

# **Sistema Híbrido Inteligente Baseado em Sistemas Nebulosos e Estratégias Evolutivas Aplicado à Previsão do Índice da BOVESPA: Fundamentos e Estudo de Caso**

Autoria: Leandro dos Santos Coelho, André Alves Portela Santos

Resumo: Este artigo apresenta o projeto de um sistema nebuloso do tipo Takagi-Sugeno-Kang. Neste projeto é utilizada uma estratégia evolutiva com mecanismos de auto-adaptação para otimização da parte antecedente das regras do sistema nebuloso, enquanto o algoritmo dos mínimos quadrados é utilizado no projeto da parte conseqüente das regras. O sistema nebuloso-evolutivo é avaliado no estudo de caso de identificação e previsão dinâmica um passo a frente para o *Índice da Bolsa de Valores de São Paulo* (Ibovespa). Os resultados de simulação com sistema nebuloso com representação matemática do tipo NARX (*Nonlinear AutoRegressive with eXogenous inputs*) indicam que a proposta de otimização baseada em estratégia evolutiva é uma abordagem promissora para aplicações de identificação não-linear e previsão de séries temporais financeiras.

## **1 INTRODUÇÃO**

A área financeira, a partir dos anos 1980, sofreu modificações radicais devido a revolução tecnológica e com o conseqüentemente estreitamento das relações entre os mercados globais. Neste contexto, diversos métodos numéricos e econométricos têm sido propostos na literatura, sendo estes, muitas vezes, capazes de determinar padrões e relacionamentos nos dados presentes no mercado financeiro e mostram-se relevantes na identificação de oportunidades mercantis.

A concepção de modelos matemáticos para a representação de sistemas dinâmicos complexos e/ou previsão de mercado financeiro é um procedimento relevante com aplicação prática. Entretanto, em geral, não é uma tarefa simples a construção de modelos matemáticos adequados para os propósitos de previsão e identificação de comportamento dinâmico do mercado financeiro. Nas últimas décadas diversas concepções de algoritmos de identificação e previsão de comportamento de sistemas complexos têm sido propostos na literatura, tais como: métodos freqüenciais, técnicas baseadas em estimativas de modelos Wiener, Hammerstein, bilinear e Volterra, regressão linear (Van Gestel et al., 2001), regressão não-linear (Haber & Unbehauen, 1990), wavelets (Billings & Coca, 1999), identificação recursiva (Ljung, 2001), teoria do caos (Schweizer & Schimming, 2001) e inteligência artificial (Gilbert, 1995; Varetto, 1998; Chen et al., 2001).

Uma abordagem relevante, entre tantas outras da inteligência artificial, para a previsão do comportamento de sistemas complexos em finanças é a dos sistemas nebulosos. Nos últimos anos, os pesquisadores (Tseng & Tzeng, 2002; Von Altrock, 2002; León et al., 2002) têm proposto um variado espectro de metodologias de identificação e previsão não-linear baseadas em sistemas nebulosos, a citar os modelos nebulosos de Mamdani, Larsen, Tsukamoto, relacional e Takagi-Sugeno-Kang (TSK), para lidarem com sistemas não-lineares, “mal definidos” e que apresentam incertezas.

Os modelos nebulosos de TSK (Takagi & Sugeno, 1985; Sugeno & Kang, 1988) apresentam características que os tornam promissores para aplicações de identificação de sistemas complexos, entre as quais a de serem aproximadores universais de funções com precisão arbitrária como algumas abordagens de projeto com redes neurais artificiais (Ying, 1998).

Além disso, os modelos nebulosos de TSK apresentam alto poder de interpolação com um número reduzido de regras de produção.

Existe um crescente interesse na integração de sistemas nebulosos e algoritmos evolutivos. A classificação bibliográfica de 562 referências tratando a combinação de sistemas nebulosos e algoritmos evolutivos nas mais diversas áreas foi realizada por Cordón et al. (1995). Este artigo apresenta o projeto de um sistema híbrido inteligente composto por um sistema nebuloso de TSK baseado em otimização através de uma estratégia evolutiva (EE) com adaptação eficiente do sistema independente de coordenadas da distribuição normal do operador de mutação. A EE é utilizada para o projeto da parte antecedente, enquanto o algoritmo dos mínimos quadrados é utilizado na otimização da parte conseqüente do sistema nebuloso. Este sistema híbrido nebuloso-evolutivo é avaliado no estudo de caso de identificação e previsão dinâmica um passo a frente para o *Índice da Bolsa de Valores de São Paulo* (Ibovespa).

O artigo é organizado da seguinte forma. Na seção 2, os fundamentos do sistema nebuloso de TSK e a otimização baseada em uma EE e algoritmo dos mínimos quadrados são apresentados. A descrição dos dois estudos de caso de identificação de sistemas é abordada na seção 3. As simulações e a análise dos resultados obtidos da aplicação da abordagem nebuloso-evolutiva no estudo de caso mencionado são apresentadas na seção 4. Na seção 5 são apresentadas a conclusão e os futuros trabalhos.

## 2 SISTEMA HÍBRIDO INTELIGENTE

O progresso da tecnologia dos sistemas inteligentes é motivado pela necessidade do desenvolvimento de estratégias flexíveis e eficientes com o propósito de tratar aplicações complexas do mundo real. Cada metodologia da inteligência computacional (*soft computing*) —uma subárea da inteligência artificial que abrange um consórcio de metodologias, incluindo as redes neurais, os sistemas nebulosos e os algoritmos genéticos—possui potencialidades e limitações que as tornam mais adequadas para algumas aplicações particulares. Por exemplo, enquanto as redes neurais são adequadas às tarefas de reconhecimento de padrões, não são eficientes para interpretação lingüística. Os sistemas nebulosos podem tratar informação imprecisa e são adequados em explanar suas decisões. Os algoritmos evolutivos são métodos de otimização para o tratamento de projetos complexos, onde, usualmente, empregam-se procedimentos heurísticos do tipo tentativa e erro, mas demandam uma elevada complexidade computacional.

Estas limitações são a principal motivação da criação de sistemas híbridos inteligentes, onde duas (ou mais) técnicas são combinadas para superar as limitações das técnicas quando tratadas individualmente. Os sistemas híbridos inteligentes são relevantes quando consideram-se domínios complexos que apresentam problemas com componentes diferentes, os quais requerem diversos tipos de processamento (Goontilake & Treveaven, 1995). No contexto deste artigo é aplicada uma abordagem de sistema nebuloso-evolutivo para aplicações de previsão no mercado financeiro. A abordagem proposta é apresentada a seguir.

### 2.1 Sistema nebuloso de Takagi-Sugeno-Kang

Os fundamentos teóricos dos conjuntos nebulosos foram propostos por Zadeh (1965), professor da Universidade de Califórnia (Berkeley, Estados Unidos), como uma forma

alternativa de modelar os sistemas complexos, se diferenciando das técnicas convencionais. Alguns trabalhos precursores aos de Zadeh foram de suma importância para o desenvolvimento dos sistemas nebulosos, entre os quais:

- (i) década de 20: os desenvolvimentos da lógica multivalorada pelo matemático polonês J. Lukasiewicz;
- (ii) década de 30: os estudos do filósofo quântico Max Black que apresentou as noções de funções de pertinência e as definições de estruturas vagas;
- (iii) década de 50: Menger, um matemático francês, propôs o termo *ensemble flou*, uma contrapartida ao termo *fuzzy set*, que seria proposto posteriormente por Zadeh (1965, 1973).

