

ASSIMILAÇÃO DE DADOS UTILIZANDO REDES NEURAS ARTIFICIAIS EM UM MODELO ATMOSFÉRICO DE CIRCULAÇÃO GERAL

Rosângela S. Cintra¹ e Haroldo F. de Campos Velho¹

¹Laboratório de Computação e Matemática Aplicada – LAC
Instituto Nacional de Pesquisas Espaciais (INPE)– São José dos Campos, SP
{rosangela.cintra, haroldo}@lac.inpe.br

Resumo: Assimilação de Dados é uma tarefa importante em Centros Operacionais de Previsão Numérica de Tempo (PNT). Vários métodos da assimilação de dados são aplicados em modelos de dinâmica atmosféricos e oceânicos. Métodos de assimilação de dados utilizando Redes Neurais Artificiais (RNA) vêm sendo propostos apresentando resultados consistentes quanto aplicação, em modelos simples. Este trabalho apresenta uma abordagem, onde se aplica uma RNA para emular um Filtro de Kalman por Conjunto (Ensemble Kalman Filter - EnKF) como método de assimilação de dados em conjunto com Modelo de Circulação Geral Atmosférico dinâmico não linear (MACG), como dado de PNT. Os resultados preliminares da análise RNA em comparação à análise do EnKF são encorajadores para continuidade da pesquisa com o novo método para a assimilação de dados. Uma vantagem do uso de redes neurais é o melhor desempenho computacional.

Abstract. Atmospheric Data Assimilation is an important task in Operational Centers of Numerical Weather Prediction (NWP). Several methods of data assimilation are applied in models of atmospheric and oceanic dynamics. Methods using Artificial Neural Networks (ANN) have been proposed showing consistent results regarding implementation in simple models. This paper presents an approach, which applies an ANN to emulate an Ensemble Kalman Filter (EnKF) as a method of data assimilation in conjunction with Atmospheric General Circulation Model (AGCM) nonlinear dynamical, as PNT data. Preliminary results of RNA analysis in comparison to the EnKF analysis are encouraging for continued research with the new method for data assimilation. One advantage of using neural networks is the best computational performance

Palavras Chave: assimilação de dados, redes neurais artificiais, filtro de Kalman por conjunto.

1. INTRODUÇÃO

A assimilação de dados utiliza um modelo dinâmico, em combinação com dados, inexatos e irregularmente distribuídos no espaço e no tempo, para inferir o estado inicial para o desenvolvimento de um modelo atmosférico. O processo da assimilação é uma tarefa importante em previsão de tempo operacional (Daley, 1991) e simulação oceânica. Os métodos clássicos da assimilação de dados podem ser descritos como um procedimento que usa dados de observação para melhorar uma previsão recente fornecida por um modelo matemático. Tipicamente, a assimilação de dados feita com métodos clássicos pode ser delineada em dois passos: Passo de previsão $x_n^f = F[x_{n-1}^f]$, Passo de análise $x_n^a = F[x_n^f] + d_n$ onde x_n representa o vetor de estado do modelo no passo de tempo t_n , $F[.]$ e o modelo

matemático de previsão, os sobrescritos f e a denotam respectivamente os valores preditos e de análise, finalmente d_n é o incremento da análise. O vetor incremento de análise é calculado como sendo um produto entre uma matriz de ponderação e uma função que mede a discrepância entre a previsão do modelo e as observações y_n^o : $d_n = M(y_n^o - Hx_n^f)$, sendo M a matriz de ponderação ou matriz “ganho”, e H o operador observação. Sabe-se que existem erros no modelo e nas observações e diversos métodos de assimilação existentes procuram uma estratégia de análise formulada como problema de otimização, ou seja, um método que minimize a diferença entre análise e o estado “verdadeiro” em determinado instante de tempo. Para isto, é necessário representar matematicamente os erros dos estados, modelados estatisticamente usando conceitos probabilísticos. Diferentes algoritmos de assimilação podem ser derivados desta abordagem, com aproximações próprias para cada algoritmo. O desafio computacional para a metodologia clássica de assimilação de dados reside na dimensão das matrizes envolvidas em modelos operacionais de PNT, atualmente na ordem de um milhão de equações (o que equivale a matrizes cheias da ordem de 10^{12} elementos!). Neste cenário a aplicação de Redes Neurais Artificiais (RNA) em assimilação de dados foi testada em Nowosad e Campos Velho (2001) e Härter e Campos Velho (2004, 2005), como proposta à solução deste desafio.

