

REDE NEURAL RBF COM TREINAMENTO BASEADO EM FILTRO DE KALMAN E EVOLUÇÃO DIFERENCIAL APLICADA À PREVISÃO DO PREÇO DA SOJA

RODRIGO CLEMENTE THOM DE SOUZA E LEANDRO DOS SANTOS COELHO

*Laboratório de Automação e Sistemas, Grupo Produtônica
Pós-Graduação em Engenharia de Produção e Sistemas, PUCPR / PPGEPS
Imaculada Conceição, 1155, CEP 80215-901 Curitiba, PR
E-mails: usinadesolucoes@gmail.com; leandro.coelho@pucpr.br*

Abstract Radial Basis Function Neural Networks (RBF-NNs) are function approximation models that can be trained by examples to implement a desired input-output mapping (interpolation). The performance of RBF-NNs depends on the number and centers of the radial basis functions, their shapes, and the method used for learning the input-output mapping. This paper presents a RBF-NN model using Gaussian functions and an extended Kalman filter training approach based on differential evolution optimization. The differential evolution is an evolutionary algorithm which uses a rather greedy and less stochastic approach to problem solving than do classical evolutionary algorithms such as genetic algorithms, evolutionary programming, and evolution strategies. The simulation results indicate the potentialities of this RBF-NN learning approach to one step ahead forecasting of soy price.

Keywords Radial basis function neural networks, Kalman filter, differential evolution, forecasting of time series.

Resumo As Redes Neurais Função de Base Radial (RNs-RBF) são modelos de aproximação de funções que podem ser treinados por exemplos para implementar um mapeamento (interpolação) de entrada e saída desejado. O desempenho das RNs-RBF depende do número e centros das funções radiais de base, suas formas, e o método utilizado para aprender o mapeamento de entrada-saída. Este artigo apresenta uma abordagem de RN-RBF usando funções Gaussianas e treinamento usando uma abordagem de filtro de Kalman estendido baseada em otimização através de evolução diferencial. A evolução diferencial é um algoritmo evolutivo que utiliza uma abordagem mais gulosa e menos estocástica para resolver problemas do que fazem os algoritmos evolutivos clássicos, tais como algoritmos genéticos, programação evolutiva e estratégia evolutiva. Os resultados de simulação indicam as potencialidades desta abordagem de aprendizado de RN-RBF para previsão de curto prazo (um passo à frente) do preço da soja.

Palavras-chave Redes neurais função de base radial, filtro de Kalman, evolução diferencial, previsão de séries temporais.

1 Introdução

Uma rede neural que pode ser aplicada com sucesso à previsão de séries temporais é a rede neural de base radial (RN-RBF, *Radial basis function neural network*).

Uma parte do interesse pelas RNs-RBF é proveniente do fato delas possuírem fundamentação estatística e são consideradas aproximadores universais de funções com dada precisão (Chen *et al.*, 1991), realizando um mapeamento das entradas de rede de dimensão vetorial m , em uma saída de dimensão vetorial n de qualquer função vetorial contínua.

As RNs-RBF apresentam potencialidades de generalização local, em que a interferência no aprendizado é minimizada e o aprendizado é relativamente rápido.

As potencialidades da utilização de RNs-RBF para a previsão de séries temporais é difundida na literatura com a utilização de diferentes abordagens de métodos de treinamento, tais como algoritmos genéticos (Doganis *et al.*, 2006), mínimos quadrados ortogonal (Billings e Hong, 1998), métodos de agrupamentos de dados (Guerra, 2006), entre outros.

Na literatura, abordagens de treinamento de redes neurais usando filtro de Kalman têm sido

propostas, tais como Puskorius e Feldkamp (1994), Birgmeier (1995), Simon (2002), Xiong e Zhou (2003), Bai e Guo (2004), Schauer *et al.* (2005) e Zhang e Luh (2005).

A contribuição deste trabalho é avaliar as potencialidades de usar uma abordagem de treinamento de RNs-RBF baseada em filtro de Kalman estendido (EKF, *Extended Kalman Filter*) com otimização usando evolução diferencial (Storn e Price, 1995; Storn, 1997) para previsão de séries temporais do preço da soja.