Os sistemas nebulosos possui um formalismo para a representação do conhecimento e inferência de novos conhecimentos que é similar à maneira utilizada pelos seres humanos para expressarem o conhecimento e raciocínio, ao contrário dos sistemas baseados em lógica clássica.

Os sistemas nebulosos têm, recentemente, emergido como uma abordagem atrativa para o controle de processos complexos. A motivação deve-se ao conhecimento do sistema e ao comportamento dinâmico qualitativo e com incertezas, pois a teoria dos sistemas nebulosos providencia uma representação adequada para tal conhecimento. Além disso, os sistemas nebulosos são capazes de manipular simultaneamente tanto os dados numéricos quanto o conhecimento lingüístico.

As características dos sistemas nebulosos, que os tornam uma alternativa atraente de projeto, nos mais variados campos do conhecimento são:

- (i) a habilidade de tratar processos não lineares: característica relevante à aplicação em identificação e previsão quando da presença de sistemas dinâmicos complexos, com ordem elevada, atraso de transporte da informação e sistemas “mal” definidos;
- (ii) a possibilidade da utilização de termos e expressões da linguagem natural;
- (iii) a possibilidade da implementação do conhecimento, dos aspectos intuitivos e da experiência do especialista em uma determinada área do conhecimento (por exemplo, mercado financeiro), utilizando regras lingüísticas (conhecimento qualitativo e estruturado) e entradas imprecisas (Navarro & Albertos, 1996);
- (iv) a facilidade de entendimento, a simplicidade de implementação, a rápida prototipagem e o custo inexpressivo de desenvolvimento (Wang, 1994).

Existem dois tipos de conhecimento, que podem ser utilizados à resolução de problemas reais, que são: (i) *conhecimento objetivo* (representado por modelos matemáticos), e (ii) *conhecimento subjetivo* (representado por informação lingüística, que é, praticamente impossível de quantificar pela matemática convencional). Estas duas formas de conhecimento podem ser coordenadas de maneira eficiente através da utilização dos sistemas nebulosos. As formas de tratar o conhecimento via sistemas nebulosos são: abordagem baseada em modelo e livre de modelo. Neste contexto, a estrutura básica de um modelo nebuloso pode ser vista com constituindo-se de: (i) *base de regras*, que contém um conjunto de regras nebulosas; (ii) *base de dados*, que define as funções de pertinência das regras nebulosas; e (iii) *mecanismo de inferência*, que executa os procedimentos de inferência sobre as regras e condição para obtenção de uma saída.

Os antecedentes das regras nebulosas dividem o espaço de entrada em um número de regiões nebulosas locais, enquanto os conseqüentes descrevem o comportamento dentro de uma

região através dos seus constituintes. Os componentes dos conseqüentes resultam em tipos diferentes de modelos (sistemas) nebulosos, mas seus antecedentes são essencialmente os mesmos.

Os componentes dos conseqüentes resultam em tipos diferentes de modelos nebulosos, mas seus antecedentes são essencialmente os mesmos. Uma classificação baseada nos componentes dos conseqüentes pode ser usada para classificar os sistemas nebulosos em baseados em (Yager & Filev, 1994):

- (i) *modelos lingüísticos*: são baseados em regras *se-então* apresentando predicados vagos e utilizando raciocínio nebuloso. Nestes modelos, as quantidades nebulosas são associadas aos termos lingüísticos, e o modelo nebuloso é essencialmente uma expressão qualitativa do sistema. Os modelos, deste tipo, formam a base de modelagem qualitativa, que descreve o comportamento do sistema através da utilização da linguagem natural. O representante mais significativo desta classe é o modelo ou sistema do tipo Mamdani (Mamdani & Assilian, 1975);
- (ii) *modelos interpolativos*: são formados por regras lógicas por uma parte nebulosa antecedente e um conseqüente funcional. Estes modelos são, essencialmente, uma combinação de modelos nebulosos e não-nebulosos. Os representantes mais significativos dos modelos interpolativos são os modelos ou sistemas do tipo Takagi-Sugeno-Kang e do tipo Tsukamoto. A representação do conhecimento, nestes modelos, caracterizam-se pelo potencial de expressar informações quantitativas. Entretanto, nos modelos interpolativos perde-se a informação qualitativa.

Neste caso, a configuração de sistemas de inferência nebulosos, usualmente, envolve três estruturas, ilustradas pela figura 1. A primeira estrutura é relacionada com a informação do processo contida no conjunto de treinamento. A informação é providenciada pelo conhecimento sobre as características físicas do processo, por alguns dados numéricos adquiridos dele, e pelo conhecimento lingüístico, expresso através das regras. Esta estrutura envolve o *conjunto das variáveis*, que regem o comportamento do processo, e a *partição nebulosa* atribuída a cada uma destas variáveis.

A segunda estrutura é o algoritmo de aprendizado, baseado nas características da primeira estrutura. Nesta estrutura, a informação adquirida é interpretada e codificada em um conjunto de regras. A terceira estrutura executa o procedimento de inferência considerando as regras extraídas e realiza as ações sobre o processo, no caso de sistemas dinâmicos, completando o ciclo de identificação nebulosa (Branco & Dente, 1997).

