

NOVA IMPLEMENTAÇÃO EM FILTRO DO KALMAN ESTENDIDO PARA ASSIMILAÇÃO DE DADOS COM REDES NEURAIS

ROSÂNGELA S. C. CINTRA¹, HAROLDO F. DE CAMPOS VELHO¹, RICARDO TODLING²

¹Laboratório de Computação e Matemática Aplicada, Instituto Nacional de Pesquisas Espaciais, Av. Astronautas, 1758 – 12.227-101 -São José dos Campos, São Paulo, Brazil

²Global Modelling and Data Assimilation Office, GMAO-NASA/Goddard Space Flight Center 20770 - Greenbelt/Maryland, EUA

E-mail: rosangela.cintra@lac.inpe.br, haroldo@lac.inpe.br, todling@gmao.gsfc.nasa.gov.

Resumo – Assimilação de Dados é uma metodologia que combina dados de um modelo matemático de evolução com observações para se obter a melhor condição inicial possível para modelos de previsão. Diversos métodos de assimilação de dados estão atualmente em uso nas ciências atmosféricas e oceânicas, onde procuram implantar um algoritmo que mais se aproxime do estado verdadeiro da dinâmica do processo. Métodos sequenciais ótimos são baseados em teoria de estimativa formal que minimiza os erros dos dados de acordo com a dinâmica do modelo. Métodos de assimilação de dados utilizando Redes Neurais Artificiais (RNA) vêm sendo propostos recentemente e têm apresentado resultados consistentes, eficientes computacionalmente e eficazes quanto à aplicação. Este trabalho apresenta uma nova abordagem na aplicação de Assimilação de Dados utilizando redes Perceptron de Múltiplas Camadas que propõe a redução da complexidade computacional do problema. Os estudos foram feitos com o Filtro de Kalman Estendido em conjunto com o Sistema de Lorenz (1963) com sua dinâmica caótica não linear e iniciou experimento com o Modelo meteorológico Dynamo. A RNA apresentou resultados muito próximos às trajetórias das variáveis dos modelos dinâmicos comprovando a eficácia do método para o problema de assimilação de dados atmosféricos com algoritmo menos complexo.

Palavras-chave – Assimilação de dados, Perceptron Multicamadas, Modelos Dinâmicos não Lineares, Filtro de Kalman.

1. Introdução

O problema de determinar as melhores condições iniciais para os modelos de Previsão Numérica de Tempo (PNT) é de grande importância prática, e assunto de muitos estudos por pessoas de cenários diferentes. Os modelos de PNT são simulações computacionais da atmosfera que tomam a análise como ponto de partida e desenvolvem o estado da atmosfera no tempo usando a compreensão de física e dinâmica do fluido atmosférico. A análise meteorológica é produzida pelo processo de assimilação de dados, onde a informação recebida da observação é usada em conjunto com a previsão mais recente de um modelo numérico no tempo que a observação foi feita. A qualidade da Previsão do Tempo é fruto da formulação dos modelos matemáticos e físicos e da quantidade e qualidade dos dados observados.

A atmosfera se comporta como um sistema caótico, pois é muito sensível às pequenas variações nas condições iniciais. Por esta razão, existe a busca constante de obter a melhor análise, que é a melhor estimativa do estado atual da atmosfera para iniciar o ciclo de PNT. Para maior compreensão e maiores detalhes em assimilação de dados atmosféricos em [1] e [4].

Diferentes algoritmos de assimilação podem ser derivados de uma fonte comum, com aproximações próprias para cada algoritmo, por exemplo: métodos sequenciais como Interpolação Ótima (OI) e Filtro de Kalman, métodos variacionais em três dimensões (3D-VAR) e em quatro dimensões (4D-VAR) e métodos que utilizam Redes Neurais Artificiais [3]. Atualmente nos grandes centros mais de um milhão de observações são utilizados nos horários de análise. Estas observações são selecionadas para que seja possível a execução dos sistemas de assimilação de dados existentes. Neste cenário, se insere metodologias que utilizam Redes Neurais Artificiais (RNA) que podem ser competitivos. A abordagem de redes neurais foi testada em Nowosad [11-12] e em [5-6].