O restante do artigo é organizado da seguinte forma. A descrição e a abordagem de treinamento da RN-RBF usando filtro de Kalman e evolução diferencial é detalhada na seção 2. Uma breve descrição da série histórica da soja e a análise dos resultados obtidos são detalhadas na seção 3. Finalizando, a conclusão e comentários em relação aos resultados obtidos são apresentados na seção 4.

2 Rede Neural de Base Radial

Para uma RN-RBF, as unidades ocultas fornecem um conjunto de “funções” que consistem em uma “base” arbitrária para os padrões (vetores) de entrada, quando eles são expandidos sobre o espaço oculto: estas funções são denominadas de funções de base radial, das quais a rede deriva seu nome. Cada

uma dessas funções é centrada em uma particular coordenada do espaço multidimensional dos pontos que compõem o espaço de dados de entrada. Cada uma destas coordenadas particulares caracteriza-se por definir o centro de uma (entre várias possíveis) região de maior aglomeração de pontos ou grupos (*clusters*), do espaço de dados de entrada (Haykin, 1996).

A arquitetura da RN-RBF, em sua estrutura mais básica, é constituída de três camadas, com propósitos totalmente diferentes:

- *Camada de entrada*: é composta por nós de fonte (unidades sensoriais) que conectam a rede ao seu ambiente;
- *Camada oculta*: é única, aplica uma transformação não-linear do espaço de entrada para o espaço oculto (na maioria das aplicações, o espaço oculto é de alta dimensionalidade);
- *Camada de saída*: é linear, fornecendo a resposta da rede ao padrão (sinal) de ativação aplicada à camada de entrada.

O processamento da entrada de uma rede neural se realiza de acordo com sua topologia e o tipo da ativação dos neurônios. Em uma RN-RBF clássica, este aspecto consiste na alimentação adiante (*feedforward network*) de uma entrada de dimensão m em uma camada oculta, que possui c neurônios ocultos que implementam uma função de ativação baseada em uma medida de similaridade entre os vetores de entrada e os protótipos. Cada neurônio da camada oculta possui um protótipo. As saídas da camada oculta são interligadas às unidades da camada de saída através de conexões com pesos sinápticos que ponderam cada saída dos neurônios presentes na camada oculta. Cada unidade da camada de saída implementa uma soma ponderada das saídas da camada oculta. A resposta das camadas de saída forma os elementos do vetor de resposta da RN-RBF.

Uma função de base radial, f_c , apresenta uma saída simétrica em torno de um centro associado, \mathbf{m}_c , tal que $f_c(x) = f(\|x - \mathbf{m}_c\|; \mathbf{s})$, onde $\|\cdot\|$ é a norma vetorial. Uma condição suficiente para uma função ser candidata à função de base radial é a de ser radialmente simétrica e possuir primeira derivada monotônica. A função Gaussiana é o tipo de função de base radial utilizada neste artigo. As funções Gaussianas são caracterizadas por um parâmetro de escala (ou comprimento), \mathbf{s} , e isto é verdade para muitas outras classes de funções de base radial. Tais classes podem ser representadas como:

$$f_c(x) = f(\|x - \mathbf{m}_c\|; \mathbf{s}) \quad (1)$$

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

$$y(x) = \sum_{j=1}^M w_j f(\|x - \mathbf{m}_c\|; \mathbf{s}) \quad (2)$$

onde w_j são os pesos sinápticos e M é o número de protótipos. Para identificação de sistemas, as RN-RBFs modelam uma função implícita de um fenômeno físico. O modelo neural de identificação é paramétrico do tipo “caixa-preta”, pois a RN-RBF contém parâmetros livres sem significado físico a serem ajustados para que a rede consiga representar a função implícita do processo.

