
UMA METODOLOGIA DE MODELAGEM EMPÍRICA UTILIZANDO O INTEGRADOR NEURAL DE MÚLTIPLOS PASSOS DO TIPO ADAMS-BASHFORTH

Regina Paiva Melo Marin*
reginam@ita.br

Paulo Marcelo Tasinaffo*
tasinafo@ita.br

*Divisão de Ciência da Computação - Instituto Tecnológico de Aeronáutica(ITA)
Pça Mal. Eduardo Gomes, 50,Vila das Acácias, 12.228-900
São José dos Campos,São Paulo,Brasil

ABSTRACT

An Empirical Modeling Methodology Using Multiple Steps Neural Integrator of Adams-Bashforth

This paper presents and develops an alternative empirical methodology to model and get instantaneous derivative functions for nonlinear dynamic systems by a supervised training using multiple step neural numerical integrator of Adams-Bashforth. This approach, neural network plays the role of instantaneous derivative functions and it is coupled to numerical integrator structure, which effectively is the responsible for execute the propagations in time through a linear combination of feedforward neural networks with delayed responses. It is an important fact that only numerical integrators of highest order effectively learn instantaneous derivative functions with suitable precision, which proves the fact that those of first order can only learn mean derivatives. This approach is an alternative to the methodology that deals with the problems of neural modeling in simple step integration structures of high-order Runge-Kutta type, and this, which is more robust and complex in determining the backpropagation, which requires - in this case - the employment of the chain rule for compounded functions. At the end this paper numerical simulation results of Adam-Bashforth neural integrator are presented in three study cases: 1) nonlinear pendulum without variables control; 2) an abstract model with variables control and 3) Van der Pol system.

KEYWORDS: Multiple Step Neural Networks, Feedforward Nets, High Order Numerical Integrators, Ordinary Differential Equations, Dynamic Systems Neural Modeling, Back-Propagation Algorithm.

RESUMO

Este artigo apresenta e desenvolve uma metodologia empírica alternativa para modelar e obter as funções de derivadas instantâneas para sistemas dinâmicos não-lineares através de um treinamento supervisionado utilizando integradores numéricos neurais de múltiplos passos do tipo Adams-Bashforth. Esta abordagem a rede neural desempenha o papel das funções de derivadas instantâneas que é acoplada à estrutura do integrador numérico, que efetivamente, é o responsável em realizar as propagações no tempo apenas através de uma combinação linear de redes neurais *feedforward* com respostas atrasadas. É um fato importante que somente os integradores numéricos de mais alta ordem aprendem efetivamente as funções de derivadas instantâneas com precisão adequada, o que comprova o fato de que os de primeira ordem somente conseguem aprender as derivadas médias. Esta abordagem é uma alternativa à metodologia que trata os problemas de modelagem neural em estruturas de integração de passo simples do tipo Runge-Kutta de alta-ordem, sendo esta, mais robusta e complexa na determinação da retropropagação, que exige - neste caso - o emprego da regra da cadeia para funções compostas. Ao final deste artigo são apresentadas simulações de resultados numéricos dos integradores neurais de Adams-Bashforth em três estudos de caso: 1) pên-

Artigo submetido em 12/02/2009 (Id.: 00951)

Revisado em 19/03/2009, 17/08/2009, 14/10/2009, 02/02/2010, 02/03/2010

Aceito sob recomendação do Editor Associado Prof. Luis Antonio Aguirre

dulo não-linear sem variáveis de controle; 2) um modelo abstrato com controle e 3) sistema de Van der Pol.

PALAVRAS-CHAVE: Redes Neurais de Múltiplos Passos, Redes *Feedforward*, Integradores Numéricos de Alta Ordem, Sistemas de Equações Diferenciais Ordinárias, Modelagem Neural para Sistemas Dinâmicos, Algoritmo de Retropropagação.

1 INTRODUÇÃO

Redes neurais artificiais são consideradas aproximadores universais de funções conforme Hornik et al. (1989), Cybenko (1988) e Zurada (1992). Nesse contexto, foi desenvolvida com sucesso a tecnologia dos integradores neurais que são redes neurais acopladas às estruturas de integração numérica.

Existem basicamente dois métodos de integração numérica para resolução de sistemas de equações diferenciais ordinárias (EDO): as estruturas de passo simples (Euler e Runge-Kuttas) e múltiplos passos (Adams-Bashforth, Preditivo-Corretor, entre outros). Existem boas bibliografias sobre resolução numérica e analítica de sistemas de EDO em Wilson (1958), Henrici (1964); Sokolnikoff and Redheffer (1966) e Braun (1983).

A literatura apresenta também uma vasta bibliografia de aplicações da modelagem de sistemas dinâmicos não-lineares por redes neurais artificiais em teoria de controle, mas priorizando a metodologia NARMAX (*Non Linear AutoRegressive Moving Average with exogenous inputs*) como em Narendra and Parthasarathy (1990); Hunt et al. (1992) e Narendra (1996). Trabalhos específicos sobre modelagem e identificação de sistemas dinâmicos através da utilização exclusiva de redes neurais também são vastos Chen and Billings (1992), Rios Neto (2001a).

A modelagem de sistemas dinâmicos não-lineares tratada como um sistema de equações diferenciais ordinárias combinando as técnicas de estruturas de integração numérica e redes neurais artificiais teve seu início com o trabalho de Wang and Lin (1998). Neste artigo surgiu o termo Redes Neurais de Runge-Kutta que trabalha com funções compostas, fato que dificulta enormemente a determinação das derivadas exigidas pelo algoritmo *backpropagation*. Dada a complexidade de se lidar com as redes de Runge-Kutta outras metodologias, em princípio, mais simples do que estas foram desenvolvidas, entre elas, a aplicação de integradores de simples e múltiplos passos desenvolvidos em Rios Neto (2001b); Melo (2008) e Melo and Tasinaffo (2008).

A metodologia das derivadas médias descrita em Tasinaffo (2003) e Tasinaffo and Rios Neto (2005) também surgi-

ram com o intuito de simplificar a abordagem envolvendo redes neurais e estruturas de integração numérica. A modelagem neural de sistemas dinâmicos envolvendo a resolução de equações diferenciais parciais já é mais rara He et al. (2000) e Mai-Duy and Tran-Cong (2001). Assim, podemos dizer que atualmente existem três metodologias para representação e modelagem de sistemas dinâmicos não-lineares regidos por equações diferenciais ordinárias: 1) metodologia NARMAX; 2) derivadas instantâneas e 3) derivadas médias. Sendo que estas duas últimas ainda podem ser divididas em metodologia direta e metodologia indireta ou empírica.

Do ponto de vista das estruturas de integração numérica, tem-se duas metodologias distintas, uma envolvendo as funções de derivadas instantâneas e outra as funções de derivadas médias. A metodologia das derivadas instantâneas tem como característica a utilização do passo de integração variável com estruturas de integração de qualquer ordem, entretanto resultados precisos são alcançados somente com integradores de alta ordem. A metodologia das derivadas médias utiliza passo de integração fixo, sendo aplicável somente ao integrador do tipo Euler de primeira ordem, cuja precisão é semelhante a obtida pela primeira metodologia.

Por outro lado, do ponto de vista do treinamento supervisionado das redes *feedforward* pode-se classificar as estruturas de integração neurais como metodologias diretas e empíricas, esta também conhecida como indiretas. As metodologias diretas utilizam o integrador numérico somente na simulação dos resultados das redes *feedforward*, e as metodologias empíricas usam o integrador numérico na aquisição de dados, durante o próprio treinamento e também na fase de simulação da rede. Já as metodologias de treinamento supervisionado empíricas podem adquirir os padrões de treinamento de duas maneiras: da mesma forma que as metodologias de treinamento diretas, ou por aquisição explícita dos padrões de treinamento do ambiente (origem do nome empírica) quando não se tem um modelo analítico suficientemente preciso para representação dos sistemas. Utiliza-se modelos analíticos quando é mais dispendioso e complicado de lidar do que os experimentos reais.

Dessa forma, este artigo propõe modelar empiricamente as funções de derivadas instantâneas por meio de integradores neurais que não trabalhem com funções compostas. Isto pode ser conseguido com os integradores de múltiplos passos, pois estes apenas utilizam uma combinação linear das funções de derivadas instantâneas, para facilitar a determinação das derivadas exigidas pelo algoritmo da retropropagação na fase de treinamento neural e evitar assim a utilização excessiva e dispendiosa da regra da cadeia sobre a estrutura de integração numérica.

O artigo encontra-se organizado em cinco seções. Após a

sua introdução, apresenta-se a definição do problema abordado. Em seguida, descreve-se a metodologia de modelagem empírica das funções de derivadas instantâneas. A seção seguinte, relata as aplicações e análises de resultados. Ao final apresenta-se algumas considerações gerais e as conclusões.

2 DEFINIÇÃO DO PROBLEMA

O problema consiste em modelar empiricamente as funções de derivadas instantâneas por meio de integradores neurais que não trabalhem com funções compostas. Esta problemática pode ser resolvida utilizando os integradores de múltiplos passos, pois estes apenas trabalham com a combinação linear das funções de derivadas instantâneas, para facilitar a determinação das derivadas exigidas pelo algoritmo da retro-propagação na fase de treinamento neural e evitem assim a utilização excessiva e dispendiosa da regra da cadeia sobre a estrutura de integração numérica.

3 METODOLOGIA DE MODELAGEM EMPÍRICA DAS FUNÇÕES DE DERIVADAS INSTANTÂNEAS

3.1 Definição da Função de Derivadas Instantâneas

Seja o sistema autônomo de ODE:

$$\dot{y} = f(y) \quad (1)$$

onde,

$$y = [y_1, y_2, \dots, y_n]^T \quad (2)$$

$$f(y) = [f_1(y), f_2(y), \dots, f_n(y)]^T \quad (3)$$

A resolução numérica de uma equação de primeira ordem consiste em calcular o valor da variável de estado numa sequência discreta de instantes a partir da função $f(y)$, que é a derivada instantânea.

3.2 Geração dos Padrões de Treinamento

A geração de padrões de treinamento dos dados de entrada e saída da rede *feedforward* é realizada com base no número de condições iniciais, que varia de acordo com a complexidade do modelo analítico que se deseja representar, e o número de propagações atrasadas ou adiantadas através de um integrador numérico de média ou alta complexidade.

Nesta fase são definidos os parâmetros necessários para montar a arquitetura da rede para o treinamento da dinâmica do

sistema, tais como: o número de variáveis de estado, o número de variáveis de controle, o número de entradas atrasadas para os estados e os controles, a discretização do tempo, a ordem do integrador, entre outros.

A Figura 1 mostra a metodologia aplicada ao integrador Adams-Bashforth de quarta ordem. Nesta figura, **P** compõe o conjunto de condições iniciais propagadas três vezes no integrador de Runge-Kutta de média complexidade e **T** o conjunto de estados futuros devido a quarta e última propagação.

Esta figura apresenta a construção dos padrões de treinamento do integrador Adams-Bashforth de quarta ordem. Em **P** tem-se uma combinação linear de quatro redes neurais onde, no tempo t_{n-4} , tem-se a quarta rede neural, em t_{n-3} a terceira rede, em t_{n-2} a segunda rede e em t_{n-1} a primeira rede. A saída da estrutura neural será então, no horizonte t_n .

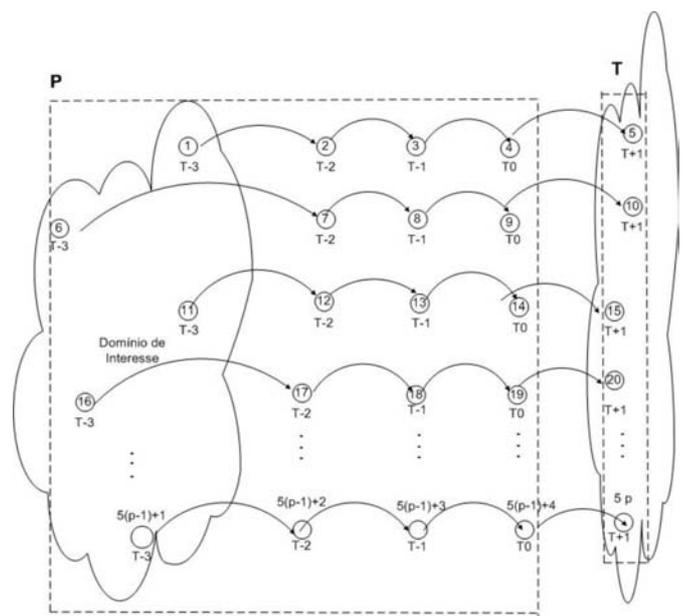


Figura 1: Propagações das Condições Iniciais em um Integrador de Quarta Ordem.

De maneira generalizada, sendo o a ordem do integrador utilizado e m o total de variáveis de estados, pode-se constatar que se existem p padrões de treinamento, efetua-se $o \cdot m$ propagações para cada uma das p condições iniciais. De acordo com a ordem do integrador escolhido são formados os vetores referentes ao número de variáveis de estado e controle definidos, que servirão como padrões de treinamento supervisionado para a rede *feedforward* combinada com qualquer integrador de múltiplos passos.

