

# MÉTODOS DE PREVISÃO USANDO A SÉRIE ECONOMÉTRICA DE NELSON-PLOSSER: UM ESTUDO COMPARATIVO

**Pedro Paulo Balestrassi**

EFEI –Escola Federal de Engenharia de Itajubá - e-mail: [pedro@eps.ufsc.br](mailto:pedro@eps.ufsc.br)

**Robert Wayne Samohyl**

UFSC - Departamento de Engenharia de Produção e Sistemas - e-mail: [samohyl@eps.ufsc.br](mailto:samohyl@eps.ufsc.br)

**Roberto Meurer**

UFSC – Departamento de Ciências Econômicas - e-mail: [rmeurer@mbox1.ufsc.br](mailto:rmeurer@mbox1.ufsc.br)

**Abstract:** *The time series forecasting is one of the most important objective in econometric and production Engineering and many mathematical methods have been used in this domain. The Artificial Neural Networks (ANN) and the Econometric Models are two of these proeminent methods. The main goal of this paper is to compare these methods using the classical econometric Nelson-Plosser data.*

**Área:** *10-Gestão da Informação (10.1-Redes Neurais)*

**Keywords:** *Neural Network, Econometric, Time Series*

## 1. Introdução

A previsão de séries temporais é um dos fatores mais importantes em econometria e em engenharia de produção e diversos métodos matemáticos tem sido empregados nessa tarefa. As Redes Neurais Artificiais (RNA) e os Modelos Econométricos são dois desses métodos.

Duas dessas redes neurais, a rede *Multilayer Perceptron* treinada por *Backpropagation* (representada aqui simplesmente por *Backpropagation*) e a rede com Função de Base Radial (RBF), são aqui utilizadas (conforme 2ª seção). Os modelos autoregressivos, usando testes de raízes unitárias e co-integração, são também aqui empregados comparativamente (3ª seção). Na 4ª seção, resultados e conclusões, os valores de previsão são fornecidos para os métodos utilizados e algumas considerações são feitas. É importante ressaltar que existe um paralelo que pode ser feito entre métodos estatísticos e redes neurais, utilizados nesse texto, conforme mostra a tabela 1.

<b>Estatística</b>	<b>Redes Neurais</b>
Modelo	Rede
Estimação	Aprendizagem
Regressão	Aprendizagem supervisionado
Interpolação	Generalização
Observações	Conjunto de treinamento
Parâmetros	Pesos
Variáveis independentes	Entradas
Variáveis dependentes	Saídas
Regressão de Ridge	Decaimento dos pesos

Tabela 1: Terminologia correspondente entre estatística e redes neurais

A série clássica de Nelson-Plosser, representando os dados macroeconômicos dos EUA (obtida em [1]), possui as seguintes variáveis:

RGNP	Produto Interno Bruto Real
GNP	Produto Interno Bruto
PCRGNP	Produto Interno Bruto Real Per Capita
IP	Produção Industrial
EMP	Emprego
UN	Taxa de Desemprego
PRGNP	Deflator do Produto Interno Bruto
CPI	Preço ao Consumidor
WG	Salário
RWG	Salário Real
M	Estoque Monetário
VEL	Velocidade
BND	Produção Vinculada
SP500	Preços de Estoque Comum

Vale ressaltar que a variável escolhida para ter seus valores previstos foi a CPI (Preço ao Consumidor) tendo suas 4 últimas observações utilizadas para avaliação dos métodos.

Todo esse trabalho foi desenvolvido no sentido de que seus resultados possam ser facilmente reproduzidos, usando os mesmos programas computacionais, e estimular um estudo de *Benchmark* com respeito aos métodos de previsão disponíveis.

## 2. Modelo Econométrico

A utilização de modelos econométricos de uma forma dinâmica por meio de regressões é atualmente descrito como um dos tópicos mais avançados em previsão de séries temporais [1]. Diversos fatores são levados em consideração nesse tipo de análise: heteroscedasticidade, multicolinearidade, raízes unitárias, correção de erros, co-integração, etc. Para a série de dados em questão foi utilizada a metodologia do programa computacional PcGive 8.0.