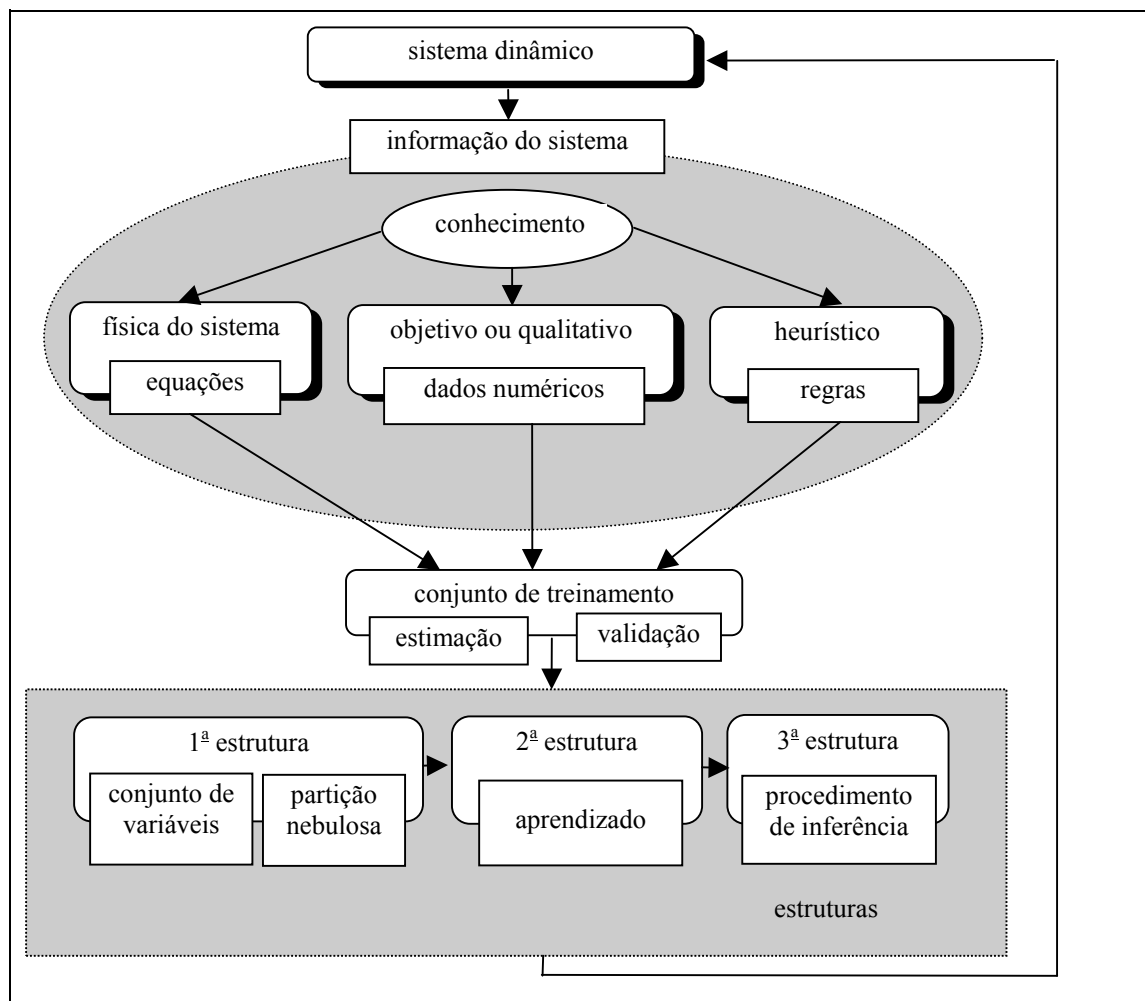


Figura 1. Esquema de configuração de identificação (ou previsão) nebulosa e suas estruturas.

Entre as características de aproximação, as seguintes são relevantes à utilização dos sistemas nebulosos em identificação de processos (Zeng & Singh, 1995): (i) *o mecanismo de aproximação nebuloso*: As regras lingüísticas são utilizadas para aproximar a saída desejada. Para uma dada entrada, as saídas correspondentes do sistema nebuloso resultam da combinação de diversas regras e consistem numa aproximação da saída desejada; (ii) *a aproximação global limitada*: a aproximação dos sistemas nebulosos é baseada em informação local, pois a base de regras é projetada de acordo com o conhecimento local; e (iii) *a propriedade de convergência*: quando um sistema nebuloso não satisfaz a uma precisão de aproximação definida, uma idéia intuitiva, para modificar o sistema e melhorar o desempenho, é dividir o espaço de entrada em mais regiões (modificação da granularidade).

A identificação de sistemas através de modelos nebulosos de TSK é realizada utilizando-se os dados de entrada(s) e/ou saída(s) do sistema dinâmico em estudo como entradas do sistema nebuloso para obtenção de uma previsão da(s) saída(s) atual(is) do sistema dinâmico. O procedimento de identificação, neste caso, é composto de duas partes: (i) identificação da estrutura, e (ii) identificação dos parâmetros. Os parâmetros do conseqüente são coeficientes de equações lineares. Na figura 2, os passos para identificação de um modelo nebuloso do tipo TSK são apresentados.

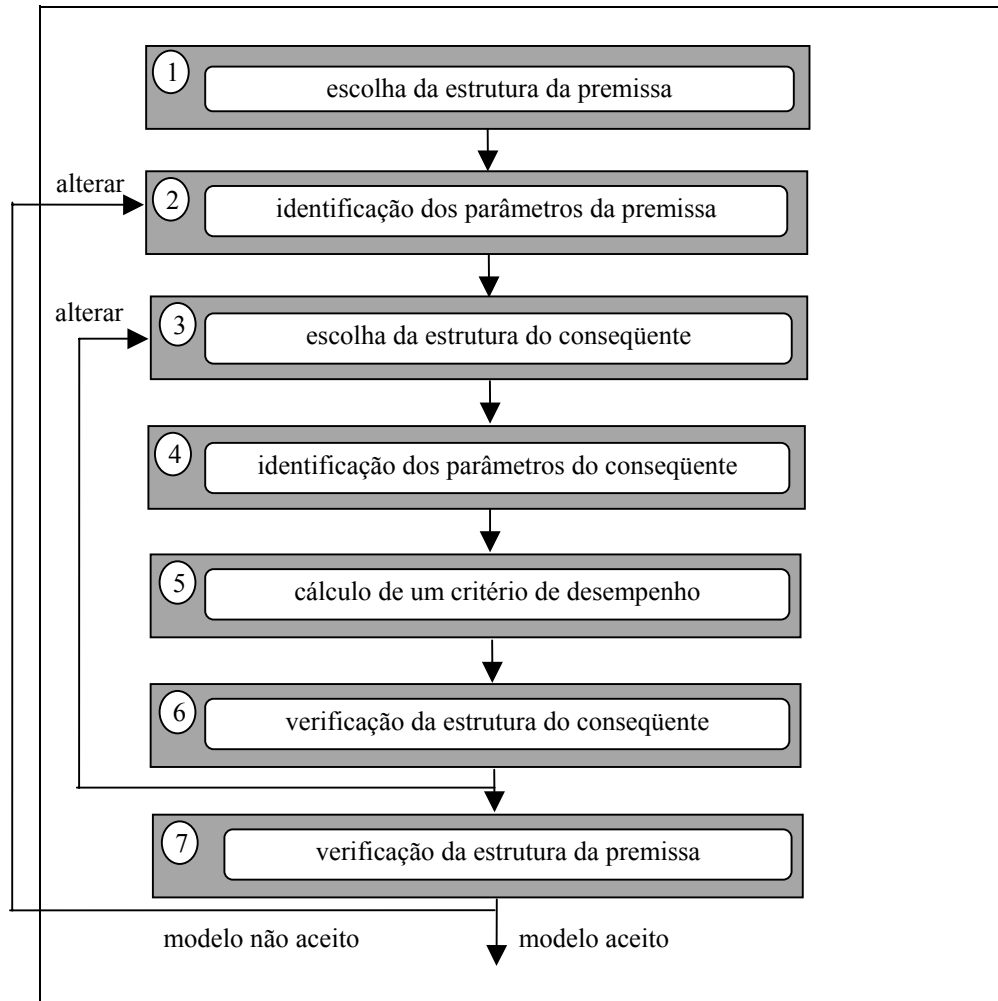


Figura 2. Passos para identificação usando um sistema nebuloso do tipo *TSK*.