3.3 Treinamento Neural

No treinamento de uma rede *feedforward* a aprendizagem por correção de erro é a técnica mais utilizada para ensinar a rede a aproximar uma tarefa específica. Nesta aprendizagem, os vetores do valor desejado \bar{y} e de saída \hat{y} de uma determinada rede no instante t são considerados e o erro de saída da rede pode ser representado por:

$$r^i(t) = \bar{y}(t) - \hat{y}^i(t) \quad (4)$$

onde,

$$\bar{y}(t) = [\bar{y}_1(t), \bar{y}_2(t), \dots, \bar{y}_{n_e}(t)]^T \quad (5)$$

$$\hat{y}^i(t) = [\hat{y}_1^i(t), \hat{y}_2^i(t), \dots, \hat{y}_{n_e}^i(t)]^T \quad (6)$$

Os índices i e t denotam, respectivamente, a i -ésima iteração do aprendizado supervisionado, o t -ésimo padrão de treinamento, $t=1, \dots, p$ e n_e é o número total de estados. No aprendizado por correção de erro, o vetor $r^i(t)$ aciona um mecanismo de controle que produz uma seqüência de ajustes nos parâmetros da rede. Os ajustes têm a propriedade de corrigir, passo a passo, o sinal de saída \hat{y}^i em relação à resposta desejada \bar{y}^i para $i=1, \dots, I$ até a rede alcançar um erro desejável.

Este objetivo pode ser alcançado, em geral, minimizando uma função de custo ou índice de desempenho $J^i(t)$, dado pelo produto matricial da equação (7) ou pelo produto escalar da equação (8) a seguir:

$$J^i(t) = \frac{1}{2} r^i(t)^T \cdot r^i(t) \quad (7)$$

ou,

$$J^i(t) = \frac{1}{2} r^i(t) \bullet r^i(t) \quad (8)$$

Para este funcional quadrático, é válida a seguinte relação em Braga et al. (2000):

$$\frac{\partial J^l(t, i)}{\partial w_{jk}^l} = -r^i(t) \bullet \frac{\partial \hat{y}^i(t)}{\partial w_{jk}^l} \quad (9)$$

Na equação (9) o índice l representa a l -ésima camada da rede, o índice k o k -ésimo neurônio da camada anterior ($l-1$) e o j o j -ésimo neurônio da camada l atual. Na equação 10 aparece o índice m chamado índice da soma, que admite valores inteiros entre 1 (limite inferior) e n_e (limite superior), sendo possível também representar as derivadas parciais do funcional $J^l(t)$ da seguinte forma:

$$\frac{\partial J^l(t, i)}{\partial w_{jk}^l} = - \sum_{m=1}^{n_e} [\bar{y}_m(t) - \hat{y}_m^i(t)] \cdot \frac{\partial \hat{y}_m^i(t)}{\partial w_{jk}^l} \quad (10)$$

$$\frac{\partial J^l(t, i)}{\partial w_{jk}^l} = -[\bar{y}(t) - \hat{y}^i(t)] \bullet \frac{\partial \hat{y}^i(t)}{\partial w_{jk}^l} \quad (11)$$

Se $\hat{y}(t)$ for aproximado por uma estrutura de integração neural de múltiplos passos, tem-se uma combinação linear da retropropagação convencional podendo ser aplicada à mesma rede ou em redes distintas sobre o funcional quadrático.

A equação analítica das derivadas instantâneas para integradores de múltiplos passos foi deduzida com base no encadeamento para trás por uma questão de facilidade na elaboração matemática conforme a equação (12), sendo importante salientar que a ordem o do integrador será também o número total de combinações lineares das redes *feedforward* a serem empregadas. Por definição, $t_n = t + n \cdot \Delta t$ e,

$$\hat{y}^i(t_n) = y^i(t_{n-1}) + \frac{h}{\alpha} \sum_{i=1-o}^0 \beta_i \hat{f}[y(t_{n-1-i})] \quad (12)$$

onde α, β são os coeficientes de integração em função da ordem o do integrador numérico, m o número total de variáveis de estado, h o passo de integração, \hat{f} é a saída da rede neural, \hat{y} é o modelo do sistema dinâmico.

Substituindo a equação (12) em (11), tem-se:

$$\frac{\partial J^l(t_n, i)}{\partial w_{jk}^l} = -[\bar{y}(t_n) - \hat{y}(t_n)] \bullet \frac{\partial}{\partial w_{jk}^l} \{y(t_{n-1}) + \frac{h}{\alpha} \cdot \sum_{i=1-o}^0 \beta_i \hat{f}[y(t_{n-1-i})]\} \quad (13)$$

$$\frac{\partial J^l(t_n, i)}{\partial w_{jk}^l} = -[\bar{y}(t_n) - \hat{y}(t_n)] \bullet \left\{ \frac{h}{\alpha} \cdot \sum_{i=1-o}^0 \beta_i \frac{\partial}{\partial w_{jk}^l} \hat{f}[y(t_{n-1-i})] \right\} \quad (14)$$

Sabe-se que o produto escalar possui as seguintes propriedades segundo Apostol (1963):

1. lei comutativa: $A \bullet B = B \bullet A$;
2. lei distributiva: $A \bullet (B + C) = A \bullet B + A \bullet C$;
3. para m escalar: $m(A \bullet B) = (mA) \bullet B = A \bullet (mB) = (A \bullet B)m$;

Combinando as propriedades (2) e (3), tem-se:

$$A \bullet (\alpha B + \beta C) = \alpha(A \bullet B) + \beta(A \bullet C) \quad (15)$$

onde, $A = (a_1, \dots, a_n)$, $B = (b_1, \dots, b_n)$ e $C = (c_1, \dots, c_n)$.

Em nível de comparação, quanto aos aspectos de complexidade algorítmica, a seguir a expressão (16) define a retropropagação convencional, enquanto a equação (17) é a aplicação da propriedade (15) em (14), resultando assim na retropropagação estendida.

Verifica-se que a utilização de uma estrutura de integração neural de múltiplos passos do tipo Adams-Bashforth de ordem o , resultará na combinação linear de o retropropagações atrasadas, durante a fase de treinamento neural.

$$\frac{\partial J^l(t_n, i)}{\partial w_{jk}^l} = -[\bar{y}(t_n) - \hat{y}^i(t_n)] \bullet \frac{\partial \hat{f}[y(t_n)]}{\partial w_{jk}^l} \quad (16)$$

$$\frac{\partial J^l(t_n, i)}{\partial w_{jk}^l} = -\frac{h}{\alpha} \sum_{i=1-o}^0 \beta_i [\bar{y}(t_n) - \hat{y}^i(t_n)] \bullet \frac{\partial \hat{f}[y(t_{n-1-i})]}{\partial w_{jk}^l} \quad (17)$$

A equação (17) pode ser utilizada para determinar a estrutura matricial do jacobiano $\frac{\partial J^l(t, i)}{\partial w_{jk}^l}$, onde n_1 é o número total de neurônios na camada l , e n_{l-1} é o número total de neurônios na camada $l-1$, conforme a equação (18).

$$\left[\frac{\partial J^l(t, i)}{\partial \bar{w}^l} \right]_{n_1 \times n_{l-1}} = \begin{bmatrix} \frac{\partial J^l(t, i)}{\partial w_{11}^l} & \frac{\partial J^l(t, i)}{\partial w_{12}^l} & \dots & \frac{\partial J^l(t, i)}{\partial w_{1n_{l-1}}^l} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial J^l(t, i)}{\partial w_{n_11}^l} & \frac{\partial J^l(t, i)}{\partial w_{n_12}^l} & \dots & \frac{\partial J^l(t, i)}{\partial w_{n_1n_{l-1}}^l} \end{bmatrix} \quad (18)$$

O algoritmo da retropropagação ou do gradiente apenas exige que os pesos das conexões da rede *feedforward* sejam atualizados recursivamente, através da expressão (19), onde $l = 1, 2, \dots, s$, para i igual a i -ésima iteração do algoritmo do gradiente e s o número total das camadas da rede.

$$\bar{W}_{i+1}^l = \bar{W}_i^l - \alpha \frac{\partial J^l(t, i)}{\partial \bar{w}^l} \quad (19)$$

Graficamente, a definição de $\hat{f}[y(t_{n-1-i})]$ na forma vetorial é apresentada conforme mostrado, na Figura 2. Na ilustração,

n_e é o número total de estados, n_c o número total de controles, o a ordem do integrador, i o número de atrasos em relação ao horizonte $(n-1)$ e r o número de neurônios na camada interna. Por exemplo, para o integrador Adams-Bashforth de quarta ordem, tem-se $o=4$ e $i=3,2,1,0$.

Assim, se a combinação linear de quatro entradas atrasadas sobre a rede neural é dada por $F_{n-4}, F_{n-3}, F_{n-2}$ e F_{n-1} . É utilizado no cálculo do integrador de quarta ordem a fórmula $y(t_n) = y(t_{n-1}) + \frac{h}{24} \cdot [55 \cdot F_{n-4} - 59 \cdot F_{n-3} + 37 \cdot F_{n-2} - 9 \cdot F_{n-1}]$.

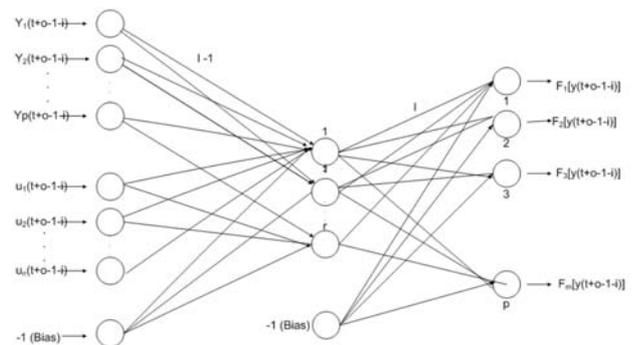


Figura 2: Rede neural Adams-Bashforth projetada com a função de derivadas instantâneas. Quanto maior a ordem o do integrador utilizado na fase de treinamento neural, mais preciso será o aprendizado da rede com relação as derivadas instantâneas do modelo empírico, para um horizonte genérico t_n .

A Figura 3 mostra o esquema gráfico para fazer a combinação numérica da estrutura do integrador com quatro entradas atrasadas, que neste caso são as quatro entradas simultâneas sobre a mesma rede *feedforward*. De forma explicativa $y(t)$ é o valor padrão de treinamento obtido *off line* pelo integrador numérico de elevada precisão, utilizado para resolver o sistema real resultando na próxima propagação $y(t_1)$, que é o valor exato do sistema para aquela entrada de dados e $\hat{y}(t_1)$ é a resposta do sistema dinâmico discretizado pelo conjunto integrador e rede. É importante ressaltar que \hat{g} representa o treinamento da rede neural para aproximar a função de derivadas dado por \hat{f} que representa a saída da rede, \hat{w} é a matriz de peso estimado pela rede, β igual a (55, -59, 37, -9) e α igual a (24) representam coeficientes de integração numérica.

3.4 Simulação Neural

Na fase de simulação neural, os dados precisaram ser normalizados no cálculo dos coeficientes da estrutura de integração de Adams-Bashforth e desnormalizados na propagação do estado futuro.

A normalização dos dados ou dos padrões de treinamento

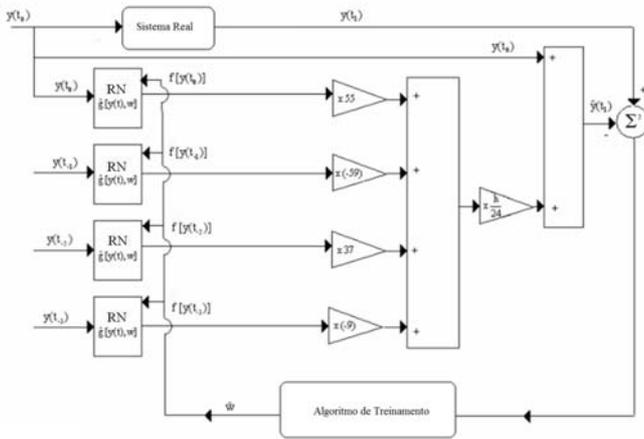


Figura 3: Interpretação do Algoritmo do Gradiente para o Integrador Neural Adams-Bashforth de Quarta Ordem.

apresentados para a rede *feedforward* através de um aprendizado supervisionado, é muito importante para evitar que a rede trabalhe com valores elevados dos pesos das conexões. Pois, além de dificultar o treinamento, pode gerar erros de arredondamento uma vez que o computador trabalha com sistema de ponto flutuante.

Na normalização divide-se os valores originais de entrada e saída pelos módulos dos valores máximos. Na desnormalização multiplica-se os valores originais de entrada e saída pelos módulos dos valores máximos. O fator de normalização dos padrões da entrada da rede é dado por f_{ent} e o fator de normalização dos padrões da saída da rede é calculado por f_{sai} .

