

**INSTITUTO BRASILEIRO DE AVALIAÇÕES E PERÍCIAS DE ENGENHARIA  
XII COBREAP - Congresso Brasileiro de Engenharia de Avaliações e Perícias**

**PREDIÇÃO DE VALORES USANDO MODELOS DE ESPECIALISTAS GLOBAIS,  
LOCAIS E DE COMPOSIÇÃO DE ESPECIALISTAS LOCAIS**

**Melo, Brício<sup>A</sup>; Lopes, Benedito Arruda Ribeiro<sup>B</sup>**

<sup>A</sup>Engenheiro Civil, CREA 39.154/D - MG, IBAPE (DF) n.º 107  
ITA/EAM (Engenharia Aeronáutica e Mecânica)  
Pça Marechal Eduardo Gomes, 50 – Vila das Acácias  
CEP 12.228-900 - São José dos Campos (SP)  
Tel (012)3947-4152 e Fax (012)3947-5815  
bricio@mec.ita.br

<sup>B</sup>Engenheiro Civil, CREA 76.049/D - SP, IBAPE (DF) n.º 25  
Banco do Brasil S/A – Diretoria de Infra-estrutura  
Divisão de Avaliações de Bens e Empreendimentos  
SBS Qd. 4 Bl. A Lt. 25 – Edifício Sede I - 12º andar  
CEP 70.073-100 - Brasília (DF)  
Tel (061)310-2670 e Fax (061)310-4550  
arruda@bb.com.br

**Resumo.** É apresentado o assunto *Composição de Especialistas Locais ou MLEM* (do inglês "Mixture of Local Expert Models") que engloba análise exploratória de dados e modelagem em uma única técnica, útil, por exemplo, em predição de valores. Neste trabalho lidaremos com os seguintes especialistas: *Redes Neurais Artificiais, Análises de Regressão Múltipla e Homogeneização*. A idéia básica é primeiramente particionar o espaço de entrada usando uma Rede Neural de Kohonen, obtendo assim os chamados "mapas auto-organizáveis". Para cada região do espaço de entrada são feitos testes de desempenho entre os diferentes especialistas, denominados *Especialistas Locais*, com o intuito de selecionar o especialista mais adequado. Para comparação, também é apurado o melhor modelo entre os especialistas globais, assim chamados quando o espaço de entrada não é particionado. Neste caso, os diversos modelos obtidos pelos melhores especialistas locais são combinados por uma Rede Supervisora (do inglês "Gating Network") que atua como mediadora entre os especialistas considerando a posição do ponto de entrada em relação ao centro de gravidade de cada uma das regiões em que foi dividido o espaço de entrada. Um estudo de caso ilustra o uso da composição de especialistas locais em problemas de predição.

**Palavras-chave:** *Predição, Composição de especialistas, Redes neurais de Kohonen.*

**Abstract.** The issue of this paper is *Mixture of Local Expert Models (MLEM)* which congregates data exploratory analysis and mathematical modeling in the same technique and has many possible applications. We use different modeling techniques to construct prediction's models and show how the output of the different models can be combined. The following modeling techniques are used: *Artificial Neural Networks, Multiple Regression Analysis and Homogenization*. We begin by clustering the points from the entry data set using Kohonen Neural Networks. Then the different modeling techniques are applied to each cluster and tests are conducted in order to select the best model for each cluster which is then named the local expert. All local experts are combined by a Gating Network, which plays the role of a mediator among the local experts by taking into account the distance between the given input point and the center of each cluster. For comparison purposes, we also evaluate the performance of each modeling technique acting as a Global Expert, i.e., when the technique has access to the entire data set. A case study of MLEM is developed to show how this system solves the prediction's problems.

**Keywords:** *Prediction, Mixture of Experts, Kohonen Neural Networks.*

## 1. Introdução

A fase de modelagem matemática de um problema de predição (chamada de “treinamento” na área de redes neurais artificiais) pode ser vista como um problema de otimização, onde é desejado minimizar uma medida do erro entre os valores observados e aqueles fornecidos pelo modelo. Existem várias técnicas para modelagens baseadas em propriedades estatísticas já consolidadas pelo uso. Mais recentemente, as técnicas de redes neurais artificiais têm despertado interesse de pesquisadores. Algumas modelagens são executadas sob forte supervisão e outras com necessidade reduzida de informações providas pelo supervisor, ambas ainda necessitando de validação por critérios estatísticos (Nascimento Jr. & Yoneyama, 2000).