A técnica desenvolvida por Nowosad (2001) e Harter (2004) utiliza as redes neurais para implementar a função: $x_n^a = F_{ma}(y_n^o, x_n^f)$ onde F_{ma} representa o processo de assimilação. Esta técnica foi testada em sistemas caóticos não lineares como: Sistema de Lorenz (1963) e o simulador meteorológico DYNAMO (baseado em equações de águas rasas). As redes neurais nestas aplicações emularam o Filtro de Kalman Estendido (FKE).

Neste trabalho a mesma técnica de RNA em assimilação de dados foi testada em um modelo atmosférico mais realista emulando um método de assimilação de dados mais recente. O modelo conhecido como SPEEDY (Simplified Parametrizations, primitivE-Equation DYnamics) (Myoshi, 2005, Molteni, 2003) é um modelo atmosférico de circulação geral (MACG) de complexidade intermediária, baseado nas equações primitivas em dinâmica espectral. O método de assimilação escolhido foi o LETKF (Local Ensemble Kalman Filter) (Ott et al, 2004, Kalnay, 2003), baseado em filtro de Kalman por conjunto. A metodologia utilizada para obtenção da análise com redes neurais descreve características do SPEEDY, do LETKF e da RNA Perceptron Multicamadas e uma estratégia de treinamento da rede.

2.METODOLOGIA

O filtro de Kalman linear é a solução ótima do problema de estimação do estado para sistemas dinâmicos lineares (Kalman, 1960); e via linearização do problema para sistemas não linear, o KF foi adaptado para o Extended Kalman filter. Embora a metodologia do Filtro de Kalman tenha sido extremamente bem sucedida para sistemas de baixa dimensão, a aplicação direta para sistemas de alta dimensão é problemática devido às operações com matrizes muito grandes. Jazwinski (1970) descreve uma introdução e discussão matemática sobre o Filtro de Kalman e Boutier (1998) introduz o filtro de Kalman no contexto de assimilação de dados atmosféricos.

Várias técnicas alternativas foram propostas usando a técnica de Filtro de Kalman, uma delas é o Filtro de Kalman por conjunto (Ensemble Kalman Filter - EnKF). O EnKF é um método sequencial de assimilação de dados, proposto por Evensen(1994). Aplica-se um conjunto de estados do modelo para representar as estatísticas de erro da estimativa do modelo, através de um conjunto de integrações do modelo usadas para prever as estatísticas de

erro no tempo. Utiliza um esquema de análise que atua diretamente sobre o conjunto de estados do modelo quando as observações são assimiladas. O conjunto de análises é obtido pela assimilação de observações (perturbadas) para cada membro do conjunto do modelo de referência. Devido a diferentes tipos de perturbações usadas para representar a matriz de covariância da análise, muitos esquemas se derivaram da abordagem do EnKF: o Local Ensemble Kalman Filter (Ott et al., 2004) é um deles, onde a "localização da covariância" é feita em pequenos subgrupos de variáveis definidos em pontos de grade do modelo para regiões locais menores que toda a superfície da Terra.

A idéia básica do LETKF é executar a análise em cada ponto de grade simultaneamente, utilizando as variáveis do estado e todas as observações na região local centrada naquele ponto. Cada membro do conjunto obtém sua previsão $x_{n-1}^{f(i)} : i = 1, 2, 3 \dots k$, sendo k o número de membros no tempo t_n . H e R representam respectivamente o operador observação e a matriz de erros de covariância da observação. Para estimar o vetor estado do modelo de referência é usada a média do conjunto de previsões: $\bar{x}^{-f} = k^{-1} \sum_{i=1}^k x^{f(i)}$, a matriz

de covariância do estado: $P^f = (k-1)^{-1} \sum_{i=1}^k (x^{f(i)} - \bar{x}^{-f})(x^{f(i)} - \bar{x}^{-f})^T$, a matriz de covariância

da análise $P^a = P^f [(P^f)^{-1} x^f + H^T R^{-1} H]^{-1}$ e um conjunto de análise $x^{a(i)} : i = 1, 2, 3 \dots k$ com a média e a covariância de erros do conjunto.