Na solução do algoritmo de análise foi implantada uma RNA para obter o incremento de análise a ser adicionado ao modelo previsto e obter a condição inicial para integrar o modelo numérico de PNT. A rede utilizada foi o Perceptron de Múltiplas Camadas (PMC) com o algoritmo de treinamento de retro-propagação do erro. A nova metodologia foi aplicada ao Sistema de Lorenz [8]. O sistema de assimilação de dados utilizado como treinamento da RNA foi o Filtro de Kalman Estendido (FKE). Boutier [1] introduz o Filtro de Kalman no contexto de assimilação de dados atmosféricos. Grande parte da carga computacional do FKE reside na avaliação da evolução de tempo da matriz de covariância de erros dos estados. Neste ponto, a investigação e utilização da RNA atuam na eliminação desta carga computacional comparada ao FKE. Não será abordado a complexidade do “software” ou as questões da validação algorítmica e científica. O desempenho do sistema de assimilação

será discutido, baseada na formulação matemática da assimilação de dados, que apresenta a RNA como algoritmo para aumentar o desempenho computacional no contexto da assimilação de dados.

2. Metodologia

Assimilação de dados é um meio manter o estado do modelo próximo à natureza pela assimilação de observações. No modelo de previsão, a representação do estado atmosférico, é uma matriz coluna chamada vetor de estado x . Alguns métodos de assimilação de dados são uma estimação de máxima verossimilhança usando toda informação disponível: denota-se a previsão como x^f e observações como y^o . Obtém-se o estado de análise x^a pela média ponderada linear entre x^f e y^o , seguindo a formulação:

$$x^a = x^f + K(y^o - H(x^f)) \quad (2.1)$$

onde K denota a matriz peso ou ganho da análise:

$$K = P^f H^T (H P^f H^T + R)^{-1} \quad (2.2)$$

Na expressão acima, P^f é a matriz de covariância do erro de modelagem (ou erro de previsão), R é a matriz de covariância do erro de medida, H representa o sistema de observação e o sobrescrito T indica matriz transposta. Na prática, observações diretas não são feitas em todos os pontos de grade para todas as variáveis prognosticas, os vetores y^o e x estão em diferentes espaços. Quando o operador H converte as variáveis prognosticas x para o espaço dos dados de observação y^o é chamado "operador de observação". H não é um mapeamento linear. As diferenças entre observações e modelo de previsão (conhecidas como *vetor inovação*), com o operador H define o espaço das variáveis nos pontos de observação. A diferença entre análise e modelo de referência (conhecidos como *incremento de análise*) define pelo operador H o espaço das variáveis na grade do modelo de previsão. É necessário distinguir entre a própria realidade e a melhor possível representação da realidade então existe um vetor estado hipotético, chamado x^t . O estado x^t define os erros da análise, previsão, e de observação. E com estes erros obtém-se a matriz de covariância do erro da análise, P^a , a matriz de covariância do erro da previsão P^f e a matriz de covariância do erro da observação R .

O Filtro de Kalman Estendido (FKE) é um método de assimilação de dados sequencial utilizado para obter a análise dos modelos numéricos testados e os resultados serviram para emulação da análise pelas redes neurais artificiais. O Filtro de Kalman é muito usado em problemas de estimação e controle, desde as primeiras aplicações no domínio aeroespacial, seu emprego ampliou-se a outras áreas e tem sido aplicado em meteorologia, oceanografia e hidrologia.

O FK tem a formulação do problema de estimação de verossimilhança acima (equação 2.1). Para sistemas não lineares, o modelo dinâmico do sistema é $x_n = f(x_{n-1}, \eta_n)$, $\eta_n \approx N(0, Q_n)$ e o modelo de medição: $z_n = h(x_n, \xi_n)$, $\xi_n \approx N(0, R_n)$, onde x_n representa o vetor de estado do modelo no passo de tempo t_n , $f(\cdot)$ e $h(\cdot)$ o modelo matemático de previsão e medição, bem como Q_n e R_n são os erros de modelagem e medição respectivamente. Neste caso usa-se o Filtro de Kalman Estendido (FKE) onde o modelo "linear" do sistema é: $x_n = \tilde{x}_n + A(x_{n-1} - \tilde{x}_{n-1}) + W(\eta_n)$ e o modelo de medição $Z_n = \tilde{z}_n + H(x_n - \tilde{x}_n) + V(\xi_n)$, onde A e H são os Jacobianos de f e h , em ordem a x , W e V são os Jacobianos de f e h , em ordem a η e ξ respectivamente.