Em geral, sabe-se que nenhuma técnica de modelagem é completa. Algumas técnicas apresentam alta velocidade de convergência, mas podem apresentar falhas de generalização detectadas na validação cruzada. Outras técnicas conseguem boa generalização ao custo de baixa velocidade de convergência.

Neste trabalho é proposto um sistema capaz de combinar os resultados obtidos de vários tipos de especialistas, i. e., várias técnicas matemáticas usadas para predição de valores, onde cada especialista tem atuação localizada. A meta é conseguir, através da composição de diferentes tipos de especialistas, um resultado melhor do que aquele que seria obtido por cada especialista trabalhando de forma isolada.

Neste artigo a modelagem obtida através da composição de especialistas tem como objetivos:

- propor uma arquitetura modular para a tarefa de modelagem onde uma rede supervisora define a área de atuação de cada especialista (particionador do espaço de entrada);
- comparar os resultados obtidos através da composição de especialistas locais com os obtidos pelo uso de especialistas globais operando isoladamente.

A grande maioria de técnicas convencionais para predição utiliza *modelos lineares* os quais, pela sua própria natureza, não são capazes de capturar certos comportamentos *não-lineares* dos dados. A aplicação de redes neurais artificiais à análise dos dados permite o desenvolvimento de modelos *não-lineares* de predição, os quais podem ou não resultar em modelos com melhor desempenho quando comparados com os resultados de técnicas com abordagem *linear* (Makridakis, S. *et al.*, 1998).

A proposta deste artigo é combinar, nas tarefas de predição de valores, técnicas com abordagem *linear* e com abordagem *não-linear*, capturando assim o melhor desempenho de cada técnica. Para tal, este trabalho apresenta uma alternativa simples para a composição de especialistas locais e para a definição da área de atuação de cada especialista.

## 2. Sistema de composição de especialistas

O sistema de composição de especialistas aqui apresentado segue a proposta elaborada por Jacobs *et al.* (1991) no qual a idéia básica é: (a) dividir o espaço de entrada em regiões, (b) designar um especialista para responder em cada região do espaço de entrada, e (c) implementar a composição dos vários especialistas usando uma rede supervisora (do inglês "*Gating Network*") que decide como ponderar as saídas de cada especialista na tarefa de modelagem.

A figura 1 mostra o arranjo dos componentes do sistema proposto para a composição de especialistas.

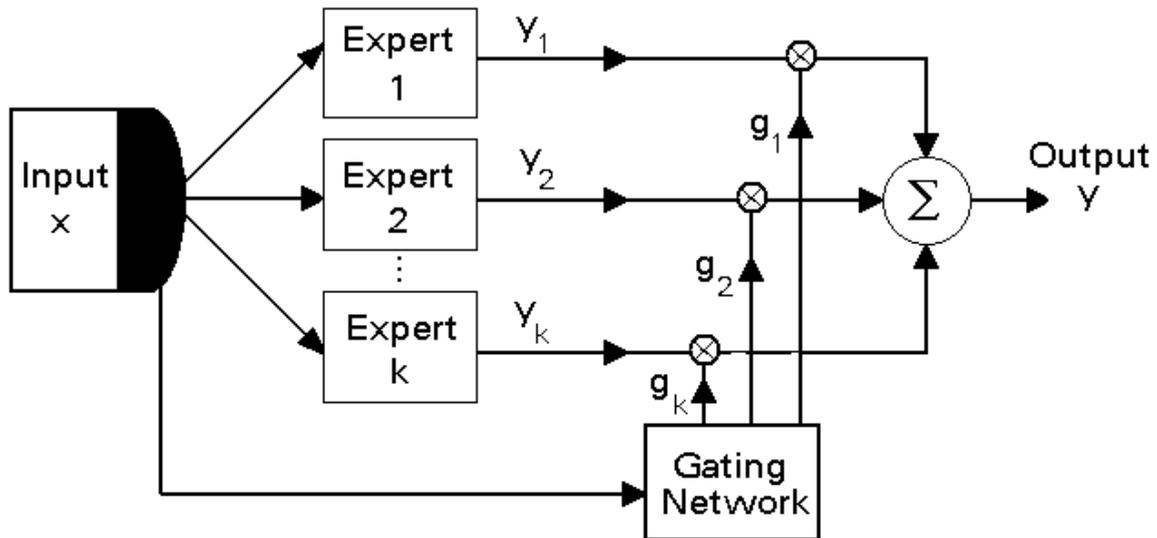


Figura 1 - Sistema de composição de especialistas.