O problema de normalização na metodologia de treinamento das derivadas instantâneas, representada pela estrutura de ordem o para os integradores Adams-Bashforth é expresso pela equação (20):

$$y_t(t_1) = y_t(t_0) + \frac{h}{\alpha} \sum_{i=1-o}^0 \beta_i f_t[y(t_i)], \text{ para } t = 1, 2, \dots, m \quad (20)$$

O problema de normalização e desnormalização é constituído pelas seguintes equações, para $t = 1, 2, \dots, m$ e $i = -(o-1), -(o-2), \dots, -2, -1, 0$. É importante esclarecer, que na utilização do sub-índice N de normalização o sub-índice t dos estados torna-se sobre-índice.

$$\hat{y}_N^t(t_1) = y_N^t(t_0) + \frac{h}{\alpha} \sum_{i=1-o}^0 \beta_i \hat{f}_N^t[y_N^t(t_i)] \quad (21)$$

$$\hat{y}_t(t_1) = y_t(t_0) + \frac{h}{\alpha} \sum_{i=1-o}^0 \beta_i \hat{f}_t[y(t_i)] \quad (22)$$

$$\hat{y}_N^t(t_i) = \frac{\hat{y}_t(t_i)}{f_{ent}^t} \quad (23)$$

$$y_N^t(t_1) = \frac{y_t(t_1)}{f_{sai}^t} \quad (24)$$

Substituindo as equações (21) e (22) em (23) e (24) para $t = 1, 2, \dots, m$, tem-se:

$$\hat{y}_t(t_1) = f_{sai}^t \cdot \{\hat{y}_N^t(t_1)\} \quad (25)$$

$$y_t(t_1) = f_{sai}^t \cdot \{y_N^t(t_0) + \frac{h}{\alpha} \sum_{i=1-o}^0 \beta_i \hat{f}_N^t[y_N^t(t_i, u_i)]\} \quad (26)$$

$$\hat{y}_t(t_1) = \frac{f_{sai}^t}{f_{ent}^t} \cdot y_t(t_n) + f_{sai}^t \frac{h}{\alpha} \sum_{i=1-o}^0 \beta_i \hat{f}_N^t \left[\frac{y(t_i)}{f_{ent(m)}}, \frac{u(t_i)}{f_{ent(n)}} \right] \quad (27)$$

para,

$$\frac{y(t_i)}{f_{ent(m)}} = \left(\frac{y_1(t_i)}{f_{ent}^1}, \frac{y_2(t_i)}{f_{ent}^2}, \dots, \frac{y_m(t_i)}{f_{ent}^m} \right) \quad (28)$$

$$\frac{u(t_i)}{f_{ent(n)}} = \left(\frac{u_1(t_i)}{f_{ent}^{m+1}}, \frac{u_2(t_i)}{f_{ent}^{m+2}}, \dots, \frac{u_n(t_i)}{f_{ent}^{m+n}} \right) \quad (29)$$

$$f_{ent} = [f_{ent(m)} : f_{ent(n)}] \quad (30)$$

As expressões (28) até (30) desnormalizam os valores normalizados de saída da estrutura de integração Adams-Bashforth de ordem o . Observe que, neste caso, não é possível obter explicitamente a relação existente entre $\hat{f}_N^t[y_N^t(t_i)]$ e $\hat{f}_t[y_t(t_i)]$ para $t = 1, 2, \dots, m$.

4 APLICAÇÃO PRÁTICA E ANÁLISE DOS RESULTADOS

4.1 Plataforma

Para a implementação da metodologia proposta foi utilizado o software e a linguagem MATLAB versão 7.0 utilizando o

Sistema Operacional Windows XP. A configuração do equipamento utilizado é descrita a seguir:

- Processador AMD Athlon XP 2800;
- Velocidade de 2.08 GHz;
- Memória RAM de 1 Gigabyte (GB); e
- HD de 120 GB.

4.2 O Protótipo

Na Figura 4 apresenta-se de forma explicativa a evolução dos trabalhos envolvendo integradores neurais. A principal contribuição deste trabalho, bem como as demais implementações encontram-se em negrito. O foco principal é a implementação da *metodologia empírica* com derivadas instantâneas aplicadas a sistemas dinâmicos que trabalham somente com variáveis de estado, ou que trabalham com variáveis de estado e de controle. A metodologia empírica com derivadas médias também é explorada.

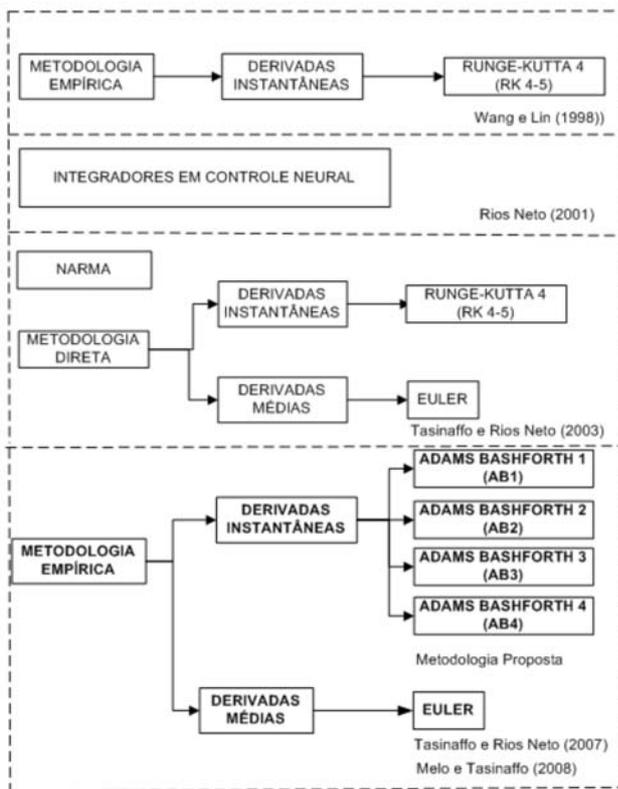


Figura 4: Evolução das Metodologias Envolvendo Integradores Neurais.

Em todas as simulações apresentadas neste artigo utiliza-se como referência da precisão desejada a função de derivada

teórica calculada pelo integrador numérico Runge-Kutta de quarta ordem (RK-4,5), que é um integrador estável e simples. Na maioria dos treinamentos os padrões de treinamentos são normalizados. Na fase de simulação as condições iniciais são atualizadas da seguinte forma: a cada 30 passos de propagações quando o passo de integração é igual a 0,1; a cada 300 passos de propagações quando o passo de integração é igual a 0,01 e a cada 3.000 passos de propagações quando o passo de integração é igual a 0,001. O erro máximo a ser alcançado pelo treinamento da rede igual a 10^{-8} , pois com um erro dessa magnitude as simulações de redes treinadas foram quase perfeitas. A porcentagem de padrões de treinamento deixados para validar ou testar o aprendizado da rede igual a 20 por cento e 80 por cento dos padrões foram utilizados no treinamento da rede.

Com o objetivo de validar o protótipo implementado, nas outras seções são apresentados, em detalhes, três estudos de caso. O primeiro estudo de caso é o pêndulo simples não-linear, um sistema simples que propicia um treinamento supervisionado rápido de forma que o refinamento de possíveis problemas nas implementações são corrigidos rapidamente. Já no segundo e terceiro estudos de caso, considera-se sistemas dinâmicos de maior complexidade, permitindo fazer uma análise mais precisa do desempenho da metodologia proposta.

4.3 Primeiro Estudo de Caso: Pêndulo Simples Não-Linear

O pêndulo simples não-linear é um sistema autônomo de segunda ordem. O modelo matemático pode ser descrito pela equação de segunda ordem a seguir:

$$l\ddot{\Theta} + g \cdot \sin\Theta = 0 \quad (31)$$

Transformando em equações de primeira ordem, descrito por (32) e (33) (Ogata, 1967):

$$\dot{\Theta} = w \quad (32)$$

$$\dot{w} = -\frac{g}{l} \sin\Theta \quad (33)$$

Na definição dos domínios dos valores limites das variáveis de estado, são estabelecidos:

- O ângulo Θ , uma variação no intervalo de $[-\frac{\pi}{2}, \frac{\pi}{2}]$ (rad); e
- A velocidade angular (w), uma variação no intervalo de $[-6, +6]$ ($\frac{\text{rad}}{\text{s}}$).

Para definição dos valores limites das constantes, são estabelecidos:

- A aceleração da gravidade g , será considerado um valor de $9,81 \left(\frac{m}{s^2}\right)$; e
- O comprimento do pêndulo l , será considerado um valor de $0,30$ (m).

Na realização do treinamento da rede neural ajustou-se os seguintes parâmetros:

- Ajuste empírico do número total de neurônios da camada interna igual a 20;
- A taxa de aprendizado do algoritmo de treinamento do gradiente igual a 0,02 na metodologia empírica e 0,01 na metodologia direta, porque a metodologia empírica exige um processamento maior que a metodologia direta;
- O valor de lambda (λ) das funções de ativação sigmoideal da camada interna igual a 2. O lambda é um parâmetro escalar numérico responsável pela inclinação da curva da função; e
- Condições iniciais igual a 500.

4.3.1 Análise dos Resultados

• Metodologia Direta com Derivadas Instantâneas

Nesta simulação, os integradores neurais de Runge-Kutta e Adams-Bashforth acompanham com precisão a trajetória de solução do sistema. Este resultado está relacionado diretamente com a precisão do integrador numérico e com a diminuição do erro quadrático médio de treinamento neural, pois como regra geral, a precisão de um integrador numérico aumenta com a diminuição do passo de integração (Δt). Na Figura 5, a média dos erros quadráticos de treinamento e teste alcançados foram, respectivamente, $3,6724 \cdot 10^{-5}$ e $4,3496 \cdot 10^{-5}$, o erro da média das normas das propagações é de $0,38737$.

• Metodologia Direta com Derivadas Médias

Na Figura 6, o integrador neural de Euler e de Runge-Kutta acompanham com precisão a trajetória de solução do sistema. Como esperado, a utilização do integrador numérico simples de Euler sobre a função de derivadas instantâneas teórica diverge da trajetória de referência, pois o passo de integração utilizado demonstrou-se muito grande para este

integrador simples de primeira ordem. Analisando ainda a Figura 6, esta metodologia é de passo fixo, não permitindo variação do passo de integração entre o treinamento e a simulação para um $\Delta t=0,1$. A média dos erros quadráticos de treinamento e teste alcançados foram, respectivamente, $1,2017 \cdot 10^{-5}$ e $1,3912 \cdot 10^{-5}$. O erro da média das normas das propagações igual a $0,33913$.

• Metodologia Empírica com Derivadas Instantâneas

Na Figura 7 apresenta-se o resultado da simulação da obtenção das derivadas instantâneas utilizando o integrador Adams-Bashforth de primeira ordem. Nesta metodologia empírica as redes neurais utilizaram um $\Delta t=0,01$. A média dos erros quadráticos de treinamento e teste alcançados foram, respectivamente, $3,0298 \cdot 10^{-7}$ e $3,2128 \cdot 10^{-7}$. O erro da média das normas das propagações igual a $1,395$. Os integradores neurais de Adams-Bashforth de primeira ordem alcança uma precisão similar ao integrador de alta ordem uma vez que utilizam a função de derivadas média neural. Neste exemplo, demonstra-se a importância de escolher um passo de integração adequado para o Adams-Bashforth, visando não distorcer a função de derivadas que se deseja aprender. Valores muito grandes para Δt resultará no aprendizado de uma função de derivadas médias muito distante das derivadas instantâneas. Valores muito pequenos também poderão não resultar numa aproximação precisa para as derivadas instantâneas, uma vez que o integrador utilizado no exemplo da Figura 7 é de baixa ordem.

Na Figura 8 apresenta-se o resultado da simulação da obtenção das derivadas instantâneas utilizando o integrador Adams-Bashforth de segunda ordem. A arquitetura desta rede neural é resultado do arranjo de duas combinações lineares sobre a mesma rede neural. A média dos erros quadráticos de treinamento e teste alcançados foram, respectivamente, $3,1188 \cdot 10^{-7}$ e $2,9921 \cdot 10^{-7}$. O erro da média das normas das propagações igual a $2,0728$. Numa estrutura de integração de Adams-Bashforth de segunda ordem a rede neural não consegue aprender a função de derivadas instantânea de forma precisa.

Na Figura 9 apresenta-se o resultado da simulação da obtenção das derivadas instantâneas utilizando o integrador Adams-Bashforth de terceira ordem. A arquitetura desta rede neural é resultado do arranjo de três combinações lineares sobre a mesma rede neural. A média dos erros quadráticos de treinamento e teste alcançados foram, respectivamente, $5,7112 \cdot 10^{-7}$ e $5,9581 \cdot 10^{-7}$. O erro da média das normas das propagações igual a $2,6142$. Numa estrutura de integração de Adams-Bashforth de terceira ordem a rede neural ainda não consegue aprender a função de derivadas instantânea de forma precisa.