Na identificação da estrutura do sistema nebuloso é necessário identificar a estrutura do antecedente e do conseqüente. Os parâmetros do conseqüente são coeficientes de equações lineares, sendo estes parâmetros obtidos pelo algoritmo dos mínimos quadrados. Esta forma de desenvolvimento de sistema nebuloso é inspirada na teoria clássica de sistemas e, alguns desenvolvimentos no campo das redes neurais. Os modelos nebulosos de *TSK* consistem de regras de produção IF-THEN (SE <condição> ENTÃO <ação>) que podem ser representadas na forma geral:

$$R^{(j)} : \text{IF } z_1 \text{ IS } A_1^j \text{ AND } \dots \text{ AND } z_m \text{ THEN} \quad (1)$$

$$g_j = w_0^j + w_1^j u_1^j + \dots + w_{qj}^j u_{qj}^j$$

onde o antecedente *IF* define a parte antecedente (premissa) enquanto as funções da regra *THEN* constituem-se na parte conseqüente do sistema nebuloso;  $\underline{z} = [z_1, \dots, z_m]^T$  é o vetor de entradas da parte premissa; e  $A_i^j$  são labels dos conjuntos nebulosos;  $\underline{u} = [u_1^j, \dots, u_{qj}^j]^T$  representa o vetor de entradas para parte conseqüente de  $R^{(j)}$  que é constituída de  $q_j$  termos;  $g_j = g_j(\underline{u}^j)$  denota a  $j$ -ésima regra de saída que é um polinômio linear dos termos de entrada

do conseqüente  $u_i^j$ ; e  $\underline{w}^j = [w_0^j, w_1^j, \dots, w_{qj}^j]^T$  são os coeficientes polinomiais que formam o conjunto de parâmetros dos conseqüentes. A cada label lingüístico  $A_i^j$  é associada com a função de pertinência; enquanto  $\mu_{A_i^j}(z_i)$  é descrita por

$$\mu_{A_i^j}(z_i) = \exp \left[ -\frac{1}{2} \frac{(z_i - m_{ij})^2}{\sigma_{ij}^2} \right] \quad (2)$$

onde  $m_{ij}$  e  $\sigma_{ij}$  são o valor médio e o desvio padrão da função de pertinência do tipo Gaussiana, respectivamente. Neste caso, para cada função de pertinência selecionada para uma entrada do sistema nebuloso existem 3 parâmetros para serem otimizados pela estratégia evolutiva. A união de todos estes parâmetros define o conjunto de parâmetros da premissa. O poder da regra  $R^{(j)}$  representa seu grau de excitação e é regido pela equação

$$\mu_j(\underline{z}) = \mu_{A_1^j}(z_1) \cdot \mu_{A_2^j}(z_2) \cdots \mu_{A_m^j}(z_m) \quad (3)$$

Os conjuntos nebulosos que dizem respeito a uma regra formam uma região nebulosa (cluster) dentro do espaço da premissa,  $A_1^j \times A_2^j \cdots \times A_m^j$ , com uma distribuição de pertinência descrita pela equação (3). Dado os vetores de entrada  $\underline{z}$  e  $\underline{u}^j$ ,  $j = 1, \dots, M$ , a saída final do sistema nebuloso é inferido pela média ponderada das saídas locais  $g_j(\underline{u}^j)$ , onde

$$y = \sum_{j=1}^M v_j(\underline{z}) \cdot g_j(\underline{u}^j) \quad (4)$$

onde  $M$  denota o número de regras e  $v_j(\underline{z})$  é a intensidade do poder normalizado de  $R^{(j)}$  que é definida por

$$v_j(\underline{z}) = \frac{\mu_j(\underline{z})}{\sum_{j=1}^M \mu_j(\underline{z})} \quad (5)$$

Um exemplo, na figura 3, ilustra o sistema de inferência do tipo *TSK* com duas regras.

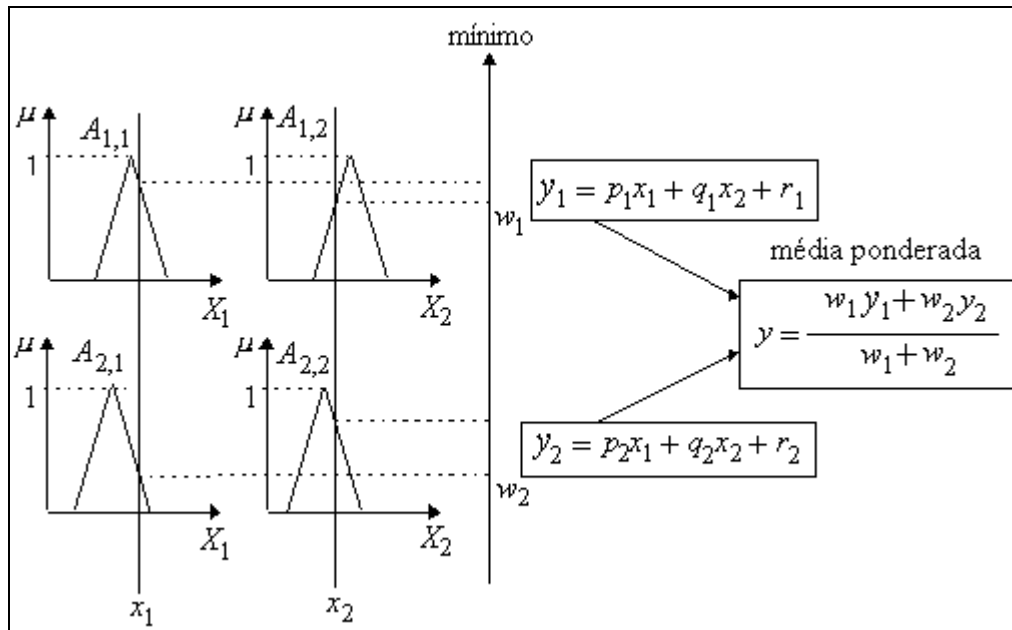


Figura 3. Exemplo do sistema nebuloso de TSK.

## 2.2 Otimização do sistema nebuloso de TSK através de uma estratégia evolutiva

As EEs, inicialmente, foram desenvolvidas para a resolução de problemas de otimização em engenharia. A primeira EE desenvolvida foi a EE-(1+1), proposta por I. Rechenberg e H. P. Schwefel, nos anos 60, no *Hermann Föttinger Institute for Hydrodynamics* (Universidade Técnica de Berlin, Alemanha), em experimentos com um processo túnel de vento. A EE-(1+1) original utiliza somente o operador de mutação, onde apenas um ancestral produz um único descendente.

A EE-(1+1) foi progressivamente generalizada em variantes do número de ancestrais (pais),  $\mu > 1$ , e número de descendentes (filhos),  $\lambda > 1$ , por geração. As EEs com múltiplos membros têm o embasamento biológico relacionado às características de poligenia e pleiotropia. Estas EEs são divididas de acordo com o mecanismo de seleção em:

- (i) estratégia soma (*plus strategy*) —EE-( $\mu+\lambda$ ) —Os  $\mu$  ancestrais geram  $\lambda$  descendentes. Após os  $\mu$  ancestrais e os  $\lambda$  descendentes competem pela sobrevivência;
- (ii) estratégia vírgula (*comma strategy*) —EE-( $\mu,\lambda$ ) —Os  $\lambda$  descendentes competem para sobreviver e os ancestrais são completamente substituídos a cada geração.

O desempenho de uma EE comparado com o de outros algoritmos evolutivos (AEs) depende da configuração adequada dos seus parâmetros internos de controle. As EEs apresentam facilidades no ajuste de tais parâmetros através da utilização de procedimento de auto-adaptação. Enquanto, nos algoritmos genéticos, outro AE, os parâmetros de controle são ajustados, usualmente, através de métodos heurísticos de tentativa e erro.