Uma das suposições mais frequentes que se faz a respeito de uma série temporal é de que ela seja estacionária, ou seja, ela se desenvolve no tempo aleatoriamente em torno de uma média constante, refletindo alguma forma de equilíbrio estável. Todavia, a maior parte das séries que encontramos na prática apresentam alguma forma de não-estacionaridade. Os modelos de auto-regressão, podem ser utilizados para descrever de maneira satisfatória séries estacionárias. Séries temporais que não se comportam de forma estacionária devem ser transformadas. A transformação mais comum consiste em tomar diferenças sucessivas da série original, até que se obtenha a estacionaridade. As duas principais razões para se efetuar transformações são: estabilizar a variância e conseguir a normalidade.

Uma rápida inspeção visual dos gráficos da série de Nelson-Plosser permite concluir que as variáveis são não-estacionárias. Para investigar a estacionaridade de uma série três métodos são frequentemente utilizados: (i) o correlograma, (ii) o teste da raiz unitária e (iii) a estatística de Durbin-Watson. O correlograma é essencialmente visual e representa a função de correlação serial das variáveis. Ele nos fornece sinais sobre o tipo de transformação a ser feita sobre os dados.

Tomando-se a primeira diferença de todas as variáveis (representado aqui pela letra D acrescida à variável) e a segunda diferença da variável M (representado por DDM), o teste da raiz unitária de Dickey-Fuller (DF) e Dickey-Fuller Ampliado (ADF) rejeita a hipótese nula da não-estacionaridade ( $H_0$ : existe raiz unitária) a um nível pequeno de significância (5%). Os resultados do teste encontram-se na tabela 2.

VARIÁVEIS	I(0)				I(1)			
	S/ Tendência		C/ Tendência		S/ Tendência		C/ Tendência	
	DF	ADF(1)	DF	ADF(1)	DF	ADF(1)	DF	ADF(1)
DRGNP	-4.6425	-3.8476	-5.0875	-4.3532				
DGNP	-2.8586	-1.8799	-4.6948	-3.6125				
DPCRGNP	-5.0266	-4.4000	-5.0818	-4.4855				
DIP	-6.0244	-4.7869	-6.5003	-5.4028				
DEMP	-4.7189	-4.8808	-4.7298	-4.9133				
DUN	-5.3988	-5.4426	-5.3489	-5.3935				
DPRGNP	-4.5900	-3.6882	-4.8535	-3.9763				
DCPI	-3.9228	-4.2780	-4.0624	-4.4487				
DWG	-3.1131	-2.8691	-4.3670	-4.3660				
DRWG	-4.8484	-4.3090	-4.8756	-4.3712				
DM	-1.7586	-1.8369	-2.7079	-2.8689	-7.2250	-4.9821	-7.1527	-4.9034
DVEL	-6.4946	5.5140	-6.4775	-5.5209				
DBND	-5.4523	-2.2874	-6.1080	-2.8543				
DSP500	-5.4123	-4.4858	-5.5756	-4.8246				

Tabela 2- Teste de raízes unitárias (Nota: Todas as equações incluem um termo constante. Os valores críticos ao nível de significância de 5% são: sem tendência, -2.9109; com tendência -3.4862; período de 1909-1970.)

Na tabela 3, são calculados os coeficientes autoregressivos do modelo após a correção da estacionaridade. Foram feitos sucessivos testes a partir de um modelo genérico com 2 lags para cada variável diferenciada. Os termos irrelevantes do modelo foram gradualmente retirados (usando o critério F de Fisher) e algumas estatísticas de controle foram apresentadas