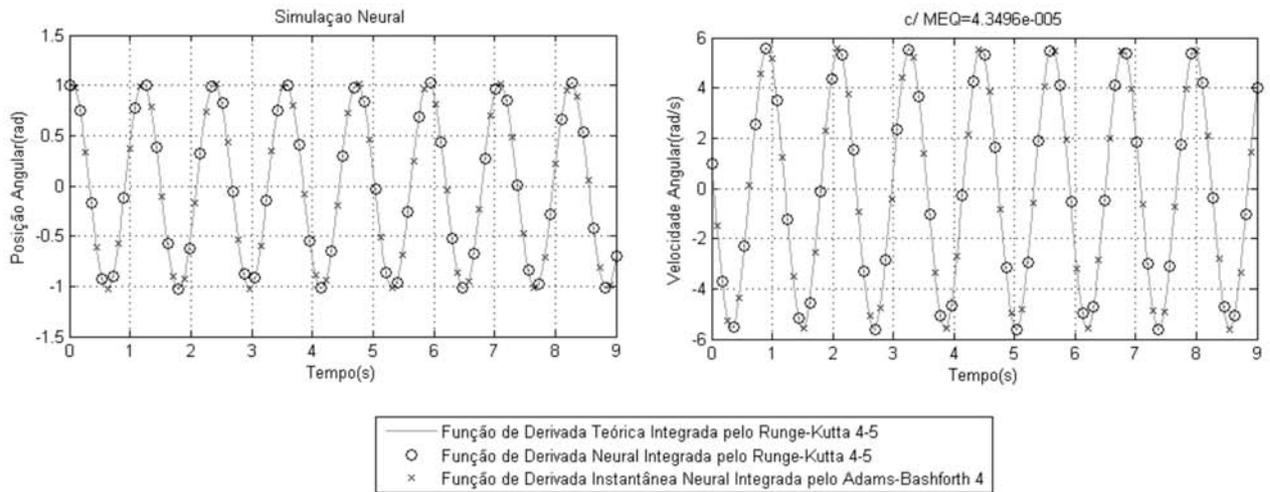


Figura 5: Simulação Obtida da Metodologia Direta das Derivadas Instantâneas pelo IN de Adams-Bashforth de Quarta Ordem ($\Delta t=0,01$).

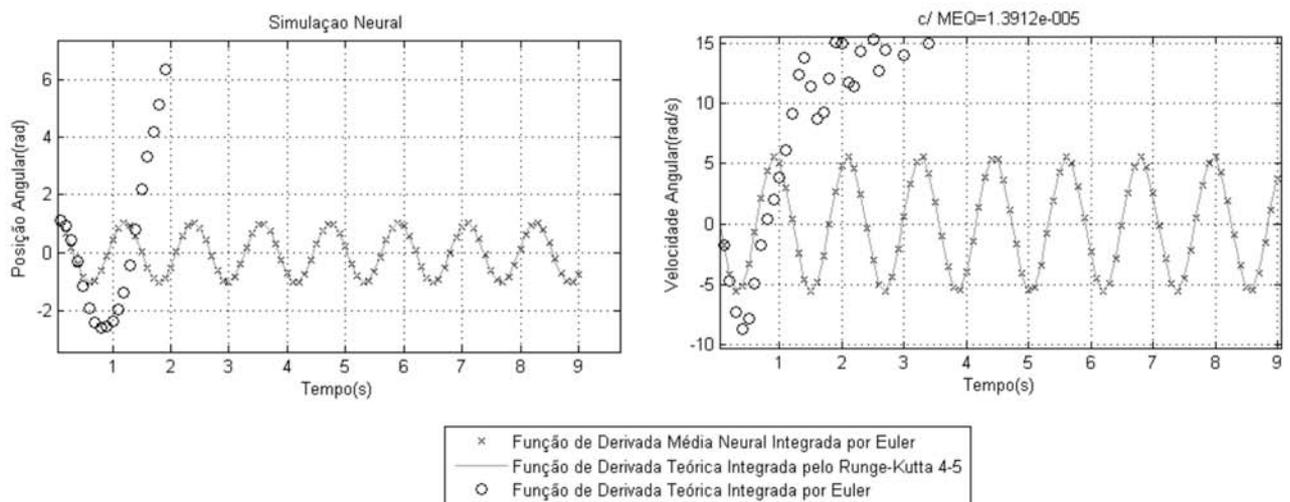


Figura 6: Simulação Obtida da Metodologia Direta das Derivadas Médias pelo IN de Euler ($\Delta t=0,1$).

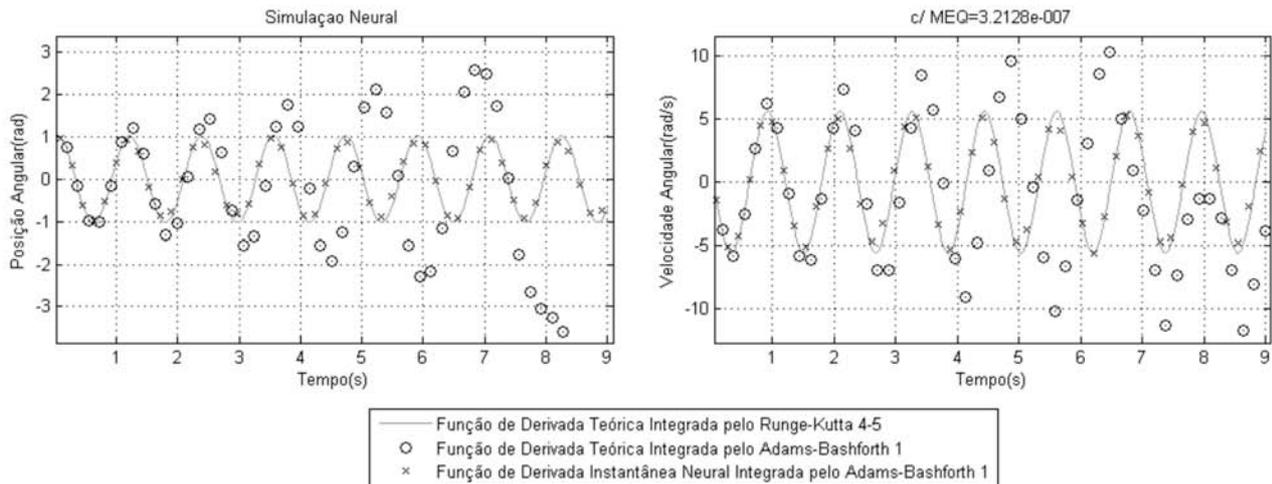


Figura 7: Simulação Obtida da Metodologia Empírica das Derivadas Instantâneas pelo IN de Adams-Bashforth de Primeira Ordem ($\Delta t=0,01$).

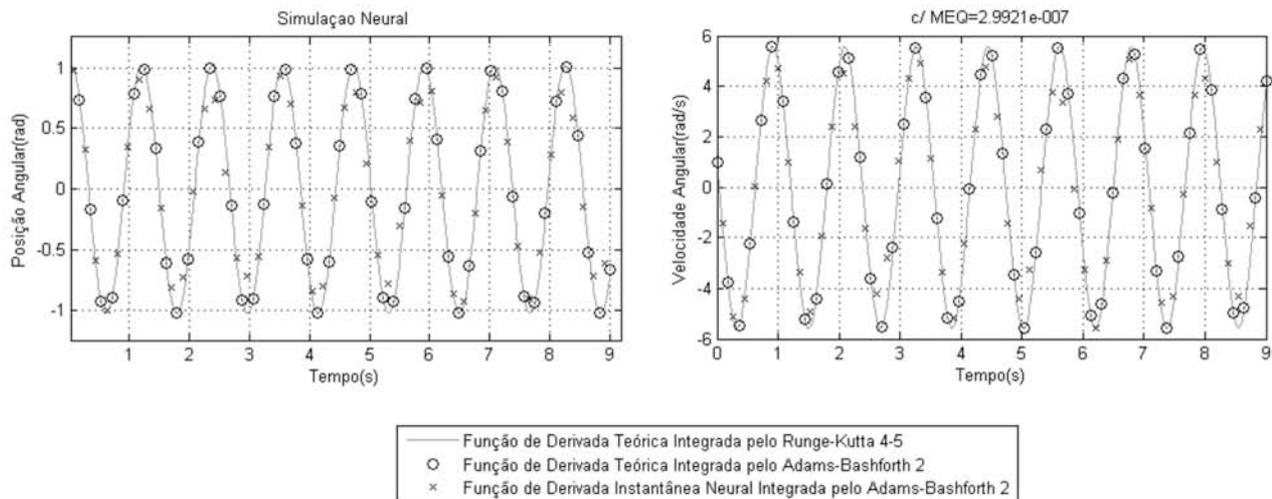


Figura 8: Simulação Obtida da Metodologia Empírica das Derivadas Instantâneas pelo IN de Adams-Bashforth de Segunda Ordem ($\Delta t=0,01$).

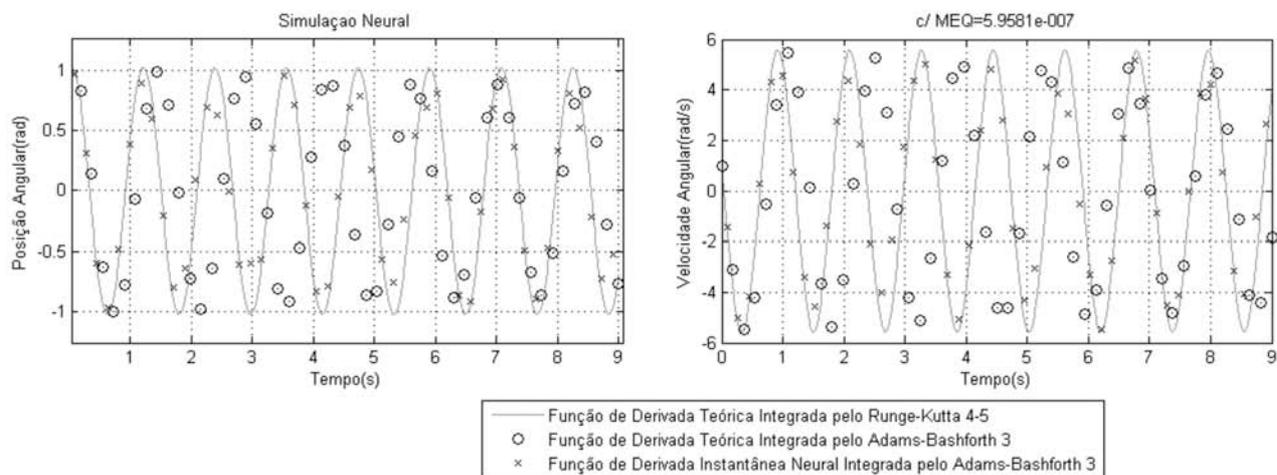


Figura 9: Simulação Obtida da Metodologia Empírica das Derivadas Instantâneas pelo IN de Adams-Bashforth de Terceira Ordem ($\Delta t=0,01$).

Na Figura 10 apresenta-se o resultado da simulação da obtenção das derivadas instantâneas utilizando o integrador Adams-Bashforth de quarta ordem. A arquitetura desta rede neural é resultado do arranjo de quatro combinações lineares sobre a mesma rede neural. A média dos erros quadráticos de treinamento e teste alcançados foram, respectivamente $4,8413 \cdot 10^{-7}$ e $8,1719 \cdot 10^{-7}$. O erro da média das normas das propagações igual a 1,2816. O integrador neural Adams-Bashforth de quarta ordem e o integrador numérico de alta ordem de Runge-Kutta obtiveram resultados bem-sucedidos. Como o integrador numérico de Adams-Bashforth é de quarta ordem, consegue-se aproximar da solução mesmo utilizando a derivada teórica.

Em princípio a metodologia empírica das derivadas instantâneas aplicada a integradores de alta ordem, por exemplo o Adams-Bashforth de quarta ordem deveria permitir alterar o passo de integração, no sentido de sempre melhorar a solução numérica discreta com a diminuição do passo de integração. Entretanto, isto não acontece na simulação da Figura 11.

As equações da normalização apresentadas na seção 3.3 impedem a desnormalização adequada das saídas geradas em simulações com Δt diferente daquele que foi utilizado na fase de treinamento conforme o postulado 1. A aplicação da heurística 1, é comprovada quando utiliza-se a normalização dos padrões de treinamento na Figura 10 se tornando inviável a diminuição do passo de integração na simulação da Figura 11.

Postulado 1: Na metodologia empírica é impossível desnormalizar de forma única as funções de derivadas instantâneas para integradores de múltiplos passos com ordem maior que 1 (um). Somente a desnormalização global da saída $y(t)$ pode

ser realizada.

Heurística 1: Na metodologia empírica o passo de integração é fixo, se normalizados os integradores de ordem maior ou igual a 2 (dois) para integradores de múltiplos passos. Para variar o passo de integração em princípio não é possível normalizar os padrões de treinamento diretamente. Assim, formula-se as seguintes heurísticas:

1. Se normalizado os padrões de treinamento nas estruturas de integração neural com ordem maior ou igual a 2 (dois), não é possível variar o passo de integração na fase de simulação neural.
2. Se desnormalizados os padrões de treinamento sobre as mesmas condições de 1, então poder-se-á variar o passo de integração na simulação neural.

Na seção 4.2 estipula-se que para um $\Delta t=0,001$ deveria-se atualizar as condições iniciais a cada 3.000 passos, mas com o objetivo de melhor visualizar os resultados na Figura 11 as atualizações ocorreram a cada 300 iterações. O erro da média das normas das propagações igual a 827461,9502.