2.1 Algoritmo EKF

O EKF apresenta o seguinte algoritmo:

- Fase de previsão, a partir do modelo:

Previsão de estado:

$$x_n^f = f_n[x_{n-1}^f, 0] \quad (2.9)$$

Previsão de covariância:

$$P_n^f = A_n P_n^a A_n^T + W_n Q_n W_n^T \quad (2.10)$$

- Fase de atualização:

Calculo da Matriz Ganho:

$$K_n = P_n^f H_n^T [H_n P_n^f H_n^T + V_n R_n V_n^T]^{-1} \quad (2.11)$$

Atualização do estado (análise):

$$x_n^a = x_n^f + K_n (z_n - h[x_n^f, 0]) \quad (2.12)$$

Atualização da covariância da análise:

$$P_n^a = [I - K_n H_n] P_n^f \quad (2.13)$$

Para um caso concreto é necessário definir o vetor de variáveis de estado x_n , identificar a matriz A que representa a dinâmica de transição de estado, identificar a matriz H que representa a relação entre as variáveis mensuráveis e as variáveis de estado do sistema. Avaliam-se as incertezas no modelo do sistema e dos instrumentos de medição e devem-se exprimi-las sob a forma de covariâncias das variáveis respectivas. O FK constitui um processo recursivo eficiente, uma vez que o erro quadrático médio é "minimizado", ou seja, é uma solução eficiente do método de mínimos quadrados, seria a estimação de mínimos quadrados com hipóteses gaussianas. O método FK tem as suas raízes no controle de sistemas, a sua aplicação em assimilação de dados já se faz há décadas. Jazwinski [1] provê uma introdução e discussão matemática mais precisa.

O FKE foi escolhido neste trabalho para se obter os dados de entrada para utilização da RNA e para avaliação de resultados do método proposto.

O processo de assimilação de dados, como já especificado, em alguns métodos consiste em um algoritmo que combina previsões de um modelo numérico, para um determinado instante, com dados observacionais, para o mesmo instante, a fim de se obter uma melhor estimativa do estado da atmosfera. Tal processo pode ser expresso de forma concisa pela Equação (2.1) e (2.2).

A nova abordagem é um método híbrido onde o algoritmo resolve uma parte da formulação matemática e a outra parte é resolvida pela RNA. A formulação matemática algébrica geral é baseada na estimação de mínimos quadrados [13]. As incertezas do modelo, das observações e da análise, assumem modelos de erros, calculadas por Função de Distribuição de Probabilidade (FDP) para cada tipo de erro. Os erros são modelados como "gaussianos". As matrizes de covariância destes erros são pré-determinadas. Esta matriz do modelo de previsão denota-se B e das observações denota-se R . Note que a matriz de covariância do erro de modelagem passa a ser denominada B por ser pré-estabelecida, enquanto que a notação em FKE é P^f , porque é determinada em tempo de integração do modelo.

A metodologia proposta utiliza uma Rede Neural Artificial (RNA), um modelo com conexões associada a "pesos" que armazenam o conhecimento e servem para ponderar a entrada recebida por cada neurônio da rede. A RNA que possui interconexões da entrada à camada de saída, e tem pelo menos uma camada de neurônios intermediária, tipicamente denominada de camada escondida, é o Perceptron Multicamadas (PMC), o modelo mais simples e conhecido [7].

Uma PMC realiza um mapeamento complexo $y = \Psi(w, x)$, do vetor de entradas $x = [x_0, x_1, x_2, x_3, \dots, x_N]$ ao vetor de saídas $y = [y_0, y_1, y_2, y_3, \dots, y_N]$ parametrizado pelos pesos sinápticos w . Redes Perceptron Multicamadas são ferramentas que, dado um número suficiente de neurônios conseguem aproximar qualquer função linearmente contínua. PMC possui a capacidade de aprender através de exemplos e fazer interpolações e extrapolações do que aprenderam. Um conjunto de procedimentos bem definidos é utilizado para adaptar os parâmetros de uma RNA a fim de produzir uma *saída desejada*. O processo de treinamento ou aprendizado supervisionado consiste em minimizar o *erro* entre a saída da rede para um determinado padrão de entrada e a resposta esperada para aquele mesmo padrão. Neste trabalho utilizou-se o algoritmo de *retro-propagação do erro* com o processo de treinamento citado.