Os vetores de protótipos e pesos sinápticos são os parâmetros da estrutura da RN-RBF. Deve-se ressaltar que os valores desses parâmetros não acrescentam nada de conhecimento a respeito do sistema, porém devem cumprir o determinado: aproximar a resposta do sistema a uma determinada entrada. Para tanto a estimação dos parâmetros das redes neurais em geral, incluindo as RNs-RBF pode ser concebido como um problema de aprendizado com uma “possível” interpretação biológica e também com equivalência matemática (problema de otimização não-linear na estimação dos parâmetros do modelo segundo um critério de otimalidade) (Haykin, 1996).

Alguns pesquisadores têm treinado as RNs-RBF, escolhendo os centros das unidades ocultas aleatoriamente no conjunto de treinamento. Outros têm utilizado métodos não-supervisionados (por exemplo, agrupamento pelo método de k -médias e fuzzy c -médias), como também algoritmos supervisionados para a seleção dos centros (Simon, 2002). Os métodos de treinamento separam as tarefas de determinação dos protótipos e a otimização dos pesos. A vantagem dessa abordagem é a rapidez no treinamento da rede como resultado da separação das tarefas e aplicação de um método híbrido de treinamento em alternativa ao treinamento global e a otimização não-linear explícita da descida de encosta (incluindo o algoritmo de retro-propagação do erro) (Chen *et al.*, 1991).

A separação das tarefas no treinamento da rede é possível devido à estrutura da RN-RBF e seu tipo de função de ativação que age localmente no espaço de entrada. Foi dada uma considerável atenção na literatura das RNs-RBF em como obter um método de treinamento baseado em técnicas lineares explorando características estruturais das RNs-RBF. Deve-se enfatizar que se os protótipos são fixados e a função da camada escondida $g(\cdot)$ conhecida, a rede é linear em seus pesos sinápticos w . Assim, a estrutura da RN-RBF possibilita a separação de tarefas no processo de aprendizado, com a vantagem que na estimação dos pesos w o treinamento da rede se torna uma tarefa rápida de ser resolvida através de métodos do tipo mínimos quadrados.

Segundo Chen *et al.* (1991), “o problema do mal-condicionamento nas estratégias híbridas pode, por exemplo, ocorrer devido aos centros estarem próximos, resultando em uma dependência linear

entre ambos os centros”. Assim o mal-condicionamento numérico da rede pode ser contornado pelos métodos de agrupamento (*clustering*) que faz a escolha dos protótipos como sendo representantes de populações encontradas nos dados de entrada por aprendizado não supervisionado. Observa-se nos métodos que exploram a localidade da RN-RBF que somente as entradas similares aos protótipos serão capazes de contribuir de forma efetiva na resposta da rede e também não usam toda informação fornecida dos exemplos de entrada(s) e saída do conjunto de treinamento na fase de estimação dos parâmetros, acarretando em uma pobre generalização no aprendizado da rede. Neste contexto, o algoritmo do filtro de Kalman é descrito na subseção apresentada a seguir.

2.1 Filtro de Kalman estendido

Criado por Rudolph Emil Kalman em 1960, o filtro de Kalman foi desenvolvido inicialmente como uma solução recursiva para filtragem linear de dados discretos (Kalman, 1960). Para isto, utiliza equações matemáticas que implementam um estimador preditivo de estados, buscando corrigir interativamente a resposta de um determinado sistema através de múltiplas variáveis relacionadas a ele. Suas áreas de aplicação são muito diversificadas, tais como: processamento de imagem, supervisores de eventos discretos, processamento de sinais, sistemas de inferência, entre outros.

O filtro de Kalman original (Kalman, 1960; Kalman e Bucy, 1961) foi concebido para ser aplicado em sistemas com dinâmica linear. Para um sistema dinâmico linear com ruído branco de processo de medição, o filtro de Kalman é um estimador ótimo dos estados do sistema. A fim de tornar a aplicação do filtro de Kalman em sistemas não-lineares, pesquisadores propuseram o filtro de Kalman estendido (Kuchner, 1967; Gordon *et al.*, 1993), o que se mostrou eficiente para muitos sistemas cuja não-linearidade é pouco significativa. Para sistemas não-lineares com ruído colorido, o filtro de Kalman pode ser estendido pela linearização da matriz de transição não-linear da descrição em variáveis de estado do sistema em torno dos parâmetros estimados correntes (Haykin, 2001). Neste contexto, o filtro de Kalman estendido atualiza os parâmetros de forma a ser consistente com todos os dados medidos do processo e geralmente converge em poucas interações (Simon, 2002; Welch e Bishop, 1995).