A característica central do sistema proposto é o uso de vários tipos de especialistas, assumindo que diferentes especialistas apresentam melhor desempenho nas diferentes regiões do espaço de entrada.

A etapa inicial de construção do modelo de previsão de valores usando a composição de especialistas locais é implementada da seguinte maneira:

- inicialmente o espaço de entrada ( $X$ ) é particionado em várias regiões ( $X_i$ ),
- em cada região são calibrados os modelos associados a todos os diferentes tipos de especialistas utilizando apenas os dados daquela região,
- para cada região, segundo algum critério específico, é eleito o melhor especialista que passará a ser considerado o único daquela região.

A utilização do sistema MLEM, uma vez eleitos os especialistas locais, é feita da seguinte forma:

- a entrada  $x$  é repassada para o especialista eleito para cada região do espaço de entrada, o qual então calcula a sua saída denotada por  $y_i$ ,
- é calculado o coeficiente de ponderação  $g_i$  em função da distância entre a entrada  $x$  e o centro da região  $i$ ,
- é calculada a saída do modelo de previsão  $y$  como a somatória ponderada das saídas  $y_i$  de cada especialista usando os respectivos coeficientes  $g_i$  como fatores de ponderação, ou seja:

$$y = \sum_{i=1}^k g_i y_i \quad (1)$$

onde  $k$  denota o número de especialistas.

### 3. Particionamento do espaço de entrada e cálculo dos coeficientes de ponderação $g_i$

Uma das funções da rede supervisora é implementar o particionamento do espaço de entrada. Para tal uma rede neural de Kohonen é treinada pelo algoritmo *SOM* (do inglês "*Self-Organizing Map*" ou mapas auto-organizáveis) usando os dados disponíveis para a construção do modelo de predição de valores (Kohonen, 1989).

O treinamento da rede neural de Kohonen tem como objetivo descobrir similaridades nos dados colocados na sua entrada durante o treinamento. A partir deste princípio é possível mostrar que a rede neural de Kohonen implementa o agrupamento ("*clustering*") dos dados de entrada de tal forma que dados de entrada  $x$  que estão próximos (segundo algum critério de distância) no espaço de entrada serão alocados ao mesmo grupo ("*cluster*") (Nascimento Jr. & Yoneyama, 2000; Haykin, 1999).

Define-se então que os dados de cada grupo definem uma região do espaço de entrada. Note que os dados de treinamento do modelo de predição são agrupados de tal forma que cada dado pertence a apenas um dos grupos.

Como a cada grupo de dados corresponde uma única unidade da rede neural de Kohonen (denominada de unidade vencedora ou "bem casada") e o peso desta unidade corresponde ao centro do grupo, então o treinamento da rede neural de Kohonen, além de separar o espaço de entrada em regiões, também fornece o centro de cada região que denotaremos por  $c_i$  (Nascimento Jr. & Yoneyama, 2000).

Os coeficientes de ponderação  $g_i$  são então calculados como a seguir:

$$d_i = \exp\left[-\frac{1}{2}\|x - c_i\|^2\right] \quad (2)$$

$$g_i = \frac{d_i}{\sum_{j=1}^k d_j} \quad (3)$$

### 4. Escolha dos Especialistas

Os especialistas devem ser escolhidos de forma a representar diferentes técnicas de modelagem, como as que dão ênfase à *não-linearidade* (p. ex., as *Redes Neurais Artificiais* denominadas de RNA) e à *linearidade* (como as técnicas do tipo *Análise de Regressão Múltipla - ARM*).

Para o papel de RNA escolhemos o *perceptron* de múltiplas camadas que utiliza o treinamento supervisionado e que, quando aplicado ao problema de predição, pode ser visto como um modelo de regressão não-linear. O *perceptron* de múltiplas camadas foi escolhido devido ao grande número de problemas que pode ser resolvido por ele e a existência de vários algoritmos que utilizam princípios diferentes para o seu treinamento. Dentre estes, o algoritmo selecionado foi o *Back-Propagation* que tem o objetivo de minimizar o erro quadrático médio da saída da rede neural em relação aos pontos observados e utiliza informações do gradiente da função erro quadrático médio (Nascimento Jr. & Yoneyama, 2000).

