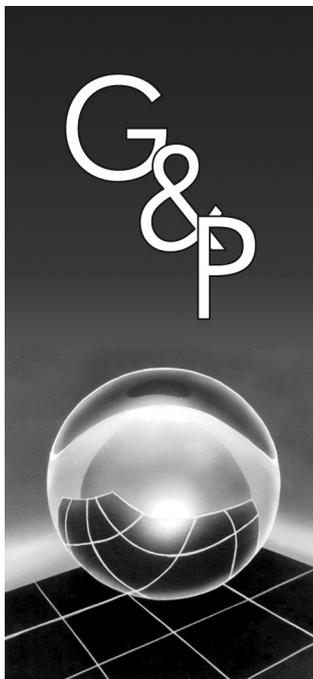


Previsão não-linear dos preços de troncos de eucalipto baseada em uma abordagem neuroevolutiva

Leandro dos Santos Coelho

Wesley Vieira da Silva

Roberto Max Protil



Resumo

As ferramentas de identificação de sistemas e previsão de séries temporais permitem a concepção de modelos matemáticos baseados em dados numéricos. O problema essencial, nestes casos, é determinar o modelo matemático apropriado. Esse artigo apresenta o projeto de uma rede neural função de base radial (RN-RBF) para a previsão de séries temporais. Na utilização da RN-RBF para previsão de sistemas não-lineares é difícil determinar um conjunto apropriado de centros e aberturas para as funções de ativação Gaussianas para obter uma boa estrutura. Neste trabalho, a configuração da RN-RBF é baseada em uma abordagem híbrida baseada em método de agrupamento de dados de Gustafson-Kessel e procedimento de otimização usando evolução diferencial. O projeto de RN-RBF é validado para previsão de um passo à frente dos preços de troncos de eucalipto para celulose e serraria para ilustrar a eficiência da abordagem híbrida proposta. Além disso, o desempenho do projeto de RN-RBF baseado nos resultados de previsão é apresentado e discutido neste artigo.

Palavras-chave: Séries temporais. Previsão. Redes neurais. Evolução diferencial. Agrupamento de dados.

1 Introdução

A identificação de sistemas não-lineares é uma área relevante de pesquisa em diversos campos do conhecimento. Um objetivo da identificação de sistemas é a obtenção de um modelo matemático de um sistema dinâmico desconhecido, para propósitos de previsão e/ou compreensão do comportamento desse sistema dinâmico.

A representação dinâmica de um sistema dinâmico é essencial para o desenvolvimento de sistemas computacionais de identificação e previsão de comportamento dinâmico de sistemas complexos. As técnicas de previsão de séries temporais clássicas – Holt, Winters, Holt-Winters, modelos Auto-Regressivos (AR), Box-Jenkins, *Generalised Auto-Regressive Conditionally Heteroscedastic* (GARCH), entre outras – são implementadas usualmente considerando a hipótese do “bom” conhecimento da dinâmica das séries temporais e a possível presença de características de tendência, sazonalidade e nível (MORETTIN; TOLOI, 1987; BOX et al., 1994; JENKINS, 1994; BERENSON; LEVINE, 1999). Estas técnicas, no entanto, muitas vezes, podem não gerar resultados satisfatórios quando aplicadas a

sistemas dinâmicos não-lineares e/ou com comportamento caótico, estes modelados como processos lineares (KANTZ; SCHREIBER, 1997).

A teoria e a prática de identificação de sistemas lineares estão bem estabelecidas na literatura (LJUNG, 1987), entretanto a utilização de modelos matemáticos lineares para representar sistemas não-lineares é limitada, pois tais representações não podem reproduzir comportamentos dinâmicos, tais como histerese, bifurcações, incertezas, acoplamentos não-lineares, que resultam em interações não-lineares. Neste caso, técnicas de previsão de séries temporais não-lineares são necessárias para modelar tais comportamentos (BILLINGS; COCA, 1999).

A identificação e/ou previsão de comportamento dinâmico de um sistema não-linear é usualmente uma tarefa difícil. Entretanto, diversas abordagens têm sido propostas na literatura recentemente para esta tarefa (ZHU, 2000; SWAIN; BILLINGS, 2001; KABOUDAN, 2001; PREVIDI, 2002; PILGRAM et al., 2002; COULIBALY; BALDWIN, 2005).

Os seres humanos são fascinados com a idéia da construção de máquinas inteligentes, por duas razões: imitar

o comportamento inteligente e providenciar soluções funcionais para tarefas complexas, que não podem ser tratadas por outras metodologias. A abordagem tradicional da produção de *software*, para a resolução de problemas particulares, não pode ser aplicada facilmente. As técnicas de previsão não-lineares nas quais as metodologias da inteligência computacional – uma subárea da inteligência artificial que abrange as redes neurais artificiais, os sistemas nebulosos (*fuzzy systems*), os algoritmos evolucionários ou evolutivos, a inteligência coletiva (*swarm intelligence*) e o raciocínio probabilístico – estão inclusas, podem ser úteis e eficientes à realização de identificação de sistemas não-lineares e/ou previsão de séries temporais com comportamento dinâmico complexo.

A complexidade inerente de muitos sistemas presentes em engenharia (multivariáveis, não-lineares e variantes no tempo) dificulta a aplicação de técnicas convencionais de identificação. Este aspecto tem motivado o desenvolvimento de técnicas de identificação avançadas, baseadas em inteligência computacional, que sejam aptas a: i) gerenciar a incerteza e o conhecimento de especialistas; ii) acomodar mudanças significativas no sistema e seu ambiente; iii) incorporar metodologias para aprendizado de informações incertas; e iv) combinar o conhecimento existente com um procedimento de aprendizado.

Neste contexto, um substancial interesse tem sido focado na utilização de redes neurais para identificação de sistemas não-lineares (GORINESKY, 1995; TAN et al., 1995; MEDEIROS; VEIGA, 2000; LEONG et al., 2002; DRIOLI; ROCCHESO, 2003; AL-AJLOUNI et al., 2004; GAO; ER, 2005) e previsão de séries temporais (TANG; FISHWICK, 1993; AZOFF, 1994; COTTREL et al., 1995; BILLINGS; HONG, 1998; NELSON et al., 1999; HIBBERT et al., 2000; VIRILI; FREISLEBEN, 2000; TSENG et al., 2002; RANK, 2003; ZEMOURI et al., 2003; HANSEN; NELSON, 2003; ZHANG, 2003; TEMIZEL; AHMAD, 2005; ZHANG; QI, 2005; TERÄSVIRTA et al., 2005; CHEN et al., 2005; GHIASSI et al., 2005). Algumas das abordagens apresentadas na literatura de previsão de séries temporais mostram redes neurais híbridas com sistemas nebulosos (CHO; WANG, 1996; NIE, 1997; AKHMETOV et al., 2001; KUO et al., 2002; CASTILLO; MELIN, 2002; KASABOV; SONG, 2002; LIN; CHEN, 2006) e algoritmos evolutivos (BILLINGS; ZHENG, 1995; YAO et al., 1996; YAO, 1999; RIVAS et al., 2004; ARMANO et al., 2005; DOGANIS et al., 2006).

As redes neurais são metodologias da inteligência computacional inspiradas por estudos do cérebro humano e o sistema nervoso. As redes neurais podem ser consideradas como sistemas dinâmicos não lineares consistindo de elementos de processamento (neurônios), que são interconectados entre si, com ponderações ajustáveis. Entre as características relevantes das redes neurais têm-se:

processamento paralelo, aprendizado, memória associativa e distribuída. Estas características são inspiradas nas redes neurais biológicas, mesmo que rudimentarmente. As redes neurais providenciam, usualmente, conhecimento quantitativo não-paramétrico e são adequadas para previsão, aprendizado e adaptação de sistemas dinâmicos não-lineares.

Em particular, as redes neurais de função de base radial (RNs-RBF) formam uma classe de redes neurais que possuem algumas vantagens em relação a outros tipos de redes neurais, incluindo melhor capacidade de aproximação, estruturas simples e algoritmos de aprendizado rápido (SARIMVEIS et al., 2002).

Neste artigo, é apresentada uma abordagem neuroevolutiva para previsão de séries temporais. Esta abordagem é composta de uma RN-RBF que utiliza o método de agrupamento de Gustafson-Kessel (GUSTAFSON; KESSEL, 1979) para o ajuste dos centros das funções Gaussianas (funções de ativação) e otimização dos centros (busca local) e das aberturas (*spreads*) das funções Gaussianas por meio de evolução diferencial (STORN; PRICE, 1995). As etapas abordadas, neste artigo, para a previsão de séries temporais pela RN-RBF incluem a determinação de uma classe de modelos candidatos (estrutura do modelo neural), um critério para verificação da aproximação entre os dados reais e o modelo matemático obtido, e o desenvolvimento de rotinas matemáticas de validação dos modelos matemáticos resultantes. O projeto da RN-RBF é validado para previsão, a curtíssimo prazo, do preço de troncos de eucalipto para celulose. O desempenho do projeto otimizado da RN-RBF baseado nos resultados de previsão é apresentado e discutido na seção de apresentação dos resultados de previsão.

O artigo está organizado da seguinte maneira. O estudo de caso de previsão dos preços de troncos de eucalipto para celulose é descrito brevemente na seção 2. Os conceitos básicos da RN-RBF e da otimização por agrupamentos e evolução diferencial são detalhados na seção 3. Os resultados obtidos da previsão de curto prazo são apresentados e analisados na seção 4 e a conclusão é mencionada na seção 5.

2 Previsão dos preços de troncos de eucalipto para celulose e serraria

O cenário atual do setor florestal brasileiro demonstra os reflexos da exploração predatória das florestas naturais causadas, principalmente, pela expansão da fronteira agropecuária, pelas atividades de mineração e pela produção de carvão vegetal a partir do cerrado, notadamente, em Minas Gerais e estados do Centro-Oeste do Brasil. A falta de uma política de monitoramento da exploração, via manejo sustentado, também contribuiu para agravar a situação. O exemplo maior deste fato é a atual demanda de

troncos da madeira serrada de eucalipto no mercado, em substituição às madeiras consideradas nobres, advindas da Mata Atlântica e Floresta Amazônica.

As florestas de eucalipto implantadas no Brasil, é sempre bom lembrar, visavam, até há pouco tempo, atender principalmente à indústria siderúrgica (carvão vegetal) e à de papel e celulose. Vale ressaltar que, existem vários problemas na utilização desta floresta na produção de madeira serrada, pois as características selecionadas das árvores visavam maximizar a produção de carvão vegetal e de papel e celulose.