Os primeiros resultados do método de assimilação de dados com RNA aplicadas a modelos não-lineares são apresentados em [12]. A mesma técnica de assimilação com RNA recorrentes apresentada em [5]. O método de assimilação de dados baseado em RNA utilizada por Nowosad [12] e Härter [5] pode ser descrito como: $x_n^a = F_{RNA}(x_n^f, y_n^o)$ onde F_{RNA} representa o processo de assimilação por RNA, onde as entradas da rede são o modelo de previsão e as observações. A saída esperada apresentada no treinamento é a análise de outro método de assimilação.

Neste trabalho, a assimilação de dados é feita com a estimação do *incremento de análise* através da RNA, com um processo de dois passos:

Passo de previsão:
$$x_n^f = F[x_{n-1}^f]$$

Passo de análise:
$$x_n^a = x_n^f + d_n$$

O vetor *incremento de análise* é o produto entre uma matriz de ponderação e uma função que mede a discrepância entre a previsão do modelo e as observações: $d_n = M_{rna}(y_n^o - x_n^f)$ sendo M_{rna} a função de estimação do vetor *correção*, equivalente às equações 2.12 e 2.13 do FKE

$$d = BH^T (HBH^T + R)^{-1} (y^o - H(x^f)) .$$

2.2 Algoritmo PMC

No procedimento de assimilação de dados com RNA proposto neste trabalho, utilizando uma PMC com uma camada estendida, temos o seguinte algoritmo:

1. Cálculo do *vetor inovação*

$$(y^o - H(x^f)) ; \tag{2.14}$$

2. Obtenção do vetor incremento de análise d com a ativação da RNA a cada iteração de n (número de entradas):

$$d_n = d_{RNA} \tag{2.15}$$

3. Adição do incremento de análise ao modelo de previsão x^f obtendo a análise x^a :

$$x_n^a = x_n^f + d_n \tag{2.16}$$

A ativação da RNA substituiu o cálculo da equação 2.11, ou seja, o cálculo da matriz ganho, formulação da inversa desta matriz e o retorno ao espaço do modelo (eq. 2.12). O algoritmo de treinamento teve como dados de entrada obtidos da execução prévia do FKE com os modelos teste de uma realização. A PMC de uma camada escondida teve como entradas o vetor *inovação*, observações (y^o) menos o vetor de previsão (x^f) e o vetor *incremento de análise* (d), para treinamento. O espaço de trabalho do *vetor inovação* é o espaço da observação, no instante da observação. O incremento de análise d foi obtido do cálculo da equação 2.21 do FKE, cujo espaço das variáveis é a grade do modelo. O *incremento de análise* a ser adicionado ao modelo de previsão, nesta nova abordagem proposta, é obtido com a generalização da RNA e resultará no espaço do modelo.

3. Resultados e discussões

O primeiro teste da técnica de RNA-PMC foi feito com o Sistema de Lorenz.

Edward Lorenz em sua pesquisa dos sistemas dinâmicos elegeu um sistema de três equações acopladas, hoje conhecidas como Equações de Lorenz. O modelo descreve um sistema relativamente simples com um padrão de complexidade infinita, onde se verificou que a partir de estados iniciais ligeiramente diferentes, o sistema de equações diferenciais (representando o estado de um fluido em convecção térmica utilizado então como protótipo do estado atmosférico), resultava em soluções completamente diferentes entre si. O sistema de Lorenz consiste de três equações diferenciais ordinárias não lineares de primeira ordem [8]:

$$\frac{\partial X}{\partial t} = -\sigma(X - Y) \tag{3.1}$$

$$\frac{\partial Y}{\partial t} = \rho X - Y - XZ \tag{3.2}$$

$$\frac{\partial Z}{\partial t} = XY - \beta Z \tag{3.3}$$

Quando os parâmetros σ , ρ e β assumem os valores 10, 28 e 8/3, respectivamente, o sistema de Lorenz apresenta uma dinâmica caótica. Com o sistema de Lorenz utilizou-se uma RNA, com três entradas relativas ao vetor de inovação das variáveis X, Y e Z e o vetor d_n das respectivas variáveis utilizadas como "saída alvo" durante o treinamento, que teve sua convergência no mínimo erro entre a saída esperada e a saída calculada pela RNA.

Na estimação do vetor d_n (incremento de análise), a PMC teve oito neurônios na camada escondida e três na camada de saída referente ao incremento de análise de X, Y e Z. O treinamento foi feito com 32.000 passos de tempo.