O especialista ARM, através do método dos mínimos quadrados ordinários, procura uma função linear que melhor se ajusta aos dados. Na modelagem por este especialista, diversas

hipóteses são necessárias para garantir a fundamentação científica do modelo calibrado. Assim, Pindyck & Rubinfeld (1998) e Gujarati (2000) ressaltam a necessidade, quando se usam quaisquer modelos de regressão, de observarem os seus pressupostos básicos, principalmente no que concerne à normalidade, homoscedasticidade, ausência de multicolinearidade e de autocorrelação serial e inexistência de pontos atípicos ("outliers") e influenciadores, com o objetivo de obter previsões não-tendenciosas, eficientes e consistentes.

Outro tipo de especialista escolhido é a homogeneização. Esta técnica fornece como predição de valor a média dos valores amostrais homogeneizados, ou seja:

$$\hat{Y} = \frac{\sum_{i=1}^n Y_i}{n} \quad (4)$$

onde n é igual ao número de dados efetivamente utilizados para predição.

## 5. O uso dos modelos de especialistas globais, locais e de composição de especialistas locais para predição de valor

Apresentamos a seguir um estudo de caso em que o objetivo é a predição de valores usando o modelo de composição dos especialistas, a partir de 505 dados. Da amostra de 505 dados, 15 foram separados para validação dos modelos (subconjunto de validação dos modelos), cada elemento deste subconjunto é equivalente a um bem avaliando, e 490 dados foram usados para construção dos modelos (subconjuntos de treinamento e teste).

Para todos os especialistas, foi usada uma janela *inicial* na entrada de comprimento 12, ou seja, a partir de  $X_1, X_2, \dots, X_{12}$  deseja-se a predição de  $Y$ . Neste caso as 12 variáveis independentes foram escolhidas em função da variável dependente estar, a priori, por hipótese, sendo explicada por estas variáveis.

Os três passos para uma análise exploratória dos dados e obtenção dos modelos iniciaram-se separando aleatoriamente os 490 elementos do banco de dados para construção dos modelos da seguinte forma: 1) o subconjunto de treinamento foi formado por 478 elementos, b) o subconjunto de testes foi formado por 12 elementos. Em seguida, através das redes de Kohonen, estes 490 dados são separados em 3 regiões, sendo 399 dados na região I, 86 dados na região II e 5 na região III. A tabela 1 sumariza a divisão dos dados de construção dos modelos:

Dados para construção dos modelos	Região I	Região II	Região III	Total
Treinamento	387	86	5	478
Teste	12	-		12
Total	399	86	5	490

Tabela 1 - Divisão dos dados de construção dos modelos.

Os dados das regiões I e II são submetidos aos seguintes 3 (três) especialistas locais:

RNA I: *Multi Layer Perceptrons* sem ligação direta *Input - Output* com arquitetura  $[12, N_h, 1]$ , significando 12 unidades na camada de entrada,  $N_h$  unidades na camada escondida e 1 unidade na camada de saída;

RNA II: *Multi Layer Perceptrons* com ligação direta *Input - Output* com arquitetura  $[12, N_h, 1]$ , significando 12 unidades na camada de entrada,  $N_h$  unidades na camada escondida e 1 unidade na camada de saída;

ARM: Neste caso o modelo para prever  $Y$  foi uma combinação linear dos regressores  $X_1, X_2, \dots, X_{12}$  e das transformações inversa e logarítmica de cada um dos regressores.

Os dados da região III são submetidos ao especialista em que é possível, devido à quantidade mínima de dados (5), tratamento compatível com este fato, assim o escolhido para tal é o HMZ com sua especificação conforme descrito na seção 4.

A principal distinção entre os dois tipos de RNA utilizados é a presença ou não da ligação direta entre a camada de entrada e de saída. Segundo Weigend & Gershenfeld (1994), ligações diretas entre cada unidade da camada de entrada e a de saída podem extrair a *linearidade*, caso exista nos dados, já ao início do processo de aprendizagem e liberar sua capacidade *não-linear* para ser empregada onde ela é realmente necessária.

