retropropagação do erro encontrado em cada iteração.

O algoritmo *backpropagation* foi desenvolvido por vários pesquisadores de maneira independente e em épocas distintas. Em 1974, Werbos (WERBOS, 1974) apresentou o algoritmo enquanto desenvolvia sua tese de Doutorado, denominando o mesmo de “Algoritmo de Realimentação Dinâmica”. Parker, ao introduzir novos estudos do mesmo algoritmo, o denominou de “Algoritmo de Aprendizado Lógico”. Contudo, o trabalho de Rumelhart, Hinton e Willians (RUMELHART & MCCLELLAND, 1986) divulgou e popularizou a técnica de aprendizado supervisionado com o Algoritmo de denominação fixa até os dias atuais de *backpropagation*.

3.5.1 Perceptrons de Múltiplas Camadas (MLP's)

Com o passar dos anos, estudos realizados no âmbito de RNA's, avaliaram que para melhor condicionamento de ajuste da rede, a inserção de uma camada de neurônios como camada escondida ou camada sensorial poderia superar o problema do aprendizado não-linearmente separáveis (LUDWIG e COSTA, 2007).

A principal característica da implementação da camada escondida ou camada oculta, é que seus elementos se organizam de tal modo que aprendam a reconhecer elementos diferentes da camada de entrada e a rede venha, então, a decidir as características que devem ser extraídas do conjunto de saída.

As redes de Múltiplas Camadas, como também são conhecidas as redes que possuem camada oculta, apresentam um poder computacional muito maior do que aquelas que não apresentam as camadas intermediárias (BRAGA et al., 2000).

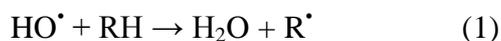
Um dos principais aspectos relacionados às redes MLP's diz respeito à função de ativação utilizada. Estas funções devem ser não-lineares e diferenciáveis, já que as funções lineares limitam a rede de modo que ela não possa generalizar, ou seja, ela está sujeita as interferências do ambiente ao qual ela se insere. O fato delas serem diferenciáveis se deve ao fato do seu gradiente poder ser calculado, direcionando os pesos e ajustes. A maioria destas funções é não-decrescente, sendo a mais utilizada a função logística ou sigmoideal logística, porém ainda são utilizadas a tangencial hiperbólica e a linear, as quais serão testados neste trabalho.

3.6 Processos Oxidativos Avançados (POA's)

Devido ao alto poder oxidativo do peróxido de hidrogênio, atrás apenas do flúor, os POA's têm delineado uma importante alternativa de tratamento de efluentes contaminados com poluentes persistentes ou com alta carga de matéria orgânica.

O processo Oxidativo Avançado (POA) surgiu da necessidade de se manter um padrão de cautela em termos de salubridade humana. Além disso, o método apresenta-se como uma tecnologia de menor custo na desinfecção bem como na geração de resíduos. Tal processo é definido como a geração de radicais livres reativos, gerados por diferenciados métodos, podendo reagir com uma gama de compostos químicos devido à sua alta reatividade que, em meio ácido possui um potencial oxidativo ainda maior.

Neste processo podem ser geradas espécies transientes que possuem alto poder oxidativo, porém baixa seletividade. O processo adotado para degradação de matéria orgânica está descrita pelas reações, sendo estas:





A técnica de POA's tem recebido mérito pela conversão durante a desinfecção, de poluentes químicos inócuos em matéria mineralizada, como é o caso do CO₂.

Embora hajam reconhecidas vantagens nos POA's, sua adoção como método de tratamento ainda é limitada pelo fato de seus reagentes serem onerosos e pelos custos de operação. Outra desvantagem é quanto aos custos com a luz ultravioleta (GALINDO, 2001).

Os principais POA's usam agentes oxidantes na presença de um catalizador apropriado e/ou luz ultravioleta como H₂O₂/UV; O₃/UV; O₃-H₂O₂/UV; [TiO₂]/UV; Fe (III) [TiO₂]/UV_{visível}; UV_{vácuo}; reação de Fenton ou H₂O₂-Fe (II); e reação de foto-Fenton ou H₂O₂[Fe(II)/Fe(III)]/UV (LENGRINI et al., 1993).

De acordo com Adreozzi et al. (2003) a absorção de luz sob radiação ultravioleta comumente utilizada se encontra na faixa de 250 – 400 nm. Esta técnica, por utilizar radiação cujas faixas são componentes da radiação solar, é muito apreciada em regiões tropicais e subtropicais. Entretanto, em áreas com pouca radiação solar, reatores convencionais equipados com lâmpadas ultravioletas ou associados a reatores solares podem ser adotados.

Os processos, apesar de apresentarem características que o diferem uns das outras, apresentam similaridade no que se refere ao mecanismo de ação devido à participação de radicais hidroxila (HO[•]) na maior parte dos mecanismos de reação.

As reações de liberação dos radicais hidroxila, que, como já foi mencionado, possuem um elevado potencial de oxidação, geram radicais orgânicos intermediários, que promovem outras reações resultando em produtos finais (P_{oxid}) provenientes de uma série de reações em cadeia da molécula inicial (R) (BORGES, 2009).



Considerando a elevada reatividade e instabilidade dos radicais hidroxila, é necessário que ocorra a sua geração contínua no reator por reações químicas e fotoquímicas. A seguir os principais processos para obtenção desses radicais, enfatizando que às reações Fenton e foto-Fenton serão abordadas com mais profundidade, já que será a técnica utilizada para o tratamento fotoquímico.

3.2.1 Fotólise de H_2O_2

Tal processo é efetuado pela adição de peróxido de hidrogênio (H_2O_2) e irradiação do efluente a ser tratado com luz ultravioleta em comprimentos de onda inferiores a 280 nm, gerando radicais hidroxila em solução aquosa. Este processo é aplicado ao tratamento de efluentes com compostos orgânicos que absorvam radiação abaixo de 300 nm (LEGRINI et al., 1993).

3.2.2 Fotocatálise com TiO_2

A fotocatalise heterogênea é caracterizada por um semicondutor que é excitado pela absorção de fótons gerando elétrons que migram da banda de valência para a banda de condução. Essas espécies podem se ligar a outras espécies ou migrarem formando uma reação de redox.

Ao haver a migração de elétrons em bandas, formam-se lacunas que liberam o íon H^+ , formando hidroxilas pela reação com partículas do semicondutor.

As principais vantagens da fotocatalise são:

- Ampla faixa de compostos podem ser mineralizados;
- Elimina a adição de oxidantes químicos;
- O catalisador pode ser reutilizado;
- Processo de baixo custo;
- A radiação solar pode ser empregada como fonte de luz.

A fotocatalise utiliza oxigênio como ativador de reação. A maior parte dos estudos utiliza óxidos metálicos de dimensões nanométricas (OLLIS, 2003).

A maior parte dos processos utiliza dióxido de Titânio (TiO_2), pela sua ampla capacidade de remoção de poluentes.

3.2.3 Processos oxidativos em meio salino

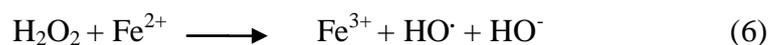
Os processos oxidativos de compostos orgânicos são inibidos de diversas formas pela presença de íons inorgânicos como cloreto, fosfato, sulfato, fluoreto, brometo, entre outros. Tais ânions podem ser encontrados normalmente em efluentes industriais ou podem ser formados como produtos de decomposição durante o processo de oxidação (BORGES, 2009).