Em decorrência deste fato, nota-se uma movimentação das empresas florestais, instituições de pesquisa e Universidades especializadas no ramo, na procura de meios de viabilização da utilização destas florestas, assim como na redução das perdas no processamento. Por conseguinte, planejar e implantar florestas, com características desejáveis para produção de madeira serrada, é um desafio para o futuro (MADEIRA, 2003).

A administração florestal é um processo de decisão, o qual, pelo tempo gasto na produção florestal, possui características próprias de decisão. Os administradores florestais, por exemplo, precisam realizar decisões que influenciarão a espécie, qualidade e quantidade das essências florestais dos futuros recursos florestais. Conseqüentemente, muitas das decisões tomadas no presente serão julgadas em um ambiente futuro diferente do atual (THOMPSON, 1968). O problema de incerteza na gestão de florestas é freqüentemente discutido na literatura florestal, todavia há poucos estudos sobre a utilização de métodos não convencionais na avaliação de risco no planejamento florestal (PROTIL, 1993).

A questão de previsão de preços de troncos de madeira para o mercado brasileiro foi tratada por Prottil (1994), o qual desenvolveu um sistema computacional para avaliação de riscos de investimentos florestais. O sistema possuía um módulo de previsão dos preços de troncos de eucalipto para a indústria de madeira compensada baseado no modelo de oscilação aleatória de ativos financeiros proposto por Fama (1965).

O desenvolvimento de métodos de previsão de séries temporais dos preços do eucalipto podem ser úteis ao planejamento estratégico das empresas a curto e a médio prazos, de forma a desenvolverem um cronograma sustentável de plantio e corte de árvores de forma a maximizar a produção e alcançar o uso múltiplo das florestas, que recebem cuidados especiais pelo plantio de reposição.

As séries temporais avaliadas, neste trabalho, são dos preços, em US\$ médio, de troncos de eucalipto, para: i) serraria (em m³ em pé); ii) serraria (em m³ estéreo); iii) celulose (em m³ estéreo); e iv) celulose com troncos menores que 15 cm de diâmetro (em m³ em pé).

3 Rede neural de função de base radial com otimização usando evolução diferencial

O desenvolvimento da tecnologia de processamento de informação e a inteligência computacional constituem uma solução alternativa para problemas que necessitam aspectos relativos à incorporação de características inspiradas na natureza e na inteligência humana em problemas complexos, a exemplo de previsão do comportamento de sistemas dinâmicos.

Estes sistemas de previsão, ditos “inteligentes”, tentam imitar a maneira de tomada de decisão humana e a representação de conhecimento, motivando a atenção crescente de pesquisadores de diversas áreas. Entre as vantagens potenciais oferecidas por estas técnicas, em relação aos esquemas convencionais de previsão linear, pode-se ressaltar: i) menor dependência de modelos quantitativos (vantagem dos sistemas nebulosos); ii) algoritmos estruturados de maneira simples para a tomada de decisões; iii) capacidade de aprendizagem; e iv) maior grau de autonomia.

Neste contexto, as redes neurais artificiais (RNs) são abordagens emergentes da inteligência computacional, estas motivadas por sistemas neurais biológicos, com intenção de simular, mesmo que rudimentarmente, a forma como o cérebro aprende, recorda e processa informações. As RNs são projetadas como um sistema interconectado de elementos de processamento (neurônios), cada um dos quais com um número limitado de entradas e saídas.

Entre as características que melhor explanam o potencial das redes neurais sobressaem-se a adaptabilidade, paralelismo, multidisciplinaridade e tolerância a falhas, o que proporciona uma ferramenta matemática promissora para aplicações práticas. As RNs proporcionam, usualmente, conhecimento quantitativo não-paramétrico e são adequadas para previsão de séries temporais e aprendizado de sistemas complexos, pois são aproximadores universais de funções (LO, 1998).

Uma função de base radial, ϕ , em que a saída é simétrica em torno de um centro, μ_c , tal que $\phi_c = \phi(\|x - \mu_c\|)$, sendo $\|\cdot\|$ um vetor norma. Muitas funções são utilizadas como funções de base, na camada intermediária de uma rede neural de base radial, tais como funções de aproximação cúbica, multiquadrática e *thin plate spline*.

Uma condição suficiente para uma função ser candidata à função de base radial é a de ser radialmente simétrica e a sua primeira derivada deve ser monotônica (POGGIO e GIROSI, 1990). As funções Gaussianas são as utilizadas neste trabalho. As funções Gaussianas são caracterizadas por um parâmetro de escala ou abertura (*spread*), σ , e podem ser representadas por:

$$\phi_c(x) = \phi(\|x - \mu_c\|; \sigma) \quad (1)$$

em que a função de ativação Gaussiana é dada por:

$$\phi = \exp\left(-\frac{\|x - \mu_c\|^2}{\sigma^2}\right) \quad (2)$$

Um conjunto de funções de base radial pode servir como uma base para a representação de uma vasta classe de funções que são expressas como combinações lineares de funções de base radial escolhidas, tal que:

$$\hat{y}(x, t) = \sum_{j=1}^{M_f} \omega_j \phi(\|x - \mu_c\|) \quad (3)$$

sendo M_f o número de funções (geralmente não-lineares) arbitrárias, conhecidas como funções de base radial, o índice c representa os diferentes centros, e $\|\cdot\|$ representa uma norma, geralmente Euclidiana. Uma RN-RBF, conforme apresentada na Figura 1, é um agrupamento de equações na forma de uma rede neural *feedforward* com três camadas: a(s) entrada(s), uma camada oculta (intermediária ou escondida) e uma camada de saída. Cada unidade oculta representa uma função de base radial simples com comprimento e posição do centro associados. As unidades ocultas são, algumas vezes, denominadas de centróides ou núcleos. Cada unidade de saída executa uma soma ponderada de unidades ocultas, utilizando os valores de ω_j como pesos.

A abordagem de determinar os componentes individuais, sendo dada uma superposição de Gaussianas, foi explorada nos anos 1960. Subseqüentemente, a utilização de funções de base radial para interpolação numérica e aproximação de funções foi demonstrada por Powell (1985). O projeto da RN-RBF adotado, neste artigo, é o do treinamento em dois estágios, que são os seguintes: i) determinar o número de centros e dos valores de μ_c e σ_j ; e ii) determinar os pesos das unidades da saída para os centros e comprimentos obtidos no estágio i).

No primeiro estágio, pode-se utilizar um algoritmo de aprendizado supervisionado ou não-supervisionado (por exemplo, mapas auto-organizáveis de Kohonen). Neste artigo, adota-se um método da computação evolutiva denominado evolução diferencial (STORN e PRICE,

1995) para otimização dos centros (junto ao método de agrupamento de dados de GUSTAFSON-KESSEL) e largura das funções de ativação Gaussianas. No segundo estágio, o ajuste dos pesos das unidades de saída é realizado pela técnica clássica da pseudo-inversa (GOLUB; VAN LOAN, 1983). A seguir são descritos brevemente o método de agrupamento de Gustafson-Kessel e a técnica de otimização denominada evolução diferencial.

3.1 Método de agrupamento de Gustafson-Kessel

Da seqüência de dados avaliada da série temporal, uma matriz de regressão X de dimensão N e um vetor de saída y são construídos, em que:

$$X = [x_1, \dots, x_N]^T, y = [y_1, \dots, y_N]^T \quad (4)$$

Neste caso, $N \gg n$, sendo N o número de amostras usadas para identificação e n o número de entradas do sistema nebuloso.

Neste artigo, a análise de *cluster* (grupo) é a metodologia empregada na procura de *clusters* correspondentes às classes presentes em dados; um método de agrupamento divide todas as amostras em pequenos subgrupos e os classifica de acordo com a similaridade entre eles. Uma análise de *cluster* extrai a informação encapsulada nos dados pela extrapolação da pertinência às classes. Existem diversas técnicas de agrupamento entre os métodos não-supervisionados, tais como agrupamento nebuloso adaptativa, k -médias, decomposição de mistura Gaussiana e c -médias (YEN; LANGARI, 1999).

Um agrupamento nebuloso é aplicado para partição dos dados de treinamento em regiões características em que o comportamento do sistema é aproximado por modelos lineares. No presente estudo, o conjunto de dados a ser agrupado é formado pela combinação X e y

$$Z = [X, y]^T \quad (5)$$

Dado o conjunto de dados de treinamento Z e o número de grupos (*clusters*) K , o algoritmo de agrupamento de Gustafson-Kessel (GK) é aplicado para calcular a matriz de partição U . Os conjuntos nebulosos do antecedente das regras são obtidos da matriz de partição U , em que o i -ésimo elemento de $\mu_{ik} \in [0, 1]$ é o grau de pertinência dos dados z_k ao *cluster* i .

O algoritmo de GK é regido pelas seguintes etapas. Dado Z , escolhe-se $1 < K < N$, $m > 1$ e $\varepsilon > 0$. Iniciar $U^{(0)}$ de forma aleatória com distribuição uniforme.

Repetir para $l = 1, 2, \dots$ (sendo l a iteração atual)

Etapa 1: Calcular a média do *cluster*;

$$v_i^{(l)} = \frac{\sum_{k=1}^N [\mu_{ik}^{(l-1)}]^m z_k}{\sum_{k=1}^N [\mu_{ik}^{(l-1)}]^m} \quad i = 1, 2, \dots, K. \quad (6)$$

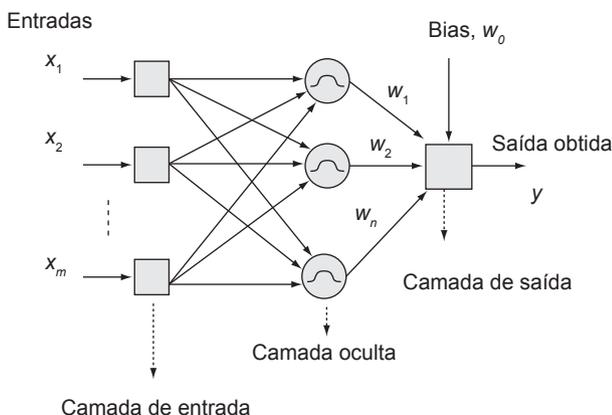


Figura 1. Uma RN-RBF com uma camada intermediária.