Variável	Coeficiente	Erro-padrão	t-value	t-prob	R <sup>2</sup> Parcial
Constante	0.45395	0.13233	3.430	0.0016	0.2516
DCPI_1	0.68266	0.10502	6.500	0.0000	0.5469
DRGNP_1	-0.17477	0.058290	-2.998	0.0050	0.2044
DGNP_1	0.10343	0.035895	2.881	0.0067	0.1917
DPCRGNP	0.0050872	0.0021221	2.397	0.0220	0.1410
DPCRGNP_1	0.016074	0.0055445	2.899	0.0064	0.1936
DIP_2	-0.18105	0.038799	-4.666	0.0000	0.3835
DEMP	-0.42899	0.14903	-2.879	0.0068	0.1914
DEMP_2	0.23737	0.069310	3.425	0.0016	0.2510
DUN	-0.17560	0.066886	-2.625	0.0127	0.1645
DPRGNP	0.46504	0.047633	9.763	0.0000	0.7314
DPRGNP_1	-0.28090	0.074879	-3.751	0.0006	0.2868
DWG	0.012291	0.0017186	7.152	0.0000	0.5937
DWG_1	-0.017193	0.0036356	-4.729	0.0000	0.3899
DRWG	-0.70149	0.12062	-5.816	0.0000	0.4914
DRWG_1	0.91608	0.22316	4.105	0.0002	0.3250
DVEL_1	.6448	0.96312	2.746	0.0095	0.1773
DBND_1	0.80827	0.30781	2.626	0.0127	0.1646
DDM_1	-0.064376	0.024935	-2.582	0.0142	0.1600

Tabela 3: Modelamento de DCPI por Mínimos Quadrados Recursivos

Os valores encontrados para o teste de Fisher,  $F(18, 35)=130.89$ , indicam significativamente que pelo menos uma das variáveis independentes do modelo explica a variável dependente DCPI. O teste Durbin-Watson,  $DW=2.04$ , mostra a ausência de correlação serial, estando o seu valor bem próximo de seu valor ideal.

Os coeficientes calculados mostram-se de acordo com a teoria econômica. O teste de estabilidade dos parâmetros de Chow,  $F(4, 35)=1.6729$  [0.1783], revela a não-existência de uma diferença significativa entre os valores observados e projetados.

Quanto a normalidade dos resíduos, um aspecto importante nesse tipo de modelamento, os valores da estatística de aderência aceitam a hipótese de que a distribuição dos resíduos é gaussiana ( $\chi^2(2)=0.067622$  [0.9668]). Além disso não houve indícios de heteroscedasticidade condicional autoregressiva, ou seja, a variância dos resíduos permanece constante ao longo do tempo (ARCH 1  $F(1,33)=0.000239$  [0.9880]). A forma funcional encontrada para o modelo pode ser escrita como uma combinação linear das variáveis independentes (RESET  $F(1,34)=1.9754$  [0.1689]).

Foram ainda realizados os testes de co-integração verificando-se que apenas a variável: DDM não era co-integrada com a sua variável dependente. Feito o modelo de mecanismo de correção de erros, os resultados de previsões encontrados (por serem previsões a curto prazo) não diferiram do modelo formulado sem o termo de correção de erro. Finalmente, os resultados de previsão, objeto desse trabalho, podem ser vistos na figura 1. Os valores numéricos são avaliados na seção 4.

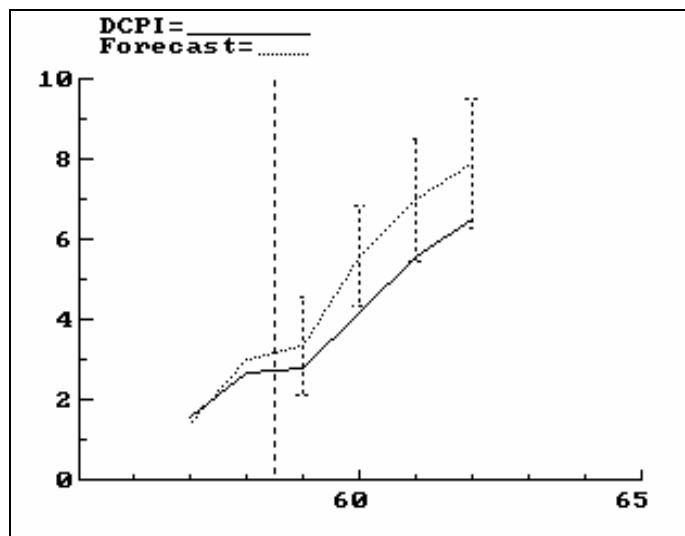


Figura 1: Intervalos de previsão dos 4 últimos anos

### 3. Redes Neurais Artificiais

Os resultados da previsão da série de Nelson-Plosser através de RNA, foram aqui obtidos usando *Backpropagation* e RBF. Para tanto foi utilizado o programa MatLab 4.2 com pequenas modificações em algumas rotinas da *Toolbox* de redes neurais. MatLab é marca registrada da The MathWorks Inc. Não é objetivo deste texto descrever a teoria matemática de RNA que pode ser encontrada facilmente em livros da área.