No caso do processo de foto-Fenton, a inibição ocorre pela presença de certos íons que são advindos da precipitação de ferro, consumo de radicais hidroxila ou por sua associação ao Fe(II) formando complexos menos reativos.

3.2.4 Reações Fenton e foto-Fenton

O processo de foto-Fenton é resultante de reações que utilizam como catalizadores óxidos de ferro, inclusive na remoção de corantes. Resíduos sólidos com alto teor de ferro podem ser adequadas fontes de catalisadores em sistemas heterogêneos, diminuindo o consumo de sais de ferro, além da possibilidade de seu reuso.

A reação de Fenton consiste na geração de radicais hidroxila pela reação entre H_2O_2 e íons Fe^{2+} em meio ácido, segundo apresentado na Equação 6 (ESPLUGAS et al., 2002):



Tal método é considerado como um dos mais simples para obtenção de hidroxilas, de acordo com Andreozzi et al. (2003), além de ser vantajoso, pois o ferro é um elemento muito abundante na natureza e não tóxico, e o H_2O_2 tem um manuseio prático e ambientalmente seguro.

O processo pode ser acelerado pela incidência de radiação UV em comprimentos de onda menores que 300 nm. Nessas condições ocorre a regeneração de Fe^{3+} a Fe^{2+} com ocorrência de Fenton, devido à presença de H_2O_2 (ANDREOZZI et al., 1995). Este processo apresenta como principais vantagens o alto contato entre o agente oxidante e o poluente e sua sensibilidade à luz em comprimentos de onda de até 600 nm. Como desvantagem a necessidade de remoção de ferro após a reação.

4.0 MATERIAIS E MÉTODOS

Para a realização deste projeto, as atividades experimentais que serviram de fonte de dados foram realizadas na Universidade de São Paulo pela pesquisadora Fúlvia Jung Borges (2009), cuja tese foi apresentada a Escola politécnica de São Paulo, em parceria com o *Laboratoire de Génie Chimique* de Toulouse, França. Os experimentos realizados no âmbito de Processos Oxidativos Avançados ocorreram em um reator em batelada com volume interno de 1 litro, conectado a um tanque de recirculação de vidro borossilicato, munido de agitador mecânico. A temperatura da solução foi controlada, por meio de banho termostático, a 25 °C. A solução circulou através de um reator numa vazão de aproximadamente 1,5 litros.

4.1 Equipamento

O reator munido de lâmpada de vapor de mercúrio de média pressão imersa em um poço de quartzo resfriado com água teve seus experimentos iniciados com os processos de foto-Fenton e UV/H₂O₂ pela adição de uma solução aquosa de peróxido de hidrogênio (H₂O₂) a uma taxa de 0,83 mL/min e o acionamento da lâmpada. Para as reações de foto-Fenton, adicionou-se FeSO₄.7H₂O (sulfato ferroso heptahidratado). Amostras de 10 mL cada foram coletadas para análise em intervalos de tempo apropriados durante o tempo de irradiação (em média 2h).

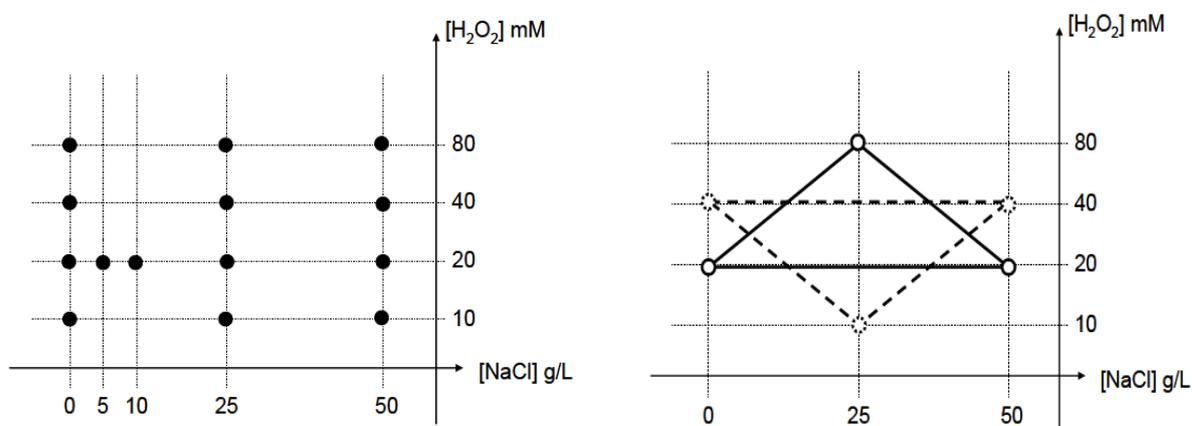
As reações foram iniciadas pela adição de uma solução aquosa de H₂O₂, a uma taxa de 0,83 mL/min e o acionamento da lâmpada. As concentrações de H₂O₂ citadas se referem a quantidade total adicionada dividida pelo volume total do sistema. No entanto essa quantidade é adicionada a uma vazão constante durante 90 minutos iniciais do experimento. Para as reações foto-Fenton, FeSO₄.7H₂O (sulfato ferroso heptahidratado) foi adicionada ao sistema.

Amostras de 10 mL foram retiradas para análise em intervalos de tempo apropriadas durante o tempo de irradiação, em média 2 h. Solução inibidora contendo 10 mL de uma mistura de 0,05 M KI; 0,05 M Na₂SO₃ e, 0,05 M NaOH é adicionada nas amostras contendo o processo foto-Fenton e UV/ H₂O₂, para interromper a reação; posteriormente as amostras são filtradas em membranas de 0,45 µm (Milipore Durapore) para remover as espécies de ferro precipitadas e posteriormente, analisar a quantidade de carbono orgânico total (COT). Em cada experimento utilizam-se 2,0 L de

solução aquosa de 100 ppm C de fenol e concentrações diferentes de NaCl (0 a 50 g/L). O pH é ajustado para o valor desejado e mantido constante durante o experimento, usando-se as soluções aquosas de H_2SO_4 ou NaOH. O pH foi mantido em 3 para minimizar os efeitos da concentração de NaCl no processo de oxidação (BORGES, 2009).

Para a execução do experimento foram utilizadas as concentrações de H_2O_2 que variaram entre 50 e 80 mM, entre 0 e 50 g/L para NaCl e 0,1 e 0,4 para Fe (II). Estas concentrações foram determinadas pelo planejamento experimental baseado no projeto Doehlert (De AGUIAR e MASSART, 1998).

Figura 4. 3 - Esquema do Plano Experimental



Fonte : BORGES (2009)

4.2 Modelagem Matemática

4.2.1 Coleta e seleção dos dados

Segundo Ludwig e Costa (2007), uma coleta criteriosa dos dados relativos ao problema torna-se a pedra fundamental para o desempenho da rede neural. Portanto, foi realizado um tratamento criterioso dos dados, sendo iniciado pela transmutação de sua forma de gráficos para tabelas.

A extração de dados experimentais teve como início a permuta da forma de gráficos para tabelas numéricas. Através do programa *Engauge 2.0*, os dados foram digitalizados, ponto a ponto, de curvas e convertidas em valores possíveis de serem

trabalhados, ou seja tabelas com valores numéricos de modo a representar, através de suas variáveis de entrada, os sistema.

