



**Assimilação de Dados Via Método 3D-Var em Dinâmica Caótica do Modelo de Lorenz**  
Variational Data Assimilation in Chaotic Regime by Lorenz Model

*Fabício Pereira Härter; Yoshihiro Yamasaki & Vinicius Carvalho Beck*

*Universidade Federal de Pelotas - Faculdade de Meteorologia*

*Avenida Engenheiro Ildefonso Simões Lopes, 2751*

*Arco-Íris 96060-290 Pelotas, RS – Brasil*

*E-mails: fabricio.harter@ufpel.edu.br; yamasaki07722@gmail.com; vonoco@gmail.com*

Recebido em: 22/08/2014    Aprovado em: 27/02/2015

DOI: [http://dx.doi.org/10.11137/2015\\_1\\_73\\_80](http://dx.doi.org/10.11137/2015_1_73_80)

## Resumo

Neste trabalho investiga-se a assimilação de dados pelo método 3D-Var no modelo de Lorenz. Este modelo foi escolhido, porque representa os diferentes modos do escoamento da atmosfera definidos em função de determinados parâmetros. Modelos simplificados, permitem avaliações do comportamento do escoamento atmosférico mais dispendiosas de serem realizadas em modelos de equações primitivas, com suas várias interações não lineares em diversas escalas. O trabalho indica a necessidade de aplicar técnicas de assimilação de dados em modelos altamente sensíveis às condições iniciais. Mostra-se que o método 3D-Var é bastante dependente do número de ciclos durante processo de minimização da função custo. São necessárias no mínimo 1.000 interações para que a técnica seja eficiente. Todavia, esta estratégia falha para condições iniciais acima de determinado grau de ruído, no caso avaliado, 40%. Mostrou-se também, a dificuldade de se realizar assimilação de dados quando os graus de liberdade do modelo são muitos maiores que o número de observações a serem assimiladas.

**Palavras-chave:** Assimilação de Dados; Variacional; Lorenz

## Abstract

This paper explores some aspects of data assimilation by 3D-Var method implemented in the Lorenz model. This system was chosen because, although simple, has similar structure to the atmosphere under certain parameters. Conceptual models allow difficult judgments to be made in primitive equation models with their various nonlinear interactions at various scales. It is shown the need to apply data assimilation techniques in highly sensitive models to initial conditions. It was also shown that 3D-Var method is very dependent on the number of cycles during the process of minimizing of the cost function. In the experiment performed on this work was found that at least 1,000 interactions are necessary to solve this minimization problem. However the technique fail to 40% noise added to initial conditions, comparing to control. Additional experiments show the difficulty of performing high quality data assimilation in chaotic dynamics for underdetermined systems, such as the Earth's atmosphere.

**Keywords:** Data Assimilation; 3D-Var; Lorenz

## 1 Introdução

A previsão de tempo realizada por diversos centros nacionais de vários países, utilizam Modelos Numéricos de Previsão de Tempo (MNPT). Estes modelos são representados por conjuntos de equações diferenciais parciais acopladas, portanto, dependentes das condições iniciais, obtidas por análise objetiva (Kalnay, 2003). Na construção da análise, podem ser utilizadas técnicas de assimilação de dados com o objetivo de combinar um campo a priori, em geral oriundo de modelagem do escoamento, chamado “*background*” em inglês. Atualmente, utilizam-se como *background* ou *informação a priori* previsões numéricas. Estas previsões apresentam incertezas devido a insuficiência física representadas pelas equações do modelo em relação a realidade do escoamento atmosférico e pela aproximação numérica empregada na solução das equações diferenciais acopladas. As observações apresentam erros de diferentes causas, dentre as quais, a limitação de precisão dos instrumentos, de calibração, de exposição, falhas, etc.

Assimilação de Dados é um procedimento no qual se combina previsão de curto prazo com observações para a construção da análise, considerando-se o conhecimento da estatística dos erros de previsão e observação. Os estados da arte em assimilação de dados incluem os métodos Variacionais tridimensionais (3D-Var) e quadridimensionais (4D-Var) e aplicação de filtro de Kalman (KF) com suas variantes (Kalnay, 2003).

Todavia, mesmo que os modelos fossem perfeitos e as observações exatas, a natureza caótica do sistema impediria que a previsão fosse exata (100% de acertos). Além disso, a assimilação de dados em modelos atmosféricos tem a dificuldade associada ao número de graus de liberdade dos modelos de previsão sinóptica ( $\sim 10^7$ ). Contudo, simplificações nos métodos associados a alta dimensão do sistema tornam-se menos importantes, à medida que a capacidade computacional aumenta.