No LETKF a assimilação é feita com "localização" espacial geralmente é feita explicitamente, fazendo uma análise separada para cada ponto da grade espacial do modelo. A análise local permite diferentes combinações lineares dos membros do conjunto em diferentes regiões, e a análise global explora uma dimensão espacial maior. Para implementação local separa-se grupos de observações vizinhas a um ponto central para uma determinada região da grade do modelo.

O algoritmo do LETKF pode ser: 1. Calcula-se um conjunto de previsões globais para obter a média global do conjunto do estado referência. 2. Para cada região local e cada membro do conjunto referência formam-se vetores do estado da atmosfera da região. 3. Em cada região local, projete o vetor "local" obtido no passo 2, ao subespaço dimensional que melhor represente o conjunto naquela região. 4. Faz-se assimilação de dados em cada região local, obtendo a análise. 5. Com as análises locais obtidas no passo 4, forma-se uma análise global. 6. Voltar ao passo 1.

O modelo de "Simplified Parametrizations, primitive-Equation Dynamics" (SPEEDY) é um recente desenvolvimento de MCGA com um núcleo de equações primitivas (Bourke, 1972, Held e Suarez 1994) e um conjunto simplificado de esquemas de parametrizações físicas. A resolução do modelo é T30L7 (truncamento horizontal spectral de 30 números de ondas e sete níveis verticais, com camadas (meios-níveis) nos valores sigma de 0.08, 0.20, 0.34, 0.51, 0.68, 0.83, 0.95.). O pacote de parametrizações é baseado nos mesmos princípios físicos adotados em esquemas dos MCGAs mais complexos, tais como: convecção (esquema simplificado do fluxo de massa), condensação em larga-escala, nuvens, radiação de ondas curtas (duas bandas espectrais), radiação de ondas longas (quatro ondas espectrais), fluxos de superfície de momentum e energia (fórmula aerodinâmica), e difusão vertical. Detalhes do esquema de parametrização física simplificada podem ser encontrados em Molteni (2003). O modelo SPEEDY com dinâmica spectral é um modelo hidrostático, em coordenadas sigma, modelo transformada spectral na forma vorticidade-divergência descrita por Bourke (1974), com tratamento semi-implícito das ondas de gravidade. As variáveis

prognósticas básicas são vorticidade (Vor), divergência (Div), temperatura absoluta (T) e logaritmo de pressão na superfície (log(ps)); e uma variável adicional a umidade específica (Q). As saídas pós-processadas do modelo estão numa grade de tamanho 96 x 18 x 7, em níveis de pressão (100, 200, 300, 500, 700, 850, 925 hPa).

Redes Neurais Artificiais (RNA) são sistemas computacionais com processamento altamente paralelo e distribuído e que apresentam a capacidade de aprender e armazenar conhecimento experimental. Compostas por unidades de processamentos simples (nós ou neurônios) que computam determinadas funções matemáticas (normalmente não lineares). Tais unidades são interligadas por um grande número de conexões associadas a "pesos" que armazenam o conhecimento representado no modelo e servem para ponderar a entrada recebida por cada neurônio da rede. Perceptrons Multicamadas (PMC) (Gardner, 1998; Haykin, 1991) é uma arquitetura de RNA onde as interconexões das entradas à camada de saída têm pelo menos uma camada de neurônios intermediária, denominada de camada escondida. Definido um vetor de entradas $x = [x_0, x_1, x_2, x_3, \dots, x_N]$ e um vetor de saídas $s = [s_0, s_1, s_2, s_3, \dots, s_N]$, uma rede neural multicamadas realiza um mapeamento complexo $s = \Psi(w, x)$, parametrizado pelos pesos sinápticos w . Um conjunto de procedimentos bem definidos é utilizado para adaptar os pesos w de uma RNA a fim de produzir uma saída desejada chamado algoritmo de aprendizado de *retro-propagação do erro*. Este algoritmo é um processo supervisionado, onde a rede recebe vetores de entrada com sua respectiva resposta ou saída desejada. O processo de treinamento determina os pesos sinápticos com a minimização do *erro* entre a saída calculada pela rede e a resposta esperada para os vetores de entrada. Após o aprendizado, o conjunto de pesos sinápticos é capaz de ativar os neurônios da rede e obter resultados para entradas que estavam fora do conjunto de dados de treinamento, generalizando a informação aprendida.