Posteriormente, os dados foram submetidos a uma análise criteriosa com a finalidade de minimizar ambiguidades e erros. Verificou-se o domínio dos dados perante o problema, ou seja, se a coleção é capaz de abranger amplamente as condições do cenário a ser modelado matematicamente, os dados foram, então, divididos em três categorias:

- Dados de treinamento, usados para o treinamento da rede;
- Dados de validação da rede, usados para a verificação do desempenho da rede,e;
- Dados para testes de generalização da rede.

O conjunto foi dividido em 60% dos dados para treinamento, 20% dos dados para validação e 20% dos dados para testes.

Os dados foram reordenados aleatoriamente para evitar tendências associadas à ordem de sua coleta.

A rede teve como dados de entrada uma coleção de 110 dados, sendo que este número variou ao longo do estudo a fim de fazer cumprido o estudo referente à variação do número de dados. Tal variação passou de 15 dados, 20 dados, 50 dados e, como já esperado, a série de 110 dados. As variáveis selecionadas foram extraídas de modo a representar melhor o sistema a ser modelado. Portanto muitas outras variáveis poderiam ser usadas, como a intensidade de radiação e o pH, porém as selecionadas para este estudo foram classificadas como variáveis de estado, pois elas caracterizam o sistema em sua diversas variações. Para a caracterização da resposta do sistema delimitado pelo Fenol, foi selecionado o Carbono Orgânico Total (COT), já que se trata de um composto orgânico.

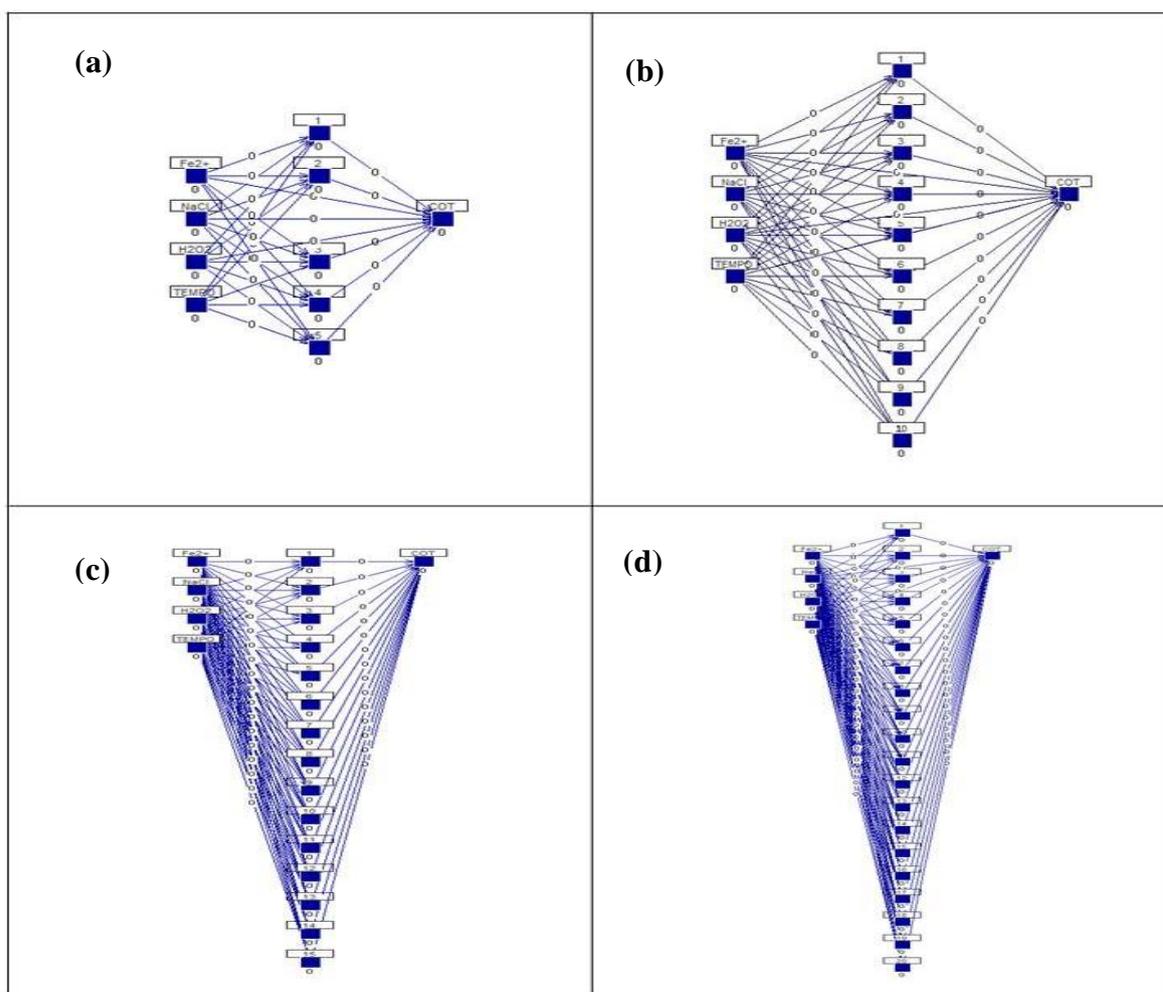
4.2.2 Configuração da rede

A rede foi desenvolvida seguindo a metodologia elaborada por Ludwig e Costa (2007), dividindo-se em quatro etapas:

- **Seleção do paradigma da rede mais apropriada** - sendo esta o Perceptron de Múltiplas Camadas;

- **Determinação da topologia da rede** - a atribuição mais adequada é a descrita na Figura 4.3, onde tem-se 4 neurônios na camada de entrada. No presente estudo foram realizados testes com 5, 10, 15 e 20 neurônios na camada oculta. A Figura 4.3 apresenta uma das topologias estudadas, contendo 10 neurônios na camada oculta;
- **Determinação do algoritmo de treinamento** - o algoritmo escolhido foi o de *backpropagation*, pela sua disseminação nos estudos já desenvolvidos em RNA's;
- **Determinação da função de transferência** – nesta etapa foram abordados os estudos referentes à sua variação em três tipos principais, os quais são: função linear, sigmoideal e tangente hiperbólica.

Figura 4.4 - Topologia da rede: 5 neurônios na camada oculta (a); 10 neurônios na camada oculta (b); 15 neurônios na camada oculta (c) e 20 neurônios na camada oculta (d).



4.2.3 Treinamento

Nesta etapa foram ajustados os pesos das conexões. A posteriori, ocorre a determinação de valores aleatórios uniformemente distribuídos em um intervalo definido.

O algoritmo de aprendizagem é ajustado de acordo com o tipo da rede, sendo um processo empírico. O tempo da rede pode ser adotado de acordo com alguns indicadores, porém o tempo de ciclos é adotado tendo como indicador o erro quadrático médio, que para este estudo teve um máximo de 0,001, ou seja, a rede não poderá ultrapassar este valor.

4.2.4 Teste

Durante esta fase, o conjunto de dados separados para validação foi utilizado para determinar o desempenho da rede com dados que não foram apresentados à mesma.