O princípio da auto-adaptação é facilitar o controle implícito dos parâmetros da EE pela incorporação deste princípio na representação do indivíduo com a evolução usual das variáveis-objeto (possíveis soluções do problema). O termo denominado parâmetros da estratégia (ou parâmetros de controle) referem-se aos parâmetros que controlam o procedimento de busca evolutiva, tais como: taxa de mutação, variância da mutação e taxa de recombinação de um AE (Bäck et al., 1997).



Muitas das pesquisas relacionadas aos princípios de auto-adaptação em AEs tratam de parâmetros relacionados com operador de mutação. A técnica de auto-adaptação é geralmente empregada com sucesso nos ajustes de variâncias e de covariâncias em relação a uma distribuição normal n-dimensional.

Angeline (1995) afirma que é possível adaptar-se dinamicamente os aspectos de processamento de um AE antecipando as regularidades do ambiente, aprimorando o procedimento de otimização e enfatizando a rapidez na busca dos parâmetros. Os AEs que apresentam mecanismos adaptativos (AEMAs) distinguem-se pela configuração dinâmica dos parâmetros selecionados ou mesmo operadores durante o ciclo evolutivo de otimização. Os AEMAs têm uma vantagem sobre os AEs básicos, pois são mais reativos em antecipar as particularidades do problema, ou mesmo em algumas formulações podem dinamicamente adquirir informação sobre as regularidades no problema e explorá-las.

Segundo Angeline (1995), os AEMAs podem ser separados em três níveis onde os parâmetros adaptativos estão presentes, que são:

- (i) *nível populacional*: os métodos adaptativos ajustam dinamicamente os parâmetros, que são globais à população inteira;
- (ii) *nível individual*: os métodos adaptativos modificam a maneira que um indivíduo da população é afetado pelos operadores de mutação;
- (iii) *nível de componente*: os métodos adaptativos alteram a forma pela qual os componentes de cada indivíduo são manipulados independentemente dos outros indivíduos da população.

Os mecanismos de auto-adaptação, no âmbito de componente dos parâmetros da estratégia adaptativa, providenciam uma das características principais do sucesso das EEs. As EEs utilizam princípios de busca no espaço de variáveis-objeto e estratégia interna de controle dos parâmetros, simultaneamente (Beyer, 1995). A abordagem de EEs com adaptação do sistema independente de coordenadas para o operador de mutação, tratada neste artigo, foi proposta por Hansen & Ostermeier, (1996) e Ostermeier & Hansen (1999).

O operador de mutação é o operador principal de uma EE e sem a mudança na distribuição do operador de mutação durante a sequência de gerações do ciclo evolutivo, existe uma diminuição na probabilidade da solução evoluir para uma solução adequada.

O caminho de evolução —“caminho” de distribuição da população no espaço de busca ao longo de um número de gerações —revela informações do ciclo evolutivo, principalmente pelas correlações entre os passos de mutação que são sucessivamente selecionados na sequência de gerações. Se os passos de mutação selecionados são correlacionados paralelamente (produto escalar maior que zero), ou seja, os passos de evolução estão na mesma direção, o caminho de evolução é comparativamente longo. Se, por outro lado, os passos de mutação são correlacionados de forma anti-paralela (produto escalar menor que zero), o caminho de evolução é comparativamente mais curto. Conseqüentemente, para realizar-se passos de mutação mais eficientes é melhor não se possuir uma correlação entre os passos de mutação selecionados no caminho de evolução (Hansen & Ostermeier, 1996).

Hansen & Ostermeier (1996) sugerem o princípio da adaptação fundamental para remover a correlação entre os passos de mutação selecionados sucessivamente, que diz: “Uma adaptação aceitável necessita reduzir a diferença entre as distribuições do caminho de evolução atual e

um caminho de evolução, através de uma seleção aleatória com relação aos parâmetros adaptados”.

Uma abordagem baseada no princípio da adaptação fundamental é adotada neste artigo. A abordagem é denominada de adaptação da matriz de covariância e é aplicada a uma  $(\mu, \lambda)$ -EE, onde os  $\lambda$  descendentes competem para sobreviver e os  $\mu$  ancestral(is) é(são) completamente substituído(s) a cada geração. A seguir são apresentadas equações que regem esta abordagem de  $(\mu, \lambda)$ -EE para  $\mu=1$ . As equações são apresentadas de forma detalhada em Hansen & Ostermeier (1996). O passo de mutação para o vetor variável objeto (vetor solução),  $x$ , para cada descendente  $k = 1, \dots, \lambda$  é dado pela equação,

$$x^{Dk} = x^A + \delta^A B^A z_k \quad (6)$$

onde  $x = (x_1, \dots, x_n)^t \in \mathfrak{R}^n$  é o vetor variável-objeto (variáveis que são as soluções para o problema de otimização) a ser otimizado,  $n$  é a dimensão do problema;  $A$  é o índice para o Ancestral;  $D_k$  é o índice para o Descendente  $k = 1, \dots, \lambda$ ;  $z = (z_1, \dots, z_n)^t : z_i$  são variáveis independentes com distribuição normal,  $N(0,1)$ ,  $z_k \in \mathfrak{R}^n$  ( $k = 1, \dots, \lambda$ ) são realizações independentes de  $z$ ;  $\delta > 0$  é o tamanho de passo global;  $B \in \mathfrak{R}^{n \times n}$  é a matriz que transforma linearmente  $z$  (a matriz  $B$  pode ser vista como a base em que a distribuição normal opera).

A adaptação da distribuição da mutação é separada em duas etapas. Primeiro, a matriz de covariância é adaptada, depois toda a variância é adaptada. A soma acumulada em  $s$  e a adaptação da matriz covariância,  $C$ , são dadas pelas equações:

$$s^{Dk} = (1 - c) \cdot s^A + c_u B^A z_k \quad (7)$$

$$C^{Dk} = (1 - c_{\text{cov}}) \cdot C^A + c_{\text{cov}} \cdot s^{Dk} (s^{Dk})^T \quad (8)$$

onde  $c \in ]0,1[$  determina o tempo acumulado para  $s$ ;  $c_u = \sqrt{c \cdot (2 - c)}$  normaliza a variância de  $s$  para resolver a equação  $1^2 = (1 - c)^2 + c_u^2$ ;  $s^{\text{inicio}} = 0$ ;  $C \in ]0,1[$  é a matriz de covariância da distribuição de mutação. O valor de  $C$  determina  $B$ , ainda que  $Bz \sim N(0, C)$  seja mantido, ou seja,  $C = BB^T$ ; e  $c_{\text{cov}} \in ]0,1[$  determina o tempo da média das distribuições  $ss^T$  sobre uma sequência de gerações.

