MAIKO FERNANDES BUZZI

UMA NOVA PROPOSTA PARA O TREINAMENTO NÃO SUPERVISIONADO EM REDES NEURAIS DE BASE RADIAL PARA PREVISÃO DE SÉRIES TEMPORAIS

CURITIBA
2012

MAIKO FERNANDES BUZZI

UMA NOVA PROPOSTA PARA O TREINAMENTO NÃO SUPERVISIONADO EM REDES NEURAIS DE BASE RADIAL PARA PREVISÃO DE SÉRIES TEMPORAIS

Tese apresentada como requisito parcial à obtenção de grau de Doutor em Ciências, Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Setor de Ciências Exatas e de Tecnologia, Universidade Federal do Paraná.

Orientadora: Prof.ª Dr.ª Andrea Sell Dyminski.
Co-Orientador: Prof. Dr. Eduardo Parente Ribeiro.

CURITIBA

2012

TERMO DE APROVAÇÃO

Maiko Fernandes Buzzi

Uma Nova Proposta para o Treinamento não Supervisionado em Redes Neurais de Base Radial para Previsão de Séries Temporais

Tese aprovada como requisito parcial para obtenção do grau de Doutor no Programa de Pós-Graduação em Métodos Numéricos em Engenharia – Área de Concentração em Programação Matemática, Setores de Tecnologia e de Ciências Exatas da Universidade Federal do Paraná, pela seguinte banca examinadora:

Orientadora: Prof. Andrea Sell Dyminski, D.Sc.
Programa de Pós-Graduação em Métodos Numéricos em Engenharia
PPGMNE da UFPR

Prof. Celso Romanel, D.Sc.
Pontifícia Universidade Católica do Rio de Janeiro – PUC-RIO

Profª. Angela Olandoski Barboza, D.Sc.
Universidade Tecnológica Federal do Paraná - UTFPR

Prof. Eduardo Parente Ribeiro, D.Sc.
Departamento de Engenharia Elétrica – UFPR

Prof. Jair Mendes Marques, D.Sc
Programa de Pós-Graduação em Métodos Numéricos em Engenharia
PPGMNE da UFPR

Prof. Paulo Henrique Siqueira, D.Sc.
Programa de Pós-Graduação em Métodos Numéricos em Engenharia
PPGMNE da UFPR

Curitiba, 05 de novembro de 2012.

AGRADECIMENTOS

A Deus, por permitir que eu chegasse até aqui.

A minha orientadora, Andréa Sell Dyminski, pela orientação, pelo incentivo constante, pela dedicação, pelos conhecimentos transmitidos durante esse curso de doutorado, pela confiança depositada em meu trabalho e principalmente pela oportunidade de estudar e desenvolver esta tese.

Ao Departamento de segurança estrutural da TRANSPETRO, pelo apoio incondicional e por permitir total acesso aos dados instrumentais dos taludes OLAPA e OSPAR apresentados neste trabalho.

Aos funcionários do CESEC, principalmente à Maristela, dona de uma disposição e simpatia singular.

Aos colegas de curso da UFPR, com os quais compartilhei inesquecíveis experiências acadêmicas e de vida.

A CAPES e a PETROBRAS, pelos auxílios concedidos, fundamentais para que este trabalho pudesse ser desenvolvido.

RESUMO

Análise de séries temporais possui frutíferas aplicações em diversas áreas, com destaque à previsão de demanda, processamento de sinais, estimativa de desempenho em obras de engenharia, entre outros. Acerca das metodologias empregadas a esta referida análise, destacam-se as redes neurais de função base radial (redes RBF – *Radial Basis Function*). Entre os entraves limitantes à aplicação das redes RBF à modelagem de séries temporais, ressalta-se a determinação do número e posicionamento adequado dos neurônios na camada intermediária, que dependendo da dimensão dos dados, demandam elevado tempo de processamento na busca por uma topologia ideal. Para tratar este problema, o presente estudo propõe, em detrimento do uso de técnicas de agrupamento não hierárquico, o emprego de um método de agrupamento hierárquico aplicado aos dados de entrada da rede RBF, para determinar o número de neurônios e seus respectivos posicionamentos. O método proposto foi submetido a séries temporais conhecidas na literatura e em séries de leituras de sistemas de instrumentação de duas grandes obras, a Usina Hidrelétrica de ITAIPU e Dutos enterrados da TRANSPETRO.

Experimentos computacionais apontam para a redução significativa do número de topologias a serem testadas, produzindo desempenho igual ou superior a outras abordagens existentes na literatura, com significativa redução no tempo de processamento, além da abordagem do método em dois problemas reais, onde a previsão pode contribuir com o avanço nas práticas que visam a segurança dessas importantes obras.

Palavras chaves: Séries Temporais, Redes Neurais de Base Radial (RBF), Análise Multivariada, Segurança de Barragens, Segurança de Dutos, Instrumentação.

ABSTRACT

Temporal series analysis has fruitful applications in several areas, with emphasis on demand forecasting, signal processing, performance estimation in engineering works, among others. About the methodologies used to this that review, stand out the neural networks of radial basis function (RBF networks - Radial Basis Function). Among the obstacles that limit the application of RBF networks to the modeling of time series barriers, underscores the determination of the number of neurons and proper positioning in the middle tier, which depending on the size of the data, require high processing time in the search for an ideal optimal topology. To address this problem, this study, instead of using the techniques of nonhierarchical clustering, proposes the use of a hierarchical clustering method applied to the input data of RBF network to determine the number of neurons and their respective positions. The proposed method was subjected to temporal series known in the literature and readings of instrumentation systems series two great works, the Itaipu Dam Hydroelectric and Transpetro's buried Pipelines.

Computational experiments point to a significant reduction in the number of topologies to be tested, producing performance equal or superior to other approaches in the literature, with a significant reduction in processing time, beyond the approach of the method in two real problems, where the prevision can contribute with the advancement in practices aimed at safety of these important work.

.

Keywords: Neural Network, Radial Basis Neural Networks (RBF), Multivariate Analysis, Dam Safety, Pipeline Safety, Instrumentation.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE SIGLAS, SÍMBOLOS E ABREVIATURAS

ANCOLD - Australian Comittee on Large Dams

AG – Algoritmo Genético

ARIMA – Auto Regressive Integrated Moving Average / Modelo Auto-Regressivo Integrado de Média Móvel

AS – Similated Annealing

DGP – Data GererationProcess / Processo Gerador do Dados

EM – Expectation Maximization

FIR – Finite Impulse Response / Resposta de Finitos Impulsos

GRBF– Redes Neurais Artificiais com Função de Crescimento de Base Radial

HRDGA– Ranking Hierárquico de Densidade do Algoritmo Genético

IA – Inteligência Artificial

ICOLD – International Comission on Large Dams

IBGE – Instituto Brasileiro de Geografia e Estatística

ILP – Inductive Logic Programming / Programação LógicaIndutiva

JIT – Just-in-Time

LIFO – Last in First Out / Último que Entra é o Primeiro que Sai

MATLAB– Matrix Laboratory – Software para Cálculo Numérico

MDL – Modelos Dinâmicos Lineares

MEST – Modelos Estruturais de Série de Tempo

MLP –Multilayer Perceptron

MQO– Método de Mínimos Quadrados Ordinários

MRE – Mean Relative Error / Erro Relativo Médio

MSE – MeanS quared Error / Erro Quadrático Médio

NMSE – Normalized Mean Square Error / Erro Médio Quadrático Normalizado

PDP –Parallel Distributed Processing / Distribuição de Processamento Paralelo

PGBF – Função Pseudo-Gaussiana de Base Radial

RBF– Redes Neurais Artificiais de Funções de Bases Radiais

RMSE – Root Mean SquaredError / Raiz do Erro Quadrático Médio

RNA – Redes Neurais Artificiais

SDR – Support Vector MachineRegression

SOM – Self Organizing Maps / Maps Auto Organizáveis

SSE– Some of SquaredError / Soma dos Erros Quadrático

SVD– Decomposição em Valores Singulares

TI – Tecnologia da Informação

TS – Tabu Search

SUMÁRIO

1 INTRODUÇÃO

Para se avaliar a segurança de uma estrutura, o princípio básico é ter em mãos informações relevantes sobre seu projeto e comportamento ao longo do tempo, e com elas procurar entender se essa estrutura encontra-se estável ou não.

Por isso, o monitoramento de grandes estruturas feita a partir de um sistema de instrumentação é uma das principais ferramentas na avaliação de suas condições de segurança, acompanhando diferentes características do sítio e tendo finalidades diversas, dependendo da etapa da obra que se deseja analisar. Ao longo da vida útil da estrutura, o monitoramento pode detectar variações nas condições de segurança, como resultado de processos de seu envelhecimento e alterações ambientais. O conhecimento do nível de segurança é importante para a definição do conjunto de ações que devem ser tomadas caso ocorra alguma alteração significativa neste nível de segurança (SARÉ, *et al*. 2006). Quando se diz sistema de instrumentação, trata-se de um conjunto de instrumentos instalados nessas obras, que medem grandezas de engenharia de interesse para a análise da segurança da estrutura, tais como, por exemplo, precipitação, poro-pressões e deslocamentos.

Além das informações atuais geradas por esses instrumentos, que podem ser utilizadas na análise da segurança de uma estrutura, podem-se usar informações do passado para se obter uma projeção dos valores de controle da instrumentação, o que pode ser fundamental para que se tome ações preventivas em tempo hábil. Nem sempre esta avaliação é trivial e depende da estimação de variáveis muitas vezes complexas. Nessa área de previsão de séries temporais, as redes neurais vêm sendo utilizadas como uma alternativa promissora.

Uma rede neural pode fazer uma previsão muitas vezes bastante adequada destes valores, uma vez que pode ser considerada como um aproximador universal e possui capacidade de auto aprendizagem. O mapeamento completo dos pares de entradas e saídas de um determinado problema pode ser construído por exemplos, o que reduz a complexidade em torno da seleção do modelo de previsão ideal a ser aplicado. A rede neural de função de base radial (RBF) é um bom candidato entre as redes neurais para a previsão, que devido à sua capacidade de aprendizagem rápida. Tem sido aplicada com êxito para a modelagem não linear de séries temporais e aplicações de previsão (CHNG, C. *et al*, 1996), (LEUNG, L. *et al*, 2001),
(LI, Y. *et al*, 2004), (WANG *et al*., 2009) e (Lee, 2010).

Normalmente em redes RBF são utilizadas funções de ativação Gaussianas na camada oculta, cujos parâmetros (centros e largura do campo receptivo) são ajustados a partir dos dados de entrada. O bom desempenho da rede RBF depende em grande parte da qualidade desse ajuste. Com relação às diferentes estratégias para a determinação dos parâmetros da rede, (CHEN, 1991) propõe a seleção aleatória, onde os centros são vetores de entrada aleatoriamente selecionados. Esta técnica demanda que os padrões de treinamento representem com acurácia todo o espaço de soluções do problema. Ainda que o método seja simples e direto, ele pode exigir grande número de unidades (neurônios) intermediárias, o que em geral implica em centros de coordenadas muito próximas umas das outras, comprometendo o funcionamento adequado da rede. Já (BISHOP 1996) propõe uma fixação dos centros em grade regular, cobrindo todo o espaço de entrada. Em geral, esta proposta exige muitas unidades intermediárias para vetores de entrada com elevada dimensão, podendo gerar um crescimento exponencial do número das unidades escondidas.

Técnicas de clusterização/agrupamento para a definição dos centros foram propostas inicialmente por (MOODY *et al*, 1999), entre as quais se destacam o algoritmo de *k*-médias e os mapas auto-organizáveis. Técnicas de clusterização/agrupamento permitem uma configuração mais adequada dos centros e, com isso, um desempenho superior em relação ao das abordagens comentadas anteriormente. Todavia, tais técnicas podem demandar um elevado custo de processamento dependendo da dimensão do problema, agravando-se ainda mais quando se executam vários treinamentos na busca de uma topologia mais adequada ao problema em questão.

Arthur e Sergei, 2007, apresentam uma proposta de melhoria para o método *k*-médias, denominada *k*-médias++. Tal melhoria reside na redução da dependência do método em relação às sementes (ou centros) iniciais escolhidas aleatoriamente. (LEE *et. al*, 2007) e (HOWARD *et. al*, 2009), empregam a técnica com bons níveis de separação dos grupos de dados. Entretanto, (PETERSON *et. al*, 2010), comparando 11 métodos de inicialização para o *k*-médias, apontam para um desempenho inferior do *k*-médias ++ em relação a outras estratégias.

Outra questão fundamental para a composição da topologia das redes RBF é a dificuldade de se determinar o número de nós ou neurônios na camada oculta.

Uma quantidade menor de neurônios do que a necessária limita o desempenho da rede, ao passo que a superestimação de neurônios demanda elevado

tempo de processamento na fase de treinamento para o ajuste dos parâmetros da rede, o que ainda não garante um bom desempenho da mesma.

Verifica-se na literatura uma vasta gama de métodos para determinar o número de neurônios na camada oculta em redes RBF (Chen, W, *et al*, 1996), (COWAN e GRANT, 1991) , (CHEN, W., *et al*, 1999), e (CHEN, W., 2006), onde nesses, destacam-se o emprego de mínimos quadrados ortogonais (OLS – *Ortoghonal Least Square*) na escolha do número de neurônios da rede RBF. (WANG *et al*., 2005), e (HUANG, S. *et al.*, 2004) propuseram um conceito da significância de um neurônio, definido como a contribuição estatística de um neurônio para o desempenho global da rede. Tal conceito é utilizado para a determinação do número de neurônios na composição da topologia da rede. Um novo neurônio é adicionado se a sua contribuição é maior do que um limiar escolhido. Inversamente, se o significado de um neurônio é inferior ao limite, o neurônio é eliminado. (LEE, 2010) adotou o conceito de (WANG *et al*., 2005), através do emprego de uma expressão heurística denomina pelo autor de M-estimador, como a contribuição estatística. Apesar de se evidenciar que já existem critérios propostos de inclusão de um novo neurônio à rede, cabe destacar que todos os métodos supracitados dependem do treinamento da rede para a conclusão desejada.

A partir da problemática apresentada, este estudo apresenta uma nova abordagem para a determinação do número de neurônios na camada escondida em redes RBF. O método proposto reside na substituição das técnicas de agrupamentos não hierárquicas pelas técnicas de agrupamento hierárquicas na etapa não supervisionada do treinamento da rede, reduzindo assim sua complexidade. Ademais, utilizam-se as informações geradas por essas técnicas de agrupamento na escolha de topologias mais adequadas ao problema de estudo em questão, reduzindo a quantidade de topologias a serem testadas. Tais inovações ensejam tempo de processamento computacional substancialmente menor em comparação com outras propostas existentes, sem comprometer o desempenho da rede quanto à qualidade da previsão.

O método proposto será testado em instâncias conhecidas na literatura e em séries temporais dos dados gerados pelos sistemas de monitoramento de importantes obras de engenharia: dutovias gerenciadas pela TRANSPETRO e a barragem de ITAIPU. Com o método atingindo o objetivo proposto, procura-se contribuir com uma

nova ferramenta que pode ser utilizada, oferecendo previsões mais adequadas e rápidas de grandezas de engenharia presentes nessas obras.

1.1 OBJETIVOS DO TRABALHO

1.1.1 Objetivo Geral

O objetivo principal desse trabalho é propor uma nova estratégia de aprendizado não supervisionado em redes neurais de base radial, tornando-a mais rápida e precisa, e aplicar esta nova abordagem à previsão de leituras de instrumentação de grandes obras.

Verificar se o método proposto possui as seguintes características em relação às redes neurais de base radial tradicional:

- processo de treinamento da rede mais eficiente/rápido computacionalmente;
- Mapear o espaço de entrada de maneira mais adequada e, com isso, melhorar a *performance* da rede;

A partir de métricas/informações obtidas na etapa não supervisionada, propor uma heurística que determine o número de neurônios a ser utilizado na camada oculta que resulte no melhor desempenho de generalização, sem a necessidade de testar todas as topologias possíveis.

1.1.2 Objetivos Específicos

No caso da previsão, serão utilizadas redes neurais de base radial, testando várias topologias de maneira automática, e para validação e/ou comparação do método proposto, confrontar com as soluções obtidas com os métodos de estimação tradicionais.

No caso das redes neurais de base radial, propor um novo método na etapa do treinamento não supervisionado, utilizando métodos de agrupamento hierárquicos na determinação da localização dos centros das bases radiais, buscando melhores resultados do que os atingidos pelos métodos referenciados para esse fim.

Aplicar o novo método em séries temporais frequentemente utilizadas nas medidas de desempenho dos métodos propostos atualmente na literatura, como por exemplo, a série temporal caótica Mackey-Glass, além de séries obtidas de

bibliografia reconhecida na área de previsão de séries. Com isso, se busca comprovar a robustez do método diante diferentes séries, e sua eficiência em relação aos métodos propostos com o mesmo fim.

Testar a nova abordagem em dois casos diferentes, sendo o primeiro o caso da usina hidrelétrica de ITAIPU, onde as séries históricas das informações de controle possuem baixa variação e tendência, o que leva a incertezas menores a cerca das informações e previsões obtidas; e em segundo dois taludes da Serra do Mar do PR, onde as séries históricas de valores de controle são influenciadas abruptamente por fatores externos, alterando tendências e consequentemente a incerteza relacionada às informações de interesse desta. Nesse caso, o teste da metodologia será executado em duas encostas (o km 57 do oleoduto OLAPA e o km 56 do oleoduto OSPAR), que devido aos históricos caracterizados pela ocorrência de processos de ruptura, possuem um volume considerável de dados.

1.2 IMPORTÂNCIA DO TRABALHO

Um método de previsão de séries temporais com precisão e computacionalmente viável tem sido o objetivo principal de vários trabalhos, em especial os relacionados às técnicas de inteligência computacional, que vem trazendo ótimos resultados em comparação aos métodos mais tradicionais de previsão, fato decorrente da alta não linearidade presente na maioria dos dados envolvidos nesse tipo de problema.

E dentro desse grupo de técnicas de inteligência computacionais aplicadas à previsão, uma muito utilizada são as redes neurais de base radial, classificado como uma técnica de aprendizado não supervisionado (agrupa os dados de entrada de acordo com suas similaridades) e supervisionado (ajusta parâmetros do modelo acordo com a saída desejada).

O que se vê na literatura, na etapa não supervisionada do treinamento de redes neurais de base radial, é a utilização métodos de agrupamento, que de acordo com os dados de entrada, são agrupados de acordo com suas similaridades, e a partir desses grupos formados, determina-se a localização mais adequada aos neurônios. Nesta etapa, o método mais citado sem dúvida é o método *k-médias,* seguido de algumas variações deste com o mesmo fim. Após esta etapa, segue-se para os ajustes dos pesos na camada de saída, onde é citada a utilização do algoritmo *back-*

propagation (Técnica de ajuste de pesos das redes perceptron multi camadas) e o método dos mínimos quadrados quando a função de ativação da camada de saída for linear, ambas referenciadas adiante.

Buscando a quantidade ideal de neurônios (arquitetura ideal), é realizado o treinamento e teste de generalização para cada quantidade de neurônios definidos a priori, e o que resultar na menor métrica de erro adotada, será a rede considerada com a arquitetura ideal.

Mas todos esses testes (quantidade de neurônios na camada escondida e suas respectivas performances) tomam um considerável tempo computacional para serem feitos, principalmente nos casos onde o número de neurônios é elevado.

Seria de grande valia um método que executasse esses testes num tempo reduzido, ou ainda, na determinação do número de neurônios sem ter que testar todas as quantidades de neurônios possíveis, gerando uma precisão igual o maior do que se encontra atualmente. E é justamente isso que esse trabalho se propõe a fazer.

Um método com boa precisão e computacionalmente viável o torna aplicável a situações que outros métodos sem essas características não o são, como por exemplo, os relacionados à segurança de uma obra/estrutura de engenharia, cuja projeção do comportamento da mesma, quanto mais acurada e rápida, torna as ações mitigadoras possíveis e/ou mais eficientes.

A necessidade dessa precisão e eficiência torna-se maior na medida da importância social e econômica que o correto e ininterrupto funcionamento dessas obras representam, e nas consequências desastrosas na ocorrência de um dano incontrolável a essa estrutura, como por exemplo, dutos e barragens.

O sistema de dutos é o meio mais seguro e econômico de transporte de petróleo e seus derivados, contribuindo para aumentar a segurança nas estradas e diminuir a poluição causada pelo tráfego pesado das carretas. Portanto, investir na ampliação, modernização e na confiabilidade da malha dutoviária brasileira é fundamental para atender às necessidades e exigências cada vez maiores da população.

A produção de petróleo e o consumo de derivados estão crescendo cada vez mais e é preciso que o cuidado com o transporte desses produtos acompanhe esse crescimento. Devido à sua importância econômica e ambiental, e aos riscos envolvidos com o dano de tais dutos, as atividades de inspeção e manutenção são de grande relevância visto que a interrupção de uma linha pode ocasionar prejuízos

elevados. Em casos extremos, a ruptura total ou parcial de dutos pode levar a graves acidentes de trabalho ou ambientais.

Nesse contexto, a previsão do comportamento dessas estruturas é de extrema importância, uma vez que, se forem obtidas com certo grau de antecedência, torna possíveis ações preventivas à sua segurança. Uma grande obra como a usina hidrelétrica de ITAIPU se enquadra na mesma situação, onde danos a estrutura pode gerar graves consequências sociais e ambientais.

Para se monitorar o comportamento destas grandes estruturas, é usado um sistema de instrumentação estrutural e geotécnica, que gera grande quantidade de dados ao longo do tempo. A interpretação destes dados nem sempre é trivial e demanda grande esforço da equipe de engenharia responsável pela segurança destas obras.

A importância da melhoria das redes neurais de base radial não se limita apenas a problemas envolvendo previsão de séries temporais, que é o objetivo principal desse trabalho, e sim para todos os problemas que possam ser abordados pela técnica. Se os objetivos propostos forem atingidos, certamente será uma contribuição importante para diversas áreas.

1.3 ESTRUTURA DO TRABALHO

O trabalho está organizado conforme descrição a seguir.

No capítulo 2, faz-se a descrição dos problemas abordados. Discutem-se todas as variáveis necessárias para o desenvolvimento das soluções e delineam-se quais são os problemas referentes à análise de segurança de grandes estruturas que serão abordados. Também, neste capítulo, encontra-se a pesquisa bibliográfica realizada referente a sistemas de instrumentação de grandes obras, técnicas de previsão de séries temporais, bem como as redes neurais e análise de agrupamento, dando especial atenção aos utilizados no trabalho.

No capítulo 3 apresenta-se a metodologia proposta para o problema abordado. O capítulo 4 contém os resultados obtidos que fornecem sustentação para o método proposto, aplicando-se a nova metodologia a séries da literatura e de problemas específicos.

No capítulo 5 são apresentadas as conclusões finais do trabalho, comentando os resultados obtidos e fazendo sugestões para continuidade da pesquisa.

2 SÉRIES TEMPORAIS E REDES NEURAIS

2.1 REDES NEURAIS ARTIFICIAIS (RNA)

Técnicas de programação elaboradas, como as de Inteligência Artificial, vêm sendo empregadas na resolução de novos e antigos problemas, pois fornecem soluções que têm apresentado resultados bastante satisfatórios, seja no produto final, como na execução da tarefa, trazendo economia de tempo e recursos computacionais.

Dentre estas técnicas, as Redes Neurais Artificiais (RNA) se destacam, pois apresentam características tais como capacidade de aprendizado e generalização. São empregadas no reconhecimento de padrões, composição musical, processamento de sinais digitais e imagens, otimização, visão computacional, robótica e previsão de séries temporais.

2.1.1 Neurônio Biológico Artificial

Uma RNA é um modelo computacional que compartilha algumas das propriedades do cérebro: consiste de muitas unidades simples trabalhando em paralelo sem um controle central.

As conexões entre unidades possuem pesos numéricos que podem ser modificados pelo elemento de aprendizado (RUSSELL *et al*, 2004). A essas unidades damos o nome de neurônios e as suas conexões, sinapses.

As RNA são semelhantes ao cérebro humano em dois aspectos, basicamente (HAYKIN, 2001):

- A experiência é a fonte do conhecimento adquirido e
- O conhecimento adquirido é armazenado nas sinapses.

O neurônio, célula constituinte do cérebro, ilustrado pela figura 2.1, é composto por um corpo celular chamado “soma”, onde se encontra seu núcleo e por axônios e dendritos.

O axônio, uma fibra nervosa de superfície lisa com poucas ramificações e maior comprimento, é responsável pela transmissão na comunicação com os outros neurônios. Os dendritos, que têm aparência de árvores, possuem superfície irregular e muitas ramificações, atuam como receptores nesta comunicação.

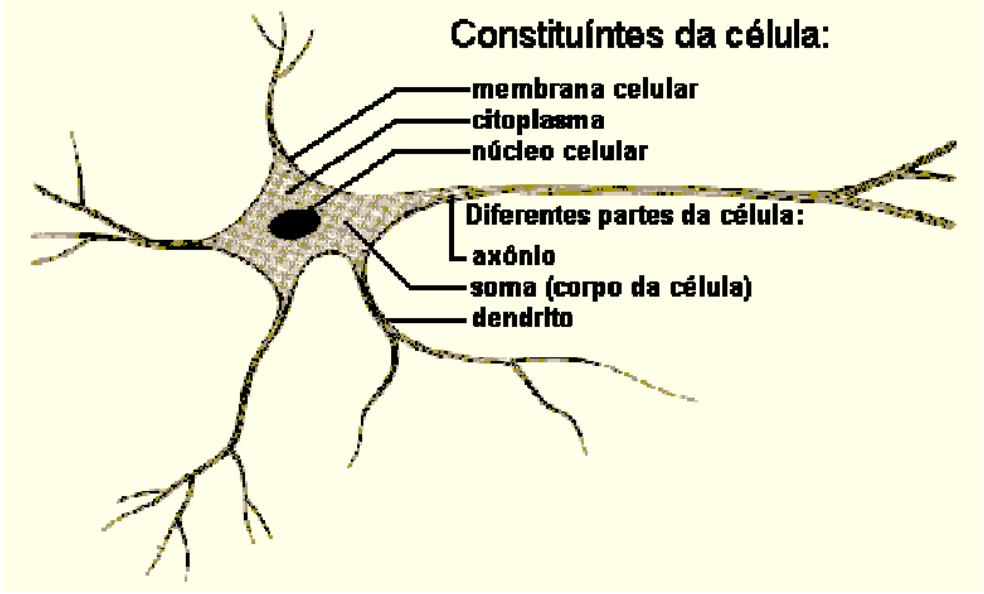


FIGURA 2.1 – NEURÔNIOBIOLÓGICO: OS CONSTITUINTES DA CÉLULA
FONTE: DIAS (2007)

Tal comunicação ou interação, sinapse, é caracterizada por um processo químico no qual são liberadas substâncias transmissoras que se difundem tal junção sináptica entre neurônios, o que causa aumento ou queda no potencial elétrico do neurônio receptor. Resumindo, uma sinapse é a conexão entre neurônios o que implica em excitação ou inibição do neurônio receptor (HAYKIN, 2001).

De forma análoga, um neurônio artificial apresenta as mesmas características. A figura 2.2 mostra um modelo onde podem ser vistos: sinapses representadas pelas entradas e pesos sinápticos, somatório e função de ativação.

Cada sinapse é caracterizada por um estímulo de entrada multiplicado pelo seu peso sináptico correspondente. Depois desta multiplicação, cada sinal de entrada é somado e o resultado é então, aplicado a uma função de ativação que restringe a saída do neurônio a um intervalo [0, 1] ou [–1, 1], dependendo da função de ativação aplicada. O neurônio artificial pode ser descrito pela equação (3.1) (HAYKIN, 2001):

O neurônio artificial pode ser descrito pela equação (2.1) (HAYKIN, 2001):

$$y_k = \varphi\left(\sum_{i=1}^{n} x_i w_{ki}\right) \qquad (2.1)$$

onde y_k é a saída do neurônio; φ é a função de ativação; x_1, x_2, ..., x_n são os sinais de entrada do neurônio; e w_{k1}, w_{k2}, ...,w_{kn}são os pesos sinápticos do neurônio em questão (neurônio k).

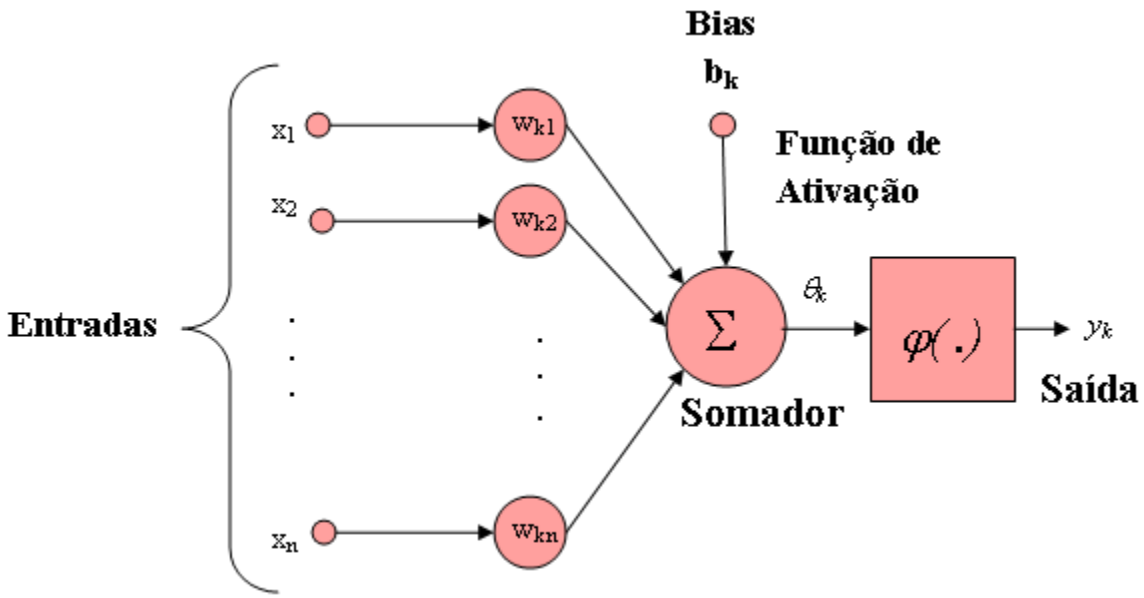


FIGURA 2.2– NEURÔNIO ARTIFICIAL.
FONTE: ADAPTADO DE HAYKIN (2001)

Portanto, o neurônio artificial imita o funcionamento do neurônio biológico por meio das entradas, pelas sinapses e pela função de ativação que simula o processo químico que libera substâncias químicas que excitarão ou inibirão os próximos neurônios.

2.1.2 Funções de Ativação

Cada neurônio realiza um processamento simples: recebe uma entrada e computa um novo nível de ativação (RUSSELL *et al*, 2004). Este processamento é composto por duas etapas: na primeira, cada entrada x_i do neurônio é multiplicada pelo peso sináptico correspondente w_{ji} (peso da entrada *i* do neurônio *j*).

O resultado de cada multiplicação é então somado. Na segunda etapa a soma é aplicada a uma função de ativação *f*, obtendo-se a saída do neurônio (*y*), conforme (2.2).

$$y = f\left(\sum x_i w_{ji}\right) \tag{2.2}$$

A função de ativação deve simular as características não lineares do neurônio biológico. As funções mais utilizadas são:

Função linear: É uma equação linear da forma (2.3):

$$f(x) = ax \tag{2.3}$$

Função degrau: É uma equação utilizada para valores binários e é da forma (2.4):

$$f(x) = \begin{cases} 1, se\ x > 0 \\ 0, se\ x \leq 0 \end{cases} \tag{2.4}$$

Função sigmóide: Também chamada de função logística, é uma função contínua que permite a transição gradual entre os dois estados. É dada por (2.5):

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2.5}$$

Função tangente hiperbólica: É uma função sigmóide que varia entre –1 e +1. É dada por (2.6):

$$f(x) = \frac{1 - e^{-x}}{1 + e^{-x}} \tag{2.6}$$

Função Gaussiana: A saída do neurônio produzirá resultados iguais para aqueles valores de potencial de ativação {*u*} que estejam posicionados a uma mesma distância de seu centro, sendo que a curva é simétrica em relação a este. É dada por (2.7) (SILVA, *et. al*, 2010).

$$f(x) = e^{-\frac{(u-c)^2}{2\sigma^2}} \tag{2.7}$$

onde c é um parâmetro que define o centro da função gaussiana e σ denota o desvio padrão associado à mesma, isto é, o quão espalhado (dispersada) está a curva em relação ao seu centro.

2.1.3 Topologia das RNA

Existe uma grande variedade de redes, cada uma produzindo diferentes resultados. Elas podem ser classificadas basicamente em alimentadas a frente (*feed-forward*) e recorrentes (RUSSELL *et al*, 2004).

Nas redes *feed-forward* os neurônios estão dispostos em camadas, podendo haver redes com uma única camada e redes com múltiplas camadas. As redes *feed-forward* são inerentemente acíclicas, ou seja, o sinal é propagado somente da entrada para a saída da rede. Também são chamadas redes sem memória.

As redes multicamadas se distinguem das redes de camada única pela presença de uma ou mais camadas ocultas. As entradas de um neurônio são as saídas dos neurônios da camada anterior, portanto não há ligação entre neurônios de uma mesma camada. Uma rede *feed-forward* pode ser representada pela notação e - o_1 - o_2-...- o_n - s, onde e representa o número de neurônios nas camadas de entrada, $o_1, o_2,..., o_n$ representam o número de neurônios nas camadas ocultas e s o número de neurônios na camada de saída.

A figura 2.3 ilustra uma rede *feed-forward* de camada única e uma rede multicamada 4-2-1 (HAYKIN, 2001) e (RUSSELL *et al*, 2004).

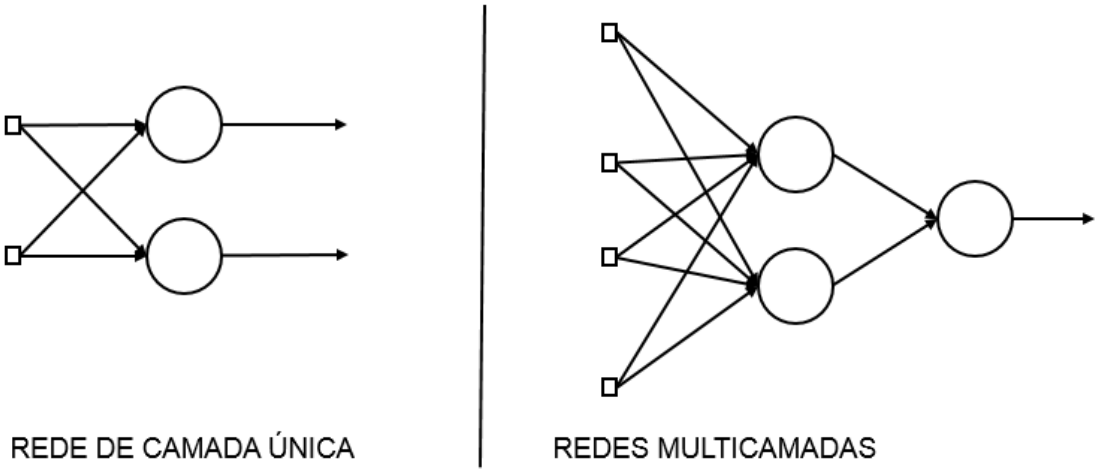


FIGURA 2.3 – EXEMPLOSDE REDES *FEED-FORWARD*
FONTE: ADAPTADO DE HAYKIN (2001)

Perceptron e o MLP (*Multi-Layer Perceptron*) são, respectivamente, exemplos de modelos de rede de camada única e rede de múltiplas camadas.

Ao contrário das redes *feed-forward*, as redes recorrentes possuem laços de realimentação, ou seja, a saída de um neurônio pode ser entrada para outro de uma camada precedente ou, no caso de auto-realimentação, para o próprio neurônio.

As redes recorrentes, chamadas de redes com memória, não possuem organização rígida e seus neurônios têm liberdade para se ligar a qualquer outro neurônio (HAYKIN, 2001) e (RUSSELL *et al*, 2004). A figura 2.4 ilustra uma rede recorrente.

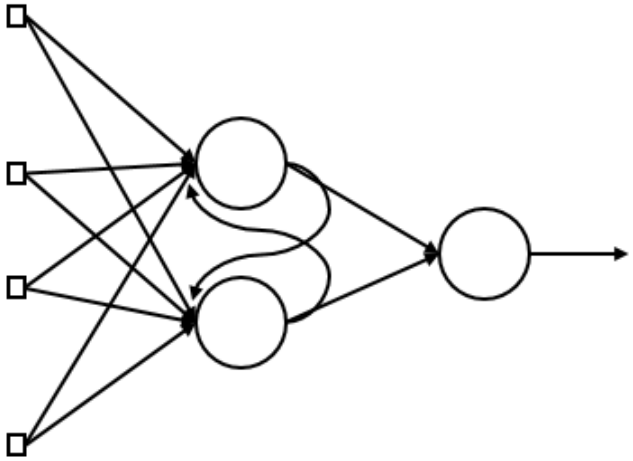


FIGURA 2.4 – ESQUEMA DE UMA REDE RECORRENTE
FONTE: DIAS (2007)

Como exemplos de modelos de redes recorrentes temos a rede de *Elman*, rede de *Hopfield*, rede de *Jordan e NARX*, dentre outras (Kohonen, etc.).

2.1.4 Aprendizado das RNA

O processo de aprendizagem ocorre através de um processo interativo de ajuste dos parâmetros livres, pesos sinápticos, por estimulação do ambiente (HAYKIN, 2001).

Os paradigmas de aprendizado são: aprendizado supervisionado e aprendizado não supervisionado, descritos a seguir.

Aprendizado Supervisionado: também chamado de aprendizado “com professor”. Esta forma de aprendizado se baseia em um conjunto de exemplos de entrada-saída que é apresentado à rede. A partir da entrada, a rede realiza seu processamento e a saída obtida é comparada com a saída esperada. Caso não sejam iguais, um processo de ajuste de pesos é aplicado buscando-se um erro mínimo ou aceitável. O algoritmo de aprendizado supervisionado mais comum é o *Backpropagation* (HAYKIN, 2001).

Aprendizado não supervisionado: é caracterizado pela ausência de algum elemento externo supervisor, ou seja, um padrão de entrada fornecido permite que a rede livremente escolha o padrão de saída a partir das regras de aprendizado adotadas. Pode ser:

- Aprendizado por reforço que consiste no mapeamento entrada-saída através da interação com o ambiente e;
- Aprendizagem auto-organizada onde, a partir de métricas de qualidade do aprendizado ocorre a otimização dos parâmetros livres da rede.

Pode, por exemplo, ser utilizada a regra de aprendizagem competitiva. Os algoritmos de aprendizado não supervisionado mais importantes são: Algoritmo de *Hopfield* e Mapas de *Kohonen* (HAYKIN, 2001).

2.1.5 REDE NEURAL DE BASE RADIAL

As Redes Neurais de Base Radial (RBF ou *Artificial Neural Networks of Radial Basis Functions*) são RNA com múltiplas camadas que não são treinadas por retropropagação (*backpropagation*) e que não têm unidades de processamento com função de ativação do tipo sigmoidal.

Redes que funcionam de acordo com esta estratégia, utilizam unidades com campos receptivos locais (*local receptive fields*), nos quais as unidades que recebem entradas diretamente da entrada do sistema estão habilitadas a “ver” apenas parte destas entradas.

Esta abordagem emprega, na maioria dos casos, treinamento supervisionado e não-supervisionado. As redes são muito empregadas como interpoladores/aproximadores e em tarefas de classificação. Algumas extensões do método são aqui mostradas.

Esta abordagem é inspirada na propriedade de alguns neurônios biológicos chamada de resposta localmente sintonizada (*locally tuned response*). Tais células nervosas respondem seletivamente a um intervalo finito do espaço de sinais de entrada.

Os modelos de RBF de hoje se diferenciam dos primeiros, pois são de natureza adaptativa que permite a utilização, em muitas situações, de um número relativamente menor de unidades de processamento localmente sintonizadas.

2.1.5.1 Arquitetura e processamento das redes RBF

As redes RBF's são redes de alimentação direta (*feedforward*) consistindo tipicamente de três camadas: entrada, escondida e saída, sendo que a primeira apenas propaga as entradas. A saída de uma rede RBF com N neurônios na camada intermediária, dada uma entrada x_i, a saída $\widehat{y_i}$ é dada por:

$$\widehat{y_i} = \sum_{j=1}^{N} F(x_i, \mu_j, \sigma_j) w_j \qquad ((2.8)$$

onde o vetor $x_i = [x_{i1}\, x_{i2} \ldots x_{im}]^T$ é uma das possíveis entradas, m é o número de dados em cada entrada (dimensão do dado de entrada), $\mu_k = [\mu_{k1}\, \mu_{k2} \ldots \mu_{kN}]$ é a coordenada do centro da RBF k, σ_k é a largura do campo receptivo relativa ao centro da RBF k, esses são os parâmetros de entrada na função F, que é dada por:

$$F(x_i, \mu_j, \sigma_j) = \frac{1}{\sigma_j \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\sqrt{(x_i - \mu_j)^T (x_i - \mu_j)}}{\sigma_j} \right)^2} \qquad ((2.9)$$

Essa função é do tipo gaussiana, a mais utilizada como função de ativação em RBF's, porém, existem outras opções como, por exemplo, a logística e a multiquadrática inversa, citadas adiante. O termo $\sqrt{(x_i - \mu_j)^T (x_i - \mu_j)}$ da função (2.9) calcula a distância euclidiana entre uma entrada x_i e um centro μ_j (podem ser utilizadas outras métricas de distância, citadas adiante), uma entrada x_i próxima a um centro μ_j, aplicadas a função F (vide figura 2.6) retorna um valor próximo de 1, a medida que uma entrada x_i se afasta do centro, a função F retorna valores menores, a velocidade dessa redução está inversamente ligada ao valor de μ_j. O papel da função F é fornecer uma métrica de quão semelhante o padrão de entrada i é ao grupo representado pelo grupo j.

A fase não supervisionada da rede neural RFB consiste em mapear os dados de entradas, sendo que um número definido de centros/neurônios são posicionados no espaço de interesse de acordo com a dispersão dos dados, ou seja, define-se os μ_j e os σ_j. A figura 2.5 ilustra o posicionamento ideal dos centros (coordenadas) em relação aos dados de entradas (caso dados de entrada bidimensional e a camada oculta da rede RBF possuir quatro neurônios/centros). Esse posicionamento pode ser obtido por técnicas de clusterização hierárquica dos dados de entrada como, por exemplo, o k-médias. Além das coordenadas do centro, normalmente sendo a média das entradas associadas ao grupo, também é calculada a largura do campo receptivo

de cada centro (normalmente definido como o desvio padrão do centro em relação aos dados de entrada pertencente a esse centro/grupo).

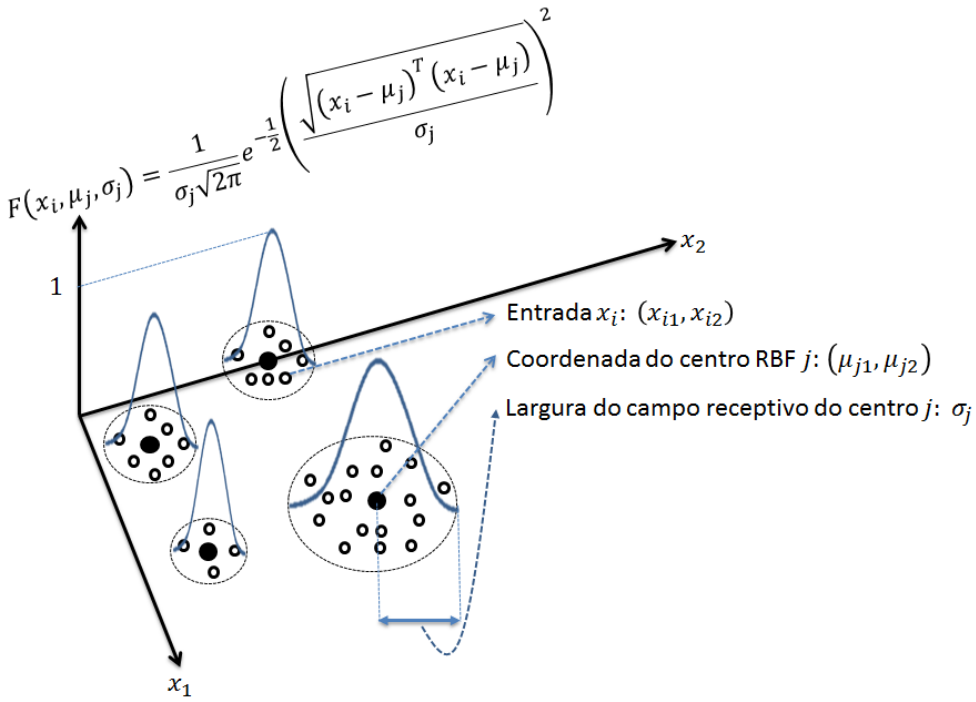


FIGURA 2.5 - DISPERSÃO DOS DADOS EM RELAÇÃO DOS CENTROS DAS RBF'S

A fase supervisionada consiste em ajustar os elementos do vetor w, de forma a minimizar o mse (erro médio quadrático, podendo ser utilizadas outras métricas de erro) em relação aos dados de saída reais do conjunto de treinamento (dado por y_i), que é dado por:

$$mse = \frac{1}{n}\sum_{k=1}^{n}(\|y_i - \hat{y}_i\|)^2 \qquad ((2.10)$$

O ajuste do vetor w pode ser feito, por exemplo, utilizando o algoritmo *backpropagation*, caso a função de ativação na camada de saída seja não linear, ou pelo ajuste usando o método dos mínimos quadrados (método sintetizado pela matriz pseudo inversa), caso a função de ativação na camada de saída seja linear. Nesse último caso, w é calculado pela seguinte expressão:

$$w = \underbrace{\left[\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}^T \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}\right]^{-1} \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{bmatrix}}_{matriz\ pseudo\ inversa} \quad (2.11)$$

Após o ajuste dos parâmetros μ_j, σ_j *e* w a partir dos dados de entrada e saída (n é o número de dados de entrada), se finalizou a etapa que é chamada de treinamento da rede RBF. O teste da rede RBF consiste em verificar o *mse* dado por:

$$mse = \frac{1}{n}\sum_{k=1}^{n}\left(\left\|y_i' - \widehat{y_i}'\right\|\right)^2 \quad (2.12)$$

onde y_i' e são os dados de saída reais do conjunto de teste e $\widehat{y_i}'$ as saídas simuladas pela rede neural cuja entradas são os dados de entrada do conjunto de teste.

2.1.5.2 Parâmetros e equações das redes RBF

Redes RBF realizam aproximação de uma função $F(x_i, \mu_j, \sigma_j)$ por superposição de funções de base radial não-ortogonais que têm forma de sino. O grau de precisão pode ser controlado por três parâmetros:

- Número de funções de base usadas;
- Localização e
- Largura do campo receptivo.

As Funções de ativação $F(x_i, \mu_j, \sigma_j)$ das unidades escondidas da rede RBF, mais comuns, são:

Função de base Gaussiana em (2.13):

$$F(x_i, \mu_j, \sigma_j) = \exp\left(-\frac{\left\|x_i - \mu_j\right\|^2}{2\sigma_j^2}\right) \quad 2.13$$

onde μ_j é a média do campo receptivo da unidade *j*, σ_j é o desvio padrão do campo receptivo da unidade *j* e $\left\|x_i - \mu_j\right\|$ é uma norma Euclidiana.

Função de base logística em (2.14):

$$F(x_i, \mu_j, \sigma_j) = 1 + \exp\left(-\frac{\|x_i - \mu_j\|^2}{\sigma_j^2}\right) - \theta_j \qquad 2.14$$

onde θ_j é um *"bias"* ajustável.

Função multiquadrática inversa em (2.15):

$$F(x_i, \mu_j, \sigma_j) = \frac{1}{\sqrt{\|x_i - \mu_j\|^2 + \sigma_j^2}} \qquad 2.15$$

Redes RBF's são adequadas para aproximar mapeamentos de valores reais contínuos, ou contínuos por partes $F: R^N \rightarrow R^M$, para *N* suficientemente pequeno. Problemas de classificação se constituem em um caso particular daquela classe de problemas.

2.1.5.3 Aprendizado em redes RBF

O método de treinamento, como os demais modelos supervisionados, deve reduzir o erro na saída da rede (E) a valores aceitáveis através da adaptação dos parâmetros livres na rede RBF:

- Os centros dos campos receptivos (μ_j);
- As larguras dos campos receptivos (σ_j) e
- Os pesos entre a camada escondida e a camada de saída (w_{lj}).

A aprendizagem pode ser supervisionada, não-supervisionada ou híbrida (que ocorre na maioria dos casos).