4.2.5 O Algoritmo

O Algoritmo de *backpropagation* pode ser descrito como a seguir:

1. Arbitram-se valores de pesos sinápticos e níveis de *bias* com distribuição uniforme cuja média deve ser zero;
2. Apresenta-se uma época de exemplos à rede, sendo que, para cada exemplo, realiza-se a propagação dos sinais e a retropropagação dos erros com a correção de pesos e *bias*;
3. Aplica-se à camada de entrada da rede o vetor de sinais de entrada $x(n)$ e calcula-se o campo local induzido e o sinal de saída para todos os neurônios, onde é obtido o vetor de sinais funcionais da rede $y(n)$. Calcula-se, portanto, o sinal de erro $e_f(n)$ para cada neurônio da camada de saída, pela comparação de $y(n)$ com o vetor de sinais de saída desejados $d(n)$. Calcula-se, por fim, o erro instantâneo e , com este, o erro médio global para teste de finalização.
4. Faz-se os cálculos dos gradientes locais para todos os neurônios da camada de saída:

$$\delta_j(n) = -e_j(n)\varphi_j(v_j(n)) \quad (7)$$

Em seguida, calculam-se os ajustes para os pesos daquela camada, bem como dos *bias* os quais devem ser somados valores atuais, de acordo com a taxa de aprendizagem:

$$\Delta w_{ij} = \eta e_j(n)\varphi_j(v_j(n))y_i(n) = -\eta\delta_j(n)y_i(n) \quad (8)$$

$$\Delta b_j = \eta e_j(n)\varphi_j(v_j(n))y_i(n) = -\eta\delta_j(n)y_i(n) \quad (9)$$

Calcula-se o gradiente local para os neurônios da penúltima camada ou camada oculta:

$$\delta_i(n) = \varphi_i(v_i(n)) \sum_{j=1}^J (\delta_j(n)w_{ij}(n)) =$$

$$\Delta w_{ki} = -\eta\varphi_i(v_i(n))y_k(n) \sum_{j=1}^J (\delta_j(n)w_{ij}(n)) = -\eta\delta_i(n)y_k(n),$$

$$\begin{aligned} \Delta b_i &= -\eta\varphi_i(v_i(n))y_k(n) \sum_{j=1}^J (\delta_j(n)w_{ij}(n)) \\ &= -\eta\delta_i(n)y_k(n) \end{aligned} \quad (10)$$

O processo prossegue de forma análoga em todas as camadas, sendo que após a camada de entrada da rede o valor $y_k(n)$ por $x_k(n)$.

Iteram-se as computações apresentando novas épocas de exemplos de treinamento para a rede de forma aleatória de época para época, até que seja satisfeito o critério de parada, que pode ser o número máximo de iterações ou um valor limite para o erro global médio da rede (LUDWIG e COSTA, 2007).

4.2.6 A propagação do sinal de erro

Tendo em vista que todo o processamento de uma rede neural pelo método de *backpropagation* se baseia no sinal de erro, e que este está fundamentado nos pesos e

sinais de entrada e que, por conseguinte, ambos dependem da taxa de aprendizagem, temos que o sinal de erro do neurônio de saída j na iteração n é definido por:

$$e_j(n) = d_j(n) - y_j(n) \quad (11)$$

O erro de cada neurônio é elevado ao quadrado para evitar um erro negativo (LUDWIG e COSTA, 2007).

$$E(n) = \frac{1}{2} \sum_1^n e_j^2(n) \quad (12)$$

O erro global médio da rede para todo o conjunto de treinamento, com um número de iterações N para apresentar um ciclo de treinamento ou época, é definido pela média aritmética dos erros globais instantâneos, conforme expressa a Equação (13):

$$\varepsilon = \frac{1}{N} \sum_{n=1}^N E(n) \quad (13)$$

O teste de incerteza foi realizado com o objetivo de verificar o ajuste da rede neural e sua capacidade de generalização, mesmo diante de ruídos externos.

Para sua melhor representatividade, foram gerados dados randomizados que variaram entre 0,01 e -0,01 totalizando em 20 dados sem repetições. Estes dados randomizados foram somados a uma alíquota do conjunto total de 20 dados experimentais, os mesmos usados para a formulação do modelo matemático. Diante desta contaminação de dados, foi treinada novamente a rede e avaliados os resultados obtidos desta nova etapa.

5.0 RESULTADOS

São apresentados neste capítulo os resultados referentes ao modelo desenvolvido para o POA.

O primeiro teste refere-se ao ajuste de 5, 10, 15 e 20 neurônios na camada oculta, aplicando-se a função tangente hiperbólica como função aditiva e função de ativação, obtendo-se a Tabela 5.1:

Tabela 5.1 - Valores de coeficiente de determinação em função da arquitetura para função tangente hiperbólica nas camadas oculta e de saída.

R^2				
<i>Arquitetura</i>	<i>Treinamento</i>	<i>Validação</i>	<i>Teste</i>	<i>Total</i>
5	0,99818	0,99766	0,99715	0,99792
10	0,9965	0,98459	0,99139	0,99285
15	0,9994	0,99626	0,99717	0,99866
20	0,9994	0,99768	0,99707	0,99863

Observa-se que para a função tangente hiperbólica em ambas as camadas, o coeficiente de correlação possuiu um excelente ajuste para todas as topologias de rede (ou arquiteturas). Entretanto, os coeficientes para o treinamento aumentam proporcionalmente ao aumento do número de neurônios na camada oculta, o que se torna evidente já que se trata do instante em que a rede aprende a operar o sistema. Porém, os coeficientes variam nas etapas seguintes, sendo possível verificar que a melhoria do desempenho aumenta com o número de neurônios da camada oculta.

Para o teste realizado mantendo-se a função aditiva como tangente hiperbólica e a função de ativação como sendo sigmoideal, obteve-se a Tabela 5.2:

Tabela 5.2 - Valores de coeficiente de determinação em função da arquitetura para função tangente hiperbólica na camada oculta e sigmoidal na camada de saída.

R^2				
<i>Arquitetura</i>	<i>Treinamento</i>	<i>Validação</i>	<i>Teste</i>	<i>Total</i>
5	0,9521	0,88317	0,97148	0,94475
10	0,94175	0,95487	0,92808	0,94259
15	0,78278	0,94025	0,87584	0,82341
20	0,93737	0,91743	0,95922	0,93705

O desempenho da rede não demonstrou melhorias em relação ao primeiro teste, observando-se que mesmo com o aumento do número de neurônios a rede manteve a média entre os pontos na dispersão.

Como um terceiro teste, foi feita a aplicação da tangente hiperbólica na camada oculta, e função linear na camada de saída.

Tabela 5.3 - Valores de coeficiente de determinação em função da arquitetura para função tangente hiperbólica na camada oculta e linear na camada de saída.

R^2				
<i>Arquitetura</i>	<i>Treinamento</i>	<i>Validação</i>	<i>Teste</i>	<i>Total</i>
5	0,96956	0,97959	0,98767	0,97448
10	0,97075	0,98777	0,97843	0,97466
15	0,96862	0,97821	0,99254	0,97399
20	0,97269	0,9798	0,9748	0,97286

Para tais resultados, também não são observadas melhorias atribuídas a mudança das funções. Entretanto, o coeficiente de correlação manteve um bom ajuste estatístico.

A Tabela 5.4 apresenta os valores de R^2 para a função aditiva e função de ativação sigmoidal.

Tabela 5. 4 - Valores de coeficiente de determinação em função da arquitetura para função sigmoidal na camada oculta e camada de saída.

R^2				
<i>Arquitetura</i>	<i>Treinamento</i>	<i>Validação</i>	<i>Teste</i>	<i>Total</i>
5	0,94037	0,89231	0,96391	0,93551
10	0,92357	0,98407	0,97229	0,93966
15	0,92984	0,93091	0,91583	0,92775
20	0,93799	0,98224	0,88841	0,93547

Percebe-se mediante a apresentação de tais valores que os coeficientes se mantiveram relativamente constantes, com resultados inferiores aos testes anteriores. A Tabela 5.5 apresenta os valores relativos aos testes realizados para a função aditiva e função de ativação linear.