A simulação para assimilação de dados foi feita nos intervalos de 10, 20 e 40 passos de tempo, os resultados podem ser vistos na Figura 1. Verificou-se que para as três frequências de obtenção de dados de observação realizadas neste experimento, as condições iniciais geradas pela ativação da rede neural, acompanham a trajetória do modelo verdadeiro. No modelo de Lorenz para as três variáveis, não houve perturbação da inserção da observação no espaço do modelo. A convergência foi mais lenta para o treinamento da rede neural com 40 passos de tempo, mas este é um comportamento esperado, pois devido à frequência de observação ser muito maior do que nos dois experimentos prévios, a perturbação na condição inicial pode gerar uma trajetória que divergisse do modelo verdadeiro.

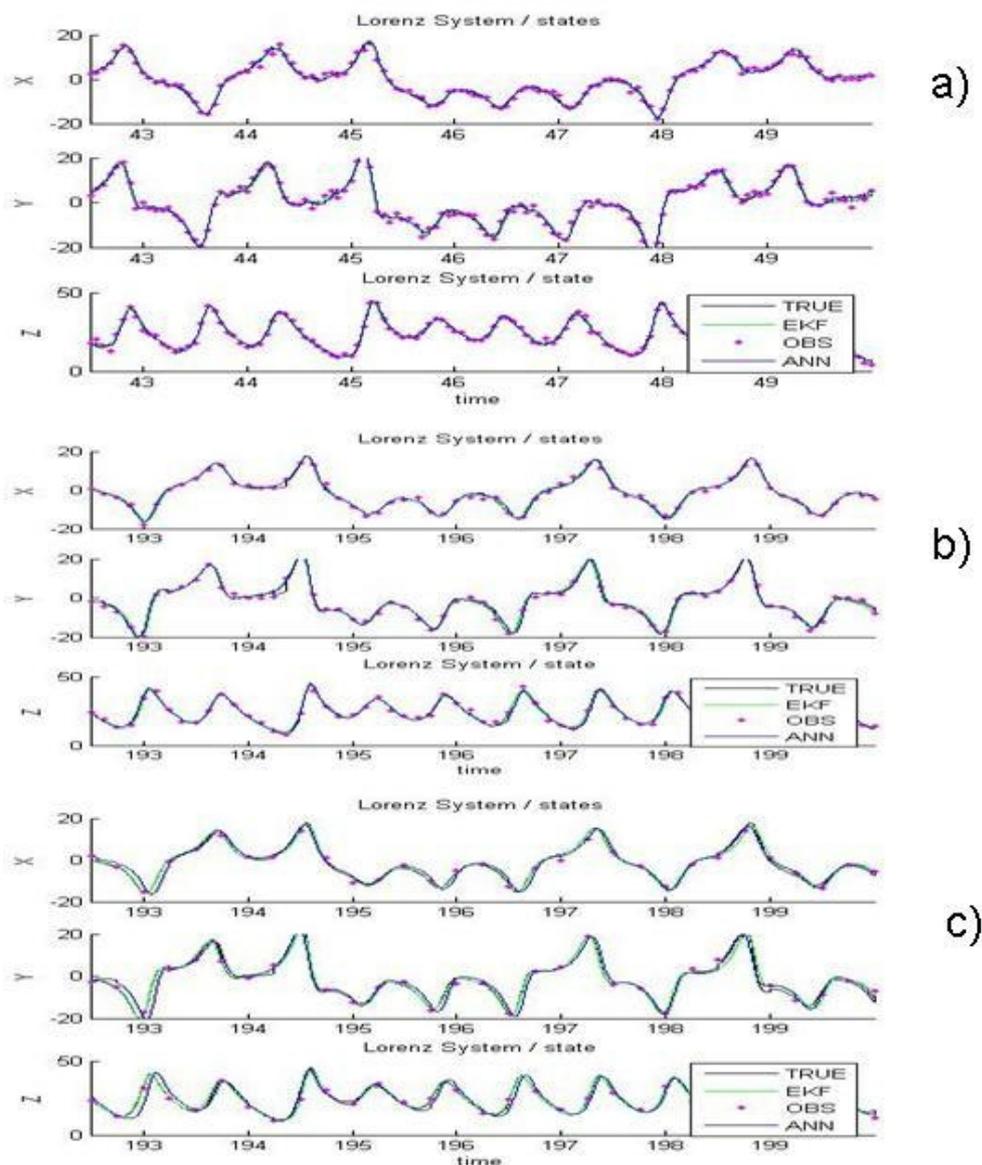


Figura 1 – Sistema de Lorenz: (a) assimilação a cada **dez** passos de tempo; (b) assimilação a **vinte** passos de tempo; (c) assimilação a **quarenta** passos de tempo; observações (pontilhado magenta), estado verdadeiro (linha preta), estado com EKF (linha verde), estado RNA2 (linha azul).