O treinamento híbrido combina aprendizagem não-supervisionada com supervisionada. A primeira treina a camada escondida, definindo os parâmetros desta camada (localização dos centros e larguras dos campos receptivos). A segunda etapa define os valores dos pesos entre as camadas escondidas e de saída.

Nesta segunda fase, os parâmetros definidos na primeira fase não se modificam. Este é o tipo de treinamento mais empregado pois, em geral, não se sabe que saídas se desejam para a camada escondida. Resumidamente, tem-se:

O treinamento não-supervisionado consiste em mapear os dados de entradas, que definido o número de centros/neurônios, são posicionados de acordo com a dispersão dos dados, esse posicionamento pode ser obtido por:

- seleção aleatória;
- distribuição sobre uma grade regular;
- técnica de agrupamento (*clustering*);
- outro algoritmo;
- Determina-se a largura do campo receptivo através de uma heurística.

Já no Treinamento supervisionado:

Determinam-se os pesos (elementos do vetor w) por um método que resolva o problema de minimização do erro, que podem ser:

- método dos mínimos quadrados;
- método da regra delta;
- matriz pseudo-inversa.

O treinamento não-supervisionado para a camada escondida compreende a determinação dos centros das funções de base radial e das larguras dos campos receptivos.A seguir são apresentados métodos para determinação dos centros das funções de base radial.

Os centros da camada escondida podem ser selecionados através das seguintes estratégias:

a) Seleção aleatória (CHEN, 1991):

Nesta estratégia, os centros são vetores de entrada aleatoriamente selecionados. Esta técnica demanda que os padrões de treinamento representem acuradamente todo o espaço de soluções do problema. Este método é simples e direto, no entanto pode exigir grande número de unidades intermediárias, escolher centros muito próximos uns dos outros que podem acarretar funcionamento inadequado da rede;

b) Fixação em grade regular (BISHOP, 1996):

Neste caso, os centros são fixados em uma grade regular, cobrindo todo o espaço de entrada. Em geral, este método exige muitas unidades intermediárias para vetores de entrada com dimensão alta (“maldição” da dimensionalidade: crescimento exponencial do número das unidades escondidas);

c) Técnicas de agrupamento (MOODY *et al*, 1989):

Os centros são definidos por técnicas de agrupamento, entre as quais se destacao algoritmo de *k*-médias (a seguir) e mapas auto-organizáveis;

O algoritmo das *k*-médias divide os padrões de treinamento em *nv* grupos, encontrando o ponto central de cada um deles através da expressão (2.16):

$$\mu_j = \frac{1}{nv_j} \sum_{x_p \epsilon S_j} x_p \qquad (2.16)$$

onde nv_j é o número de vetores contidos no agrupamento S_j.

Os primeiros centros são inicializados arbitrariamente. Em seguida, os padrões de entrada vão trocando de centro (de acordo com a distância euclidiana) até se chegar a uma situação estável. O número de centros é determinado de acordo com a estratégia de treinamento que pode ser, por exemplo, a validação cruzada.

d) Mapas auto-organizáveis

Os mapas auto-organizáveis (*SOM*,*Self Organizing Maps*) caracterizam-se por agrupar padrões espacialmente próximos que compartilhem micro características. No início, os centros são aleatoriamente atribuídos. O centro que apresentar maior produto escalar com um dado vetor de entrada adiciona uma versão ponderada deste vetor de entrada ao seu grupo.

As Heurísticas para determinação da largura do campo receptivo são apresentadas a seguir. O valor de (σ_j) pode ser único para todas as unidades ou pode ser diferente para cada unidade escondida. Algumas das principais heurísticas são:

a) Utilização da distância euclidiana média entre centros conforme (2.17) (MOODY *et al*, 1989):

$$\sigma = \frac{1}{ng}\sum_{j=1}^{ng}\left\|\mu_j - \mu_{j(mprox)}\right\| \quad (2.17)$$

onde *ng* é o número de grupos que serão formados e $\boldsymbol{\mu}_{j(mprox)}$ é o centro com menor distância euclidiana com relação ao centro $\boldsymbol{\mu}_j$. Este método produz um único valor de raio.

b) Utilização da distância euclidiana entre centro e vetor de entrada conforme (2.18) (SAHA et al, 1990):

$$\sigma_j^2 = \frac{1}{nv}\sum_{x_p \in \psi_j}\left\|\mu_j - \mu_p\right\|^2 \quad (2.18)$$

onde Ψ_j é o conjunto dos *nv* vetores de entrada com menor distância euclidiana para o centro $\boldsymbol{\mu}_j$.

c) Utilização da distância euclidiana entre centros como em (2.19) (HASSOUN, 1995):

$$\sigma_j = \alpha\left\|\mu_j - \mu_{j(mprox)}\right\| \quad (2.19)$$

onde $\boldsymbol{\mu}_{j(mprox)}$ é o centro com menor distância euclidiana com relação ao centro $\boldsymbol{\mu}_j$ e o parâmetro geralmente se situa $1{,}0 \leq \alpha \leq 1{,}5$.

Utilização da distância euclidiana entre os centros determinados pelo método *k*-médias como em (2.20):

$$\sigma_j^2 = \frac{1}{na}\sum_{x_p \in S_j}\left\|\mu_j - x_p\right\|^2 \quad (2.20)$$

onde S_j é o agrupamento contendo *na* vetores de entrada.

O treinamento supervisionado para a camada de saída é a etapa compreende a determinação dos pesos entre a camada escondida e a de saída. A primeira etapa neste processo é o cálculo do erro. Este é função da resposta dada pela rede comparada com a resposta que se deseja dela. Existem algumas maneiras diferentes de se calcular o erro:

Soma dos erros quadráticos conforme (2.21) (SSE - *sum of squared error*)

$$SSE = \sum_{i=1}^{npad} \| y_d^{(i)} - y_0^{(i)} \|^2 \quad (2.21)$$

Erro quadrático médio conforme (2.22) (MSE - *mean squared error*)

$$MSE = \frac{1}{K} \sum_{i=1}^{npad} \| y_d^{(i)} - y_0^{(i)} \|^2 \quad (2.22)$$

Erro relativo médio conforme (2.23) (MRE – *mean relative error*)

$$MRE = \frac{1}{K} \sum_{i=1}^{npad} \left\| \frac{y_d^{(i)} - y_0^{(i)}}{y_d^{(i)}} \right\|^2 \quad (2.23)$$

Raiz do erro quadrático médio conforme (2.24) (RMSE – *root mean squared error*)

$$RMRE = \sqrt{\frac{1}{K} \sum_{i=1}^{npad} \left\| y_d^{(i)} - y_0^{(i)} \right\|^2} \quad (2.24)$$

onde $y_d^{(i)}$ e $y_0^{(i)}$ são o *i*-ésimo padrão desejado e obtido respectivamente e *K* é o número total de padrões.

Escolhida a métrica do erro, este é minimizado por procedimentos tais como o algoritmo *backpropagation*, caso função de ativação na camada de saída seja não linear, ou pelo ajuste usando o método dos mínimos quadrados (método sintetizado pelo cálculo da matriz pseudo-inversa), caso função de ativação na camada de saída seja linear.

2.1.5.4 Estratégias de treinamento

O compromisso entre precisão e generalização deve ser obtido através da aprendizagem. Para tal, duas estratégias de treinamento podem ser empregadas:

Leave-one-out e validação cruzada (*cross-validation*) (HAYKIN, 2001). Sucintamente, pode-se dizer que:

Validação cruzada: Neste procedimento o conjunto de padrões é dividido em três grupos: treinamento, validação e teste. Cada topologia tem, com seus centros, seu desempenho testado com respeito aos três conjuntos.

Leave-one-out: Esta é uma estratégia típica de situações onde todos os padrões devem ser considerados para o treinamento (normalmente, pois eles são poucos). Neste caso, dividem-se os padrões em *ng* grupos. Seleciona-se aleatoriamente *ng*-1 conjuntos para treinamento e testa-se a rede com aquele conjunto que não foi selecionado. Este processo deve prosseguir até que todos os conjuntos tenham sido usados para testes. A partir daí, calcula-se o erro (*E*) que pode ser dado através de (2.25):

$$E = \frac{1}{ng}\sum_{i=1}^{ng} E_i \qquad (2.25)$$

2.1.5.5 Desempenho das redes RBF

As redes RBF têm sido aplicadas com sucesso na aproximação de funções e em problemas de classificação. Em tarefas difíceis de aproximação/interpolação (por exemplo, predição da série caótica de *Mackey-Glass T* instantes de tempo no futuro, *T*>50), redes RBFs que empregam a técnica de agrupamento no posicionamento dos campos receptivos podem alcançar desempenho comparável ao das redes de retropropagação (redes de alimentação direta com unidades escondidas sigmoidais e treinadas por retropropagação), enquanto requerem tempo de treinamento algumas ordens de grandeza menor (HAYKIN, 2001).

Em tarefas difíceis de classificação, redes RBFs empregando um número suficiente de padrões de treinamento e de unidades escondidas podem superar o desempenho de redes retropropagação, obtendo melhores taxas de classificação e menores erros de classificações positivas falsas (HASSOUN, 1995).

As redes RBF têm tempo de treinamento muito menor porque apenas uma pequena fração de unidades escondidas responde a um dado padrão de entrada (pois

são unidades localmente sintonizáveis, sensíveis apenas a padrões próximos de seus campos receptivos). Isto permite o uso eficiente de algoritmos auto-organizáveis no ajuste dessas unidades no modo de treinamento, que não envolve a camada de saída da rede. Por outro lado, todas as unidades de uma rede retropropagação são avaliadas e têm seus pesos ajustados para cada vetor de entrada. Outro fator que contribui para a velocidade de treinamento das redes RBF é o esquema de treinamento separado da camada escondida e da camada de saída.

Quando utilizadas em aproximação de funções, as redes do tipo retropropagação conseguem maior capacidade de generalização (ou extrapolação) do que as RBF, pois ajustam globalmente os padrões de entrada enquanto estas últimas fazem um ajuste local. Pelo mesmo motivo, em problemas de classificação, redes RBF cometem menos erros de falsa classificação positiva do que as redes do tipo retropropagação. Pode-se dizer que, em geral, é melhor o uso de redes do tipo *feed-forward* treinadas por retropopagação (do erro) quando os padrões de entrada difícieis de se obter e/ou quando a velocidade de recuperação, considerando-se a implementação em máquinas seriais, é crítica. Esse tipo de rede, em geral, é menor, requer menos memória e leva a maiores velocidades de recuperação do que as RBF.

No entanto, se os dados são “baratos” e abundantes, e se é necessário treinamento *on-line* (como no caso de processamento de sinal adaptativo e controle adaptativo, onde os dados são adquiridos em altas taxas e não podem ser salvos), então as redes RBF são superiores (HAYKIN, 2001).

2.1.5.6 Análise de agrupamento aplicada ao treinamento não supervisionado em RBF

Antes de ajustar os pesos entre a camada de intermediária e a camada de saída de uma rede neural de base radial, é realizado o posicionamento das funções de base radial, para isso, é utilizada alguma técnica de agrupamento/clusterização, e de acordo com os grupos formados, é realizado o posicionamento, normalmente sendo a média dos dados que formam cada grupo, esse posicionamento depende unicamente dos dados de entrada, por isso essa etapa é considerada não supervisionada, e utiliza-se principalmente o método denominado *k-médias* para esse fim.

O *k-médias* é classificado como um método de agrupamento não hierárquico, pois se define *a priori* o número de grupos/cluster's a serem formados antes da execução do método. Muitos trabalhos propõe outros métodos não hierárquicos e/ou variações desses, todos buscando melhorar o posicionamento das funções de base radial e, com isso, melhorar a *performance* da rede neural.

Este item começa explicando o funcionamento do métodos *k-médias* e os principais métodos de agrupamento hierárquicos da bibliografia, dando enfoque aos utilizados nesse trabalho, para que no próximo capítulo seja descrito como que as técnicas de agrupamento hierárquicas serão incorporadas na etapa não supervisionada da rede neural de base radial.

2.1.5.7 Análise de agrupamentos (cluster)

A análise de agrupamentos, também conhecida como análise de conglomerados, classificação ou cluster, tem como objetivo dividir os elementos da amostra, ou população, em grupos de forma que os elementos pertencentes a um mesmo grupo sejam similares entre si com respeito às variáveis (características) que neles foram medidas, e os elementos em grupos diferentes sejam heterogêneos em relação a estas mesmas características.

Várias são as situações nas quais a análise de agrupamentos se faz presente. Em psicologia, onde é utilizada na classificação de pessoas de acordo com seus perfis de personalidade (SPEECE *et al*., 1985), em pesquisa de merdado, na identificação do posicionamento de produtos (ou serviços) em relação aos concorrentes de mercado e na segmentação de clientes de acordo com perfis de consumo (PUNJ *et al*, 1983), (BERRY, 1997) e (LINOF 2005), em ecologia, na classificação de espécies (MCGARIGAL *et al*, 2000), em geoquímica, na caracterização de conteúdo de minerais (JUNIOR et al, 1998), na ergonomia, na seleção de assentos para uso em determinadas atividades (WERNER et al, 2008), em geografia, na classificação de cidades, estados ou regiões de acordo com suas variáveis físicas, demográficas e econômicas. A análise de agrupamentos também é muito popular em Data Mining (mineração de dados), que está relacionada com a análise de dados e uso de ferramentas computacionais na busca de padrões em conjunto de dados (HAND, 1998) e (Neto L., 2006), estando implementado em alguns software estatísticos.

Uma questão importante refere-se ao critério a ser utilizado para se decidir até que ponto dois elementos do conjunto de dados são semelhantes ou não. Uma maneira é considerar medidas que descrevam a similaridade entre elementos amostrais de acordo com as características que neles foram medidas, considerando que cada elemento amostral têm-se informações de m variáveis armazenadas em um vetor. Calculando as distâncias entre os vetores/observações, agrupa-se aqueles de menor distância, ou seja, os mais similares.

Os coeficientes de similaridade mais usuais, obtidos num espaço multidimensional, podem ser subdivididos em três categorias:

a) os que medem a distância, ou a separação angular, entre pares de pontos;

b) os que medem a correlação entre pares de valores;

c) os que medem a associação entre pares de caracteres qualitativos.

Existem diversas publicações que discutem esses diversos tipos de medidas como, por exemplo, (SNEATH *et. al.,* 1973), (EVERITT, 1990), (PRENTICE 1992), (GORDON, 1993), (SMITH G., 1995) e (PIELOU 1998).

Medidas de distâncias expressam o grau de similaridade como distância em um espaço multi-dimensional. Quanto maior a distância, maior o grau de similaridade e vice-versa. A distância D entre dois pontos, cuja localização é especificada num sistema de coordenadas cartesianas, é fornecida, segundo o teorema de Pitágoras, representada pela equação 2.26.

$$D_{1.2} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \qquad 2.26$$

onde, $x_1, x_2, y_1, y_2,$ são valores das coordenadas dos dois pontos.

Para a distância entre dois pontos, num espaço n-dimensional, é dada pela fórmula generalizada 2.27.

$$D_{ij} = \sqrt{\sum_{k=1}^{n} (x_{ik} - x_{jk})^2} \qquad 2.27$$

onde D_{ij} é a distância entre os pontos i e j, x_{ik} x_{jk} as coordenadas k dos pontos i e j respectivamente. Esse é um tipo de distancia (euclideana), outros podem ser utilizados, como comendado a seguir.

Pode, também, ser utilizado o coeficiente cosseno-teta, que é uma medida de proporcionalidade expressando o grau de similaridade em termos de separação angular dos pontos p e q:

$$Cos\ \theta_{pq} = \frac{\sum x_{ip} x_{iq}}{\left(\sum x_{ip}^{\ 2} \sum x_{iq}^{\ 2}\right)^{1/2}}, \qquad \text{p e q = valores comparados} \qquad 2.28$$

onde x_{ip} e x_{pi} são as coordenadas k dos pontos p e q respectivamente.
Quando a similaridade é completa, a separação angular é 0º e cos q = 1, quando não ocorre similaridade nenhuma, a separação angular é 90° e o cos q = 0.

2.1.5.8 Técnicas de agrupamento não-hierárquico

Esses métodos procuram diretamente uma partição dos *n* objetos, de modo que satisfaçam às duas premissas básicas: coesão interna de isolamento dos grupos. Portanto, estas técnicas exigem a prefixação de critérios que produzam medidas sobre a qualidade da partição produzida. O uso dos métodos de partição pressupõe também o conhecimento do número de k partições desejadas.

Os algoritmos de agrupamento não hierárquicos diferem um do outro pela escolha diferente de um ou mais dos seguintes procedimentos:

(a) Método de iniciar os agrupamentos

(b) Método de designar objetos aos agrupamentos iniciais;

(c) Método de resignar um ou mais objetos já agrupados para outros agrupamentos.

O método das *k-medias* é uma técnica de agrupamento não hierárquica, muito conhecida e utilizada. Grande parte dos trabalhos, com poucas exceções, utilizam esse método para determinar as coordenadas dos centros dos neurônios das redes neurais base radial.

O algoritmo *k-medias* é um método não-supervisionado de classificação que tem como objetivo particionar *K* registros em k agrupamentos, onde $k < K$. Seu funcionamento é descrito a seguir.

Dado um valor inicial de k médias, os registros são separados em agrupamentos, onde esses pontos k médias representam o centro de cada agrupamento. Normalmente, as coordenadas iniciais desses centróides são determinadas de forma aleatória. Em seguida, cada registro é associado ao cluster cujo centro está mais próximo, seguindo uma métrica de distância. Existem diversas métricas de distância, sendo a mais comum a euclidiana, outras métricas serão citadas adiante. Quando todos os registros estiverem classificados, os k centros são recalculados como as médias aritméticas dos registros de cada cluster. Então, os registros são novamente associados a um agrupamento segundo sua distância à média do cluster e os centros são novamente calculados. Esse passo se repete até que as médias dos clusters não se desloquem consideravelmente. O processo visa minimizar a soma dos quadrados dada por pela equação (2.29), supondo que existam K amostras no total e deseja-se encontrar os k vetores μ_j onde $j = 1, \dots, k$, e se deseja particionar as amostras x^i em k vetores subconjuntos S_j contendo contendo K_j amostras.

$$J = \sum_{j=1}^{k} \sum_{n \in S_j} \left\| x^i - \mu_j \right\|^2 \qquad 2.29$$

onde μ_j é o centro das amostras do conjunto S_j e é dado pela equação (2.30):

$$\mu_j = \frac{1}{K_j} \sum_{n \in S_j} x^n \qquad 2.30$$

Foi proposto em 2007 por David Arthur e Vassilvitskii Sergei, um método de melhoria do k-medias, batizado como k-médias ++, cuja tentativa principal foi minimizar uma das maiores deficiências do método: a sensibilidade no desempenho do método em relação as sementes (ou centros) iniciais escolhidas aleatoriamente. Alguns artigos citaram esse método como, por exemplo, (LEE et. al, 2007) e (HOWARD, J., 2009), indicando bons resultados. No entanto, num estudo muito completo realizado por (ANNA, D. P. *et al*, 2010) comparando 11 métodos de inicialização para o k-médias, verificou-se um desempenho inferior do k-medias ++.

Há outras técnicas de agrupamento não hierárquicas que são utilizadas em redes neurais RBFs como, por exemplo, o algoritmo *fuzzy* c-médias desenvolvido inicialmente e melhorado por (DUNN, J. C., 1973) e (BEZDEK, J. C. *et all*,1981) respectivamente; o algoritmo de Gustafson-Kessel, uma extensão do algoritmo c-médias nebuloso (GUSTAFSON, D. E. *et al*., 1979); o algoritmo de Gath-Geva

proposta por (BEZDEK, J. C. *et al*, 1985); assim como as redes neurais aplicadas à analise de *cluster* (SCHREER *et al*. 1998) e (MANGIAMELLI et al. 2006).

Pelo fato do método k-médias por ser um dos mais empregados na etapa não supervisionada, este será utilizado para comparações nesse trabalho.

2.1.5.9 Técnicas de agrupamento hierárquico

Quando não se tem uma ideia clara de quantos grupos distintos existem em uma base de dados, são comumente utilizados métodos de agrupamento hierárquico, pois eles fornecem informações que podem ajudar na definição do número de grupos mais adequado aos dados. Ou seja, não se determina o número de grupos *a priori*; o método escolhido fornece parâmetros para que se faça uma análise de número de grupos que são distintos, e quais seriam os elementos que formam cada grupo.

As técnicas hierárquicas aglomerativas partem do princípio de que no início do processo de agrupamento tem-se k conglomerados, ou seja, cada elemento do conjunto de dados observado é considerado como um conglomerado isolado. Em cada passo do algoritmo, os elementos dos dados amostrais vão sendo agrupados, formando novos conglomerados até o momento no qual todos os elementos considerados estão num único grupo. Portanto, no estágio inicial do processo de agrupamento, cada elemento amostral é considerado como um *cluster* de tamanho um e no último estágio de agrupamento tem-se um cluster constituído de todos os elementos amostrais. Em termos de variabilidade, no estágio inicial, tem-se a partição com a menor dispersão interna possível, já que todos os elementos amostrais tem um único elemento e, logo, a variância de cada um deles é zero. No estágio final, tem-se a maior dispersão interna possível, já que todos os elementos amostrais estão em um único cluster. Em cada estágio do procedimento de agrupamento, os grupos são comparados através de alguma métrica de sensibilidade ou (dissimilaridade) previamente definida.

Em cada estágio do algoritmo, o novo conglomerado formado sempre é um agrupamento de conglomerados nos estágios anteriores, ou seja, uma vez unidos os elementos, não poderão ser separados. Devido a essa característica, é possível criar um gráfico chamado de dendrograma, um forma gráfica usada para representar o resultado final dos métodos de agrupamentos hierárquicos, como o exemplo mostrado na figura 2.6. Nele estão dispostas linhas ligadas segundo os níveis de dispersão que

agruparam pares de variáveis. Como este gráfico é uma simplificação em duas dimensões de uma relação n-dimensional, é inevitável que algumas distorções quanto à similaridade apareçam. A medida de tal distorção pode ser obtida por um coeficiente de correlação, dito "cofenético", entre os valores da matriz inicial de similaridade e aqueles derivados do dendrograma.

Visualmente isso pode ser também verificado por meio da construção de um sistema de eixos ortogonais. Nele os valores dos coeficientes de similaridade originais estarão no eixo das abcissas e os coeficientes de similaridade a partir do dendrograma em ordenadas. Se ambas as matrizes forem idênticas, os pontos cairão sobre uma linha reta que passa pela origem do sistema. Desvios dos pontos em relação a essa reta indicarão as distorções. Se situadas acima da reta, indicarão coeficientes de similaridade apontados pelo dendrograma mais altos que os originais e vice-versa.

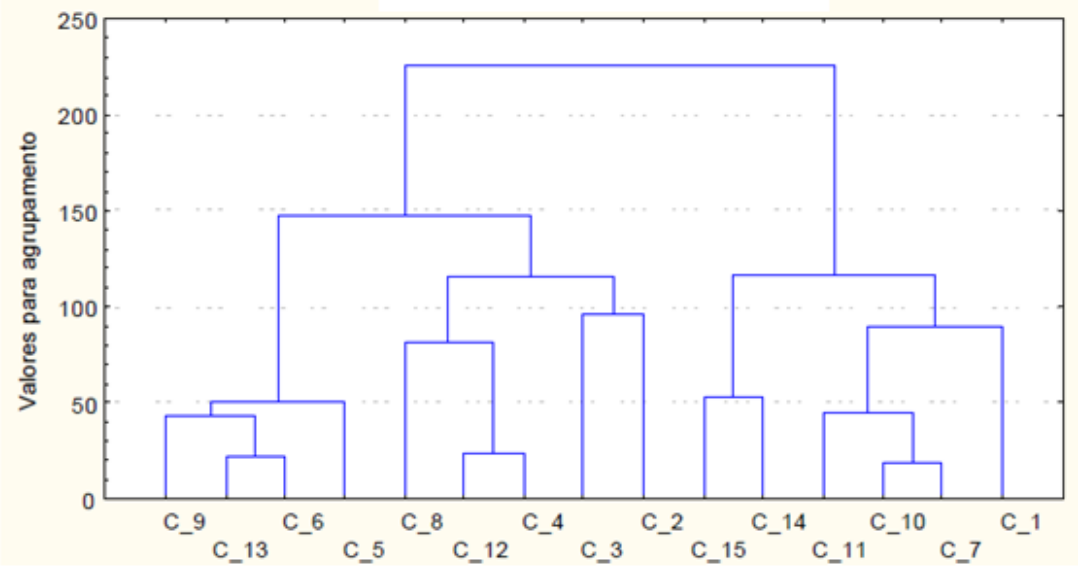


FIGURA 2.6 - EXEMPLO DE UM DENDROGRAMA (ANDERBERG, 2003)

Várias técnicas de agrupamento tem sido propostas, e os métodos mais utilizados são: “ligação simples” (*single linkage method* ou *nearest neighbor*); “ligação completa” (*complete linkage method* ou *farthest neighbor*); ”agrupamento pareado proporcionalmente ponderado" (*weighted pair-group method*, WPGM); “agrupamento pareado igualmente ponderado” (*unweighted pair-group method*”, UPGM); “variância mínima*”* (*minimum variance clustering* ou *Ward’s method of sum-of-squares method* (ANDERBERG, 2003).

No método de ligação simples, os grupos iniciais são determinados pelos mais altos coeficientes de associação mútua. Para admissão de novos membros aos grupos, é suficiente encontrar quais os que representam os maiores coeficientes de

associação com um dos elementos de determinado grupo. A ligação será estabelecida a esse nível de associação com todo o grupo.

No método de ligação completa, os grupos são determinados pelos mais baixos coeficientes de associação mútua, este método é também denominado de método do elemento mais distante, sendo uma das técnicas de hierarquização aglomerativa de maior aplicação na análise de agrupamento (GAMA, 2004). Como no método de ligação simples, não é exigida a fixação, *a priori*, do número de agrupamentos.

Conforme Bussab et al. (1990), no método da ligação completa, também, a dissimilaridade entre dois grupos é definida como sendo aquela apresentada pelos indivíduos de cada grupo que mais se parecem, ou seja, formam-se todos os pares com um membro de cada grupo, e a similaridade entre os grupos é definida pelo par que mais se parece. Este método, geralmente, leva a grupos compactos e discretos, tendo os seus valores de dissimilaridade relativamente grande.

Kaufmann & Rosseeuw (1990) citam as seguintes características desse método:

- Apresenta bons resultados na maioria das métricas de distâncias utilizadas;
- Tendência a formar grupos compactos;
- Os ruídos demoram para serem incorporados ao grupo.

No método de agrupamento pareado também procuram-se inicialmente pelos mais altos coeficientes de associação mútua. Em seguida, esses pares de casos fornecerão valores médios originando um novo elemento singular. No "método de agrupamento pareado igualmente ponderado", para o cálculo dos valores médios, atribui-se sempre o mesmo peso aos dois elementos que estão sendo integrados.

No método de agrupamento pareado proporcionalmente ponderado, para cada agrupamento, é dado um peso proporcional ao número de objetos que o constitui, de tal modo que a incorporação de um novo elemento a um grupo baseia-se no nível médio de similaridade desse elemento com todos os que fazem parte do grupo. Tanto num caso como no outro, alternativamente, em vez de obter valores médios entre os casos podem ser utilizados centroides e verificadas as distâncias entre os mesmos.

No método de agrupamento pela variância mínima, o enfoque é sobre a variabilidade que existe dentro de cada caso e os agrupamentos são efetuados ao se determinar que pares de casos, quando tomados em conjunto, apresentam o menor acréscimo de variabilidade.

No método de ligações singulares, as ligações tendem a ocorrer a níveis mais altos do que nos métodos de agrupamento pareado. No método de agrupamento pareado igualmente ponderado, como cada membro adicionado ao agrupamento tem sempre o mesmo peso, isso traz como efeito que os últimos elementos a se integrarem têm maior influência que os primeiros.

As técnicas de agrupamentos da análise multivariada podem ser utilizadas como pré-processamento e no processamento das redes neurais, na escolha das variáveis de entrada da rede neural, contribuindo com o aperfeiçoamento da mesma, outro uso dessas técnicas com o mesmo objetivo será proposto nesse trabalho.

2.2 PREVISÃO DE SÉRIES TEMPORAIS

Uma série temporal é uma sequência de valores, ordenados no tempo, de uma variável de interesse particular. Também denominada de série histórica, é uma sequência de dados obtidos em intervalos regulares de tempo durante um período específico (SILVA *et al*., 2007). Modelos de séries temporais realizam previsões baseadas em uma série de dados observados em intervalos de tempo regulares, buscando padrões no passado para prever o futuro. Esse tipo de modelagem é especialmente útil quando há pouco conhecimento da base teórica sobre o processo em que os dados foram gerados ou quando a complexidade da explicação sobre o processo é muito alta, como no caso de grandezas físicas.

Apesar de sua grande aceitação, as técnicas de previsão de séries temporais possuem diversas limitações. A mais visível delas é o fato de as causas que agem sobre as variáveis previstas serem, em geral, ignoradas. As forças externas, tais como fatores ambientais, nem sempre são levadas em conta de maneira adequada. Outra deficiência é que os padrões históricos que geraram as séries podem mudar com o tempo, e algumas técnicas não detectam tais mudanças. Assim, estas técnicas podem resultar previsões com baixa acurácia, especialmente em longo prazo (LOPES, 2002).

Entretanto, para obter informações suficientes para elaborar um planejamento de ações e estimar um valor futuro mais próximo do real, como em leituras de instrumentação de obras, algumas técnicas de previsão de séries temporais demonstram grande aplicabilidade.

A preocupação, neste tipo de aplicação, é prever com certo grau de precisão o que vai acontecer, com isso, as técnicas mais avançadas de previsão de séries temporais, os algoritmos matemáticos e estatísticos, são utilizadas mesmo que o esforço computacional seja maior do que o usual, de modo a encontrar valores futuros que permitam antecipar as ações preventivas.

A literatura é vasta quanto à utilização de algoritmos de previsão de séries temporais. Nela encontram-se evidências de que, mesmo não considerando as variáveis ambientais, podem-se ajustar os modelos de modo a compreender e prever períodos de sazonalidades, tendências a queda de leituras ou de elevação.

A figura 2.7 esquematiza os principais métodos aplicados a previsão de séries temporais, alguns deles são detalhados em seguida.

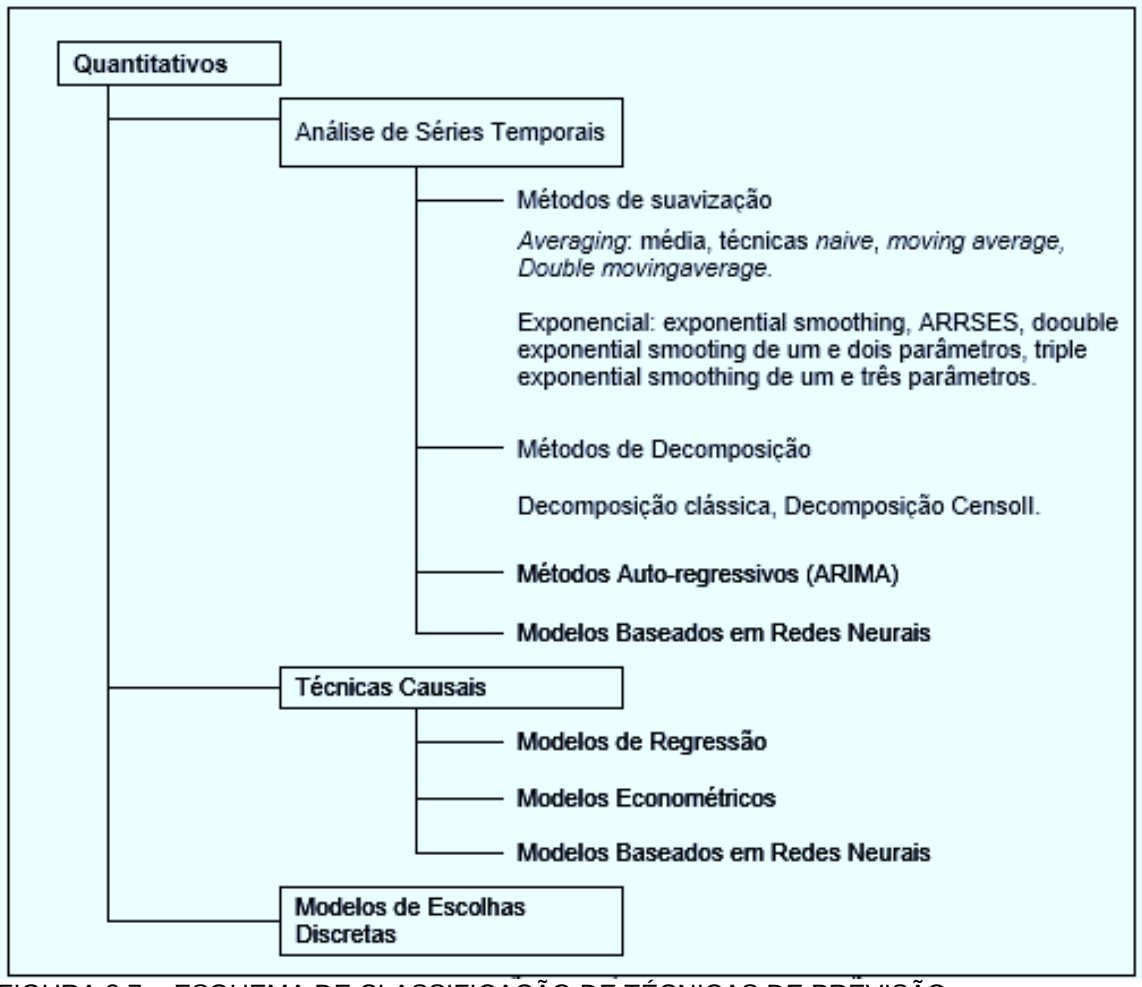


FIGURA 2.7 - ESQUEMA DE CLASSIFICAÇÃO DE TÉCNICAS DE PREVISÃO
FONTE: ADAPTADO DE PASSARI (2003)

As últimas décadas trouxeram mudanças dramáticas na maneira como os pesquisadores analisam séries temporais, a blibliografia na área evolui muito. O livro de Hamilton (1994) sintetiza alguns destes avanços e os torna acessíveis a estudantes e pesquisadores iniciantes. Traz em seu livro tratamentos adequados e inovações importantes na análise séries temporais, como auto regressão vetorial, Método dos Momentos Generalizados, as conseqüências econômicas e estatísticas e os modelos da série não-linear do tempo. Além disso, apresenta ferramentas básicas para análise de sistemas dinâmicos (incluindo representações lineares, funções de auto covariância, análise espectral e o filtro de Kalman) de uma forma que integra teoria econômica com as dificuldades práticas de análise e interpretação dos dados do mundo real.

A existência de observações discordantes com as restantes é de fácil determinação em amostras univariadas, em uma série temporal. Algumas vezes, por observação dos valores que constituem a amostra ou pela análise de alguns gráficos,

é fácil identificar as observações que se afastam da maioria. Em outros casos, é necessária a aplicação de técnicas mais sofisticadas. Em ambos os casos, esta análise prévia tem de ser seguida de testes apropriados para confirmar as suspeitas de existência de observações *outliers* (elementos que não obedecem a um padrão do conjunto de dados ao qual eles pertencem).

Um procedimento utilizado é a intervenção simples em que um parâmetro extra é adicionado para a média da observação em questão. Quando os modelos de séries temporais contêm mais de uma variação (ATHINSON *et al*, 1997), o efeito da intervenção é medido pela mudança de variações das variáveis individuais. Em seu artigo, (ATHINSON *et al*, 1997) estudam o efeito sobre os parâmetros estimados para se fazer a intervenção ao longo da série temporal. Os problemas computacionais envolvidos são superados pelo uso de estatísticas de filtragem e suavização. A interpretação do tempo, resultando em parcelas da série de diagnósticos, é auxiliado por programas de simulação.

Faria e outros (2008) mostram o uso das técnicas estatísticas Alisamento Exponencial Simples e Médias Móveis na previsão de séries temporais, verificando que, apesar de serem métodos simples de previsão, os resultados foram razoáveis. Verificam, também, que aumentando a janela de tempo de observações no método de Médias Móveis, aumenta também a medida de erros de previsão.

O objetivo do livro de Montgomery e Johnson (1976) é destilar e integrar esses resultados através de metodologias coesas e compreensíveis, e de proporcionar uma abordagem simplificada para análise de séries temporais e previsão. O uso de computadores e *softwares* é essencial em qualquer análise quantitativa moderna, mais ainda, na análise de séries temporais, onde algoritmos complexos e extensos cálculos são muitas vezes necessários. Com a velocidade e capacidade dos computadores modernos, em muitas situações é preferível escolher uma metodologia com cálculos mais complexos, porém que simplifica os meios de realizar uma análise, mesmo que seja em detrimento do tempo computacional.

Já o livro de Makridakis *et al.*(1983) é destinado para os problemas práticos de previsão, a cobertura abrangente de modelos estatísticos e como implementá-los na prática dentro de um ambiente empresarial da época. Segundo os autores, explicar o passado é importante, mas não é o suficiente para prever com precisão o futuro.

A proposta apresentada por Dias (1998) é prescrever um processo de previsão para empresas de bens de consumo para elaboração de suas previsões de

curto prazo. Está baseado em uma pesquisa acadêmica que abordou a elaboração da previsão como um processo. Desde o projeto do processo de previsão, a gestão do banco de dados, o *software* utilizado, as características da demanda, a escolha e uso das técnicas quantitativas e qualitativas de previsão, o papel da força de vendas, as revisões das previsões, e até o tratamento dos erros de previsão. Além da revisão bibliográfica, o modelo tentativo para o processo de previsão sugerido no artigo baseou-se também em duas pesquisas de campo. Uma para identificar boas práticas que possam ser incorporadas ao processo proposto e outra para analisar qual pode ser o papel dos *softwares* de previsão.

As previsões podem ser melhoradas através da combinação de previsões distintas obtidas por métodos diferentes (WANG e LAN, 2007). A natureza complementar da análise de cenários e modelos de substituição tecnológica significa que a combinação de dois métodos pode obter melhores previsões. A análise de cenários tem a força de lidar com o futuro incerto, enquanto os modelos baseados nos dados para realizar as previsões são baseados em parâmetros quantificáveis. O estudo utiliza a previsão de participação em um mercado de Fibra de Taiwan nos próximos 10 anos como um exemplo que ilustra o processo proposto de previsão.

2.2.1 Técnicas Estatísticas

A grande maioria dos sistemas de previsão acoplados aos *softwares* disponíveis são simplificados, como o “Alisamento Exponencial Simples” e as “Médias Móveis”. Estes métodos são utilizados para a previsão de séries temporais e são de fácil entendimento e aplicação. Contudo, para séries temporais com alta instabilidade, estes métodos acabam por não satisfazer o objetivo de prever com qualidade a série estudada. A previsão das leituras dos sistemas de instrumentação é uma etapa-chave na análise da segurança da estrutura a ser verificada, devido à complexidade e incertezas intrínsecas às suas atividades. Uma ação preventiva baseada nesses métodos poderia ser equivocada, ou simplesmente não ser tomada devido a uma previsão errônea indicando valores dentro da margem de segurança.

Isso pode ocorrer com grande facilidade devido à característica das variáveis envolvidas, onde a instabilidade dos dados ambientais é muito grande (WANKE, 2003). A sazonalidade, o dia da semana e o horário de medição, bem como os eventos isolados são componentes que influenciam diretamente no comportamento dessas

medições (FIGUEIREDO, 2006). No caso de sistemas de instrumentação, as séries temporais dos dados podem ser sensíveis a qualquer um desses aspectos, fazendo com que um método mais robusto deva ser utilizado para a previsão dessas leituras.

Os métodos tradicionais são aplicados quando se consegue identificar a características da série e a composição do seu componente sistemático (CHOPRA e MEIDL, 2010), (BOWERSOX et. al., 2007) e (NOVAES, 2007). O método mais conhecido e simples é o chamado Médias Móveis, onde se calcula a média de demanda dos *n* períodos anteriores e, a cada nova ocorrência, atualiza-se essa média deslocando o cálculo em uma unidade de tempo para frente e obtendo a nova previsão.

O método de Alisamento Exponencial Simples incorpora a informação da previsão do período anterior e a ocorrência do período anterior, fazendo uma combinação convexa dessas duas informações. A estimativa do método está no ajuste do fator da combinação convexa calculada, a chamada constante de alisamento. Para a aplicação deste método, a demanda não pode possuir tendência ou sazonalidade para se obterem bons resultados (CHOPRA e MEIDL, 2010).

O Método de Holt é o alisamento exponencial corrigido pela tendência, e se o alisamento exponencial for corrigido pela tendência e sazonalidade tem-se o método de Winter. Ambos métodos necessitam da estimativa de coeficientes de influência dessas características e a determinação destes coeficientes que, em geral, são feitas por interpretações do decisor sobre a influência dos dados observados recentemente (NOVAES, 2007).

Os métodos de decomposição da série temporal possuem uma grande aceitação em sua aplicação devido a sua simplicidade matemática e precisão de informações. Este método decompõe a série histórica em quatro fatores: tendência, sazonalidade, ciclo e ruído branco ou índice residual. Na prática, o valor previsto, que é a multiplicação destes quatro índices, pode se resumir à influência dos dois primeiros, tendência e sazonalidade. Para a tendência, em geral, estima-se o valor através do clássico modelo dos Mínimos Quadrados, já para a sazonalidade o valor é calculado como o quociente da demanda real em relação à demanda média, num determinado período (BALLOU, 2006).

A Análise de Regressão Múltipla é uma técnica estatística que ajuda a determinar o grau de associação entre um número de variáveis selecionadas e a demanda. A informação sobre as variáveis preditivas é convertida em uma equação

de regressão a fim de proporcionar uma previsão de demanda. (AKABAY *et al*, 1995), (WANKE, 2004), (MOREIRA, 2008) e (JUNGES *et al*, 2011) são trabalhos que utilizaram e discutiram a aplicação da Regressão Múltipla para a previsão de séries temporais quando estas, além de dependerem unicamente dos valores passados, dependem de outras variáveis que estão disponíveis e podem ser utilizadas como co-variáveis na equação de regressão, visando obter resultados superiores aos utilizando apenas dados da própria série.

2.2.2 Mineração de Dados

A interpretação dos dados em uma série temporal é uma tarefa importante para se conhecer o comportamento da variável analisada e quais são as influências a que esta variável está sujeita. A mineração de dados (*Data Mining*) é uma atividade de grande valia no auxílio dessa interpretação.

O livro interdisciplinar (com foco em Mineração de Dados) de Wong e Leung (2002) integra três áreas de pesquisa: a mineração de dados propriamente dita, a programação de algoritmos genéticos e programação lógica indutiva (*ILP*, *Inductive Logic Programming*). Em essência, a mineração de dados consiste em extrair os conhecimentos de dados de forma válida, compreensível e interessante.

Há, naturalmente, muitos tipos de métodos que podem ser utilizados para extrair certo conhecimento de dados, principalmente a partir de métodos de aprendizado de máquina (um ramo da Inteligência Artificial) e estatísticas. A diferença entre essas áreas e a mineração de dados é um pouco confusa e controversa. Há uma interseção significativa entre a mineração de dados, aprendizado de máquina e estatística. Em geral, porém, pesquisadores de mineração de dados enfatizam mais a importância da descoberta compreensível e interessante do que pesquisadores de outras áreas.

Por "interessante", Wong referem-se ao conhecimento novo para o utilizador e potencialmente útil para tomada de decisões inteligentes. A *ILP* tem ganhado maior atenção na aprendizagem de máquina e na mineração de dados, devido ao uso de uma linguagem de representação de regras com maior poder de expressividade do que o utilizado pela maioria dos algoritmos de descoberta de regras. No entanto, a maioria dos atuais algoritmos *ILP* são baseados em otimização local. Assim, a integração das poderosas linguagens de representação da *ILP* com as

potencialidades de pesquisas de otimização global robusta, realizadas por algoritmos evolucionários, parece uma direção de pesquisa promissora.

Um artigo muito completo a respeito de métodos de clusterização aplicados em redes neurais e baseados em redes neurais, produzido por (K. L,-, Du, 2010), apresenta as vantagens e desvantagens dos métodos mais conhecidos, destacando principalmente a importância dos métodos no pré-processamento dos dados para os objetivos pretendidos usando redes neurais.

2.2.3 Modelos de Previsão

Modelos de séries temporais são aplicados de várias fórmas, dependendo do objetivo, como filtragem, predição de valores futuros, simulação ou modelagem de uma série temporal, de formal geral, que forneça informações novas ao problema em estudo (Brockell e Davis, 2002).

A maior parte dos métodos propostos até então, estão baseados na teoria tradicional utilizada para a análise de séries temporais, e como esses modelos são definidos considerando séries ou processos estacionários no tempo, o que nem sempre é satisfeito, por esse motivo, os dados são transformados para que adquiram tal característica. Após essa transformação, escolhe-se as variaveis de entrada considerando relações lineares entre as variáveis dependentes e independentes. Geralmente a modelagem é baseada em modelos lineares auto- regressivos (AR) ou auto-regressivos médias móveis (ARMA) (Box et al., 1994), (Hamilton, 1994).

A metodologia empregada para a construção de modelos de séries temporais é baseada numa sequência de etapas, na qual a escolha do modelo a ser utilizado é baseada nos próprios dados (Ballini, 2000).

A sequancia das etapas são (Morettin e Toloi, 1985):

1. Uma classe de modelos é definida para análise;
2. Escolha do modelo, com base na análise da autocorrelação, autocorrelação parcial e critérios de Informação Bayesiana (BIC);
3. Fase da estimação, na qual os parâmetros do modelo escolhido são estimados;
4. Verificação do modelo ajustado em relação aos dados da série, por meio da análise dos resíduos para concluir se este é adequado ou não para a previsão.

Se a conclusão for que o modelo escolhido nãe é adequado, o procedimento é repetido, incluindo novos modelos e/ou reavaliando os não escolhidos. De acordo com a viabilidade computacional, o procedimento pode ser utilizado para identificar mais de um modelo para ser estimados e verificado, assim, a fase crítica do procedimento é a identificação do modelo (Morettin e Toloi, 1985).

Essa metodologia e os modelos estatísticos usados na análise de séries temporais fornecem uma idéia da sequência de passos a seguir para a obtenção de modelos mais adequados, a principal desvantagem está dos modelos, que possuem capacidade de lidar apenas com informações lineares entre as variáveis, e na maioria dos casos reais, as relações são não lineares.

A sequência de passos aqui citada e normalmente seguida na construção de modelos estatísticos não eram muito aplicados no campo da inteligência computacional, mas nesses últimos anos têm surgido algusn trabalhos relevantes e direcionados nesse assunto, como é trabalho descrito em (Tsay, 2005), onde um modelo linear é substituido com por um modelo de rede neural multicamadas(vide Haykin, 2001b), sendo o modelo validado e analisado de forma similar aos modelos tradicionais ajustados para o mesmo problema, apresentado inclusive, o enfoque estatístico da rede neural e suas diferenças em relação as abordagens estatísticas tradicionais.

Nesse contexto, outros métodos na área de inteligência computacional para previsão de séries foram surgindo, mas mantendo o esquema descrito anteriormente, concluido sua performance baseado em comparações de ajuste entre os métodos abordados.

2.2.4 Técnica de Redes Neurais Artificiais aplicadas a previsão de séries temporais

Dada sua capacidade de extrair informações lineares e não lineares entre as variáveis utilizadas, diversas abordagens neurais vêm sendo utilizadas, e assim, obtendo modelos de séries temporais mais eficientes. Esses modelos vem apresentando resultados promissores quando aplicadas a previsão de séries temporais, nas mais diversas áreas como economia, física, engenharia (Nie, 1997) e recursos hídricos (Zealand et al., 1999), (luna et al., 2005).

Desta forma, estes modelos, tolerantes as incertezas que podem estar nos dados, assim como a adaptação a possíveis alterações da estrurura da série, os tornaram alternativas atraentes. Características particulares de uma série de tempo, tais como: sazonalidade, tendência e ciclo podem ser aprendidas por uma RNA, possibilitando, assim, a realização de previsões.

As RNAs têm sido bastante empregadas na predição de valores e identificação de séries temporais (WAN,1994). A rede neural MLP (*Multilayer Perceptron*) é a mais utilizada dentre as RNA existentes, apesar de usar em sua maioria um mapeamento estático de entrada-saída. A melhor previsão da instância do Concurso de Santa Fé, instância de dados disponível na *internet*, foi alcançada por Wan (1994), que construiu uma rede de Respostas de Finitos Impulsos (*FIR - Finite Impulse Response*) com filtros lineares. O erro médio quadrático normalizado (*NMSE*, *Normalized Mean Square Error*), para os 100 pontos (padrões) instâncias escolhidas, obtido com a previsão foi de NMSE (100) = 0,0273.