O modelo proposto por Lorenz (1963), devido a sua simplicidade matemática e sua similaridade com padrões atmosféricos, tem sido amplamente aplicado para estudos de técnicas de assimilação de dados. Miller *et al.* (1994) aplicaram ao Modelo de Lorenz o filtro de Kalman Estendido (EKF), versão do KF para problema não-lineares. Os autores concluíram

que a assimilação apresenta bons resultados para frequências de observação relativamente altas, considerando-se o EKF com primeiro e segundo momento estatístico. Ao utilizarem terceiro e quarto momento, formulação de Monte Carlo, a trajetória é muito bem corrigida pela inserção do dado.

Huang & Yang (1996) avaliam o procedimento de assimilação com as equações de Lorenz e o método de assimilação 4D-Var. Os autores mostram que quanto maior for o ruído na CI, maior deve ser número de interações durante o minimização do gradiente da função custo. Os autores concluem também, que com observações de apenas uma das três variáveis do sistema e 40% de ruído na CI, a trajetória do modelo não é corrigida pela assimilação.

Mogensen & Huang (1998) exploram a capacidade do método 3D-Var estimar parâmetros do Modelo de Lorenz e apresentam resultados a cerca de variações na janela de assimilação e ruído nos dados observados. Os autores concluem que a estimativa de parâmetros é uma aplicação promissora para a abordagem variacional.

Miyoshi (2005) investiga variações do filtro de Kalman por Ensemble (EnKF), apresentando as seguintes conclusões baseadas no sistema de Lorenz: (1) calculando a covariância do erro de observação de maneira localizada (multiplicando a covariância do erro de observação pelo inverso de uma Gaussiana) é possível assimilar maior número de observações; (2) A performance do filtro de Kalman Localizado por Ensemble (LEKF) é igual ou melhor do que a do Ensemble do Filtro de Kalman por Raiz Quadrada (EnSR) com parâmetros similares; (3) ponderar a covariância em função do tempo e local é uma proposta promissora em assimilação de dados.

Hayden *et al.* (2011) destacam a desejável característica de rápida convergência do algoritmo 3D-Var em aplicações ao modelo de Lorenz tanto para o caso discreto como para o caso contínuo.

Law *et al.* (2013) estudaram o a performance do 3D-Var no modelo de Lorenz (1963) e concluíram que a metodologia de inflação de variância (por fator multiplicativo) tende a estabilizar o filtro, condicionando a convergência do algoritmo de minimização e melhorando a precisão da correção das trajetórias do escoamento.

Härter & Campos Velho (2008) emulam o EKF através de Redes Neurais Artificiais, com o objetivo de mostrar que no caso sobredeterminado (quando o número de observações é maior do que o número de graus de liberdade do modelo), a técnica de inteligência artificial é uma alternativa às demais técnicas.

O objetivo deste trabalho é avaliar o método 3D-Var no contexto de assimilação de dados em dinâmica não-linear. Especificamente, estimar o grau de sensibilidade do 3D-Var ao número de interações durante o processo de minimização da função custo pelo método Gradiente Descendente, ao nível de ruído nas CI do modelo e a quantidade de observações assimiladas.

## 2 Metodologia

A seguir, na seção 2.1 apresentam-se as equações diferenciais ordinárias que constituem o Modelo de Lorenz. Na seção 2.3, descreve-se o 3D-Var implementado em linguagem de programação Fortran 90.

### 2.1 Modelo de Lorenz

Saltzman (1962) interessado em movimentos atmosféricos de origem convectiva, desenvolveu um modelo simplificado (baseado nas equações do movimento, continuidade, e termodinâmica) para estudar o escoamento em uma camada de fluido com profundidade constante. Quando a diferença de temperatura entre as superfícies inferior e superior do fluido for mantida constante, o sistema permanece em equilíbrio. Neste caso se diz que o sistema possui solução estável. No entanto, se o sistema em questão for o sistema terra-atmosfera, pode-se descrever o processo que ocorre dentro da camada de fluido em um referencial não-rotante da seguinte maneira: da radiação incidente sobre a terra, emitida pelo sol, uma parte é absorvida e outra parte é refletida para a atmosfera. As camadas mais altas da atmosfera vão resfriando por perda radiativa para o espaço, de forma que na parte superior da atmosfera o ar fica mais frio e mais pesado do que nas camadas próximas a superfície, onde se concentra o ar mais quente e mais leve. Este ar mais leve tende a subir, enquanto o mais pesado tende a descer na atmosfera,

caracterizando uma circulação direta da convecção atmosférica. Esse problema de transporte, com camadas oscilantes de ar frio e quente, gera movimentos convectivos em diferentes escalas e modos característicos do escoamento, podendo ser desde a turbulência térmica caótica até sistemas altamente organizados multiescala, como furacões. Neste caso se diz que o sistema possui solução estável periódica, semi-periódica ou definitivamente aperiódica, (Lorenz, 1963).