O passo intermediário consiste das seguintes fases:

- Para cada uma das 2 (duas) regiões I e II são usados apenas os dados do subconjunto de treinamento de cada região, assim, foram treinados 3 especialistas (RNA I, RNA II e ARM), o HMZ foi utilizado na região III.
- Determina-se para cada região, qual é o melhor especialista usando os dados do subconjunto de *treinamento*, através do indicador RMSE (do inglês "*Root Mean Squared Error*"):

$$RMSE = \sqrt{\frac{1}{NT} \sum_{t=1}^{NT} (Y_t - \hat{Y}_t)^2} \quad (5)$$

onde  $NT$  = número de pontos do subconjunto de treinamento na região.

- Determina-se para cada região I e II qual é o melhor especialista usando os dados do subconjunto de *testes*, caso exista, usando também o indicador RMSE, como na equação (5), porém substituindo  $NT$  por  $NE$  que é igual ao número de pontos do subconjunto de *testes* na região, procurando com isto obter uma *validação cruzada* dos modelos de especialistas e confirmar ou não o melhor especialista para a predição naquela região.

O passo final consiste na execução da tarefa de predição que é feita usando a composição dos especialistas eleitos conforme explicado nas seções 2 e 3.

Para comparação, foram também determinados os modelos dos especialistas globais, ou seja, as 3 técnicas RNA I, RNA II e ARM foram aplicadas considerando o espaço de entrada como sendo uma única região.

## 6. Análise dos resultados

A composição foi elaborada com os especialistas locais sendo eleitos pelo critério de menor valor do RMSE considerando o subconjunto de dados de *teste* na região. Caso o subconjunto de dados de testes para a região *seja vazio*, então o melhor especialista é eleito pelo critério de menor valor do RMSE considerando o subconjunto de dados de *treinamento* na região.

Considerando os 490 dados para construção dos modelos, a 1ª região tem 387 pontos de treinamento e 12 pontos de teste. O melhor especialista para esta região, pelo critério de menor valor do RMSE no subconjunto de teste, foi o RNA I. Além disso, o RNA I foi também o melhor especialista para a região I pelo critério de menor valor do RMSE no subconjunto de treinamento.

A 2ª região tem 86 pontos de treinamento e nenhum ponto de teste. O melhor especialista para esta região, pelo critério de menor valor do RMSE no subconjunto de treinamento, foi o ARM.

A 3ª região tem 5 pontos de treinamento e nenhum ponto de teste. O especialista, conforme citado na seção 5, foi o HMZ.

O subconjunto de 15 elementos para validação dos modelos, neste caso, equivalentes aos bens avaliados, sob a mediação da rede supervisora (do inglês "*Gating Network*") têm a predição feita pela composição dos especialistas locais eleitos para cada uma das regiões:

$$\hat{Y} = g_{RNA I} \times \hat{Y}^{RNA I} + g_{ARM} \times \hat{Y}^{ARM} + g_{HMZ} \times \hat{Y}^{HMZ} \quad (6)$$

A tabela 2 a seguir mostra os valores de RMSE medidos sobre este subconjunto de validação de modelos para os especialistas globais.

Modelo do Especialista	RMSE	RMSE relativo
Global RNA I	2,22375	1,00000
Global RNA II	1,22406	0,55045
Global ARM	1,30798	0,58819

Tabela 2 – Consolidação de resultados para especialistas globais.

Da tabela 2 podemos observar que o especialista global RNA II apresentou o melhor desempenho entre as 3 técnicas investigadas, caso seja considerado espaço de entrada único.

A tabela 3 mostra os valores de RMSE medidos sobre este mesmo subconjunto de validação de modelos para os especialistas locais eleitos e para o modelo de composição destes especialistas.

Modelo do Especialista	RMSE	RMSE relativo
Local RNA I	1,74665	0,3786
Local ARM	4,61339	1,00000
MLEM(Mixture of Local Experts)	1,02079	0,22127

Tabela 3 – Consolidação de resultados para especialistas locais e de sua composição.

Das tabelas 2 e 3 podemos observar que a composição de especialistas locais apresentou o melhor desempenho de todos os especialistas utilizados.

A figura 2 mostra as predições para 13 pontos do subconjunto de validação obtidas pelos 2 especialistas com treinamento local e pela composição das predições destes dois especialistas.