Almeida e Siqueira (2004) exploram o uso de RNA com método de retropropagação na previsão de falência de bancos brasileiros. Os resultados de classificação das redes são comparados com a técnica estatística de regressão logística. As RNAs mostraram algumas vantagens em relação à regressão logística como, por exemplo, consideração de instituições que foram desprezadas pela técnica estatística. A qualidade de classificação das redes, segundo Almeida e Siqueira, poderia eventualmente melhorar a partir da exploração de outros métodos.

Corrêa e Portugal (2004) apresentam em seu artigo uma avaliação empírica da capacidade preditiva de modelos de RNA e modelos estruturais de série de tempo (MEST) quando existe uma mudança estrutural aplicadas a séries geradas por simulação seguindo diferentes processos ARIMA com imposição de mudanças estruturais na média e tendência. Os resultados mostram uma capacidade levemente superior das RNAs em realizar previsões um passo à frente.

Martineli *et al.* (2005) descrevem o uso de algoritmos conexionistas de aprendizado, baseado em RNA, no problema da previsão de falências. Os dados sobre os bancos brasileiros são representados por 26 indicadores, esse trabalho enfatizou a importância da seleção dos atributos mais importantes para o problema visando melhorar o desempenho das RNAs e diminuir o esforço e tempo desperdiçado para escolher uma boa arquitetura.

No trabalho de Falco *et al.* (2005) é investigada a concepção eficaz de um modelo de RNA adequado para previsão de séries temporais com base em uma abordagem evolutiva. Utilizam os algoritmos genéticos para escolher a concepção de uma arquitetura de rede neural e a escolha do melhor método de aprendizagem. A principal característica desse trabalho é o fato de combinar as características topológicas das redes que resultaram nas maiores performances na definição de novas topologias a serem testadas, o que pode resultar na escolha da melhor topologia num menor tempo computacional.

Outra alternativa as redes neurais aplicadas a séries temporais têm sido a chamada combinação modular, presente no trabalho de Pai e Lin (2005), onde o modelo ARIMA é utilizado para prever a parte linear do problema, depois uma rede neural tenta melhorar à previsão final, atuando sobre os resíduos do primeiro passo, outro trabalho semelhante a este pode ser visto em (Lemos & Fogliato, 2008).

Buscando melhorar previsões mais complexas, mas que possuem um certo grau de sazonalidade, (Wong *et al, 2010*) combinaram as saídas de duas redes neurais, uma com os dados brutos e a outra com os dados filtrados usando funções trigonométricas, que juntas, são inseridas em uma outra função ajustada que combina linearemente os resultados de forma a minimizar o erro na saída.

Nesse contexto, Santos (2012), em sua tese, criou o que ele chama de Ponderação Neural de Experts (NEW – *Neural Expert Weighting*), que basicamente, aplica regressão linear múltipla, adotando como variáveis independentes as previsões obtidas pelas redes neurais, e a variável dependente o dado real da previsão. A justificativa principal do autor está no fato da previsão oriunda de uma combinação linear de várias outras previsões produz em média, erros menores do que a utilização de uma única topologia.

Outras técnicas baseadas em redes neurais foram utilizadas para a previsão, mas a partir daqui se dará enfoque na rede neural utilizada nesse trabalho, as redes neurais de bases radiais.

2.2.5 Redes Neurais Artificiais de Funções de Bases Radiais Aplicadas a Previsão de Séries Temporais

A configuração e aplicação de RBF para o procedimento de previsão de séries temporais é o tema principal abordado por Coelho e Canciglieri (2001). A RBF é avaliada em dois estudos de casos: na previsão da concentração de dióxido de carbono dos dados da fornalha de gás de Box e Jenkins e a previsão da equação diferencial de Mackey-Glass que descreve um sistema de controle fisiológico. As simulações tratam os procedimentos de estimação e validação do modelo neural.

Os resultados obtidos são encorajadores e serviram para constatar-se que a RBF constitui-se de uma ferramenta promissora em aplicações de previsão de séries temporais e mapeamentos não-lineares de sistemas dinâmicos complexos. A RBF demonstrou vantagens devido a capacidade de aproximação de funções não-lineares, rapidez e eficiência do aprendizado, possibilitando a obtenção de resultados precisos. Entretanto, segundo os autores, necessita-se de estudos mais aprofundados quanto a aspectos de aprimoramento das capacidades de interpolação, generalização e aprendizado da RBF para a previsão de séries temporais.

O artigo (ROJAS *et al*., 2004.a) descreve uma nova estrutura para criar uma RBF. Esta nova estrutura tem 4 características principais: em primeiro lugar, a arquitetura de rede especial RBF usa pesos de regressão para substituir os pesos constantes normalmente utilizados. A segunda característica é a normalização da ativação de neurônios ocultos (média ponderada) antes de agregar as ativações, que, como observado por vários autores, produz melhores resultados do que a arquitetura clássica soma ponderada. O terceiro aspecto é que um novo tipo de função não-linear é proposta: a função pseudo-gaussiana (PGBF).

Com isso, a rede neural obtém ganhos de flexibilidade, os neurônios possuem um campo de ativação que não tem necessariamente de ser simétrico em relação ao centro ou a localização do neurônio na camada de entrada. Além desta nova estrutura, Rojas propõem como a quarta e última característica, um algoritmo seqüencial de aprendizagem, que é capaz de se adaptar a estrutura da rede, com isso, é possível criar novas unidades escondidas e também de detectar e remover unidades inativas.

Outro artigo (ROJAS *et al.*, 2004.b) apresenta uma proposta de construção e formação de uma RBF para previsões de curto prazo. A estrutura das funções gaussianas é modificada usando uma função pseudo-gaussiana, no qual dois parâmetros sigma de dimensionamento são introduzidos, o que elimina a restrição de simetria e fornece os neurônios na camada oculta com uma maior flexibilidade no que diz respeito à função de aproximação. O desempenho superior da técnica proposta

sobre o sistema RBF padrão são apresentados com testes com o problema de previsão de curto prazo de séries temporais caóticas.

Um algoritmo genético com o objetivo de criar uma arquitetura para uma RBF é apresentado por (YEN, 2004). Neste trabalho, o autor apresenta a idéia de criação de um ranking hierárquico de densidade do algoritmo genético (HRDGA) o qual é usado para topologia de evolução da rede neural e seus parâmetros. Além disso, o fitness do AG é utilizado para otimizar o desempenho e a topologia da rede neural e para lidar com o conflito entre a formação, desempenho e complexidade da rede. Em vez de produzir uma única rede neural ideal, o HRDGA fornece um conjunto de Redes Neuraisperto do ideal para os tomadores de decisão para que eles possam tem mais flexibilidade para a decisão final com base em suas preferências. As RBF concebidas pelo algoritmo proposto prova ser competitivo, ou mesmo superior, a RBF para funções de Mackey-Glass na previsão de séries temporais caóticas.

Outra aplicação para as RBF são os estudos de modelagem de tráfego, onde nas quais se forem feitas considerações prévias sobre suas características podem ser problemas com a aplicação de predição. Entre estes modelos estão as Redes Neurais e modelagem *fuzzy* de (Vieira et al., 2004).

Em muitas aplicações de mineração de dados para problemas de classificação, os modelos são considerados como tarefas-chave. Isto é, as características adequadas de entrada do classificador devem ser selecionadas a partir de um determinado (e muitas vezes grande) conjunto de características possíveis da estrutura e os parâmetros do classificador devem ser adaptados com respeito a estas características e um determinado conjunto de dados.

O artigo e Buchtala, Klimek e Sick (2005) descreve um algoritmo evolutivo que executa a função de seleção simultânea do modelo para classificadores de função de base radial. A fim de reduzir o esforço de otimização, vários técnicas são integradas que acelerar e melhorar os algoritmos evolutivos significativamente: de formação híbrida de Redes Neurais RBF, avaliação preguiçosa, entre outras. A viabilidade e os benefícios da abordagem são demonstrados por meio de quatro problemas de mineração de dados, detecção de intrusões em redes de informáticas, verificação de assinatura biométrica, para aquisição de clientes com métodos de marketing e otimização de processos na produção da indústria química.

É mostrado que, em comparação com algoritmos anteriores, nos algoritmos evolutivos baseado em técnicas de otimização com funções de bases radiais o tempo

é reduzido em até 99%, enquanto as taxas de erro são reduzidas em até 86%, dependendo da aplicação. O algoritmo descrito é independente das aplicações específicas, de modo que muitas idéias e soluções podem ser transferidas para outros paradigmas para a classificação.

Duas dificuldades estão envolvidas com RBF tradicional: a configuração inicial de uma rede RBF que precisa ser determinada por um ensaio e certo método de erro e a degradação que o desempenho sofre quando a localização pretendida dos centróides não é adequada. O artigo de Song, Yu e Chen (2007) propõe alternativas para superar estas dificuldades. A nova função de base radial é usada na camada oculta, onde o número de nós é determinado automaticamente pelo teorema da amostragem de Shannon.

O algoritmo de aprendizagem correspondente geralmente leva muito menos tempo de aproximação com uma configuração de parâmetros otimizados. As localizações dos centróides da RBF são fixas. Resultados experimentais têm demonstrado que as redes RBF construídas pelo método proposto no artigo de Song, Yu e Chen têm um menor número de nós, uma velocidade mais rápida de aprendizagem e um menor erro de aproximação do que as redes produzidas por outros métodos.

O trabalho de Dias (2007) propõe a utilização de técnicas de previsões de séries temporais, para a estimação do consumo, respeitando a sazonalidade dos produtos. Dias (2007) utiliza 7 (sete) séries temporais reais, que representa a demanda dos produtos alimentícios de uma empresa brasileira, com o objetivo de ajustar o melhor modelo para que as previsões sejam feitas a curto e a longo prazo, diminuindo assim incerteza no planejamento estratégico da empresa.

O trabalho contempla estudos de previsões de series temporais através das RBF e compara seus resultados, com as previsões dos modelos ajustados através da Metodologia Box & Jenkins. Para efeito de comparação dos dois métodos Dias (2007) utiliza a medida do RMSE (Raiz do Erro Quadrático Médio). Após estudos e comparações a conclusão do trabalho de Dias (2007) é que as RBF mostram-se um ferramental com grande robustez e consistência para ser utilizada na previsão de demanda de series temporais.

O artigo de Gonzalez et al. (2007) apresenta um algoritmo evolutivo multiobjetivo para otimizar as RBF a fim de aproximar funções de um conjunto de pares de entrada e saída. O procedimento permite a aplicação de heurísticas para

melhorar a solução do problema em questão, incluindo alguns novos operadores genéticos no processo evolutivo. Estes novos operadores são baseadas em duas conhecidas transformações de matriz: decomposição em valores singulares (SVD) e o método de mínimos quadrados ordinários (MQO), que têm sido utilizados para definir novas operações de mutação que produzem modificações locais ou globais nas funções de base radial das Redes Neurais(os indivíduos da população no processo evolutivo). Depois de analisar a eficiência dos diferentes operadores, Gonzalez et al. (2006) mostra que operadores de mutação globais tem um rendimento melhor no procedimento para ajustar os parâmetros da RBF.

K. Meng e outros (2008) propuseram o uso de método k-médias nebuloso na etapa não supervisionada da RBF aplicadas visando a melhoria da a previsão de séries temporais do custo de energia elétrica e em séries conhecidas como a Mackey Glass.

Usando algoritimo genético em RBF, R.J. Kuo e outros (2009) combinam as características topológicas das RBF's que obtinham melhores resultados preditivos na criação de uma nova rede RBF, que após treinamento, obteram resultados ainda melhores que os anteriores, o método proposto foi confrontado com os métodos mais tradicionais e a RBF na forma básica.

Visando obter melhorias do aprendizado local das redes RBF sem prejudicar o desempenho do aprendizado local, (L.J. Herrera *et al*, 2011) propuseram um pós processamento, evitando a sobreposição dos campos receptivos dos centros, alterando as coordenadas do centro e a largura do campo receptivo de maneira controlada.

Zhou, P. e outros faz algo semelhante ao Herrera, porém, faz uso do método dos mínimos quadrados ortogonais para estimar o número de centros adequados e propondo uma heurística que altera o referido método, levando em consideração a complexidade dos dados, a largura do campo receptivo e o número de centros na criação de um índice, cujo valor ótimo define a topologia.

WANG e outros (2005) propuseram um conceito da significância de um neurônio, definido como a contribuição estatística de um neurônio para o desempenho global da rede. Tal conceito era utilizado para a determinação do número de neurônios na composição da topologia da rede. Um novo neurônio é adicionado se a sua contribuição é maior do que um limiar escolhido. Inversamente, se o significado de um neurônio é inferior ao limite, o neurônio é eliminado. (LEE, 2010) adotou o conceito de

(WANG *et al*., 2005), através do emprego de uma expressão heurística denomina pelo autor de M-estimador, como a contribuição estatística, o método proposto foi aplicado na série Mackey Glass com inserção de vários níveis de ruído.

Com obtejivo semelhante (Min Gan, et. al. 2012) usaram algoritmos genéticos na definição da topologia da rede e na escolha escolha de um sub conjunto de dados de de treinamento, para isso, definiram cada cromossomo como essas caraterísticas e seu respetvivo fitness como o erro no teste de previsão, e em cada geração, os melhores são escolhidos para combinarem na geração de novos cromossomos. A diferença fundamental desse método mora no fato que ele não testa topologias de maneira incremental, por tentativa esse erro.

2.2.6 Comparações Entre Métodos de Previsão

Com o intuito de melhorar os métodos de previsão de séries temporais, muitos trabalhos passaram a comparar dois ou mais métodos de previsão e tentar assim encontrar uma que forneça melhores resultados de previsão, gerando menores erros.

O artigo de Souza e Zandonade (1993) é principalmente um exercício empírico, destinado a comparar o desempenho de previsão através de RNA e métodos estatísticos tradicionais de previsão de séries temporais, como ARIMA e modelos de componentes não observáveis. No trabalho de Souza e Zandonade foram usados modelos estruturais de série de tempo e modelos ARIMA para fins de previsão, além das RNAs. Os autores indicam que modelos baseados em RNAs parecem produzir previsões mais precisas. Entretanto, os autores argumentam, ainda, que os modelos ARIMA não são uma referência adequada para a comparação com os outros modelos.

A principal contribuição do trabalho de Yamazaki (2006) em problemas de previsão de séries temporais foi a aplicação da metodologia proposta por Zandonade (1993). Foram apresentadas duas técnicas de otimização global, *Simulated Annealing*(*SA*) e *Tabu Search*(*TS*), duas abordagens híbridas envolvendo as técnicas e o conhecido algoritmo de otimização local *resilient backpropagation*ou *RPROP*. O *RPROP* apresenta a vantagem de convergir mais rapidamente que o *backpropagation*. O uso do algoritmo *RPROP* também é um diferencial deste trabalho com relação ao de Yamazaki, que utilizou o *backpropagation* como método de convergência local.

A aplicação das metodologias de treinamento híbridas mostrou-se mais eficiente que o treinamento realizado apenas com o uso do *RPROP*. Foram mostrados os resultados dos treinamentos executados com a abordagem híbrida para duas séries temporais, nele é possível comparar os resultados entre as abordagens com otimização e sem otimização. Para ambas as séries, as redes otimizadas obtiveram melhor desempenho que as redes que não passaram pelo processo de otimização. Ainda com base nos resultados obtidos, é possível também comparar as abordagens híbridas entre si.

Zandonade verifica que os treinamentos com *Tabu Search* foram, em média, sempre mais eficientes que os treinamentos com *Simulated Annealing*. Além disso, foi abordado um problema bastante comum na área científica (previsão de séries temporais), a fim de contribuir um pouco mais com esta área. A metodologia de otimização utilizada para a previsão no trabalho de Zandonade já foi usada com sucesso em problemas de classificação em outros trabalhos, fortalecendo a sua eficiência na otimização das Redes Neurais Artificiais.

O artigo de Portugal (1995) apresenta um exercício empírico de previsão econômica através de métodos tradicionais da análise de séries temporais, como ARIMA, modelos de decomposição em componentes não observáveis (UCM) e RNA. Portugal utiliza os dados brutos mensais da produção industrial para o estado do Rio Grande do Sul (Brasil) para realizar uma comparação e avaliar o desempenho relativo aos diferentes métodos de previsão. Os resultados mostraram que a RNA obtém uma previsão mais precisa do que os modelos ARIMA, mas a comparação com UCM não é tão simples. A UCM encontra uma previsão melhor que a previsão da RNA, mas o desempenho da RNA para horizontes maiores mostra previsões que, especialmente depois que uma metodologia adequada de modelagem foi estabelecida, pode ser uma ferramenta valiosa a previsão econômica.

A aplicabilidade de modelos de previsão de séries temporais como ferramenta de decisão de compra e venda de contratos futuros de boi gordo, café e soja na BM&F (Bolsa de Valores, Mercadorias e Futuros), em datas próximas ao vencimento é o foco do trabalho de Bressan (2003). Os modelos estudados em seu trabalho são: ARIMA, Redes NeuraisArtificiais, e Modelos Dinâmicos Lineares (MDL). Os dados correspondem às cotações semanais, nos mercados: físico e futuro, entre 1996 e 1999. O objetivo consiste em calcular os retornos médios dos modelos em operações no mercado futuro de boi gordo, café e soja, de modo a indicar o potencial ou limitação

dos modelos, utilizando o índice Sharpe (índice que indica a relação retorno/risco) como parâmetro de comparação.

Os resultados apresentam retornos financeiros positivos na maioria dos contratos analisados, indicando o potencial de utilização desses modelos como ferramenta de decisão em negociações de contratos para datas próximas ao vencimento, com destaque para operações fundamentadas nas previsões nos MLD e ARIMA havendo, contudo, diferenças de desempenho preditivo. Com base nos resultados obtidos, Bressan conclui que, para o período analisado, o modelo com melhor desempenho simulado nos três mercados é o ARIMA que, em função de sua rápida adaptabilidade e estrutura parcimoniosa, produz as melhores previsões em termos agregados, com médias positivas nas simulações de compra e venda de contratos futuros das três commodities.

Os Modelos Lineares Dinâmicos também apresentam desempenho satisfatório, com retornos financeiros positivos derivados de sinalizações corretas da tendência de mercado, em especial no contrato futuro de boi gordo. Já os modelos de Redes NeuraisArtificiais geram resultados financeiros positivos nos mercados de boi gordo e soja, captando com certa precisão as reversões de tendência nesses mercados. Entretanto, segundo o autor, a média negativa dos retornos para o contrato de café limitam a sua aplicabilidade como ferramenta de auxílio à tomada de decisão de agentes que transacionam esta *commodity*, que se notabiliza pela alta volatilidade de preços no mercado físico.

Observa ainda que a construção e ajuste dos modelos envolve um *trade-off* em termos da adaptação do modelo estimado à série de dados e seu poder de previsão. Ademais, os modelos Redes Neurais requerem um elevado grau de subjetividade na interpretação das variáveis envolvidas no processo de modelagem. Por esta razão, o potencial de utilização de cada um dos métodos está associado à facilidade ou dificuldade na compreensão e aplicação dos mecanismos que determinam sua estimação, por parte dos agentes tomadores de decisão.

Um estudo comparativo entre métodos estatísticos e RNAs foi apresentado por Lima e Almeida, em 2008, visando explorar a possibilidade de usar uma metodologia capaz de decompor uma série temporal, via Wavelets, conjuntamente com os modelos já existentes de previsão e comparar a qualidade das previsões obtidas. Os resultados mostraram que as RNAs possuem um desempenho superior quando há períodos de menor volatilidade no mercado financeiro.

Um artigo intitulado “Rumo a previsão automática utilizando redes neurais” comparou o método batizado de *Regression Neural Network* (GRNN), o autor, Yan W. (2012) cita que o referido método obteve o melhor resultado numa competição diante cerca de 60 modelos de previsão diferentes apresentados por estudiosos de todo o mundo.

2.3 SEGURANÇA DE GRANDES ESTRUTURAS

Obras de grande porte são concebidas para os mais diversos fins. Entre esses podem ser citadas as usinas hidrelétricas, metrôs, túneis, minas, barragens, estádios, entre outras. Todas têm um papel vital na sociedade atual. Por outro lado, os riscos inerentes à criação e manutenção dessas obras existem e nem sempre conhecidos.

A monitoração de estruturas deve ser feita constantemente, através da observação de pontos estratégicos tanto na obra em si como fora dela. A correta instrumentação no controle de estruturas tem o objetivo de aumentar a segurança, minimizar prejuízos econômicos e danos ao meio-ambiente, além de salvar vidas.

Os próximos tópicos do trabalho citam dois casos de importância nessa pesquisa: as barragens e as dutovias.

2.3.1 Barragens

A construção de barragens para os mais variados fins acontece há longa data, mas a oficialização de quesitos sobre a segurança deste tipo de obra vem acontecendo desde um intervalo de tempo muito menor. Apenas alguns estados dos EUA criaram algum tipo de regulamentação a respeito antes de 1900 (Reed 1987, Walz 1990a).

À medida que a engenharia evoluía, a concepção de barragens se tornou mais ousada, aumentando seu potencial de perigo e a ocorrência de acidentes devidos a falhas de projeto, ocasionado desastres gigantescos (ARMY, 1996) como, por exemplo, a ruptura abrupta da barragem de Teton Idaho, EUA, ocorrida em 1976, com a perda de 11 vidas e muitos milhões de dólares (Teton Dam Failure, 2005).

Segundo Gutiérrez (2003), o histórico de rupturas de barragens revelou que um longo período de operação normal das obras não é garantia de condições futuras de

segurança, uma vez que tem havido casos de ruptura brusca após 10 e 20 anos de operação normal.

Tais problemas fizeram com que profissionais de diversas áreas do conhecimento viessem a discutir o tema segurança de barragens. Desta forma, em 1928, foi criado a ICOLD – *International Comission on Large Dams*, uma instituição não governamental que visa promover um fórum permanente de discussão e troca de conhecimento e experiências entre profissionais do mundo todo a respeito de engenharia de barragens. Atualmente, a ICOLD têm Comitês Nacionais em 83 países, incluindo o Brasil, onde é representado pelo Comitê Brasileiro de Barragens (CBDB). (AIVEC, 2005). O CBDB foi criado com a finalidade de ser um agente facilitador no processo de assegurar que a realização e a operação de barragens e hidrelétricas seja técnica, ambiental e socialmente adequada ao máximo benefício da sociedade brasileira (CBDB, 2006).

Desde a década de 1960, os temas de maior ênfase que a ICOLD tem abordado são relacionados à segurança de barragens, o seu monitoramento, re-análise da estabilidade de obras antigas, estudo de efeitos de envelhecimento e impactos ambientais gerados por barragens (AIVEC, 2005).

A cada três anos, o ICOLD promove o Congresso Internacional de Grandes Barragens, cuja primeira edição foi em 1933, na cidade de Estocolmo, na Suécia. Em cada congresso, o ICOLD lança algumas questões que podem ser respondidas pelos diversos profissionais em engenharia de barragens, através da publicação e apresentação de trabalhos técnicos. Nas últimas edições deste congresso houve questões relacionadas ao tema “Segurança de Barragens”, sendo que na vigésima e penúltima edição, ocorrida na cidade de Pequim, em 2000, três das quatro questões formuladas estavam intimamente ligadas ao tema de segurança e análise de risco em barragens (AIVEC, 2005).

Segurança de barragens se define pela capacidade para satisfazer exigências de comportamento, visando evitar a ocorrência de acidentes e/ou incidentes (RSB, 1990).

Segundo Cardia (2004), de acordo com Comissão Internacional das Grandes Barragens, acidente é qualquer situação que possa afetar a segurança e incidente é qualquer situação que possa afetar a funcionalidade da barragem. Também segundo o mesmo autor, ruptura de barragem é qualquer ocorrência na estrutura da barragem,

fundação, órgão de segurança ou reservatório, que provoque (para jusante), liberação não controlada de elevado volume de água.

De acordo com Saré, *et al.* (2006), o conceito de segurança deve ser entendido em um sentido global, envolvendo aspectos de natureza geotécnica, estrutural, hidráulica, operacional e ambiental.

Atualmente existem diversas diretrizes adotadas por diferentes países no que diz respeito à segurança de barragens. Um apanhado geral sobre as mesmas, incluindo comparações entre as diferentes metodologias empregadas na priorização de riscos e tomada de decisões relacionadas a barragens, visando sua segurança, podem ser encontradas em (HARRALD *et al*, 2004).

Dentre as metodologias existentes, duas delas se destacam: as diretrizes do ANCOLD (Australian Comittee on Large Dams), (ANCOLD), e da BCUC (British Columbia Utilities Commission).

As diretrizes do ANCOLD se baseiam no princípio ALARP, "*reducing risks as low as reasonably practicable*", que se fundamenta na obrigação legal que os proprietários de barragens têm de reduzir os riscos a valores bastante baixos, considerados aceitáveis (AIVEC, 2005).

Algumas, porém, estes tipos de critérios exagerados acabam por gerar gastos de manutenção e reforma em barragens desproporcionais aos benefícios gerados pela redução, às vezes insignificante, do risco de ruptura (BOWLES, 2003).

De acordo com Silveira (2004), alguns dos principais pontos sobre a avaliação da segurança de barragens são:

- Todas as barragens devem ser classificadas quanto às conseqüências de uma ruptura em potencial, onde devem ser considerados fatores como população a jusante, danos materiais, danos ao meio ambiente, danos à infra-estrutura, etc.;

- Devem ser inspecionadas periodicamente, para detectar eventuais deteriorações;

- Devem ser instrumentadas de acordo com seu porte e riscos associados, e terem seus dados analisados, através das leituras;

- Todos os instrumentos devem ser dotados de valores de controle ou limites;

- Todas as barragens devem ser submetidas periodicamente a uma reavaliação de suas condições de segurança, segundo sua classificação quanto às conseqüências de ruptura;

- As barragens deverão ser dotadas de um plano de emergência, objetivando a preservação das pessoas residentes a jusante, em caso de acidente.

Segundo Cardia (2004), os principais tipos de acidentes que ocorrem em barragens são o galgamento, a erosão interna e os sismos. O galgamento (*overtopping*) é a situação onde o nível de água do reservatório sobe muito por algum motivo, normalmente por vazão afluente elevada, e provoca a passagem da água por cima do topo da estrutura da barragem, de montante para jusante. A erosão interna (*internal erosion*) é a formação de vazios no interior de solo ou rocha mole, causada por efeito mecânico ou químico, de remoção de material, por percolação. Também é conhecida como *piping*, que é o desenvolvimento progressivo da erosão tubular interna por percolação, surgindo à jusante na forma de cavidade, descarregando água turva por carregamento.

Para Ramos e Melo (2006), as partes de uma barragem que merecem atenção com relação à segurança são principalmente as estruturas de desvio, os vertedouros e a descarga de fundo.

De acordo com (CRUZ, 2005), os principais meios de que o engenheiro dispõe para avaliar a segurança de um empreendimento ao longo de sua vida útil são: inspeções visuais (inclusive subaquáticas, quando for o caso), auscultação geodésica de deslocamentos verticais e/ou horizontais, levantamentos batimétricos, e instrumentação de auscultação.

No Brasil, algumas das principais entidades técnicas nacionais como ABMS (Associação Brasileira de Mecânica dos Solos e Eng. Geotécnica), CBDB (Comitê Brasileiro de Barragens), ABGE (Associação Brasileira de Geologia de Engenharia e Ambiental), Ibracon (Instituto Brasileiro do Concreto) e o Clube de Engenharia colaboraram na elaboração do texto da lei 12.334/2010, que após sete anos de tramitação, foi sancionada pelo presidente da República, no dia 21 de setembro de 2010. Esta lei **e**stabelece a Política Nacional de Segurança de Barragens destinadas à acumulação de água para quaisquer usos, à disposição final ou temporária de rejeitos e à acumulação de resíduos industriais, e define responsabilidades e atribuições a respeito do cuidado com a segurança das barragens brasileiras.

Em julho de 2012, o CNRH - Conselho Nacional de Recursos Hídricos publicou duas importantes resoluções que fazem parte do processo de regulamentação da lei 12.334/2010. Tratam-se das resoluções 143/2012 e 144/2012. A primeira trata dos critérios gerais de classificação de barragens por categoria de

risco, dano potencial associado e pelo volume do reservatório. A segunda estabelece diretrizes para implementação da Política Nacional de Segurança de Barragens, aplicação de seus instrumentos e atuação do Sistema Nacional de Informações sobre Segurança de Barragens (SNISB)**.**

Essas resoluções do CNRH seguem sugestões de engenheiros que atuam em projetos, construção e operação de barragens. Percebeu-se a necessidade de se criar regras que tornassem obrigatórios os procedimentos técnicos a serem executados antes, durante e depois da construção desse tipo de estrutura.(ITAIPU, 2012)

2.3.2 Dutovias

É notório o espaço que a indústria química vem ocupando como uma das principais atividades impulsionadoras do desenvolvimento econômico e social através dos benefícios da criação de seus produtos. Mas esse espaço começa a ser questionado perante a sociedade com o crescimento da consciência social sobre as questões ambientais. A tese de doutorado de Carlos André Vaz Junior (2010) relata uma série de acidentes amplamente divulgados pela mídia nas décadas de 1970 e 1980, que colocou o assunto ainda mais em evidência, como os acidentes como Sevezo na Itália em 1976, onde um vazamento de tetracloro-dibenzo-dioxina deixou mais de 200 mil feridos; Bhopal na Índia com 3800 mortos após vazamento de metil isocianato; Cidade do México com 490 mortos e 7000 feridos devido à explosão de GLP; e Cubatão com 500 mortos devido a um incêndio provocado pelo vazamento de hidrocarbonetos, estes três últimos acidentes ocorreram no mesmo ano: 1984.

Diante desses e outros casos, naturalmente ocorreu a necessidade de um maior aprofundamento das questões referentes aos riscos inerentes as atividades da indústria química, exigindo melhorias em seus programas de ações preventivas, além de um maior controle pelas entidades do governo. A legislação ambiental no Brasil vem se tornando cada vez mais rigorosa, impondo responsabilidades que antes seriam consideradas catástrofes naturais, no contexto de que quem atua numa atividade perigosa deve ser responsável direto por qualquer dano causado, (MARTINI, 2009) diz que essas empresas devem assumir todo o risco ou não exerce a função.

Assim, a responsabilidade em relação ao dano ambiental existe independente da culpa ou prática de ato ilegal (GUSMÃO, 2009). A Lei 9.605, de 12 de fevereiro de

1998 dispõe sobre as sanções penais e administrativas derivadas de condutas e atividades lesivas ao meio ambiente.

A indústria química, visando reduzir custo e aumentar sua agilidade, sempre procurou executar o transporte de matérias-primas e produtos por longas distâncias através de dutovias, dos mais diversos produtos, tais como petróleo, seus derivados e gases combustíveis. Dutovias são formas baratas, seguras, confiáveis e muito eficientes para o transporte de fluidos.

Embora confiáveis, dutovias empregadas no transporte de produtos tóxicos e/ou inflamáveis apresentam um perigo operacional inerente, ocasionado pela conjunção da operação sob pressão e vazões normalmente elevadas com a sempre possível manifestação de fadiga e colapso de material das tubulações, que dependendo de sua localização e magnitude, têm óbvia conotação desastrosa, gerando danos ambientais, danos materiais, perdas de inventário, incêndios, explosões e a possibilidade de perda de vidas humanas. Evitar tais falhas e, quando da sua ocorrência, minimizar a extensão de suas consequências, torna-se uma preocupação constante da indústria, agências reguladoras, e da sociedade como um todo (SILVA et al., 1996).

A legislação ambiental brasileira exige a implantação de sistemas de detecção de vazamentos em dutovias. Segundo a Resolução CONAMA nº1, de 23 de Janeiro de 1986, (CONAMA, 1986) a instalação de oleodutos e gasodutos depende da elaboração do Estudo do Impacto Ambiental, a Resolução nº 293 de 12 de Dezembro de 2001 (CONAMA, 2001) define a elaboração do Plano de Emergência Individual para acidentes que gerem vazamentos em instalações como dutos, terminais e plataformas.

Em uma dutovia construída e operada de forma adequada, falhas em sensores não irão ocasionar acidentes, porém podem causar problemas operacionais que resultem na interrupção da operação, o que acarreta prejuízos que alcançam facilmente a faixa dos milhões de reais, seja por lucro cessante ou por multas aplicadas por clientes não atendidos, o que torna as dutovias mais instrumentadas alvo desse tipo de problema, com isso, detectar falhas do tipo “anomalias na instrumentação” ou a ocorrência de vazamentos e outras grandezas antecipadamente são preocupações constantes na operação de uma dutovia (VAZ, C. A., 2010).

2.4 INSTRUMENTAÇÃO DE BARRAGENS E DUTOVIAS

Um sistema de monitoração, baseado em dados de instrumentação confiáveis, auxilia na prevenção de acidentes e incidentes (CRUZ, 2005). Geralmente, esse sistema é composto por diversos instrumentos, instalados em locais estratégicos do sítio de interesse (Dunnicliff, 1993). Tal prática é bastante comum em diferentes tipos de estruturas geotécnicas, tais como barragens (Silveira, 2003 e 2006), fundações (Reese et al, 2005) e taludes (Cornforth, 2005).

Os dispositivos de instrumentação são usados de forma suplementar às inspeções visuais para avaliar o desempenho e a segurança das operações. A monitoração cuidadosa dos dados da instrumentação em uma base continuada pode revelar uma condição crítica possível ou dar meios para assegurar que uma condição observada não é séria e não requer medidas corretivas imediatas (DEPARTMENT OF THE ARMY U.S. ARMY CORPS OF ENGINEERS, 1996).

“A instrumentação deve ser monitorada, analisada e mantida, para garantir a operação segura da barragem”. (COMISSÃO REGIONAL DE SEGURANÇA DE BARRAGENS, 1999).

O sistema de instrumentação é feito através de um plano que se baseia em um projeto obtido primeiramente através dos resultados do estudo detalhado das características geológico-geotécnicas da região e do entorno onde estará locada a construção, através do qual são definidas seções e blocos "chaves" a serem observados e instrumentados (GUTIÉRREZ, 1996).

A instrumentação monitora o desempenho estrutural e funcional e é instalada nos locais onde as condições complexas ou incomuns do local foram encontradas ou onde há uma probabilidade elevada de que a falha poderia resultar na perda de vida ou danos de propriedades extensivas (FEDERAL GUIDELINES FOR DAM SAFETY, 2004).

De acordo com Duarte (2006), entre as medidas necessárias para garantir um nível de segurança adequado para a obra, a instalação de um sistema de instrumentação geotécnica é uma das mais importantes, pois constituirá o meio de acompanhar durante a vida útil da obra se o nível de segurança se mantém dentro das premissas estabelecidas em projeto, permitindo, caso sejam detectadas anomalias, intervenções tempestivas para a manutenção da integridade e segurança da obra.

Segundo Cruz (2005), o instrumento ideal deveria ter as seguintes características:

- confiabilidade;
- alta durabilidade;
- não provocar, durante ou após a instalação, alterações no valor da grandeza que pretende medir;
- robustez;
- alta precisão;
- alta sensibilidade;
- não ser influenciável por outras grandezas, que não a de interesse;
- instalação simples;
- não causar interferência na leitura de outros instrumentos.

As principais razões para o uso de instrumentação em grandes obras de engenharia, segundo CELERI (1995), são:

A instrumentação é diferente para cada tipo de estrutura a ser instrumentada, as grandezas de maior importância a serem medidas variarão de acordo com o tipo de estrutura (DUARTE J. *et. al*., 2006).

As principais grandezas a serem medidas em barragens de concreto, segundo (SILVEIRA J.F.A., 2003), são: recalques, medidos por extensômetros e marcos superficiais; deslocamentos horizontais da crista, detectados por pêndulos diretos, marcos superficiais, inclinômetros, fitas de cisalhamento; distensão a montante, medida por extensômetros múltiplos; deslocamentos diferenciais entre blocos, detectados por bases de alongâmetros e medidores triortogonais; temperatura, medida por termômetros internos e de superfície.

No caso dos taludes/encostras/dutos, sabe-se que as poro-pressões do subsolo têm fundamental importância no conhecimento de sua dinâmica, pois a sua variação pode desencadear movimentações do terreno. Dessa forma, o monitoramento constante das poro-pressões se faz necessário para a previsão de possíveis deslocamentos do solo. Uma das formas mais utilizadas de se monitorar as poro-pressões *in situ* é através de piezômetros. Além dos piezômetros, é comum a utilização de inclinômetros nos taludes em estudo. Esses instrumentos são capazes de detectar os deslocamentos sofridos pelo terreno ao longo da profundidade, no

ponto (ou furo) em que os mesmos estão instalados. Também se faz uso dos pluviômetros, para medir a precipitação no sítio de interesse.

2.5 ANÁLISE DOS DADOS DE INSTRUMENTAÇÃO

O barateamento e o avanço da tecnologia micro-eletrônica permitiu que o número de equipamentos voltados para instrumentação de grandes estruturas aumentasse em variedade e os mesmos se tornassem mais viáveis financeiramente, o que permitiu aplicação de instrumentos em situações para as quais anteriormente seus custos eram proibitivos ou simplesmente não existiam. Um grande número de sensores, tais como de pressão, vazão e temperatura, passaram a ser empregados mesmo nas instalações mais simples.

O aumento do número total de sensores instalados nesses locais teve como consequência direta um aumento na quantidade de dados acerca do comportamento do processo. Essa quantidade de dados, se devidamente aproveitada, permite aumentar o controle operacional, reduzindo perdas, prevendo ou evitando acidentes.

A desvantagem do alto número de instrumentos é o aumento na probabilidade de ocorrência de falhas nos mesmos. Sensores que medem dados físicos podem resultar em leituras (dados) extremas (altos e baixos), o que pode ser consequência dessas possíveis falhas, levando a conclusões equivocadas a respeito da segurança da estrutura em análise. Além da interferência do correto diagnóstico acerca do estado da estrutura, dados com essas características também prejudicam a eficiência dos métodos computacionais e estatísticos, pois valores extremos interferem fortemente na média e variância estimada.

Esses dados extremos, chamados de *outliers*, podem ser resultado de alguns fatores (BUZZI M. F., 2004).

- Erro durante a medição: como muitas medições nesses instrumentos são feitas manualmente, pode ocorrer erro humano, ou o instrumento pode estar com algum tipo de falha;
- Eventos muito particulares: um dado atípico pode ser consequência de eventos que não representam a real situação ao longo de todo o tempo, por exemplo, uma obra sendo executada próxima ao instrumento de medição, ou um evento climático abrupto;

- Troca de instrumento ou manutenção: no momento da troca de um instrumento ou na sua manutenção, pode ocorrer uma mudança no marco inicial de medida. Situações como essa podem ocorrem durante a recalibragem de um instrumento, deixando os valores em intervalos diferentes dos originais, ou até valores muito diferentes no caso de uma substituição onde o anterior estava danificado;
- Ausência de leitura: aqui se considera ausência de leituras como um *outlier.* Alguns instrumentos podem estar localizados em pontos de difícil acesso, o que pode impossibilitar que o leiturista chegue ao aparelho caso ocorra uma adversidade, como por exemplo, chuva, deslizamentos, etc.

Em instrumentos automatizados, pode ocorrer falha momentânea na comunicação, seja por quebra de cabos, ou na comunicação via rádio.

Para identificar esses *outlier's* (menos no caso de ausência de leituras) será utilizado o Boxplot, ferramenta muito conhecida em estatística na análise exploratória dos dados, quando se deseja detectar possíveis *outlier's*.

O critério considera valores extremos (candidatos a *outlier's*) os que não pertence ao intervalo $[Q_1 - 3L, Q_3 + 3L]$, onde Q_1 e Q_3 são o primeiro e terceiro quartil respectivamente, e L é a diferença entre o quartil 3 e quartil 1.

Com o objetivo de se garantir uma operação eficiente de qualquer processo, é importante promover prontamente a detecção de equipamentos que apresentem mau funcionamento (KANO et al. , 2000).

Se por um lado a presença de um número elevado de sensores em uma instalação industrial eleva o número anual de falhas, essa grande quantidade de instrumentos pode gerar dados redundantes e é exatamente essa redundância que vem sendo fortemente explorada por diferentes metodologias de detecção de anormalidades (BUZZI M. F., 2004). As séries temporais geradas pelos sensores apresentam-se em geral altamente correlacionadas quando obtidas sob condições normais de operação. Tal correlação advém dos fenômenos físicos e químicos que governam a operação do processo, tais como transferência de massa e energia. Consequentemente, a falha de um instrumento pode ser caracterizada pela perda de correlação com os demais sensores, permitindo a identificação da origem da falha (VAZ, C. A., 2010).

Além da mudança na correlação, as previsões de séries temporais podem ser utilizadas para o mesmo fim. Caso a previsão obtida fique muito distante do valor que

efetivamente ocorreu, significa que o modelo de previsão está equivocado ou o instrumento está em falha, ou um fenômeno pontual ocorreu. Alguns estudos sobre previsão de leituras de instrumentos de obras instrumentadas já foram realizados, entre eles podem ser citados (CARVALHO J. V., 2005) e (WAHL T.L., 2004).

2.6 TÉCNICAS PARA A PREVISÃO DE LEITURAS EM SISTEMAS DE INSTRUMENTAÇÃO

O processo de previsão de leituras da instrumentação constitui-se numa das importantes atividades em um sistema de segurança e, diante dessa perspectiva, é considerada a base do processo do planejamento de ações futuras. Toda organização deve proceder de maneira cuidadosa e responsável em relação ao processo em questão, pois uma previsão precária pode acarretar em decisões equivocadas ou desnecessárias, podendo gerar sérios prejuízos. Uma previsão acertada fornece a possibilidade aos gestores da segurança de tomarem uma decisão correta no momento certo, com as informações que estiverem disponíveis.

Dada as características inerentes aos sistemas de instrumentação de obras geotécnicas, a previsão de séries temporais a partir dos dados de instrumentação encontra dificuldades, como por exemplo, imprecisão dos dados, irregularidades no intervalo de tempo entre leituras (BUZZI M. F., 2004).

Outro problema comumente encontrado é a falta de conhecimento acerca das relações existentes entre as variáveis, pois com estas, seria possível o uso de co-variáveis, o que pode aumentar a precisão das estimativas dos métodos. Nesse sentido, vários trabalhos foram desenvolvidos especificamente nesse campo, como o trabalho da Villwock (2009), em sua tese de doutorado, que aplicou análise fatorial nos dados de instrumentação da usina hidrelétrica de Itaipu e comparou os resultados com uma nova técnica baseada em colônias de formigas. O principal resultado foi a formação de grupos de variáveis que são relacionadas, hierarquizando a importância das variáveis numa análise.

A correlação dos dados nem sempre ocorre de forma imediata, por exemplo, a água da chuva leva algum tempo para se infiltrar no solo e provocar em aumento da poro-pressão detectável por piezômetros. Este intervalo de tempo para a infiltração da água deve ser considerado na análise de correlação das leituras piezométricas e

pluviométricas. Dyminski et al (2006) desenvolveram um trabalho sobre a previsão de leituras piezométricas baseadas em leituras anteriores destes instrumentos e dados atrasados de pluviometria a respeito de um dos taludes a serem estudados.

Informações sobre as características do sítio também podem fornecer co-variáveis importantes para os métodos usados nas análises de séries temporais. No Brasil, já foram desenvolvidos trabalhos de mapeamento geotécnico usando diferentes abordagens, tais como interpolação linear, *splines*, vizinhança mais próxima, geoestatística e redes neurais, a fim de caracterizar tridimensionalmente sítios com formações bastante distintas. O maciço da Usina Nuclear de Angra-2, no litoral do Rio de Janeiro, foi mapeado utilizando-se informações obtidas através de ensaios SPT (Romanel et al, 2003). O subsolo da Região Central de Curitiba, perfazendo uma área de 9 km^2, foi modelado por Miqueletto e Dyminski (2004). O subsolo do sítio do Porto de Navegantes, no litoral catarinense, teve sua estratigrafia e N-SPT mapeados tridimensionalmente nos trabalhos de Dyminski *et* al (2006) e Ribeiro et al (2007). Os resultados obtidos foram considerados satisfatórios e muito dependentes das heterogeneidades dos materiais envolvidos, bem como da quantidade e qualidade das informações disponíveis para o mapeamento.

Quanto às previsões de séries temporais de dados de instrumentação propriamente dita, vários trabalhos no Brasil foram desenvolvidos. Entre estes, pode-se citar Gutiérrez (2003), que preveu os dados da Instrumentação da Barragem Corumbá I por Redes Neurais e Modelos de Box & Jenkins. Um trabalho completo foi desenvolvido pela Janaina Veiga Carvalho (2005), que utilizando redes neurais e métodos estatísticos tradicionais fez a Modelagem temporal das medidas de vazão de drenos na Barragem de Funil (RJ), indicando a superioridade das redes neurais.

Outro trabalho semelhante, foi o executado por Sanches (2004), baseado em redes neurais artificiais (RNA's) e geoestatística, para a análise de leituras de instrumentos de auscultação da usina hidrelétrica de Itaipu, visando um melhor entendimento da distribuição espacial e espaço-temporal, usando dados de piezometria da barragem, avaliando leituras em diferentes regiões do maciço de fundação, procurando identificar tendências de leituras a níveis críticos. Redes neurais recorrente de Elman foram utilizadas por Romanel (2007) na previsão de séries temporais da barragem de Funil.

Como alternativa aos métodos tradicionais que não traziam resultados de previsão satisfatórios, (Kagoda, P. A. *et al*, 2010) aplicaram redes neurais de funções de base radial na previsão de vazão nas bacias na Africa.

O trabalho produzido por (Maier, R. H. et al, 2010) , apresentou o principais métodos baseados em redes neurais aplicado a previsão de recursos hídricos desenvolvidos nos últimos 15 anos, inclusive as RBF's.

Han H. G. et al (2011), usaram os pesos sinápticos obtidos no treinamento, e a variação do erro da aproximação como critério de poda ou incluisão de novos neurônios, seu método converge para a topologia procurada quando não há pesos abaixo de um valor definido heuristicamente, e o aumento de um centro RBF a mais aumente significativamente o aumento de precisão.

Mustafa et al (2012), usando redes neurais de funções de base radial para predição de pororessão de solos, adotaram como critério de definição da arquitetura mais adequada a função de autocorrelação da própria poropressão, e na correlação cruzada com pluviometria. Os autores compararam o desempenho da arquitetura previamente escolhida por esses critérios com o método exaustivo (teste de todas as topologias possíveis).

3 METODOLOGIA

Nesse capítulo estão descritas todas as etapas do novo método proposto, iniciando pelo pré-processamento dos dados, com intuito de eliminar possíveis dados duvidosos nas séries a serem previstas, seguido da remoção de tendências de crescimento e decrescimento da série temporal.

Com os dados sem *outliers* e tendências removidas, inicia-se a formatação dos dados de entrada e saída para o treinamento e teste das redes neurais utilizadas nesse trabalho.

Parte da nova proposta de modificação da rede neural de base radial reside na execução de métodos de agrupamento hierárquicos sobre os dados de entrada, antes da execução dos treinamentos das redes, identificando *a priori* topologias mais aptas ao conjunto de dados de entrada.

Com a lista de topologias aptas, inicia-se o treinamento da rede neural de base radial, e nesse momento surge outra parte da nova proposta, que é a utilização do dendrograma formado pelo método de agrupamento hierárquico na etapa do treinamento não supervisionado em substituição aos métodos de agrupamentos não hierárquicos.

A fim de validar o método proposto, são aplicados outros métodos baseados em RBF nas séries em estudo, e comparados com a previsão do novo método proposto, tanto em séries temporais reais quanto teóricas. O tempo de execução dos métodos também são confrontados. As próximas seções desse capítulo detalham cada etapa.

3.1 PRÉ- PROCESSAMENTO DOS DADOS

Buscando tornar os dados obtidos aptos para execução dos métodos de previsão, sejam esses métodos os já conhecidos na literatura ou no novo método proposto, algumas medidas devem ser tomadas. Nesta seção, descrevem-se quais foram os procedimentos adotados com esse fim.

Após a identificação de um *outlier*, por qualquer um dos motivos mencionados em 2.5, deve-se utilizar alguma metodologia para repor, substituir ou manter esse dado.

No caso de um dado faltante, esse deve ser estimado, pois sem esse, fica inviável a previsão de séries temporais das mesmas. No caso de um valor extremo ou duvidoso, esse não necessariamente deve ser substituído por um valor estimado, pois pode ser importante numa série e deve ser substituído somente caso seja confirmado como errôneo.

Pelas características inerentes ao estudo de caso, que possui várias leituras/dados obtidos em mesmas datas, é possível a utilização da regressão linear múltipla para interpolar esses dados faltantes ou atípicos, e como comparação e validação serão utilizadas comparativamente a *spline* cúbica.

A interpolação de leituras usando regressão linear múltipla só pode ocorrer a partir das séries temporais relacionadas linearmente em relação a série que terá seus dados interpolados, uma maneira de determina-las é calculando coeficiente de correlação amostral entre as séries.