Um segundo teste foi realizado com o modelo meteorológico Dynamo, para atestar a eficiência da técnica de PMC-RNA na aplicação de modelo meteorológico de uma dimensão.

Segundo Lynch [10] (ver também [2]), o modelo Dynamo simula movimentos atmosféricos em escala sinótica. Apesar de sua grande simplicidade, o código Dynamo é capaz de reproduzir importantes fenômenos da dinâmica atmosférica. Este modelo é uma aproximação das equações de águas rasas, que considera a atmosfera como um fluido incompressível e em equilíbrio hidrostático na vertical, com o campo de velocidades independente da componente z ; admite-se atmosfera seca e fronteira livre para a altura geopotencial.

A superfície é considerada plana ($z = 0$) e com condições de contorno periódicas para todas as variáveis. Partindo das equações de águas rasas, num sistema de referência não inercial, temos:

$$\frac{du}{dt} - fv + \frac{\partial \Phi}{\partial x} = 0 \quad (3.4)$$

$$\frac{dv}{dt} + fu + \frac{\partial \Phi}{\partial y} = 0 \quad (3.5)$$

$$\frac{d\Phi}{dt} + \Phi(u_x + v_y) = 0 \quad (3.6)$$

onde d/dt é a derivada substantiva; x, y são as coordenadas no sentido oeste e norte; t é o tempo; $\Phi = gh$ geopotencial, sendo h a altura do fluido acima da superfície plana; u, v são as velocidades zonal e meridional. Considera-se o parâmetro de Coriolis na forma $f = f_0 + \beta y$ com f_0, β constantes. Por último, os subscritos significam derivação parcial. Supondo que o vento zonal seja constante e esteja em balanço geostrófico com o geopotencial, $f\bar{u} = -\partial\bar{\Phi}/\partial y$, e os desvios sejam variáveis unidimensionais [5]: $u = \bar{u} + U(x, t)$; $v = V(x, t)$; $\Phi = \bar{\Phi}(y) + \varphi(x, t)$, chegam-se as equações do modelo Dynamo.

O modelo Dynamo foi integrado no passo de tempo Δt de 60 segundos. Os dados foram inseridos a cada 360 passos de tempo. A rede PMC para este modelo possui três entradas relativas ao vetor de inovação das variáveis: U, V e Φ e o vetor incremento d das respectivas variáveis. No treinamento, estas entradas foram coletadas da execução prévia do modelo Dynamo com FKE. A PMC foi implantada com quinze neurônios na camada escondida e foi treinada com 200 exemplos. Na generalização, a ativação da RNA foi feita a cada 360 passos de tempo.

O tempo total de integração foi feito para 1200 horas, ou 50 dias. Os resultados mostram que com a assimilação de dados a cada 6 horas, as condições iniciais geradas pela RNA fizeram com que a trajetória do modelo de previsão coincidissem a trajetória do modelo verdadeiro, para as três variáveis, os resultados podem ser vistos na Figura 2 abaixo. Para cada gráfico da evolução das variáveis do modelo, há um gráfico da diferença logarítmica entre os modelos.