A formulação dos filtros de Kalman está fundamentada na representação em espaço de estados de um sistema dinâmico linear discreto de dimensão finita, ou seja, o filtro de Kalman é descrito em termos de equações a diferenças, porém se o sistema for não-linear, pode-se estender o uso do filtro de Kalman com um procedimento de linearização de um sistema da forma

$$\begin{aligned} \mathbf{q}_{k+1} &= f(\mathbf{q}_k) + \mathbf{w}_k \\ y_k &= h(\mathbf{q}_k) + \mathbf{u}_k \end{aligned} \quad (3)$$

onde o vetor \mathbf{q}_k é o estado do sistema no instante de tempo K , \mathbf{w}_k é o ruído provocado pela própria dinâmica do processo, y_k é o vetor de observação, \mathbf{u}_k é o ruído de observação (ou de medida), e $f(\cdot)$ e $h(\cdot)$ são funções não-lineares.

A idéia básica do filtro de Kalman estendido consiste em linearizar a equação (3) a cada instante de tempo em torno do estado estimado recente. Se as não-linearidades são suficientemente suaves, pode-se fazê-lo aplicando a aproximação de Taylor de primeira ordem nas funções não-lineares v e $h(\cdot)$ em torno do estado estimado $\hat{\mathbf{q}}$ para obter

$$f(\mathbf{q}_k) = f(\hat{\mathbf{q}}_k) + F_k(\mathbf{q}_k - \hat{\mathbf{q}}_k) + \text{termos de ordem elevada} \quad (4)$$

$$h(\mathbf{q}_k) = h(\hat{\mathbf{q}}_k) + H_k^T(\mathbf{q}_k - \hat{\mathbf{q}}_k) + \text{termos de ordem elevada} \quad (5)$$

sendo adotada a seguinte notação para as derivadas parciais:

$$F_k = \left. \frac{\partial f(\mathbf{q})}{\partial \mathbf{q}} \right|_{\mathbf{q}=\hat{\mathbf{q}}_k} \quad (6)$$

$$H_k^T = \left. \frac{\partial h(\mathbf{q})}{\partial \mathbf{q}} \right|_{\mathbf{q}=\hat{\mathbf{q}}_k} \quad (7)$$

Desconsiderando os termos de ordem mais elevada na equação (5), um sistema linear aproximado é obtido de (4), onde

$$\mathbf{q}_{k+1} = F_k \mathbf{q}_k + \mathbf{w}_k + \mathbf{f}_k \quad (8)$$

$$y_k = H_k^T \mathbf{q}_k + v_k + \mathbf{j}_k \quad (9)$$

Neste caso, \mathbf{f}_k e \mathbf{j}_k são definidos por

$$\mathbf{f}_k = f(\hat{\mathbf{q}}_k) - F_k \hat{\mathbf{q}}_k \quad (10)$$

$$\mathbf{j}_k = h(\hat{\mathbf{q}}_k) - H_k^T \hat{\mathbf{q}}_k \quad (11)$$

A obtenção da aproximação linear do sistema representado pela equação (4) é o primeiro passo para utilização do filtro de Kalman. O problema do filtro de Kalman consiste em resolver simultaneamente e de maneira ótima as equações de processo e de medida para um estado desconhecido.