Tabela 5. 5 - Valores de coeficiente de determinação em função da arquitetura para função linear na camada oculta e camada de saída.

R^2				
<i>Arquitetura</i>	<i>Treinamento</i>	<i>Validação</i>	<i>Teste</i>	<i>Total</i>
5	0,86178	0,93743	0,90163	0,8803
10	0,87572	0,88699	0,90048	0,88017
15	0,86327	0,93065	0,90046	0,88038
20	0,87081	0,89885	0,93691	0,88003

Os valores da Tabela 5.5 refletem o já esperado. Dentre as funções de ativação, a que possui menor capacidade de restrição dos valores de saída, pois esta funciona adequadamente em classificação de padrões, e o presente trabalho se baseia na criação de modelos de predição.

A Tabela 5.6 apresenta os valores concernentes aos testes para a função aditiva linear e função de ativação tangente hiperbólica.

Tabela 5.6 - Valores de coeficiente de determinação em função da arquitetura para função função aditiva linear e função de ativação tangente hiperbólica.

R^2				
<i>Arquitetura</i>	<i>Treinamento</i>	<i>Validação</i>	<i>Teste</i>	<i>Total</i>
5	0,96956	0,97959	0,98767	0,97448
10	0,97075	0,98777	0,97843	0,97466
15	0,96862	0,97821	0,99254	0,97399
20	0,97269	0,97989	0,9748	0,97286

A Tabela 5.6 demonstra que há uma melhoria nos valores de R^2 mediante a combinação de funções de ativação linear e tangente hiperbólica. A Tabela 5.7 apresenta os valores para os testes relativos a função de ativação linear e junção aditiva sigmoidal.

Tabela 5.7 - Valores de coeficiente de determinação em função da arquitetura para função aditiva sigmoidal e função de ativação linear.

R^2				
<i>Arquitetura</i>	<i>Treinamento</i>	<i>Validação</i>	<i>Teste</i>	<i>Total</i>
5	0,99664	0,9924	0,98869	0,99464
10	0,99416	0,97606	0,97325	0,98853
15	0,98202	0,97068	0,9824	0,97891
20	0,99861	0,99091	0,98576	0,99465

A Tabela 5.7 apresenta resultados favoráveis de modo que o modelo está bem ajustado em todas as fases, entretanto, nota-se que os resultados correspondentes ao total das etapas de ajuste do modelo ainda apresentam oscilações nítidas. A Tabela 5.8 apresenta os valores associados aos testes com função aditiva sigmoidal e função de ativação tangente hiperbólica.

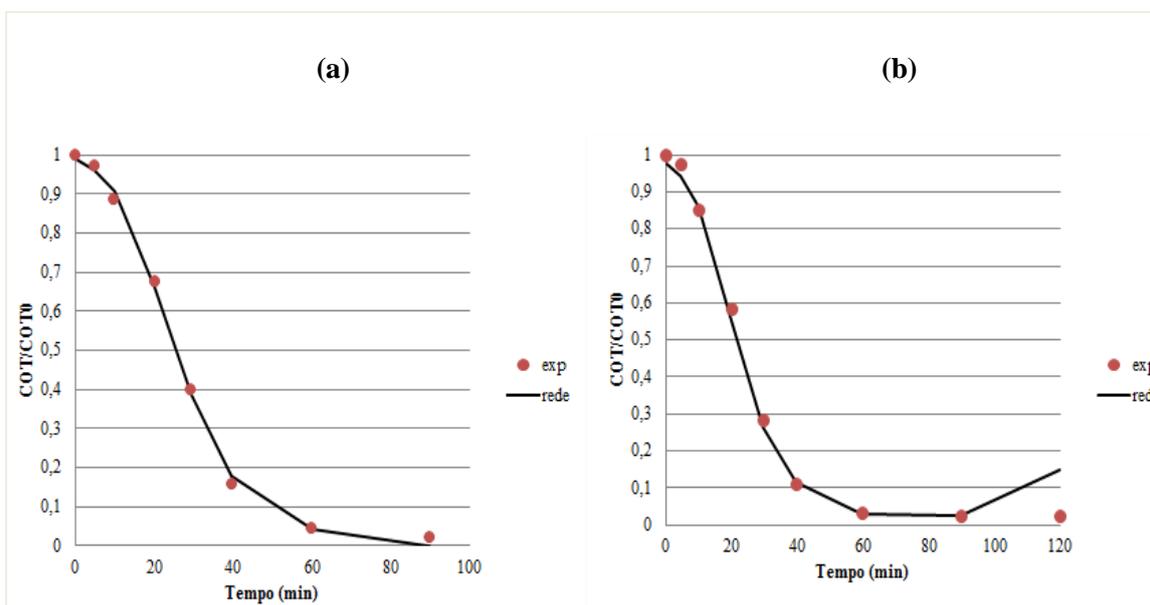
Tabela 5.8 - Valores de coeficiente de determinação em função da arquitetura para função junção aditiva sigmoideal e função de ativação tangente hiperbólica.

R^2				
<i>Arquitetura</i>	<i>Treinamento</i>	<i>Validação</i>	<i>Teste</i>	<i>Total</i>
5	0,99702	0,99786	0,99846	0,99739
10	0,99917	0,99903	0,99825	0,99891
15	0,99931	0,99582	0,99913	0,99881
20	0,99934	0,99597	0,99544	0,99821

Os resultados dos testes delineados na Tabela 5.8 demonstram resultados exímios, os quais se assemelham com os da Tabela 5.1, para testes com funções tangente hiperbólicas em ambas as camadas, tornando expresso que o emprego de tal combinação de funções resultará num adequado ajuste para o modelo.