Ainda em relação a aplicação da heurística 1, testa-se o mesmo estudo de caso em um novo treinamento do Adams-Bashforth de quarta ordem *sem normalizar* os padrões de treinamento visando obter resultados encontrados conforme o aumento do passo de integração na Figura 12 e a diminuição do passo na Figura 13. A média dos erros quadráticos de treinamento e teste alcançados foram, respectivamente, $9,3932 \cdot 10^{-6}$ e $6,6256 \cdot 10^{-6}$, pois decorrem do mesmo treinamento, que utilizou um passo de integração de 0,01. O erro

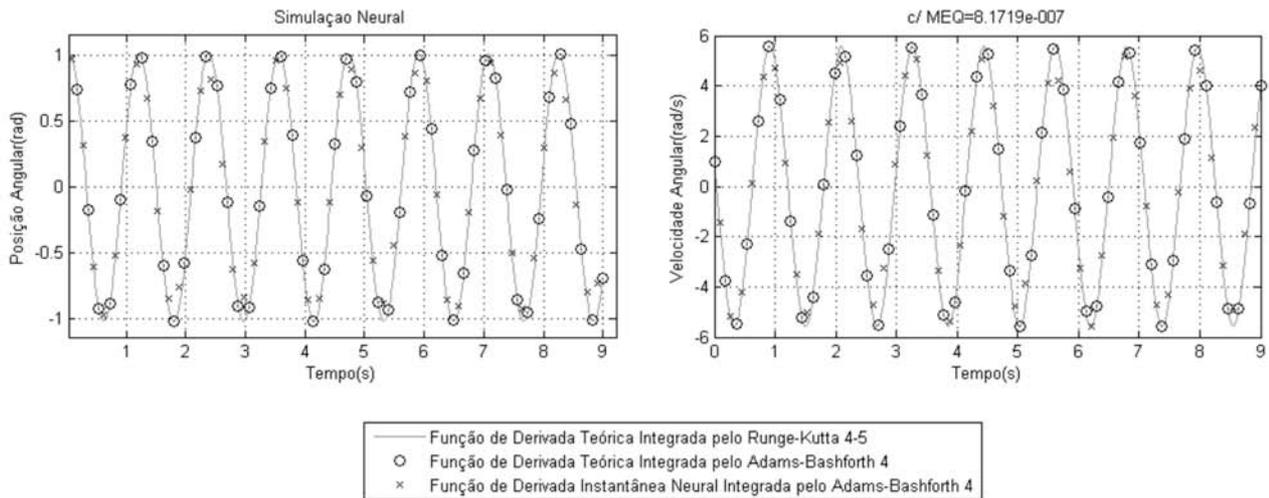


Figura 10: Simulação Obtida da Metodologia Empírica das Derivadas Instantâneas pelo IN de Adams-Bashforth de Quarta Ordem ($\Delta t=0,01$).

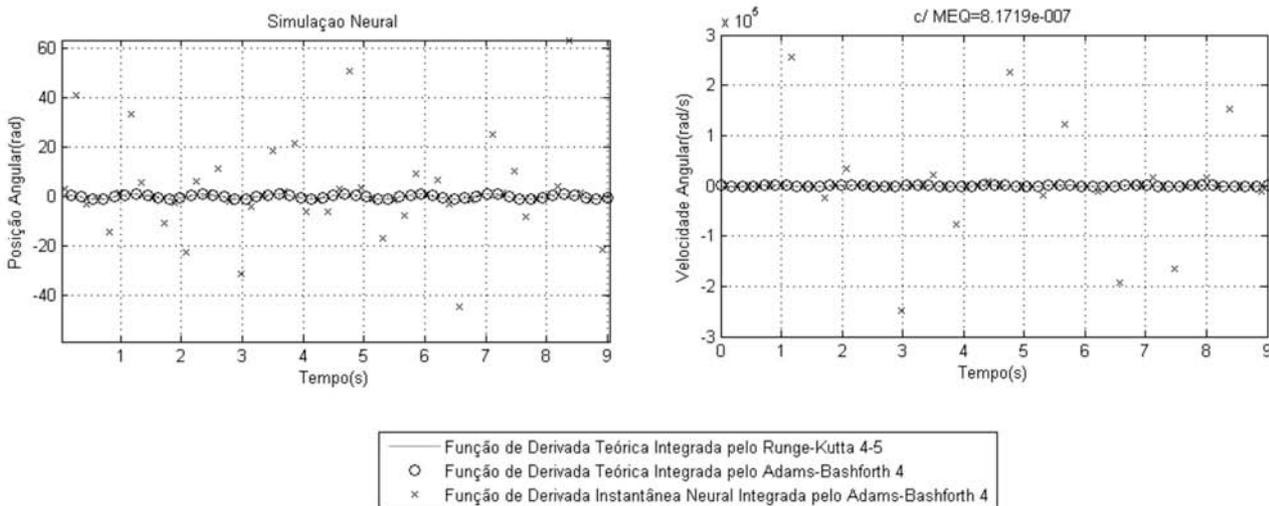


Figura 11: Simulação Obtida da Metodologia Empírica das Derivadas Instantâneas pelo IN de Adams-Bashforth de Quarta Ordem ($\Delta t=0,001$).

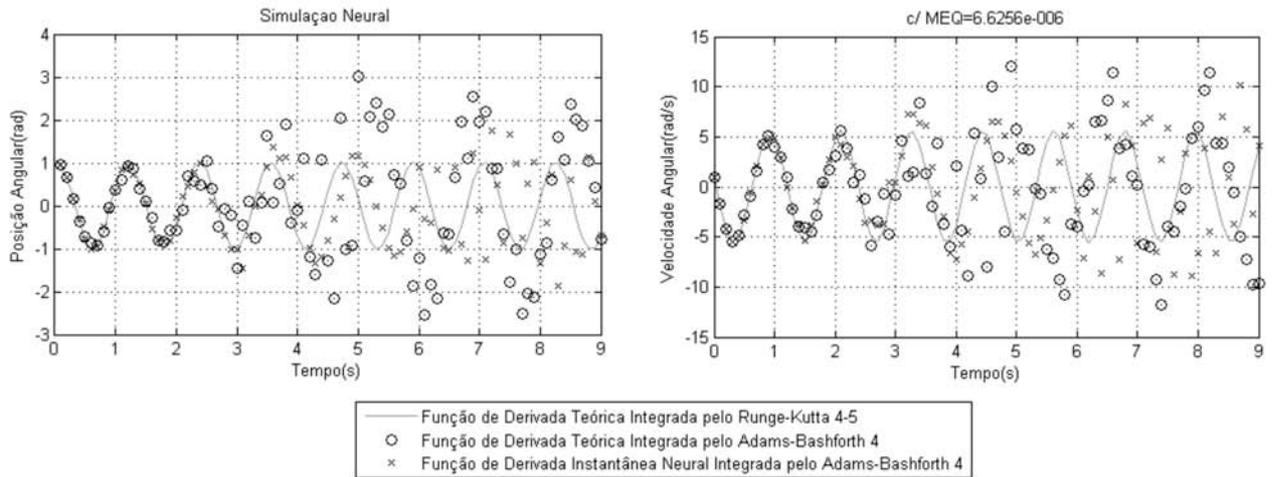


Figura 12: Simulação Obtida da Metodologia Empírica das Derivadas Instantâneas pelo IN de Adams-Bashforth de Quarta Ordem ($\Delta t=0,1$).

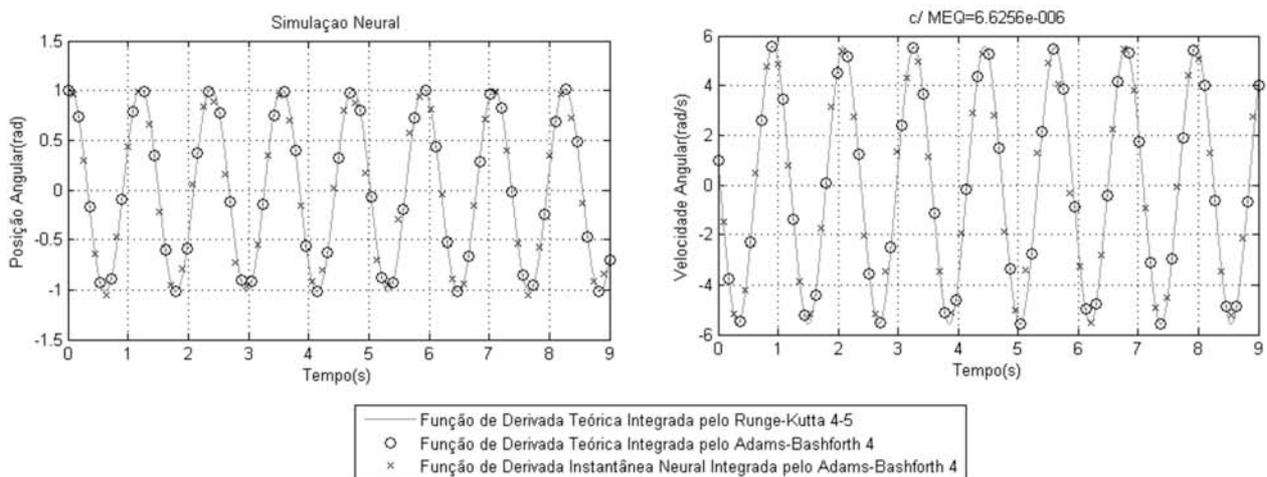


Figura 13: Simulação Obtida da Metodologia Empírica das Derivadas Instantâneas pelo IN de Adams-Bashforth de Quarta Ordem ($\Delta t=0,001$).

da média das normas das propagações na Figura 12 igual a 14, 1571 e na Figura 13 de 1, 9218.

- **Metodologia Empírica com Derivadas Médias**

Na Figura 14 apresenta-se o resultado da simulação obtida das derivadas médias utilizando o integrador Euler de primeira ordem. A rede foi treinada e simulada com passo de integração 0,1. É evidente que o integrador neural de Euler alcança uma precisão equivalente ao integrador de Runge-Kutta de alta precisão. A utilização do integrador numérico simples de Euler diverge da trajetória de referência, pois conforme visto nas Figuras 6 e 7, os métodos de primeira ordem são numericamente menos eficientes. A média dos erros quadráticos de treinamento e teste alcançados durante o treinamento e teste foram, respectivamente, $6,4667 \cdot 10^{-6}$ e $1,0168 \cdot 10^{-5}$. O erro da média das normas das propagações igual a 0,55325.

4.4 Segundo Estudo de Caso: Sistema Dinâmico Abstrato

O segundo estudo de caso é proposto com o intuito de validar a metodologia empírica para sistema que trabalham com variáveis de estado e de controle. Considere o sistema dinâmico abstrato não linear adaptado de Vidyasagar (1978), com duas variáveis de estado e uma variável de controle, descrito pelas equações 34 e 35.

$$\dot{y}_1 = 10^{-1} \cdot y_1 + y_2^2 - \left\{ \frac{2}{1 + \exp[-\frac{1}{2}(2y_2 - u)]} - 1 \right\} \quad (34)$$

$$\dot{y}_2 = \sin y_1 - y_2 + u \quad (35)$$

As variáveis de estado deste problema são: y_1, y_2 . A variável de controle é dada por u .

Na definição dos domínios dos valores limites das variáveis de estado, são estabelecidos:

- \dot{y}_1 , uma variação no intervalo de [-6,+6]; e
- \dot{y}_2 , uma variação no intervalo de [-2,+2].

Para definição dos valores limites da variável de controle, é estabelecida:

- u , foi gerada aleatoriamente no intervalo de [-3,+3].

4.4.1 Análise dos Resultados

Na implementação deste protótipo, o modelo teórico é constituído por variáveis de estado e variáveis de controle. A construção da arquitetura da rede *feedforward* é dada por 3 entradas, sendo 2 variáveis de estado e 1 de controle, 41 neurônios na camada intermediária foram obtidos de forma experimental e empírica e 2 neurônios da camada de saída equivalente as 2 variáveis de estado.

Na realização do treinamento da rede neural ajustou-se vários parâmetros, tais como:

- Ajuste empírico do número total de neurônios da camada interna igual a 41;
- A taxa de aprendizado do algoritmo de treinamento do gradiente igual a 0,02;
- O valor de lambda (λ) das funções de ativação tangente hiperbólica sigmóide(tansig) da camada interna igual a 2; e
- Condições iniciais igual a 1.400.

- **Metodologia Direta com Derivadas Instantâneas**

É apresentado os resultados da simulação da obtenção das derivadas instantâneas utilizando o integrador neural de Runge-Kutta de quarta ordem. Na Figura 15 obteve-se resultados precisos quando a rede foi simulada com passo de integração igual a 0,001, e o erro da média das normas das propagações é 0,33341.

- **Metodologia Direta com Derivadas Médias**

Na Figura 16, ilustra-se a simulação da obtenção das derivadas médias utilizando o integrador de Euler de primeira ordem. A média dos erros quadráticos de treinamento e teste alcançados foram, respectivamente, $3,5926 \cdot 10^{-6}$ e $4,3952 \cdot 10^{-6}$, e o erro da média das normas das propagações igual a 0,17114. Constata-se que o integrador neural implementado consegue um nível de precisão equivalente ao integrador de alta complexidade, como o Runge-Kutta de quarta ordem. A utilização do integrador numérico simples de Euler sobre a função de derivadas instantânea teórica neste caso, converge com a trajetória de referência em virtude da simplicidade do sistema.