Pode-se formular o problema de filtragem de um sistema linear achando um valor ótimo de $\hat{\mathbf{q}}_k$, tal que o sistema deve satisfazer as restrições impostas pelos teoremas do estimador médio condicional e o do princípio da ortogonalidade. Finalmente, pode-se mostrar que a estimação desejada $\hat{\mathbf{q}}_k$ pode ser obtida recursivamente pelas equações do filtro de Kalman:

$$\hat{\mathbf{q}}_k = f(\hat{\mathbf{q}}_{k-1}) + K_k [y_k - h(\hat{\mathbf{q}}_{k-1})] \quad (12)$$

$$K_k = P_k H_k (R + H_k^T P_k H_k)^{-1} \quad (13)$$

$$P_{k+1} = F_k (P_k - K_k H_k^T P_k) F_k^T + Q \quad (14)$$

onde K_k é o ganho de Kalman e P_k a matriz de covariância do erro de estimação do estado \mathbf{q}_k . As matrizes Q e R são as matrizes de covariância dos ruídos artificiais \mathbf{w}_k e \mathbf{u}_k , respectivamente. Como o sistema descrito nas equações (8) e (9) é uma aproximação de um sistema não-linear, o filtro não é ótimo e a estimação é somente uma aproximação da média condicional.

O problema de otimização da matriz de pesos sinápticos W e as coordenadas dos vetores protótipos v_j sob o ponto de vista da estimação ótima de parâmetros com filtro de Kalman de uma RN-RBF pode ser tratado como um problema de minimização do tipo mínimo quadrados ponderado. Neste caso, para aplicar o EKF, faz-se com que os elementos da matriz de pesos W e os elementos dos vetores de protótipos v_j constituam o estado do sistema discreto não-linear em que o filtro de Kalman é aplicado.

2.2 Evolução diferencial

A evolução diferencial (ED) é um paradigma da computação evolutiva desenvolvido por Rainer Storn e Kenneth Price (Storn e Price, 1995; Storn, 1997) para problemas de otimização não-linear contínua. Basicamente, a ED realiza mutações nos vetores pela adição ponderada de diferenças aleatórias entre eles. A escolha da ED para problemas de otimização é baseada nas características de que a ED:

- é um algoritmo de busca estocástica que é motivado pelos mecanismos de seleção natural;
- é menos susceptível a mínimos (ou máximos) locais, pois busca a solução ótima global pela manipulação de uma população de soluções candidatas, ou seja, busca um número de diferentes áreas simultaneamente no espaço de busca;
- não requer informação de derivadas para o cálculo da função de aptidão;
- geralmente não necessita de populações grandes para funcionar eficientemente.

Na ED clássica, cada variável (indivíduo) é representada por um valor real (ponto flutuante). A variante implementada neste trabalho foi a ED/rand/1/bin.

3 Série Histórica e Resultados de Simulação

3.1 Série histórica da soja

O Indicador de preço de soja CEPEA/ESALQ, (Centro de Estudos Avançados em Economia Aplicada/ESALQ) antes denominado ESALQ/BM&F, é uma média aritmética ponderada dos preços observados no Paraná para soja em grão a granel, tipo exportação, conforme padrão Concox: até 14% de umidade, até 2% de impurezas, e limites máximos de 8% para grãos avariados (até 5% de ardidos) e 30% de grãos quebrados (CEPEA, 2007).

O Indicador da Soja é divulgado diariamente pelo CEPEA desde agosto de 1997 e é utilizado por agentes e instituições que lidam com o mercado físico como uma das principais referências de preços. Apesar do Indicador considerar apenas os preços do Paraná, um estado importante na produção, industrialização e exportação de soja, a equipe CEPEA consulta agentes também nos estados do Rio Grande do Sul, Mato Grosso, Mato Grosso do Sul, Goiás, Minas Gerais e São Paulo (CEPEA, 2007).

Para a obtenção do Indicador, o estado do Paraná foi dividido em cinco regiões: Porto de Paranaguá, Ponta Grossa, Sudoeste, Oeste e Norte. Em cada região é consultada diariamente uma amostra representativa de compradores, vendedores, corretores e comerciantes que fornecem os preços, livres de impostos, que comporão o Indicador deste dia. Os cálculos do desconto de pagamento à vista são feitos a partir da NPR (Nota Promissória Rural) coletada diariamente junto a vários bancos (CEPEA, 2007).