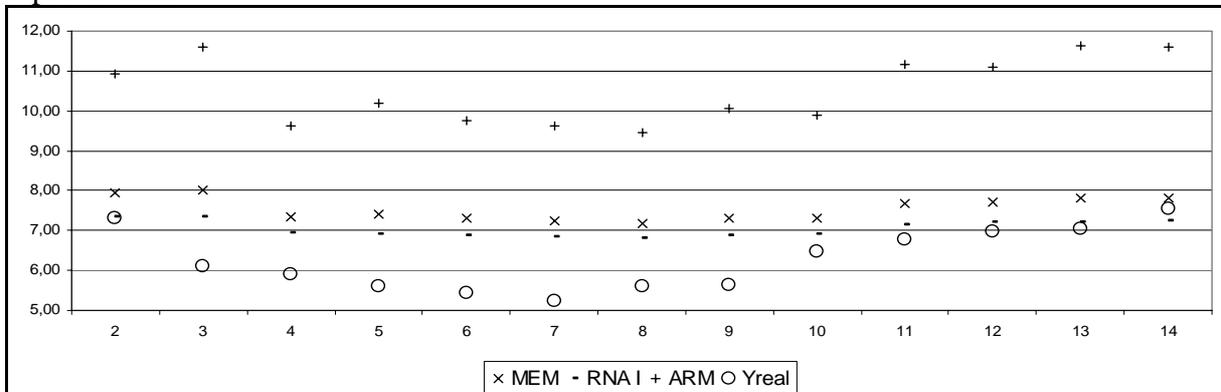


Figura 2 - Predições para o subconjunto de validação obtidas pelos 2 especialistas locais e pela composição destes 2 especialistas.

Pela análise da figura 2, pode-se deduzir que a maioria dos pontos do subconjunto de validação pertencem à região I e também observar que o sistema de composição de especialistas locais, automaticamente, considerou como mais importante a predição do especialista da região I, neste caso o RNA I.

Para melhor ilustração dos elementos avaliados 1 e 15, componentes do subconjunto de validação, têm seus resultados detalhados. Sob a mediação da rede supervisora (do inglês "Gating Network") quando cada elemento, neste caso o primeiro do subconjunto de validação, a exemplo de um único bem avaliando, é apresentado à rede supervisora, é designado o especialista HMZ para sua predição, automaticamente, pois a distância deste  $x$  é muito pequena em relação ao centro  $c_i$  deste cluster, no caso o da região III, assim a predição da composição de especialistas locais e o da homogeneização são bastante próximos. Os resultados de predição para este avaliando são apresentados na figura 3, os de HMZ, de composição dos especialistas locais, além dos obtidos pelos demais especialistas globais, em especial, o ARM que comumente e indiscriminadamente é utilizado sobre uma massa única de dados sem análises exploratórias prévias.