Etapa 2: Calcular as matrizes de covariâncias do *cluster* i ;

$$F_i = \frac{\sum_{k=1}^N [\mu_{ik}^{(l-1)}]^m [z_k - v_i^{(l)}][z_k - v_i^{(l)}]^T}{\sum_{k=1}^N [\mu_{ik}^{(l-1)}]^m}, \quad i = 1, 2, \dots, K. \quad (7)$$

O determinante $|F_i|$ da matriz F_i resulta no volume do *cluster*. Por F_i ser uma matriz simétrica e positiva definida, ela pode ser composta tal que $F_i = Q_i \Lambda_i Q_i^T$, em que Q_i é ortonormal e Λ_i é uma matriz diagonal com elementos diferentes de zero $\lambda_{i1}, \dots, \lambda_{im}$. Neste caso, os autovalores e autovetores da matriz de covariância do *cluster* geram informações, respectivamente, sobre a forma e orientação do *cluster*. Detalhes sobre o cálculo das matrizes de covariância são apresentados em Krishnapuram e Kim (1999) e Kaymak e Setnes (2002).

Etapa 3: Calcular as distâncias;

$$D_{ik}^2 = [z_k - v_i^{(l)}]^T \left\{ [\det(F_i)]^{1/(n+1)} F_i^{-1} \right\} [z_k - v_i^{(l)}] \quad (8)$$

$$i = 1, 2, \dots, K, \quad k = 1, 2, \dots, N.$$

Nesta etapa é utilizada a distância de Mahalanobis dada por F_i^{-1} , em vez da distância Euclidiana (esta usada no método c-médias) como medida de distância.

Etapa 4: Atualizar a matriz de partição,

Se $D_{ik} > 0$ para $1 \leq i \leq K, 1 \leq k \leq N$

$$\mu_{ik}^{(l)} = \frac{1}{\sum_{j=1}^K (D_{ik}/D_{jk})^{2/(m-1)}} \quad (9)$$

senão

$$\mu_{ik}^{(l)} = 0 \text{ se } D_{ik} > 0, \text{ e } \mu_{ik}^{(l)} \in [0, 1] \text{ com } \sum_{i=1}^K \mu_{ik}^{(l)} = 1 \quad (10)$$

até $\|U^{(l)} - U^{(l-1)}\| < \epsilon$. Adotou-se neste trabalho $\epsilon = 10^{-5}$.

No caso da Equação 9, os valores de pertinência ao *cluster* são obtidos da distância dos dados de entrada para os centros dos *clusters* modificados por um fator de ponderação que depende da tal distância. De forma a prevenir valores de pertinência menores que zero, é empregada a restrição imposta pela Equação 9.

Os centros das funções Gaussianas da RN-RBF são obtidos da matriz de partição U , em que o i -ésimo elemento de $\mu_{ik} \in [0, 1]$ é o centro da função de ativação dos dados z_k no *cluster* i .

3.2 Evolução diferencial

Os algoritmos evolutivos ou evolucionários (AEs) são ferramentas computacionais promissoras para busca, otimização, aprendizado de máquina e para resolução de problemas de projeto. Estes algoritmos utilizam a evolução simulada buscando soluções para problemas complexos (BÄCK et al., 1997). Os AEs são baseados

em uma população de indivíduos, em que cada um representa um ponto de busca no espaço de soluções potenciais para um dado problema. Os AEs possuem alguns procedimentos de seleção baseados na aptidão dos indivíduos (vetores solução), operadores de cruzamento e mutação.

Os AEs incluem algoritmos genéticos, programação evolucionária, estratégias evolutivas, programação genética, entre outras variantes (BÄCK et al., 1997). A ED é um AE desenvolvido por Storn e Price (1995) e visa à busca por melhores resultados com uma abordagem um pouco diferente da utilizada nos algoritmos genéticos, mas com similaridades de implementação com as estratégias evolutivas. Na ED a mutação não é realizada por uma função densidade de probabilidade. Algumas das potencialidades da ED são: i) a rapidez de convergência da otimização; e ii) a facilidade de implementação.

Neste artigo enfoca-se a ED, proposta originalmente por Storn e Price (1995), que apesar de apresentar conceitos simples é de fácil implementação, robusta e eficiente para a minimização de funções não-lineares e não-diferenciáveis no espaço contínuo. Na ED, os parâmetros da função a ser otimizada são codificados com variáveis representadas usando ponto flutuante na população e são realizadas mutações simples com uma operação aritmética simples. Storn (1997) relatou resultados impressionantes que mostram que a ED supera outros AEs (*simulated annealing* adaptativo, Nelder e Mead com *annealing*, algoritmo genético, estratégia evolutiva e equações diferenciáveis estocásticas) para abordagens de minimização em relação ao número de avaliações necessárias localizando o mínimo global de diversas funções teste consolidadas na literatura.

A escolha da ED para otimização da RN-RBF foi baseada nas seguintes características (CHENG; HWANG, 2001):

- é um algoritmo de busca estocástica que apresenta menor tendência de se concentrar em mínimos locais, pois a busca pelo ótimo global é feita pela manipulação de uma população de soluções ou, em outras palavras, por uma busca simultânea em diferentes áreas do espaço de soluções;
- é eficiente para problemas de otimização de funções objetivo que não requerem informações relativas às derivadas;
- permite que os parâmetros de entrada e saída sejam representados como ponto flutuante sem nenhum esforço computacional adicional (como é o caso da representação binária dos algoritmos genéticos); e
- não necessita manter um tamanho grande de população.

Existem duas variantes da ED que têm sido reportadas na literatura, a ED/*rand/1/bin* e a ED/*best/2/bin* (Storn, 1997). As diferentes variantes são classificadas usando a seguinte notação: ED/ $\alpha/\beta/\delta$, em que α indica o método de

seleção dos indivíduos genitores que formarão o vetor base que passa pela operação de mutação, o β indica o número de vetores de diferenças utilizados para perturbar os indivíduos do vetor base, e δ indica o mecanismo de cruzamento utilizado para geração da população de indivíduos descendentes. A sigla *bin* indica que a operação de cruzamento é controlada por uma série de experimentos binomiais.

A implementação de ED adotada neste artigo, é a $ED/rand/1/bin$. Neste caso, cada variável de um indivíduo (vetor) da população é representada por um valor real (ponto flutuante) e o seu procedimento de otimização é regido pelas seguintes etapas:

i) **gerar a população inicial:** gerar de forma aleatória com distribuição uniforme uma população inicial (geração $G=0$) de soluções factíveis, em que é garantido por regras de “reparo” que os valores atribuídos às variáveis estão dentro das fronteiras delimitadas pelo projetista;

$$v_i^j(G=0) = \lim_{\inf}(v_i^j) + rand_i[0,1] * \{\lim_{\sup}(v_i^j) - \lim_{\inf}(v_i^j)\} \quad (11)$$

sendo $j = 1, \dots, M$ indivíduos, M é o tamanho da população ($M > 4$), n é a dimensão da solução, $\lim_{\inf}(v_i^j)$ e $\lim_{\sup}(v_i^j)$ são os limites inferior e superior de valores admissíveis para a variável x_j , respectivamente, e $rand_i[0,1]$ gera um número aleatório, com distribuição uniforme, no intervalo entre 0 e 1;

ii) **avaliar a população:** para cada indivíduo é avaliada a função de aptidão (a ser maximizada), F^j , $j = 1, \dots, M$. Neste artigo, adotou-se uma função de aptidão dada pela média harmônica do índice de correlação múltipla, R^2 , obtidos nas fases de estimação e validação, tal que:

$$F^j(G) = \frac{2R_{est}^2 R_{val}^2}{R_{est}^2 + R_{val}^2} \quad (12)$$

em que R_{est}^2 e R_{val}^2 são os índices de correlação múltipla das fases de estimação e validação da previsão das séries temporais dos preços de troncos de eucalipto tanto para serraria quanto celulose. Neste caso, o coeficiente de correlação múltipla, R^2 , entre a saída real, $y(t)$, e saída estimada (prevista), $\hat{y}(t)$ é dado por:

$$R^2 = 1 - \frac{\sum_{k=1}^N [y(t) - \hat{y}(t)]^2}{\sum_{k=1}^N [y(t) - \bar{y}]^2} \quad (13)$$

em que \bar{y} é a média das N amostras medidas. Quando o valor de R^2 é igual a 1,0 indica uma exata adequação do modelo para os dados medidos do histórico da série temporal;

iii) **selecionar os indivíduos:** um indivíduo $v_i^j(G)$, da i -ésima unidade geradora na G -ésima geração, é selecionado de forma aleatória, para ser substituído e outros três diferentes indivíduos r_1, r_2, r_3 (valores inteiros do indi-

víduo i da população) são selecionados aleatoriamente como genitores;

iv) **operação de mutação:** um destes três indivíduos é selecionado como genitor principal (por exemplo, r_1) e com alguma probabilidade, cada variável do genitor principal é modificada. Neste caso, pelo menos uma variável deve ser alterada. A modificação é realizada adicionando ao valor atual da variável um fator de mutação, f , regida pela diferença entre dois valores desta variável nos outros dois genitores, ou seja:

$$v_i(G+1) = v_{i,r_1}(G) + f \cdot [v_{i,r_2}(G) - v_{i,r_3}(G)] \quad (14)$$

Em outras palavras, o vetor denominado genitor principal é modificado baseado no vetor de variáveis de dois outros genitores;

v) **operação de cruzamento:** seguindo a operação de mutação, o cruzamento é aplicado na população. Para cada vetor mutado, $v_i(G+1)$, um índice $rnbr(i) \in \{1, 2, \dots, n\}$ é aleatoriamente escolhido, e um *vetor tentativa*, $vt_i(G+1)$ é gerado tal que:

$$vt_{i,q}(G+1) = \begin{cases} v_{i,q}(G+1), & \text{se } randb(q) \leq CR \text{ ou } q = rnbr(i) \\ v_{i,q}(G), & \text{se } randb(q) > CR \text{ ou } q \neq rnbr(i) \end{cases} \quad (15)$$

sendo CR a taxa de cruzamento; $i = 1, \dots, n$ um índice de indivíduo da população e $q = 1, \dots, n$ é a posição dentro do indivíduo n -dimensional. Para decidir se o vetor $vt_i(G+1)$ deve ser um membro da população na próxima geração, este é comparado com o vetor $v_i(G+1)$ correspondente. Se o vetor $vt_i(G+1)$ apresenta uma função de aptidão melhor que a escolhida à substituição, ele a substitui; caso contrário, o vetor escolhido $v_i(G+1)$ para ser substituído é mantido na população; e

vi) **critério de parada:** incrementar uma unidade ao valor de G ; após ir para o passo ii) até que um critério de parada seja atendido. Neste artigo o critério de parada é o número máximo de gerações, G_{max} .