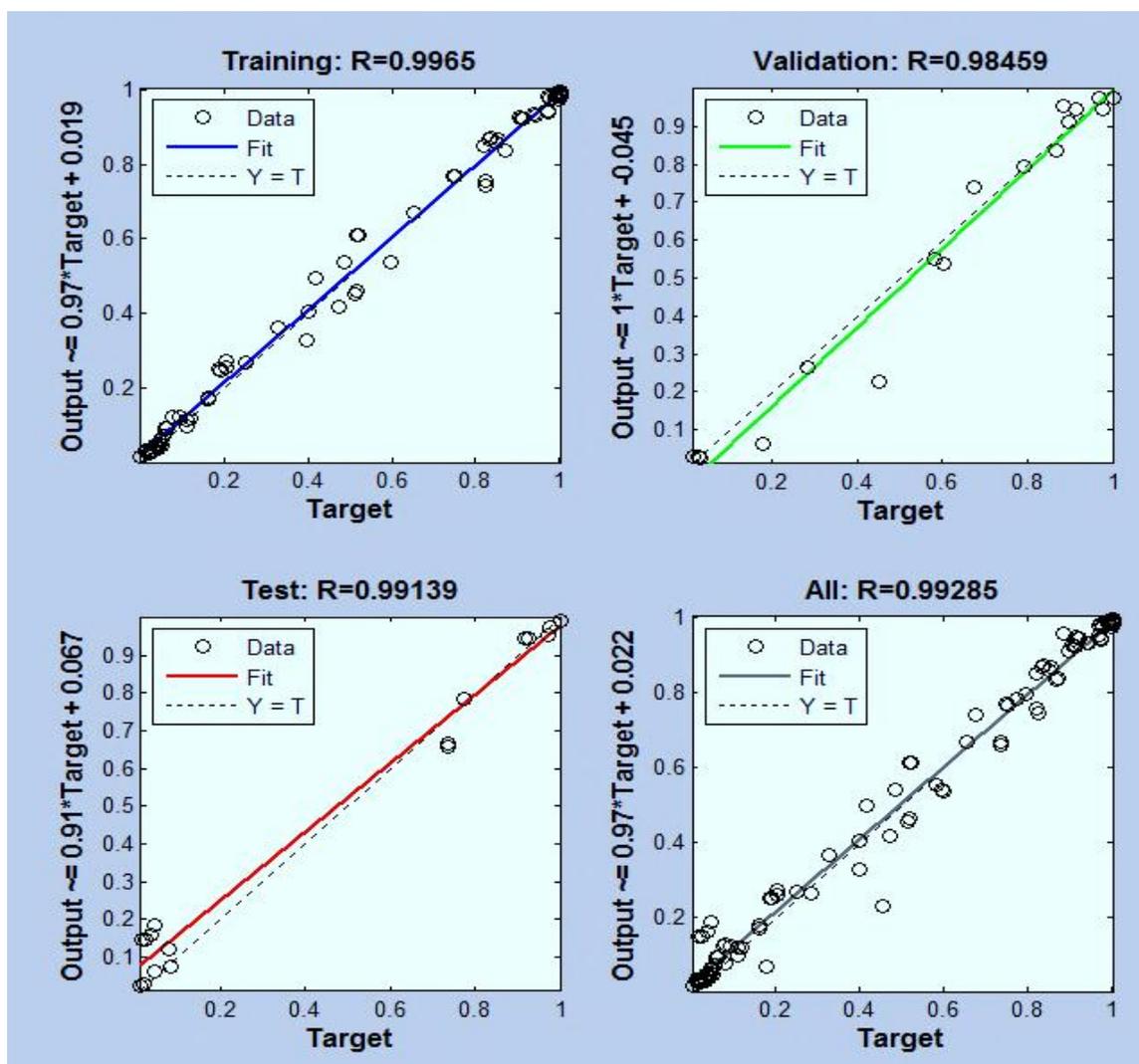
Perante os resultados supracitados, é possível ratificar que o melhor ajuste está estabelecido com a aplicação das funções de ativação e junção aditiva tangente hiperbólica em ambas as camadas (camada de entrada e camada oculta), a partir dos quais resultaram a geração dos gráficos apresentados na Figura 5.4 a e b respectivamente, que apresenta a comparação entre os resultados experimentais e os resultados calculados para os experimentos para 80 mM de peróxido de hidrogênio; 25 g/L de cloreto de sódio e 0,2 mM de Ferro II, bem como com 80 mM de peróxido de hidrogênio; 50 g/L de cloreto de sódio e 0,2 mM de Ferro II. Verifica-se um ajuste ideal, onde a curva correspondente à rede está sempre próxima aos dados experimentais. O fato de a tangente hiperbólica ter apresentado um ajuste adequado ao modelo deve-se ao fato de a inclinação da curva ser maior do que a curva sigmoideal, por exemplo, apresentado uma resposta maior para desvios pequenos. Outro feito se deve a tangente hiperbólica apresentar signos diferentes para as entradas e respostas, enquanto que as outras funções utilizadas para nossos testes apresentam funções positivas para todos os valores de entrada, esta característica de função tangente hiperbólica permite que a rede já comece treinada para os casos nominais e necessite ser treinada apenas para os casos de desvios.

Figura 5.5 - Gráfico de resultados experimentais versus calculados em função do tempo. (a) 80 mM de peróxido de hidrogênio; 25 g/L de cloreto de sódio e 0,2 mM de Ferro II. (b) 80 mM de peróxido de hidrogênio; 50 g/L de cloreto de sódio e 0,2 mM de Ferro II



A Figura 5.5 contém a representação gráfica da rede com topologia de 10 neurônios e funções de ativação e função aditiva tangente hiperbólica, onde se percebe o coeficiente de regressão linear para tal teste, em cada fase de execução da rede.

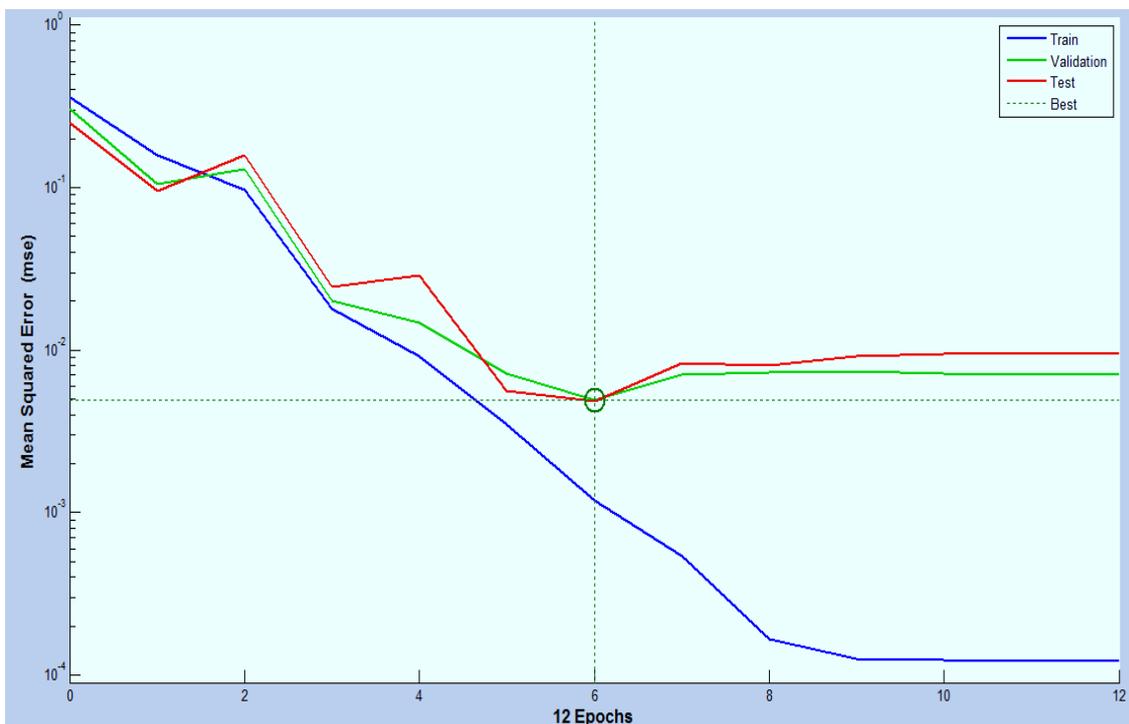
Figura 5.6 - Comportamento da rede com 10 neurônios e funções de ativação e junção aditiva tangente hiperbólica.



É, portanto, constatado a partir da linha de tendência e também da linha que representa o ajuste $Y=T$ que as respostas da rede são, cada vez mais, semelhantes aos dados experimentais, o ajuste do modelo em coeficientes que variam de 98% a 99% conforme exposto na Tabela 5.1.

A Figura 5.6 ilustra o desempenho da rede, exprimindo que a rede obteve melhor ajuste ao executar 12 ciclos. Isto significa que a rede pôde repetir em 12 épocas o processo de transmissão para frente (*feedforward*) e para trás (*backforward*) até obter um erro mínimo, porém, na época 6 seu desempenho em validação do modelo foi ultimada.

Figura 5.7 - Performance da rede, com melhor validação na época 6.