A equação (8) é bastante similar a regra de atualização de métodos quase-Newton utilizados em otimização clássica. Em ambos os casos, a estimação de segunda ordem da topologia do problema é realizada. O tamanho de passo global,  $\delta$ , é adaptado de uma forma similar ao proposto em Ostermeier et al. (1994) e é denominada de “adaptação do tamanho de passo acumulado”. Neste caso, a soma dos tamanhos de mutação e adaptação de  $\delta$  são regidos pelas seguintes equações:

$$s_{\delta}^{Dk} = (1 - c) \cdot s_{\delta}^A + c_u B_{\delta}^A z_k \quad (9)$$

$$\delta^{Dk} = \delta^A \cdot \exp \left\{ \beta \cdot \left( \|s_{\delta}^{Dk}\| - \hat{\chi}_n \right) \right\} \quad (10)$$

onde  $B_{\delta}^A \in \Re^{n \times n}$  é igual a  $B^A$  com colunas normalizadas;  $c \in ]0,1]$  determina o tempo acumulado para  $s$ ;  $c_u = \sqrt{c \cdot (2 - c)}$  normaliza a variância de  $s$  para resolver a equação  $1^2 = (1 - c)^2 + c_u^2$ ;  $s_{\delta}^{início} = 0$ ;  $\beta$  é um parâmetro para diminuição da variação do tamanho de passo entre as sucessivas gerações;  $\hat{\chi}_n = \sqrt{n} \left( 1 - \frac{1}{4n} + \frac{1}{21n^2} \right)$  aproxima a expectativa da distribuição de  $\chi_n$ , em que é a distribuição do comprimento de uma vetor aleatório com distribuição  $N(0, I)$  em  $\Re^n$ .

A equação (9) é quase igual a equação (7). A diferença é a normalização das colunas em  $B_{\delta}$  que é importante para derivar o comprimento esperado de  $s_{\delta}$ . A equação (10) adapta  $\delta$  para reduzir a diferença entre o comprimento do caminho de evolução “normalizado” atual  $\|s_{\delta}\|$  e seu comprimento esperado. Os parâmetros de projeto adotados no procedimento de otimização do sistema nebuloso de TSK são os recomendados por Hansen & Ostermeier (1996), ou seja,  $c = 1/\sqrt{n}$ ,  $\beta = 1/n$  e  $c_{cov} = 2/n^2$ . A parte conseqüente do sistema nebuloso é realizada pelo algoritmo dos mínimos quadrados em batelada. Os detalhes deste algoritmo são apresentados em Box et al. (1994).

### 3 DESCRIÇÃO DO ESTUDO DE CASO

O *Índice da Bolsa de Valores de São Paulo* (Ibovespa) constitui-se no indicador mais importante de desempenho médio do mercado de ações brasileiro. A carteira teórica do Ibovespa é composta pelas ações que atenderam cumulativamente aos seguintes critérios, com relação aos doze meses anteriores à formação da carteira: (i) estar incluída em uma relação de ações cujos índices de negociabilidade somados representem 80% do valor acumulado de todos os índices individuais; (ii) apresentar participação, em termos de volume, superior a 0,1% do total e (iii) ter sido negociada em mais de 80% do total de pregões do período.

A participação de cada ação na carteira tem relação direta com a representatividade desse título no mercado à vista – em termos de número de negócios e volume financeiro – ajustada ao tamanho da amostra. Essa representatividade é obtida pelo índice de negociabilidade da ação, calculado pela relação  $IN = \sqrt{N} \cdot V$  onde  $N$  é o número de negócios da ação em relação ao número total de negócios do mercado e  $V$  é o volume financeiro negociado da ação em relação ao volume financeiro do mercado.

O valor do índice no período  $t$  seria o somatório dos pesos (quantidade teórica da ação multiplicada pelo seu preço no período  $t$ ) das ações integrantes de sua carteira teórica, ou seja:

$$Ibovespa_t = \sum_{i=1}^{na} P_{i_t} * Q_{i_t} \quad (11)$$

onde  $na$  é o número de ações que compõem a carteira teórica do índice,  $P_{i_t}$  é o preço da ação  $i$  no período  $t$  e  $Q_{i_t}$  é a quantidade teórica da ação  $i$  do período  $t$ .

Uma série temporal pode ser definida como uma função de uma variável independente (período  $t$ ), vinculada a um sistema dinâmico em que uma descrição matemática é desconhecida (ou considerada como tal). A característica principal de tais séries é que o seu comportamento futuro não pode ser previsto exatamente, como pode ser previsto de uma função determinística conhecida em um período  $t$ . Contudo, o comportamento de uma série temporal, no caso o Ibovespa, pode algumas vezes ser antecipado através de procedimentos estocásticos. Nas simulações realizadas adota-se 1254 amostras para aprendizado (fase de estimação) do sistema nebuloso-evolutivo (análise da capacidade de interpolação e do desempenho do algoritmo baseado em EE) e outras 1255 amostras para validação do sistema nebuloso (análise da capacidade de generalização) obtido para o Ibovespa (os dados usados são os obtidos a partir do ano 1995).

#### 4 APLICAÇÕES E RESULTADOS

O procedimento de análise dos resultados obtidos nos problemas de identificação não-linear é dividido nas seguintes etapas: (i) obtenção dos dados do sistema dinâmico (dados de entrada e saída do estudo de caso), (ii) escolha da estrutura utilizada para representar o modelo (modelo nebuloso de TSK com configuração série-paralelo), (iii) determinação de um procedimento de otimização adequado para o tratamento do modelo do processo (usando uma estratégia evolutiva e algoritmo dos mínimos quadrados), (iv) estimação dos parâmetros do modelo matemático (fase de estimação), e (v) a validação do modelo matemático (fase de validação ou de testes).

O critério de desempenho avaliado é a soma do erro quadrático, ISE, entre a saída real do Ibovespa,  $y(t)$ , e saída estimada,  $\hat{y}(t)$ , pelo sistema nebuloso-evolutivo dividido pelo número de amostras (estimação: 1254 e validação: 1255). Os resultados obtidos para o modelo nebuloso de TSK com otimização através da estratégia evolutiva e mínimos quadrados são resumidos na tabela 1, respectivamente, usando-se  $\mu=1$  e  $\lambda=50$  e número de gerações da EE igual a 500.

Tabela 1: Resultados obtidos para um número diferente de funções de pertinência (fps) Gaussianas para cada entrada do sistema nebuloso através do sistema nebuloso-evolutivo para o Ibovespa (entradas:  $y(t-1)$  e  $y(t)$  e saída:  $y(t+1)$ )

fps	fase de estimação (aprendizado)				fase de validação (teste)			
	ISE	erro máximo	erro médio	desvio padrão do erro	ISE	erro máximo	erro médio	desvio padrão do erro
2	0,0270	-0,0015	0,1643	1,1045	0,4156	0,2953	0,5739	2,3692
<b>3</b>	0,0257	-0,0014	0,1608	1,0782	<b>0,1359</b>	<b>0,1259</b>	<b>0,3458</b>	<b>1,7138</b>
4	0,0257	-0,0028	0,1603	1,0731	0,2269	0,1978	0,4340	1,8877
5	0,0252	-0,0014	0,1589	1,1476	0,3134	0,2497	0,5017	2,1023
6	0,0239	-0,0014	0,1546	0,9673	0,6762	0,3668	0,7370	3,6750

7	0,0240	<b>-0,0013</b>	0,1550	1,0447	0,3172	0,2553	0,5027	2,0966
8	<b>0,0227</b>	-0,0022	<b>0,1507</b>	<b>0,8842</b>	0,6993	0,3755	0,7482	3,4384
9	0,0234	-0,0018	0,1530	0,8759	0,7646	0,3899	0,7838	3,3435

Os resultados para o sistema nebuloso de TSK com a estratégia evolutiva e método dos mínimos quadrados foram precisos para identificação do estudo de caso (ver tabela 1). Nas figuras 4 e 5 são apresentados os melhores resultados na fase de estimação (com 3 funções de pertinência para cada variável de entrada do sistema nebuloso) e na fase de validação (com 8 funções de pertinência para cada variável de entrada do sistema nebuloso) obtidos com o sistema nebuloso-evolutivo para o Ibovespa, respectivamente.