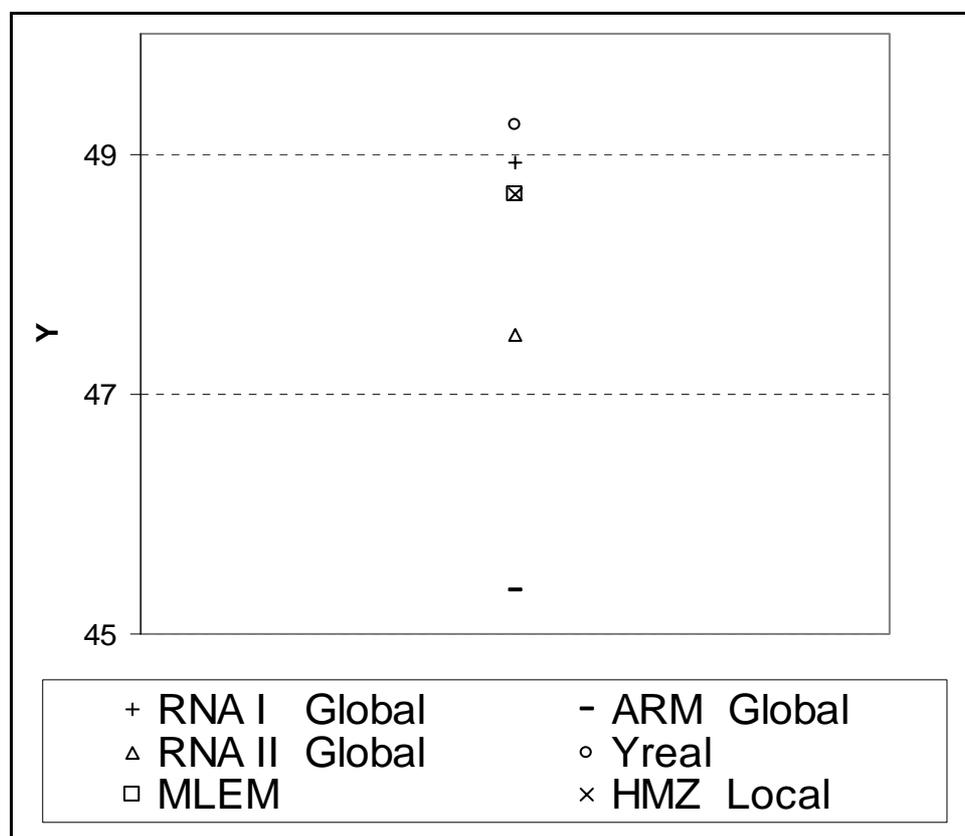


Figura 3 – Resultados de predição para avaliando 1.

Nota-se pela análise da figura 3 que para o elemento avaliando 1, o especialista global ARM fornece a pior predição e o HMZ juntamente com o da composição de especialistas locais, as melhores predições à exceção apenas do RNA I global. Em face da praticidade de cálculo, o especialista HMZ apresenta, neste caso, uma vantagem favorável quando comparados com os demais já que a precisão dos resultados entre estes dois especialistas (RNA I e HMZ) e de composição de especialistas locais, neste caso, é praticamente a mesma e o cálculo por RNA I é por demais *dispendiosa*, assim com o detalhamento destes resultados, constata-se a vantagem de modelos de composição de especialistas locais frente aos demais especialistas.

Conforme-se observa na figura 4 os resultados de predição para o avaliando 15 apontam para as melhores predições as do modelo de composição de especialistas locais e, contrariamente ao obtido para o avaliando 1, a do especialista global ARM. Este elemento apresenta-se localizado entre os centros de clusters da região I e II, ligeiramente mais próximo do centro da II já que o valor final obtido pelo modelo de composição é a média ponderada de valores apurados por RNA I e ARM e os pesos fornecidos pela rede supervisora são iguais, respectivamente a 0,44 e 0,56.

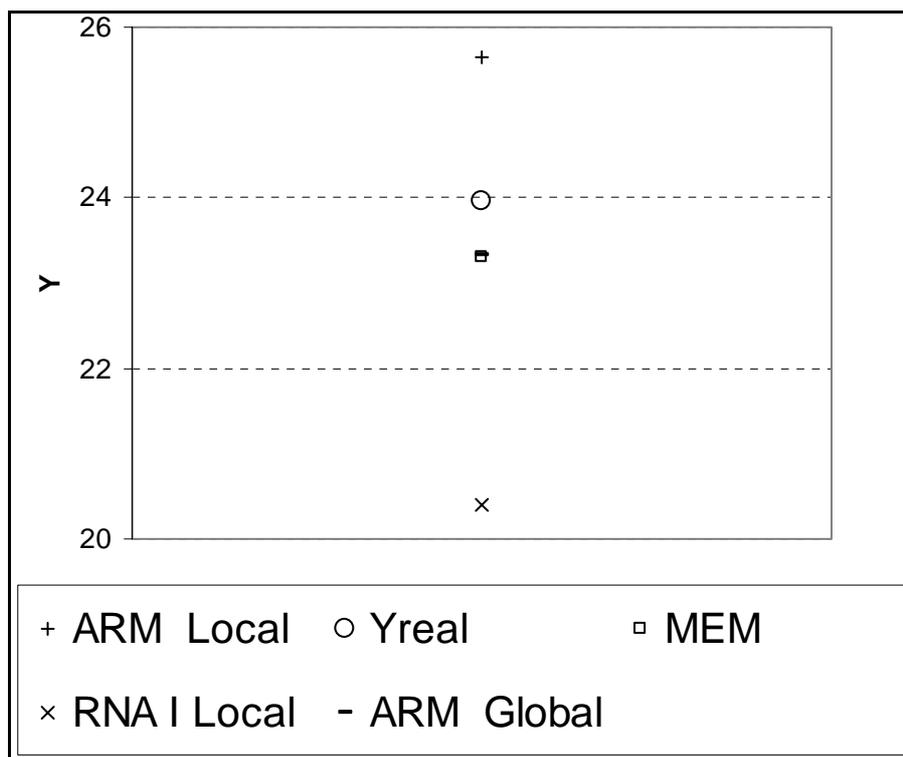


Figura 4 – Resultados de predição para avaliando 15.