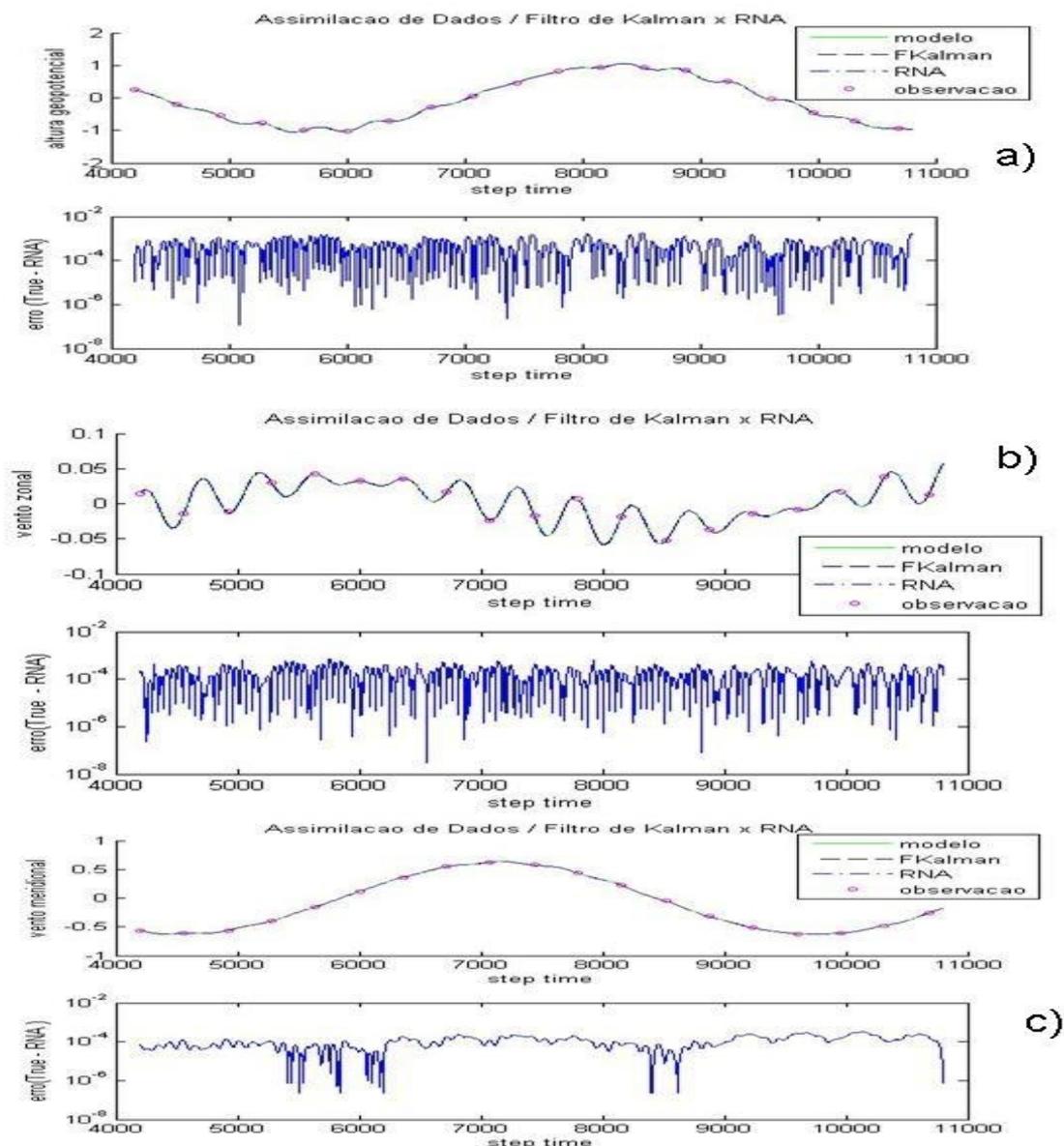


Figura 2 - Sistema Dynamo-1D: (a) assimilação a cada 360 passos, para variável ϕ , (b) para a variável u ; (c) para a variável v ; sendo observações (magenta), estado verdadeiro (linha preta), estado com EKF (linha verde), estado iniciado com RNA (linha azul). E respectivamente a diferença entre o estado verdadeiro de o modelo com assimilação RNA.

O sistema de Lorenz é um dos exemplos mais citados de dinâmica caótica. Dinâmica caótica ocorre em sistemas dinâmicos determinísticos em que a diferença entre duas trajetórias cresce exponencialmente com o tempo. Devido a esta grande sensibilidade, aliado a sua simplicidade, o sistema de Lorenz é usado para testar técnicas de assimilação dados.

É pertinente ressaltar que o treinamento da RNA para o sistema de Lorenz foi feito com incorporação das observações com uma frequência de inserção a cada 10 passos de tempo. Nos testes de generalização, os experimentos foram realizados com diferentes frequências de inserção dos dados de observação. Em todos estes experimentos, o sistema de Lorenz iniciado com a condição gerada pela RNA, seguiu a trajetória da dinâmica verdadeira.