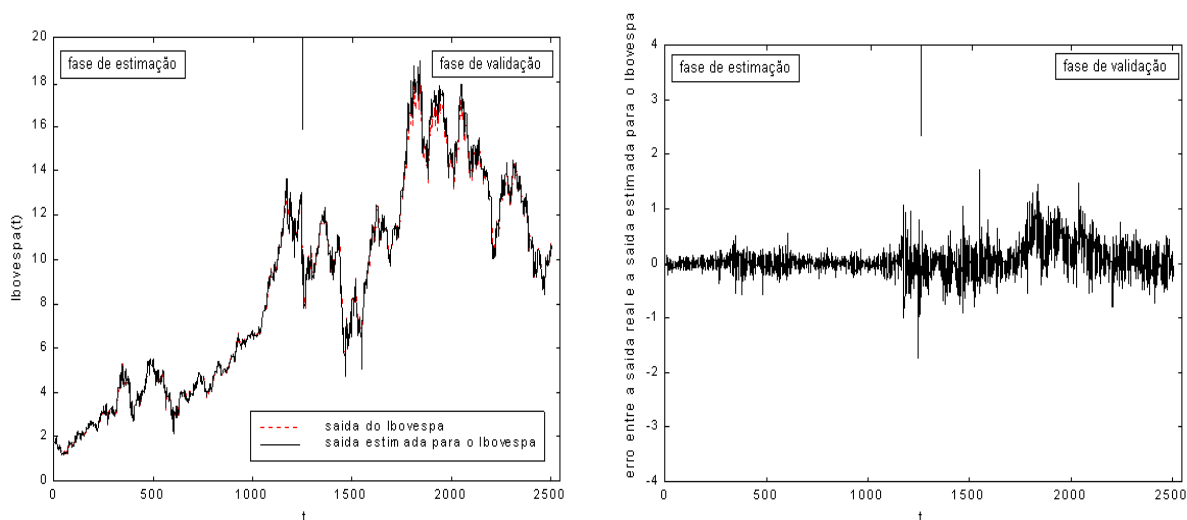


Figura 4. Resultado da previsão do Ibovespa com 3 funções de pertinência para cada variável de entrada..

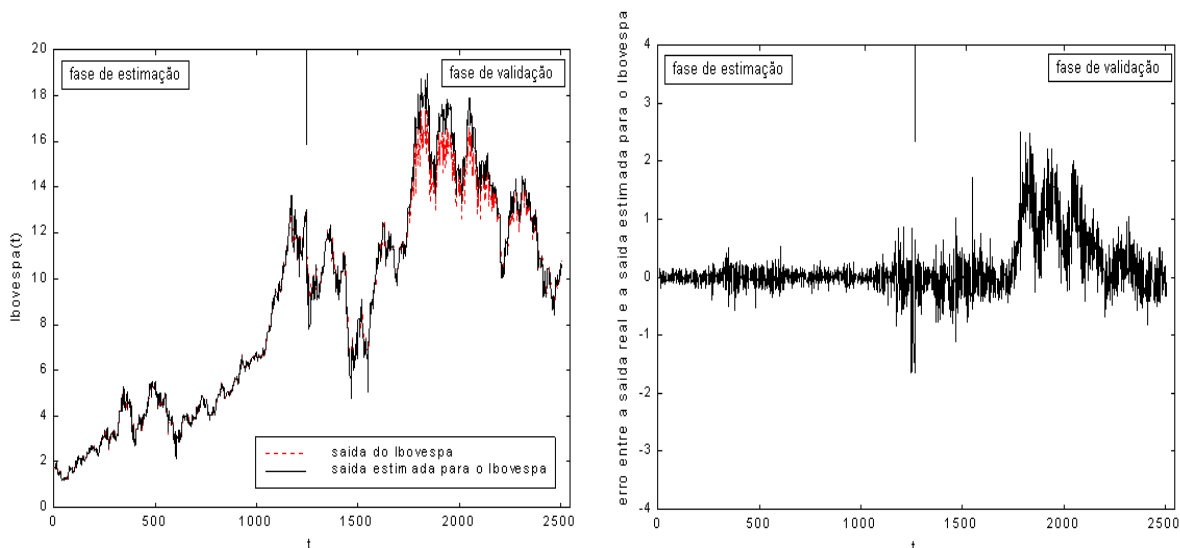


Figura 5. Resultado da previsão do Ibovespa com 8 funções de pertinência para cada variável de entrada.

Os resultados apresentados na tabela 1 mostraram que quanto maior a precisão (sobre treinamento) obtida na fase de estimação menor é a precisão obtida na fase de validação. Em outras palavras, tem-se um compromisso entre precisão e generalização, pois quanto mais preciso o modelo matemático menos genérico ele se torna.

Neste contexto, quando um sistema nebuloso aprende muitas relações de entrada e saída ou o número de parâmetros do sistema nebuloso é grande (número maior de graus de liberdade), o sistema nebuloso pode memorizar (interpolar) os dados da fase de estimação, porém pode ser menos apta a generalizar. Uma forma de evitar este problema de sobre-treinamento é utilizar um conjunto de dados de treinamento amplo, que abranja toda a região de interesse e limitar o número de parâmetros, tão estritamente quanto possível. Neste caso, a utilização de uma quantidade inadequada de dados de estimação levará a um problema de estimação de parâmetros mal-condicionado.

Outros questionamentos podem ser realizados quanto a flexibilidade dos sistemas nebulosos do tipo TSK. A solução do problema de seleção estrutural que está presente em um sistema nebuloso de TSK depende dos valores atribuídos para os centros da função de base, em termos de quantidade e posição. Neste caso, a estratégia evolutiva com operador de mutação adaptativo mostrou-se promissora, apresentando uma minimização do critério de desempenho ISE adequada para o Ibovespa.

## 5 CONCLUSÃO

Os sistemas nebulosos são usualmente denominados de estimadores livres de modelo, pois os sistemas nebulosos estimam relações de entrada e saída sem a necessidade de um modelo analítico de como as entradas dependem das saídas. Além disso, codificam a informação amostrada em uma estrutura distribuída de forma paralela denominada de estrutura nebulosa. Existem diversas abordagens de sistemas nebulosos utilizados para propósitos de identificação e previsão, tais como os modelos nebulosos relacionais, modelos nebulosos baseados em regras e modelos nebulosos interpolativos.

Neste artigo é apresentada a aplicação de sistemas nebulosos do tipo TSK com otimização evolutiva à identificação do comportamento dinâmico do Ibovespa um passo a frente (previsão um período de tempo a frente).

O sistema nebuloso de TSK apresentou resultados precisos para a previsão de séries temporais, devido a sua capacidade de aproximar funções não-lineares e eficiência do aprendizado. Entretanto, necessita-se de estudos mais aprofundados em trabalhos futuros quanto aos aspectos de aprimoramento das potencialidades do sistema nebuloso de TSK. Neste aspecto deseja-se utilizar procedimentos de otimização com múltiplos objetivos para obtenção de um melhor compromisso entre interpolação, generalização e aprendizado dos sistemas nebulosos visando-se aplicações de identificação multivariável, com comportamento caótico e análise para horizontes de previsão com  $n$  passos de tempo a frente.