O Indicador será determinado pela média aritmética geral das médias aritméticas simples regionais ponderadas pelas participações relativas das regiões na capacidade estática total de esmagamento do estado do Paraná, no ano imediatamente anterior. Por exemplo, para 1997, as ponderações regionais para o indicador de preços de soja foram de: Porto de Paranaguá (19,13%), Ponta Grossa (25,49%), Norte (34,72%), Oeste (9,73%) e Sudoeste (10,93%). O percentual de participação de cada região é determinado pela sua capacidade instalada de esmagamento (CEPEA, 2007).

Essas participações serão atualizadas, desde que a média móvel dos últimos dois anos referente à capacidade estática de processamento do estado do Paraná, com dados fornecidos pela ABIOVE (Associação Brasileira das Indústrias de Óleos Vegetais), apresente mudanças estatisticamente significativas. Em caso de não existirem preços em certa região, esta localidade não será considerada e sua participação será distribuída proporcionalmente entre as demais para cálculo do Indicador (CEPEA, 2007).

3.2 Análise dos resultados

O procedimento de análise da identificação através de RN-RBF é dividido nas seguintes etapas: (i) escolha da estrutura da RN-RBF para representar o sistema a ser identificado, (ii) otimização da RN-RBF usando agrupamento de dados e pseudo-inversa; (iii) estimação dos parâmetros do modelo matemático (*etapa de estimação*), e (iv) a validação do modelo matemático (*etapa de validação*).

O espaço de busca usado na ED para otimização dos três parâmetros de projeto do EKF foi o seguinte: $P \in [170;300]$, $Q \in [170;300]$ e $R \in [0,1;6]$.

Neste caso a função de aptidão usada na ED foi o da maximização do coeficiente de correlação múltipla, também denominado de R^2 , tal que

$$R^2 = \frac{\sum_{t=1}^N [y(t) - \hat{y}(t)]^2}{\sum_{t=1}^N (y(t) - \bar{y}(t))^2} \quad (15)$$

onde $y(t)$ é a saída real do sistema, $\hat{y}(t)$ é a saída estimada, e $\bar{y}(t)$ é a média das N amostras medidas (média da saída real). Quando o valor de R^2 é igual a 1, indica uma exata adequação do modelo para os dados medidos do sistema. O valor de R^2 entre $0,9 \leq R^2 \leq 1$ é considerado suficiente para a maioria das aplicações práticas em previsão de séries temporais.

Foram usadas as amostras 1 a 1400 do preço da soja (preços diários entre 29/07/1997 e 21/03/2003) na fase de estimação da RN-RBF, e as amostras de 1401 a 2000 na fase de validação (preços diários de 24/03/2003 a 12/08/2005).

Os valores dos termos da diagonal principal das matrizes P , Q e R foram obtidos usando uma abordagem de ED com uma população de 5 e critério de parada de 15 gerações totalizando 75 avaliações da função objetivo R^2 por rodada de otimização.

Foram realizadas 30 simulações com cada abordagem de projeto da RN-RBF, ou seja, com número de funções Gaussianas, M , na camada oculta entre 3 e 7 para previsão um passo à frente (previsão de curtíssimo prazo) do preço da soja. Na Tabela 1 é apresentado o resultado para o projeto do EKF (executado apenas em uma época de aprendizado) baseado ED (30 simulações) para o treinamento da RN-RBF. As tabelas 1 e 2 apresentam, respectivamente os resultados de previsão do preço da soja pela RN-RBF e os melhores projetos do EKF para o treinamento da RN-RBF. O melhor resultado obtido pela ED para o projeto do EKF usado na RN-RBF foi para o caso de $M = 7$. Entretanto, nota-se pelos resultados da tabela 1 que os resultados para valores de M entre 4 e 8 são próximos.