Na ED, $f > 0$ é um parâmetro real, que controla a amplificação da diferença entre dois indivíduos e para evitar a estagnação é usualmente projetado no intervalo $[0, 1; 2]$; $randb(q)$ é a q -ésima avaliação de um número aleatório com distribuição uniforme no intervalo $[0, 1]$; CR é a taxa de cruzamento (*crossover rate*) no intervalo $[0, 1]$. O desempenho da ED depende de três variáveis: o tamanho de população M , o fator de mutação f e a taxa de cruzamento CR .

4 Resultados de previsão obtidos para o estudo de caso

Os dados históricos de preços de troncos de eucalipto foram divididos em três conjuntos de dados. Os dados

para serem usados na fase de estimação (treinamento) da RN-RBF e os dados a serem usados nas fases de validação e teste, para verificação da capacidade de generalização do modelo de previsão da RN-RBF, conforme apresentado na Tabela 1.

O procedimento de previsão de séries temporais por uma RN-RBF foi configurado para utilizar um modelo auto-regressivo (usando saídas reais passadas) com estrutura série-paralelo para previsão um passo à frente, os detalhes desta estrutura são apresentados em Van Gorp (2000).

A RN-RBF projetada para previsão de preços de madeira de eucalipto consiste de 3 neurônios de entrada compostos dos vetores $\{y(t-3), y(t-2), y(t-1)\}$ e 1 neurônio na camada de saída composto da saída estimada $\hat{y}(t+1)$. Esta configuração de projeto da RN-RBF foi adotada depois da realização de vários testes. Na Figura 2 são apresentados os detalhes do projeto e otimização da RN-RBF.

Neste estudo de caso foram utilizadas as primeiras 42 amostras (da amostra 1 à 42) para a fase de estimação do modelo de previsão, com otimização com método de agrupamento de Gustafson-Kessel (GK) e ED. As 30 amostras seguintes (da amostra 43 à 70) para a fase de validação do modelo de previsão de preços e outras 12 amostras para a fase de testes do modelo. Foram testados diferentes números de funções Gaussianas na

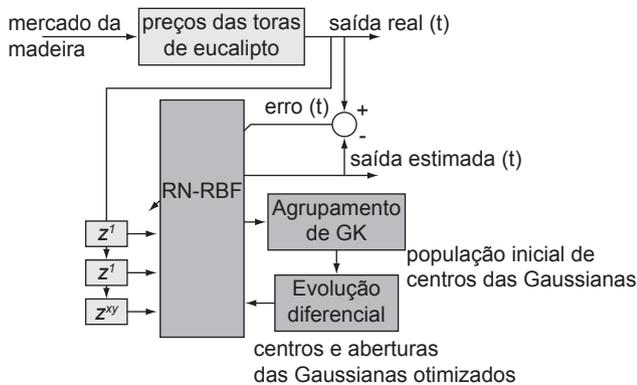


Figura 2. RN-RBF usando estrutura série-paralelo para previsão um passo à frente.

camada oculta da RN-RBF e os melhores resultados dos experimentos são resumidos nas Tabelas de 1 a 3. O tempo computacional (em segundos) apresentado nas Tabelas de 1 a 3 foi obtido usando um computador com processador Pentium IV 3.2 GHz com 2 GB de RAM e ambiente computacional Matlab 6.5.

Para seleção de uma configuração da ED, foram realizados diversos experimentos com diferentes valores de M, f, CR , espaço de busca e critério de parada. E decidiu-se por determinada configuração de ED com tamanho de população pequena com valores de f e CR recomendados por Storn (1997). Neste contexto, nas simulações de otimização por meio da ED foram utilizados os seguintes parâmetros:

- tamanho de população, M : 20 indivíduos;
- taxas de mutação e cruzamento, respectivamente, com valores $f = 0,4$ e $CR = 0,8$;
- espaço de busca dos parâmetros (centros e larguras das funções de ativação Gaussianas) entre -15 e 15 para os centros das funções Gaussianas e entre $0,0001$ e 15 para as aberturas das Gaussianas;
- experimentos: 30 (com diferentes sementes de números aleatórios para gerar a população); e
- critério de parada da otimização usando ED: 200 gerações.

Nota-se, dos resultados apresentados nas Tabelas de 1 a 4, que a RN-RBF apresentou resultados para a função custo no intervalo $0,6913 \leq F^i(200) \leq 0,9286$ com valores pequenos de desvio padrão mostrando a robustez do GK com ED para otimizar o projeto da RN-RBF. Nota-se também que os valores de R^2_{est} e R^2_{val} foram promissores para os casos das Tabelas de 1 a 3. Entretanto, uma análise detalhada, em trabalho futuros, se faz necessária para o caso da Tabela 4. Neste caso, a previsão dos preços de troncos de eucalipto para a celulose, em que a RN-RBF na fase de estimação obteve valores pouco precisos, sendo $0,5744 \leq R^2_{est} \leq 0,7487$. Observa-se também que os resultados da fase de estimação foram inferiores aos da fase de validação.

Foram testados projetos de RN-RBF com 2 a 10 funções de ativação Gaussianas, na camada oculta. Os melhores resultados foram obtidos com a RN-RBF mais complexa, ou seja, com 10 funções Gaussianas para três casos (ver

Tabela 1. Dados usados nas previsões um passo à frente da série temporal.

Série temporal	Dados da fase de estimação (mês/ano)	Dados da fase de validação (mês/ano)	Dados da fase de teste (mês/ano)
Serraria (em m ³ em pé)	08/1997 a 12/2000	01/2001 a 12/2003	01/2004 a 12/2004
Serraria (em m ³ estéreo)	08/1997 a 12/2000	01/2001 a 12/2003	01/2004 a 12/2004
Celulose (em m ³ estéreo)	08/1997 a 12/2000	01/2001 a 12/2003	01/2004 a 12/2004
Celulose com troncos menores que 15 cm de diâmetro (em m ³ em pé)	01/2000 a 06/2002	07/2002 a 12/2003	01/2004 a 12/2004

Tabela 2. Resultados da previsão de preços usando RN-RBF para a série temporal de troncos de eucalipto para serraria (em m³ em pé).

Gaussianas na camada oculta da RN-RBF		R^2_{est} (fase de estimação do modelo de previsão)				
	Máximo	Média	Mínimo	Desvio padrão		
2	0,8786	0,8784	0,8771	4,3757.10 ⁻⁴		
3	0,8864	0,8824	0,8779	3,6031.10 ⁻³		
4	0,8866	0,8861	0,8840	7,8710.10 ⁻⁴		
5	0,8874	0,8866	0,8857	4,4933.10 ⁻⁴		
6	0,8882	0,8870	0,8861	5,3071.10 ⁻⁴		
7	0,8890	0,8876	0,8865	9,2401.10 ⁻⁴		
8	0,8885	0,8875	0,8866	6,7177.10 ⁻⁴		
9	0,8899	0,8885	0,8863	9,7170.10 ⁻⁴		
10	0,8898	0,8891	0,8820	5,8489.10 ⁻⁴		
Gaussianas na camada oculta da RN-RBF		R^2_{val} (fase de validação do modelo de previsão)				
	Máximo	Média	Mínimo	Desvio padrão		
2	0,9456	0,9450	0,9419	1,1233.10 ⁻³		
3	0,9573	0,9516	0,9456	5,4515.10 ⁻³		
4	0,9614	0,9596	0,9564	1,4176.10 ⁻³		
5	0,9675	0,9627	0,9603	2,1836.10 ⁻³		
6	0,9699	0,9676	0,9631	2,2312.10 ⁻³		
7	0,9724	0,9676	0,9649	2,1293.10 ⁻³		
8	0,9703	0,9689	0,9673	1,0311.10 ⁻³		
9	0,9727	0,9699	0,9671	1,5520.10 ⁻³		
10	0,9711	0,9702	0,9691	6,3344.10 ⁻⁴		
Gaussianas na camada oculta da RN-RBF		R^2_{val} (fase de testes do modelo de previsão)				
	Máximo	Média	Mínimo	Desvio padrão		
2	0,6385	0,6305	0,5871	1,5957.10 ⁻²		
3	0,6771	0,6186	0,5524	3,7913.10 ⁻²		
4	0,6115	0,6108	0,6011	1,1200.10 ⁻²		
5	0,6989	0,6296	0,5818	4,1169.10 ⁻²		
6	0,7064	0,6789	0,6495	2,0141.10 ⁻²		
7	0,7083	0,6920	0,6742	1,1529.10 ⁻²		
8	0,7107	0,6990	0,6808	9,3794.10 ⁻³		
9	0,7079	0,6865	0,6693	1,2703.10 ⁻²		
10	0,7021	0,6893	0,6467	1,6231.10 ⁻²		
Gaussianas na camada oculta da RN-RBF		Melhor função de aptidão: $F^j(G) = \frac{2R^2_{est} R^2_{val}}{R^2_{est} + R^2_{val}}$				
	Máximo	Média	Mínimo	Desvio padrão	Tempo*	
2	0,9109	0,9105	0,9084	7,5254.10 ⁻⁴	11,2	
3	0,9203	0,9157	0,9107	4,3953.10 ⁻³	11,7	
4	0,9225	0,9214	0,9188	1,0405.10 ⁻³	12,1	
5	0,9255	0,9231	0,9220	1,0656.10 ⁻³	12,8	
6	0,9265	0,9256	0,9236	9,7440.10 ⁻⁴	13,3	
7	0,9279	0,9259	0,9242	1,1041.10 ⁻³	13,6	
8	0,9274	0,9264	0,9257	4,4831.10 ⁻⁴	14,1	
9	0,9286	0,9274	0,9262	6,7789.10 ⁻⁴	14,8	
10	0,9283	0,9279	0,9272	3,9462.10 ⁻⁴	15,3	

Tabelas de 2 a 4), e com 9 funções Gaussianas para um caso (ver Tabela 1). Em outras palavras, a complexidade da otimização por ED cresce com o aumento do número de funções de ativação utilizadas no modelo de previsão,

pois a ED necessita, neste caso, otimizar mais centros e larguras das funções de ativação Gaussianas. Nas Figuras de 3 a 6, são apresentados os melhores resultados obtidos com as RNs-RBF testadas.

Tabela 3. Resultados da previsão de preços usando RN-RBF para a série temporal de troncos de eucalipto para serraria (em m³ estéreo).