O experimento foi realizado com o modelo SPEEDY na configuração citada foi executado por um período longo de integrações para criar o estado "verdadeiro" para iniciar as integrações do modelo. A assimilação de dados do LETKF foi executada com 415 estações de observações sintéticas das cinco variáveis prognósticas para gerar os vetores de análise e obter a saída esperada para treinar as redes neurais. As execuções foram feitas para cada dia do mês de janeiro de 1982, com assimilação de observações a cada 6 horas (00, 06, 12, 18 UTC). O conjunto de previsões do LETKF possui 30 membros e as "perturbações" consistem de números randômicos com distribuição Gaussiana, e as observações com adição de ruídos aleatórios adicionados ao estado "verdadeiro". Destas integrações foram obtidos os vetores de entrada para a rede neural, a previsão obtida pelo SPEEDY e as observações inseridas no LETKF. Neste processo foram inseridas observações em pontos de grades vizinhos com influência dos pontos de grade com observações. Este cálculo foi baseado na distância de seu

vizinho: $y_m^o = \left(\sum_{m=1}^N y_m^o \cdot \frac{1}{r_m^2} \right)$ onde N é o número de pontos de grade vizinhos sem observações

mas dentro de um quadrante Δ_r , a distância de cada ponto de observação), m é o ponto de grade dentro de $\Delta_r = (x_i^f - x_i^o)^2 + (x_j^f - x_j^o)^2 + (x_k^f - x_k^o)^2$, os subscritos i, j, k , são as localizações dos pontos de grade nas dimensões (latitude, longitude, nível) respectivamente. Foram coletados 345.120 pontos de observações ps e 1.477.956 pontos de observações u, v, t e q como entradas das respectivas redes neurais para toda a área global do modelo.

Foram desenvolvidas cinco PMC para cada variável prognóstica (ps, u, v, t e q) com duas entradas (modelo e previsão), um neurônio na camada de saída (análise), treze neurônios em uma camada escondida com função de ativação sigmóide para garantir a não linearidade do problema. A rede foi treinada inserindo valores de entrada de um ponto de

grade a cada vez, ou seja, a análise é feita em cada ponto de grade onde tem observação. A generalização foi iniciada com um ciclo de assimilação feita pela RNA no primeiro horário de 00 UTC em primeiro de fevereiro de 1982, gerando previsão com o modelo SPEEDY e a condição inicial gerada pela RNA para as rodadas posteriores até o final do mês. O LETKF foi rodado para comparação no mesmo período.

3.RESULTADOS E CONCLUSÕES

Os processos de entrada e saída dos valores dos pontos de grade em tempo intermitente foram desenvolvidos para integrações de curto prazo do ciclo previsão análise para as variáveis prognósticas: componentes da velocidade do vento zonal e meridional (u, v), temperatura(T), umidade específica (q) e pressão na superfície (ps). Os resultados apresentam campos de análise (q, T, e ps) gerados pela RNA e campos de análise gerados pelo LETKF do uma realização (1/fevereiro/1982 às 12 UTC) para variáveis prognósticas. Veja figuras de campos globais de número 1 a 3; mostram que a aplicação de RNA-PMC como sistema de assimilação de dados consegue reproduzir o sistema de assimilação LETKF para um tempo curto de integração (12) e generalização.