A importância de se avaliar o desempenho do método de assimilação com redes neurais com o modelo Dynamo é que este modelo está mais próximo da realidade dos modelos meteorológicos operacionais: é um modelo não-linear, descreve a propagação de ondas meteorológicas importantes (ondas de Rossby e de gravidade, por exemplo) e as variáveis tem dependência temporal e espacial. Nos testes apresentados, a observação foi gerada sinteticamente somando-se um ruído aleatório às variáveis de previsão do modelo numérico (outro modo, seria adicionar uma variável aleatória diretamente nas equações do sistema de Lorenz), ou seja, a curva de observação é acoplada à dinâmica calculada pelo modelo numérico. Isto significa que após o processo de filtragem, as curvas estimadas pelas RNA e pelo FKE devem coincidir com a curva resultante do modelo de PNT. No experimento com modelo Dynamo, o treinamento e a generalização foram feitos com a mesma

freqüência de observação. O erro das trajetórias entre o modelo verdadeiro e o estimado é da ordem de 10^{-4} a 10^{-6} . Testes com diferentes freqüências de observação serão realizados em futuro próximo.

5. Conclusão

Sabe-se que existem erros no modelo e nas observações, e os sistemas de análise atuais (métodos de assimilação de dados) devem conter um algoritmo automático que minimize a diferença entre a análise e a “verdade”. A nova abordagem da utilização da RNA tem o propósito de redução do esforço computacional para se obter a condição inicial de modelos numéricos. Nesta abordagem, não existe dependência do conhecimento dos erros da modelagem e das observações. O método pode ser implementado independentemente do espaço das variáveis (observação ou modelo). A generalização da RNA para obtenção do incremento de análise d substitui o cálculo da inversão das matrizes de covariância de erro para obter a matriz “ganho” da análise e não possui o cálculo da soma do incremento de análise para a grade do modelo, pois já está neste espaço, com isso executa menos operações computacionais o que diminui a complexidade do algoritmo.

Agradecimentos:

Os autores agradecem ao Dr. Fabrício P. Härter pela concessão das rotinas referentes ao modelo Dynamo-1D, e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro à pesquisa.

6. Referências Bibliográficas:

- [1] F. Boutier, P. Courtier. Data Assimilation concepts and methods. In: **Meteorological Training Course**, Reading, UK, 1998.
- [2] H.F. Campos Velho. A Comprehensive Analysis of a Barotropic Limited Area Model Using the Nonmodal Matrix Technique. **Revista Brasileira de Meteorologia**, 2, 12(1997), 41–50.
- [3] H.F. Campos Velho, N. Vijaykumar, S. Stephany, A. J. Preto, A. G. Nowosad. A Neural Network Implementation for Data Assimilation using MPI, **Applications of High-Performance Computing in Engineering**, WIT Press, 2002, 280 p
- [4] R. Daley. **Atmospheric data analysis**. Cambridge University Press, 1991. 457 p
- [5] F. P. Hartër. Redes neurais recorrentes aplicadas à assimilação de dados em dinâmica não linear, **Dissertação de Doutorado** INPE, São José dos Campos, 2004..
- [6] F. P. Harter, H. F. Campos Velho. Recurrent and feedforward neural networks trained with cross correlation applied to the data assimilation in chaotic dynamic, **Revista Brasileira de Meteorologia**, 20, 3(2005), 411-420.
- [7] S. Haykin. **Redes neurais: princípios e práticas**. Porto Alegre: Bookmann, 2001
- [8] A. Jazswinski. **Stochastic processes and filtering theory**. New York: Academic, 1970. 376 p.
- [9] E. Lorenz. Deterministic nonperiodic flow. **Journal of the Atmospheric Sciences**, 20, 2(1963), 130-141.
- [10] P. Lynch. Initialization of a barotropic limited area model using the Laplace Transform technique. *Monthly Weather Review*, 113, 8(1985), 1338-1344.
- [11] A. G. Nowosad. Novas abordagens para assimilação de dados meteorológicos. **Dissertação de Doutorado**, INPE, 2001
- [12] A. Nowosad, H. F. Campos Velho, A. Rios Neto.. Neural network as a new approach for data assimilation. In: **Brazilian Congress On Meteorology**, SBMET, November (2000), 3078-3086.
- [13] O. Talagrand. Assimilation of observations, an introduction. **J. Meteor. Soc. Japan**, 75(1997), 191-209.