Gaussianas na camada oculta da RN-RBF	R^2_{est} (fase de estimação do modelo de previsão)				
	Máximo	Média	Mínimo	Desvio padrão	
2	0,7644	0,7610	0,7585	1,8079.10 ⁻³	
3	0,7742	0,7694	0,7630	3,5494.10 ⁻³	
4	0,7753	0,7727	0,7706	1,7678.10 ⁻³	
5	0,7785	0,7756	0,7714	2,5887.10 ⁻³	
6	0,7801	0,7740	0,7694	3,7115.10 ⁻³	
7	0,7776	0,7732	0,7697	2,6367.10 ⁻³	
8	0,7778	0,7724	0,7698	2,6115.10 ⁻³	
9	0,7811	0,7772	0,7724	2,8313.10 ⁻³	
10	0,7857	0,7808	0,7766	3,0157.10 ⁻³	
Gaussianas na camada oculta da RN-RBF	R^2_{val} (fase de validação do modelo de previsão)				
	Máximo	Média	Mínimo	Desvio padrão	
2	0,8852	0,8840	0,8824	1,0257.10 ⁻³	
3	0,9067	0,8999	0,8838	7,0325.10 ⁻³	
4	0,9084	0,9054	0,8917	5,0747.10 ⁻³	
5	0,9100	0,9066	0,9047	1,6884.10 ⁻³	
6	0,9180	0,9104	0,9033	5,1280.10 ⁻³	
7	0,9194	0,9144	0,9080	3,6593.10 ⁻³	
8	0,9202	0,9172	0,9131	2,2795.10 ⁻³	
9	0,9215	0,9175	0,9133	2,7186.10 ⁻³	
10	0,9265	0,9179	0,9130	3,8437.10 ⁻³	
Gaussianas na camada oculta da RN-RBF	R^2_{val} (fase de testes do modelo de previsão)				
	Máximo	Média	Mínimo	Desvio padrão	
2	0,4404	0,4198	0,4000	1,1923.10 ⁻²	
3	0,4282	0,3992	0,3689	2,0341.10 ⁻²	
4	0,4180	0,3974	0,3417	2,4532.10 ⁻²	
5	0,4360	0,4045	0,3806	1,7692.10 ⁻²	
6	0,5316	0,4417	0,3812	5,1445.10 ⁻²	
7	0,5212	0,4819	0,4317	2,9916.10 ⁻²	
8	0,5339	0,5033	0,4800	1,6708.10 ⁻²	
9	0,5590	0,5400	0,5032	1,5827.10 ⁻²	
10	0,5960	0,5426	0,5031	2,6560.10 ⁻²	
Gaussianas na camada oculta da RN-RBF	Melhor função de aptidão: $F^j(G) = \frac{2R^2_{est} R^2_{val}}{R^2_{est} + R^2_{val}}$				
	Máximo	Média	Mínimo	Desvio padrão	Tempo*
2	0,8192	0,8179	0,8159	9,2102.10 ⁻⁴	11,2
3	0,8335	0,8296	0,8192	4,2528.10 ⁻³	11,7
4	0,8361	0,8338	0,8276	2,3295.10 ⁻³	12,5
5	0,8376	0,8360	0,8342	1,0690.10 ⁻³	12,6
6	0,8383	0,8367	0,8344	1,1508.10 ⁻³	13,2
7	0,8389	0,8379	0,8365	7,1113.10 ⁻⁴	13,7
8	0,8407	0,8386	0,8376	9,2714.10 ⁻⁴	14,4
9	0,8428	0,8415	0,8403	8,9267.10 ⁻⁴	14,9
10	0,8466	0,8438	0,8421	1,5766.10 ⁻³	15,5

Na Tabela 5 é apresentado um comparativo do método proposto usando GK-ED, com o treinamento da RN-RBF usando: i) apenas GK; ii) apenas ED para seleção dos centros e larguras das funções Gaussianas (é apresentado na Tabela 5 apenas o máximo valor da função

de aptidão usando critério de parada de 200 gerações); e iii) o método clássico de agrupamento *k*-médias (KM), este descrito em Yen e Langari (1999).

Nota-se, pelos resultados da Tabela 5, que o projeto de RN-RBF utilizando GK-ED apresenta os melhores

Tabela 4. Resultados da previsão de preços usando RN-RBF para a série temporal de troncos de eucalipto para celulose (em m³ estéreo).

Gaussianas na camada oculta da RN-RBF	R^2_{est} (fase de estimação do modelo de previsão)				
	Máximo	Média	Mínimo	Desvio padrão	
2	0,8253	0,8241	0,8227	7,3416.10 ⁻⁴	
3	0,8439	0,8338	0,8242	7,3961.10 ⁻³	
4	0,8441	0,8428	0,8399	1,5916.10 ⁻³	
5	0,8455	0,8442	0,8426	9,7537.10 ⁻⁴	
6	0,8456	0,8446	0,8428	8,1233.10 ⁻⁴	
7	0,8460	0,8452	0,8443	6,2053.10 ⁻⁴	
8	0,8459	0,8445	0,8432	8,8034.10 ⁻⁴	
9	0,8490	0,8458	0,8434	1,6518.10 ⁻³	
10	0,8503	0,8479	0,8441	2,3561.10 ⁻³	
Gaussianas na camada oculta da RN-RBF	R^2_{val} (fase de validação do modelo de previsão)				
	Máximo	Média	Mínimo	Desvio padrão	
2	0,8538	0,8531	0,8518	5,9855.10 ⁻⁴	
3	0,8972	0,8782	0,8532	1,7274.10 ⁻²	
4	0,8983	0,8940	0,8899	3,2533.10 ⁻³	
5	0,9027	0,8951	0,8885	4,5802.10 ⁻³	
6	0,9034	0,8991	0,8964	2,5012.10 ⁻³	
7	0,9048	0,9021	0,8989	2,0967.10 ⁻³	
8	0,9074	0,9038	0,9006	2,5534.10 ⁻³	
9	0,9091	0,9056	0,9027	2,4194.10 ⁻³	
10	0,9165	0,9096	0,9039	3,6889.10 ⁻³	
Gaussianas na camada oculta da RN-RBF	R^2_{val} (fase de testes do modelo de previsão)				
	Máximo	Média	Mínimo	Desvio padrão	
2	0,3555	0,3459	0,3377	6,2812.10 ⁻³	
3	0,4681	0,4082	0,3527	3,6573.10 ⁻²	
4	0,4552	0,4325	0,4081	1,9427.10 ⁻²	
5	0,4616	0,4337	0,3975	2,0669.10 ⁻²	
6	0,4645	0,4454	0,4310	9,0616.10 ⁻²	
7	0,5184	0,4654	0,4414	2,2888.10 ⁻²	
8	0,5095	0,4787	0,4632	1,5697.10 ⁻²	
9	0,5121	0,4963	0,4751	1,2810.10 ⁻²	
10	0,5914	0,5298	0,4786	3,4842.10 ⁻²	
Gaussianas na camada oculta da RN-RBF	Melhor função de aptidão: $F^j(G) = \frac{2R^2_{est} R^2_{val}}{R^2_{est} + R^2_{val}}$				
	Máximo	Média	Mínimo	Desvio padrão	Tempo*
2	0,8385	0,8384	0,8376	2,7957.10 ⁻⁴	11,2
3	0,8672	0,8553	0,8386	1,1751.10 ⁻²	12,0
4	0,8702	0,8676	0,8656	1,5965.10 ⁻³	12,2
5	0,8730	0,8689	0,8656	2,3372.10 ⁻³	12,4
6	0,8723	0,8710	0,8699	8,7063.10 ⁻⁴	13,0
7	0,8738	0,8727	0,8714	7,9790.10 ⁻⁴	13,3
8	0,8753	0,8732	0,8720	1,0648.10 ⁻³	13,8
9	0,8767	0,8747	0,8739	87834.10 ⁻⁴	14,3
10	0,8801	0,8776	0,8756	1,6091.10 ⁻³	15,3

resultados em termos de precisão que os obtidos pelos projetos baseados em GK e ED quando empregados de forma isolada. Entretanto, o tempo computacional gasto pelo projeto da RN-RBF usando somente GK ou mesmo KM foi muito menor. Neste contexto, observa-se que a RN-RBF utilizando GK-ED foi útil para o ajuste fino dos

centros e aberturas das funções Gaussianas, o que não é uma tarefa realizada pelo GK quando usado de forma isolada. Nota-se também, pela Tabela 5, que os resultados usando método de agrupamento KM foram próximos em termos de precisão e tempo computacional que os obtidos pela RN-RBF com GK.

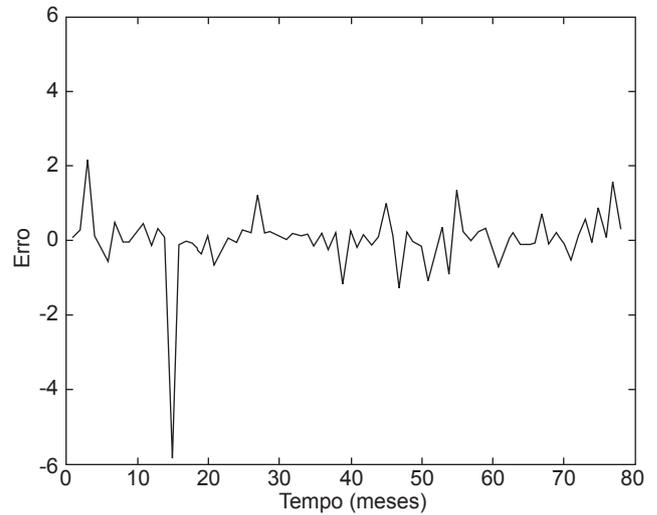
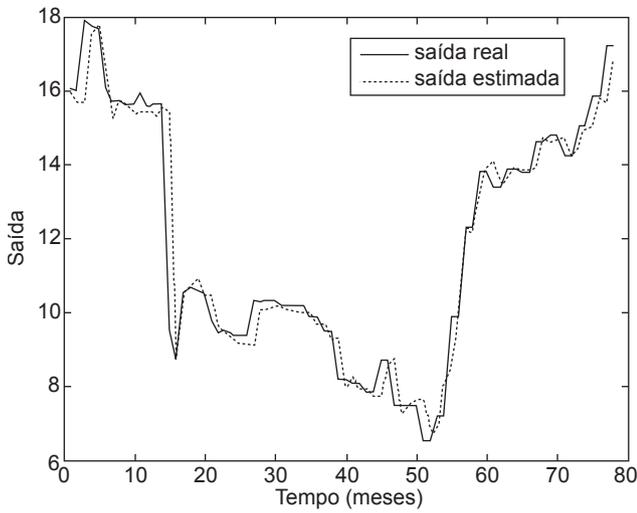


Figura 3. Melhor resultado de previsão pela RN-RBF usando 9 Gaussianas na camada oculta para previsão do preço de troncos de eucalipto para serraria (m^3 em pé).