Para estimar a correlação, serão utilizados n pares de valores (X, Y) , provenientes da população de todos os pares possíveis. Como há duas variáveis envolvidas, esta população é denominada bidimensional e muitas vezes ela apresenta uma distribuição normal bidimensional.

Assim, pode-se pensar no coeficiente de correlação de uma população teórica representada por ρ, que é estimado a partir do coeficiente de correlação amostral $\hat{\rho}$. O coeficiente de correlação amostral entre X e Y é definido pela equação (3.2).

$$\hat{\rho}(X,Y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \qquad (3.2)$$

onde $\bar{x}$ é a média aritmética das amostras de X , $\bar{y}$ é a média das amostras de Y.

Após o cálculo da correlação amostral entre todas as séries temporais disponíveis, localizam-se as cinco séries de maior correlação com a série temporal a ser estimada. A partir dessas aplicada-se a regressão múltipla, que é um procedimento que estabelece uma equação linear entre as 5 variáveis independentes (as com maiores correlações a série a ser interpolada) e a variável dependente (série à ser estimada). Esse cálculo fornece os valores a_1, a_2, a_3, a_4, a_5 e a_6 da expressão

(3.3), pelo método dos mínimos quadrados, sintetizado pelo cálculo da seguinte expressão (3.4).

$$\widehat{inst6}_x = a_1.inst1_x + a_2.inst2_x + a_3.inst3_x + a_4.inst4_x + a_5.inst5_x + a_6$$

ou

$$\widehat{inst6} = [inst1_x \quad inst2_x \quad inst3_x \quad inst4_x \quad inst5_x \quad 1]\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix} \tag{3.3}$$

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix} = \left(\begin{bmatrix} inst1_1 & inst2_1 & inst3_1 & inst4_1 & inst5_1 & 1 \\ inst1_2 & inst2_2 & inst3_2 & inst4_2 & inst5_2 & 1 \\ inst1_3 & inst2_3 & inst3_3 & inst4_2 & inst5_3 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ inst1_n & inst2_n & inst3_n & inst4_n & inst5_n & 1 \end{bmatrix}^T . \begin{bmatrix} inst1_1 & inst2_1 & inst3_1 & inst4_1 & inst5_1 & 1 \\ inst1_2 & inst2_2 & inst3_2 & inst4_2 & inst5_2 & 1 \\ inst1_3 & inst2_3 & inst3_3 & inst4_2 & inst5_3 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ inst1_n & inst2_n & inst3_n & inst4_n & inst5_n & 1 \end{bmatrix} \right)^{-1} . \tag{3.4}$$

$$\begin{bmatrix} inst1_1 & inst2_1 & inst3_1 & inst4_1 & inst5_1 & 1 \\ inst1_2 & inst2_2 & inst3_2 & inst4_2 & inst5_2 & 1 \\ inst1_3 & inst2_3 & inst3_3 & inst4_3 & inst5_3 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ inst1_n & inst2_n & inst3_n & inst4_n & inst5_n & 1 \end{bmatrix}^T . \begin{bmatrix} inst6_1 \\ inst6_2 \\ inst6_3 \\ \vdots \\ inst6_n \end{bmatrix}$$

onde, $\widehat{inst6}_x$ representa a estimativa do instrumento a ser interpolado, $inst1_x$, $inst2_x$, $inst3_x$, $inst4_x$ e $inst5_x$ representam os valores das leituras conhecidas na data da interpolação pretendida do $\widehat{inst6}_x$, os $inst1_1, inst1_2, \ldots, inst1_n$ são as n leituras disponíveis do instrumento 1 e assim respectivamente, $inst1_1$ e $inst2_1$ são leituras do instrumento 1 e instrumento 2 ocorridas na mesma data.

Para verificar a qualidade do ajuste da equação da regressão, adota-se o coeficiente R^2, dada pela fórmula 3.5. O resultado do R^2 varia entre 0 e 1. Quanto mais próximo de 1, melhor é a qualidade do ajuste. No trabalho, propõe-se aceitar a equação de regressão para $n \geq 30$ e $R^2 \geq 0{,}9$.

$$R^2 = \frac{\sum_{i=1}^{n}\left(\widehat{inst6}_i - \overline{inst6}\right)^2}{\sum_{i=1}^{n}(int6_i - \overline{inst6})^2} \tag{3.5}$$

Caso R^2 resulte num valor inferior a 0,9 ou $n < 30$, adota-se como interpolador a spline cúbica, que nada mais é que a interpolação dos pontos $inst6$ por meio de uma função definida por partes $\left(S_3(x)\right)$ com polinômios de grau três $(s_k(x), k = 1, \dots, n)$, $S_3(x)$ devem satisfazer a cinco condições impostas em (3.6):

1. $S_3(x) = s_k(x)$ para $[x_{k-1}\ x_k]$, $k = 1, \dots, n$;
2. $S_3(x_i) = inst6_i\ , i = 0, 1, \dots, n$;
3. $s_k(x_k) = s_{k+1}(x_k), k = 1, 2, \dots, (n-1)$; (3.6)
4. $s_k(x_k)' = s_{k+1}(x_k)', k = 1, 2, \dots, (n-1)$;
5. $s_k(x_k)'' = s_{k+1}(x_k)'', k = 1, 2, \dots, (n-1)$.

A condição 1 impõe que o valor retornado da função por partes seja o mesmo que o retornado pela função individual, a condição 2 garante que a função por partes passe pelos pontos reais, a últimas 3 condições garante a passagem de uma função para outra de maneira adequada, com continuidade e suavidade.

Com o polinômio definido, o valor de $\widehat{inst6}_x$ é calculado:

$$\widehat{inst6}_x = S_3(x) \tag{3.7}$$

Em linhas gerais, o fluxograma da figura 3.1 esquematiza os passos aplicados desde aquisição dos dados até a interpolação dos dados.

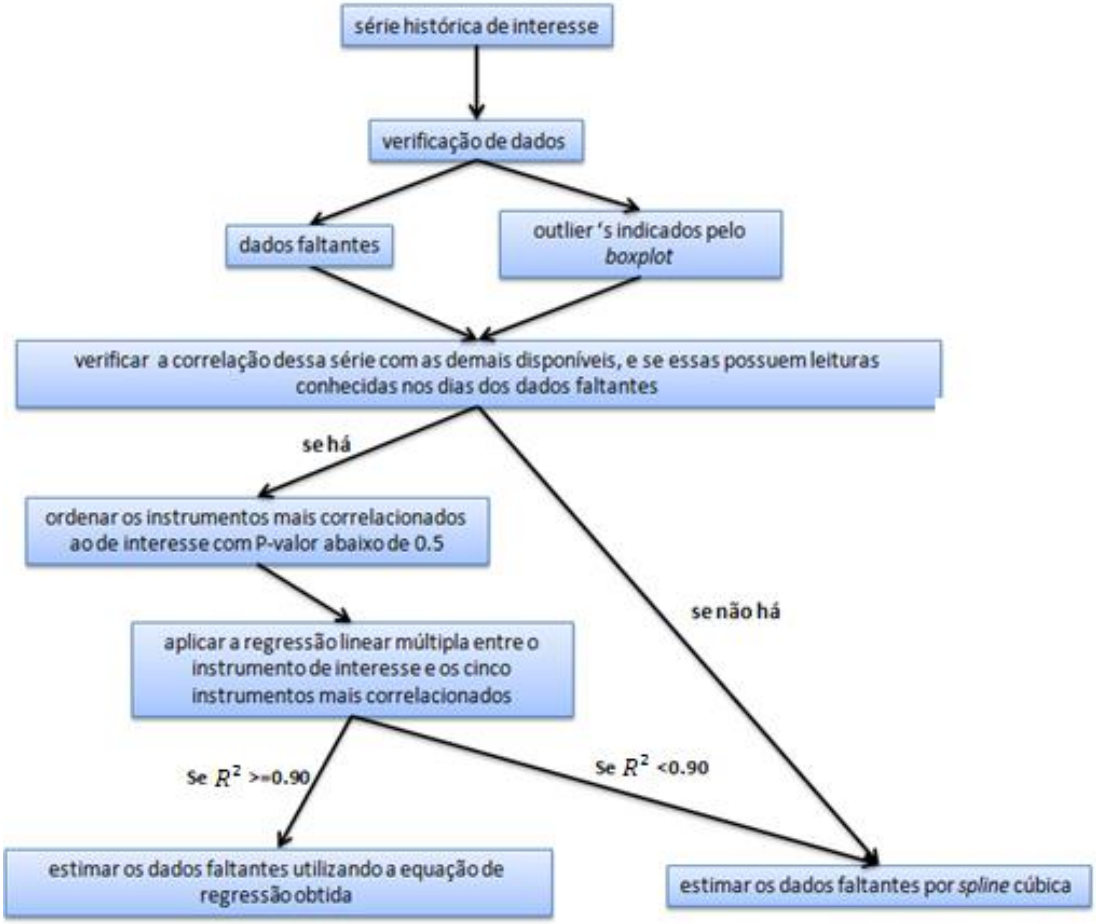


FIGURA 3.1: PASSOS EXECUTADOS NA RENOÇÃO DE OUTLIERS
FONTE: O AUTOR

3.2 REMOÇÃO DE TENDÊNCIA DE CRESCIMENTO E DECRESCIMENTO EM SÉRIES TEMPORAIS

As redes neurais não conseguem detectar tendências de crescimento e decrescimento em séries temporais, pois na etapa pré-processamento, os dados de entrada e saída são normalizados, o que faz com que uma rede treinada gere saídas apenas dentro desse intervalo normalizado. Caso ocorram entradas e novas saídas, fora do intervalo normalizado, a rede se comporta de maneira indesejada.

A fim de contornar esse problema, propõe-se o ajuste do conjunto de dados usando a funções do primeiro grau, e definir um novo conjunto de dados como sendo a diferença entre os dados originais e o valor da função ajustada no ponto da série. Realizando-se então o treinamento da rede com o novo conjunto de dados. A fim de ilustrar esse procedimento, suponha que uma série temporal seja composta pelos

dados discretos $v(i)$, $0 \leq i < m-1$, onde m é o número de dados da série temporal e i a referência temporal da série. Ajusta-se então uma função aos pontos $\big(t(i), v(m)\big)$, onde $t(i)$ representa o instante de tempo real associado ao valor da série na sequência i. Na figura 3.2, representa-se graficamente o exemplo de uma função linear ajustado ao conjunto de pontos.

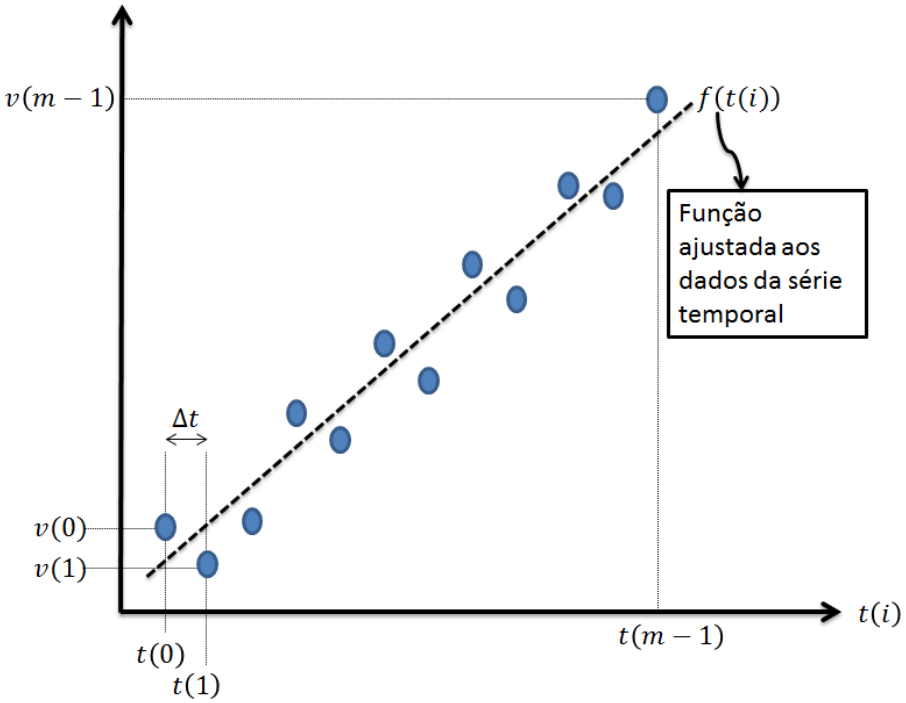


FIGURA 3.2: AJUSTE DE DADOS COM TENDÊNCIA LINEAR
FONTE: O AUTOR

Com a função definida, aplica-se a expressão $v'(i) = v(i) - f\big(t(i)\big)$, e novos pontos $\big(t(i), v'(i)\big)$ são definidos. Na figura 3.3, pode-se observar que a tendência de crescimento dos dados, estando os dados agora mais adequado para a utilização de redes neurais.

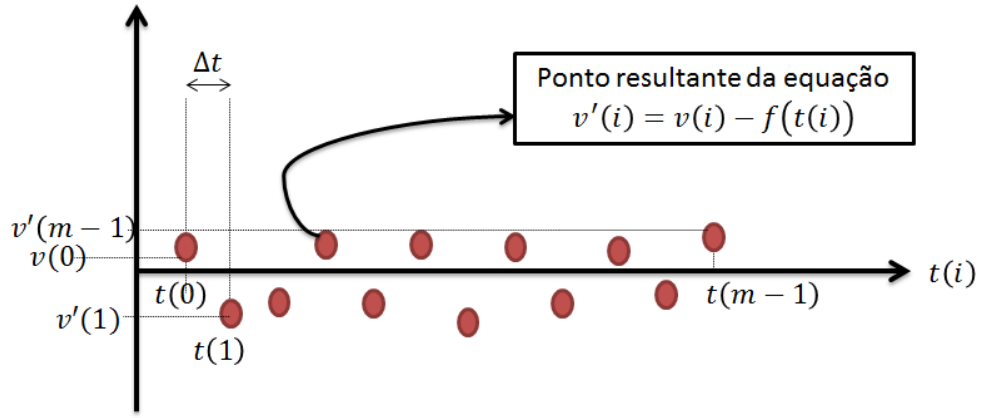


FIGURA 3.3: DADOS COM REMOÇÃO DE TENDÊNCIA LINEAR
FONTE: O AUTOR

Com esses dados transformados, as redes neurais serão treinadas, e independentemente da topologia utilizada e do formato dos dados de entrada. Para a obtenção de uma predição em relação aos dados reais, utilizou-se a equação (3.3)

$$aproximação\ real(i) = saídadarede(i) + f(t(i)). \quad (3.3)$$

3.3 PREPARAÇÃO E SELEÇÃO DOS DADOS DE ENTRADA E SAÍDA DA REDE NEURAL

Após a recuperação dos dados faltantes usando *spline* cúbica e/ou regressão linear múltipla, seguido da remoção de tendência dos dados, devem-se preparar os dados para a aplicação dos métodos de previsão propostos.

No caso da rede neural de base radial aplicada à previsão, a entrada e a saída necessárias para a execução do método podem ser organizadas de várias formas. Nesse trabalho, utilizou-se como padrão de cada entrada, n leituras ocorridas, e como saída desejada a leitura *n+1* prevista. Esse valor n é definido como o tamanho da janela de tempo usada na rede. A figura 3.4 abaixo ilustra um exemplo como se dá a preparação dos dados para o treinamento da rede utilizando uma janela de tempo igual a quatro. A primeira entrada na rede seria um vetor contendo as quatro primeiras leituras ocorridas, e como saída desejada a quinta leitura ocorrida; o segundo padrão de entrada seria da segunda leitura e assim sucessivamente, até que a quinta leitura seria a saída desejada referente a essa entrada, e assim por diante.

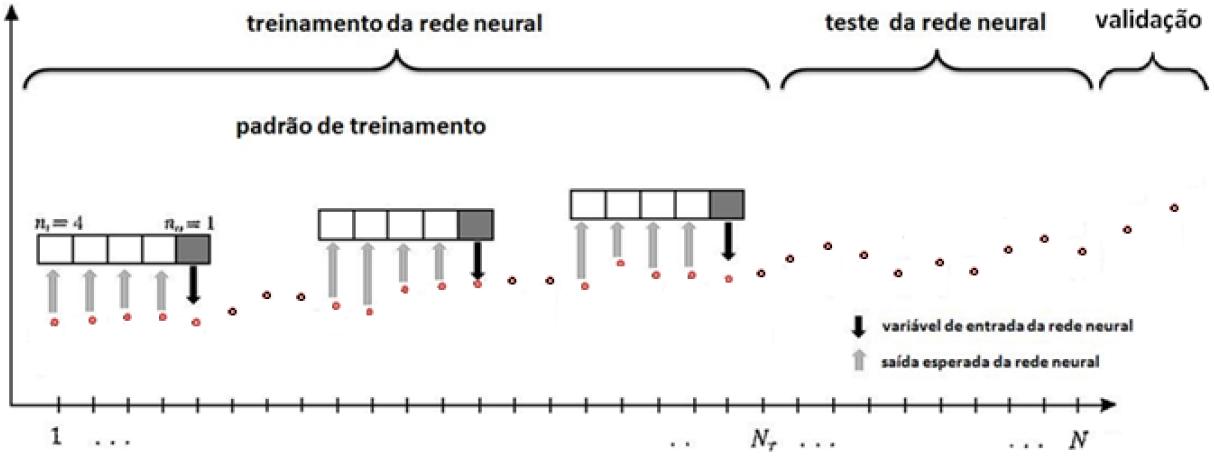


FIGURA 3.4: PREPARAÇÃO DOS DADOS PARA A REDE NEURAL
FONTE: O AUTOR

Nessa mesma figura, apresenta-se como se poderia dar a separação dos dados de teste e de validação para a rede. Os dados de teste não entram no grupo dos dados que ajustam os pesos da rede neural. São utilizados apenas como comparação com a saída da rede, com o intuito de verificar a capacidade de ajuste sobre dados não apresentados à rede durante o treinamento. Os dados de validação serão os últimos, que serão testados de maneira análoga ao conjunto de teste, visando verificar a capacidade de previsão da rede neural. Neste trabalho, não ocorre o agrupamento dos dados necessariamente na ordem que essa figura mostra, com exceção ao conjunto de validação, que sempre serão os últimos dados disponíveis, ou seja, os mais recentes, da série temporal.

São definidas várias porcentagens para a separação dos dados de treinamento e teste, e dada cada porcentagem, serão escolhidas aleatoriamente as entradas e saídas que irão compor os dados de treinamento e de teste.

Por ser definido por sorteio, haverá dados separados para treinamento e teste ao longo de toda a série. Com isso a rede pode "aprender" o comportamento da série ao longo de todo o domínio. Optou-se por esta estratégia porque, caso ocorra uma separação dos conjuntos como a figura 3.4 sugere, teria-se um treinamento em que a rede aprenderia bem o comportamento do trecho inicial da série, porém podendo ter um desempenho insatisfatório caso a série mostre uma mudança de comportamento de um período para o outro.

3.4 POSSÍBILIDADE DE INSERÇÃO DE CO-VARIÁVEIS AMBIENTAIS NA REDE NEURAL RBF

Séries temporais complexas podem ter sua previsão aperfeiçoada por um dado método, se esse levar em consideração outras variáveis relacionadas a série a ser predita em relação a um método que utiliza apenas a própria série.

Os eventos hidro-meteorológicos podem influenciar no comportamento da barragem e/ou talude e, conseqüentemente, no resultado da leitura de vários instrumentos. O princípio da coincidência de datas de leituras para se calcular a correlação direta entre dados de variáveis ambientais e as leituras dos instrumentos é o mesmo descrito em 3.1, ou seja, são inseridas nas análises apenas leituras ocorridas em mesma data.

Mas sabe-se que a influência de condições ambientais na leitura dos instrumentos nem sempre ocorrem imediatamente. Isto acontece pelo fato das estruturas em análise possuirem grandes dimensões, e os instrumentos estarem instalados em diferentes posições das mesmas, retardando de maneira diferenciada o tempo de resposta a modificações ambientais e, dependendo de sua posição, sofrendo maior ou menor influência destas variações. Algumas medidas ainda estão relacionadas com um acúmulo desses eventos, como por exemplo, a dilatação do concreto componente da barragem, que depende da frente de propagação de calor ao longo do corpo da mesma. Em uma parte interna da barragem, caso ocorra um dia de alta temperatura ambiente, não ocorrerá imediatamente tal dilatação. A propagação de calor ocorre de maneira lenta, e a dilatação resultante dependerá da quantidade de calor absorvida durante vários dias consecutivos.

Sabe-se que picos de temperaturas ou períodos de ocorrência de altas ou baixas temperaturas geram picos de deformação após algum tempo, detectados pela instrumentação em diferentes períodos de tempo. Desta forma, julgou-se importante quantificar este tempo de resposta à variação da temperatura e como esta variação influencia nas leituras dos instrumentos.

Para verificar a influência dos eventos passados nos resultados das leituras dos instrumentos da barragem, foram observados dois aspectos:

a) Correlação entre a leitura do instrumento e temperatura ambiente em dias anteriores:

Nesse tipo de análise, busca-se verificar a correlação entre a leitura de um instrumento e a temperatura ambiente de dias anteriores a essa leitura. Para isso, foram agrupadas em uma matriz, as leituras do instrumento de interrese, e a cada leitura desse instrumento, foi agrupada a temperatura de dias ateriores. Para cada dia de atraso abordado é calculada a correlação.

Um exemplo a seguir ajudará a entender a metodologia empregada.

Na figura 3.5, pode-se observar as leituras do instrumento base de alongâmetro JSF51, onde a coluna 2 mede abertura ou fechamento e coluna 3 deslizamento. Para as leituras de temperatura ambiente, coluna 2 mede a mínima, coluna 3 a média e coluna 4 a máxima. Caso se queira calcular a correlação entre determinada leitura de JSF51 e a temperatura do dia imediatamente anterior. Agrupa-se numa matriz de dados às leituras de JSF51 com a temperatura do dia anterior, e a partir dessa matriz, calculada-se a correlação.

Leituras JSF51		
...	...	...
5/2/1996	-2567	522
12/3/1996	-2565	505
16/4/1996	-2571	505
14/5/1996	-2551	530
11/6/1996	-2571	545
9/7/1996	-2552	598
12/8/1996	-2579	615
10/9/1996	-2580	640
15/10/1996	-2574	591
12/11/1996	-2550	561
10/12/1996	-2566	550
18/2/1997	-2566	505
11/3/1997	-2559	511
15/4/1997	-2571	529
12/5/1997	-2565	515
...	...	...

Leituras Tempetaturas			
...	...	...	...
2/2/1996	21,2	25,7	31,8
3/2/1996	23,7	25,5	27,8
4/2/1996	23,5	27,2	31,0
5/2/1996	20,2	25,0	31,0
6/2/1996	18,5	24,9	31,0
...	...	...	...
9/3/1996	21,6	27,2	33,6
10/3/1996	23,6	26,9	33,6
11/3/1996	21,9	26,4	31,6
12/3/1996	21,8	26,9	32,0
13/3/1996	20,0	24,4	31,0
...	...	...	...
13/4/1996	22,5	26,0	31,4
14/4/1996	22,0	26,0	31,4
15/4/1996	23,5	26,1	30,0
...	...	...	...

Matriz Agrupada				
...	...	...	...	...
-2567	522	23,5	27,2	31,0
-2565	505	21,9	26,4	31,6
-2571	505	23,5	26,1	30,0
...	...	...	..	...

FIGURA 3.5: EXEMPLO DE AGRUPAMENTO DE DADOS PARA O CÁLCULO DE CORRELAÇÃO ATRASADA.

Foi criada uma rotina no Matlab que permitiu calcular várias correlações, uma para cada dia de atraso, até 150 dias, verificando para qual atraso encontrou-se a maior correlação. Tal resultado poderia explicar qual é o atraso (*delay*) que um instrumento apresenta a responder à variação pontual da temperatura.

b) Correlação entre a leitura do instrumento e a temperatura ambiente acumulada de dias passados:

Nesse tipo de análise, busca-se verificar a correlação entre a leitura de um instrumento e a temperatura acumulada de dias anteriores a essa leitura, para isso, foram agrupadas em uma matriz as leituras do instrumento de interresse, e a cada leitura desse instrumento, foi agrupado o somatório da temperatura ambiente de dias anteriores. Para cada intervalo de acúmulo abordado é calculado a correlação.

Um exemplo que ajuda a entender a metodologia empregada é o descrito a seguir.

Na figura 3.6, pode-se observar as leituras do instrumento base de alongâmetro JSF51 e os da temperatura ambiente. Caso se queira calcular a correlação entre determinada leitura de JSF51 e a temperatura dos últimos três dias, agrupa-se a leitura do dia, e, como dado “temperatura” usa-se a soma dos últimos três dias (incluído o dia em questão) . A partir dessa matriz, é calculada a correlação.

JSF51 AB/F	
...	...
5/2/1996	522
12/3/1996	505
16/4/1996	505
14/5/1996	530
11/6/1996	545
9/7/1996	598
12/8/1996	615
10/9/1996	640
15/10/1996	591
12/11/1996	561
10/12/1996	550
18/2/1997	505
11/3/1997	511
15/4/1997	529
12/5/1997	515
16/6/1997	559
14/7/1997	558
...	...

Temp. Média	
...	...
2/2/1996	26
3/2/1996	26
4/2/1996	27
5/2/1996	25
6/2/1996	25
...	..
9/3/1996	27
10/3/1996	27
11/3/1996	26
12/3/1996	27
13/3/1996	24
...	...
13/4/1996	26
14/4/1996	26
15/4/1996	26
16/4/1996	20
17/4/1996	15
...	...

Matriz Agrupada		
...	..	...
5/2/1996	522	78
12/3/1996	505	80
16/4/1996	505	72
...	...	...

FIGURA 3.6: EXEMPLO DE AGRUPAMENTO DE DADOS PARA O CÁLCULO DE CORRELAÇÃO ACUMULADA

Foi criada uma rotina no Matlab que calcula várias correlações, uma para intervalo acumulado, até 150 dias, e verificando para qual intervalo acumulado resultou a maior correlação. Tal resultado pode explicar a influência acumulada da

temperatura ambiente nas leituras dos instrumentos, e quantificar o intervalo de tempo de temperaturas acumuladas que melhor se relaciona com as leituras dos instrumentos.

c) Correlação entre a leitura do instrumento e a temperatura acumulada de dias anteriores (janela de tempo deslizante)

É uma combinação dos itens a) e b), onde se verifica a correlação entre o acúmulo de temperaturas não necessariamente anterior à leitura, como mostra o exemplo da figura 3.7, com uma representação da correlação acumulada de três dias de temperaturas, contadas a partir de um dia de atraso em relação a data atual.

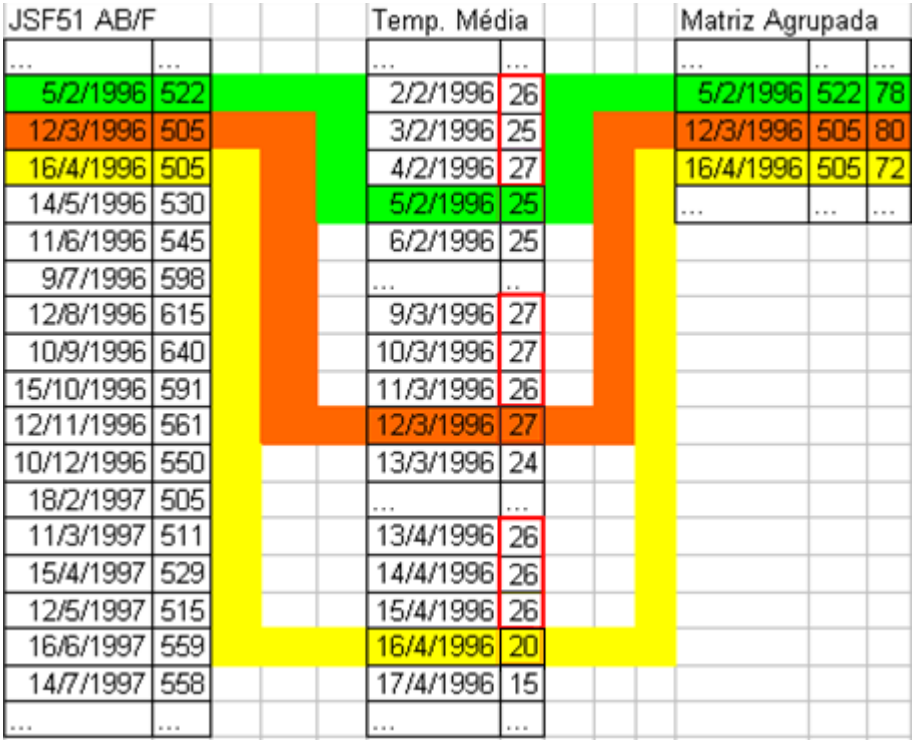


FIGURA 3.7: EXEMPLO DE AGRUPAMENTO DE DADOS PARA O CÁLCULO DE CORRELAÇÃO ATRASADA.

Foi criada uma rotina no Matlab que calcula várias correlações, uma para intervalo acumulado e atraso, até 150 dias, e verificando para qual intervalo acumulado resultou a maior correlação. Tal resultado pode explicar a influência acumulada da temperatura ambiente nas leituras dos instrumentos, bem como o atraso de reação desses dias acumulados.

3.5 AGRUPAMENTO HIERÁRQUICO APLICADO AO TREINAMENTO NÃO SUPERVISIONADO EM REDES RBF

Em redes neurais, geralmente são testadas várias topologias e, analisando-se os erros de treinamento e de teste para cada uma, escolhe-se a mais adequada, de acordo com algum critério pré-estabelecido. No caso das redes neurais de base radial, o mais comum é começar com um neurônio na camada oculta, treinando e verificando o erro de treinamento e teste. Os neurônios são incrementados um-a-um na rede, sendo que a cada neurônio inserido na rede, todo treinamento e teste são refeitos, ajustando-se os pesos sinápticos entre a camada de entrada e a camada interna (fase não supervisionada) e entre a camada interna e a da saída (fase supervisionada).

O método de agrupamento deve ser executado para cada topologia (número de neurônios) testada. Por exemplo, para uma rede neural de função de base radial com dois neurônios, durante a fase não supervisionada, se executa algum método de clusterização/agrupamento para classificação dos dados em dois grupos e consequentemente o centróide de cada grupo. Em uma rede neural de base radial com seis neurônios, durante a fase não supervisionada, se executa o método de clusterização/agrupamento para seis grupos.

Para essas duas topologias e outras que se coloque em teste, se propõe a substituição dos métodos de agrupamento não hierárquicos para cada topologia em teste pelas técnicas de agrupamentos hierárquicos, utilizando o dendrograma gerado por métodos de agrupamento hierárquico na extração dos grupos. Dendrograma é uma representação matemática e ilustrativa de todo o procedimento de agrupamento através de uma estrutura de árvore (EVERITT *et al*. 2001) (isso já foi definido anteriormente. Retirar.).

Isso se justifica nas sucessivas execuções dos métodos de agrupamento não-hierárquicos, se tornando um dos gargalos computacionais em redes neurais de funções de base radial, principalmente quando se deseja testar muitas topologias. A vantagem de se utilizar agrupamento hierárquico em relação aos métodos não hierárquicos é o fato de se utilizar o mesmo dendrogama para estabelecer os vários grupos que a rede neural necessita formar na determinação os centros durante a fase

do treinamento não supervisionado, independente do número de neurônios/centros que a rede utiliza.

Como comentado anteriormente, no caso dos agrupamentos hierárquicos, não se define um número de grupos a priori, e sim os grupos que se formariam a partir de uma métrica de distância e o método de ligação pré - estabelecidos.

A fim de tornar mais claro como se dá a escolha de topologias candidatas a ideal, é apresentado o exemplo da (equação 3.4).

$$y = sen\,(x), 0 \leq x_i \leq 2\pi, onde\ \ x_i - x_{i-1} = 0.1 \qquad 3.4$$

Executando a análise de agrupamento dos pontos (x_i, y_i), e usando a métrica distância euclidiana no método da ligação completa, geraram-se os agrupamentos representados pelo dendrograma da figura 3.8.

Os nós do dendrograma representam agrupamentos, sendo compostos pelos grupos ou objetos (grupos formados apenas por ele mesmo) ligados a ele (nó). Ao se cortar o dendrograma em um nível de distância desejado, obtêm-se os grupos existentes nesse nível e os indivíduos que os formam. A partir desta seção (Seção 1), é possível obter quais observações/dados formam os grupos quase que imediatamente. Por exemplo, para separação dos dados da função seno em dois grupos, tem-se como observações do grupo 1 os pontos $\{(x_1, y_1), (x_2, y_2)\ , \dots, (x_{30}, y_{30})\}$ e do grupo 2 os pontos $\{(x_{31}, y_{31}), (x_{32}, y_{32})\ , \dots, (x_{62}, y_{62})\}$. Calculando-se a média das coordenadas em cada um dos dois grandes grupos, obtém-se como o centro do grupo 1 a coordenada $\left(\frac{\pi}{2}, 0.5\right)$, e do grupo 2, a coordenada $\left(\frac{3\pi}{2}, -0.5\right)$. Com o mesmo dendrograma, pode-se obter a separação em 6 grupos (Seção 2), ou para qualquer número de grupos que se deseja separar.

É essa característica dos métodos hierárquicos que pode tornar a etapa não supervisionada do treinamento mais rápida, pois uma vez formado o dendrograma, se tem todas as separações prontas e, consequentemente, as coordenadas dos centros dos N grupos formados. Se fosse usado o método do k-médias ou outro método de agrupamento não hierárquico, seria necessária a execução de um método de agrupamento não hierárquico para cada topologia testada.

Para tornar a proposta viável, foi necessária a criação de uma rotina/programa computacional cuja entrada é o dendrograma formado e o número de grupos que se deseja formar, e a saída os elementos/observações/entradas que formam cada grupo.

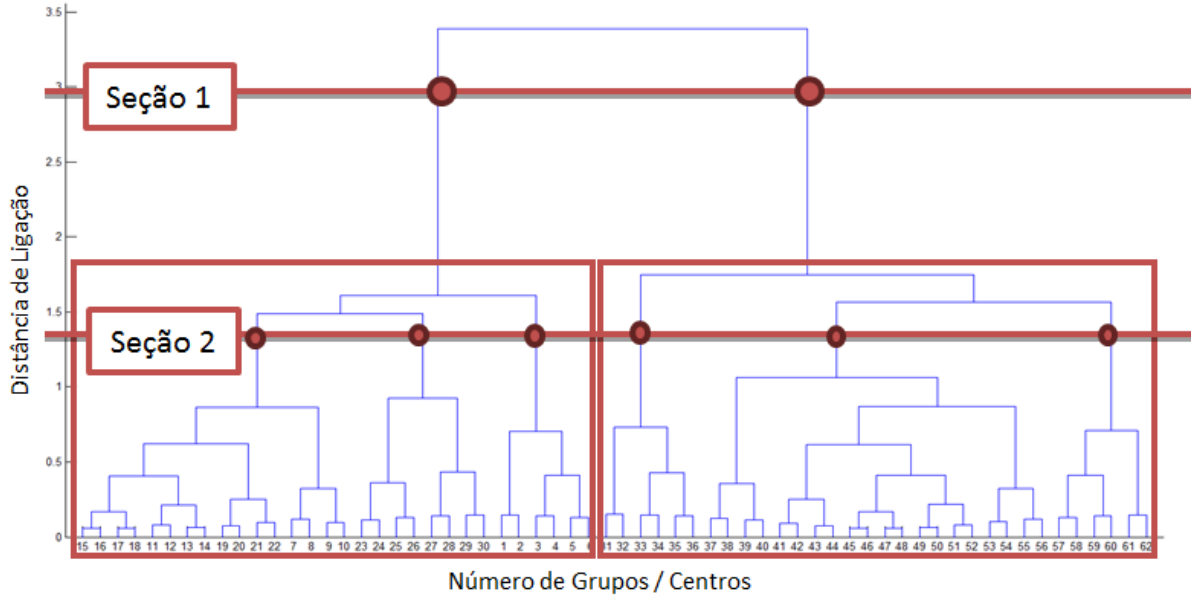


FIGURA 3.8: DENDROGRAMA GERADO PARA O EXEMPLO Y = SEN(X)
FONTE: O AUTOR

3.6 MÉTODOS PARA ENCONTRAR O NÚMERO *K* DE *CLUSTERS OU NEURÔNIOS* DA PARTIÇÃO

Uma questão importante nessa proposta é como se deve proceder para escolher o número final de ***k*** grupos, que definirá o número de centros ou de neurônios da RBF. Assim, pode-se gerar a melhor partição ou o mapeamento do conjunto de dados analisado, ou, de outra forma, em quais passos ***k*** o algoritmo de agrupamento formou grupos bem separados, pois esses apontam as topologias candidatas a ideal, ou com os melhores desempenhos. Existem alguns critérios que podem auxiliar na decisão final, como mostrado a seguir.

3.6.1 Critério da Análise do Comportamento do Nível de Fusão (Distância)

À medida que se avança no algoritmo de agrupamento, ou seja, passa-se de um estágio k para o estágio $k+1$, a similaridade entre os conglomerados ou grupos

que estão combinados nos respectivos passos vai decrescendo e, consequentemente, a distância entre eles vai aumentando. Desse modo, se for feito um gráfico do passo (número de grupos) *versus* o nível de distância (nível de fusão) do agrupamento de cada estágio do processo, pode-se visualizar se há "pontos de salto" relativamente grandes em relação aos demais valores de distância. Neste caso, procura-se detectar pontos nos quais há um decrescimento acentuado na similaridade dos conglomerados unidos, pontos estes que indicam que o algoritmo está formando grupos "bem separados ou diferentes".

Em geral, a escolha de valores de similaridade acima de 90% resulta num número de grupos muito elevado (FELIX, 2004). Com isso, deve-se escolher *a priori* a faixa de similaridade para a busca da solução. Logo, se a função apresentar vários "pontos de saltos", pode-se delimitar uma região de prováveis valores de número de grupos k que deveriam ser mais investigados por algum outro procedimento.

A figura 3.17 apresenta a relação entre o número de grupos e a distância entre grupos do dendrograma mostrado na figura 3.9, mostrando a evolução da construção do referido dendrograma. Por exemplo, há 62 grupos formados para distância de ligação igual a zero, que são as próprias observações ou variáveis e, à medida que a distância de ligação aumenta, o número de grupos diminui. Ainda, analisando esse gráfico, nota-se uma grande variação da distância para formação de dois para três grupos, ou seja, mesmo diminuindo muito a distância de ligação entre os grupos, persistem os dois grupos formados. Isso significa que esses dois grupos possuem um alto grau de dissimilaridade, ou que a distância entre os dados do mesmo grupo é muito baixa (ou ainda que a densidade dos pontos em cada grupo é alta). O mesmo ocorre na formação de 6 para 7 grupos, indicando a formação de 6 grupos menores bem separados entre si.

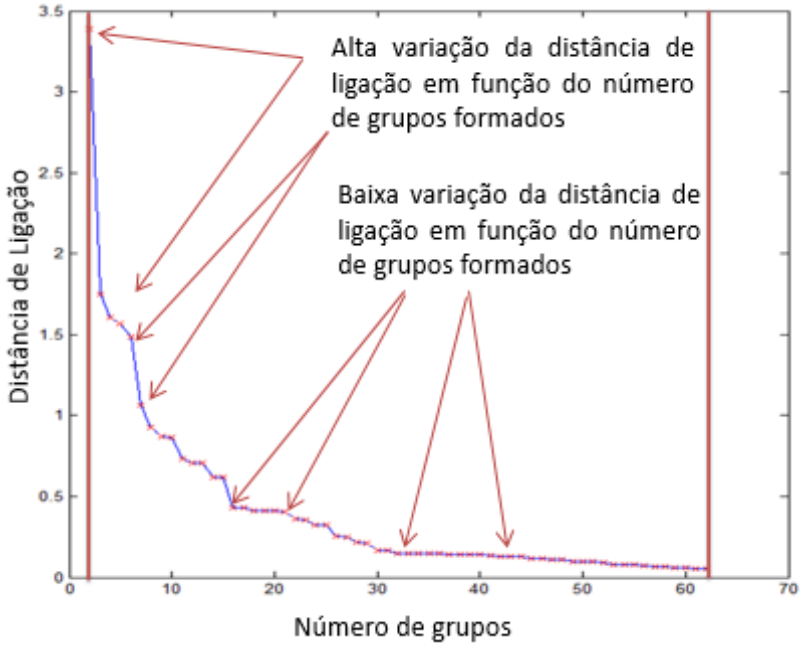


FIGURA 3.9: EVOLUÇÃO DAS LIGAÇÕES EM FUNÇÃO DA DISTÂNCIA DE LIGAÇÃO
FONTE: O AUTOR

Genericamente, a figura 3.10 abaixo sintetiza como os dados desse exemplo e outros poderiam estar separados: a linha (a) do gráfico se enquadra justamente no caso da separação dos dados do exemplo. Essa é uma situação ideal para uma boa execução da etapa não supervisionada da rede neural de base radial, pois com poucos grupos bem separados, a rede neural de base radial tende a ter um bom poder de generalização global com poucos neurônios (no mapeamento dos dados de entrada).

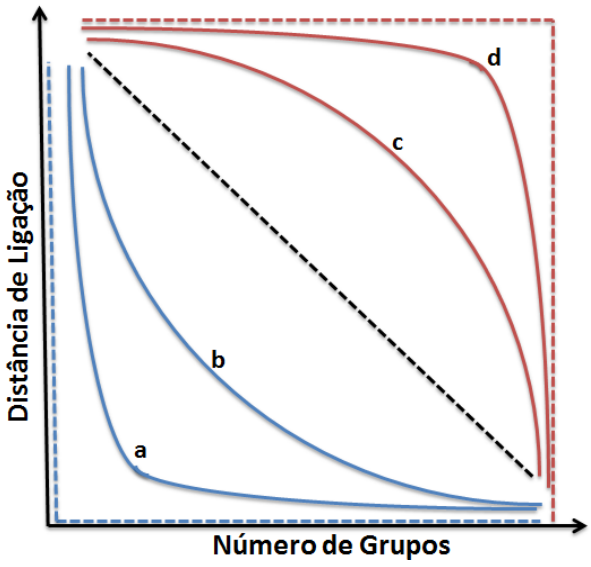


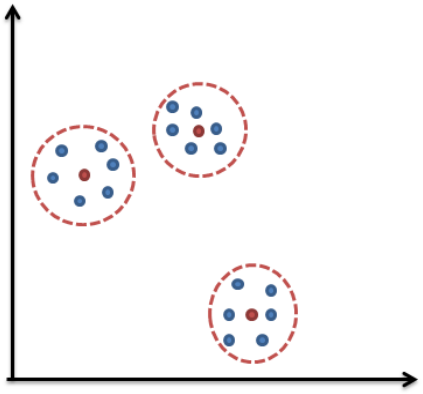
FIGURA 3.10: EVOLUÇÃO DAS LIGAÇÕES EM FUNÇÃO DA DISTÂNCIA DE LIGAÇÃO
FONTE: O AUTOR

Ainda sobre a figura 3.10, a linha(d) representa um cenário desfavorável para aplicação de redes neurais de base radial, pois indica baixas taxas de variações com altas distâncias para números pequenos de grupos, separando grandes grupos muito próximos entre si. Também indica as altas taxas de variações em baixas distâncias para grupos numerosos, mostrando que todos os elementos ficam separados. A redução do número de grupos formados está relacionada à alta variação na distância. Isso significa que os dados estão muito espalhados, o que não seria ruim caso existissem subgrupos bem separados, o que não ocorre, pois olhando para o trecho inicial do gráfico, nota-se que o aumento do número de grupos está associado a uma pequena diminuição na distância de ligação. Isso indica que há grandes grupos com pequenas distâncias entre si, ou seja, individualmente os dados possuem altas distâncias entre si, e grandes grupos com baixa distância entre si. As linhas (b) e (c) indicam situações intermediárias em relação aos casos (a) e (d).

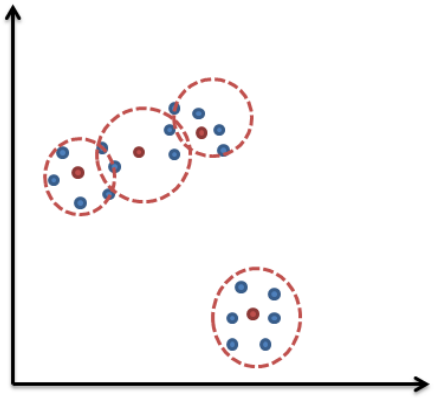
A figura 3.11 (a) exemplifica um bom posicionamento de três centros em relação aos dados, o que provavelmente estaria associado a um dendrograma com alta variabilidade de distância de ligação na formação de 3 para 4 grupos, ou seja, três grupos bem separados. Já utilizando quatro centros (figura 3.11 (b)), ocorre uma sobreposição das regiões do domínio de cada grupo, o que não é bom para utilização de redes neurais, pois pontos em sobreposição levam a estímulos semelhantes nas funções de base radial onde ocorrem essas sobreposições. Com isso, perde-se capacidade de distinção entre duas entradas diferentes, o que prejudica o desempenho da rede neural. O caso 3.10 (b) estaria associado a um dendrograma com baixa variabilidade de distância de formação de 4 para 5 grupos. A figura 3.11 (c) exemplifica uma boa separação em 6 grupos de um determinado conjunto de dados.

Quando se propõe analisar a taxa de variação das distâncias de ligação dos grupos, é justamente para captar qual número de grupos é o mais adequado, de forma que esses fiquem bem separados. Quando ocorre uma alta taxa de variação da distância de ligação associada à variação do número de grupos formados, indica-se uma separação bem definida entre os grupos. Uma dada série em estudo, cujo agrupamento das observações em função do número de grupos fosse a mesma dessa figura, a ideia principal seria criar uma heurística que indique o não treinamento da

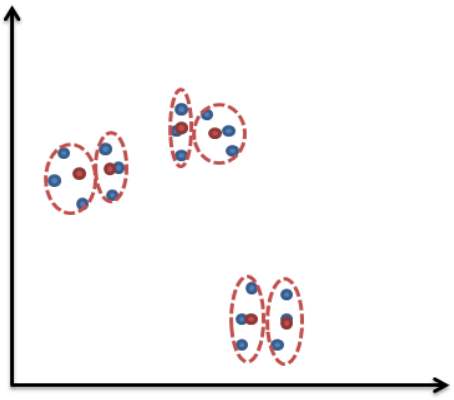
rede com quatro neurônios, pois eles não resultarão em um bom posicionamento dos centros.



(a) Utilização de três centros



(b) Utilização de quatro centros



(c) Utilização de seis centros

FIGURA 3.11 - POSSÍVEIS AGRUPAMENTOS DE DADOS EM FUNÇÃO DO NÚMERO DE CENTROS
FONTE: O AUTOR

As informações oriundas do gráfico da evolução da distância de ligação *versus* número de grupos formados que darão subsídios para atender um dos objetivos propostos, que é criar uma heurística que busque evitar treinamentos cujas posições dos centros das funções de base radial fiquem muito sobrepostas e/ou e a partir de quantos neurônios utilizados, não haja melhoria significativa no desempenho de generalização da rede neural com o aumento no número de neurônios.

3.6.2 Critério da Análise da Soma dos Quadrados Entre os Grupos: Coeficiente R^2

Em cada passo do algoritmo de agrupamento é possível calcular a soma dos quadrados entre os grupos (*clusters*) e dentro dos grupos da partição correspondente. Seja $X = x$ e $X'_{IJ} = \left(X_{I1J}\ X_{I2J}\ \ldots\ X_{ImJ}\right)$ o vetor de medidas observadas para o *I-ésimo* elemento amostral do *J-ésimo* grupo, $\bar{X}'_J = \left(\bar{X}_{1J}\ \bar{X}_{2J}\ \ldots,\ \bar{X}_{mJ}\right)$ o vetor de médias do *J-ésimo* grupo, e $X' = (\bar{X}_{.1}\ \bar{X}_{.2}\ \ldots,\ \bar{X}_{.m})$ o vetor de médias global, sem levar em conta qualquer partição onde:

$$\bar{X}_{.l} = \frac{1}{n}\sum_{J=l}^{k}\sum_{I=1}^{n_J} X_{IlJ}\,, l = 1,2,\ldots,m \qquad ((1)$$

(i) Soma dos quadrados Total corrigida para a média global em cada variável:

$$SSTc = \sum_{J=1}^{k}\sum_{I=1}^{n_J}(X_{IJ} - \bar{X})'(X_{IJ} - \bar{X}) \qquad ((2)$$

(ii) Soma de Quadrados Total Dentro dos Grupos da partição (Soma de quadrados Residual)

$$SSR = \sum_{J=l}^{k}\sum_{j=1}^{n_J}\left(X_{ij} - \bar{X}_j\right)'(X_{ij} - \bar{X}_j) \qquad ((3)$$

(iii) Soma dos Quadrados Total entre os k grupos da partição:

$$SSB = \sum_{j=1}^{N_J} n_J\,(X_J - \bar{X})'(X_j - \bar{X}) \qquad ((4)$$

Define-se o coeficiente R^2 da partição como:

$$R^2 = \frac{SSB}{SST_c} \qquad ((5)$$

Quanto maior for o valor de R^2, maior será a soma de quadrados entre grupos SSB e menor será o valor da soma dos quadrados residual SSR, esse valor varia entre 0 e 1, valores próximos de 1 indica que cada um dos grupos formados possuem uma alta densidade (elementos/observações/entradas semelhantes).