Para a variação do número de dados foi engendrada uma tabela que constitui, no modelo de melhor ajuste, com função logística (tangente hiperbólica) e 10 neurônios na camada oculta, a variação do conjunto de dados entre 50, 26 e 17 dados que demonstram através de seus coeficientes de correlação que o ajuste da rede se manteve primoroso, admitindo que mesmo havendo a necessidade de se realizar um número de experimentos reduzido, o modelo responderá corretamente.

Tabela 5.9 - Valores referentes a variação do número de dados

Arquitetura	Função de ativação	Função aditiva	Número de dados na entrada	Coeficiente de correlação			
				Treinamento	Validação	Teste	Total
10	Tangente Hiperbólica	Tangente Hiperbólica	50	0,99891	0,99814	0,99835	0,9865
10	Tangente Hiperbólica	Tangente Hiperbólica	26	0,99969	0,99951	0,99969	0,99961
10	Tangente Hiperbólica	Tangente Hiperbólica	17	0,99985	0,97163	0,99744	0,99829

Para a verificação do ajuste da rede foi realizado um teste de incerteza, com uma série de 20 dados contaminados, que foram usados para as etapas de treinamento, validação e teste subdivididos em parcelas como outrora descritos na Metodologia. A partir deste teste, foi verificado que o ajuste da rede se manteve mesmo diante de tais contaminações, o que fundamenta sobre a capacidade de generalização do modelo. Porém, para a etapa de validação do modelo, ocorreu uma queda no coeficiente de correlação, manifestando que a contaminação dos dados interferiu em tal etapa. A Figura 5.7 apresenta os gráficos do teste de incerteza, cujos valores de coeficiente de regressão linear compõem a idéia original de que a rede neural deve ser capaz de aprender e, então, generalizar o melhor possível, ainda que exposta a ruídos. A Figura 5.8 ilustra o desempenho da rede, alcançou ajuste com primor na época 4, com um erro médio de 0,01.

Figura 5.8 - Comportamento da rede com dados contaminados com 1% de erro.

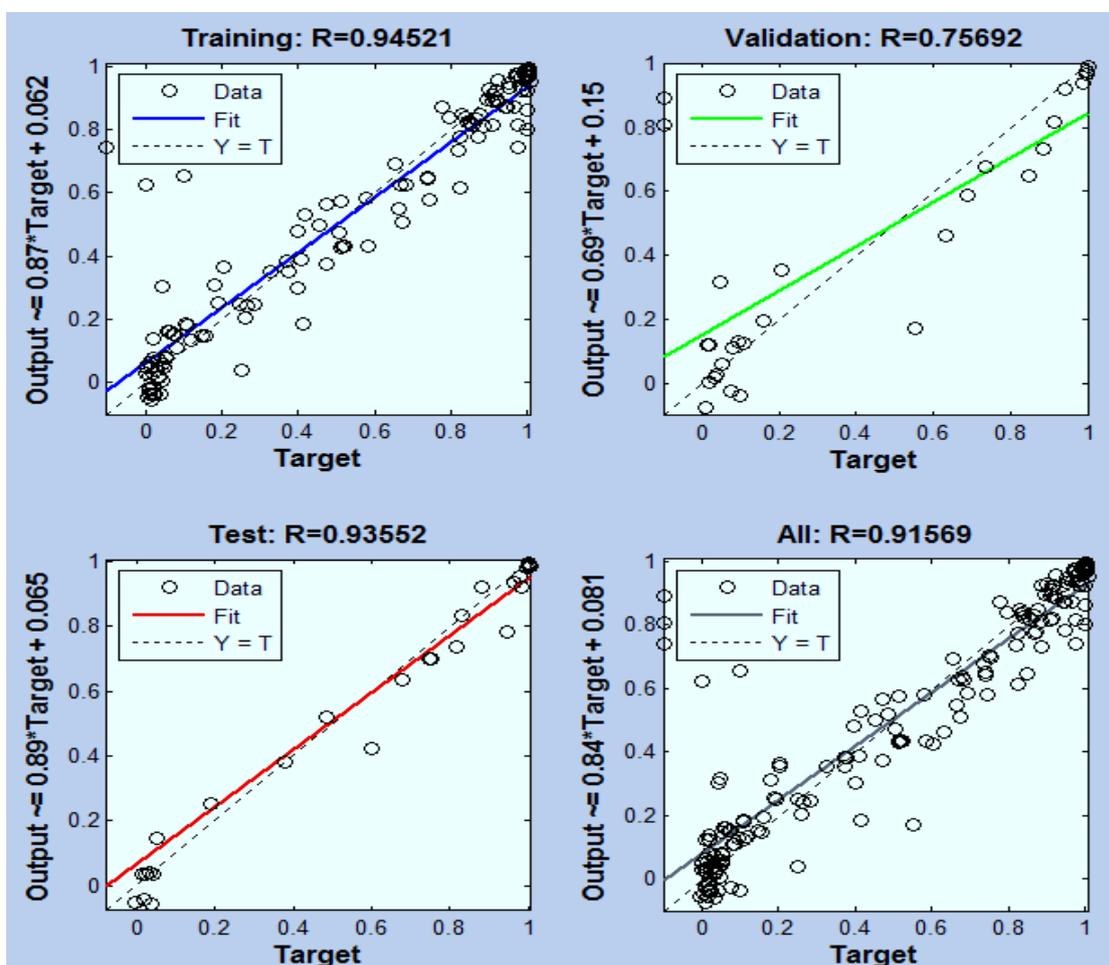
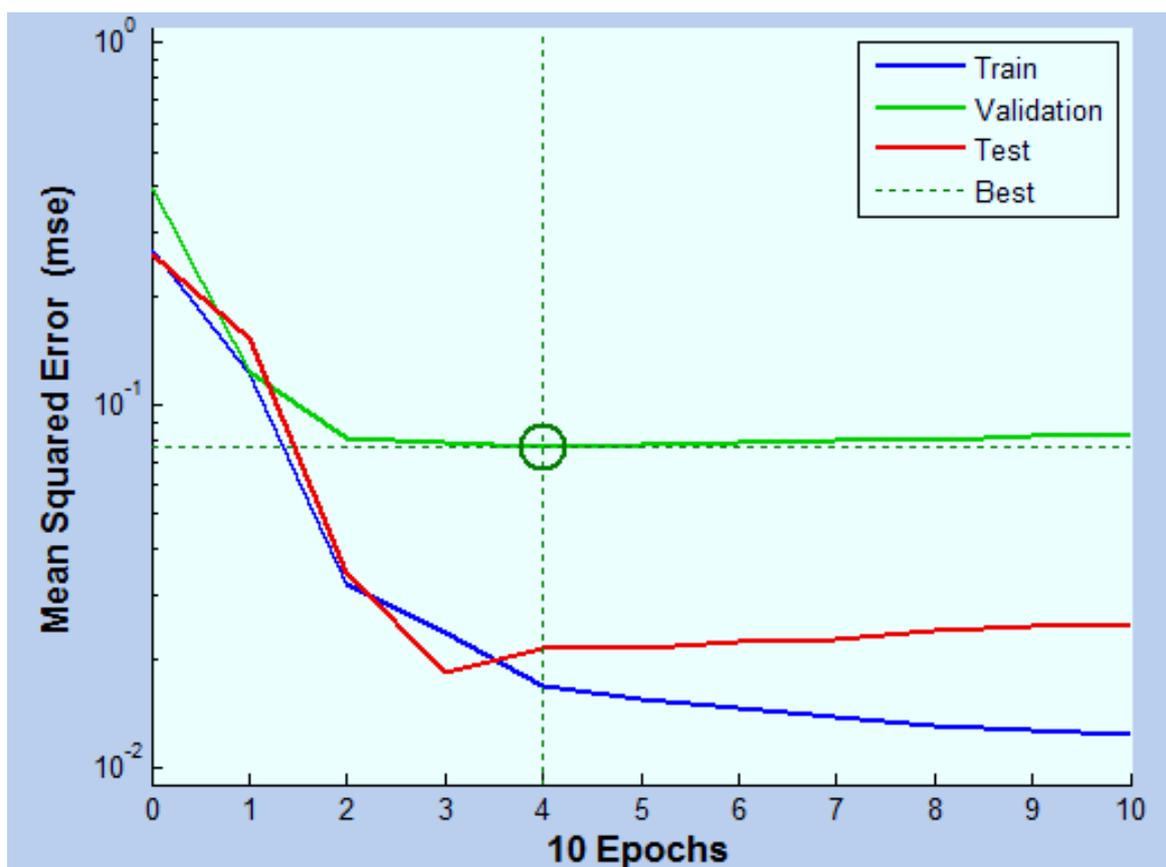