O formato dos dados (de treinamento e teste) das redes neurais está exemplificado abaixo (houve aqui uma transformação linear da série):

Entrada	Saída (para treinamento)
0.0240 (RGNP)	0.0193 (CPI)
0.0066 (GNP)	
0.0918 (PCRGNP)	
0.0248 (IP)	
0.0690 (EMP)	Os dados foram transformados linearmente
0.1308 (UN)	obtendo-se a amplitude de cada série e
0.0188 (PRGNP)	dividindo-se cada observação por essa
0.0119 (WG)	amplitude.
0.0366 (RWG)	
0.0077 (M)	
0.7557 (VEL)	
0.3037 (BND)	
0.0289 (SP500)	

Algumas observações a respeito da aplicação de redes neurais são aqui ressaltadas (para um maior aprofundamento uma boa referência é dada em [4]):

- É demonstrado que as redes neurais do tipo *feedforward* são aproximadores de função universais (teoremas de Kolmogorov, Sprecher e Hecht-Nielsen [4]). Esse resultado teórico é acompanhado por investigações numéricas que tem mostrado experimentalmente a grande capacidade de redes neurais relativamente simples em aproximar mapas não-lineares. Esses achados estimulam a aplicação de redes neurais no campo de séries temporais não lineares.
- No projeto de uma rede neural os resultados finais de desempenho, e especialmente a sua capacidade de generalização da solução, dependem da configuração inicial dos parâmetros da rede neural. Isso é, via de regra, baseado em heurística e muitas vezes pode tornar o método extremamente complexo.
- Uma rede neural pode construir um modelo linear a partir de uma função de transferência não-linear. Se a equação de uma rede neural for (por exemplo) dada pela equação:

$$x_t = w_0 + \sum_{l=1}^H w_l \tanh\left(\sum_{k=1}^I a_{lk} x_{t-k} + a_{l0}\right)$$

( $H$ : número de camadas ocultas,  $I$ : número de entradas da rede)

para pequenos valores de  $|x|$ ,  $\tanh(x)$  é aproximadamente igual a  $x$  e a equação acima pode ser re-escrita em forma de um modelo autoregressivo de ordem  $i$ :

$$x_t = w_0 + \sum_{k=1}^I \left( \sum_{l=1}^H w_l a_{lk} \right) x_{t-k} + \sum_{l=1}^H w_l a_{l0}$$

### ***Backpropagation***

A rede neural apresentada na figura 2 pode ser usada como um aproximador universal de funções. Ela pode aproximar qualquer função com um número finito de descontinuidades.