Assim, o seguinte procedimento a seguir pode ser utilizado na escolha do número K de grupos/neurônios candidatos. Vale lembrar que serão os grupos que serão utilizados pelo treinamento e pelo teste:

Passo 1: Estabeleça a relação entre o número de grupos ***k*** versus R^2;

Passo 2: Calcule a diferença absoluta da distância de ligação quando *k* passa para ***k+1;***

Passo 3: Ordene de forma decrescente os valores obtidos no passo 2;

Passo 4: As K primeiras topologias escolhidas para treinamento serão aquelas com R^2 acima de um valor estabelecido a priori.

3.7 FLUXOGRMA DA METODOLOGIA ADOTADA E PSEUDO-CÓDIGO DO MÉTODO PROPOSTO

A figura 3.12 organiza num fluxograma a sequência de ações e as relações de dependência entre esses que compõe a proposta desse trabalho.

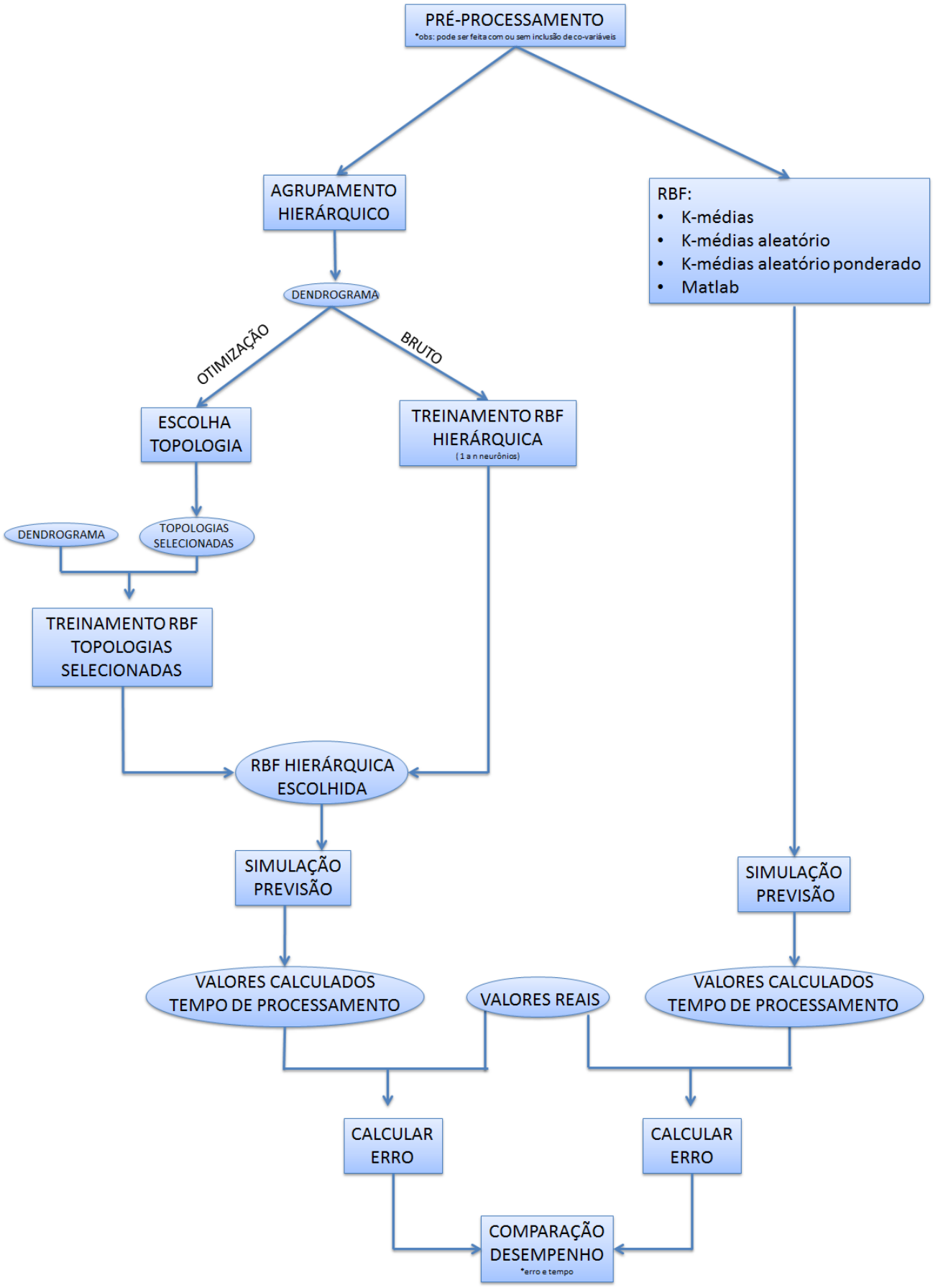


FIGURA 3.12 - FLUXOGRAMA GERAL DO TRABALHO

O pseudo código abaixo descreve todas as etapas executadas pelo método proposto:

- Dados brutos/série histórica de interesse
  - Verificação de *outliers* (*boxplot*), e caso ocorra:
  - Verificar a correlação linear dessa série com as demais disponíveis, e se essas possuem dados nos dias classificados como *outliers.*
  - Se há:
    - Ordenar os instrumentos mais correlacionados ao instrumento de interesse
    - A partir da regressão linear múltipla, estabelecer uma relação linear da série de interesse com os mais correlacionados a ele
    - Se $R^2 \geq 0.90$
      - Estimar os dados faltantes utilizando as leituras dos outros instrumentos
    - Se $R^2 < 0.90$
      - Estimar os dados faltantes por *spline* cúbica
  - Se não há
    - Estimar os dados faltantes por *spline* cúbica
  - Inserir na série histórica esses dados aproximados

- Dados pré-processados (série temporal) com m elementos
  - $t(i), i = 0,1,2,\dots,m-1$, onde $\Delta t = t_{i+1} - t_i$***o instante de tempo que ocorreu cada item da série temporal**
  - $v(i) i = 0,1,2,\dots,m-1$***o valor de cada item da série temporal**
- Ajustar a série temporal pelo método dos mínimos quadrados sobre os dados $\big(t(i), v(i)\big)$ com as seguintes funções:
  - Polinômios do primeiro ao terceiro grau
  - Para a função com o maior coeficiente de explicação R^2, aplicar a seguinte expressão sobre cada elemento da série temporal
    - $v'(i) = v(i) - f\big(t(i)\big)$
- Definir a porcentagem de treinamento e teste
- Definir o conjunto de tamanhos de janela de tempo JT e *dellay* (atraso) D a serem testados
- Definir a quantidade máxima de neurônios n utilizados na camada escondida
  - Para cada janela de tempo e *dellay* do conjunto, fazer:
    - A partir do $v'(i)$, gerar os dados de entrada $e(j)$ e saída $s(j)$ de acordo com a janela de tempo e *dellay* verificado, $j = 0,1,2,\dots,m-1-JT-D$
    - separar o conjunto de treinamento e teste como o especificado
    - aplicar o método de agrupamento hierárquico (ligação completa) nos dados de entrada, e com o dendrograma formado:
      - Armazenar as distâncias utilizadas na separação dos grupos
      - Estabelecer a relação entre o número de grupos ***k*** versus R^2;
      - Calcular a variação absoluta da distância de ligação quando *(k-1)* passa para ***k***

    - Escolha os K primeiros com maior nível de fusão (distância entre os grupos)
    - Os K que serão utilizados serão aqueles com R^2 acima de um valor estabelecido a priori.
    - Para cada K utilizado
        - determinar o centro de cada grupo pela média de cada coordenada do grupo
        - determinar a variância de cada grupo (raiz quadrada da distância de cada entrada em relação ao centro do grupo) e pela variância única sugeria por Haykin 2004.
        - ajustar os pesos da camada de saída (pseudo inversa por decomposição em valores singulares)
        - com a rede r ajustada fazer
            - $saídadarede = r + v'$
        - verificar o erro médio quadrático do conjunto de treinamento
        - verificar o erro médio quadrático do conjunto de teste
            - $previsãoumpassoafrente = r([m-1-JT-D,\dots,m-1-D,m-1]) + f\big(t(m)\big)$
- Armazenar os desempenhos no treinamento e no teste
- Armazenar o tempo computacional (tempo de agrupamento e ajustes dos pesos)

- de k=1 até n fazer
    - realizar agrupamento utilizando k-médias aleatório para k grupos
    - Para cada um desses grupos formados
        - determinar o centro de cada grupo pela média de cada coordenada do grupo
        - determinar a variância de cada grupo (raiz quadrada da distância de cada entrada em relação ao centro do grupo) e pela variância única sugeria por Haykin 2004.
        - ajustar os pesos da camada de saída (pseudo inversa por decomposição em valores singulares)
        - com a rede r ajustada fazer
            - $saídadarede = r + v'$
        - verificar o erro médio quadrático do conjunto de treinamento
        - verificar o erro médio quadrático do conjunto de teste
            - $previsãoumpassoafrente = r([m-1-JT-D,\dots,m-1-D,m-1]) + f\big(t(m)\big)$
    - Armazenar os desempenhos no treinamento e no teste

    - Armazenar o tempo computacional (tempo de clusterização e ajustes dos pesos)

- Verificar as performances e o tempo de execução obtidos nos demais métodos para todas as topologias que os dados permitem

4 RESULTADOS

A fim de demonstrar o potencial do método proposto frente aos tradicionais, serão expostos aqui os resultados obtidos na análise de séries temporais teóricas e reais. A primeira é a conhecida série temporal caótica Mackey-Glass, a segunda é uma série gerada artificialmente, com parâmetros controlados, usando um software gerador de séries, a terceira trata-se de um fenômeno químico. Além dessas séries "controladas", visando testar de maneira mais completa o método, foram analisadas, a título de estudo de caso, séries temporais obtidas dos sistemas de instrumentação da Usina Hidrelétrica de Itaipu e de uma encosta com dutos da TRANSPETRO, a descrições mais precisas dessas séries são apresentadas na sequência. Também são apresentados alguns dos resultados obtidos nos procedimentos da remoção de *outliers* e na remoção das tendências quando elas ocorrem.

4.1 SÉRIES TEMPORAIS TEÓRICAS

4.1.1 MACKEY-GLASS

A série Mackey-Glass, descrita pela equação diferencial (4.1)

$$\frac{d\big(x(t)\big)}{dt} = \frac{0{,}2x(t-\tau)}{1+x^{10}(t-\tau)} - 0{,}1x(t) \qquad 4.1$$

É uma série muito sensível às condições iniciais, não converge nem diverge. Para se obter os pontos discretos dessa série temporal, essa equação diferencial foi resolvida aplicando-se o método Runge-Kutta de quarta ordem para t de zero até 1200. Essa série é muito utilizada para verificar o desempenho dos métodos de inteligência artificial (WEIGEND, A. S. 1994).

Como exemplo, utilizando $\tau = 17$, com a condição inicial $x(0) = 1.2$, se obtém a seguinte série temporal apresentada na figura 4.1.

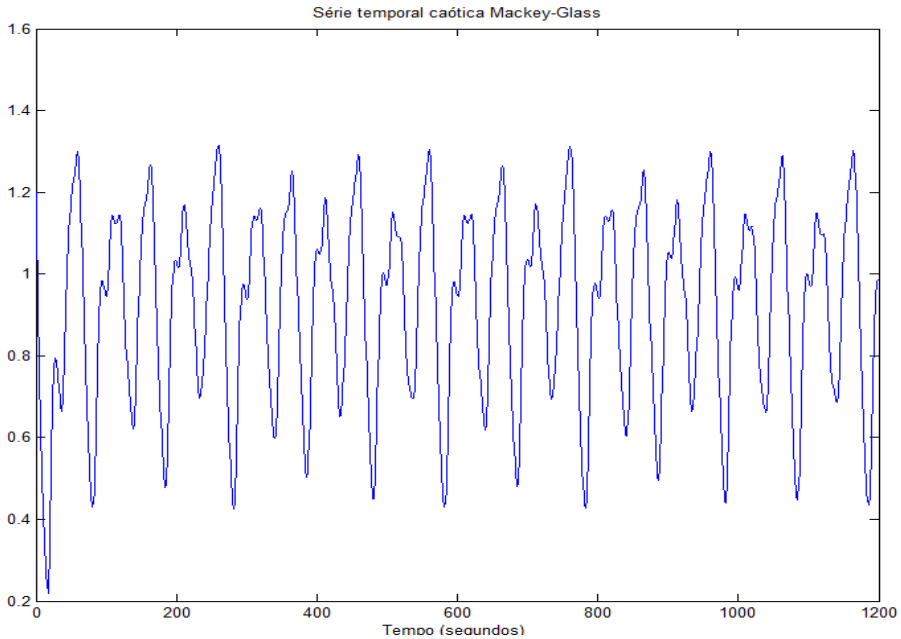


FIGURA 4.1: SÉRIE TEMPORAL CAÓTICA MACKEY-GLASS.
FONTE: (WEIGEND, A. S. 1994)

A partir dessa série, 1000 dados de entrada e saída foram gerados, onde as entradas são dadas por $[x(t)\ x(t-6)\ x(t-12)\ x(t-18)]$ e as saídas por $x(t+6)$. Os primeiros 500 pares de entrada e saída são separados para o conjunto de treinamento, e os demais 500 para o conjunto de teste.

Executando o método de agrupamento hierárquico com ligação completa sobre os dados de entrada do conjunto de treinamento, tem-se o dendrograma representado pela figura 4.2. Desse dendrograma extrai-se o centróide dos grupos formados, ou seja, o número de neurônios utilizados na rede.

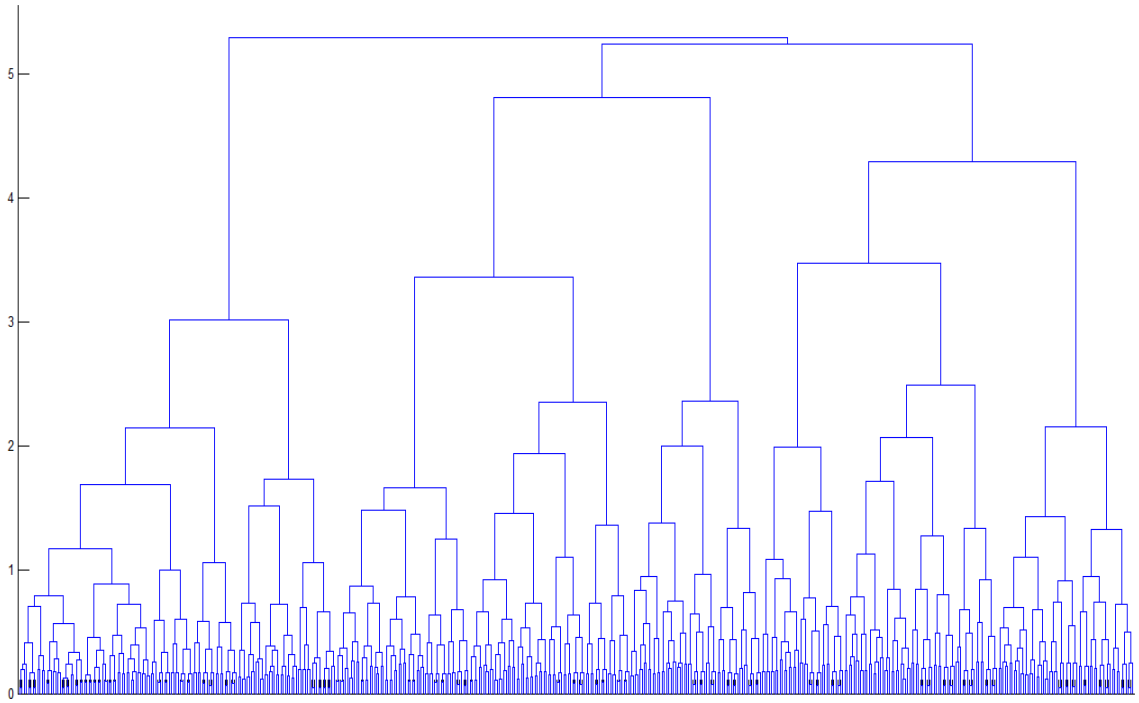


FIGURA 4.2: DENDROGRAMA DA LIGAÇÃO COMPLETA APLICADA AOS DADOS DE ENTRADA DA REDE NEURAL DE BASE RADIAL
FONTE: O AUTOR (2011)

A partir do dendrograma da figura 4.2, obtém-se a distância de ligação para cada k grupos/centros e as observações/entradas referentes a esses grupos/centros. Com isso, calcula-se o R^2 para cada k e a variação absoluta da distância de ligação quando *k* passa para *k+1,* e escolhendo os valores de *k* que tem o valor R^2 acima de 0.9, valor esse definido a priori. Ordenando o que atendem essas condições, de acordo com a variação de distância de ligação, de forma decrescente leva a concluir que somente grupos/centros com3 k>13.

Com a variação da distância de ligação, se escolhe $K = 10$ com os maiores valores dessa variação. A tabela 4.1 abaixo mostra os primeiros escolhidos que definem as topologias a serem testadas, que são os cujas linhas estão sombreadas, por exemplo, o primeiro escolhido foi com 18 grupos, pois seu R^2 está acima de 0.9, a variação da distância de ligação que forma 18 é 0.4396, quando passa de 18 para 19 grupos, a variação da distância é 0.0464, a maior variação de todas, os próximos sombreados foram os escolhidos pelo método proposto.

TABELA 4.1 - INFORMAÇÕES OBTIDAS APÓS O AGRUPAMENTO HIERÁRQUICO DOS DADOS DE ENTRADA DA SÉRIE MACKEY-GLASS

k Grupos	Distância	R^2	Variação da distância
...	...	...	...
18	0,4396	0,9316	0,0464
21	0,3761	0,9496	0,0333
13	0,4874	0,9042	0,0174
14	0,4700	0,9131	0,0168
33	0,2826	0,9703	0,0168
40	0,2395	0,9766	0,0137
26	0,3251	0,9621	0,0137
17	0,4516	0,9256	0,0121
34	0,2659	0,9727	0,0109
31	0,3000	0,9683	0,0107
41	0,2259	0,9772	0,0077
22	0,3428	0,9514	0,0071
20	0,3829	0,9453	0,0068
...	...	...	...

FONTE: O AUTOR

A figura 4.3 mostra o gráfico do Erro (MSE), de 2 até 50 neurônios, e a distância de ligação de 2 até 50 grupos. Não foram escolhidos grupos abaixo de 13 pelo fato de eles terem um R^2 abaixo de 0.9. Porém, na figura 4.3 (a), se evidencia uma relação entre a variação de distância de ligação com a variação do erro. Na figura 4.3 (b) foram escolhidas duas topologias, trecho onde começam a ter grupos com R^2 acima de 0.9. Nota-se que na figura 4.3 (c) foram escolhidos aquele que possuem um alto R^2 e/ou alta variação da distância. O mesmo ocorre na figura 4.3 (d). Nota-se que, nesse trecho, as distâncias de ligações começam a se estabilizar, ocorrendo o mesmo com o MSE. O grande diferencial desse método é o fato de ele escolher um número reduzido de topologias a serem testadas, evitando topologias que teoricamente resultariam no MSE semelhante a esses escolhidos. Observa-se ainda, nesse mesmo gráfico, que as 10 topologias escolhidas englobam grande parte da variação do MSE.

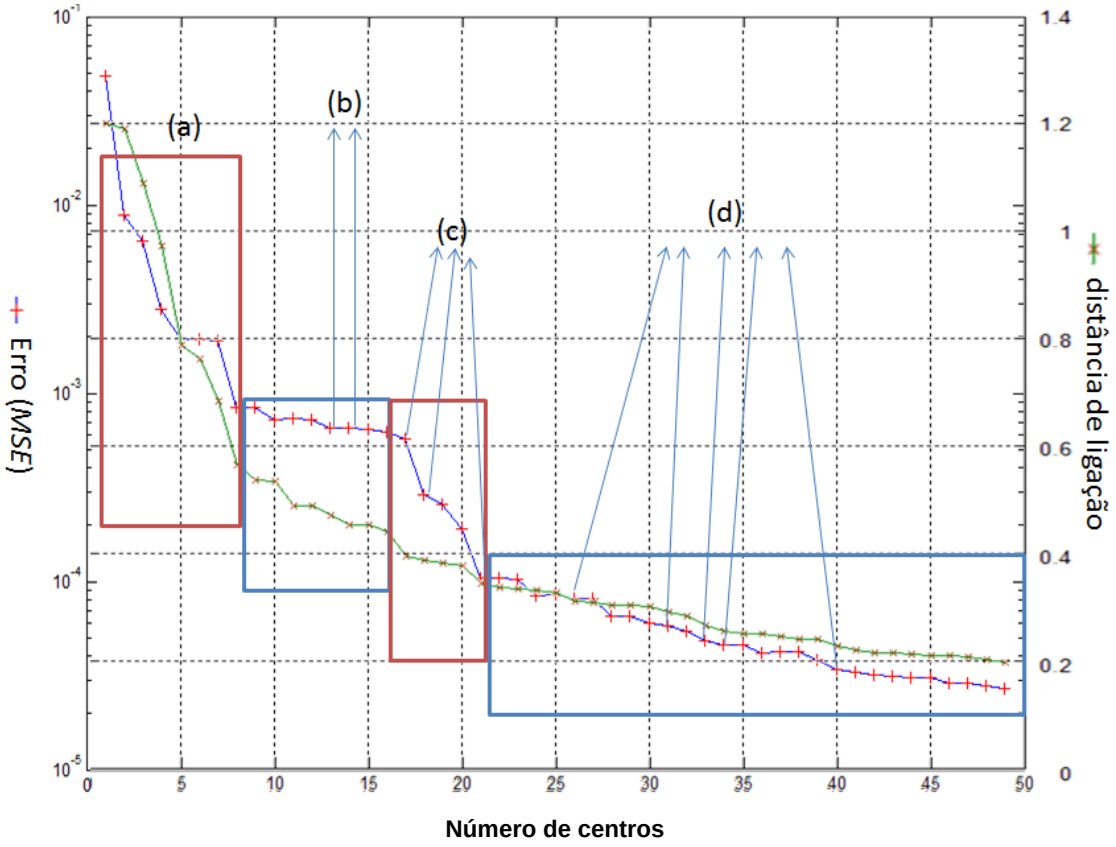


FIGURA 4.3 - EVOLUÇÃO COMPARATIVA ENTRE O MSE E A DISTÂNCIA DE LIGAÇÃO (MSE), DESTACANDO QUAIS TOPOLOGIAS FORAM ANTECIPADAMENTE ESCOLHIDAS

A tabela 4.2 mostra um comparativo do desempenho do método proposto com outros métodos referenciados na literatura aplicados para os mesmos. O método proposto mostra-se eficaz, sendo inferior apenas ao LLWNN + hibrid, de Chen, Yang & Dong (2006). Porém, deve-se destacar o número reduzido de topologias testado que garantam mesmo assim um bom desempenho, pois se aproximou dos melhores resultados testando um número muito inferior de topologias. Caso se aceitasse um número superior de topologias a serem testadas, talvez 100 topologias, o resultado poderia ser superior a todos apresentados nessa tabela.

TABELA 4.2 - COMPARAÇÃO DOS RESULTADOS DA METODOLOGIA PROPOSTA COM DIFERENTES MÉTODOS APLICADOS À SÉRIE MACKEY-GLASS

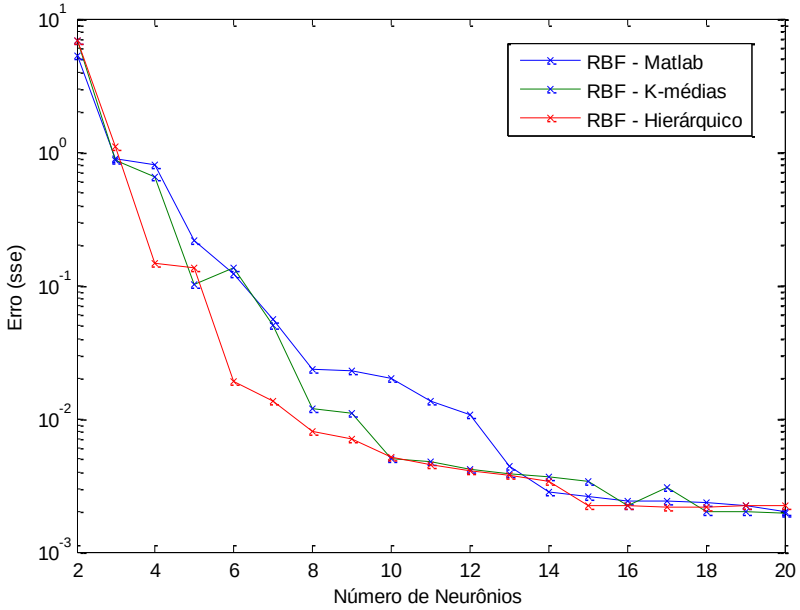
Método	**Erro de predição (RMSE)**	**Número de iterações ou épocas (quando for o caso)**
Clássical RBF (with 23 neurons) Cho and Wang (1996)	0.0114	
ANFIS and fuzzy system, Jang et al. (1997)	0.007	500
LLWNN + hibrid Chen, Yang & Dong (2006)	0.0036	3000
Neural tree Chen, Yang & Dong (2006)	0.0069	
Backpropagation NN	0.02	500
Auto regressive model	0.19	500
Genetic algorithm and fuzzy system (Ensemble) Kim and Kim (1997)	0.026243	500
M-Estimator in RBF Lee, C. C, et all (2009)	0.005541	500
RBF Hierárquico (10 topologias testadas)	**0.005287**	

4.1.2 Mackey-Glass – Caso 2

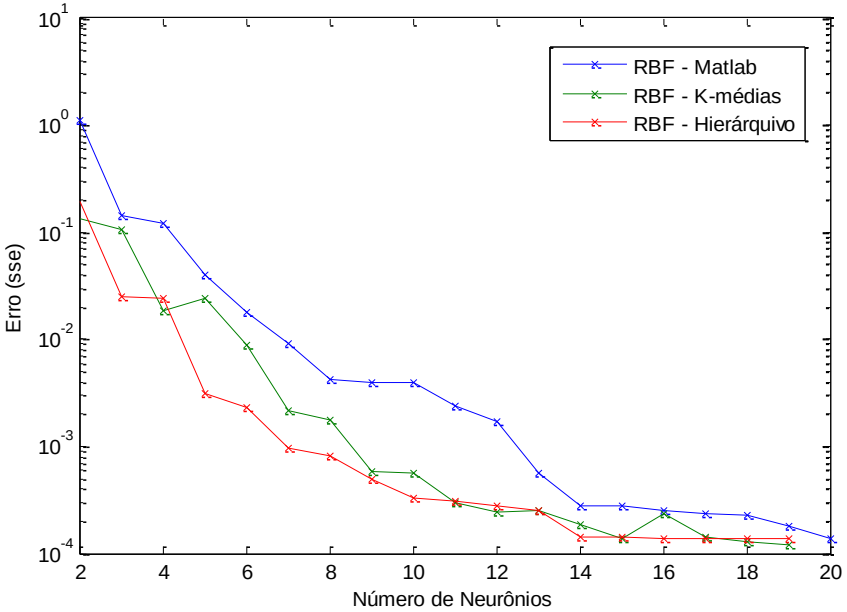
Os dados de treinamento usados no caso 1 já são conhecidos, dado o elevado número de artigos que os sugerem, o que facilita a obtenção de bons índices de desempenho pelos métodos. Mas deve-se discutir até que ponto o formato dos dados de entrada contribuíram para tal desempenho. Para responder essa pergunta mudou-se o formato dos dados de entrada dessa série, usando um conjunto de entrada para 1000 dados de entrada e saída, onde as entrada são dadas por $[x(t)\ x(t-1)\ x(t-2)\ x(t-3)]$ e a saída por $x(t+4)$. Os primeiros 1000 pares de entrada e saída são separados para o conjunto de treinamento, e os demais 197 para o conjunto de teste.

A FIGURA 4.4 ilustra os resultados obtidos pelos métodos utilizados nesse trabalho, que foram: a RBF utilizando *K-médias* na fase não supervisionada, a RBF do pacote computacional *Matlab*, e a RBF proposta nesse trabalho, usando o método

da ligação completa aplicada à fase não supervisionada, cada item dessa figura será comentada na sequência.



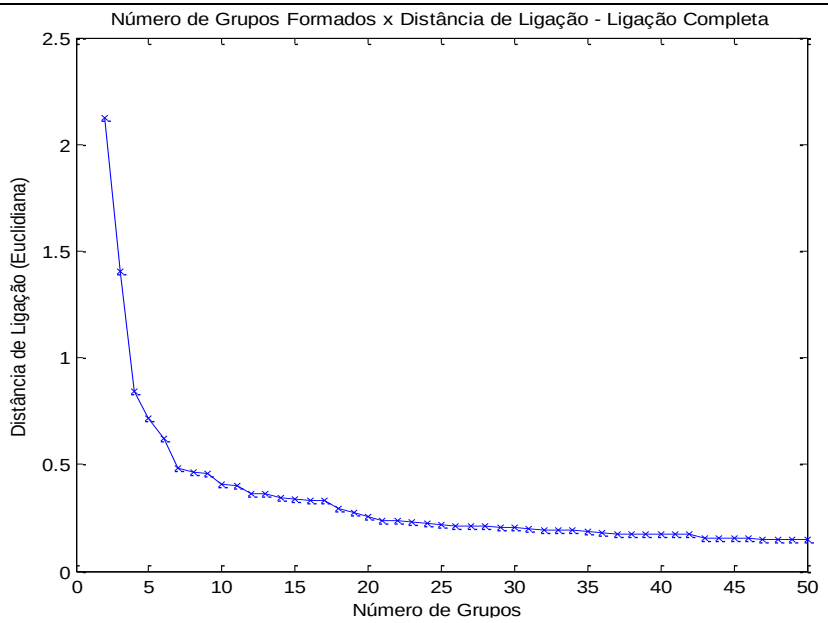
(a) Desempenho Médio das Redes no Treinamento



(b) Desempenho Médio das Redes no Teste

	Tempo de execução do treinamento das redes (segundos)				
Número de neurônios	5	10	20	50	100
RBF K-médias	0.375	1.515	6.438	35.687	170.891
RBF Hierárquico	0.531	1.36	4.328	27.313	135.906

(c) Tempo de Execução Médio do Treinamento de Cada método: Incrementando Um a Um Neurônio, n Neurônios



(d) Gráfico: Número de grupos x Distância de Ligação

FIGURA 4.4 - SÍNTESE GERAL DO DESEMPENHO DAS REDES NEURAIS TESTADAS.
FONTE: O AUTOR (2011)

Analisando o gráfico da evolução da distância de ligação em função do número de grupos formados na figura 4.4 (d), observam-se altas variabilidades da distância de ligação em função dos grupos para formações abaixo de 10 grupos, mesmo intervalo onde o desempenho de generalização no treinamento e teste tem altas variações ((FIG 4.4: (a) e (FIG 4.4: (b)). Pois esses grupos formados possuem uma boa separação. Porém, a partir de 20 grupos, nota-se baixa variabilidade da distância. Isso significa que aumentar ou diminuir o número de grupos em regiões como essa, de baixa variabilidade da distância, contribuirá pouco para o aumento do desempenho de generalização, pois não propiciarão separações entre grupos mais distintas em relação ao número de grupos atual. Essa conclusão se confirma analisando comparativamente o gráfico da evolução do erro em relação ao número de neurônios utilizado na rede neural. Aproximadamente a partir de 20 neurônios, não há ganho significativo da capacidade de generalização em função do aumento do número de neurônios utilizados na rede neural.

Quanto aos resultados em si, a RBF proposta teve um desempenho satisfatório frente à RBF utilizando K-médias na etapa não supervisionada e a RBF da pacote computacional *Matlab*.

Tanto na RBF Hierárquica (proposta) como na RBF com *K-médias*, utilizou-se a matriz pseuso inversa para os ajustes dos pesos da camada de saída, e em ambos os casos observou-se o fato que a matriz pseudo inversa a ser calculada estar muito próxima da singularidade, o que torna o resultado dos cálculos da matriz pseudo inversa imprecisos numericamente e, com isso, prejudicar o desempenho da rede. Uma maneira de se contornar esse problema foi calcular a matriz pseuso inversa por decomposição em valores singulares, método já utilizado em outros trabalhos da área como comentado anteriormente, e realmente o ganho de precisão foi notado.

Em relação ao tempo computacional despendido, era esperado que o tempo de execução do método proposto fosse menor pois, uma vez formado o dendrograma, tem-se todos os grupos formados. Já na RBF por *k-médias*, é feita uma clusterização para cada topologia testada. A figura 4.3 (c) lista o tempo de execução de cada um dos métodos. O tempo da RBF do Matlab não foi incluso pelo fato do programa possuir outros recursos que aumentam o tempo de execução, qualquer comparação de tempo em relação a este seria muito duvidoso.

4.1.3 Série Temporal Artificial

Em virtude da conhecida compatibilidade das redes neurais aplicadas à previsão da série Mackey-Glass, foi gerada artificialmente uma segunda série para análise. O modelo tradicional mais adequado a esta série é o ARMA(1,0,2), com um desvio padrão de 0,1. A série em questão possui 100 observações, onde as entradas são dadas por $[x(t)\ x(t-4)\ x(t-8)\ x(t-12)]$ e a saída $x(t+4)$. Os primeiros 90 pares de entrada e saída são separados para o conjunto de treinamento, as últimas 6 observações foram utilizadas para teste de previsão. A série em questão, bem como os intervalos de dados utilizados no treinamento e teste da rede neural estão apresentados na figura 4.5.

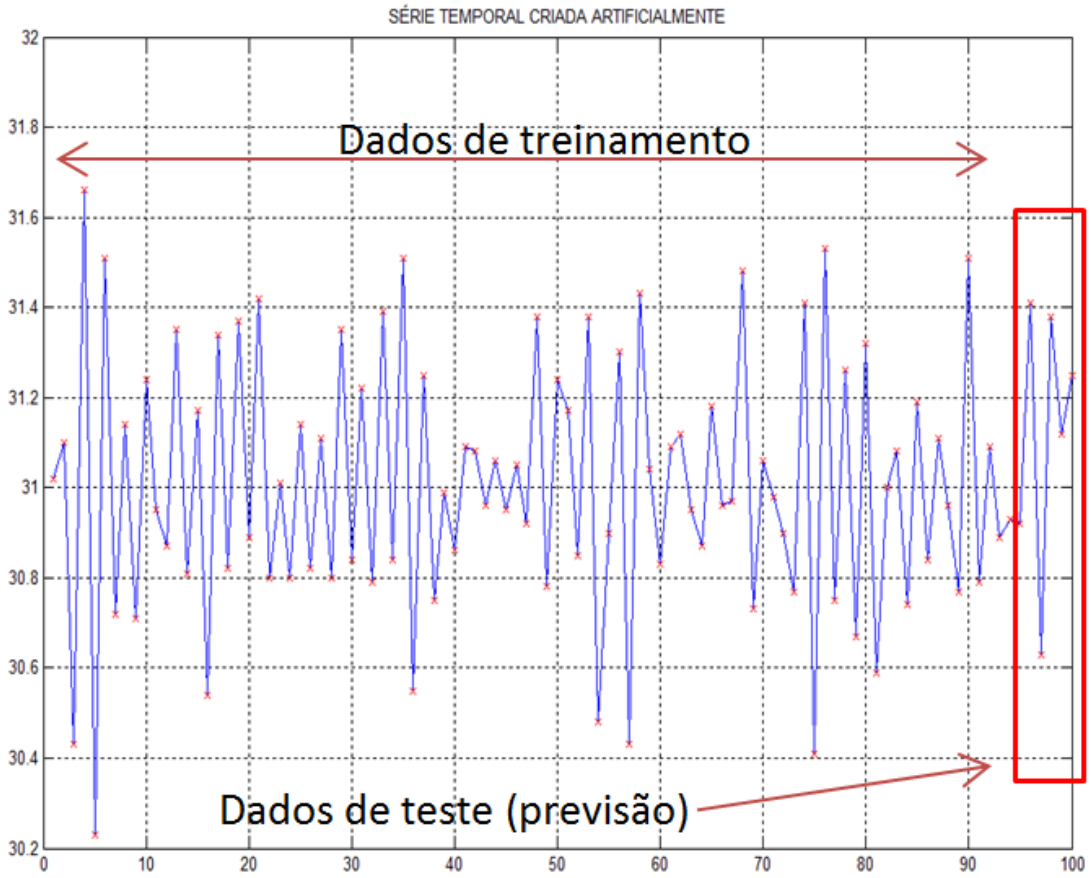


FIGURA 4.5 - SÉRIE TEMPORAL ARTIFICIAL

Sobre os dados de entrada, juntamente com o dendrograma formado, aplicou-se o método proposto, avaliando ao todo 12 topologias. A escolha foi baseada nos resultados obtidos na tabela 4.3, que apresenta o número de grupos/centros extraídos

do dendrograma em relação a distância de ligação necessária para ocorrer uma nova separação, por exemplo, com uma distância ligação igual 2,41 formam-se 2 grupos, mas se diminuir essa distância em 1,01, o número de grupos aumentaria para 3, quanto maior esse valor, maior é o indício que esses grupos estão separados, as topologias relacionadadas nessa tabela foram ordenadas em ordem decrescente em relação a esse valor, e os selecionados são as 15 primeiras topologias dessa lista cujo R^2 ficou acima de 0,9 (os destacados em cinza).

TABELA 4.3 - INFORMAÇÕES OBTIDAS PELA APLICAÇÃODO MÉTODO PROPOSTO NA SÉRIE ARTIFICIAL

k grupos centros neurônios	Distância	R^2	Variação distância de ligação ***k-(k+1)***
2	2,4125	0,5050	1,0091
6	1,2321	0,6876	0,1692
8	1,0411	0,7588	0,1223
4	1,3796	0,5955	0,1076
9	0,9187	0,7751	0,0935
17	0,6051	0,8816	0,0616
15	0,6633	0,8669	0,0557
11	0,7850	0,8071	0,0469
20	0,5286	0,9005	0,0434
33	0,3844	0,9552	0,0406
10	0,8253	0,7966	0,0403
5	1,2720	0,6396	0,0400
13	0,7109	0,8359	0,0370
12	0,7380	0,8204	0,0271
36	0,3360	0,9621	0,0258
3	1,4034	0,5465	0,0238
71	0,1746	0,9960	0,0220
7	1,0629	0,7226	0,0219
88	0,0707	0,9999	0,0197
23	0,4689	0,9178	0,0184
29	0,4256	0,9413	0,0174
37	0,3102	0,9637	0,0157
22	0,4839	0,9120	0,0150
72	0,1526	0,9964	0,0137
31	0,4082	0,9483	0,0133
47	**0,2650**	**0,9779**	**0,0132**
89	0,0510	1,0000	0,0123
84	0,1068	0,9995	0,0119
49	0,2408	0,9805	0,0113
48	0,2517	0,9795	0,0109

A partir do gráfico de desempenho figura 4.6 e relacionando as informações do mesmo com a tabela 4.3, seria suficiente perto de 12 topologias diferentes para a obtenção da topologia ideal, o que corresponde a cerca de 15% de todas as topologias possíveis. Independente da topologia o menor erro ser considerado um caso de overfitting, o método se mostrou funcional, indicando a topologia compu menor erro no teste.

Dessa maneira, é possível constatar que o método proposto apresenta um resultado compatível com os métodos baseados em redes neurais de base Radial, com a vantagem de um número inferior de topologias a serem testadas.

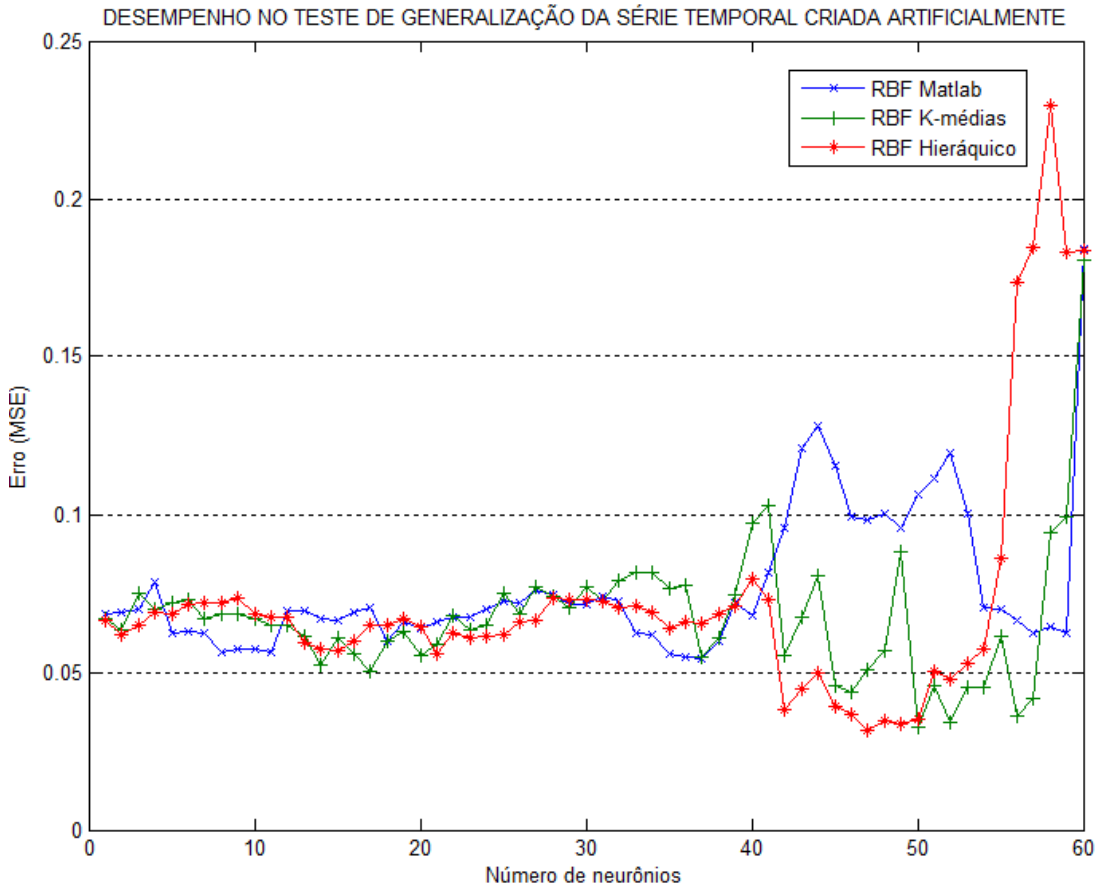


FIGURA 4.6 - DESEMPENHO OBTIDO NO TESTE DAS RBF'S EM FUNÇÃO DOS NÚMEROS DE NEURÔNIOS.

Na tabela 4.4 são apresentados os melhores desempenhos obtidos em cada método, bem como suas respectivas topologias, na previsão da referida série considerando-se seis passos adiante. O método proposto obteve o melhor resultado com o teste de apenas 15 topologias. O RBF K-Médias apresentou um resultado

semelhante ao método proposto, contudo exige um treinamento e teste de, no mínimo, 50 topologias. A FIGURA 4.7 apresenta a previsão nos testes obtida por cada um dos quatro métodos comparados em relação à série temporal avaliada.

TABELA 4.4 - MELHORES RESULTADOS OBTIDOS EM CADA MÉTODO UTILIZADO

Método	Configuração	Erro (MSE)
RBF Matlab	37 Neurônios	0.0548
RBF K-Médias	50 Neurônios	0.0349
RBF Hierárquico	**47 Neurônios**	**0.0318**
ARIMA	(1,2)	0.1052

O fato das RFB's serem melhores que o modelo arIma provavelmente é pelo ruído inserido em apenas 100 pontos.

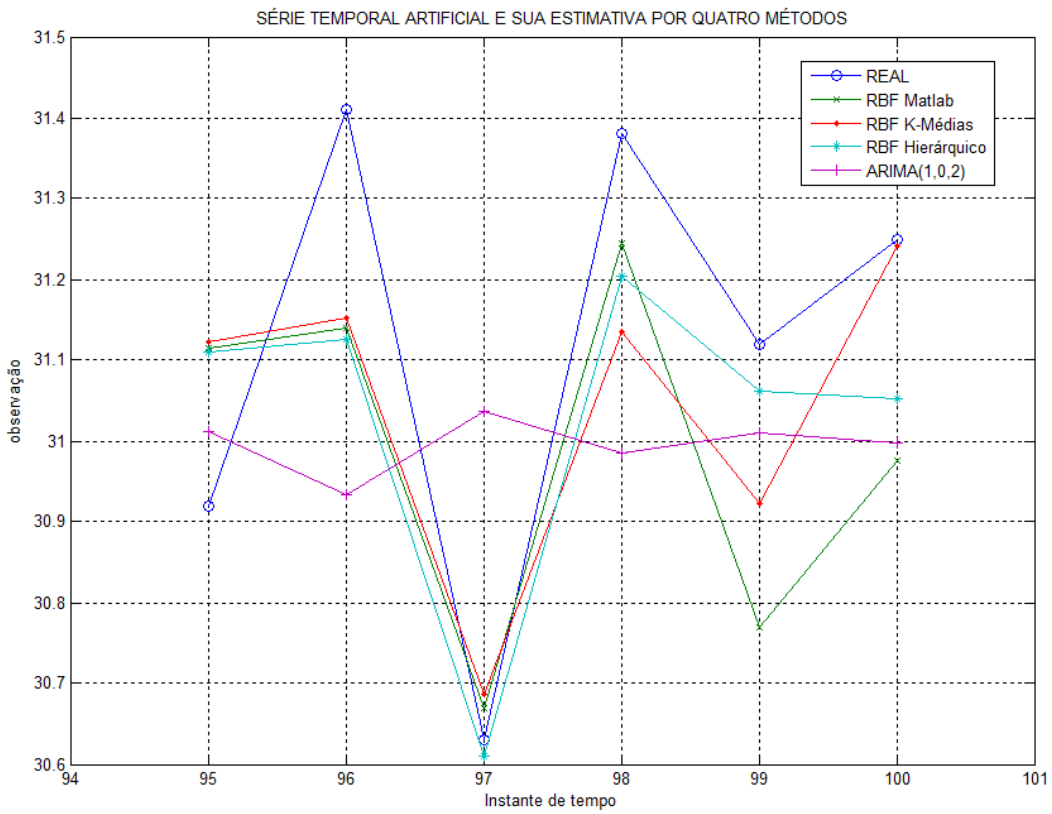


FIGURA 4.7 - SÉRIE TEMPORAL E SUA ESTIMATIVA PELO MÉTODO PROPOSTO E COMPARAÇÕES. FONTE: O AUTOR (2012)

4.1.4 Série Temporal de um Problema Envolvendo Reação Química

A terceira série abordada trata da medição da temperatura em função do tempo numa reação química, retirada do livro *Time Series Analysis,* dos autores Box e Jenkins (1994). A série possui 226 observações, e pode ser observado no gráfico 4.7. Foram retiradas as primeiras 214 observações para formar o conjunto de treinamento, e as observações restantes para teste. Aqui foi utilizado o mesmo tamanho de janela de tempo das séries anteriores, formado por 4 observações.