Tabela 1. Resultados da RN-RBF para previsão do preço da soja (30 simulações).

| M | R^2 (durante a etapa de estimação) | | | |
|-----|--------------------------------------|-----------------|--------|---------------|
| | Mínimo (pior) | Máximo (melhor) | Média | Desvio padrão |
| 3 | 0,7326 | 0,8369 | 0,7726 | 0,078 |
| 4 | 0,9443 | 0,9960 | 0,9831 | 0,023 |
| 5 | 0,9303 | 0,9801 | 0,9303 | 0,040 |
| 6 | 0,9177 | 0,9872 | 0,9461 | 0,031 |
| 7 | 0,9780 | 0,9976 | 0,9927 | 0,085 |
| 8 | 0,9914 | 0,9974 | 0,9944 | 0,029 |

Tabela 2. Melhores projetos do EKF para o treinamento da RN-RBF.

| M | R^2 | P | Q | R |
|-----|--------|----------|----------|--------|
| 3 | 0,8369 | 230,3714 | 150,0126 | 0,6095 |
| 4 | 0,9960 | 258,5130 | 240,3800 | 0,3077 |
| 5 | 0,9801 | 197,1772 | 209,8963 | 0,4519 |
| 6 | 0,9872 | 209,1205 | 77,6371 | 0,1164 |
| 7 | 0,9974 | 226,1313 | 109,4076 | 0,1206 |
| 8 | 0,9976 | 280,1469 | 139,9563 | 0,1366 |

Na Figura 1 é apresentado o resultado de previsão de curto prazo para o preço da soja. O erro máximo de previsão, neste caso, foi -3,37 e o R^2 na fase de validação (amostras 1401 a 2000) foi de 0,9893.

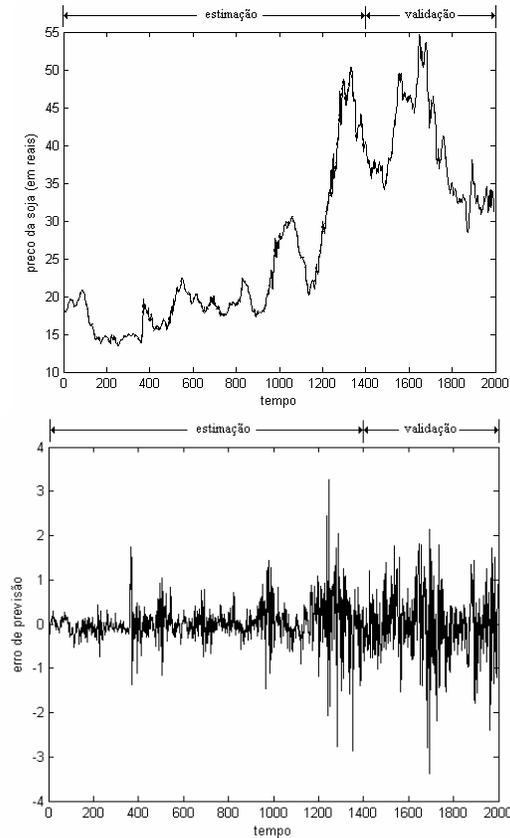


Figura 1. Estimativa de curto prazo para o preço da soja usando uma RN-RBF com $M=7$.

4 Conclusão e Futura Pesquisa

Este trabalho apresentou uma abordagem de RN-RBF baseada em treinamento por EKF híbrido com ED. Os resultados preliminares de simulação indicam as potencialidades desta abordagem de rede neural para identificação de curto prazo (um passo à frente) do preço da soja.

O filtro de Kalman possibilitou lidar com informações ruidosas do sistema dinâmico, utilizando informações matemáticas para estimar o estado atual do ambiente no qual está inserido. Uma dificuldade no projeto do filtro de Kalman foi o projeto das matrizes P , Q e R . Neste contexto, utilizou-se a técnica de evolução diferencial para otimizar os termos da diagonal principal das matrizes de P , Q e R . Os resultados de empregar filtro de Kalman com ED no projeto da RN-RBF obtidos foram satisfatórios em termos de R^2 .

Em pesquisa futura, deseja-se realizar um estudo comparativo e uma análise estatística de diferentes abordagens de treinamento de RNs-RBF usando métodos dos mínimos quadrados, ED, filtro de Kalman, pseudo-inversa com fator de regularização,

entre outros, visando aplicações de previsão de séries temporais a médio e longo prazos na área de agro-negócio.