Lorenz (1963), mais interessando na natureza não periódica das soluções do modelo de Saltzman do que nos modos oscilatórios do problema convectivo, expandiu as variáveis de estado do modelo em série de Fourier e, retendo apenas os termos de baixa ordem, obteve o seguinte sistema acoplado de três equações diferenciais ordinárias de primeira ordem não-linear

$$\frac{dX}{d\tau} = -\sigma(X-Y), \quad \frac{dY}{d\tau} = rX - Y - XZ, \quad \frac{dZ}{d\tau} = XY - bZ,$$

onde  $\tau = \pi H^2(1+f^2)kt$  é o tempo adimensional, com  $H$ ,  $f$ ,  $k$  e  $t$  sendo respectivamente altura (ou profundidade) da camada, número de onda, condutividade térmica e tempo.  $\sigma = k^{-1}\nu$  é o número de Prandtl, uma quantidade adimensional que depende da natureza do fluido e, em menor extensão, de sua temperatura;  $\nu$  é a viscosidade cinemática do fluido. O parâmetro  $r = R/R_c$  é o número de Rayleigh (representa a diferença de temperatura entre as superfícies), com  $R = g\alpha H^3 \delta T \nu^{-1} k^{-1}$ , sendo  $R_c = \pi^4 f^{-2}(1+f^2)^3$ , onde  $g$  é a aceleração da gravidade e  $R_c$  é o número de Rayleigh crítico. O mínimo valor de  $R_c$  é  $27\pi^4/4$ , ocorrendo quando  $f^2 = 0,5$ .

Nas equações do sistema de Lorenz,  $X$  é proporcional a intensidade do movimento convectivo,  $Y$  é proporcional a diferença de temperatura entre as correntes ascendentes e descendentes do fluido. Sinais similares de  $X$  e  $Y$  significa que o fluido quente está em movimento ascendente e o fluido frio em movimento descendente. A variável  $Z$  é proporcional a perturbação do perfil de temperatura vertical, sendo que os valores positivos indicam forte gradiente próximo a fronteira.

## 2.2 Método Variacional Tridimensional

No método 3D-Var é definida uma função custo, proporcional ao quadrado da distância entre a análise e o background e entre a análise e as observações. O mínimo desta função (equação abaixo) é a análise.

$$J_o(W) = \frac{1}{2} \{ [y_o - H(W)]^T \mathbf{R}^{-1} [y_o - H(W)] + (W - W_b)^T \mathbf{B}^{-1} (W - W_b) \}$$

onde  $W = W(X, Y, Z)$  é a matriz de estados,  $W_b$  é o campo de *background* (integração curta do modelo ou climatologia),  $y_o$  é o vetor de observações,  $H$  é o operador de observação (projeta a matriz de estado no espaço do vetor de observações),  $B$  é a matriz de covariância dos erros de estimativa e  $R$  é a matriz de covariância dos erros de observação.

A função custo  $J$ , mede:

- distância entre o campo  $W$  e as observações (primeiro termo);
- distância entre a verdade e o background;

O mínimo de  $J$  é obtido para  $W = W_a$ , ou seja, a análise. A minimização é realizada através de processos iterativos por algoritmos como *steepest descent* (gradiente descendente), como:

$$W_0^{f,i+1} = W_0^{f,i} - \alpha \nabla J(W_0^{f,i}),$$

onde  $i$  é o número de iterações e  $\nabla J$  é o gradiente da função custo em relação ao estado inicial  $w_0^{f,i}$ . Se a iteração converge,  $w_0^{f,i}$  aproxima-se do estado inicial  $w_0^{f,\infty}$ , que satisfaz  $J = \min(J)$ .

## 3 Resultados

Com o propósito de avaliar o método de assimilação proposto, foram conduzidos 4 experimentos. No primeiro experimento mostra-se o comportamento do sistema ao se inserir 10% de ruído nas condições iniciais do modelo em estado caótico e não-linear, Experimento 1 (EXP1). No Experimento 2 (EXP2) avalia-se os resultados do 3D-Var ao se variar os ciclos de minimização da função custo. No Experimento 3 (EXP3) avalia-se a sensibilidade do método de assimilação aplicado

as equações de Lorenz com 20% e 40% de ruído na CI e no Experimento 4 (EXP4) avalia-se a solução do 3D-Var com apenas uma observação no passo de tempo 180 na variáveis  $X$ ,  $Y$  e  $Z$  do sistema e em apenas na componente  $Y$  do sistema. Em todos os experimentos o modelo foi integrado com um simples métodos de diferenças finitas avançado no tempo e passo de tempo adimensional de 0,01.