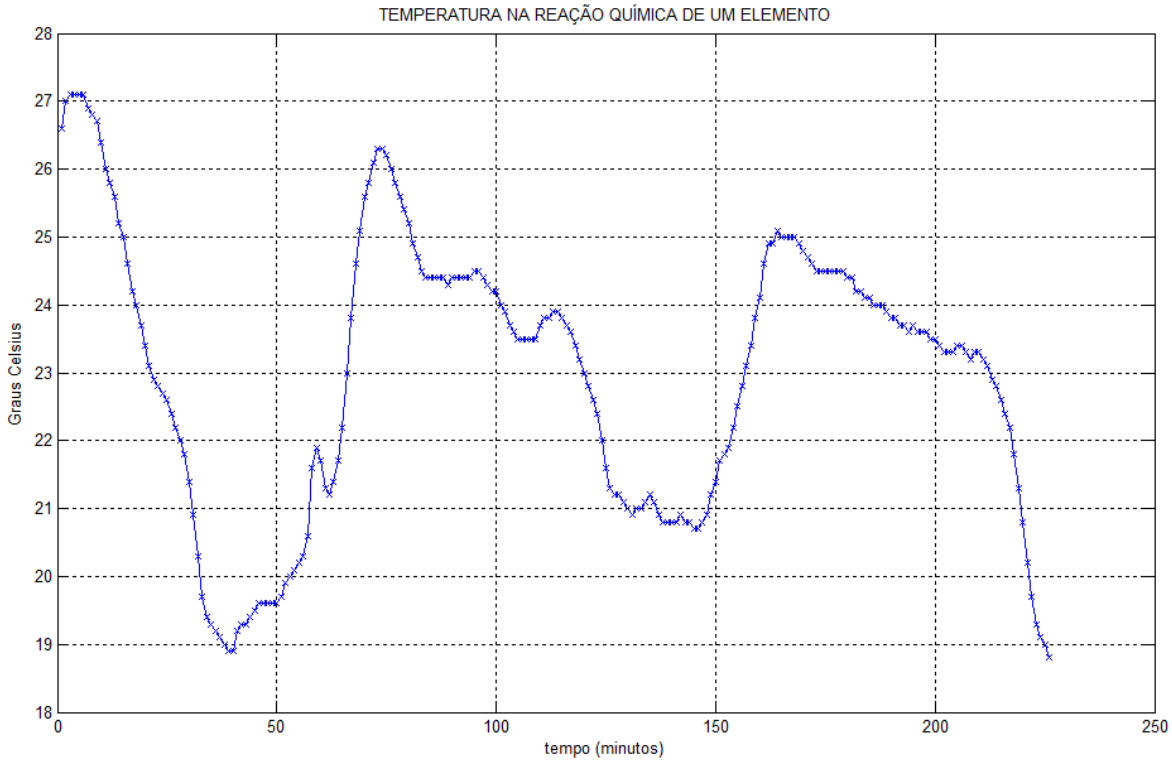


FIGURA 4.8 - SÉRIE DE TEMPERATURAS MEDIDAS A CADA MINUTO NUMA REAÇÃO QUÍMICA. FONTE: O AUTOR (BOX & JENKINS, 2008)

Executando o método de agrupamento hierárquico (ligação completa com a distância euclidiana) sobre os dados de entrada do treinamento, juntamente com o dendrograma formado, aplicou-se o método proposto. Na tabela 4.5 apresentam-se os resultados obtidos para topologias cujo R^2 resultou acima de 0.9. As linhas em negrito nessa mesma tabela indicam as melhores topologias obtidas caso se limitasse a 10 e 20 topologias a serem testadas, respectivamente. O gráfico do desempenho no teste da rede neural, que inclui essas topologias em negrito, está na figura 4.9. Esses desempenhos são para previsões de passo 1, ou seja, se conhecem os valores reais das quatro últimas leituras, e essas são as utilizadas na previsão.

TABELA 4.5 - INFORMAÇÕES OBTIDAS PELA APLICAÇÃO DO MÉTODO PROPOSTO NA SÉRIE TEÓRICA (DADOS DE ENTRADA)

k Grupos	Distância	R^2	Variação da distância
...	...	...	...
7	4,0484	0,9550	0,8877
5	5,2583	0,9312	0,7527
6	4,5055	0,9467	0,4570
13	2,6627	0,9775	0,3346
18	2,0322	0,9854	0,3146
27	**1,4106**	**0,9922**	**0,2023**
16	2,2693	0,9835	0,1860
9	3,1368	0,9682	0,1771
10	2,9597	0,9723	0,1615
24	1,6522	0,9898	0,1356
11	2,7982	0,9733	0,0982
26	1,4798	0,9920	0,0691
15	2,3280	0,9808	0,0587
30	1,1575	0,9933	0,0530
17	2,0832	0,9846	0,0510
54	0,7549	0,9972	0,0478
44	0,8831	0,9959	0,0465
29	1,2	0,9927	0,0424
34	**1,0630**	**0,9941**	**0,0383**
12	2,7	0,9757	0,0372
25	1,5165	0,9913	0,0367
...	...	...	...

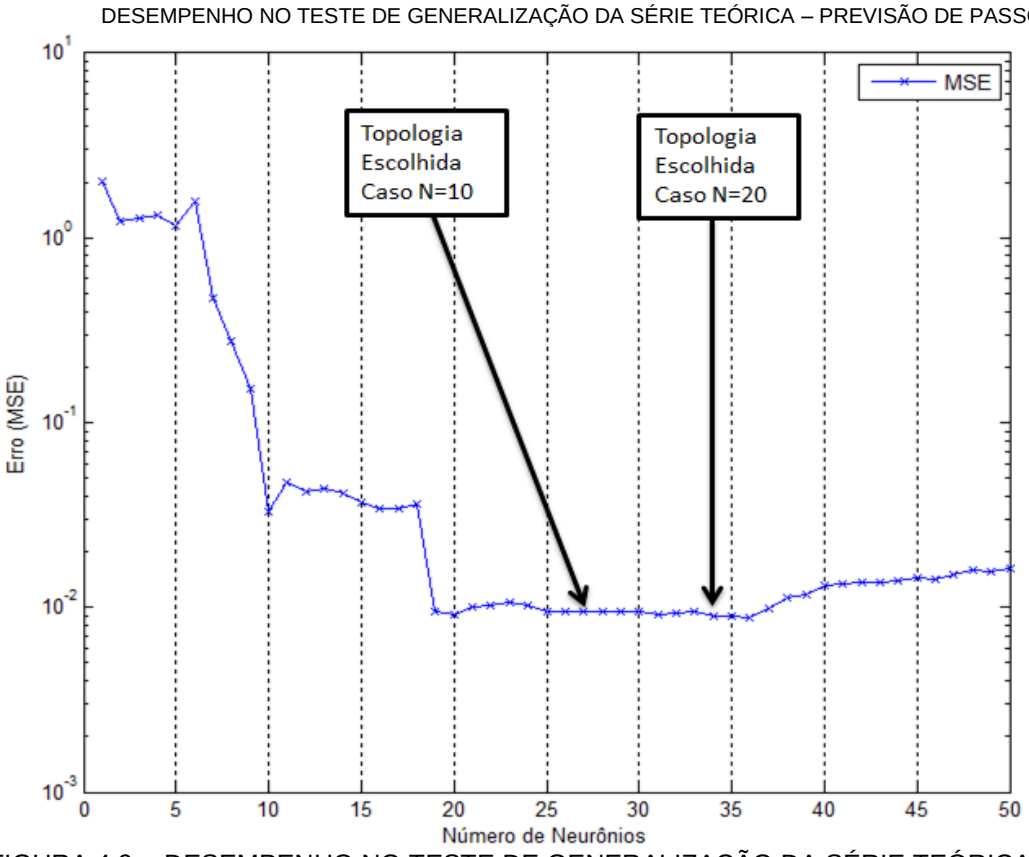


FIGURA 4.9 - DESEMPENHO NO TESTE DE GENERALIZAÇÃO DA SÉRIE TEÓRICA – PREVISÃO DE PASSO 1

O método novamente se mostrou funcional, pois com apenas 10 topologias testadas, foi incluída pelo método a topologia com 27 neurônios, que obtém um desempenho no teste muito próximo à topologia de maior performance, que é a de 34 neurônios. Esta topologia mais robusta estaria indicada caso fossem testadas 20 topologias pelo método proposto, ou seja, menos de 10% de todas as topologias possíveis, o que reduz o tempo de processamento consideravelmente.

Aplicando a melhor topologia em cada caso (rede RBF do MATLAB e a Rede RBF proposta) e o método tradicional de melhor desempenho nos dados de teste, se obtém o comparativo das previsões apresentadas na figura 4.10. Todos os métodos se comportaram de maneira satisfatória nas 9 primeiras previsões. Nas últimas 3 previsões, os métodos RBF MATLAB e o ARIMA(2,0,0) começam a divergir. A relação do *MSE* em cada método se encontra na tabela 4.6, indicando o método proposto com o melhor desempenho.

SERIE TEMPORAL TEÓRICA E SUA PREVISÃO POR TRES MÉTODOS DIFERENTES

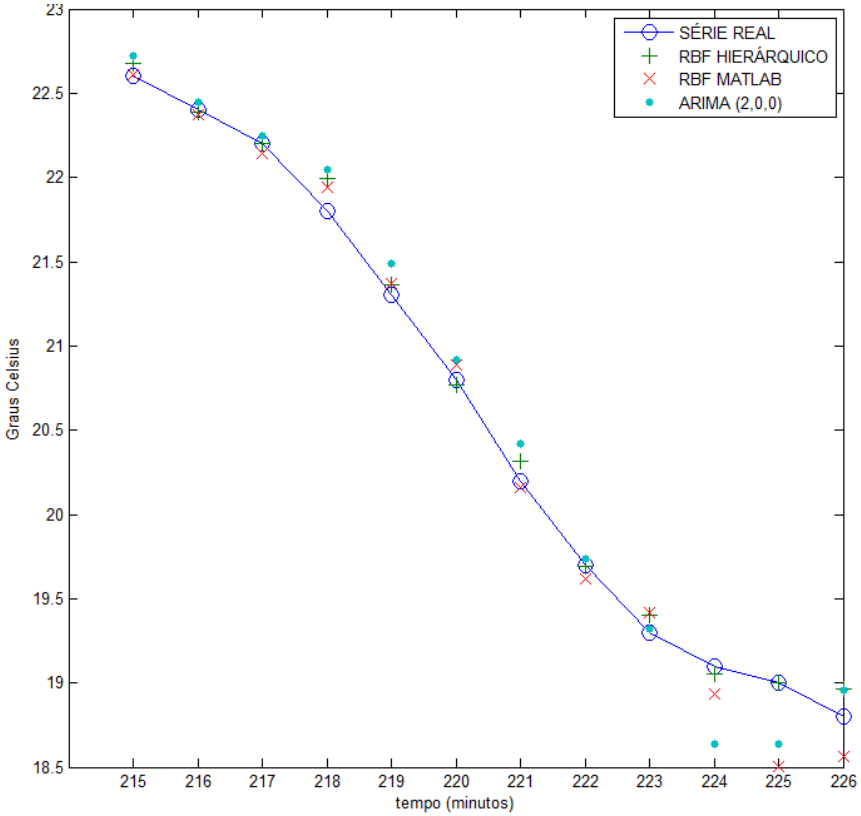


FIGURA 4.10 - SÉRIE TEMPORAL TEÓRICA E SUA ESTIMATIVA PELO MÉTODO PROPOSTO E COMPARAÇÕES. FONTE: O AUTOR (2012)

TABELA 4.6 - MELHORES RESULTADOS OBTIDOS EM CADA MÉTODO UTILIZADO – SÉRIE TEÓRICA: PREVISÃO DE PASSO 1

Método	Configuração	Erro (MSE)
RBF Matlab	57 Neurônios	0.0326
RBF Hierárquico	**34 Neurônios**	**0.0085**
ARIMA	(2,0,0)	0.0326

4.1.5 SÉRIE TEÓRICA – PREVISÃO DE PASSO 7

Usando a mesma série e o mesmo conjunto de treinamento e teste da série estudada anteriormente (4.1.4) , porém com a previsão de 7 passos adiante. Por exemplo, para gerar a previsão de passo 2, são necessárias as últimas 3 observações do conjunto de treinamento e a previsão de passo 1 obtida; para a previsão de passo 3, são utilizadas 2 previsões anteriormente obtidas, e assim por diante. Estas previsões possuem erros que naturalmente irão se propagar nas

previsões de passo 3, 4, 5, 6 e 7. Com as mesmas topologias e métodos aplicados em 4.1.4, se obtiveram as previsões indicadas pela figura 4.11, e se nota um bom desempenho dos três métodos na previsão até passo 3. A partir disso, o modelo ARIMA começa a divergir. A partir da previsão de passo 5, todos os métodos começaram a divergir. No geral, o RBF MATLAB obteve um desempenho levemente superior ao RBF hierárquico. Não foi apresentada até previsão de passo 12, pois como aqui comentado, todas divergiram acentuadamente a partir do passo 5.

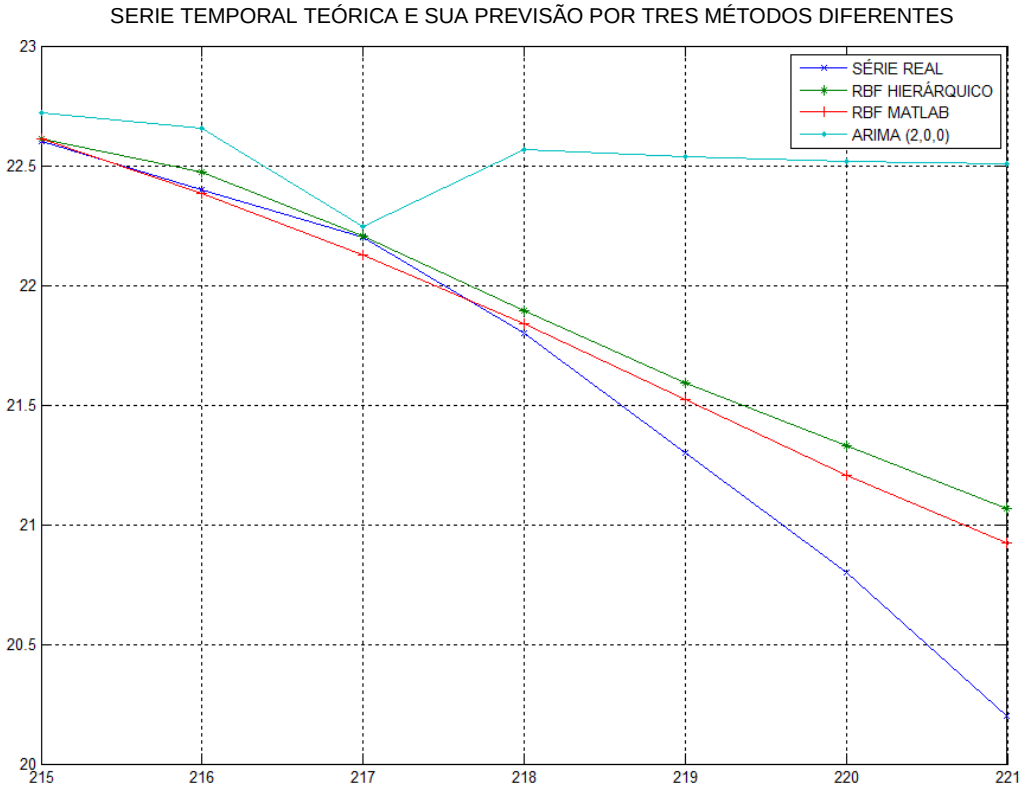


FIGURA 4.11 - SÉRIE TEMPORAL TEÓRICA E SUA ESTIMATIVA 7 PASSOS ADIANTE PELO MÉTODO PROPOSTO E COMPARAÇÕES. FONTE: O AUTOR (2012)

4.2 ESTUDOS DE CASO

Como estudos de caso, são utilizados dados reais advindos de dois tipos de obras instrumentadas: dois taludes em que se localizam dutovias e uma barragem de concreto sobre maciço rochoso, descritos brevemente a seguir.

Por se tratar de obras complexas e dados reais, problemas adicionais nos dados foram encontrados nos estudos de caso, que tiveram de ser solucionados. Após a descrição dos estudos de caso, são apresentados esses problemas e como eles foram

solucionados, para então serem apresentadas as séries temporais e o desempenho da previsão usando o método proposto.

4.2.1 Caso 1: Oleodutos da Transpetro

O oleoduto denominado de Ospar (Oleoduto Santa Catarina – Paraná) abastece a Refinaria do Paraná (REPAR), e distribui produtos para os estados do Paraná e de Santa Catarina (Transpetro, 2010), como mostra a figura 4.12.

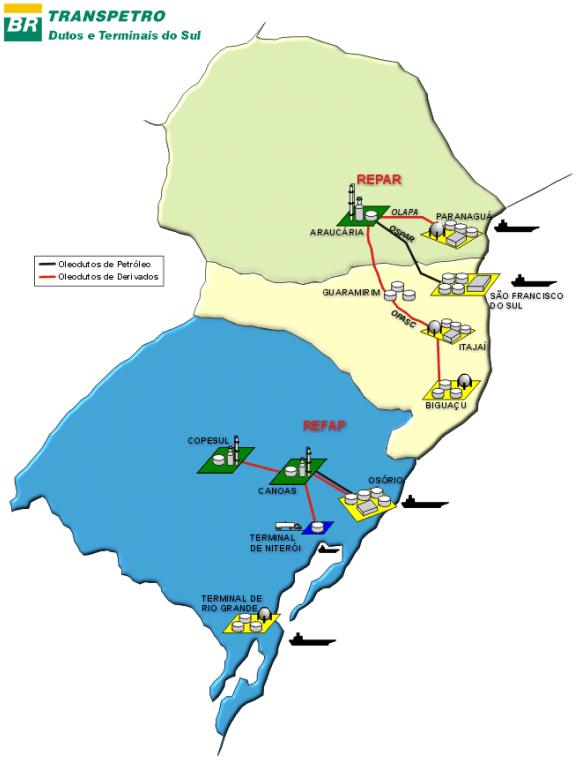


FIGURA 4.12 - TRAÇADO DE OLEODUTOS OPERADOS PELA TRANSPETRO NA REGIÃO SUL, JUNTAMENTE COM AS REFINARIAS E TERMINAIS.

Este oleoduto atravessa, em seu trajeto, diversos trechos da Serra do Mar. Um destes trechos é caracterizado por uma importante encosta coluvionar de geologia complexa, de mais de 100 m de altura. Este sítio localiza-se no município de Guaratuba, Paraná, ao longo da rodovia BR-376, próximo do km 55 + 80 cm. Nesta encosta se encontram apoiados dois oleodutos (OSPAR e OPASC, que ligam a REPAR - Refinaria de Araucária, PR a São Francisco do Sul, SC), o gasoduto Brasil-Bolívia (GASBOL) e uma linha de transmissão de energia elétrica de alta tensão, localizando-se ainda no pé do talude a própria BR-376.

Esta encosta vem apresentando sinais de movimentação desde 1995, quando das obras de duplicação da rodovia BR-376. Em janeiro de 1997, durante um período de

fortes chuvas, ocorreu um novo escorregamento da porção inferior do talude, desencadeando uma série de escorregamentos sucessivos, chegando a atingir a faixa do OSPAR/OPASC, situada a quase 300 m de distância e cerca de 80 m acima do nível da rodovia. Foram então realizadas investigações de campo (sondagens SPT) e obras de estabilização da encosta.

Próximo à faixa do OSPAR/OPASC, foram executados drenos sub-horizontais profundos e placas atirantadas. Junto à rodovia, foi construída uma cortina atirantada pelo DNER.

Foram instalados 12 piezômetros e 10 inclinômetros no local. Desde essa época (1997), o comportamento desta encosta vem sendo sistematicamente acompanhado, sendo que os movimentos observados até o momento encontram-se dentro de limites aceitáveis. Em 2003, foram instalados *strain-gages* nos oleodutos, a fim de serem medidos os deslocamentos (e conseqüentemente esforços) nos mesmos. Até 2011, duas medições deste tipo foram executadas nestas estruturas.

Atualmente, neste sítio, a instrumentação geotécnica consiste de: 21 extensômetros, 13 inclinômetros, 3 medidores de nível de água, 21 piezômetros e 90 drenos. A seguir, tem-se as principais informações a respeito destes instrumentos.

a) Extensômetros

Com base nas informações fornecidas pela Transpetro, foram instalados na encosta 21 extensômetros na faixa do OSPAR e 21 na faixa do OPASC com o objetivo de fornecer dados para avaliar o estado de segurança da obra. As leituras foram realizadas entre os anos de 2003 e 2009, com alguma diferença no período de sua realização. A tabela 4.7 mostra um resumo das informações dos extensômetros, juntamente com o período e quantidade de leituras obtidas.

TABELA 4.7- DADOS DOS EXTENSÔMETROS DA ENCOSTA OSPAR

NUMERO	COORDENADA E(m)	COORDENADA N (m)	ALTITUDE TOPO (m)	PERÍODO DE LEITURAS	QUANTIDADE DE LEITURAS
EXT-4350.02-055.800-071-OPASC	707643,031	7133451,579	382,182	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-072-OPASC	707642,953	7133451,502	381,991	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-073-OPASC	707643,109	7133451,656	381,991	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-061-OPASC	707665,752	7133429,583	372,626	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-062-OPASC	707665,673	7133429,507	372,435	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-063-OPASC	707665,831	7133429,659	372,435	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-051-OPASC	707682,041	7133412,565	369,226	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-052-OPASC	707681,963	7133412,488	369,035	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-053-OPASC	707682,119	7133412,642	369,035	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-041-OPASC	707699,224	7133395,139	365,852	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-042-OPASC	707699,133	7133395,050	365,661	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-043-OPASC	707699,315	7133395,228	365,661	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-031-OPASC	707718,673	7133375,009	363,687	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-032-OPASC	707718,594	7133374,933	363,496	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-033-OPASC	707718,752	7133375,085	363,496	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-021-OPASC	707740,027	7133352,273	365,188	15/10/2003 a 29/12/2009	30
EXT-4350.02-055.800-022-OPASC	707739,946	7133352,199	364,997	15/10/2003 a 29/12/2009	30
EXT-4350.02-055.800-023-OPASC	707740,108	7133352,347	364,997	15/10/2003 a 29/12/2009	30
EXT-4350.02-055.800-011-OPASC	707762,515	7133329,034	368,188	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-012-OPASC	707762,436	7133328,957	367,995	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-013-OPASC	707762,594	7133329,110	367,995	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-071-OSPAR	707645,552	7133453,872	382,081	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-072-OSPAR	707645,791	7133454,099	381,509	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-073-OSPAR	707645,313	7133453,644	381,509	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-061-OSPAR	707667,561	7133431,600	372,043	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-062-OSPAR	707667,800	7133431,828	371,471	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-063-OSPAR	707667,322	7133431,372	371,471	15/10/2003 a 29/12/2009	30
EXT-4350.02-055.800-051-OSPAR	707684,355	7133414,449	369,033	15/10/2003 a 19/02/2009	27
EXT-4350.02-055.800-052-OSPAR	707684,595	7133414,676	368,462	15/10/2003 a 19/02/2009	26
EXT-4350.02-055.800-053-OSPAR	707684,115	7133414,222	368,462	15/10/2003 a 19/02/2009	28
EXT-4350.02-055.800-041-OSPAR	707701,218	7133397,118	365,673	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-042-OSPAR	707701,446	7133397,356	365,101	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-043-OSPAR	707700,990	7133396,879	365,101	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-031-OSPAR	707720,817	7133376,956	363,423	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-032-OSPAR	707721,048	7133377,191	362,852	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-033-OSPAR	707720,586	7133376,721	362,852	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-021-OSPAR	707742,220	7133354,397	364,554	15/10/2003 a 29/12/2009	30
EXT-4350.02-055.800-022-OSPAR	707742,448	7133354,636	363,983	15/10/2003 a 29/12/2009	30
EXT-4350.02-055.800-023-OSPAR	707741,992	7133354,158	363,983	15/10/2003 a 29/12/2009	30
EXT-4350.02-055.800-011-OSPAR	707764,816	7133331,178	367,818	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-012-OSPAR	707765,058	7133331,402	367,246	15/10/2003 a 29/12/2009	31
EXT-4350.02-055.800-013-OSPAR	707764,574	7133330,954	367,246	15/10/2003 a 29/12/2009	31

b) Inclinômetros

Encontram-se instalados 13 inclinômetros. Os dados de instalação dos inclinômetros são apresentados nas Tabelas 4.8 e 4.9.

TABELA 4.8 - DADOS DOS INCLINÔMETROS DA ENCOSTA OSPAR

NUMERO	COORDENADA E(m)	COORDENADA N (m)	ALTITUDE TOPO (m)	ALTURA (m)	ALTITUDE TERRENO (m)	ORIENTAÇÃO (AZ)
INC-4350.02-055.800-001	707718,436	7133350,095	370,800	0,295	370,505	205°00'
INC-4350.02-055.800-002	707824,688	7133433,361	342,693	0,315	342,378	51°00'
INC-4350.02-055.800-003	707661,877	7133419,906	377,875	0,280	377,595	195°00'
INC-4350.02-055.800-004	707752,436	7133431,813	353,746	0,340	353,406	240°00'
INC-4350.02-055.800-005	707862,759	7133508,560	325,680	0,390	325,290	90°00'
INC-4350.02-055.800-006	707690,810	7133378,435	372,333	0,320	372,013	244°00'
INC-4350.02-055.800-007	707760,326	7133481,976	346,879	0,460	346,419	36°00'
INC-4350.02-055.800-008	707685,752	7133453,849	368,681	0,310	368,371	37°00'
INC-4350.02-055.800-009	707767,588	7133366,027	358,693	0,315	358,378	22°00'
INC-4350.02-055.800-010	707653,218	7133367,382	391,391	0,420	390,971	267°00'
INC-4350.02-055.800-011	707654,163	7133327,626	387,880	0,430	387,450	216°00'
INC-4350.02-055.800-012	707732,871	7133331,394	380,378	0,410	379,968	338°00'
INC-4350.02-055.800-013	707632,890	7133407,767	397,169	0,405	396,764	33°00'

TABELA 4.9 - PERÍODOS E QUANTIDADE DE LEITURAS DOS INCLINÔMETROS DA ENCOSTA OSPAR.

NUMERO	PERÍODO DE LEITURAS	QUANTIDADE DE LEITURAS
INC-4350.02-055.800-001	07/04/2001 a 11/03/2010	12
INC-4350.02-055.800-002	22/02/2002 a 11/03/2010	13
INC-4350.02-055.800-003	07/04/2001 a 11/03/2010	12
INC-4350.02-055.800-004	07/04/2001 a 11/03/2010	14
INC-4350.02-055.800-005	07/04/2001 a 11/03/2010	12
INC-4350.02-055.800-006	02/04/2000 a 11/03/2010	14
INC-4350.02-055.800-007	07/04/2001 a 11/03/2010	14
INC-4350.02-055.800-008	07/04/2001 a 11/03/2010	16
INC-4350.02-055.800-009	07/04/2001 a 11/03/2010	14
INC-4350.02-055.800-010	08/04/2001 a 11/03/2010	12
INC-4350.02-055.800-011	23/04/2002 a 11/03/2010	7
INC-4350.02-055.800-012	02/04/2000 a 11/03/2010	11
INC-4350.02-055.800-013	23/04/2002 a 09/05/2005	7

c) Medidores de Nível de Água

Encontram-se instalados medidores de nível de água. Os dados de instalação dos medidores de nível d´água, bem como os períodos de leituras estão resumidos nas Tabelas 4.10 e 4.11.

TABELA 4.10 - DADOS DOS MEDIDORES DE NÍVEL D´ÁGUA DA ENCOSTA OSPAR

NUMERO	COORDENADA E(m)	COORDENADA N (m)	ALTITUDE TOPO (m)	ALTURA (m)	ALTITUDE TERRENO (m)
MNA-4350.02-055.800-001	707753,323	7133430,037	353,619	0,100	353,519
MNA-4350.02-055.800-002	707790,355	7133458,417	342,814	0,285	342,529
MNA-4350.02-055.800-003	707824,618	7133485,843	335,557	0,245	335,312

TABELA 4.11 - PERÍODOS E QUANTIDADE DE LEITURAS DOS MEDIDORES DE NÍVEL D´ÁGUA DA ENCOSTA OSPAR

NUMERO	PERÍODO DE LEITURAS	QUANTIDADE DE LEITURAS
MNA-4350.02-055.800-001	27/07/2004 a 24/02/2010	45
MNA-4350.02-055.800-002	27/07/2004 a 24/02/2010	45
MNA-4350.02-055.800-003	27/07/2004 a 24/02/2010	44

d) Piezômetros

Em relação aos piezômetros, foi constatada altas variações na periodicidade das leituras (Tabelas 4.12 e 4.13). Optou-se por realizar leituras com maior freqüência em épocas de chuva, e com menor freqüência na estação seca.

TABELA 4.12 - DADOS DOS PIEZÔMETROS DA ENCOSTA OSPAR

NUMERO	NÚMERO ANTIGO	COORDENADA E(m)	COORDENADA N (m)	ALTITUDE TOPO (m)	ALTURA (m)	ALTITUDE TERRENO (m)
PZM-4350.02-055.800-001	PZ-01	707737,206	7133285,375	374,284	0,650	373,634
PZM-4350.02-055.800-002	PZ-02	707692,103	7133312,000	380,836	0,650	380,186
PZM-4350.02-055.800-003	PZ-03	707692,868	7133268,876	364,358	0,580	363,778
PZM-4350.02-055.800-004	PZ-04	707752,722	7133430,901	353,737	0,255	353,482
PZM-4350.02-055.800-005	PZ-05	707862,653	7133507,544	325,685	0,400	325,285
PZM-4350.02-055.800-006	PZ-06	707690,137	7133379,201	372,444	0,345	372,099
PZM-4350.02-055.800-007	PZ-07	707730,509	7133483,732	357,880	0,350	357,530
PZM-4350.02-055.800-008	PZ-08	707761,918	7133408,652	354,407	0,350	354,057
PZM-4350.02-055.800-009	PZ-09	707592,747	7133461,804	397,983	0,350	397,633
PZM-4350.02-055.800-010	PZ-10	707639,200	7133355,443	393,518	0,805	392,713
PZM-4350.02-055.800-011	PZ-11	707756,562	7133310,187	374,530	0,890	373,640
PZM-4350.02-055.800-012	PZ-12	707788,526	7133329,067	367,248	0,285	366,963
PZM-4350.02-055.800-013	PD-01	707791,683	7133457,334	342,574	0,190	342,384
PZM-4350.02-055.800-014	PD-01	707791,642	7133457,286	342,477	0,090	342,387
PZM-4350.02-055.800-015	PD-02	707765,415	7133367,580	358,196	0,215	357,981
PZM-4350.02-055.800-016	PD-02	707765,350	7133367,546	358,191	0,210	357,981
PZM-4350.02-055.800-017	PD-03	707686,635	7133452,642	368,619	0,290	368,329
PZM-4350.02-055.800-018	PD-03	707686,582	7133452,601	368,519	0,190	368,329
PZM-4350.02-055.800-019	P-01	707715,197	7133401,582	365,998	0,200	365,789
PZM-4350.02-055.800-020	P-02	707729,977	7133388,224	364,664	0,230	364,431
PZM-4350.02-055.800-021	P-03	707701,833	7133415,926	368,011	0,230	367,781

.

TABELA 4.13 - PERÍODOS E QUANTIDADES DE LEITURAS DOS PIEZÔMETROS DA ENCOSTA OSPAR.

NUMERO	PERÍODO DE LEITURAS	QUANTIDADE DE LEITURAS
PZM-4350.02-055.800-001	27/07/2004 a 24/02/2010	46
PZM-4350.02-055.800-002	27/07/2004 a 24/02/2010	45
PZM-4350.02-055.800-003	27/07/2004 a 24/02/2010	47
PZM-4350.02-055.800-004	27/07/2004 a 24/02/2010	46
PZM-4350.02-055.800-005	27/07/2004 a 24/02/2010	46
PZM-4350.02-055.800-006	27/07/2004 a 24/02/2010	46
PZM-4350.02-055.800-007	27/07/2004 a 24/02/2010	44
PZM-4350.02-055.800-008	27/07/2004 a 24/02/2010	46
PZM-4350.02-055.800-009	27/07/2004 a 24/02/2010	44
PZM-4350.02-055.800-010	27/07/2004 a 24/02/2010	47
PZM-4350.02-055.800-011	27/07/2004 a 24/02/2010	46
PZM-4350.02-055.800-012	27/07/2004 a 24/02/2010	46
PZM-4350.02-055.800-013	27/07/2004 a 24/02/2010	46
PZM-4350.02-055.800-014	27/07/2004 a 24/02/2010	46
PZM-4350.02-055.800-015	27/07/2004 a 24/02/2010	46
PZM-4350.02-055.800-016	27/07/2004 a 24/02/2010	46
PZM-4350.02-055.800-017	27/07/2004 a 24/02/2010	46
PZM-4350.02-055.800-018	27/07/2004 a 24/02/2010	46
PZM-4350.02-055.800-019	18/01/2008 a 24/02/2010	18
PZM-4350.02-055.800-020	18/01/2008 a 24/02/2010	18
PZM-4350.02-055.800-021	18/01/2008 a 24/02/2010	18

e) Drenos

Como parte das obras de drenagem realizadas pela Petrobrás incluem-se a drenagem superficial, composta por canaletas, e a drenagem profunda, composta por 90 drenos sub-horizontais (DHs).

4.2.1.1 Pré-processamento dos Dados das Dutovias

A TRANSPETRO tem um sistema próprio de armazenagem dos dados, que faz parte de uma plataforma de dados chamada GeoRisco. A figura 4.13 mostra um *screenshot* (tela) do programa. Outra vantagem desse sistema é que seu acesso pode ser remoto, para usuários cadastrados pela Petrobrás.

FIGURA 4.13 - TELA DO PROGRAMA GEORISCO.
FONTE: TRANSPETRO 2010

As leituras ficam organizadas em planilhas no formato “csv”, próprios para aplicativos como o *Excel* e o *Open Office*.

O sistema não possui uma rotina para capturar leituras de vários instrumentos de maneira automática. Por exemplo, para comparar as leituras de dois instrumentos, seria necessário abrir e baixar uma planilha de dados por vez, usando o recurso de copiar e colar para deixa-las numa mesma planilha. A simples tarefa de baixar as planilhas uma a uma pode ficar demorada caso se tenha necessidade de comparar muitos instrumentos. Além disso, aumenta-se a chance de erro humano nesse processo.

Para contornar esse problema, teve-se que criar um conjunto de rotinas computacionais capazes de executar essa e outras tarefas importantes de maneira automática, como seleção de leituras apenas num intervalo de tempo determinado, agrupamento de várias leituras obedecendo a critérios, detecção de leituras vazias, erros de formato, etc.

A FIGURA 4.14 abaixo mostra uma planilha original, contendo informações dos instrumentos e o cabeçalho referente a cada coluna de dados. Os dados propriamente ditos ficam armazenados juntos em uma única planilha, o que impede a utilização imediata dos mesmo. Por isso, foi criada a rotina no Matlab chamada captura, como no comando abaixo:

drenos_olapa=captura('c:\Users\Usuario\Desktop\dados_olapa\ospar\Olapa\DRENOS\','*.csv')

Este comando captura de maneira automática todos os arquivos no formato csv dentro de uma determinada pasta especificada, separando as informações qualitativas e quantitativas, o cabeçalho referente a cada coluna de dados e os dados, tornando cada uma dessas informações acessíveis por linhas de comando, o que é fundamental para a execução dos métodos aplicados nesse trabalho. A FIGURA 4.15 mostra os dados já armazenados no ambiente computacional *Matlab*.

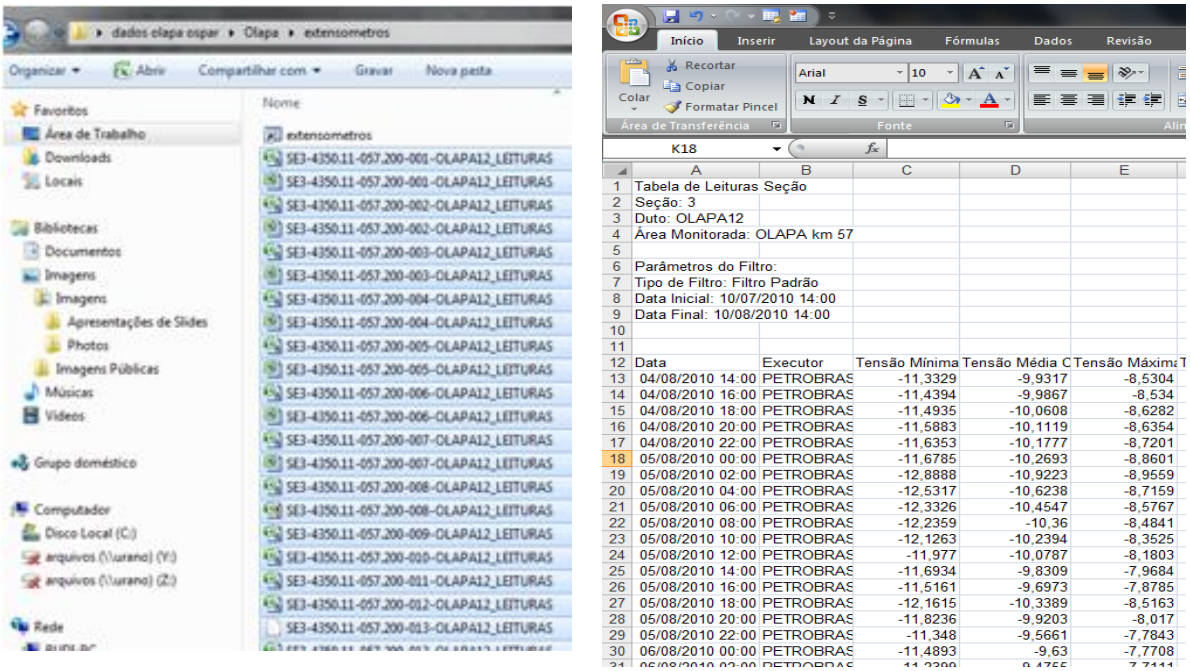


FIGURA 4.14 - DADOS EXTRAÍDOS DO GEORISCO NO FORMATO *.CSV

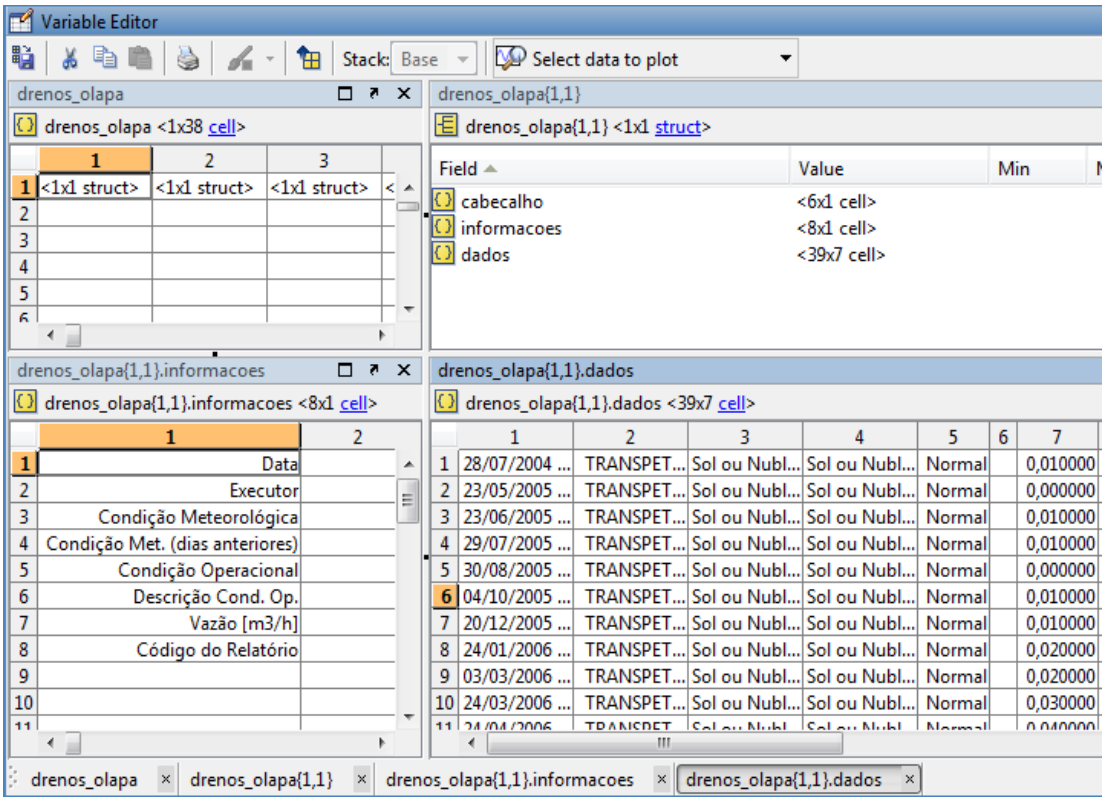


FIGURA 4.15 - RESULTANTE DA FUNÇÃO CAPTURA SOBRE OS DADOS NO FORMATO *.CSV

Nem sempre se quer trabalhar com todas as informações disponíveis. Por exemplo, os drenos da OLAPA possuem sete variáveis (informações) ao todo. O comando imediatamente abaixo

agrupa_cell(drenos_olapa,[1 7])

cria uma nova planilha apenas com as colunas 1 e 7 de todos os drenos da OLAPA, e armazena em outros campos, nesses instrumentos, essas colunas representam o instante da leitura e o valor da leitura respectivamente.

Os dados da TRANSPETRO possuem suas informações em formato texto, mesmo os números. Quando se abre uma dessas planilhas, deve-se converter coluna a coluna do formato texto para o numérico, caso contrário, não há possibilidade de efetuar operações matemáticas entre esses campos. O comando

para_numero(drenos_quantitativos,2)

converte a coluna 2 de todas as planilhas em número.

Estes dados possuem suas datas em formato dd/mm/aaaa. O comando

data_numero(drenos_quantitativos_n,1)

converte a coluna 1 de todas as planilhas em número.

Quando se quer comparar as leituras, é muito útil ter em uma única matriz os dados referentes a elas, e o mais imediato é agrupar leituras que ocorrem na mesma data. Por exemplo, o comando

quantitativos_mesma_data(drenos_quantitativos_n,0,'01/01/2004')

agrupa leituras dos drenos, com 0 dias de tolerância (leituras ocorridas no mesmo dia), ocorridas a partir do dia 01/01/2004. Quando a variabilidade de uma leitura em relação ao tempo é pequena, pode ser possível agrupar leituras com alguma defasagem. Este comando cobre essa possibilidade, bastando mudar o campo que contém o valor 0 para valor desejado.

Outras rotinas foram criadas, e serão expostas no desenvolvimento do trabalho.

Ressalta-se a importância dessas rotinas pelo fato de que a cada intervalo de tempo, novos dados são inseridos no sistema, e novamente é necessário executar todas essas conversões, o que tornaria inviável trabalhar sempre com dados atualizados.

4.2.2 Remoção de *Outliers*

Um problema muito recorrente no caso das séries reais de sistemas de instrumentação é a falta do dado num respectivo dia ou o valor dessa observação ser anômalo e, caso ele seja realmente equivocado, saber que atitude tomar . Deve-se lembrar que dados faltantes ou equivocados tornam inviável ou falha a previsão de séries temporais contendo estas características.

Após a identificação de um *outlier*, por qualquer um dos motivos mencionados em 2.5, pode-se usar os critérios de determinação de um *outlier* e a metodologia para trocar ou manter esse dado anômalo. Outra opção também pode ser a de repor dados faltantes, como descrito em 3.2.

Pelas características inerentes ao estudo de caso, que possui vários instrumentos/dados obtidos em mesmas datas, torna-se possível a utilização da regressão linear múltipla para interpolar dados faltantes ou inequivocadamente incertos, optando-se para tal a *spline* cúbica.

Para expor como é feita a interpolação de leituras usando regressão linear múltipla, suponha que existam 6 séries temporais. Nesse exemplo específico, foram usadas cinco séries temporais das leituras piezométricas (referenciadas como PZ1, PZ2, PZ3, PZ4 e PZ5), localizadas em OLAPA. Foram retiradas ao acaso três leituras de um desses piezômetros, o PZ1, e a partir das outras cinco séries disponíveis, pretende-se estimar as leituras retiradas. A ideia é verificar com que precisão esses dados retirados é recuperada.

Antes de executar a regressão linear múltipla, deve-se verificar se existe alguma relação linear entre essas séries. Para isso, é necessário calcular a correlação entre cada par de séries. Na figura 4.16 apresentam-se essas seis séries que geraram as correlações na tabela 4.24.

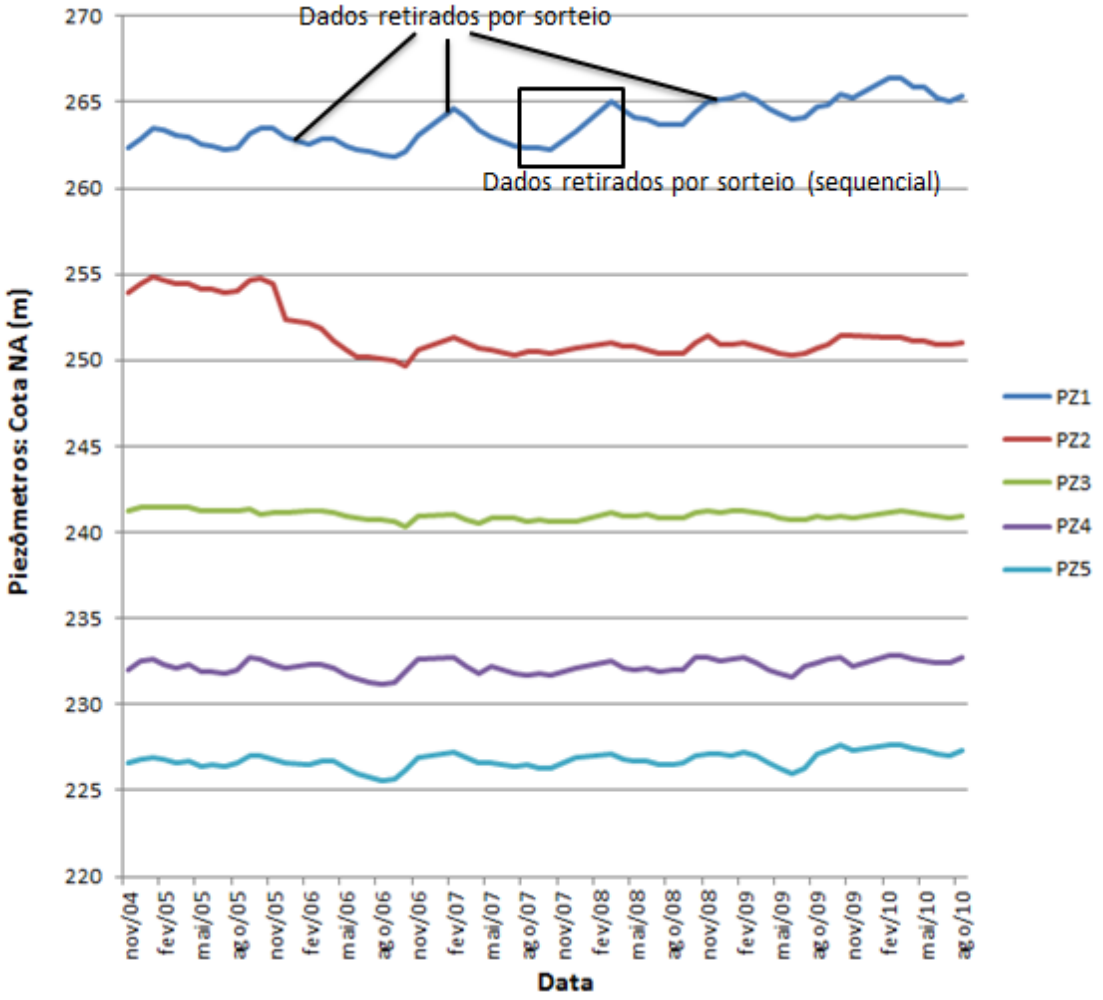


FIGURA 4.16 - SÉRIE HISTÓRICA DOS PIEZÔMETROS 001 A 005 E DO PLUVIÔMETRO DA ENCOSTA OLAPA. FONTE: TRANSPETRO (2011)

TABELA 4.14 - MATRIZ DE CORRELAÇÃO ENTRE DATAS E NA DOS INSTRUMENTOS

	Data	PZ1	PZ2	PZ3	PZ4	PZ5
Data	1	0,81	-0,65	-0,36	0,27	0,44
PZ1	0,81	1	-0,2	0,13	0,72	0,8
PZ2	-0,65	-0,2	1	0,73	0,36	0,2
PZ3	-0,36	0,13	0,73	1	0,54	0,37
PZ4	0,27	0,69	0,36	0,54	1	0,91
PZ5	0,44	0,8	0,2	0,37	0,91	1

Na matriz da Tabela 4.14, os dados com correlação significativa ao instrumento PZ1 são os piezômetros PZ4 e PZ5 e a data, pois esses possuem correlações absolutas acima de 0,7, e como eles são aos mais relacionados, são os indicados na aplicação da regressão linear. Como se deseja recuperar os dados faltantes do piezômetro PZ1, foi executada a regressão linear múltipla entre os instrumentos PZ1(variável dependente), e os PZ4, PZ5 e

data (variáveis independentes), que busca explicar a relação linear do PZ1 em relação a data, PZ4 e PZ5. Isso resultou na seguinte expressão:

$$PZ1 = -141{,}581 + 0{,}00111607 * data + 0{,}564519 * PZ4 + 1{,}01596 * PZ5 \quad (4.1)$$

Como são conhecidas as leituras dos PZ4 e PZ5 nas datas onde a leitura do PZ1 foi perdida ou retirada, usa-se e equação (4.1) para estimar quanto seria a leitura do PZ1. Na tabela 4.15, mostram-se as datas das leituras recuperadas e as leituras do PZ4 e PZ5 usadas na expressão (4.1)

TABELA 4.15 - RECUPERAÇÃO DE LEITURAS DO PZ1 USANDO REGRESSÃO

Data	PZ4	PZ5	Valor Estimado PZ1 Regressão	Valor Estimado PZ1 Spline	Valor Real PZ1
01/07/2005 – 38534	231,788	226,357	262,24	262,24	262,225
01/02/2007 – 39114	232,689	227,197	264,25	264,5	264,652
01/09/2009 – 40057	232,641	227,289	265,37	265,32	264,793

Na tabela 4.25, observa-se que os valores estimados ficaram próximos aos valores reais. No entanto, deve-se destacar as vantagem da *spline* cúbica nesse caso, pois ela depende apenas da própria série para interpolar os dados, diferente da regressão linear múltipla, que necessita das informações dos outros instrumentos com alta correlação para gerar aproximações precisas. A estimativa terceira data não teve bom resultado, em termos absolutos, olhando somente para o valor da leitura, aparentemente o erro é pequeno (Considerando-se 60 cm em 265 metros). Considerando que as leituras tem variação total em torno de 2 a 3 metros, mas deve-se destacar o objetivo dessa etapa, que é repor um dado faltante com o mínimo de perda de informação possível, se fosse substituído pela média da leitura por exemplo, o erro poderia ser de até 3 metros.