- **Metodologia Empírica com Derivadas Instantâneas**

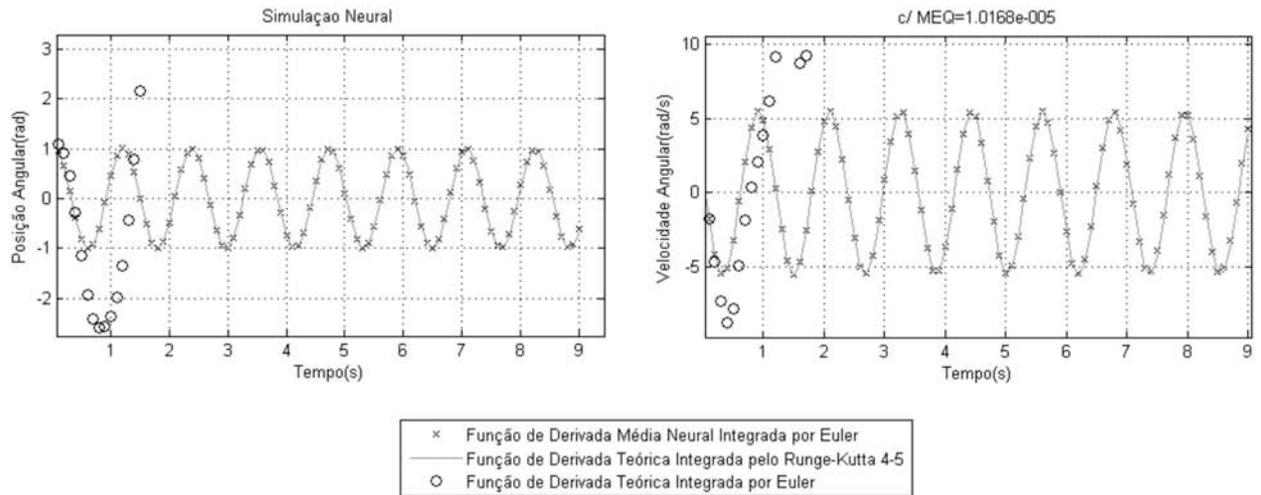


Figura 14: Simulação Obtida da Metodologia Empírica das Derivadas Médias pelo IN de Euler ($\Delta t=0,1$).

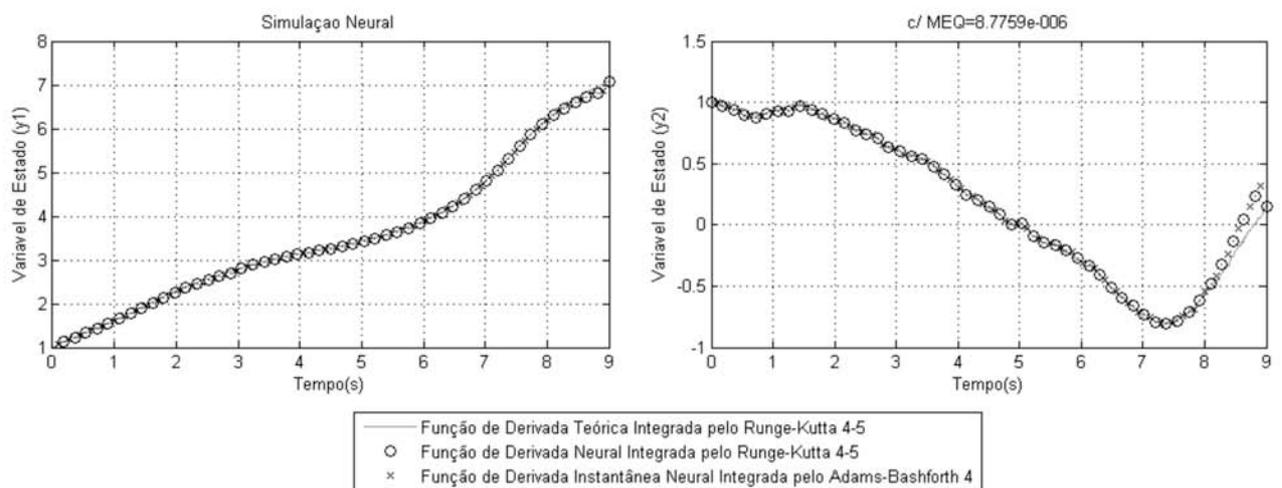


Figura 15: Simulação Obtida da Metodologia Direta das Derivadas Instantâneas pelo IN de Adams-Bashforth de Quarta Ordem ($\Delta t=0,001$).

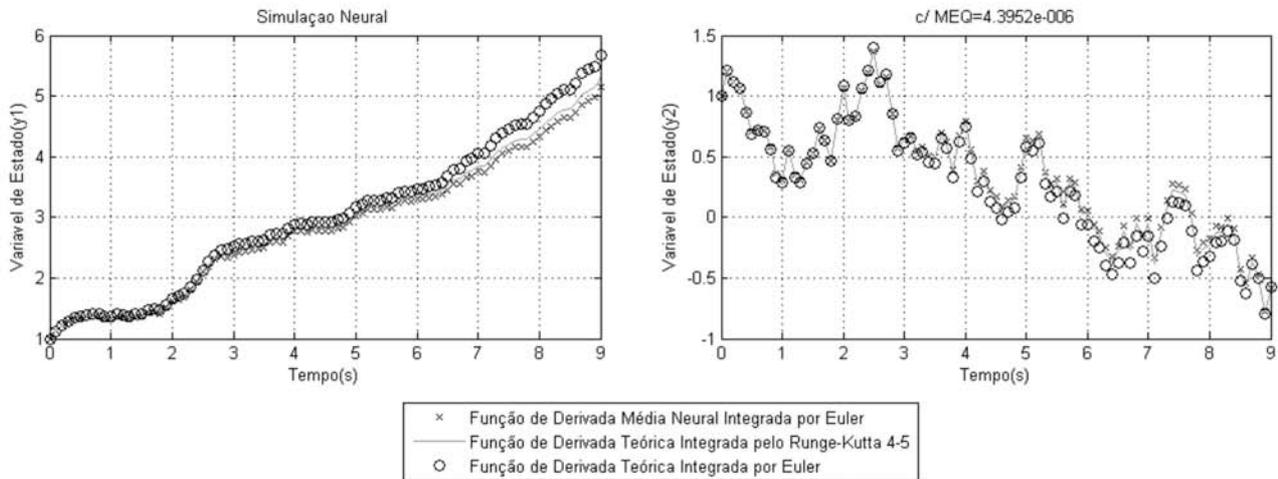


Figura 16: Simulação Obtida da Metodologia Direta das Derivadas Médias pelo IN de Euler ($\Delta t=0,1$).

Na Figura 17 vê-se o resultado da simulação da obtenção das derivadas instantâneas utilizando o integrador Adams-Bashforth de primeira ordem. Nesta metodologia empírica as redes neurais utilizaram um $\Delta t=0,01$. A média dos erros quadráticos de treinamento e teste alcançados foram, respectivamente, $1,983.10^{-7}$ e $2,2157.10^{-7}$ alcançando um erro da média das normas das propagações igual a 1,2338.

Nos treinamentos das derivadas instantâneas utilizando o integrador Adams-Bashforth de segunda, terceira ordem e quarta ordem. A média dos erros quadráticos de treinamento e teste alcançados foram insuficientes para obtenção de uma simulação que acompanhasse com exatidão a trajetória de solução numérica do sistema. Pois, estes treinamentos tem um custo computacional grande exigindo maior tempo de processamento e implicaria em um aumento do erro para a ordem de 10^{-7} . Nestes três casos tornaram-se inviáveis do ponto de vista do hardware utilizado.

• **Metodologia Empírica com Derivadas Médias**

Na Figura 18 apresenta-se o resultado da simulação obtida das derivadas médias utilizando o integrador de Euler de primeira ordem. A média dos erros quadráticos de treinamento e teste alcançados foram, respectivamente, $1,6556.10^{-6}$ e $1,6379.10^{-6}$, e o erro da média das normas das propagações igual a 0,20285. As conclusões foram equivalentes a da metodologia direta conforme apresentado na Figura 14.

4.5 Terceiro Estudo de Caso: Sistema de Van der Pol

Considere o sistema de Van der Pol descrito pela equação de segunda ordem a seguir.

$$\ddot{y} - (1 - y^2)\dot{y} + y = 0 \tag{36}$$

Transformando em equações de primeira ordem, descrito por 37 e 38:

$$\dot{y}_1 = y_2 \tag{37}$$

$$\dot{y}_2 = (1 - y_1^2)y_2 + y_1 \tag{38}$$

Na definição dos domínios dos valores limites das variáveis de estado, são estabelecidos:

- \dot{y}_1 , uma variação no intervalo de $[-2,2, +2,2]$; e
- \dot{y}_2 , uma variação no intervalo de $[-2,8,+3,5]$.

Na realização do treinamento da rede neural ajustou-se os seguintes parâmetros:

- Ajuste empírico do número total de neurônios da camada interna igual a 20;
- A taxa de aprendizado do algoritmo de treinamento do gradiente igual 0,1;

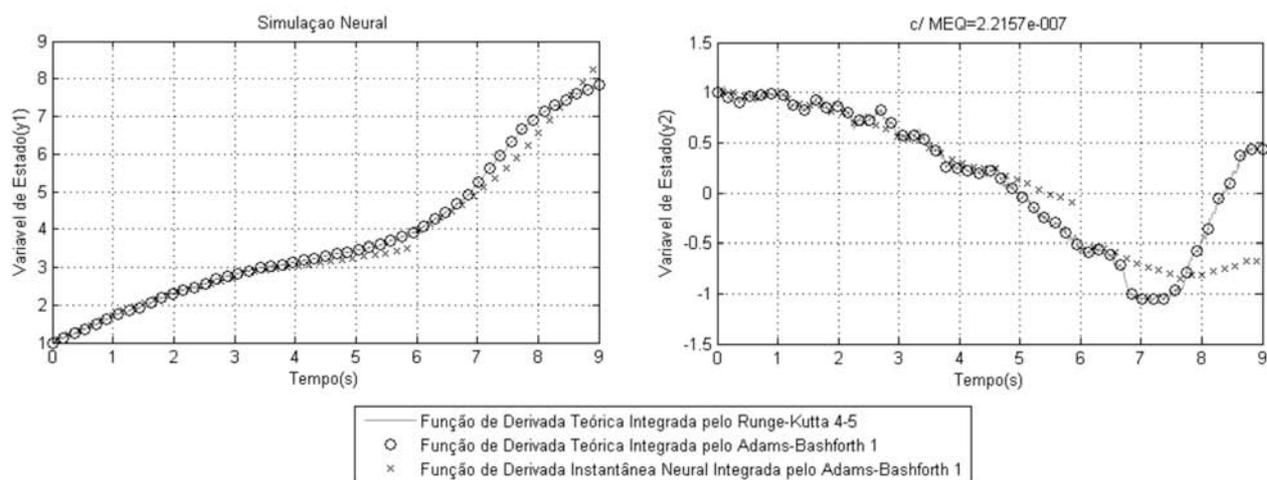


Figura 17: Simulação Obtida da Metodologia Empírica das Derivadas Instantâneas pelo IN de Adams-Bashforth de Primeira Ordem ($\Delta t=0,01$).

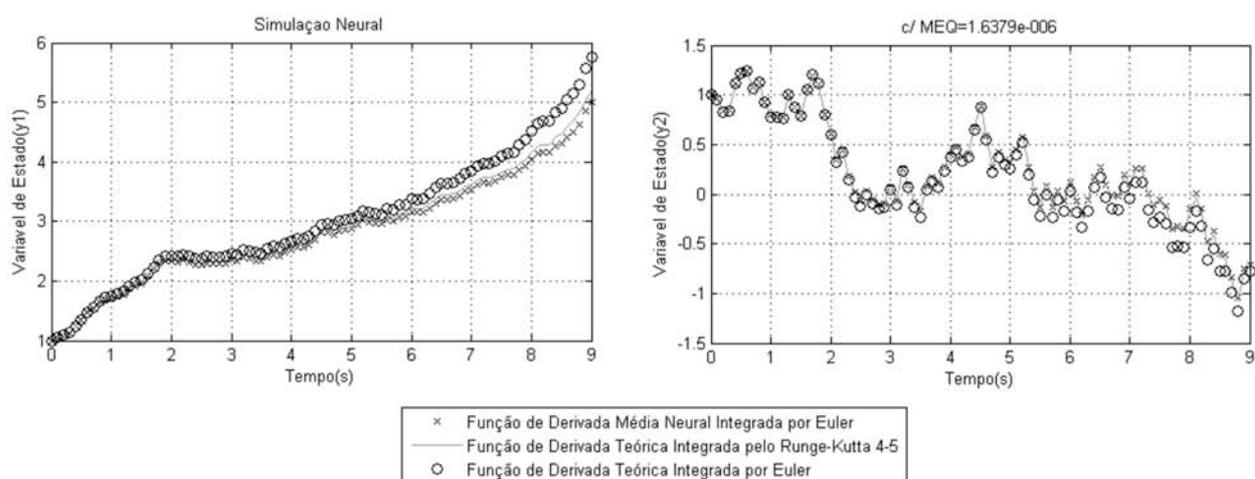


Figura 18: Simulação Obtida da Metodologia Empírica das Derivadas Médias pelo IN de Euler ($\Delta t=0,1$).

- O valor de λ das funções de ativação tangente hiperbólica sigmóide(tansig) da camada interna igual a 2; e
- Condições iniciais do sistema igual a 500.

4.5.1 Análise dos Resultados

• Metodologia Direta com Derivadas Médias

Na Figura 19, ilustra-se a simulação da obtenção das derivadas médias utilizando o integrador de Euler de primeira ordem. A média dos erros quadráticos de treinamento e teste alcançados foram, respectivamente, $9,5186 \cdot 10^{-7}$ e $9,0802 \cdot 10^{-7}$, e o erro da média das normas das propagações igual a 0,24279. Constatou-se que o integrador neural de Euler implementado consegue um nível de precisão equivalente ao integrador de alta complexidade, como o Runge-Kutta de quarta ordem, mas como esperado integrador teórico de Euler é menos eficiente.