## 7. Conclusões e recomendações finais

Através do estudo de caso visto, destaca-se no sistema de composição de especialistas locais, que a rede supervisora torna simples e interessante o uso dos diversos especialistas, cujos resultados individuais, independentemente tratados sob a mediação da mesma, permitem obter para o resultado final da modelagem menor erro de predição. Não há, a priori, prevalência de nenhum método sobre outro, o que há é preferência por determinados especialistas eleitos por melhor explicarem a variação de valores dos diversos dados e indiscriminadamente (RNA, ARM e HMZ) tornam exequível a obtenção de melhor predição, seja de forma isolada ou pela composição dos especialistas locais.

É importante registrar, todavia, que tais melhorias de desempenho obtidas pela composição de especialistas locais demonstradas através de menores erros de predição devem ser ponderadas com a complexidade extra e os custos de sua implantação e manutenção quando comparados com os de especialistas globais operando isoladamente. Este “*tradeoff*” deve ser favorável e suficientemente significativo para a aplicação desta abordagem de composição de especialistas locais (Khotanzad *et al.*, 2000).

O sistema proposto para a composição de especialistas locais tem uso potencial para uma grande variedade de problemas referentes a modelagens, principalmente na *predição* de valores. A utilização do sistema que permite obter os modelos de composição de especialistas locais contribui para melhorar o desempenho de predições, o que o habilita ser recomendado para tarefas de *previsão* em séries temporais e porque não para valores de bens que, até hoje, continuam sendo também desafios para os avaliadores. Ambas tarefas afins a área de Engenharia de Avaliações e fundamentais para análises e tomadas de decisão.

Finalmente, podemos concluir que é um erro capital *teorizar* sobre a prevalência de certa técnica sobre outra se ainda não temos os dados devidamente explorados e modelos validados. Ao máximo, podemos ter preferência, circunstancial, em face de sua escolha por apontar menores erros de predição que conforme vimos, são aleatórios, dependentes dos dados e especialistas utilizados.

## Referências Bibliográficas

- [1] GUJARATI, D. N. (2000) – *Econometria Básica*. 3ª edição, Makron Books. São Paulo.
- [2] HAYKIN, S. (1999) – *Neural Networks: A Comprehensive Foundation*. 2<sup>nd</sup> ed., Prentice Hall. New York.
- [3] JACOBS, R. A.; JORDAN, M. I.; NOWLAN, S. J. & HINTON, G. E. (1991) – Adaptive Mixture of Local Experts. *Neural Computation*. Vol. 3, No. 1, pp.79-87, MIT Press.
- [4] KHOTANZAD, A.; ELRAGAL, H.; LU, T. L. (2000) – Combination of Artificial Neural Network Forecasters for Prediction of Natural Gas Consumption. *IEEE Transactions on Neural Networks*. Vol. 11, No. 2, pp. 464-473.
- [5] KOHONEN, T. (1989) – *Self-Organization and Associative Memory*. 3<sup>rd</sup> ed., Springer-Verlag. Berlin.
- [6] MAKRIDAKIS, S.; WHEELWRIGHT, S. & HYNDMAN, R. J. (1998) - *Forecasting Methods and Applications*. 3<sup>rd</sup> ed., John Wiley & Sons. New York.
- [7] MILIDIÚ, R. L.; MACHADO, R. J. & RENTERÍA, R. P. (1999) - Time-series forecasting through wavelets transformation and a mixture of expert models. *Neurocomputing*. Vol. 28, No. (1-3), pp. 145-156.
- [8] NASCIMENTO JR, C. L. & YONEYAMA, T. (2000) - *Inteligência Artificial em Controle e Automação*. Editora Edgard Blücher. São Paulo.
- [9] PINDYCK, R. S. & RUBINFELD, D. L. (1998) – *Econometric models and economic forecasts*. 4<sup>th</sup> ed., McGraw-Hill. New York.
- [10] WEIGEND, A. S. & GERSHENFELD, N. A. (1994) – *Time Series Prediction: Forecasting the Future and Understanding the Past*; Addison Wesley. Reading.