Figura 5.9 - Performance da rede com dados contaminados.

6.0 CONCLUSÕES

Com base nos resultados, pode ser concluído que:

O modelo de Redes Neurais Artificiais é um excelente método de modelagem para processos de tratamento com reações complexas, verificando-se a possibilidade de ajustes que condizem com o experimento real.

O número de neurônios interfere no desempenho da rede de modo que na camada oculta um número de 10 neurônios foi concebido como uma quantia apropriada para a garantia de um bom ajuste.

As funções de transferência empregada para um modelo de representação de um cenário real são as funções logísticas, já que sua curva é capaz de representar melhor respostas com variações tênues.

O número de experimentos não interferiu no desempenho da rede, o que constitui vantagem, pois a diminuição do conjunto de dados implica na redução de custos na etapa experimental.

A contaminação de dados influiu na etapa de validação do modelo, o que sugere que ruídos externos interferem em 1% da atuação da rede.

REFERÊNCIAS

- ANDERSON, J.A. *An Introduction to Neural Networks*. Bradford Book, 1995.
- ANDREOZZI, R.; CAPRIO, V; MAROTTA, R.A.; VOGNA, D. *Paracetamol oxidation from aqueous solutions by means of ozonation and H₂O₂/UV system*. *Water Research*, v. 37, p. 993-1004, 2003.
- BARONE, D.A.C. *Sociedades Artificiais: a nova fronteira da inteligencia nas máquinas*, Porto Alegre: Bookman, 2003.
- BARRETO, J. M. *Redes Neurais Artificiais. Texto de Curso apresentado no II-CIPEEL, Universidade Federal de Santa Catarina*, Dept. de Engenharia Elétrica, novembro 1996.
- BARRETO, J. M. *Inteligência Artificial no Limiar do Século XXI*, 3 ed. Edições Florianópolis, 2002.
- BAUGHMAN D.R. e LIU Y.A. *Neural networks in bioprocessing and chemical engineering*. Academic press limited, California, 1995.
- BORGES, F.J. *Integração dos processos de Eletrodiálise e de Degradação Fotoquímica para o Tratamento de Efluentes Salinos contendo Fenol*. Tese de Doutorado. São Paulo, 2009.
- BRAGA, A.P; CARVALHO, FERREIRA A. P. L.; LUDERMIR, T.B., *Redes Neurais Artificiais: Teoria e Aplicações*. Rio de Janeiro: LTC, 2000.
- BURROWS, T. e NIRAJAN, M. *The use of feed-forward and recurrent networks for system identification*. Technical Report CUED F-INFENG TR158, Cambridge University Engineering Dept., Cambridge U., Cambridge, UK, Dec.1993.
- CASEY, M. *Computation dynamics in discrete-time recurrent neural networks. In Proceedings of the Annual Research Symposium*. Institute for Neural Computation. La Jolla, CA, 1993.
- CLEEREMANS, A., SERVAN-SCHREIBER, D., e MCCLELLAND, J. *Finite state automata and simple recurrent neural networks*. Neural Computation, 1989.
- ESPLUGAS, S., GIMENEZ, J., CONTRERAS, S., PASCUAL, E., RODRIGUEZ, M. *Comparison of diferente advanced oxidatin process forphenol degradation*. *Water Research*, v.36, p. 1034-1042, 2002.
- GALINDO, C.; JACQUES, P.; KALT, A. *Photochemical and photocatalytic degradation of an indigolid dye: a case study of acid blue 74 (AB74)*. *Journal of Photochemistry and Photobiology A: Chemistry*. v. 141, p. 47-56, 2001.
- GALLANT, S. I. *Neural network learning and expert systems*. MIT press, Cambridge, Massachussets, 1993.

- HAYKIN, S. *Redes neurais: princípios e prática*. Trad. Paulo Martins Engel. 2.ed. Porto Alegre:Bookman., 1994.
- HEBB, D.O. *Brain Mechanisms and Learning*. J. F..London: Delafresnaye, 1961.
- HERTZ, J.; PALMER, RG; KROGH, AS. *Introdução à teoria da computação neural*. Perseus Books, 1990.
- LEGRINI, O., OLIVEIRO E., BRAUN, A.M. *Photochemical process for water treatment Chemical. Review*, 1993.
- LIN, S. H., CHANG, C. C. *Treatment of landfill leachate by combined electro-Fenton oxidation and sequencing batch reactor method*. Water Research. V. 34, n. 17. 2000.
- LUDWIG Jr., O. e COSTA, E.M.M. *Redes Neurais: Fundamentos e Aplicações com Programas em C*. Rio de Janeiro :Ciência Moderna, 2007.
- MÁSSON, E; WANG, Y.J. *Introduction to Computation and Learning in Artificial Neural Networks. European Journal of Operational Research*, North-Holand. 1990.
- MINSKY, M. L. e PAPERT, S. A. *Perceptrons: an introduction to computational geometry*.The MIT Press, Massachussets, 1969.
- NICHOLLS, J. G. et a.l *From Neuron to Brain*. Sinauer Associates, Inc., 2000.
- OLLIS, D. *Integrating Photocatalysis and Membrane Technologies for water treatment. Annals*. New York Academy of Sciences, 2003.
- PACHECO J. R., *Estudo de certas potencialidades de Processos Oxidativos Avançado para o tratamento de percolado de aterro sanitário*. Dissertação de mestrado, Universidade Federal do Paraná. 2004.
- REZENDE, S.O. *Sistemas Inteligentes:fundamentos e aplicações*. Barueri, 2003.
- RUMELHART, D. E.; MCCLELLAND, J. L. e GROUP, P. *Parallel Distributed Processing*. vol 1: Foundations. The MIT Press, 1986.
- SHU, H.Y.; CHANG; M.C.; 2005. *Decolorization effectes of six azo dyes by O₃; UV/O₃ and UV/H₂O₂ processes*. Dyes and Pigments. 2003.
- TIRZHÁ L. P. D. *Decomposição de Peróxido de Hidrogênio em um Catalisador Híbrido e Oxidação Avançada de Efluente Têxtil por Reagente Fenton modificado*. Dissertação de Pós-graduação, Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina, 2005.
- WASSERMAN, P. D. *Neural Computing : Theory and Practice*. New York : Van Nostrand Reinhold, 1989.
- WERBOS, P. J. *Beyond regression: new tools for prediction and analysis in the behavioral sciences*. PhD thesis, Harvard University, Cambridge, MA, 1974.