• Metodologia Empírica com Derivadas Instantâneas

Na Figura 20, ilustra-se a simulação da obtenção das derivadas médias utilizando o integrador de Adams-Bashforth de primeira ordem. Nesta metodologia empírica as redes neurais utilizaram um $\Delta t=0,1$. A média dos erros quadráticos de treinamento e teste alcançados foram, respectivamente, $2,9806 \cdot 10^{-7}$ e $5,3989 \cdot 10^{-7}$, e o erro da média das normas das propagações igual a 0,059473. Nesta simulação o integrador neural de Adams-Bashforth de primeira ordem alcança uma precisão equivalente ao integrador teórico de Runge-Kutta. O integrador teórico de Adams-Bashforth de primeira ordem não acompanha adequadamente a trajetória da solução.

Na Figura 21, ilustra-se a simulação da obtenção das derivadas médias utilizando o integrador de Adams-Bashforth de segunda ordem. A média dos erros quadráticos de treinamento e teste alcançados foram, respectivamente, $4,1444 \cdot 10^{-7}$ e $4,7545 \cdot 10^{-7}$, e o erro da média das normas das propagações igual a 0,072186. Nesta simulação o integrador neural de Adams-Bashforth e o integrador teórico de Adams-Bashforth de segunda ordem alcançam uma precisão equivalente ao integrador teórico de Runge-Kutta.

Na Figura 22, ilustra-se a simulação da obtenção das derivadas médias utilizando o integrador de Adams-Bashforth de terceira ordem. A média dos erros quadráticos de treinamento e teste alcançados foram, respectivamente, $7,268 \cdot 10^{-7}$ e $6,9463 \cdot 10^{-7}$, e o erro da média das normas das propagações igual a 0,071247. Nesta simulação o integrador neural de Adams-Bashforth de terceira ordem alcança

uma precisão equivalente ao integrador teórico de Runge-Kutta. O integrador teórico de Adams-Bashforth de terceira ordem não acompanha adequadamente a trajetória da solução.

Na Figura 23, ilustra-se a simulação da obtenção das derivadas médias utilizando o integrador de Adams-Bashforth de quarta ordem. A média dos erros quadráticos de treinamento e teste alcançados foram, respectivamente, $5,4579 \cdot 10^{-7}$ e $8,1736 \cdot 10^{-7}$, e o erro da média das normas das propagações igual a 0,050604. Nesta simulação o integrador neural de Adams-Bashforth de quarta ordem alcança uma precisão equivalente ao integrador teórico de Runge-Kutta. O integrador teórico de Adams-Bashforth de quarta ordem não acompanha adequadamente a trajetória da solução.

• Metodologia Empírica com Derivadas Médias

Na Figura 24 apresenta-se o resultado da simulação da obtenção das derivadas médias utilizando o integrador de Euler de primeira ordem. A média dos erros quadráticos de treinamento e teste alcançados foram, respectivamente, $9,2606 \cdot 10^{-7}$ e $1,5481 \cdot 10^{-7}$, e o erro da média das normas das propagações igual a 0,050613. Constatou-se que o integrador neural de Euler implementado consegue um nível de precisão equivalente ao integrador de alta complexidade, como o Runge-Kutta de quarta ordem, mas como esperado integrador teórico de Euler é menos eficiente.

Na Tabela 1 apresenta-se as principais comparações quantitativas entre os métodos de treinamento neural, que são observações empíricas que se confirmam com novos treinamentos.

Em suma, na Tabela 2 apresenta-se as principais características do integrador neural desenvolvido em comparação com o integrador neural de Runge-Kutta e Euler encontrados na literatura.

5 CONCLUSÕES

Este artigo traz como principal contribuição o desenvolvimento de uma nova metodologia empírica de treinamento supervisionado ainda não desenvolvida na literatura capaz de identificar e representar sistemas dinâmicos do mundo real diminuindo o esforço computacional no treinamento neural.

Do ponto de vista científico, o trabalho realizado apresenta as seguintes contribuições, teóricas e práticas: identificação e representação de sistemas do mundo real; redução da complexidade das expressões algébricas na determinação do jacobiano, ao se trabalhar com uma combinação linear das retropropagações atrasadas do algoritmo *backpropagation* em relação ao trabalho de Wang and Lin (1998); e desenvolvimento de uma equação analítica para os integradores neurais

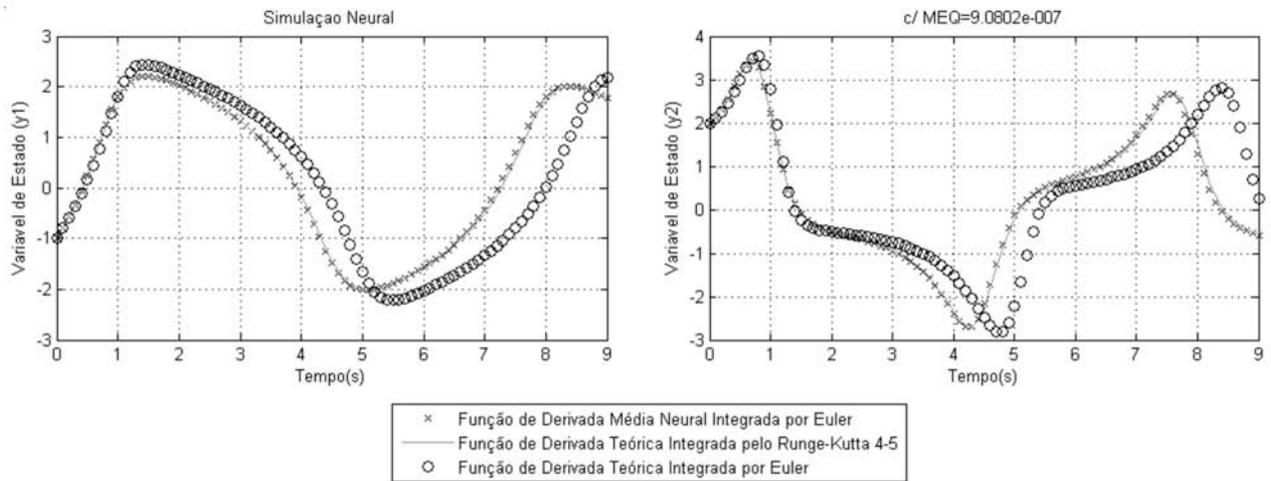


Figura 19: Simulação Obtida da Metodologia Direta das Derivadas Médias pelo IN de Euler ($\Delta t=0,1$).

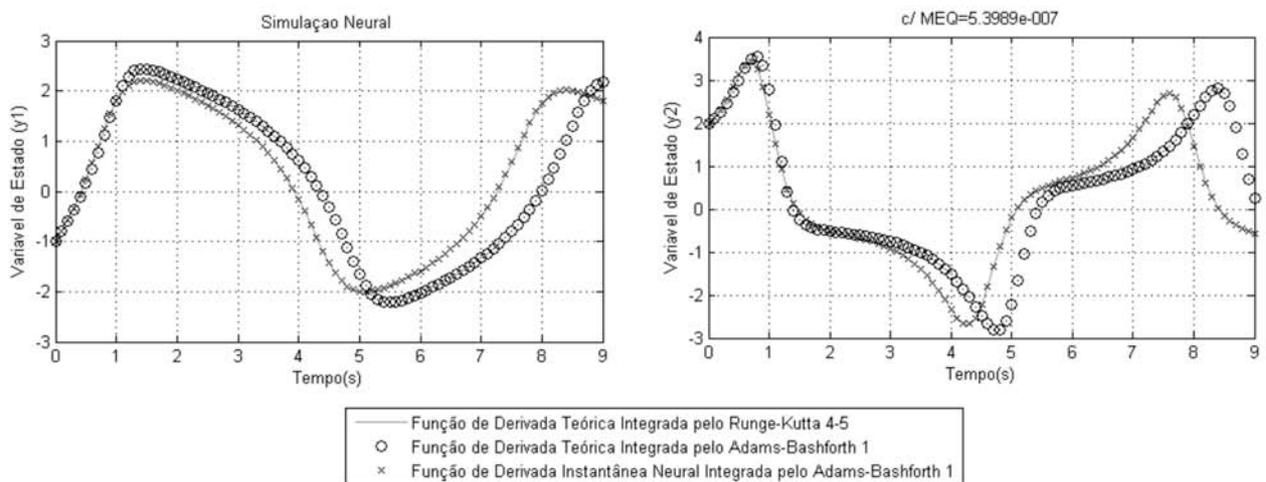


Figura 20: Simulação Obtida da Metodologia Empírica das Derivadas Instantâneas pelo IN de Adams-Bashforth de Primeira Ordem($\Delta t=0,1$).

Tabela 1: Comparações Quantitativas Entre as Metodologias de Treinamento Neural.

	NARMAX	Derivadas Instantâneas		Derivadas Médias	
		Metodologia Direta	Metodologia Empírica	Metodologia Direta	Metodologia Empírica
Média do Erro Quadrático Necessário ao Treinamento Neural	$10^{-7} - 10^{-8}$ Tasinaffo (2003); Rios Neto e Tasinaffo (2003)	$10^{-5} - 10^{-6}$ Rios Neto e Tasinaffo (2002); Tasinaffo e Rios Neto (2003)	$10^{-7} - 10^{-8}$ Melo (2008)	$10^{-5} - 10^{-6}$ Tasinaffo (2003); Tasinaffo e Rios Neto (2004); Tasinaffo e Rios Neto (2005); Tasinaffo e Rios Neto (2006); Melo (2008)	$10^{-6} - 10^{-7}$ Melo (2008); Melo e Tasinaffo (2008)
Metodologia Aplicável aos Dados do Mundo Real	Sim	Não	Sim	Sim	Sim

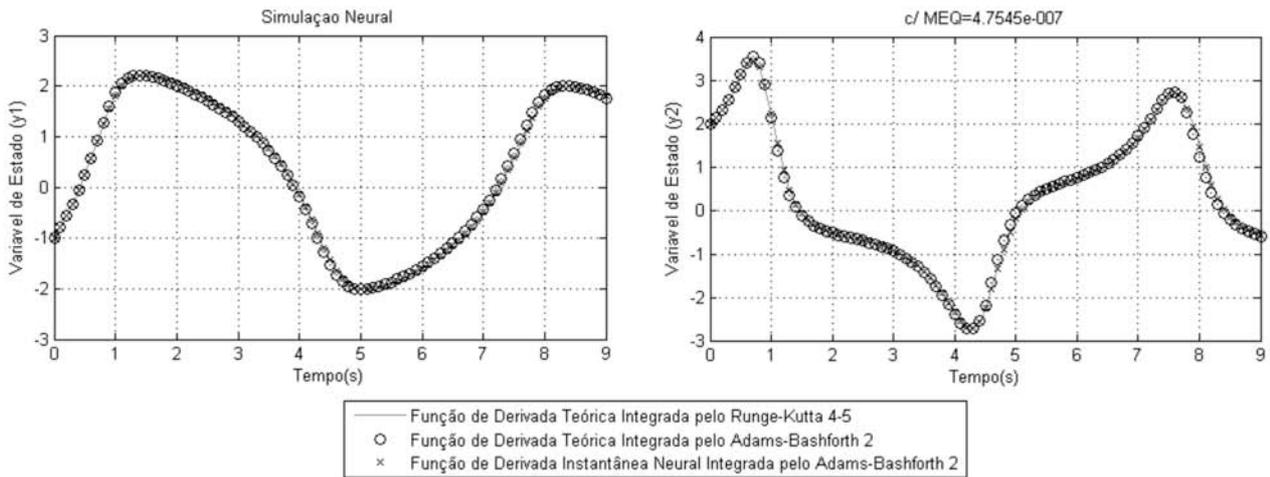


Figura 21: Simulação Obtida da Metodologia Empírica das Derivadas Instantâneas pelo IN de Adams-Bashforth de Segunda Ordem ($\Delta t=0,1$).

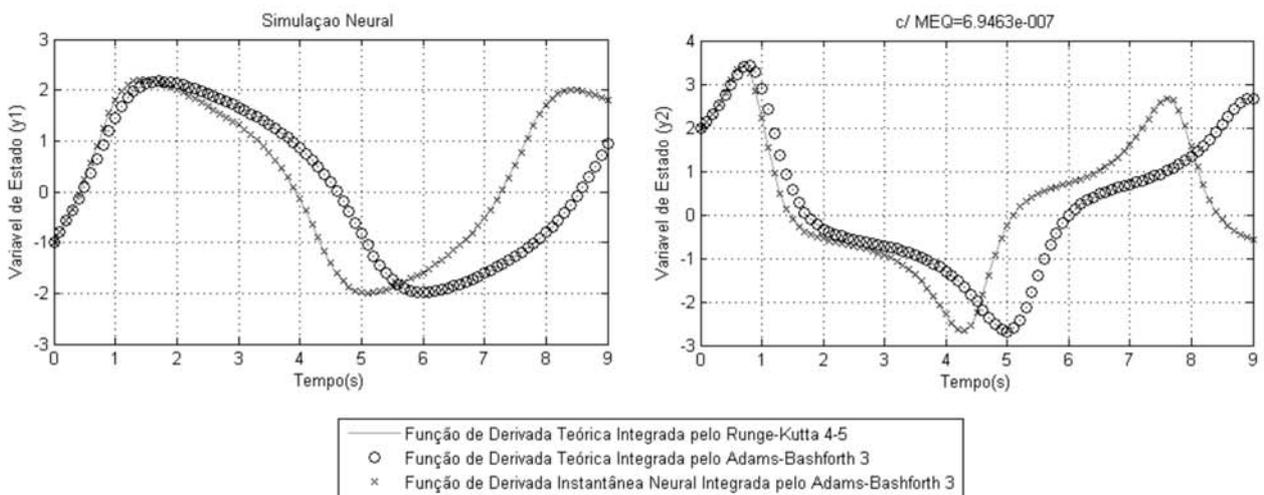


Figura 22: Simulação Obtida da Metodologia Empírica das Derivadas Instantâneas pelo IN de Adams-Bashforth de Terceira Ordem ($\Delta t=0,1$).