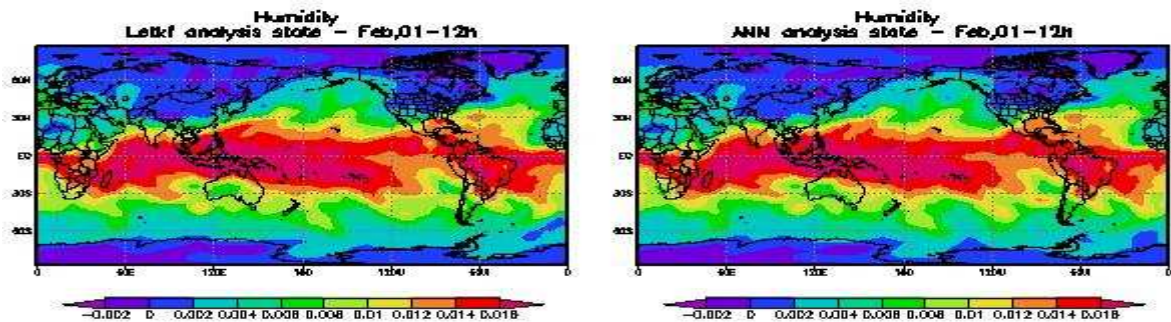


Figura 1 – Campos de análise (a) LETKF e (b) RNA para a variável umidade.

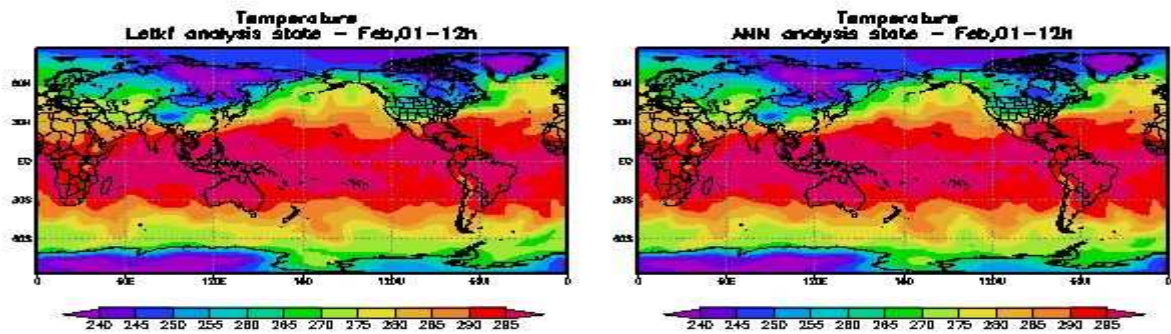


Figura 2 – Campos de análise (a) LETKF e (b) RNA para a variável temperatura.

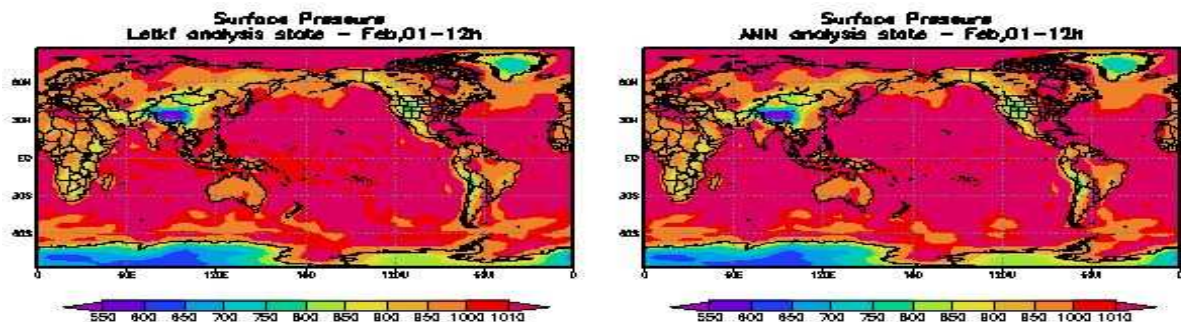


Figura 3 – Campos de análise (a) LETKF e (b) RNA para a variável pressão na superfície.