Outra abordagem a ser estudada futuramente é quanto ao desempenho do algoritmo de estimação de parâmetros do sistema nebuloso. Este pode ser aprimorado pela adição de um termo de regularização (Sjöberg, 1995) na função custo a ser otimizada (neste artigo, utilizou-se a minimização do ISE). Esta função tem um valor grande para mapeamentos “suaves” (modelos simples) e pequeno para outros (modelos complexos). A regularização prioriza a obtenção de modelos menos complexos, se utilizada para melhorar a capacidade de generalização dos sistemas nebulosos. Uma possibilidade é adicionar um termo, que consiste do erro médio quadrático dos parâmetros da sistemas nebulosos na função objetivo original. A regularização é uma característica necessária para manter a variância do erro pequena em

modelos que empregam muitos parâmetros. Neste caso, a resposta do sistema nebuloso de TSK será menos sensível ao sobre-treinamento.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Angeline, P. J. (1995). Adaptive and self-adaptive evolutionary computations, *Computational intelligence: a dynamic systems perspective*, Palniswami, M.; Attikiouzel, Y.; Marks, R.; D. Fogel & T. Fukuda (eds.), Piscataway, NJ: IEEE Press, pp. 152-163.
- Bäck, T., Fogel, D. B.; Michalewicz, Z. (eds.) (1997). *Handbook of evolutionary computation*, Bristol, Philadelphia: Institute of Physics Publishing, NY, Oxford: Oxford University Press.
- Beyer, H. G. (1995). Toward a theory of evolution strategies: self-adaptation, *Evolutionary Computation*, vol. 3, no. 3, pp. 311-348.
- Billings, S. A.; Coca, D. (1999). Discrete wavelet models for identification and qualitative analysis of chaotic systems, *International Journal of Bifurcation and Chaos*, vol. 9, no. 7, pp. 1263-1284.
- Box, G., Jenkins, G.; Reinsel, G. (1994). *Time series analysis, forecasting and control*, 3rd ed., Englewood Cliffs, NJ: Prentice Hall.
- Branco, P. J. C.; Dente, J. A. (1997). Fuzzy systems modelling in practice, *Proceedings of the 7th International Fuzzy Systems Association World Congress, IFSA'97*, Prague, Czech Republic, vol. II, pp. 360-365.
- Chen, X.; Racine, J.; Swanson, N. R. (2001). Semiparametric ARX neural-network models with an application to forecasting inflation, *IEEE Transactions on Neural Networks*, vol. 12, no. 4, pp. 674-683.
- Cordón, O.; Herrera, F.; Lozano, M. (1995). A classified review on the combination fuzzy logic-genetic algorithms bibliography, *Technical Report DECSAI-95129*, Department of Computer Science and A.I., University of Granada.
- Gilbert, J. (1995). Artificial intelligence on Wall Street: an overview and critique of applications in the finance industry, <http://gryphon.ccs.bradeis/~grath/brandeis/ai-paper>.
- Goonatilake, S.; Khebbal, S. (1995). *Intelligent hybrid systems*. Chichester: John Wiley & Sons, USA.
- Haber, R.; Unbehauen, H. (1990). Structure identification of nonlinear dynamic systems —a survey on input/output approaches, *Automatica*, vol. 26, no. 4, pp. 651-677.
- Hansen, N.; Ostermeier, A. (1996). Adapting arbitrary normal mutation distributions in evolution strategies, *IEEE International Conference on Evolutionary Computation*, Nagoya, Japan, pp. 312-317.
- León, T.; Liern, V.; Vercher, E. (2002). Viability of infeasible portfolio selection problems: a fuzzy approach, *European Journal of Operational Research*, vol. 139, pp. 178-189.
- Ljung, L. (2001). Black-box models from input-output measurements, *Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference*, vol. 1, pp. 138-146.
- Mamdani, E.; Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller, *International Journal of Man-Machine Studies*, vol. 7, no. 1, pp. 1-13.
- Navarro, J. L.; Albertos, P. (1996). Fuzzy logic implementation of industrial controllers. 13<sup>th</sup> IFAC World Congress, San Francisco, CA, USA, pp. 409-414.
- Ostermeier, A.; Gawelczyk, A.; Hansen, N. (1994). Step-size adaptation based on non-local use of selection information, *Proceedings of Parallel Problem Solving from Nature*, PPSN III, Berlin: Springer, Jerusalem, Israel, pp. 189-198.

- Ostermeier, A.; Hansen, N. (1999). An evolution strategy with coordinate system invariant adaptation of arbitrary normal mutation distributions within the concept of mutative strategy parameter control, *Genetic and Evolutionary Computation Conference, GECCO*, Orlando, FL, pp. 902-909.
- Schweizer, J.; Schimming, T. (2001). Symbolic dynamics for processing chaotic signal, II. communication and coding. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 48, no. 11, pp. 1283-1295.
- Sjöberg, J. (1995). Non-linear system identification with neural networks, *PhD thesis*, Department of Electrical Engineering, Linköping University, Linköping, Sweden.
- Sugeno, M.; Kang, G. T. (1988). Structure identification of fuzzy model, *Fuzzy Sets and Systems*, vol. 28, pp. 15-33.
- Takagi, T.; Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control, *IEEE Transactions on Systems, Man and Cybernetics*, vol. 15, no. 1, pp. 116-132.
- Tseng, F.-M.; Tzeng, G.-H. (2002). A fuzzy seasonal ARIMA model for forecasting, *Fuzzy Sets and Systems*, vol. 126, pp. 367-376.
- Van Gestel, T.; Suykens, J. A. K.; Baestaens, D.-E.; Lambrechts, G.; Vandaele, B.; De Moor, B.; Vandewalle, J. (2001). Financial times series prediction using least squares support vector machines within the evidence framework, *IEEE Transactions on Neural Networks*, vol. 12, no. 4, pp. 809-821.
- Varetto, F. (1998). Genetic algorithm applications in the analysis of insolvency risk, *Journal of Banking & Finance*, vol. 22, pp. 1421-1439.
- Von Altrock, C. (2002). Applying fuzzy logic to business and finance, *Optimus*, no. 2, pp. 38-39.
- Wang, L. -X. (1994). *Adaptive fuzzy systems and control, design and stability analysis*. Prentice-Hall, Englewood Cliffs, NJ, USA.
- Yager, R. R.; Filev, D. P. (1994). *Essentials of fuzzy modeling and control*, John Wiley & Sons, New York, NY, USA.
- Ying, H. (1998). General SISO Takagi-Sugeno fuzzy systems with linear rule consequent are universal approximators, *IEEE Transactions on Fuzzy Systems*, vol. 6, no. 4, pp. 582-587.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, vol. 8, pp. 338-353.
- Zadeh, L. A. (1973). Outline of a new approach to the analysis of complex systems and decision processes, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, pp. 28-44.
- Zeng, X. -J.; Singh, M. G. (1995). Approximation theory of fuzzy systems —MIMO case, *IEEE Transactions on Fuzzy Systems*, vol. 3, no. 2, pp. 219-235.