Referências Bibliográficas

- Bai, L. e Guo, P. (2004). A novel approach to stellar recognition by combining EKF and RBF-net, *Advances in Neural Networks*, International Symposium on Neural Networks, Dalian, China, Lecture Notes in Computer Science, vol. 3173/2004, Springer, pp. 1002-1007.
- Billings, S. A. e Xong, X. (1998). Dual-orthogonal radial basis function networks for nonlinear time series prediction, *Neural Networks*, **11**(3): 479-493.
- Birgmeier, M. (1995). A fully Kalman-trained radial basis function network for nonlinear speech modeling, *Proceedings of IEEE International Conference on Neural Networks*, Perth, Australia, Vol. 1, pp. 259-264.
- CEPEA (2007). <http://www.cepea.esalq.usp.br/> [Acesso em 10 de Maio de 2007].
- Chen, S., Cowan, C. e Grant, P. (1991). Orthogonal least squares learning algorithm for radial Basis function networks, *IEEE Transactions on Neural Networks*, **2**(2): 302-309.
- Doganis, P., Alexandridis, A., Patrinos, P. e Sarimveis, H. (2006). Time series sales forecasting for short shelf-life food products based on artificial neural networks and evolutionary computation, *Journal of Food Engineering*, **75**(2): 196-204.
- Gordon, N. J., Salmond, D. J. e Smith, A. F. M. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEEE Proceedings-F*, **140**(2): 107-113.
- Guerra, F. A. (2006). Análise de métodos de agrupamento para o treinamento de redes neurais de base radial em aplicações de identificação de sistemas, *Dissertação de mestrado*, PPGEPS/PUCPR, Curitiba, PR.
- Haykin, S. (1996). *Neural networks*, 2nd edition, Prentice-Hall, Upper Saddle River, NJ, USA.
- Haykin, S. (2001). *Kalman filtering and neural networks*, John Wiley & Sons, Inc, NY, USA.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems, *Transactions of the ASME – Journal of Basic Engineering*, **83**: 35-45.
- Kalman, R. E. e Bucy R. S. (1961). New results in linear filtering and prediction theory, *Transactions of the ASME - Journal of Basic Engineering*, **83**: 95-107.
- Kuchner, H. J. (1967). Approximations to optimal nonlinear filters, *IEEE Transactions on Automatic Control*, **12**(5): 546-556.
- Puskorius, G. V. e Feldkamp, L. A. (1994). Neurocontrol of nonlinear dynamical systems with Kalman filter trained recurrent networks, *IEEE Transactions on Neural Networks*, **5**(2): 279-297.
- Schauer, T., Negard, N.-O., Previdi, F., Hunt, K.J., Fraser, M.H., Ferchland, E. e Raisch, J. (2005). Online identification and nonlinear control of electrically stimulated quadriceps muscle, *Control Engineering Practice*, **13**(9): 1207-1219.
- Simon, D. (2002). Training radial basis neural networks with the extended Kalman filter, *Neurocomputing*, **48**(1-4): 455-475.
- Storn, R. (1997). Differential evolution — a simple and efficient heuristic for global optimization over continuous spaces, *Journal of Global Optimization*, **11**(4): 341-359.
- Storn, R. e Price, K. (1995). Differential evolution: a simple and efficient adaptive scheme for global optimization over continuous spaces, *Technical Report TR-95-012*, International Computer Science Institute, Berkeley, USA.
- Welch, G. e Bishop, G. (1995). An introduction to the Kalman filter, *Technical Report TR-95-041*, Department of Computer Science, University of North Carolina at Chapel Hill, NC, USA.
- Xiong, S. -S. e Zhou, Z. -Y. (2003). Neural filtering of colored noise based on Kalman filter structure, *IEEE Transactions on Instrumentation and Measurement*, **52**(3): 742-747.
- Zhang, L. e Luh, P. B. (2005). Neural network-based market clearing price prediction and confidence interval estimation with an improved extended Kalman filter method, *IEEE Transactions on Power Systems*, **20**(1): 59-66.