Impondo uma situação mais extrema, retiraram-se as leituras de 8 meses seguidos do instrumento PZ1, cujo intervalo de retirada está em negrito na tabela 4.16. Com os dados restantes, calculou-se a matriz de correlação entre os piezômetros, que indicou como instrumentos mais relacionados os mesmos indicados pela matriz de correlação mostrada na tabela 4.24. Isso significa que a regressão do PZ1 deve utilizar os mesmos instrumentos do caso anterior, o que é esperado.

TABELA 4.16 - LEITURAS DE CINCO PIEZÔMETROS EM OLAPA.

DATA	PZ1	PZ2	PZ3	PZ4	PZ5
...	...	...	...	...	...
07/2007	262,415	250,314	240,883	231,771	226,416
08/2007	**262,379**	250,459	240,663	231,725	226,454
09/2007	**262,294**	250,551	240,686	231,768	226,328
10/2007	**262,198**	250,448	240,661	231,705	226,336
11/2007	**263,32**	250,664	240,591	232,145	226,871
12/2007	**265**	251,057	241,154	232,529	227,167
01/2008	**264,502**	250,791	240,991	232,118	226,85
02/2008	**264,068**	250,841	240,975	231,988	226,682
03/2008	**264,018**	250,62	241,027	232,11	226,707
04/2009	263,739	250,417	240,843	231,868	226,522
...	...	...	...	...	...

FONTE: TRANSPETRO (2010)

A expressão gerada pela regressão linear múltipla é:

PZ1 = -106,915 + 0,00112294*data + 0,273942*PZ4 + 1,15963*PZ5 (4.2)

Com a expressão acima, se obtêm as interpolações mostradas mostrada na tabela 4.17, e a respectiva interpolação obtida por *spline*. Esse procedimento se mostrou muito funcional, conseguindo recuperar os dados com erro na ordem de 70 cm, uma precisão melhor do que a substituição pela média das leituras, que poderia levar a erros na ordem de 2 metros

A spline estimou melhor as datas do início do intervalo avaliado e a regressão estimou melhor do que a spline as leituras finais do intervalo. Porém, a regressão necessita de informações de outros instrumentos nas mesmas datas da recuperação desejada, e que esses possuam boas correlações; caso contrário, a equação de regressão gerada possuirá um baixo coeficiente de explicação (R^2), o que leva a aproximações imprecisas. Caso a estimativa esteja com um desvio máximo de 5% de relação ao valor suspeito (caso exista), o dado suspeito é mantido, caso contrário ou não exista, ele é substituído pela estimativa.

TABELA 4.17 - COMPARATIVOS DAS RECUPERAÇÕES OBTIDAS

DATA	PZ4	PZ5	Valor Estimado PZ1 Regressão	Valor Estimado PZ1 Spline	Valor Real PZ1
08/2007	231,725	226,454	263,293	262,2495	262,379
09/2007	231,768	226,328	263,1935	262,1944	262,294
10/2007	231,705	226,336	263,2192	262,2291	262,198
11/2007	232,145	226,871	264,0286	262,4984	263,32
12/2007	232,529	227,167	264,5792	263,1417	265
01/2009	232,118	226,85	264,1339	263,3569	264,502
02/2009	231,988	226,682	263,9371	263,5353	264,068
03/2009	232,11	226,707	264,0343	263,6712	264,018
		MSE	0,55992	0,620587	

4.3 Caso 2: Usina Hidrelétrica de ITAIPU

A hidrelétrica de ITAIPU é um empreendimento binacional desenvolvido pelo Brasil e pelo Paraguai no Rio Paraná, no trecho de fronteira entre os dois países, 14 km ao norte da Ponte da Amizade. A potência instalada da Usina é de 14.000 MW (megawatts), com 20 unidades geradoras de 700 MW cada. No ano 2000, a usina atingiu o seu recorde de produção de 93,4 bilhões de quilowatts-hora (kWh), sendo responsável pela geração de 95% da energia elétrica consumida no Paraguai e 24% de toda a demanda do mercado brasileiro. A barragem possui uma extensão de 7.700 m e altura máxima de 196 m. É composta de trechos construídos em concreto, enrocamento com núcleo de argila, e terra, como pode ser observado na figura 4.17. A área alagada de seu reservatório é de 1350 km2 (Itaipu, 2012).

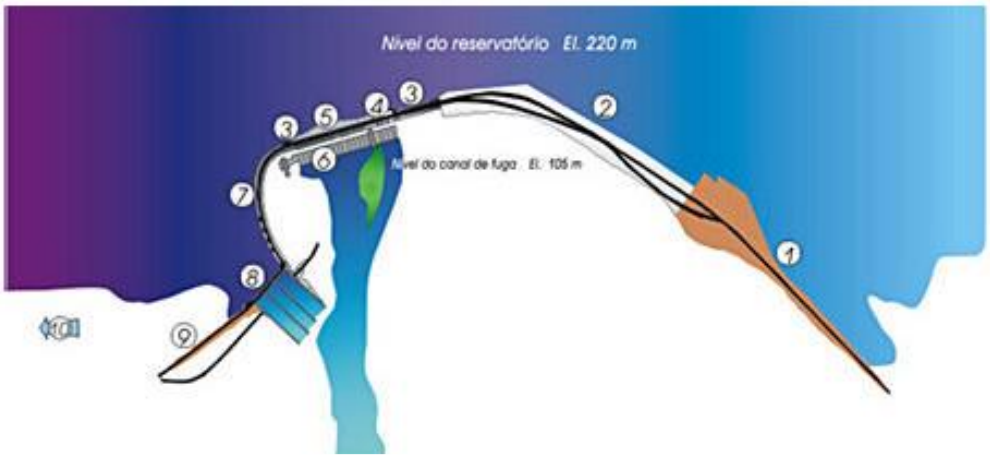


FIGURA 4.17 - PLANTA DO COMPLEXO ITAIPU.
AS ESTRUTURAS NUMERADAS SÃO: 1) BARRAGEM DE TERRA A MARGEM ESQUERDA, 2) BARRAGEM DE ENROCAMENTO, 3) BARRAGENS DE LIGAÇÃO, 4) ESTRUTURA DE DESVIO, 5) BARRAGEM PRINCIPAL, 6) CASA DE FORÇA, 7) BARRAGEM LATERAL DIREITA, 8) VERTEDOURO, 9) BARRAGEM DE TERRA DA MARGEM DIREITA. FONTE: ITAIPU (2007)

Ao longo da barragem de ITAIPU são encontrados cerca de 2300 instrumentos e 5200 drenos, no concreto e na fundação. Alguns desses instrumentos coletam dados de 1979 até os dias de hoje.

A leitura desses instrumentos é efetuada em diferentes frequências (diária, semanal, quinzenal, mensal). Conta-se também com a monitoração dos dados hidro-meteorológicos, realizada através de algumas estações que são da própria ITAIPU e de outras entidades, como a Companhia Paranaense de Energia (Copel), Agência Nacional de Águas (ANA), Operador Nacional do Sistema (ONS) e Sistema Meteorológico do Paraná (Simepar) e da paraguaia Dirección Nacional de Aeronáutica Civil (DINAC). Os instrumentos encontrados ao longo da barragem são:

a) Concreto: roseta de deformímetro, roseta de tensômetro, tensômetro, termômetro na massa, medidor de junta interno, base de alongâmetro na parede, base de alongâmetro no piso, pêndulo direto, pêndulo invertido, termômetro de superfície (RIC, 2005).
b) Fundação: piezômetro, extensômetro de haste, medidor triortogonal, medidor de nível d'água, medidor de vazão, dreno (RIF, 2005).

Em 2005, Itaipu concluiu a implantação do sistema de aquisição automática de dados, em cerca de 210 instrumentos. Esses instrumentos foram selecionados pelo corpo de engenheiros da empresa, tendo-se em vista sua importância no diagnóstico da segurança das estruturas, devido à sua localização, seus resultados no passado, entre outras razões (Itaipu, 2006).

A parte circulada na foto da figura figura. 4.18, mostra a parte de concreto com tomada de água, por onde passa a água que movimenta as turbinas para a geração de energia. Trata-se do trecho com maior altura de coluna de água da barragem e, portanto, um dos mais críticos. Esta parte é referenciada como trecho F.

O trecho F é constituído de vários blocos, sendo que cada um deles possui instrumentos que fornecem dados a respeito de seu comportamento físico, tanto na estrutura de concreto como na sua fundação. Nas tabela 4.18 e 4.19, podem-se observar os tipos e quantidades de instrumentos instalados no concreto e na fundação dos blocos do trecho F.

FIGURA 4.18 - VISTA AÉREA DA USINA HIDRELÉTRICA DE ITAIPU, CIRCULADO O TRECHO F.
FONTE: ITAIPU, 2006

TABELA 4.18 - NÚMERO DE INSTRUMENTOS NOS BLOCOS DO TRECHO F – CONCRETO

RESUMO - CONCRETO							
INSTRUMENTO	Sigla	BLOCO - F					Total por instrumentos
		5/6	13/14	15/16	19/20	35/36	
Rosetas de Deformímetros	RD	4	-	-	11	-	15
Tensômetro	TN	1	-	-	4	-	5
Rosetas de Tensômetros	RT	2	-	-	6	-	8
Medidor de Junta Interno	JM	-	-	-	7	-	7
Pêndulo Direto	PD	5	6	-	6	4	21
Pêndulo Invertido	PI	3	1	1	1	-	6
Termômetro na Massa	TM	3	-	-	17	3	23
Termômetro de Superfície	TS	2	-	-	6	2	10
Total por bloco		20	7	1	58	9	95

FONTE: ITAIPU, 2006

TABELA 4.19 - NÚMERO DE INSTRUMENTOS NOS BLOCOS DO TRECHO F – FUNDAÇÃO

RESUMO - FUNDAÇÃO																		
INSTRUMENTO		BLOCO F																Total por instrumento
		1/2	3/4	5/6	7/8	9/10	11/12	13/14	15/16	17/18	19/20	21/22	23/24	27/28	29/30	31/32	35/36	
Piezômetro Standpipe	PS	-	4	6	5	-	6	7	3	6	8	-	4	10	-	6	9	74
Piezômetro Geonor	PG	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
Extensômetro de haste	EM	4	-	1	-	-	-	3	5	4	3	1	-	4	-	-	4	29
Medidor de Aterro	MA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
Medidor triortogonal	MT	-	1	-	-	1	4	1	1	-	-	1	1	-	-	1	-	11
Célula de Pressão Total	CL	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
Medidor de vazão	MV	-	1	-	-	-	2	-	2	-	-	2	1	-	1	-	-	9
Drenos	DR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
Medidor de nível d' água	PZ	-	-	-	-	-	-	1	1	1	1	-	-	-	-	-	-	4
Total por bloco		4	6	7	5	1	12	12	12	11	12	4	6	14	1	7	13	127

FONTE: ITAIPU, 2006

4.3.1 Seleção e Obtenção das Séries Temporais do duto e barragem /Pré Processamento De Dados

No caso da ITAIPU, as séries de leituras dos instrumentos, em sua maioria, são bem comportadas, em vários casos próximos a um comportamento estacionário. Esta característica é bastante diferente do caso da TRANSPETRO, em que algumas séries possuem um grau de instabilidade elevado, mostrando mudanças abruptas em um pequeno intervalo de tempo.

Deve-se salientar ainda que são necessários procedimentos para tornar os dados de leituras acessíveis à utilização dos métodos de análise abordados neste trabalho. Cada empresa/instituição adota um padrão de armazenagem dos dados, e os programas computacionais de análise matemática e estatística normalmente exigem um único padrão. Assim, a criação de rotinas computacionais que executem a transformação dos padrões particulares de cada local/série/empresa torna-se fundamental, pois a quantia de informação é grande, o que tornaria uma tarefa manual com o mesmo fim inviável.

4.3.2 Dados da Barragem de ITAIPU

Os dados com as leituras dos instrumentos de ITAIPU encontravam-se organizados em arquivos de formato texto. Cada um dos instrumentos da barragem foi identificado durante a fase de projeto com uma sigla. Após sua instalação e posterior funcionamento, estes instrumentos receberam outra denominação, também em forma de sigla, diferente daquela usada na fase de projeto. Este fato dificultou de certa maneira a identificação inicial de cada instrumento. Desta forma, foi construído um banco de dados em formato Excel, onde se procurou relacionar as informações de projeto com as constantes nos arquivos texto das leituras dos instrumentos. Este banco de dados foi construido baseado nas pranchas do projeto original de ITAIPU e de relatórios técnicos relacionados à instrumentação da barragem. Maiores detalhes podem ser encontrados em (ANDRAOS, 2006) e (SANCHEZ, 2006).

A figura 4.19 mostra um exemplo de como os dados encontravam-se organizados originalmente, em grandes arquivos textos contendo as séries históricas das leituras de todos os instrumentos, onde as colunas armazenam respectivamente: o

código do instrumento, a data e hora da leitura e as medições obtidas; nesse caso é de um pêndulo direto, onde as duas últimas colunas são as leituras obtidas.

Também foi necessário entender como são obtidas em campo as medidas dos instrumentos aqui avaliados e o significado de cada grandeza obtida por eles.

Verificou-se que cada tipo de instrumento possuía uma periodicidade de leituras (diária, semanal, mensal e trimestral), e que são feitas campanhas de leituras de instrumentos de mesmo tipo, ou seja, primeiro se faz a leitura de todos os instrumentos tipo “a”, depois todos do tipo “b”, etc., até se obterem as leituras de todos os instrumentos. Nos procedimentos realizados neste trabalho, onde se busca comparar as leituras entre instrumentos, não houve problemas em fazer comparações entre instrumentos do mesmo tipo, já que as leituras em sua maioria eram feitas no mesmo dia, algumas com intervalo máximo de dois dias. Contudo, houve dificuldade em obter leituras de instrumentos diferentes ocorridas no mesmo dia, o que poderia inviabilizar uma comparação entre as leituras desses instrumentos.

```
COF21  ;19/05/1982;0800;    4.00;   -4.80
COF21  ;23/05/1982;0830;    4.10;   -4.70
COF21  ;26/05/1982;0840;    4.00;   -4.70
COF21  ;28/05/1982;0815;    3.90;   -4.60
COF21  ;31/05/1982;0815;    3.90;   -4.80
COF21  ;02/06/1982;0830;    4.10;   -4.80
COF21  ;04/06/1982;0830;    4.10;   -4.80
COF21  ;07/06/1982;0840;    4.10;   -4.80
COF21  ;09/06/1982;0820;    4.10;   -4.70
COF21  ;11/06/1982;0820;    4.20;   -4.90
COF21  ;14/06/1982;1005;    4.10;   -4.90
COF21  ;16/06/1982;0810;    4.20;   -4.80
COF21  ;18/06/1982;0800;    4.10;   -4.70
COF21  ;21/06/1982;0830;    4.10;   -4.70
COF21  ;25/06/1982;0845;    4.20;   -4.60
COF21  ;28/06/1982;0830;    4.10;   -5.00
COF21  ;30/06/1982;0836;    4.10;   -4.80
COF21  ;02/07/1982;0840;    4.20;   -4.80
COF21  ;05/07/1982;0845;    4.10;   -4.90
COF21  ;07/07/1982;0815;    4.00;   -5.00
COF21  ;09/07/1982;0820;    4.30;   -4.90
```

FIGURA 4.19 - PARTE DE ARQUIVO TEXTO COM AS LEITURAS DO PÊNDULO DIRETO COF21.
FONTE: ITAIPU (2006)

Mostrou-se necessária a conversão do formato texto dos dados originais para formato Excel, compatível com o formato de entrada do programa *Matlab*, que foi usado posteriormente neste trabalho. O *Matlab* foi usado para o pré-processamento das leituras, possibilitando a construção de um conjunto de rotinas capazes de selecionar e agrupar dados dessas leituras de maneira rápida e respeitando condições necessárias para a validade das análises posteriores.

4.4 Resultados dos Estudos de Caso

4.4.1 Análise de instrumentos de sítios dos oleodutos

4.4.1.1 Série piezométrica em OLAPA – Análise 1

A instalação de piezômetros objetiva determinar pressões neutras em maciços de terra ou rocha, ou subpressões em contatos com estruturas de concreto. Podem ser de tubo aberto, pneumático, hidráulico, elétrico de resistência e de corda vibrante (CRUZ, 2005).

A série a ser analisada é de um importante instrumento instalado em OLAPA, o piezômetro de leituras automatizadas 13, cuja medida pode auxiliar no conhecimento da estabilidade do talude onde ele se encontra.

O gráfico da referenciada série se encontra na figura 4.20. Nesse caso, as dados estão no formato original, cujo valor é obtido em relação a base do próprio instrumento. O formato dos dados de entrada é o mesmo da utilizada na série do capítulo anterior, onde as entrada são dadas por $[x(t)\ x(t-1)\ x(t-2)\ x(t-3)]$ e a saída $x(t+1)$, formando 1632 dados de entrada e saída, desse total, 20% foram retirados aleatoriamente para se utilizar como teste de desempenho de generalização da rede neural após o treinamento com os 80% restantes.

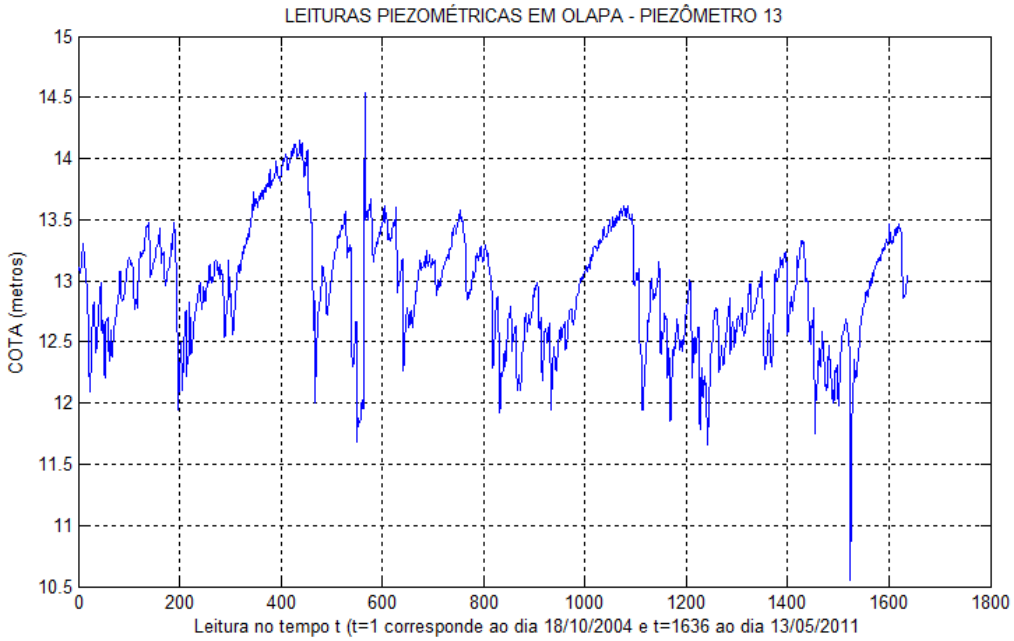


FIGURA 4.20 - SÉRIE HISTÓRICA DE LEITURAS DO PIEZÔMETRO 13

Foram comparadas as RBF do pacote computacional do MATLAB 2011, RBF com k-médias, na fase não supervisionada e o RBF hierárquico na fase não supervisionada. Até 25 neurônios, nota-se uma ligeira vantagem no RBF hierárquico, seguida da RBF com k-médias e por fim a RBF do MATLAB. De 25 até 50 neurônios, já considerado *overfitting,* os métodos RBF k-médias e do MATLAB se mostraram mais estáveis e, a partir de 50 neurônios, todas divergiram acentuadamente. O desempenho de teste de generalização das redes com diferentes arquiteturas é mostrado na figura 4.21.

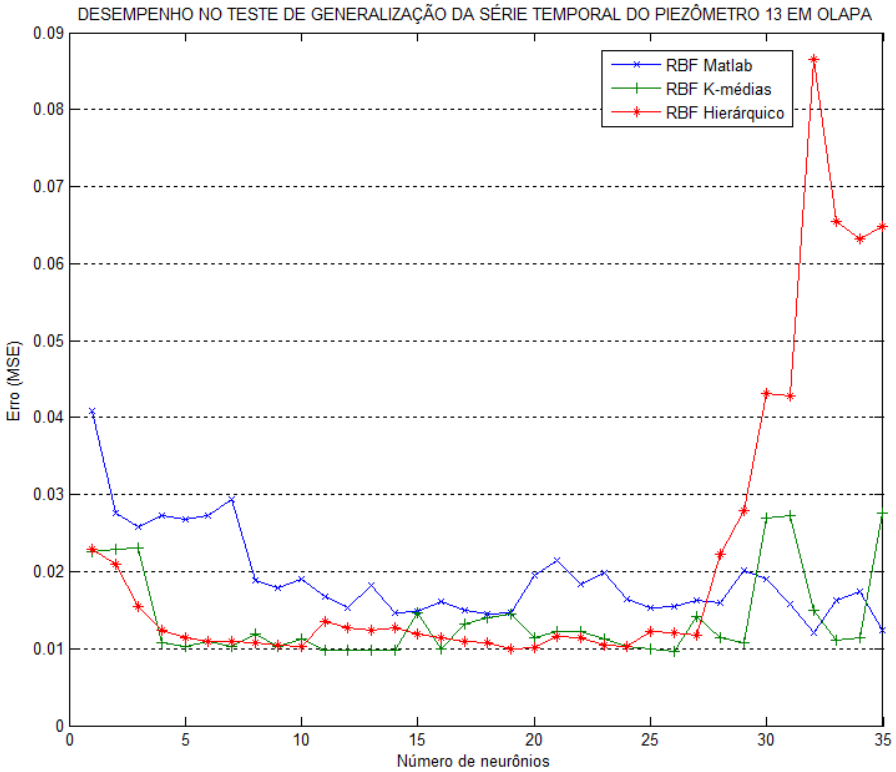


FIGURA 4.21 - DESEMPENHO DE TRÊS DIFERENTES REDES RBF'S APLICADAS A SÉRIE TEMPORAL DO PIEZÔMETRO 13

Haveria a possibilidade de se testar topologias com até 1305 neurônios, mas o estudo limitou-se a 100 neurônios por questões computacionais. Para realizar todos os testes, foram consumidos 97 segundos para o método RBF k-médias e 72 segundos para o RBF hierárquico. Com a escolha automática de arquiteturas candidatas a ideal para a série usando o método proposto, aceitando clusterização de grupos com índice de similaridade (R^2) acima de 0.9, foram testadas no máximo 10 topologias. Nesses termos ficaram estabelecidas as topologias apresentadas na tabela 4.20, que sintetiza os resultados às arquiteturas escolhidas, como a distância

de ligação para cada formação, o coeficiente R^2 dos grupos e a variação da distância de ligação de k grupos em relação $k+1$ grupos

Uma das 10 topologias indicadas pelo método foi a rede com 19 neurônios (em negrito), sendo a que mostrou melhor desempenho entre as 100 topologias testadas. Isso comprova a vantagem do método da otimização de testes de arquitetura aplicada ao RBF hierárquico , reduzindo ainda mais o tempo de processamento, passando de 100 topologias testadas para 10 topologias. Isso confere uma redução de 70% no tempo total de processamento em relação ao treinamento de todas as 100 topologias.

TABELA 4.20 - INFORMAÇÕES OBTIDAS APÓS O AGRUPAMENTO HIERÁRQUICO DOS DADOS DE ENTRADA DA SÉRIE HISTÓRICA DO PIEZÔMETRO 13

k grupos	Distância	R^2	Variação da distância em relação ao anterior
13	1,4998	0,9041	0,1515
16	1,311	0,9367	0,0835
22	1,0311	0,9506	0,0741
19	**1,1246**	**0,9479**	**0,0708**
41	0,6828	0,9777	0,0611
17	1,2275	0,9463	0,0577
25	0,9408	0,9537	0,0534
18	1,1698	0,9471	0,0452
27	0,8776	0,9543	0,0421
42	0,6217	0,9779	0,0285

4.4.1.2 Série piezométrica em OLAPA – Análise 2

Ainda analisando a série apresentada em 4.1.1.1, alterou-se o conjunto de entrada, separando as primeiras 1500 entradas e saída para treinamento e as demais 132 adiante para teste. Procurou-se então determinar a capacidade de previsão a longo prazo da rede, lembrando que as entradas são dadas por $[x(t)\ x(t-1)\ x(t-2)\ x(t-3)]$ e as saídas $x(t+1)$, formando 1632 dados de entradas e saídas. Os erros foram ligeiramente maiores do que no caso avaliado no tópico anterior, o que era de se esperar pelo fato de ser uma previsão de 132 passos à frente. A figura 4.22 mostra que o desempenho da RBF proposta, em geral,

apresentou bons resultados frente aos melhores resultados obtidos para essa série pelos métodos tradicionais. Não foi apresentado o desempenho dos modelos ARIMA por não serem competitivos nesse caso.

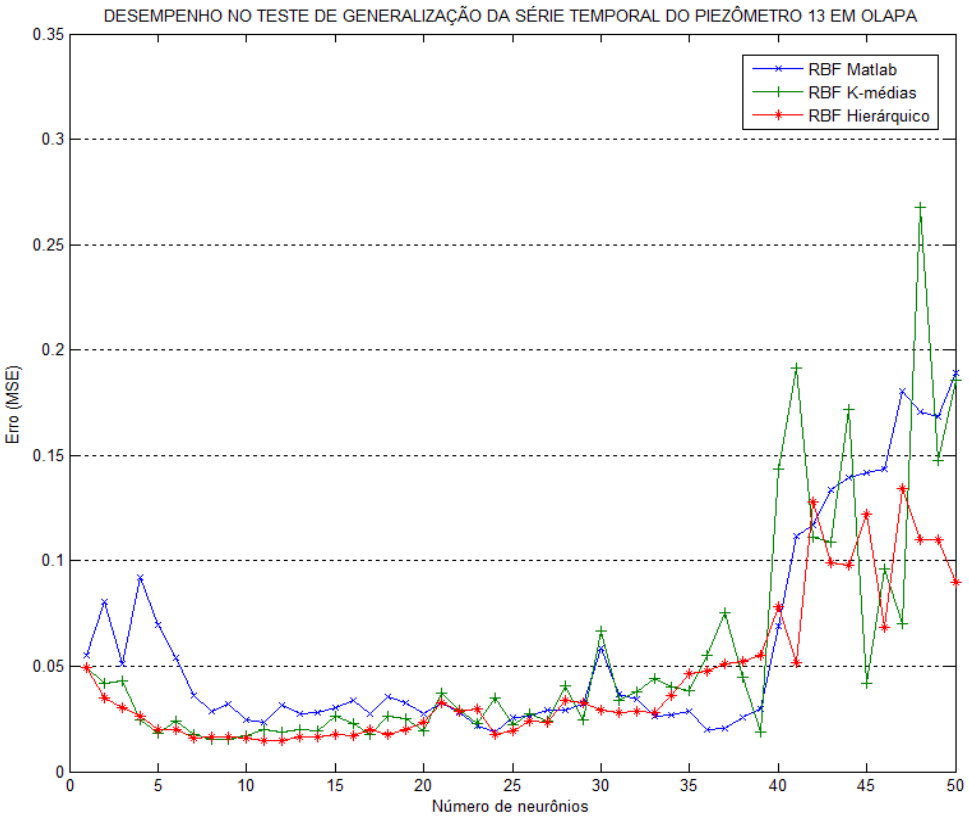


FIGURA 4.22 - DESEMPENHO DE TRÊS DIFERENTES REDES RBF'S APLICADAS A SÉRIE TEMPORAL DO PIEZÔMETRO 13 – CASO 2

O método para estabelecer topologias mais adequadas aos dados também funcionou nessa série. Isso pode ser notado na tabela 4.29, que mostra as 10 topologias candidatas. A primeira linha apresentada nessa tabela é correspondente à RBF hierárquica que obteve o melhor desempenho, como observado no ponto de erro mínimo mostado na figura 4.21.

TABELA 4.21 -INFORMAÇÕES OBTIDAS APÓS O AGRUPAMENTO HIERÁRQUICO DOS DADOS DE ENTRADA DA SÉRIE HISTÓRICA DO PIEZÔMETRO 13 – CASO 2

k grupos	Distância	R^2	Variação da distância em relação número de grupos imediatamente menor
12	**1,365296**	**0,915545**	**0,0222**
10	1,467472	0,901373	0,0072
11	1,450334	0,908596	0,0069
25	0,914635	0,951277	0,0064
26	0,824314	0,95769	0,006
29	0,784749	0,964928	0,0031
15	1,163752	0,940101	0,0028
49	0,532445	0,978295	0,0025
42	0,625591	0,973561	0,0023
13	1,17154	0,937716	0,0018

4.4.2 Resultados de Previsão em Séries de piezômetros do OLAPA

a) Previsão usando somente dados da própria série

Visando comprovar o desempenho do método proposto, foram testadas outras 24 séries de piezômetros, compostos de leituras manuais ou automatizadas, presentes no sistema de instrumentação em OLAPA, separando 20% dos dados de cada série (número de leituras, vide tabelas no capítulo 4.2), para o conjunto de teste e sem dados de validação. Foram comparadas as técnicas das redes neurais de base radial com a etapa não supervisionada usando o método de agrupamento pelo k-médias e o do pacote computacional do *Matlab,* e o RBF hierárquico (proposto). Foram calculados o erro no treinamento, erro no teste, e o número de neurônios que compõe a melhor topologia. Analisou-se ainda se o método proposto acertou na topologia que resultou no menor erro no teste. Em caso negativo, destaca-se qual topologia indicada pelo método que resultou no menor erro. Os resultados estão compilados na tabela 4.22, indicando que o método RBF Hierárquico superou os outros dois métodos em 70% das séries analisadas no quesito menor erro. Quanto à capacidade de indicar a topologia ideal, o método proposto acertou a topologia ideal em 67% dos casos, com erro de no máximo **6** neurônios em relação ao ideal (testando todas as topologias) .Pelo que se nota, em geral, o método tem dificuldades de captar topologias com número reduzido de neurônios, pois a exigência do R^2 ser acima

de 0,9 impõe uma alta densidade dos dados, e se os dados forem muito esparsos, os R^2 tendem a serem baixos para um número pequeno de grupos.

Como cada uma das séries possuíam cerca de 540 observações, as topologias poderiam variar entre 1 e 540 neurônios, e o método acertou 67% das topologias com apenas o teste de 15 topologias. Os métodos comparados possuem um erro de treinamento em geral menor que o método proposto (destacados em negrito).

Os resultados poderiam ser melhores caso se diminuísse o valor do R^2 e aumentasse o valor de K (número de topologias a serem testadas pelo método de otimização).

TABELA 4.22 - COMPARATIVOS DOS TRÊS MÉTODOS APLICADOS EM 24 SÉRIES DE LEITURAS DE INSTRUMENTOS EM OLAPA

	RBF K-MÉDIAS/MATLAB			RBF HIERÁRQUICO			O método indicou como candidato a topologia ideal? $R^2 = 0.8$, $K = 15$
	Erro (MSE) - Treinamento (30 neurônios)	Menor erro (MSE) - Teste	Número de Neurônios no Teste	Erro (MSE) - Treinamento (30 neurônios)	Menor erro (MSE) - Teste	Número de Neurônios no teste	
Série 1	**0,8378**	0,0915	3	0,9050	0,0079	12	Sim
Série 2	**0,0276**	0,1057	5	0,0385	0,0706	6	Não, 8 neurônios
Série 3	**0,5096**	0,0240	6	0,7985	0,0208	10	Sim
Série 4	**0,4303**	0,0174	20	0,5235	0,0111	3	Não, 6 neurônios
Série 5	0,0252	0,0253	9	**0,0240**	0,0287	2	Não, 6 neurônios
Série 6	**1,2247**	0,0254	9	1,3087	0,0436	8	Não, 12 neurônios
Série 7	**0,4394**	0,0276	3	0,4849	0,0274	8	Sim
Série 8	**0,0800**	0,0568	3	**0,3903**	0,0234	10	Sim
Série 9	**0,0251**	11,8261	2	0,0467	1,8750	9	Sim
Série 10	0,3784	0,0144	10	**0,3193**	0,0048	18	Sim
Série 11	**0,0873**	0,0130	5	0,1090	0,0091	8	Sim
Série 12	**0,0037**	0,0165	2	0,0043	0,0125	7	Sim
Série 13	**0,0611**	0,0757	30	0,0774	0,0342	29	Não, 35 Neurônios
Série 14	0,5668	0,5906	27	**0,3861**	1,7803	5	Não, 7 Neurônios
Série 15	3,7898	0,8738	11	3,7898	1,4227	11	Sim
Série 16	**0,1231**	0,1427	5	0,1557	0,1367	2	Não, 10 neurônios
Série 17	**0,1121**	0,3403	2	0,1740	0,2092	9	Sim
Série 18	**0,0420**	0,2994	3	0,0669	0,3484	8	Sim
Série 19	0,0535	0,3678	11	**0,0500**	0,3489	7	Sim
Série 20	3,8959	0,6110	10	3,8959	3,9040	19	Sim
Série 21	8,6237	2,5293	3	**8,6186**	0,9529	13	Sim
Série 22	0,0894	0,0086	11	**0,0795**	0,0086	9	Sim
Série 23	**0,4243**	0,0169	30	0,4858	0,0106	17	Sim
Série 24	**0,0176**	0,1990	5	0,0264	0,0532	2	Não, 8 neurônios

b) Previsão com uso de co-variável (Pluviometria) em rbf's aplicadas a previsão de séries temporais

De maneira geral, imagina-se que uma previsão pode ser aperfeiçoada se um dado método de previsão incluir outras informações além da própria série a ser predita. Essas informações adicionais devem ser escolhidas de maneira adequada na previsão da série, caso contrário, terão o efeito oposto ao desejado, perturbando o desempenho dos modelos.

No caso da previsão da série histórica de um piezômetro, sabe-se que as poro-pressões possuem uma forte relação com a precipitação pluviométrica, mas nem sempre de maneira imediata. Isso ocorre porque os piezômetros estão localizados no subsolo, e uma chuva que ocorre num dado dia demora um certo tempo para surtir efeito nas leituras piezométricas, pois é necessário que a água superficial infiltre no terreno, modifique a poro-pressão e então mobilize o equipamento. Isso pode levar à conclusão equivocada de não haver relação entre essas duas grandezas, pois o valor da correlação direta imediata é próxima de zero.

Por esse motivo, foi calculada a correlação das leituras diárias dos piezômetros e drenos com a pluviometria atrasada e acumulada em vários níveis, lembrando que, a pluviometria acumulada é a somatória da pluviometria ocorrida em n dias anteriores ao em questão. Pluviometria atrasada é simplesmente o valor da pluviometria medido n dias atrás, e nos dois casos, foram testados até 90 dias. A que resultou em maior correlação em relação aos instrumento foi a escolhida como co-variável na rede neural. O resultado das correlações diretas, as atrasadas e as acumuladas que resultaram na maior correlação entre a pluviometria e alguns instrumentos se encontra na tabela 4.23.

Os resultados apresentados nessa tabela são úteis na melhoria da previsão de uma série temporal. Por exemplo, a correlação direta da pluviometria com as leituras do piezômetro 15 é 0,2, valor considerado muito baixo, mas esse valor sobe para 0,7 se for calculada a correlação da pluviometria acumulada dos últimos 30 dias, contados a partir de dois dias anteriores à leitura do piezômetro. Então, a pluviometria deve ser usada como covariável nesse formato para aumentar as chances que essa inclusão resulte em melhoria no desempenho dos modelos neurais. Conclusões diretas podem ser obtidas com essa tabela, com o atraso de resposta de um evento pluviométrico agudo nas leituras dos piezômetros. Os tamanhos das janelas de tempo e os atrasos que resultam para cada

instrumento naturalmente são diferentes. Dependem da posição em que se encontram no sítio, da geologia local, das condições de drenagem, e outros fatores.

TABELA 4.23 - RELAÇÃO DAS CORRELAÇÕES DIRETAS, ATRASADAS E ACUMULADAS DA PLUVIOMETRIA VERSUS LEITURAS DE PIEZÔMETROS E DRENOS.

	Correlação direta com pluviometria	Correlação acumulada com pluviometria	Dias acumulados anteriores ao dia da leitura do instrumento	Dias atrasados ao dia da leitura do instrumento
Piezômetro PZM 001	0,1	0,19 (Não significativo)	61	4
Piezômetro PZM 002	0,13	0,64	9	1
Piezômetro PZM 003	0,1	0,44	47	9
Piezômetro PZM 004	0,17	0,73	30	1
Piezômetro PZM 005	0,14	0,65	7	4
Piezômetro PZM 006	0,14	0,5	31	9
Piezômetro PZM 007	0,13	0,63	38	9
Piezômetro PZM 008	0,11	0,5	14	8
Piezômetro PZM 009	0,12	0,15 (não significativo)	3	1
Piezômetro PZM 010	**0,16**	0,45	3	2
Piezômetro PZM 011	**0,2**	0,25	19	4
Piezômetro PZM 012	0,15	0,64	61	3
Piezômetro PZM 013	0,14	0,22 (não significativo)	19	11
Piezômetro PZM 014	0,12	0,13 (não significativo)	45	11
Piezômetro PZM 015	0,2	0,7	30	2
Piezômetro PZM 016	0,21	0,67	44	8
Piezômetro PZM 017	0,17	0,4	54	5
Piezômetro PZM 018	0,14	0,54	61	11
Piezômetro PZM 019	0,19	0,58	17	2
Piezômetro PZM 020	0,19	0,43	9	1
Piezômetro PZM 021	0,12	0,25 (não significativo)	3	1
Piezômetro PZM 022	0,13	0,39	44	4
Piezômetro PZM 023	0,11	0,36 (não significativo)	61	11
Piezômetro PZM 024	0,14	0,18 (não significativo)	61	8
Piezômetro PZM 025	0,29	0,64	16	2
Piezômetro PZM 026	0,17	0,2 (não significativo)	32	1
Piezômetro PZM 027	0,16	0,45	61	11
Piezômetro PZM 028	0,15	0,25 (não significativo)	6	1
Piezômetro PZM 029	0,15	0,32 (não significativo)	61	5
Dreno DHP 015a	0,47	0,73	15	1
Dreno DHP 016	0,11	0,83	4	1
Dreno DHP 047	0,24	0,69	4	1
Dreno DHP-049A	0,67	0,77	2	0

A fim de verificar se a inclusão de uma covariável realmente melhora a previsão dos métodos, foi usada como exemplo a série de leituras do piezômetro 15, cujas leituras podem ser visualizadas pela figura 4.23, este instrumento é automatizado, com leituras a cada hora, mas aqui, estão agrupados em dias (média de leituras do dia). O detalhamento das correlações acumuladas em janelas de tamanhos diferentes, contadas a partir de dois dias de atraso, se encontram na figura 4.24.

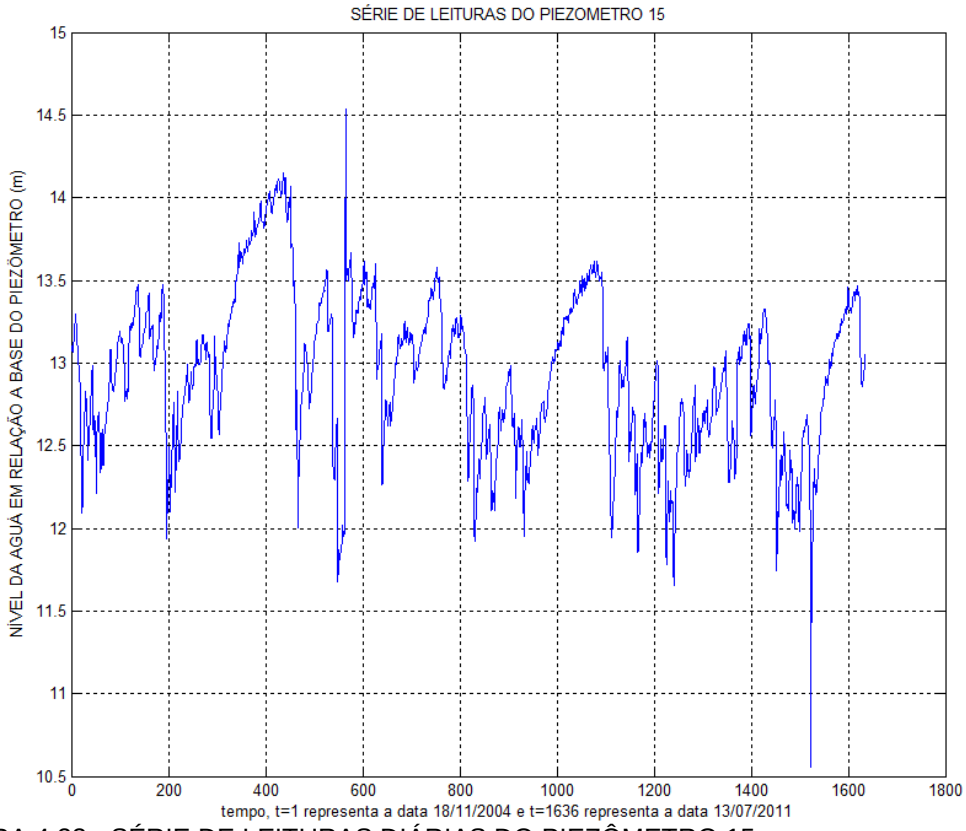


FIGURA 4.23 - SÉRIE DE LEITURAS DIÁRIAS DO PIEZÔMETRO 15

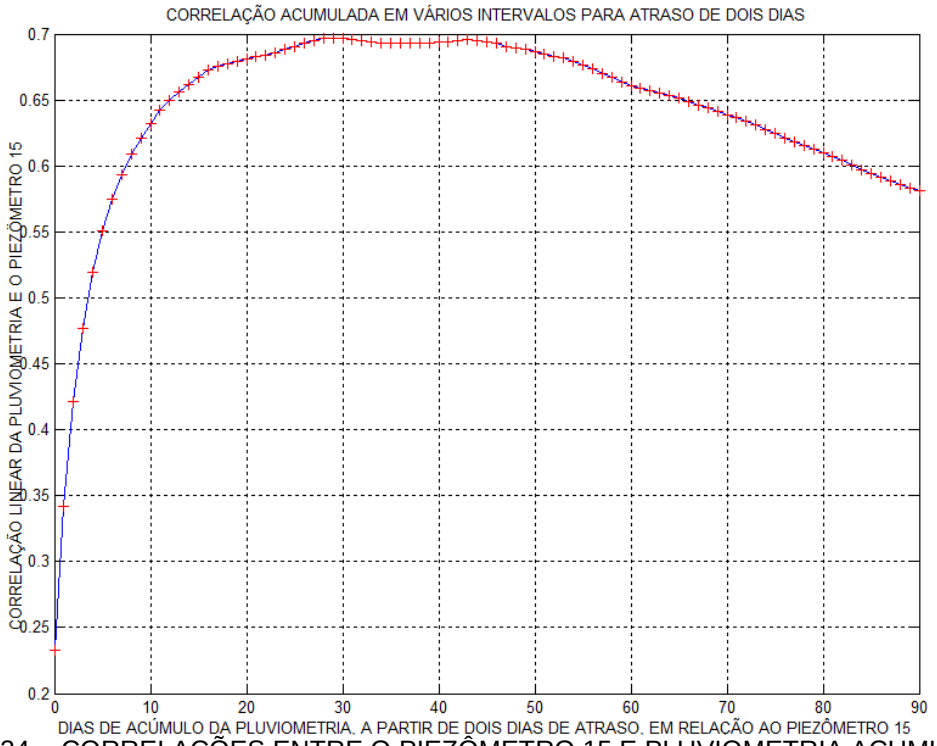


FIGURA 4.24 - CORRELAÇÕES ENTRE O PIEZÔMETRO 15 E PLUVIOMETRIA ACUMULADAS

Usou-se o mesmo formato dos dados de entrada usados nas séries temporais anteriores, e incluindo a pluviometria acumulada de 30 dias, contado a partir de dois dias anteriores ao dado a ser predito. Assim, o vetor de entrada possuia quatro elementos referente às ultimas quatro leituras anteriores à predita e um elemento com a pluviometria acumulada, totalizando um vetor de tamanmho 5. Separaram-se 80% dos dados de treinamento e 20% para o teste. Estes dados foram escolhidos aleatoriamente ao longo de toda a série. A comparação do desempenho nos dois casos se encontra na figura 4.25.

O resultado da inclusão da pluviometria acumulada como co-variável não é considerado positivo nesse caso, pois o erro no teste da RBF foi maior que o erro usando a RBF sem a co-variável (vide figura 4.25). O fato das leituras do piezômetro serem diárias torna a inclusão da pluviometria desnecessária, pois o efeito da pluviometria está inserida.

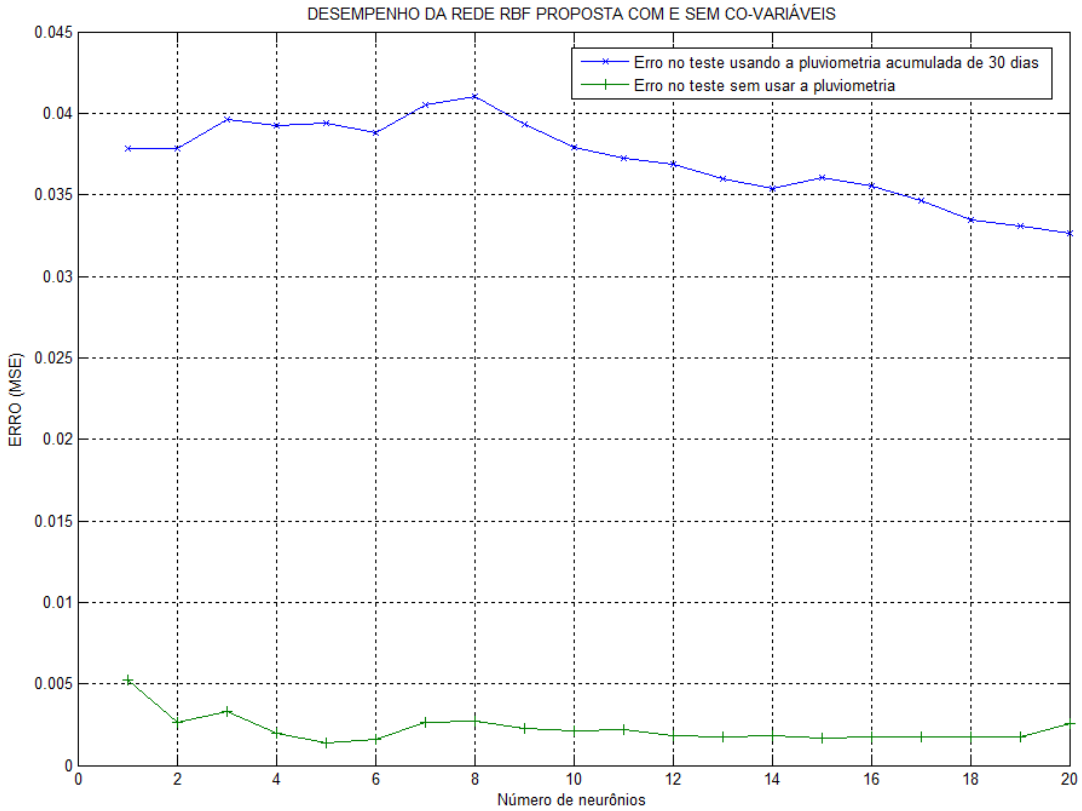


FIGURA 4.25 - DESEMPENHO NO TESTE DE GENERALIZAÇÃO PELO MÉTODO PROPOSTO COM E SEM COVARIÁVEIS

Porém, com séries temporais com leituras mais espaçadas, com leituras manuais mensais por exemplo, a inclusão da co-variável torna-se atraente. Imaginando essa mesma série de PZ-15, mas apenas com as leituras ocorridas no primeiro dia de cada mês, reduz-se a série para 72 observações (FIGURA 4.26). As duas últimas observações foram usadas para teste da rede, as primeiras 70 para treinamento, com o intuito de obter uma previsão de passo 2. Sobre os dados de treinamento foi feita a análise da correlação atrasada e acumulada com a pluviometria acumulada e atrasada, que resultou nos valores apresentados na próxima tabela (4.24). Indicando algo pouco diferente do caso do piezômetro captando as leituras diárias, 25 dias de acumulados contados a partir de 4 dias atrás resulta na maior correlação.