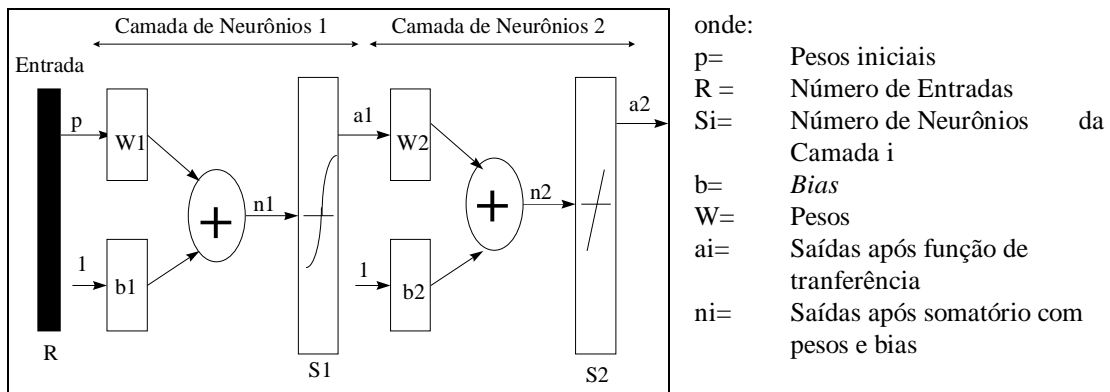


Figura 2: Rede *Multilayer Perceptron (feedforward)*

O processo de aprendizagem da rede *Backpropagation* foi implementado utilizando o código do programa em MatLab 4.2 usando as funções *INITFF*, *TRAINBPX* e *SIMUFF* como descrito a seguir.

```

load P.txt      % P define 13 entradas de treinamento com vetores de diferenças minimizados
load T.txt      % Alvo de valores minimizados (CPI)
S1= 20;        % Uma camada oculta com S1 neurônios

                % Inicialização dos pesos da rede

PIni=[0 1; 0 1; 0 1; 0 1; 0 1; 0 1; 0 1; 0 1; 0 1; 0 1; 0 1; 0 1; 0 1; 0 1]
TIni=[0 1];
[w1,b1,w2,b2] = initff(PIni,S1,'tansig',TIni,'tansig'); % Função de Inicialização da Rede

                % Treinamento da rede
df = 25;       % Frequência de atualização do gráfico de erros.
me =3000;     % Máximo número de épocas a ser treinada
eg = 0.001;   % SSE
lr = 0.01;    % Taxa de Aprendizagem
tp = [df me eg lr]; % Uso de Backpropagation com Taxa de Aprendizagem
[w1,b1,w2,b2,ep,tr] = trainbpx(w1,b1,'tansig',w2,b2,'purelin',P,T,tp); % Função de Treinamento da Rede

                % Teste da Rede (Previsão e Apresentação do Resultado)

load TesIn.txt % Vetor de teste de entrada
load Alvo.txt  % Alvo do teste (CPI)
for i=1:4
teste(i) = simuff(TesIn(:,i),w1,b1,'tansig',w2,b2,'purelin'); % Função de Teste da Rede
Previsão(i)= teste(i)*88.3+28; % Fator de Correção Linear da Minimização da série CPI
disp([Alvo(i), Previsão(i),(Alvo(i)- Previsão(i)),(Alvo(i)- Previsão(i))/Alvo(i)]);
end;
disp('Fim')

```

### **RBF**

As funções de base radiais são uma classe especial de funções. A característica específica desse tipo de função é que a resposta diminui (ou aumenta) monotonicamente com a distância de um ponto central. Tal ponto, uma distância escalar e a forma da função de base radial são parâmetros do modelo que devem ser fixados como se fossem lineares.

Desde os trabalhos de Broomhead e Lowe em 1988 as funções de base radial tem sido tradicionalmente relacionadas às RNA com apenas uma camada intermediária. Um bom

texto (com código de programa computacional) sobre RBF pode ser livremente obtido no endereço eletrônico: <http://www.cns.ed.ac.uk/people/mark>

Uma enorme vantagem desse tipo de rede sobre a *Backpropagation* é que o projeto de sua rede é muito mais simples em função da definição heurística de um número bem menor de parâmetros.

O programa em MatLab usando as funções SOLVERB e SIMURB é dado a seguir:

```
load P.txt      % P define 13 entradas de treinamento com vetores de diferenças minimizados
load T.txt      % Alvo de valores minimizados (CPI)
df = 10;        % Frequência de atualização do gráfico de erros.
mn = 20;        % Máximo Número de Neurônios
eg = 0.01;     % SSE.
sc = 100;      % Constante de Alargamento para a Função de Base Radial
tp = [df mn eg sc];
[w1,b1,w2,b2,nr,tr] = solverb(P,T,tp); % Função de Treinamento da Função de Base Radial

                % Teste da Rede (Previsão e Apresentação do Resultado)

load TesIn.txt  % Vetor de teste de entrada
load Alvo.txt   % Alvo do teste (CPI)
for i=1:4
    teste(i) = simurb(TesIn(:,1),w1,b1,w2,b2); % Função de Teste da Função de Base Radial
    Previsão(i)= teste(i)*88.3+28; % Fator de Correção Linear da Minimização da série CPI
    disp([Alvo(i), Previsão(i),(Alvo(i)- Previsão(i)),(Alvo(i)- Previsão(i))/Alvo(i)]);
end;
disp('Fim')
```

#### 4. Resultados e Conclusões

A tabela 4 apresenta os resultados de previsão para as últimas 4 observações da série de Nelson-Plosser. Pode-se perceber que os valores são bem próximos e se revelam estatisticamente iguais.