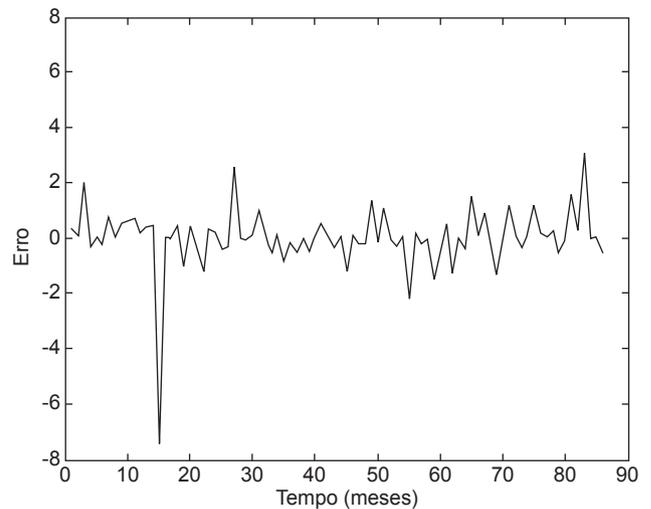
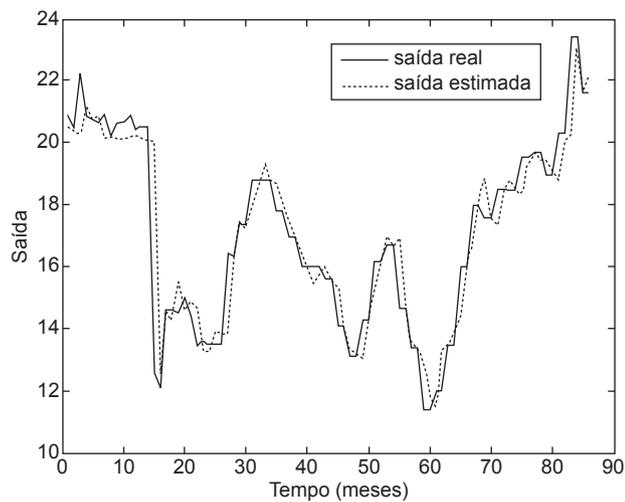


Figura 4. Melhor resultado de previsão pela RN-RBF usando 10 Gaussianas para previsão do preço de troncos de eucalipto para serraria (m^3 estéreo).

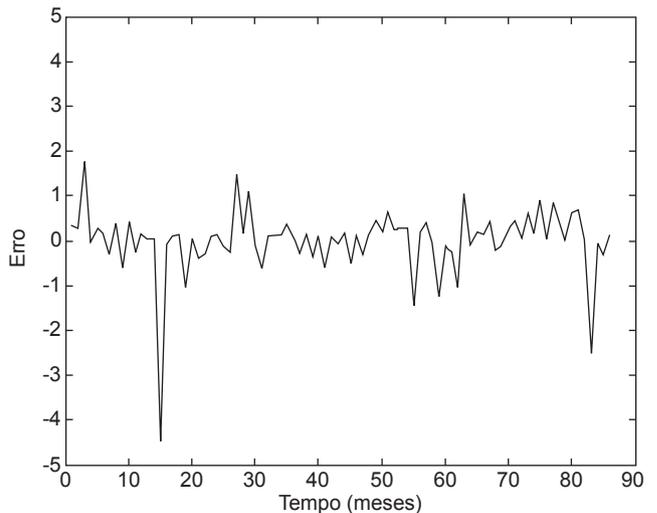
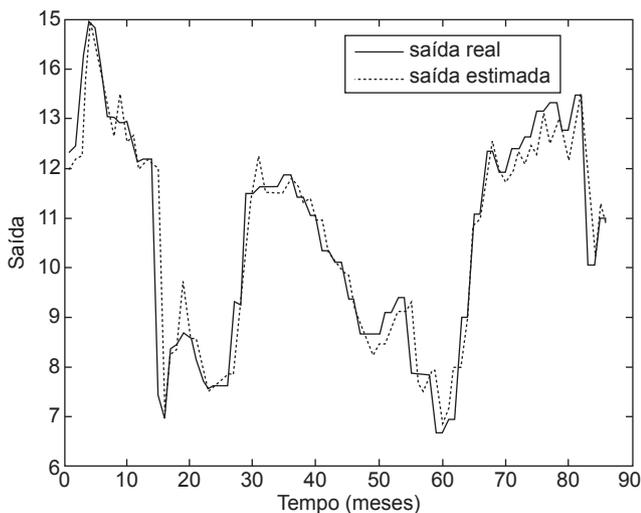


Figura 5. Melhor resultado de previsão pela RN-RBF usando 10 Gaussianas para previsão do preço de troncos de eucalipto para celulose (m^3 estéreo).

Tabela 5. Resultados da previsão de preços usando RN-RBF para a série temporal de troncos de eucalipto para celulose menores que 15 cm (m³ em pé).

Gaussianas na camada oculta da RN-RBF		R^2_{est} (fase de estimação do modelo de previsão)				
		Máximo	Média	Mínimo	Desvio padrão	
2		0,5744	0,5643	0,5447	1,1781.10 ⁻²	
3		0,6560	0,5917	0,5491	3,8644.10 ⁻²	
4		0,6578	0,6368	0,5719	2,4650.10 ⁻²	
5		0,6762	0,6387	0,6117	1,8621.10 ⁻²	
6		0,6727	0,6499	0,6242	1,5263.10 ⁻²	
7		0,6720	0,6454	0,6228	1,7660.10 ⁻²	
8		0,6736	0,6399	0,6158	2,0056.10 ⁻²	
9		0,6769	0,6355	0,6142	2,0637.10 ⁻²	
10		0,7487	0,6467	0,5958	4,4320.10 ⁻²	
Gaussianas na camada oculta da RN-RBF		R^2_{val} (fase de validação do modelo de previsão)				
		Máximo	Média	Mínimo	Desvio padrão	
2		0,8845	0,8686	0,8439	1,1965.10 ⁻²	
3		0,8904	0,8783	0,8689	5,9350.10 ⁻³	
4		0,8903	0,8817	0,8717	6,3759.10 ⁻³	
5		0,8924	0,8889	0,8819	3,5216.10 ⁻³	
6		0,8931	0,8868	0,8826	3,7716.10 ⁻³	
7		0,8980	0,8910	0,8852	4,1940.10 ⁻³	
8		0,9021	0,8921	0,8777	6,6869.10 ⁻³	
9		0,9159	0,9017	0,8925	7,6710.10 ⁻³	
10		0,9332	0,9170	0,8914	1,3839.10 ⁻²	
Gaussianas na camada oculta da RN-RBF		R^2_{val} (fase de testes do modelo de previsão)				
		Máximo	Média	Mínimo	Desvio padrão	
2		0,7782	0,7435	0,6867	2,6961.10 ⁻²	
3		0,8056	0,7693	0,7485	1,7513.10 ⁻²	
4		0,7992	0,7810	0,7570	1,5494.10 ⁻²	
5		0,8044	0,7914	0,7708	1,1682.10 ⁻²	
6		0,8014	0,7870	0,7725	1,1782.10 ⁻²	
7		0,8057	0,7926	0,7770	8,9667.10 ⁻³	
8		0,8114	0,7918	0,7611	1,3067.10 ⁻²	
9		0,8511	0,8140	0,7826	2,3420.10 ⁻²	
10		0,8951	0,8578	0,7857	3,9686.10 ⁻²	
Gaussianas na camada oculta da RN-RBF		Melhor função de aptidão: $F^j(G) = \frac{2R^2_{est} R^2_{val}}{R^2_{est} + R^2_{val}}$				
		Máximo	Média	Mínimo	Desvio padrão	Tempo*
2		0,6913	0,6840	0,6735	6,2318.10 ⁻³	10,8
3		0,7519	0,7065	0,6767	2,7997.10 ⁻²	11,1
4		0,7541	0,7293	0,6920	1,7547.10 ⁻²	11,5
5		0,7655	0,7432	0,7258	1,1383.10 ⁻²	11,8
6		0,7638	0,7500	0,7348	9,1879.10 ⁻³	12,3
7		0,7653	0,7484	0,7354	1,0679.10 ⁻²	13,0
8		0,7640	0,7450	0,7310	1,1463.10 ⁻²	13,5
9		0,7741	0,7454	0,7301	1,5148.10 ⁻²	13,9
10		0,8253	0,7578	0,7228	3,1090.10 ⁻²	14,2

5 Conclusão e futura pesquisa

A RN-RBF é uma abordagem de reconhecimento de padrões útil para o projeto de modelos matemáticos não-lineares de previsão, pois as RNs-RBF são aproximadores locais universais, conforme provado por Lo (1998). O aprendizado da RN-RBF é particularmente

simples e rápido em relação à rede neural *perceptron* multicamadas, esta muito utilizada em aplicações de identificação de sistemas complexos e previsão não-linear de séries temporais.

Neste artigo, foi apresentado um projeto de RN-RBF utilizando um procedimento de otimização baseado em

Tabela 6. Comparação de resultados usando diferentes projetos de RN-RBF para a previsão da série temporal de troncos de eucalipto.