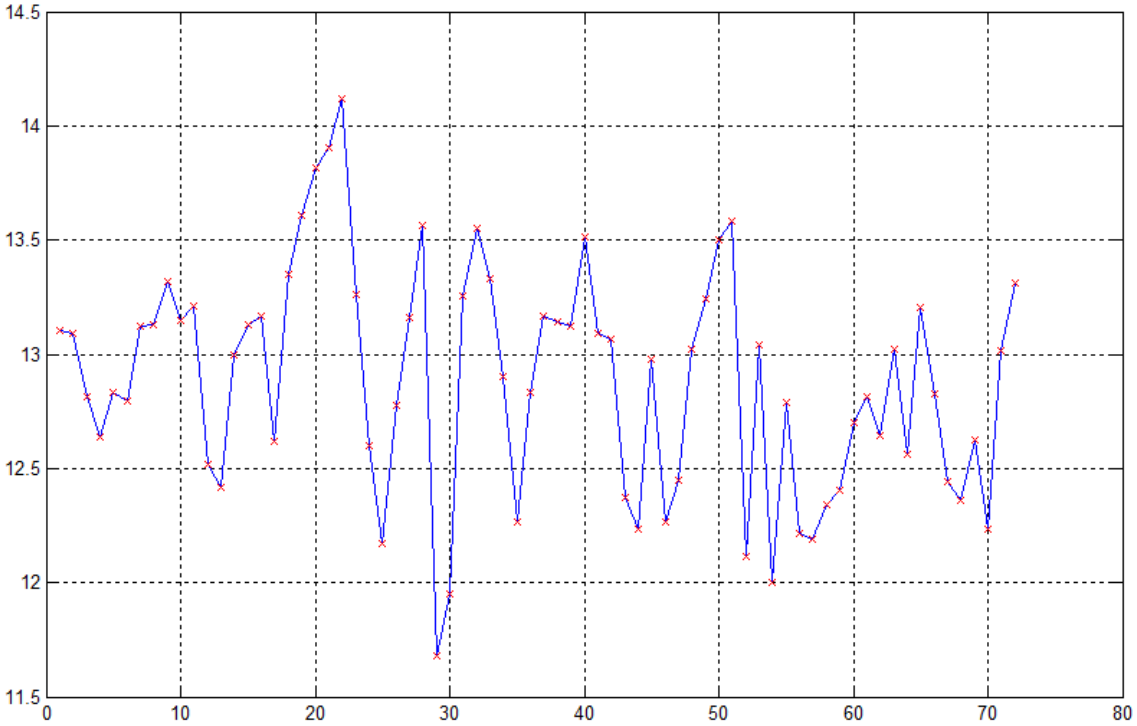


FIGURA 4.26 - SÉRIE DE LEITURAS MENSAIS DO PIEZÔMETRO 15

TABELA 4.24 - CORRELAÇÕES ATRASADAS E ACUMULADAS ENTRE A PLUVIOEMETRIA E O PIEZÔMETRO 15

dias acumulados \ dias de atraso	0	1	2	...	24	25	...	39	40
0	0,13	0,11	0,17	...	0,51	0,52	...	0,52	0,52
1	0,10	0,13	0,26	...	0,57	0,57	...	0,56	0,56
2	0,13	0,22	0,33	...	0,59	0,60	...	0,58	0,57
3	0,19	0,28	0,38	...	0,61	0,61	...	0,58	0,58
4	0,22	0,30	0,38	...	0,61	0,65	...	0,58	0,59
5	0,23	0,31	0,38	...	0,61	0,61	...	0,58	0,59
6	0,23	0,31	0,38	...	0,61	0,61	...	0,58	0,59
7	0,23	0,31	0,38	...	-0,61	0,61	...	0,58	0,59
8	0,23	0,31	0,37	...	0,60	0,60	...	0,58	0,58
9	0,24	0,31	0,37	...	0,60	0,60	...	0,58	0,58
10	0,23	0,30	0,36	...	0,59	0,59	...	0,58	0,58

Na figura 4.27, apresentam-se as correlações obtidas para quatro dias de atraso, acumulando de 0 até 40 dias, onde a maior correlação encontrada foi de 0,65, indicando como a covariável deve ser inserida na rede neural para aumentar as chances de melhoria no desempenho de previsão.

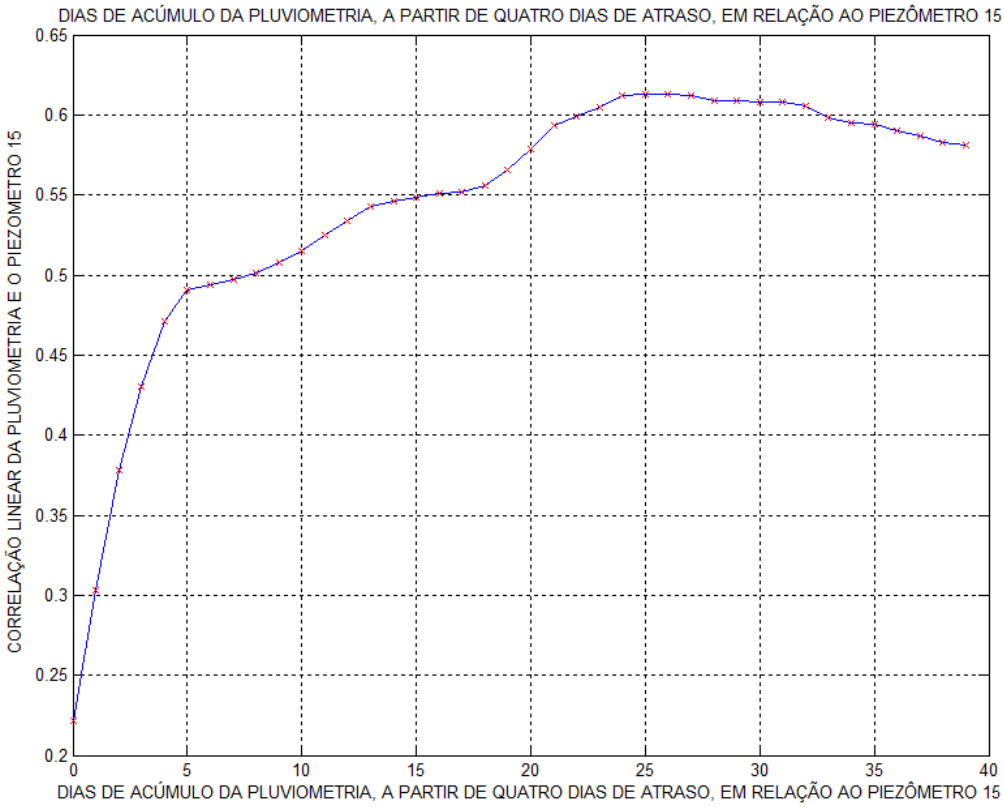


FIGURA 4.27 - CORRELAÇÕES ENTRE O PIEZÔMETRO 15 E PLUVIOMETRIA ACUMULADAS

Na figura 4.28, apresentam-se os erros obtidos no teste de previsão da série usando a RBF com e sem covariável. Nesse caso, o ganho no desempenho é evidente quando se inclui a covariável de pluviometria. A inclusão da co-variável é mais atraente quando as observações da série são longamente espaçadas, e a covariável possui mais observações intermediárias, o que aumenta a chance de se captar perturbações nas leituras do piezômetro diante da variação antecipada da pluviometria. Isso é diferente do caso de leituras piezométricas diárias, em que perturbações imediatas de precipitação não se propagam e as acumuladas já estão inseridas indiretamente nas medidas de poro-pressão.

A principal limitação da inclusão de covariáveis dessa maneira é a necessidade da utilização dos dados da pluviometria próximo ao dia da previsão, pelas correlações atrasadas e acumuladas evidenciadas, algo em torno de 9 dias antes da previsão pretendida, o que torna a previsão muito curta em relação a previsão sem covariável, que é de 30 dias, mas a inclusão pode ser útil na geração de cenários simulados,

Como se tem leituras piezométricas um mês antes da previsão pretendida, pode-se simular uma série de valor pluviométricos que poderiam ocorrem até os

dias necessários para a utilização da rede neural com as co-variáveis, respeitando os intervalos conchecidos na região, e estabelecer um intervalo de valores piezométricos que poderiam ocorrem.

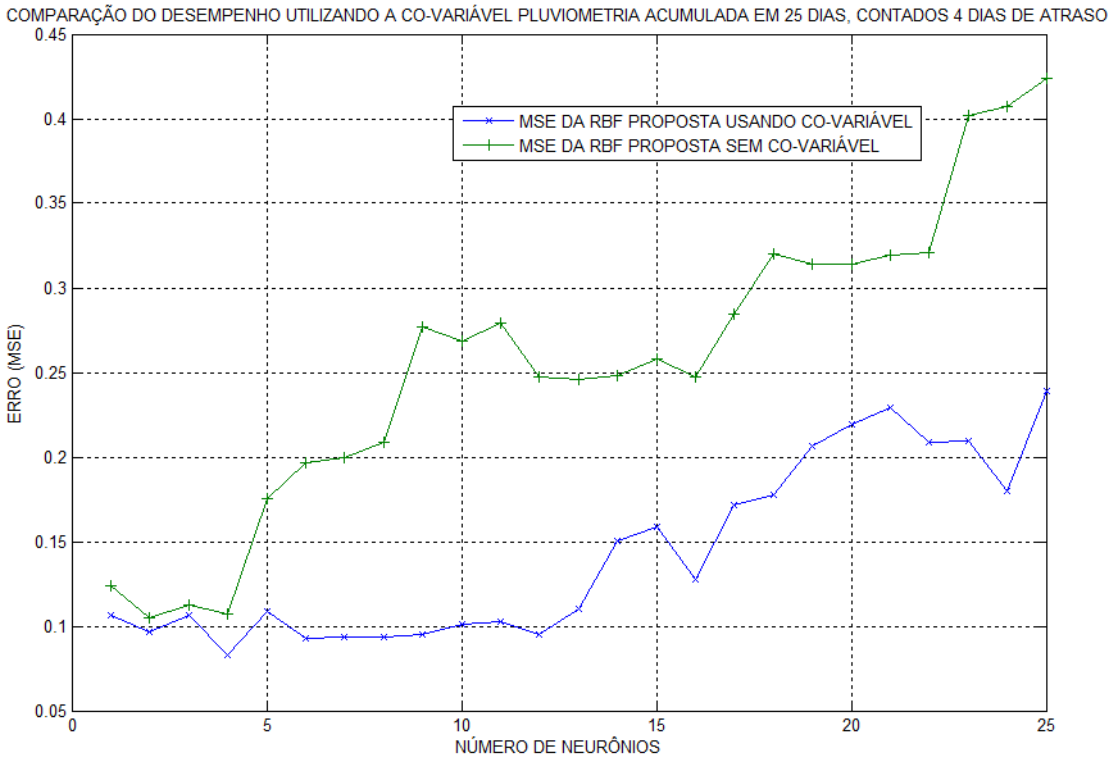


FIGURA 4.28 - COMPARAÇÃO NO DESEMPENHO DA RBF COM E SEM CO-VARIÁVEIS

4.4.3 Resultados de Série de Itaipu - coordinômetro cof22 do pendulo pdf20

4.4.3.1 Série histórica usada e resultados

O pêndulo é um instrumentos que mede dois movimentos da barragem no local onde está instalado: o deslocamento na direção do fluxo da água, e o normal ao fluxo da água. Aqui é analisada a série histórica do deslocamento na direção do fluxo da água, de um instrumento específico instalado no trecho F de Itaipu: o coordinômetro cof22 do pendulo pdf20

Apesar da disponibilidade de leituras entre 1982 e 2006, optou-se em incluir nas análises leituras ocorridas entre 1996 e 2006, a fim de evitar leituras que sofreram influências da época de enchimento da barragem. Com isso, 131 leituras constituem a série, e o gráfico da mesma se encontra na figura 4.29.

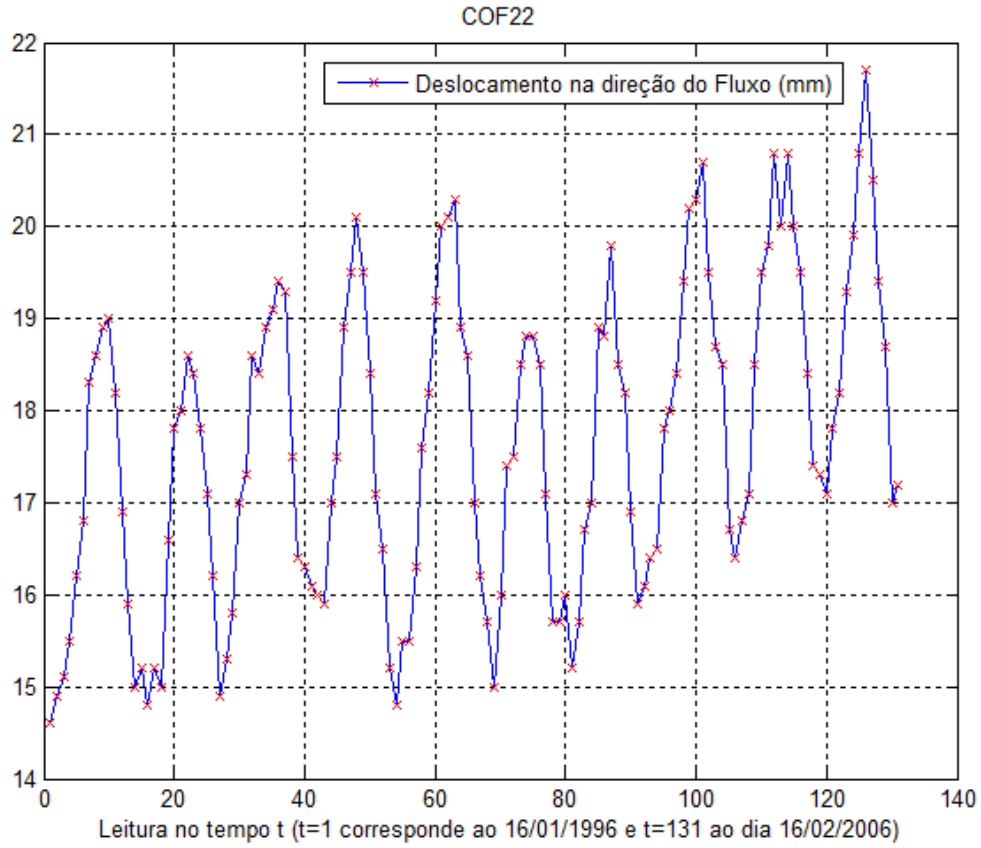


FIGURA 4.29 - SÉRIE TEMPORAL DO COORDINOMETRO COF22 DO PÊNDULO DIRETO PDF20.

A rede neural RBF hierárquica obteve o melhor resultado entre as técnicas comparadas, a figura 4.30 apresenta o gráfico do desempenho das três diferentes técnicas RBF, e pelo que se vê nesse gráfico, a rede neural RBF hierárquica com 11 neurônios obteve o melhor resultado, as entradas são dadas por $[x(t)\ x(t-1)\ x(t-2)\ x(t-3)]$ e as saídas $x(t+1)$, obtendo 116 pares de entrada e saída de a guardando as últimas 11 observações para validação, com isso, impondo uma previsão de passo 11.

Quanto à técnica proposta para determinar as topologias mais indicadas ao problema, novamente nesse exemplo a RFB hierárquica se mostra funcional, pois a rede com 11 neurônios está inclusa entre as dez topologias apontadas pelo método. A tabela 4.25 apresenta as 10 topologias escolhidas (em negrito a que obteve melhor resultado), bem como seus parâmetros.

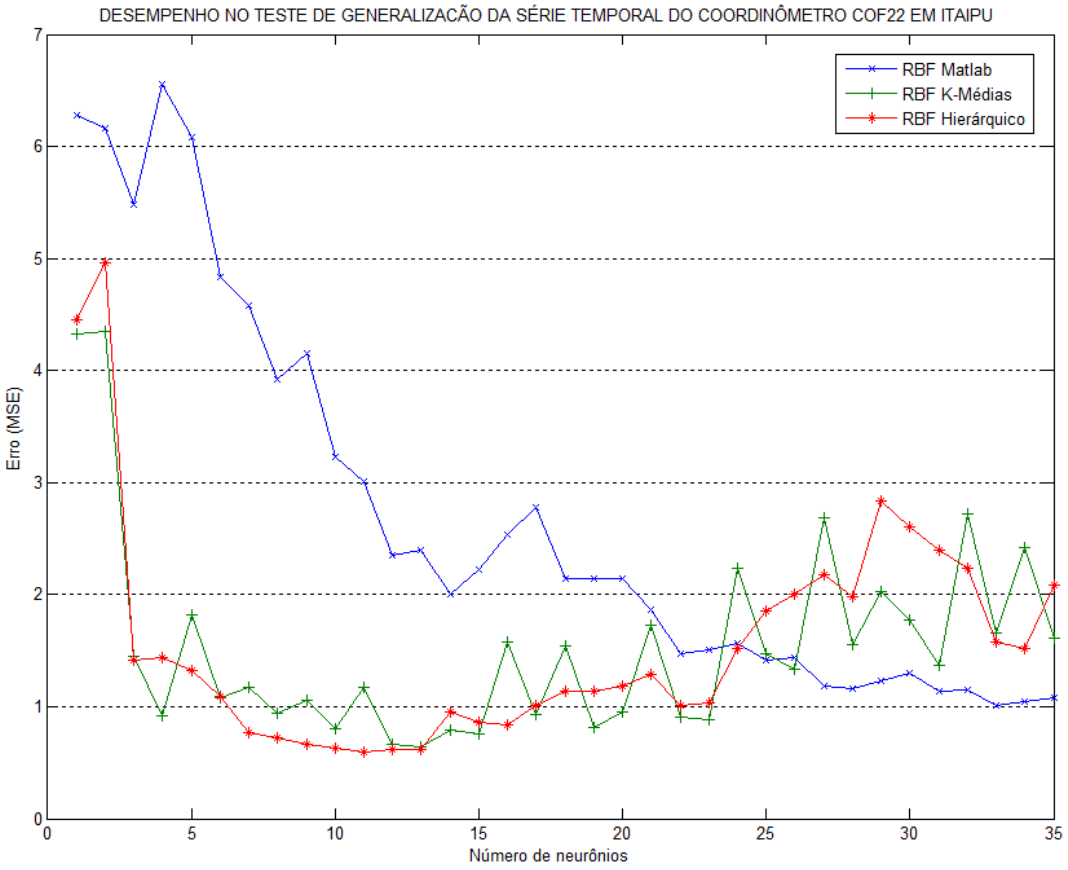


FIGURA 4.30 - DESEMPENHO DE TRÊS DIFERENTES REDES RBF'S APLICADAS A SÉRIE TEMPORAL DO COORDINÔMETRO COF22 EM ITAIPU

TABELA 4.25 - INFORMAÇÕES OBTIDAS APÓS O AGRUPAMENTO HIERÁRQUICO DOS DADOS DE ENTRADA DA SÉRIE HISTÓRICA DO COORDINÔMETRO COF22 EM ITAIPU

k grupos	Distância	R^2	Variação da distância em relação número de grupos imediatamente menor
10	3,05123	0,90363	0,01016
14	2,50799	0,92935	0,0082
11	**2,98831**	**0,91379**	**0,00589**
13	2,67021	0,92376	0,00559
12	2,83725	0,91968	0,00408
21	1,94679	0,95316	0,00326
22	1,94422	0,95642	0,00315
19	2,0567	0,94782	0,00272
16	2,24277	0,94025	0,0027
15	2,48596	0,93755	0,0027

4.4.3.2 Série histórica com remoção de tendência e resultados

Analisando visualmente o gráfico na figura 4.31, é possível perceber uma leve tendência de crescimento da série, o que induz à ideia de remover essa tendência antes de se executar o treinamento das redes neurais, pois redes neurais têm dificuldade de captar tendências. Aplicando a remoção da tendência, a nova série temporal passa a ser a representada pelo gráfico 4.31. Usando essa nova série, os modelos de previsão baseados em RBFs foram treinados e foram feitas as previsões. Estes valores foram novamente transformados, aplicando-se a inversa da função da remoção da tendência da série e, com isso, obtendo a previsão da série temporal do gráfico na figura 4.35.

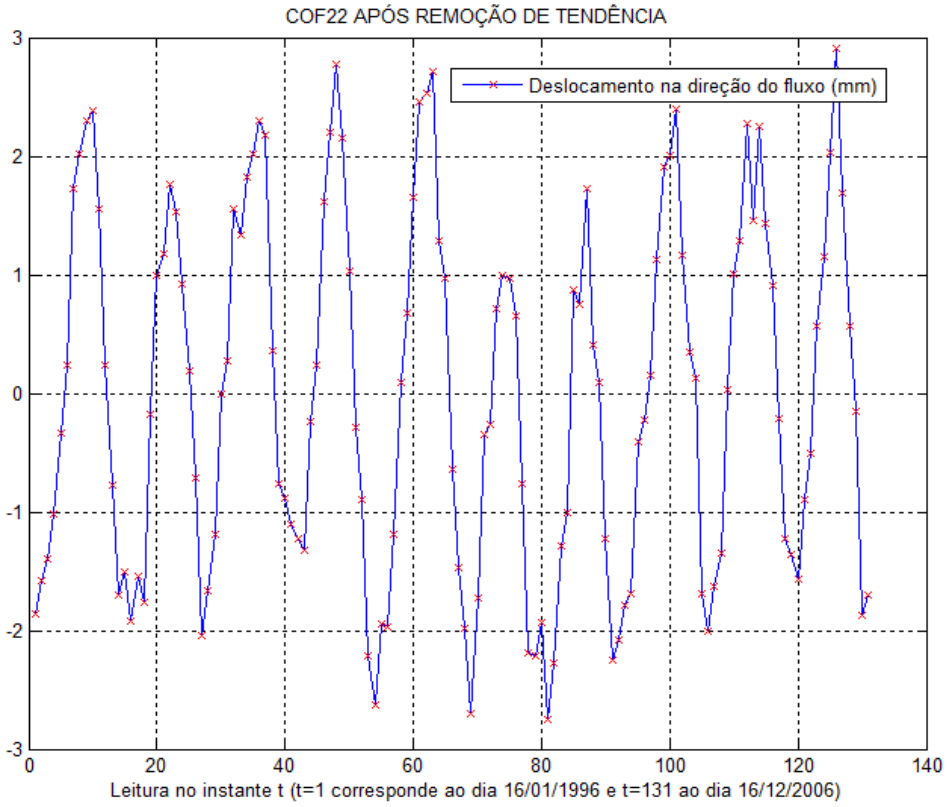


FIGURA 4.31 - SÉRIE TEMPORAL DO COORDINOMETRO COF22 DO PÊNDULO DIRETO PDF20, APÓS REMOÇÃO DE TENDÊNCIA

Analisando o gráfico do desempenho da generalização da série em questão, apresentado na figura 4.32, o desempenho foi superior ao encontrado nas redes neurais treinadas com a série sem remoção de tendência. A rede com o melhor desempenho foi a com seis neurônios, topologia que é indicada pelo método proposto se fosse reduzido o R^2 de 0.9 para 0.85.

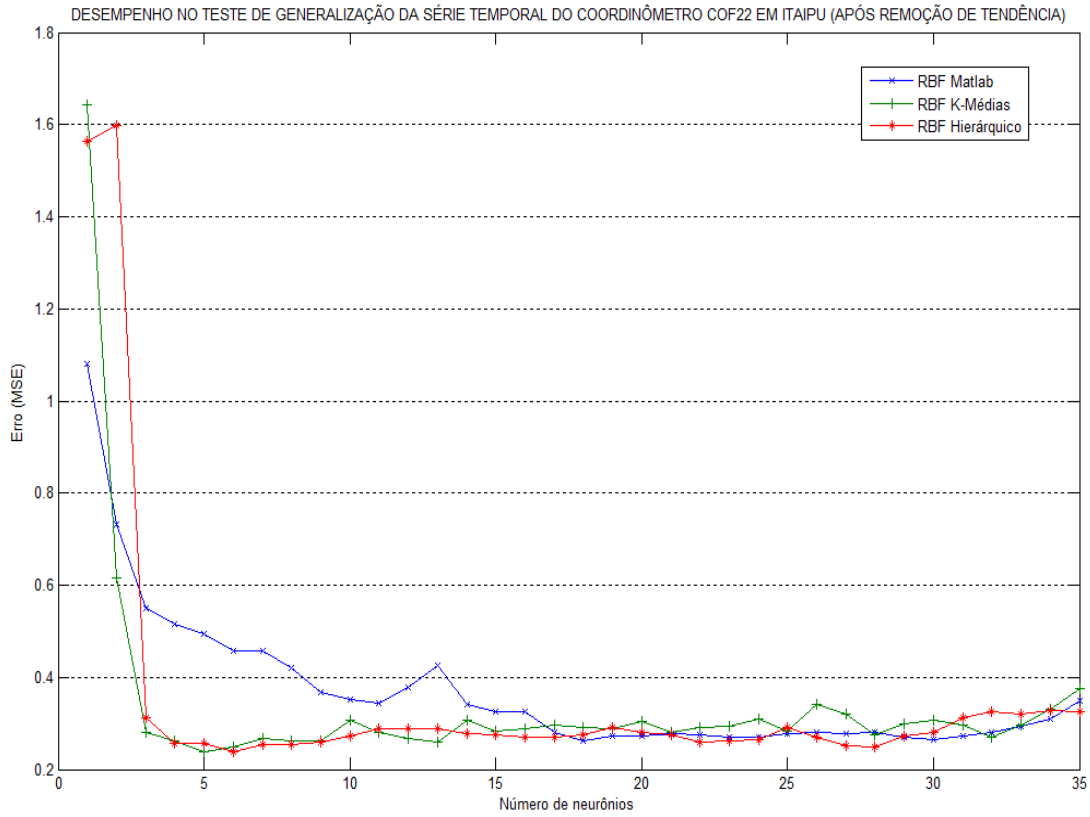


FIGURA 4.32 - DESEMPENHO DE TRÊS DIFERENTES REDES RBF'S APLICADAS A SÉRIE TEMPORAL DO COORDINÔMETRO COF22 EM ITAIPU COM TENDÊNCIA REMOVIDA

A figura 4.33 mostra as duas melhores aproximações obtidas para as previsão dessa série, a rede neural RBF hierárquica e o modelo ARIMA(2,1,2), onde a rede RBF obteve um MSE de 0.2128 e o ARIMA(2,1,2) de 0.9448. Apesar de a previsão ser de 11 passos à frente, o que aumenta a chance de ocorrer erros maiores nas aproximações, os dois métodos se mostraram estáveis, captando muito bem a tendência de crescimento de decrescimento da função. Porém, não foram capazes de modelar o pico da série, provavelmente pelo fato de esse ser o maior valor da série e nenhum semelhante ter entrado nos dados de treinamento, gerando uma situação de extrapolação.

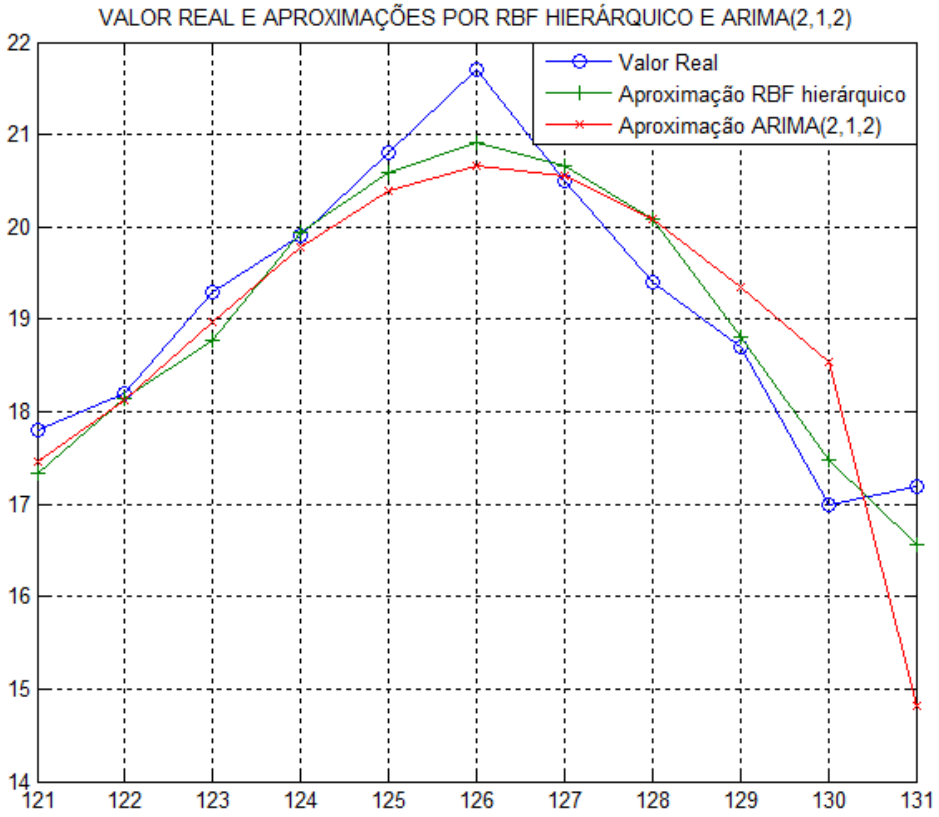


FIGURA 4.33 - PREVISÃO OBTIDA PELA RBF HIERÁRQUICA E ARIMA(2,1,2) ONZE PASSOS A FRENTE DA SÉRIE HISTÓRICA DO COORDINÔMETRO COF22

5 RESULTADOS

5.1 CONCLUSÕES

Esse trabalho apresenta uma nova proposta para a etapa não supervisionada em redes neurais de funções de base radial aplicada à previsão de séries temporais, propondo a substituição dos métodos de agrupamentos não hierárquicos pelos hierárquicos na clusterização dos dados. Também se propõe a utilização das informações geradas por esses agrupamentos hierárquicos na determinação da topologia da rede neural mais adequada aos dados em questão, problema esse muito recorrente em redes neurais.

O método proposto foi aplicado em séries temporais já estudadas na bibliografia e em séries temporais reais de sistemas de instrumentação da Usina Hidrelétrica de Itaipu e de Oleodutos da TRANSPETRO, e seus resultados foram comparados às técnicas mais utilizadas na atualidade.

Como foram utilizadas séries reais, houve a necessidade da criação de diversas rotinas computacionais para tornar possível a aplicação do método proposto, como a captura, organização, limpeza e formatação da grande massa de dados disponíveis. Outra característica comum aos dois casos reais é a ocorrência de alguns casos de leituras faltantes ou sem uma periodicidade constante, característica fundamental para correta aplicação dos modelos de previsão, e nesses casos foram utilizadas técnicas numéricas e estatísticas para interpolar esses dados em pontos de interesse.

Séries com tendência também foram um problema para a aplicação de redes neurais, o que impôs a necessidade de utilizar transformações inversíveis nessas séries para remover tais tendências, de maneira automática e integrada ao método. Foi utilizada a matriz pseudo inversa nos ajustes dos pesos sinápticos entre camada intermediária e a de saída. Por vezes ocorreu matrizes próximas da singularidade, isso leva a altos erros numéricos, prejudicando as previsões, o que foi contornado pela técnica da decomposição em valores singulares dessas matrizes.

Como o número de séries testadas foi muito elevado, todas as rotinas foram programadas de forma que, caso houvesse alguma interrupção, seja por queda de energia, dados inválidos ou problemas no *hardware,* o programa permitisse a retomada do programa de onde ele parou, não perdendo os resultados anteriores.

Após essas preparações e a aplicação do método proposto, os resultados encontrados foram promissores, pois mostram bom desempenho tanto no treinamento quanto nos testes, indicando que existe um grande potencial para o desenvolvimento e/ou melhoramento de técnicas e algoritmos matemáticos e estatísticos que possam agregar confiabilidade no processo de previsão dos dados de instrumentação.

Pôde-se gerar uma rede neural RBF, chamada de hierárquica, com poder aproximador superior ou equivalente às redes neurais RBF com os melhores desempenhos obtidos até o momento na literatura. Ainda atingiu-se um tempo de processamento inferior, que fica reduzido quando treinadas e testadas apenas as topologias indicadas pelo método proposto, ainda garantindo os melhores resultados possíveis.

A técnica proposta tem apelo prático, por envolver séries temporais de importantes à segurança de obras de engenharia, e as técnicas aqui propostas podem ser estendidas a outras obras instrumentadas.

Os problemas de perda de dados, dados duvidosos, dados em intervalos de tempo diferentes, formatação de dados, imprecisão numérica nos cálculos, que foram contornados pelas rotinas de pré-processamento de dados merecem destaque nesse trabalho, pois certamente podem ocorrem em outras aplicações.

O fato da RBF do pacote computacional *Matlab* mostrar um resultado em média inferior aos métodos confrontados pode ser por dois motivos. Primeiro, o fato do *Matlab* não calcular a matriz pseudo inversa usando decomposição em valores singulares, e/ou escalonamento, e que provavelmente cairá em erros numéricos que prejudicaram a previsão. O segundo motivo é devido à técnica de agrupamento utilizada na determinação das coordenadas dos neurônios e base radial.

A nova técnica proposta não está limitada apenas à utilização em séries temporais, e sim a qualquer problema passível de abordagem por redes neurais, se configurando como uma nova alternativa para problemas cuja determinação da topologia da rede neural RBF mais adequada tome um tempo computacional elevado.

5.2 SUGESTÕES PARA TRABALHOS FUTUROS

O método proposto pode ser paralelizado, o que o tornaria ainda mais eficiente, ficando aqui essa sugestão. Não foram utilizadas covariáveis na previsão das séries temporais de instrumentação, como por exemplo, dados hidro-metereológicos ou dados relacionados à série a ser predita, o que poderia potencializar o desempenho da previsão.

Heurísticas a serem aplicadas após a geração do dendrograma não foram completamente exploradas, certamente há espaço para novas melhorias.

Um estudo detalhado, que estabelecesse diretrizes gerais de captação dos dados de instrumentação de grandes obras de engenharia, e que tornasse mais simples a utilização deles em ambientes computacionais, poderia ser desenvolvido. Como exemplo de atributos a serem adotados neste aplicativo, estaria a captação de dados em intervalo de tempo constante, simultaneamente a outras variáveis, com possibilidade de seleção condicional de informações, de forma que essas informações ficassem numa base de dados on-line acessíveis à comunidade científica. Isso permitiria o aumento da produção de novos trabalhos associados à segurança de grandes obras.

REFERÊNCIAS

ARTHUR, D. AND VASSILVITSKII, S. (2007). **"*k*-means++: the advantages of careful seeding"**. *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*. pp. 1027–1035.

BEZDEK, J. C. **Pattern Recognition witth Fuzzy Objective Function Algorithms.** Ed. Plenum Press, New York. USA, 1981.

BISHOP, R.; ADDISON, WESLEY.**Pattern Recognition and Neural Networks**. 1996.

BOX, G. E. P. JENKINS, GM. **Time series Analysis, Forecasting and Control**, Holden-Day, San Francisco. 1976.

BOX, G. E. P. JENKINS, GM, REINSEL, G. **Time series analysis, forecasting and control**, 3rd Ed., Englewood Cliffs, NJ: Prentice Hall, 1994.

BUCHTALA, O., KLIMEK, M., SICK,B. **Evolutionary Optimization of Radial Basis Function Classifiers for Data Mining Applications**. IEEE Transactions on Systems, Man, and Cybernetics—part b: cybernetics, vol. 35, no. 5, october 2005

CHEN, J. P. **Ph.D. thesis Longitudinal and transverse response functions**, University of Virginia, 1991

COELHO, L. S.; CANCIGLIERI Jr, O. **Rede neural de base radial aplicada em previsão de séries temporais: algoritmo e aplicação.** XX ENEGEP - Construindo competências para a manufatura internacional. 2000.

CORRÊA, W.R.; PORTUGAL, M.S. **Previsão de séries de tempo na presença de mudança estrutural: Redes Neurais Artificiais e Modelos estruturais.**XVIII Internacional Symposium on Forecasting.VersãoTraduzida. Edinburg, Escócia. 1998.

DU, K. L., **Clustering: A neural network approach,** Neural Networks 23 (2010) 89_107

DUNN, J, C. A **Fuzzy Relative of the ISODATA Process and it's Use in Detecting Compact Well-Separated Clusters**. Journal of Cybernetics, vol. 3, 1973, p. 32-57.

FALCO I., CIOPPA, A. D.,IAZZETTA, A., NATALE, P. TARANTINO E. **Optimizing neural networks for time series prediction**. in Proc. 3rd On-Line World Conf. Soft Computing (WSC3). Advances in Soft Computing — Eng. Design and Manufacturing, R. Roy, T. Furuhashi, and P. K. Chawdhry, Eds, June 1998.

FARIA, E. L.; ALBUQUERQUE, M.P.; ALFONSO, J. L. G.; ALBUQUERQUE, M. P.; CAVALCANTE, J. T. P. **Previsão de séries temporais utilizando métodos estatísticos.** CBPF Index – Centro Brasileiro de Pesquisas Físicas – Novas Técnicas. 2008.

FERNANDES, L. G. L.; PORTUGAL, M.S.; NAVAUX, P.O.A. **Previsão de séries de tempo: Redes NeuraisArtificiais e Modelos Estruturais**. Pesquisa e PlanejamentoEconômico, vol 26,n. 2, p. 253-276, 1996.

GONZALEZ, J., ROJAS, I., ORTEGA, J., POMARES, H., FERNÁNDEZ, F. J., DÍAZ, A. F. **Multiobjective Evolutionary Optimization of the Size, Shape, and Position Parameters of Radial Basis Function Networks for Function Approximation**. IEEE Transactions on Neural Networks. V. 14, n. 6, Novembro 2006.
-Graduação em Engenharia de Produção, UFSC, Florianópolis, 2005.

GUSTAFSON, D. E. & KESSEL, W. C. **Fuzzy clustering witch fuzzy covariance matrix**. Proceedings of the IEEE Control and Decision Conference, San Diego, 1979, p. 761-766.

HAMILTON, J. D. **Times series analysis**. Princeton University Press.Princeton, 1994.

Han H., Chen Q., Qiao J., **An efficient self-organizing RBF neural network for water quality prediction**, Neural Networks 24 (2011) 717–725.

Kagoda P. A., Ndiritu J., Ntuli C., Mwaka B., **Application of radial basis function neural networks to short-term streamflow forecasting**, Physics and Chemistry of the Earth 35 (2010) 571–581.

HAYKIN, S. **Redes Neurais. Princípios e prática.** Porto Alegre, RS: Bookman, 2001.

HASSOUN, S. **Logic Synthesis and Verification**.Tufts University, Medford, Massachusetts, USA, 1995.

Herrera L.J., Pomares H., Rojas I., Guille´ n A., Rubio G., Urquiza J., **Global and local modelling in RBF networks**, Neurocomputing 74 (2011) 2594–2602.

http://sir-lab.usc.edu/publications/2008-ICWSM2LEES.pdf. **Descobrindo Relações entre Tags e Geotags de 2007**

KARAYIANNIS, N. B. **Reformulated radial basis neural networks trained by gradient descent**. IEEE Trans. Neural Networks, vol. 10, pp.657–671, May 1999.

KARAYIANNIS, N. B., MI, G. W. **Growing radial basis neural networks: Merging supervised and unsupervised learning with network growth techniques.** IEEE Trans. Neural Networks, vol. 8, pp. 1492–1506, Nov. 1997.

LANGARI, R., WANG, L., YEN, J. **Radial basis function networks, regression weights, and the expectation-maximization algorithm**.IEEE Trans. Syst. Man Cybern. A, vol. 27, pp. 613–623, Sept. 1997.

Lau S., Lu M., Ariaratnam S. T., **Applying radial basis function neural networks to estimate next-cycle production rates in tunnelling construction**, Tunnelling and Underground Space Technology 25 (2010) 357–365.

LIMA, F. G.; ALMEIDA, F. C. **Previsão de séries temporais financeiras com o uso das Wavelets.**XXXVIII EnANPAD 2004.Curitiba-PR. 2004.

Maier Holger R., Jain A., Dandy G. C. a, Sudheer K.P., **Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions**. Environmental Modelling & Software 25 (2010) 891- 909

MAILLARD, E. P., GUERIOT, D. **RBF neural network, basis functions and genetic algorithms**. Proc. 1997 IEEE Int. Conf. Neural Networks, vol. 4, pp. 2187–2190, 1997.

MAKRIDAKIS, S., WHEELWRIGHT, S. C. McGEE, V. E. **Forecasting: methods and applicatinos**. John Wiley & Sons, New Yourk, 1983.

MARTINELI, E., DINIZ, H., CARVALHO, A.C.P.L.F., REZENDE, S. O. **Bankruptcy Prediction Using Connectionist and Symbolic Learning Algorithms.**  IEEE World Congress on Computational Intelligence, USA, v. 1, p. 271-276, 1998.

GAN M., PENG H., DONG X., **A hybrid algorithm to optimize RBF network architecture and parameters for nonlinear time series prediction,** Applied Mathematical Modelling 36 (2012) 2911–2919.

MOODY, J.; DARKEN, C. **Neural Computation**, 1989.

MUSAVI, M. T. , AHMED, W. , CHAN, K. H., FARIS, K. B., HUMMELS, M. **On the training of radial basis function classifiers**. Neural Networks,vol. 5, no. 4, pp. 595–603, 1992.

Mustafa M.R., Rezaur R.B. Rahardjo H. Isa M.H., **Prediction of pore-water pressure using radial basis function neural network**, Engineering Geology 135–136 (2012) 40–47

PASSARI, A. F. L. **Exploração de Dados Atomizados para Previsão de Vendas no Varejo Utilizando Redes Neurais.** Dissertação de Mestrado. Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo. 2003

MONTGOMERY, D. C., JOHNSON, L. A. **Forecasting and time series analysis**.McGraw-Hill, New York, 1976.

PEDRO, L. R. M. **Como trabalhar junto com fornecedores para a redução do stock-out**. 23ª Reunião do Roundtable Brasil. CSCMP, 2006.

PORTUGAL, M. S. **Neural networks versus time series: a forecasting exercice**. Revista Brasileira de Economia, 49 (4), p. 611-629, 1995.

PORTUGAL, M. S., FERNANDES, L. G. L. **Redes Neurais Artificiais e previsão de séries econômicas: Uma introdução**. Nova Economia, vol 6, n. 4, p. 611-629, 1996.

RIBEIRO, C. de O.; SOSNOSKI, A. A. K. B.; WIDONSCK, C. A. **Redes Neurais aplicadas à previsão de preços da soja no mercado futuro.** In: Congresso Brasileiro de Economia e Sociologia Rural, 43. Anais, Ribeirão Preto: SOBER, 2005.

RIMELHART, D. E. WEIGEND, S. A. **Predicting the future: a connectionist approach**. Stanford, 25, p. 737-744. 1990.

ROJAS, I., POMARES, H., GONZÁLEZ, J., ROS, E., SALMERÓN, M., ORTEGA, J., PRIETO, A.**A new radial basis function networks structure: Application to time series prediction.** In Proc. IEEE-INNS-ENNS Int. Joint Conf. Neural Networks, S. I. Amari, C. L. Giles, M. Gori, and V. Piuri, Eds, Como, Italy: IEEE Computer Society, vol. IV, pp. 449–454, July 2000.

ROJAS, I., GONZÁLEZ, J., CAÑAS, A., DÍAZ, A., F. ROJAS, F. J., RODRIGUEZ, M. **Short-term prediction of chaotic time series by using RBF network with regression weights**. Int. J. Neural Syst., vol. 10, no. 5, pp. 353–364, 2000.

RUSSELL, S. J.; NORVIG, P. **Inteligência artificial: teoria e prática**. 2.ed. Rio de Janeiro: Campus, 2004.

SAHA, A.; KELLER, J. D.; MORGAN K. **Neural Information Processing Systems**, San Mateo, CA, 1990.

SCHARY, P. B., BECKER W. B. **The Impact of Stock-Out on Market Share:Schary, Philip B. e Martin Christopher (1979), "The Anatomy of a Stock-Out,"** Journal of Retailing,Vol. 55, No. 2, pp. 59-67, 1978.

SLOOT, L. M., PETER C. V., E PHILIP H. F**. The Impact of Brand Equity and the Hedonic Level of Products on Consumer Stock-Out Reactions**.Journal of Retailing, Vol. 81, No. 1, pp. 15–34, 2005.

SMITH, K. A.; GUPTA, J. N. D. **Neural networks in business: techniques and applications for the operations researcher**. Computers & Operations Research, p.1023-1044. Set. 2000.

SONG, S., YU, Z., CHEN, X. **A Novel radial Basis Function Neural Network for Aproximation**. Internacional Journal of Information Tecnology, v. 11, n. 9, 2005.

SOUZA, R. C., ZANDONADE, E. **Forecasting via neural networks: A comparative study**. Mimeo, Departamento de Engenharia Elétrica, PUC-RJ, 1993.

SWANSON, N. WHITE, H. **Can neural networks forecast in the bib leagues? Comparing networks forecasts to the pros**. Trabalho apresentado no XIV Internacional Symposium on Forecasting, Estocolmo, Suécia, 12 a 16 de Junho de 1994.

TAVARES, L.V., OLIVEIRA, R.C. **Investigação Operacional**, 1 ed, Alfragide, Portugal, McGraw-Hill, 1996.

TOMPKINS, J. A.; SMITH, J. D.**The warehouse management handbook**.2ª ed. Raleigh: Tompkins Press, 1998.

VIEIRA, F.H.T, LEMOS, R.P., LEE, L.L. **Aplicação de Redes NeuraisRBF Treinadas com Algoritmo ROLS e Análise Wavelet na Predição de Tráfego em Redes Ethernet**. Proceedings of the VI Brazilian Conference on Neural Networks- pp.145-150, June 2-5 - SP- Brasil, 2003.

Zhou P., Li D., Wu H., Cheng F., **The automatic model selection and variable kernel width for RBF neural networks,** Neurocomputing 74 (2011) 3628–3637

WAN, E.A. **Time series prediction by using a connectionist network with internal delay lines, In Time Series Prediction: Forecasting the Future and Understanding the past**. Addison –Wesley, pp.195-217, 1994.

WANG, M.; LAN, W. **Combined Forecast Process: Combining Scenario Analysis with the Technological Substitution Model.**Technological Forecasting & Social Change. v. 74, p. 357–378, 2007.

WASSERMAN, P. D. **Neural computing: Theory and Practice**. Van Nostrand Reinhold, New York. 1989.

WEST, M., HARRISON, J. **Bayesian forecasting and dynamic models**. Springer Verlag, segunda edição, New York, 1997.

WHITE, H. **Artificial neural networks: Aproximation and learning theory**. Blackwell Publishers, Oxford, 1992.

WHITEHEAD, B. A., CHOATE, T. D. **Cooperative-competitive genetic evolution of radial basis function centers with widths for time series prediction.** IEEE Trans. Neural Networks, vol.7, pp. 869–880, July 1996.

WONG, M. L.; LEUNG, K. S. **Data mining using grammar based genetic programming and applications**. New York: Kluwer Academic Publisher, 2002.

WOMACK, J. P.; JONES, D. T. **Lean Consumption**. Harvard Business Review, Mar 2005.

YAO X. **Evolving Artificial Neural Networks**. Proceedings of the IEEE, vol. 87, no. 9, september 1999

YEN, G. G., LU, H. **Hierarchical rank density genetic algorithm for radial-basis function neural network design**. In Proc. Congr.Evolutionary Computation (CEC), vol. 1, Honolulu, HI, pp. 25–30, 2002.

ZAMDONADE, E. **Aplicação da metodologia de Redes Neuraisvem previsão de séries temporais**. Dissertação de Mestrado. Departamento de Engenharia Elétrica, PUC-RJ. 1993.

ZINN, WALTER E PETER C LIU. **Consumer Response to Retail Stockouts**.Journal of Business Logistics, Vol. 22, No. 1, pp. 49-71, 2001.

W.K. Wong *et al*, **Adaptive neural network model for time-series forecasting**, European Journal of Operational Research 207 (2010) 807–816.

WEIGEND, A. S., AND N. A. GERSHENFELD, eds**., Time Series Prediction: Forecasting the Future and Understanding the Past**, Reading, MA: Addison-Wesley, 1994.