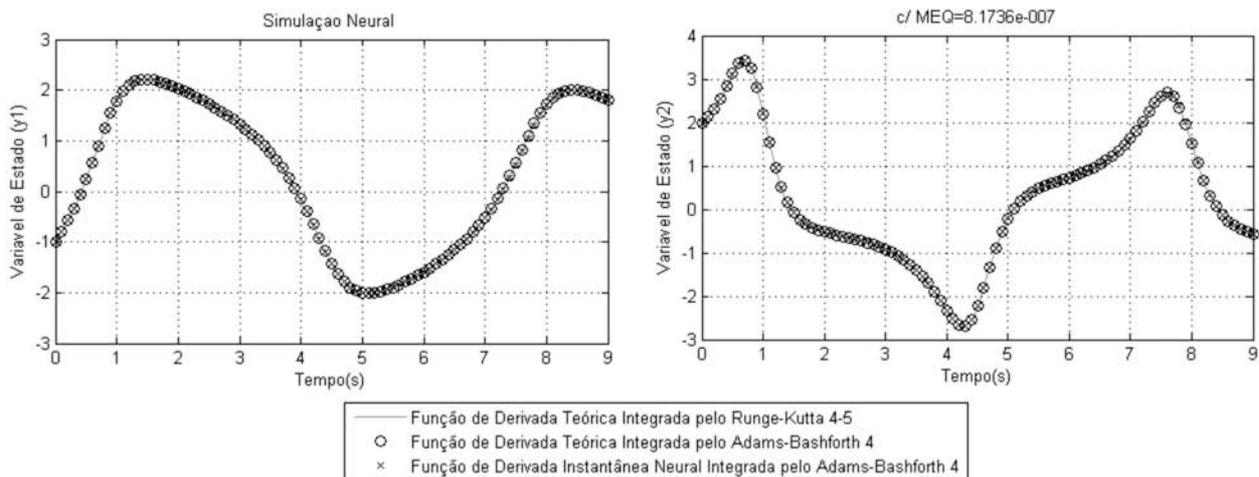


Figura 23: Simulação Obtida da Metodologia Empírica das Derivadas Instantâneas pelo IN de Adams-Bashforth de Quarta Ordem ($\Delta t=0,1$).

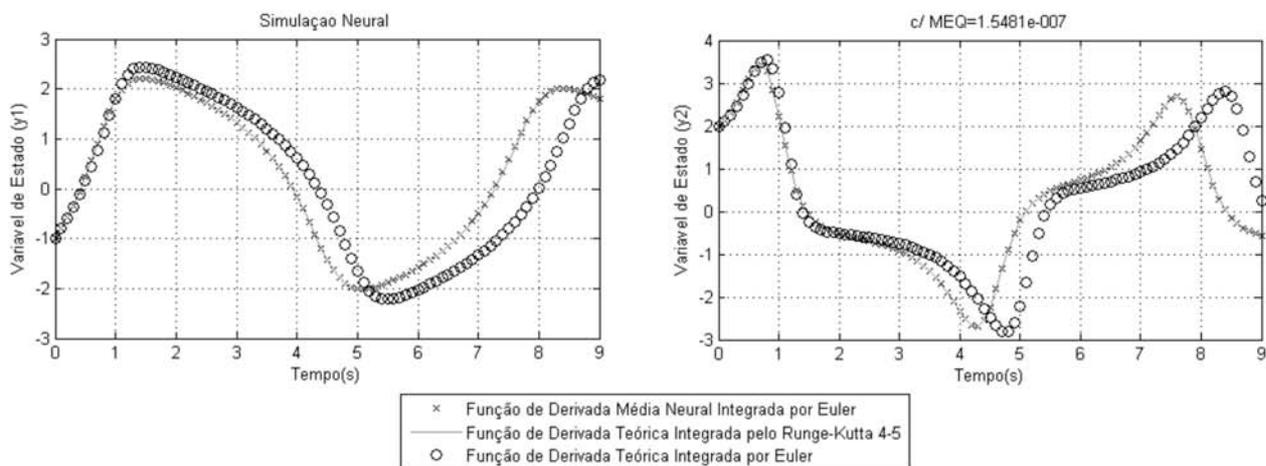


Figura 24: Simulação Obtida da Metodologia Empírica das Derivadas Médias pelo IN de Euler ($\Delta t=0,1$).

Tabela 2: Comparações Descritivas Entre os Metodologias de Treinamento Neural.

Influência do Desenvolvimento Teórico dos Integradores Utilizados Quando Acolados às Redes Feedforward.			
Metodologia das Derivadas Instantâneas		Metodologia das Derivadas Médias	
	Integradores de Passo Simples do tipo Runge-Kutta	Integradores de Múltiplos Passos do Tipo Adams-Bashforth	Integradores de Passo Simples do tipo Euler
1	Regras da cadeia cada vez mais complexas com o aumento da ordem do integrador utilizados no treinamento neural conforme Wang e Lin (1998).	Somente a presença de uma combinação linear no algoritmo <i>backpropagation</i> com o crescimento da ordem do integrador em Melo (2008).	Mais simples de todos os algoritmos simplesmente o algoritmo <i>backpropagation</i> acrescido do produto de Δt em Tasinaffo e Rios Neto (2005).
2	Boas precisões numéricas somente alcançáveis com integradores de mais alta ordem.	Boas precisões numéricas somente alcançáveis também com integradores de mais alta ordem.	Boas precisões alcançáveis simplesmente com o integrador tipo Euler projetado com o conceito de derivadas médias.
3	O número de padrões de treinamento não aumenta linearmente com a ordem do integrador utilizado.	O número de padrões de treinamento sempre aumenta linearmente com a ordem do integrador em Melo (2008).	O número de padrões de treinamento é sempre da mesma grandeza dos integradores de passo simples.
4	Método de passo variável e menos sensível à variação do mesmo.	Método de passo variável e mais sensível à variação do mesmo.	Método incondicionalmente de passo fixo. Para alterá-lo é necessário treinar uma nova rede.
5	O Runge-Kutta é obtido pelo desenvolvimento em série de Taylor da função original e, portanto, quanto maior a ordem do integrador tem-se necessariamente maior precisão numérica na propagação da solução em Henrici (1964).	O Adams-Bashforth é obtido pela substituição da função original, por uma interpolação polinomial, portanto não necessariamente, quanto maior a ordem do integrador maior será a precisão numérica da solução propagada, pois polinômios de graus mais elevados oscilam em torno da função original em Vidyasagar (1978).	Prova-se matematicamente que o integrador de Euler - projetado com derivadas médias - obtém uma precisão equivalente à qualquer integrador, de mais alta ordem projetado com derivadas instantâneas em Tasinaffo (2003).
6	De difícil obtenção matemática e em aberta sua generalização com a relação à ordem de integração. Wang e Lin (1998) aplicam esta metodologia somente para o Runge-Kutta de quarta ordem.	Facilmente obtido e generalizado do ponto de vista matemático em Melo (2008).	Facilmente obtido do ponto de vista matemático e equivalente ao Adams-Bashforth de primeira ordem.
7	Aplicação em controle preditivo extremamente difícil do ponto de vista matemático e ainda em aberto.	Facilmente aplicável na teoria de controle em Rios Neto (2001); Tasinaffo e Rios Neto (2004).	Facilmente aplicável na teoria de controle em Tasinaffo (2003).
8	Modelos do mundo real obtidos somente pela metodologia empírica.	Modelos do mundo real obtidos somente pela metodologia empírica.	Modelos do mundo real obtidos tanto pela metodologia empírica como pela metodologia direta.
9	De difícil normalização e desnormalização dos dados de treinamento do ponto de vista teórico.	De difícil normalização e subsequente desnormalização dos dados de treinamento. Ainda em estudos.	Problemas de normalização e desnormalização facilmente resolvíveis Melo (2008).

de Adams-Bashforth.

Na implementação do protótipo constata-se a diminuição do esforço computacional no treinamento neural. Ao utilizar o integrador de múltiplos passos do tipo Adams-Bashforth de quarta ordem, numa combinação linear de funções sobre a mesma rede neural, evitou-se a utilização da regra da cadeia de forma excessiva na determinação das derivadas exigidas pelo algoritmo *backpropagation*.

Realizando uma comparação de resultados entre as metodologias diretas e empíricas, constata-se que a metodologia direta é relativamente mais rápida, pois não utiliza os integradores numéricos na fase de treinamento neural, e necessita alcançar um erro maior, em torno de 10^{-5} ou 10^{-6} .

A metodologia empírica é mais lenta, pois realiza todos os cálculos dos integradores numéricos na geração, treinamento e simulação neural. Este fato, também justifica o motivo do erro da média das normas das propagações ser um pouco menor, em torno de 10^{-7} .

REFERÊNCIAS

Apostol, T. M. (1963). *Cálculo*, Vol. 31, Editorial Reverté, Rio de Janeiro.

Braga, A., Ludermir, T. and Carvalho, A. (2000). *Redes neurais artificiais*, LTC, Rio de Janeiro.

Braun, M. (1983). *Differential equations and their applications*, 3 edn, Springer-Verlag, New York.

Chen, S. and Billings, S. A. (1992). Neural networks for nonlinear dynamic system modelling and identification, *International Journal of Control* **56**(2): 319–346.

Cybenko, G. (1988). *Continuous valued neural networks with two hidden layer are sufficient*, Tufts University, Medford. (Technical Report).

He, S., Reif, K. and Unbehauen, R. (2000). Multilayer neural networks for solving a class of partial differential equations, *Neural Netw.* **13**(3): 385–396.

Henrici, P. (1964). *Elements of numerical analysis*, John Wiley and Sons, New York.

Hornik, K., Stinchcombe, M. and White, H. (1989). Multilayer feedforward networks are universal approximators, *Neural Networks* **2**(5): 359–366.

Hunt, J. K., Sbarbaro, D., Zbikowski, R. and Gawthrop, P. (1992). Neural networks for control system-a survey, *Automatica* **28**(6): 1083–1112.

Mai-Duy, N. and Tran-Cong, T. (2001). Numerical solution of differential equations using multiquadric radial basis function networks, *Neural Netw.* **14**(2): 185–199.

-
- Melo, R. P. (2008). *Metodologia de modelagem empírica utilizando integradores neurais aplicada a sistemas dinâmicos não-lineares*, Master's thesis, Instituto Tecnológico de Aeronáutica-ITA, São José dos Campos.
- Melo, R. P. and Tasinaffo, P. (2008). Uma metodologia de modelagem empírica utilizando o integrador neural de euler, *XVII Congresso Brasileiro de Automática*, Juiz de Fora-MG. 1 CD-ROM.
- Narendra, K. S. (1996). Neural networks for control, *Proceedings of the IEEE* **84**: 1385–1406.
- Narendra, K. S. and Parthasarathy, K. (1990). Identification and control of dynamical systems using neural networks, *IEEE Transactions on Neural Networks* **1**: 4–27.
- Ogata, K. (1967). *State space analysis of control systems*, Prentice-Hall, Englewood Cliffs.
- Rios Neto, A. (2001a). Design of a kalman filtering based neural predictive control method., *XIII Congresso Brasileiro de Automática*, Florianópolis-SC, pp. 2130–2134.
- Rios Neto, A. (2001b). Dynamic systems numerical integrators in neural control schemes., *V Congresso Brasileiro de Redes Neurais*, Rio de Janeiro-RJ, pp. 85–88.
- Sokolnikoff, I. S. and Redheffer, R. M. (1966). *Mathematics of physics and modern engineering*, 2 edn, Mcgraw-Hill Kogakusha, Tokio.
- Tasinaffo, P. M. (2003). *Estruturas de integração neural feedforward testadas em problemas de controle preditivo*, PhD thesis, Instituto Nacional de Pesquisas Espaciais, São José dos Campos.
- Tasinaffo, P. and Rios Neto, A. (2005). Mean derivatives based neural euler integrator for nonlinear dynamic systems modeling, *Learning and Nonlinear Models* **3**(2): 98–109.
- Vidyasagar, M. (1978). *Nonlinear systems analysis*, Prentice-Hall, Englewood Cliffs.
- Wang, Y. J. and Lin, C. T. (1998). Runge-kutta neural network for identification of dynamical systems in high accuracy, *IEEE Transactions on Neural Networks* **9**(2): 294–307.
- Wilson, E. B. (1958). *Advanced calculus*, Dover, New York.
- Zurada, J. M. (1992). *Introduction to artificial neural system*, PWS, St. Paul.