O desempenho de um sistema pode ser medido pela complexidade de seu algoritmo e também por medidas numéricas de tempo de execução que é apresentado na Figura 4. A figura apresenta 112 rodadas em ciclo de assimilação, feitas por 30 dias em com 2075 observações nos horários de (06 e 18 UTC) e 12035 observações nos horários de (00, 12 UTC), o tempo foi medido de milésimos de segundos, verifica-se que a execução da assimilação do LETKF (fora a execução do conjunto de previsões) é executada em tempo real em média de 95000 milésimos de segundo (1min, e 50 segundos), enquanto a rodada da assimilação ativando a RNA/PMC leva em média 300 milésimos de segundos. Estas medidas comprovam que o desempenho computacional da RNA testado foi superior ao desempenho apresentado pelo sistema LETKF. Resultados estes que comprovam a eficiência computacional da rede neural para o problema de assimilação de dados atmosféricos.



Figura 4 – Campos de análise (a) LETKF e (b) RNA para a variável pressão na superfície

Neste trabalho foram apresentados resultados preliminares desta pesquisa para obter análises para o MACG SPEEDY. Foi apresentada apenas uma estratégia de treinamento, considerando arquitetura simples da PMC, e a área global de modelo SPEEDY. Os testes foram realizados para um ciclo de curto prazo (10 dias), entretanto a figura 4 evidencia o ganho de eficiência computacional obtido com o novo método. Novas investigações estão sendo feitas para atestar a eficácia do método em ciclos maiores.

4. REFERÊNCIAS

- W. BOURKE; A multilevel spectral model: formulation and hemispheric integrations. *Mon. Wea. Rev.*, 102, 687-701, 1974.
- F. BOUTTIER and P. COURTIER. Data assimilation concepts and methods. Meteorological Training Course Series, ECMWF, Reading, UK, 75pp., 1998.
- R. DALEY, Atmospheric data analysis, Cambridge University Press, 1991.
- M.W. GARDNER, S.R. DORLING, Artificial neural networks: the multilayer perceptron, *Review of Applic. in the Atmospheric Sciences, Atmospheric Environment*, 32(14/15),(1998) 2627-2636.
- G. EVENSEN. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *J. Geophys. Res.*, 99(C5):10143–10162, 1994.
- S. HAYKIN, Redes neurais: princípios e práticas. 2 ed. Porto Alegre: Bookmann, 200.
- R. E. KALMAN. A new approach to linear filtering and prediction problems. *Trans. of the ASME–Journal of Basic Engineering*, 82(Series D):35–45, 1960.

M.L. Held, . L. , Suarez; A proposal for the intercomparison of dynamical cores of atmospheric general circulation models. *Bull Am Meteorol Soc* 75: 1825–1830, 1994

E. KALNAY. Atmospheric modeling, data assimilation and predictability. Cambridge University Press, New York, 2003.

A. H. JAZSWINSKI; Stochastic Processes and Filtering Theory, Academic Press, New York, USA, 1970.

T. Miyoshi, T., A review of the deterministic ensemble Kalman filtering methods and related techniques. Masters Scholarly Paper. Department of Meteorology, University of Maryland, College Park, 20pp, 2004.

F. MOLTENI. Atmospheric simulations using a gcm with simplified physical parameterizations. i: model climatology and variability in multi-decadal experiments. *Clim. Dyn.*, 20:175–191, 2003.

F.P. HARTER, "Redes Neurais Recorrentes Aplicadas à Assimilação de Dados em Dinâmica Não Linear", Tese de Doutorado, Computação Aplicada, CAP-INPE, São José dos Campos (SP), 2004.

F.P. HARTER, H.F. de 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.

A. G. NOWOSAD, Novas Abordagens para Assimilação de Dados Meteorológicos, tese de Doutorado, Computação Aplicada, CAP-INPE, São José dos Campos (SP), 2001.

E. OTT, B. R. HUNT, I. SZUNYOGH, A. V. ZIMIN, E. J. KOSTELICH, M. CORAZZA, E. KALNAY, D. J. PATIL and J. A. YORKE; A local ensemble Kalman filter for atmospheric data assimilation, *Tellus*, 56(a), 415-428, 2004