(a) série temporal de troncos de eucalipto para serraria (em m³ em pé)								
Gaussianas na camada oculta da RN-RBF					Tempo computacional*			
$F^j(G) = \frac{2R_{est}^2 R_{val}^2}{R_{est}^2 + R_{val}^2}$								
	GK-ED	GK	ED	KM	GK-ED	GK	ED	KM
7	0,9279	0,9212	0,9038	0,9210	13,6	0,3	12,8	0,2
8	0,9274	0,9246	0,9047	0,9241	14,1	0,1	14,9	0,1
9	0,9286	0,9247	0,9100	0,9242	14,8	0,1	15,9	0,1
10	0,9283	0,9245	0,9032	0,9243	15,3	0,1	16,6	0,1
(b) série temporal de troncos de eucalipto para serraria (em m³ estéreo)								
Gaussianas na camada oculta da RN-RBF					Tempo computacional*			
$F^j(G) = \frac{2R_{est}^2 R_{val}^2}{R_{est}^2 + R_{val}^2}$								
	GK-ED	GK	ED	KM	GK-ED	GK	ED	KM
7	0,8389	0,8342	0,8328	0,8340	13,7	0,3	13,9	0,2
8	0,8407	0,8341	0,8339	0,8345	14,4	0,1	14,9	0,1
9	0,8428	0,8343	0,8343	0,8340	14,9	0,1	15,2	0,1
10	0,8466	0,8355	0,8343	0,8350	15,5	0,1	16,1	0,1
(c) série temporal de troncos de eucalipto para celulose (em m³ estéreo)								
Gaussianas na camada oculta da RN-RBF					Tempo computacional*			
$F^j(G) = \frac{2R_{est}^2 R_{val}^2}{R_{est}^2 + R_{val}^2}$								
	GK-ED	GK	ED	KM	GK-ED	GK	ED	KM
7	0,8738	0,8636	0,8519	0,8612	13,3	0,1	13,9	0,1
8	0,8753	0,8696	0,8612	0,8687	13,8	0,1	14,8	0,1
9	0,8767	0,8699	0,8615	0,8698	14,3	0,1	16,2	0,1
10	0,8801	0,8717	0,8613	0,8706	15,3	0,1	16,3	0,1
(d) série temporal de troncos de eucalipto para celulose com troncos menores que 15 cm de diâmetro (em m³ em pé)								
Gaussianas na camada oculta da RN-RBF					Tempo computacional*			
$F^j(G) = \frac{2R_{est}^2 R_{val}^2}{R_{est}^2 + R_{val}^2}$								
	GK-ED	GK	ED	KM	GK-ED	GK	ED	KM
7	0,7653	0,7133	0,6992	0,6831	13,0	0,1	14,0	0,1
8	0,7640	0,6965	0,6980	0,6733	13,5	0,1	14,8	0,1
9	0,7741	0,6924	0,6921	0,6797	13,9	0,1	15,9	0,1
10	0,8253	0,6792	0,6922	0,6612	14,2	0,1	16,3	0,19

*Tempo computacional (em segundos) usando um computador com processador Pentium IV 3.2 GHz com 2 GB de RAM e *software* Matlab 6.5.

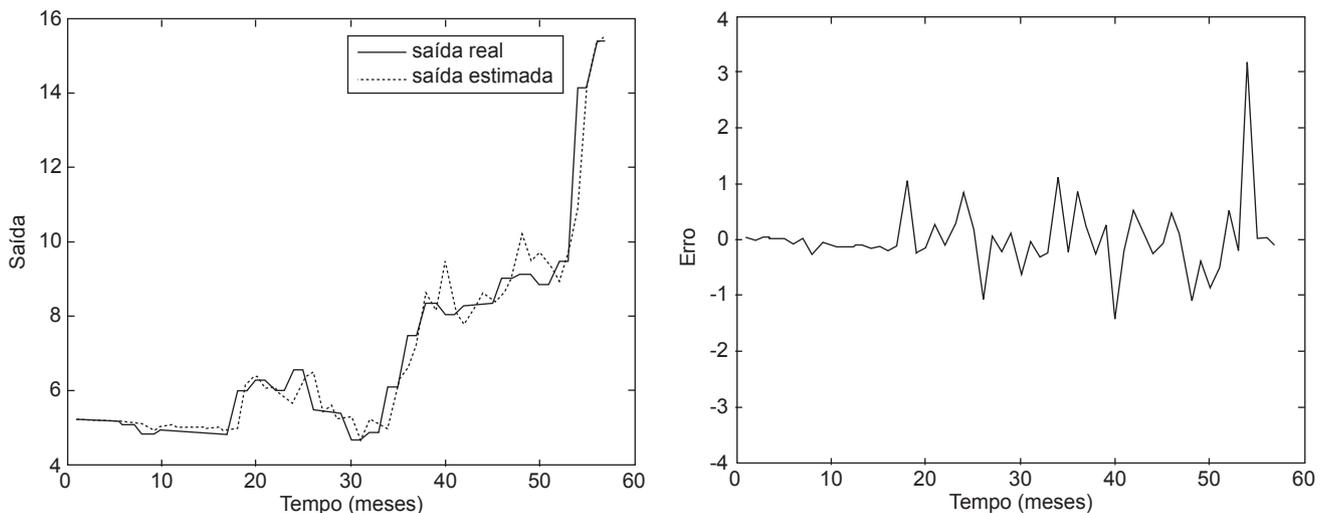


Figura 6. Melhor resultado de previsão pela RN-RBF usando 10 Gaussianas para previsão do preço de troncos de eucalipto menores que 15 cm para celulose (m³ em pé).

agrupamento de GK e ED. O estudo de caso de previsão foi da previsão a curtíssimo prazo (um passo à frente) do preço de troncos de eucalipto para celulose. O desempenho do projeto e otimização da rede neural é detalhado também no artigo.

Uma dificuldade na configuração de RNs-RBF para aplicações de previsão de séries temporais é o aspecto relativo ao domínio temporal. Se a ordem de um sistema é conhecida, a utilização do método *tapped-delay-line* permite a um sistema dinâmico ser modelado com uma RN-RBF estática. Entretanto,

quando a ordem é desconhecida, muitas aplicações podem levar o projetista a adotar um número elevado de entradas. Tal procedimento pode levar a modelos ineficientes devido ao grande número de entradas necessário.

Neste contexto, em futura pesquisa, estudos aprofundados são necessários quanto à análise da ordem (lags) de atraso na entrada (BILLINGS; VOON, 1986) da RN-RBF e também em relação à obtenção de modelos matemáticos com RNs-RBF com um melhor compromisso entre complexidade e precisão da aproximação de dinâmicas não-lineares.

Nonlinear forecasting of eucalyptus wood prices based on an evolutionary neural network approach

Abstract

Computational tools of system identification and prediction of time series allows for the conception of mathematical models based on numerical data. The key problem in these cases is to find a suitable mathematical model. This paper presents a radial basis function neural network (RBF-NN) design for forecasting time series. Using the RBF-NN for nonlinear system forecasting is quite difficult as one has to choose an appropriate set of centers and spreads for the Gaussian activation functions to achieve a good network structure. In this work, the setup of RBF-NN is based on a hybrid method based on the Gustafson-Kessel clustering method and optimization procedure by differential evolution. The RBF-NN design is validated for the one-step ahead forecasting of eucalyptus wood prices for cellulose and sawmill to illustrate the effectiveness of this hybrid approach. The performance of the RBF-NN design based on forecasting results is presented and discussed in details.

Keywords: Time series. Forecasting. Neural network. Radial basis function. Differential evolution.

Referências bibliográficas

- AKHMETOV, D. F.; DOTE, Y.; OVASKA, S. J. Fuzzy neural network with general parameter adaptation for modeling of nonlinear time-series. **IEEE Transactions on Neural Networks**, v. 12, n. 1, p. 148-152, 2001.
- AL-AJLOUNI, A. F.; SCHILLING, R. J.; HARRIS, S. L., Identification of nonlinear discrete-time systems using raised-cosine radial basis function networks. **International Journal of Systems Science**, v. 35, n. 4, p. 211-221, 2004.
- ARMANO, G.; MARCHESI, M.; MURRU, A. A hybrid genetic-neural architecture for stock indexes forecasting. **Information Sciences**, v. 170, n. 1, p. 3-33, 2005.
- AZOFF, E. M. **Neural network time series forecasting of financial markets**. Chichester; New York: Wiley & Sons, 1994.
- BÄCK, T.; FOGEL, D. B.; MICHALEWICZ, Z. (eds.). **Handbook of evolutionary computation**. Bristol, Philadelphia: Institute of Physics Publishing, NY and Oxford: Oxford University Press, 1997.
- BERENSON, M. L.; LEVINE, D. M. **Basic business statistics: concepts and applications**. 7. ed., Upper Saddle River: Prentice Hall, USA, 1999.
- BILLINGS, S. A.; COCA, D. Discrete wavelet models for identification and qualitative analysis of chaotic systems. **International Journal of Bifurcation and Chaos**, v. 9, n. 7, p. 1263-1284, 1999.
- BILLINGS, S. A.; HONG, X. Dual-orthogonal radial basis function networks for nonlinear time series prediction. **Neural Networks**, v. 11, n. 3, p. 479-493, 1998.
- BILLINGS, S. A.; ZHENG, G. L. Radial basis function network configuration using genetic algorithms. **Neural Networks**, v. 8, n. 6, p. 877-890, 1995.
- BILLINGS, S. A.; VOON, W. S. F. Correlation based model validity tests for non-linear models. **International Journal of Control**, v. 44, n. 1, p. 235-244, 1986.

- BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Time series analysis: forecasting and control**. 3. ed., Englewood Cliffs: Prentice Hall, USA, 1994.
- CASTILLO, O.; MELIN, P. Hybrid intelligent systems for time series prediction using neural networks, fuzzy logic, and fractal theory. **IEEE Transactions on Neural Networks**, v. 13, n. 6, p. 1395-1408, 2002.
- CHEN, Y. et al. Time-series forecasting using flexible neural tree model. **Information Sciences**, v. 174, n. 3-4, p. 219-235, 2005.
- CHENG, S. -L.; HWANG, C. Optimal approximation of linear systems by a differential evolution algorithm. **IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans**, v. 31, n. 6, p. 698-707, 2001.
- CHO, K. B.; WANG, B. H. Radial basis function based adaptive fuzzy systems and their applications to system identification and prediction. **Fuzzy Sets and Systems**, v. 83, n. 3, p. 325-339, 1996.
- COTTRELL, M. et al. Neural modeling for time series: a statistical stepwise method for weight elimination. **IEEE Transactions on Neural Networks**, v. 6, n. 6, p. 1355-1364, 1995.
- COULIBALY, P.; BALDWIN, C. K. Nonstationary hydrological time series forecasting using nonlinear dynamic methods. **Journal of Hydrology**, v. 307, n. 1-4, p. 164-174, 2005.
- DOGANIS, P. et al. Time series sales forecasting for short shelf-life food products based on artificial neural networks and evolutionary computing. **Journal of Food Engineering**, v. 75, n. 2, p. 196-204, 2006.
- DRIOLI, C.; ROCCHESO, D. Orthogonal least squares algorithm for the approximation of a map and its derivatives with a RBF network. **Signal Processing**, v. 83, n. 2, p. 283-296, 2003.
- FAMA, E. F. The behavior of stock-market prices. **The Journal of Business**, v. 38, p. 34-105, 1965.
- GAO, Y.; ER, M. J. NARMAX time series model prediction: feedforward and recurrent fuzzy neural network approaches. **Fuzzy Sets and Systems**, v. 150, n. 2, p. 331-350, 2005.
- GHIASSI, M.; SAIDANE, H.; ZIMBRA, D. K. A dynamic artificial neural network model for forecasting time series events. **International Journal of Forecasting**, v. 21, n. 2, p. 341-362, 2005.
- GOLUB, G. H.; VAN LOAN, C. F. **Matrix computations**. Baltimore: The John Hopkins University Press, 1983.
- GORINEVSKY, D. On the persistency of excitation in radial basis function network identification of nonlinear systems. **IEEE Transactions on Neural Networks**, v. 6, n. 5, p. 1237-1244, 1996.
- GUSTAFSON, D. E.; KESSEL, W. C. Fuzzy clustering with a fuzzy covariance matrix. In: IEEE CONFERENCE ON CONTROL AND DECISION, 1979, San Diego, CA, USA. **Proceedings...** p. 761-766.
- HANSEN, J.; NELSON, R. Time-series analysis with neural networks and ARIMA-neural network hybrids. **Journal of Experimental and Theoretical Artificial Intelligence**, v. 15, n. 3, p. 315-330, 2003.
- HIBBERT, H.; PEDREIRA, C.; SOUZA, R. Combining neural networks and ARIMA models for hourly temperature forecast. In: INTERNATIONAL CONFERENCE ON NEURAL NETWORKS, 2000, Como, Italy. **Proceedings...** p. 414-419.
- KABOUDAN, M. A. Genetically evolved models and normality of their fitted residuals. **Journal of Economic Dynamics & Control**, v. 25, n. 11, p. 1719-1749, 2001.
- KANT, H.; SCHREIBER, T. **Nonlinear time series analysis**. Cambridge: Cambridge University Press, 1997.
- KASABOV, N. K.; SONG, Q. DENFIS: dynamic evolving neural-fuzzy inference system and its application for time-series prediction. **IEEE Transactions on Fuzzy Systems**, v. 10, n. 2, p. 144-154, 2002.
- KAYMAK, U.; SETNES, M. Fuzzy clustering with volume prototypes and adaptive cluster merging. **IEEE Transactions on Fuzzy Systems**, v. 10, n. 6, p. 705-712, 2002.
- KRISHNAPURAM, R.; KIM, J. A note on the Gustafson-Kessel and adaptive fuzzy clustering algorithms. **IEEE Transactions on Fuzzy Systems**, v. 7, n. 4, p. 453-461, 1999.
- KUO, R. J.; WU, P.; WANG, C. P. An intelligent sales forecasting system through integration of artificial neural networks and fuzzy neural networks with fuzzy weight elimination. **Neural Networks**, v. 15, n. 7, p. 909-925, 2002.
- LEONG, T. K.; SARATCHANDRAN, P.; SUNDARARAJAN, N. Real-time performance evaluation of the minimal radial basis function network for identification of time varying nonlinear systems. **Computers & Electrical Engineering**, v. 28, n. 2, p. 103-117, 2002.
- LIN, C. -J.; CHEN, C. -H. A compensation-based recurrent fuzzy neural network for dynamic system identification. **European Journal of Operational Research**, v. 172, n. 2, p. 696-715, 2006.
- LJUNG, L. **System identification: theory for the user**. Prentice-Hall: New York, NY, USA, 1987.
- LO, J. T. -H. Multilayer perceptrons and radial basis functions are universal robust approximators. In: International Joint Conference on Neural Networks, 1998, Anchorage, AK, USA. **Proceedings...** v. 2, p. 1311-1314.
- MADEIRA, REVISTA DA, **Indústria brasileira de painéis de madeira**, n. 71, ano 12, maio de 2003. Disponível em: <http://www.remade.com.br/revista/index.php>.
- MEDEIROS, M. C.; VEIGA, A. A hybrid linear-neural model for time series forecasting. **IEEE Transactions on Neural Networks**, v. 11, n. 6, p. 1402-1412, 2000.
- MORETTIN, P. A.; TOLOI, C. M. C. **Séries temporais**. 2. ed., Atual: São Paulo, SP, 1987.
- NELSON, M. et al. Time series forecasting using neural networks: Should the data be deseasonalized first?. **Journal of Forecasting**, v. 18, n. 5, p. 359-367, 1999.
- NIE, J. Nonlinear time-series forecasting: a fuzzy-neural approach. **Neurocomputing**, v. 16, n. 1, p. 63-76, 1997.
- PILGRAM, B.; JUDD, K.; MEES, A. Modelling the dynamics of nonlinear time series using canonical variate analysis. **Physica D: Nonlinear Phenomena**, v. 170, n. 2, p. 103-117, 2002.
- POGGIO, T.; GIROSI, F. Networks for approximation and learning. **Proceedings of the IEEE**, v. 78, n. 9, p. 1481-1497, 1990.
- POWELL, M. J. D. Radial basis functions for multivariable interpolation: a review. In: IMA Conference on Algorithms for the Approximation of Functions and Data, Shrivensham, UK, 1985. **Proceedings...** p. 143-167.
- PREVIDI, F. Identification of black-box nonlinear models for lower limb movement control using functional electrical stimulation. **Control Engineering Practice**, v. 10, n. 1, p. 91-99, 2002.
- PROTIL, R. M. **Desenvolvimento de um sistema computacional para análise de risco em investimentos florestais**. Dissertação de mestrado, PPGA, Universidade Federal do Rio Grande do Sul, UFRGS, Porto Alegre, RS, 1993.
- PROTIL, R. M. Análise de risco em investimentos florestais. In: 32º CONGRESSO DE ECONOMIA E SOCIOLOGIA RURAL, 1994, João Pessoa, PB. **Anais...**

- RANK, E. Application of bayesian trained RBF networks to nonlinear time-series modeling. **Signal Processing**, v. 83, n. 7, p. 1393-1410, 2003.
- RIVAS, V. M. et al. Evolving RBF neural networks for time-series forecasting with EvRBF. **Information Sciences**, v. 165, n. 3-4, p. 207-220, 2004.
- SARIMVEIS, H. et al. fast and efficient algorithm for training radial basis function neural networks based on a fuzzy partition of the input space. **Industrial Engineering Chemical Research**, v. 41, n. 4, p. 751-759, 2002.
- STORN, R. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. **Journal of Global Optimization**, v. 11, n. 4, p. 341-359, 1997.
- STORN, R.; PRICE, K. **Differential evolution: a simple and efficient adaptive scheme for global optimization over continuous spaces**. Technical Report TR-95-012, International Computer Science Institute, Berkeley, CA, USA, 1995.
- SWAIN, A. K.; BILLINGS, S. A. Generalized frequency response function matrix for MIMO non-linear systems. **International Journal of Control**, v. 74, n. 8, p. 829-844, 2001.
- TAN, S.; HAO, J.; VANDEWALLE, J. Efficient identification of RBF neural net models for nonlinear discrete-time multivariable dynamical systems. **Neurocomputing**, v. 9, n. 1, p. 11-26, 1995.
- TANG, Z.; FISHWICK, P. A. Feed-forward neural networks as models for time series forecasting. **ORSA Journal on Computing**, v. 5, n. 4, p. 374-386, 1993.
- TEMIZEL, T. T.; AHMAD, K. Are ARIMA neural network hybrids better than single models?. In: INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS, 2005, Montreal, Canada. **Proceedings...** v. 5, p. 3192-3197.
- TERÄSVIRTA, T.; VAN DIJK, D.; MEDEIROS, M. C. Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: a re-examination. **International Journal of Forecasting**, v. 21, n. 4, p. 755-774, 2005.
- THOMPSON, E. F. The theory of decision under uncertainty and possible application in forest management. **Forest Science**, v. 14, n. 2, p. 156-163, 1968.
- TSENG, F. -M.; YU, H.-C.; TZENG, G.-H. Combining neural network model with seasonal time series ARIMA model. **Technological Forecasting and Social Change**, v. 69, n. 1, p. 71-87, 2002.
- VAN GORP, J. **Nonlinear identification with neural networks and fuzzy logic**. Thesis, Vrije Universiteit Brussel, Faculteit Toegepaste Wetenschappen, Brussel, Belgium, 2000.
- VIRILI, F.; FREISLEBEN, B. Nonstationarity and data preprocessing for neural network predictions of an economic time series. In: INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS, 2000, Como, Italy. **Proceedings...** p. 5129-5136.
- YAO, S.; WEI, C. J.; HE, Z. Y. Evolving wavelet neural networks for function approximation. **Electronics Letters**, v. 32, n. 4, p. 360, 1996.
- YAO, X. Evolving artificial neural networks. **Proceedings of the IEEE**, v. 87, n. 9, p. 1423-1447, 1999.
- YEN, J.; LANGARI, R. **Fuzzy logic: intelligence, control, and information**. Prentice Hall, Upper Saddle River, NJ, USA, 1999.
- ZEMOURI, R.; RACOCEANU, D.; ZERHOUNI, N. Recurrent radial basis function network for time-series prediction. **Engineering Applications of Artificial Intelligence**, v. 16, n. 5-6, p. 453-463, 2003.
- ZHANG, G. -P., 2003, Time series forecasting using a hybrid ARIMA and neural network model. **Neurocomputing**, v. 50, p. 159-175, jan. 2003.
- ZHANG, G. -P.; QI, M. Neural network forecasting for seasonal and trend time series. **European Journal of Operational Research**, v. 160, n. 2, p. 501-514, 2005.
- ZHU, Y. Identification of Hammerstein models for control using ASYM. **International Journal of Control**, v. 73, n. 18, p. 1692-1702, 2000.

Sobre os autores

Leandro dos Santos Coelho

Grupo Produtônica, Programa de Pós-Graduação em Engenharia de Produção e Sistemas – PPGEPS, Centro de Ciências Exatas e de Tecnologia – CCET, Laboratório de Automação e Sistemas, Pontifícia Universidade Católica do Paraná – PUCPR, Rua Imaculada Conceição, 1155, CEP 80215-901, Prado Velho, Curitiba, PR, Brasil, e-mail: leandro.coelho@pucpr.br

Wesley Vieira da Silva

Roberto Max Protil

Programa de Pós-Graduação em Administração, Pontifícia Universidade Católica do Paraná – PUCPR, Rua Imaculada Conceição, 1155, CEP 80215-901, Prado Velho, Curitiba, PR, Brasil, e-mail: wesley.vieira@pucpr.br, protil@ppgia.pucpr.br

Agradecimentos: Os autores agradecem os valiosos comentários do editor e dos revisores que visaram melhorar a fundamentação e a legibilidade do artigo.

Recebido em 17/8/2005
Aceito em 28/11/2006