Real	Modelo Econométrico			Backpropagation			RBF		
	Previsto	Diferença	Difer./Real	Previsto	Diferença	Difer./Real	Previsto	Diferença	Difer./Real
100.0	100.57	-0.57	-0.0057	100.10	-0.10	-0.0010	98.69	1.30	0.0130
104.2	105.60	-1.40	-0.0135	105.30	-1.10	-0.0106	103.36	0.83	0.0080
109.8	111.22	-1.42	-0.0129	110.55	-0.75	-0.0069	108.60	1.19	0.0109
116.3	117.71	-1.41	-0.0122	116.95	-0.65	-0.0056	116.96	-0.66	-0.0057

Tabela 4: Resultados de previsão

No intuito de estabelecer um estudo comparativo entre os métodos, algumas vantagens e desvantagens podem aqui ser relacionadas:

- *Redes Neurais*

- *Vantagens*

- Menor erro de previsão (não estatisticamente significativa);
    - Menor dependência do econometrista;
    - Existem poucas restrições à morfologia dos dados;
    - Rapidez nos cálculos;
    - Pode frequentemente funcionar como uma “caixa preta”;
    - Não necessidade de definição da variável endógena

- *Desvantagens*
  - Pode frequentemente funcionar como uma “caixa preta”;
  - Não há, frequentemente, um controle gráfico das variáveis;
  - Não há, frequentemente, intervalos de confiança para estimativas
- *Modelo Econométrico*
  - *Vantagens*
    - A equação do modelo é obtida;
    - Os intervalos de confiança das estimativas são fornecidos;
    - Há um acompanhamento gráfico das variáveis modeladas;
    - Rapidez nos cálculos;
  - *Desvantagens*
    - Existem inúmeras condições iniciais que devem ser observadas;
    - Existe um grande número de testes de hipóteses;
    - Tem maior erro de previsão (não estatisticamente significativo);
    - Não pode frequentemente funcionar como uma “caixa preta”;

### *Conclusões*

A rede com Função de Base Radial, sendo considerada como *aproximador universal* teve o desempenho ligeiramente superior (não significativo estatisticamente) com o menor tempo de treinamento. O número de parâmetros (definidos heurísticamente) é bem menor nesse tipo de rede.

A rede *Multilayer Perceptron* com treinamento por *Backpropagation* teve resultados similares à rede com Função de Base Radial mas com um número de parâmetros superior e um tempo de treinamento também superior.

O funcionamento do método como uma *caixa preta* pode ser ou não uma vantagem, dependendo do nível de conhecimento do usuário.

O modelo econométrico permitiu um maior conhecimento das variáveis envolvidas na série, permitindo (a posteriori) uma melhor definição sobre o tipo de transformação a ser efetuada sobre os dados para a entrada das redes neurais.

Uma importante conclusão do trabalho é que ambos os métodos atuantes conjuntamente (melhor que separadamente) formam uma poderosa ferramenta de previsão econométrica. O pré-processamento dos dados é extremamente importante em qualquer tipo de rede neural e isso pode ser viabilizado pelo modelo econométrico.

## **5. Bibliografia**

- [01] MADALLA, G.S. **Introduction to Econometrics**. New Jersey, Prentice-Hall, Inc. 1992
- [02] DOORNICK, J. A.; HENDRY, D. F. **PcGive 8.0 An Interactive Econometric Modelling System**. London, Chapman & Hall, 1995.
- [03] KUAN, C.M.; WHITE, H. “Artificial Neural Networks: An Econometric Perspective” *Econometric Reviews*, 13(1), 1994.
- [04] FAUSSET, L. V. **Fundamentals of Neural Networks** New Jersey, Prentice-Hall, Inc. 1994