### 3.1 Experimento 1 – (EXP1)

No seminal trabalho de Lorenz (Lorenz, 1963), foi demonstrado que para  $\sigma=10$  e  $b=8/3$  o número de Rayleigh crítico é  $R_c=24,74$ . Isto significa que para  $r > 24,74$  o sistema de Lorenz é caótico e para  $r \leq 24,74$  o sistema é não-linear. No regime caótica a trajetória do sistema é altamente sensível as condições iniciais. A Figura 1a ilustra a integração da componente  $X$  do modelo por 500 passos de tempo (5 unidades de tempo adimensionais) em regime não-linear ( $\sigma = 10$ ,  $b = 8/3$ ,  $r=10$ ), com as seguintes CI:  $X = 1,0$ ;  $Y = 3,0$  e  $Z = 5,0$  (linha contínua) e  $X = 1,1$ ;  $Y = 3,3$  e  $Z = 5,5$  (linha tracejada). O mesmo é refeito para a integração do modelo em regime caótico ( $\sigma = 10$ ,  $b = 8/3$ ,  $r = 32$ ).

O EXP1 mostra a sensibilidade de modelo em regime caótico a pequenas variações na CI, o que não ocorre quando o regime é não linear. O mesmo ocorre com as componentes  $Y$  e  $Z$ , não apresentadas neste texto. Para  $r = 32$ , o sistema de Lorenz deve ter o comportamento similar a atmosfera terrestre. Portanto, este experimento mostra claramente a necessidade de se escolher a CI mais precisa possível para a Previsão Numérica de Tempo (PNT), ou seja, este experimento ilustra a importância da assimilação de dados para a PNT.

### 3.2 Experimento 2 (EXP2)

Neste experimento avalia-se a resposta da 3D-Var ao se variar o número de iterações durante o processo de minimização da função custo, doravante denominado de Ciclos. O parâmetro  $\alpha$  do método de minimização é  $5 \times 10^{-4}$ . Neste experimento assimila-se 3 observações, geradas adicionando-se no máximo 10% de ruído aleatório com distribuição gaussiana ao controle. O controle consiste na integração do modelo em regime caótico sem ruído, conforme descrito na seção 3.1. Confronta-se o controle

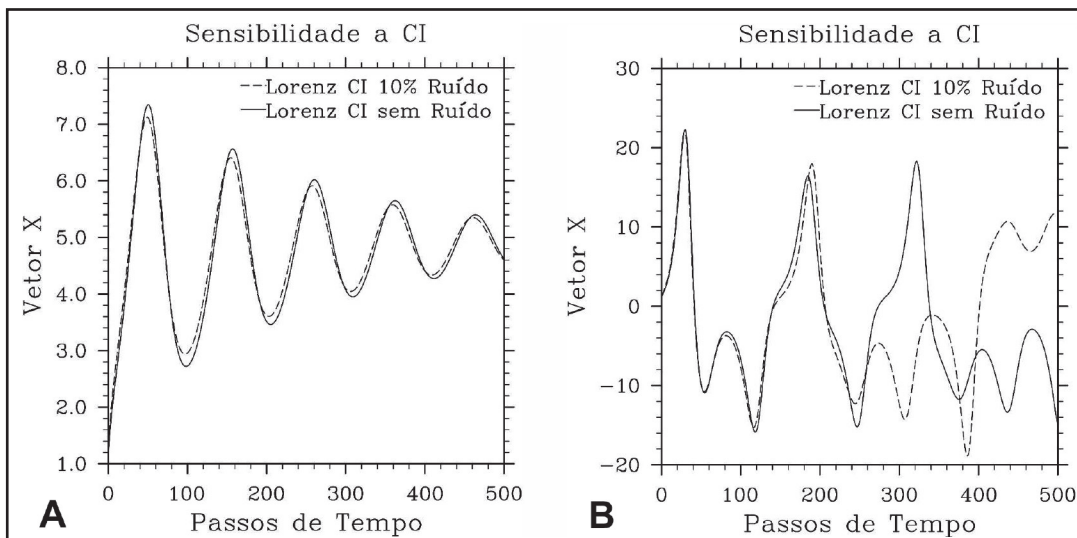


Figura 1 Sistema de Lorenz em (a) regime não-linear, (b) regime caótico. Na integração representada pelas linhas tracejadas adicionou-se 10% de ruído nas CI.

com a integração do modelo em regime caótico, com assimilação por 3D-Var durante a janela de assimilação de 200 passos de tempo (equivalente a 2 unidades de tempo adimensionais para  $\Delta t = 0,01$ ) e CI com 10% de ruído ( $X = 1,1$ ;  $Y = 3,3$  e  $Z = 5,5$ ). Avaliam-se os resultados para ciclos de 100, 200, 300, 500, 1000 e 3000 interações durante o processo de convergência do gradiente descendente.

O EXP2, Figura 2, mostra que com 100, 200, 300 e 500 interações no processo de minimização do gradiente há um desacoplamento entre a “trajetória” (linha contínua) e o 3D-Var (linha tracejada), principalmente no final do período de integração. Este desacoplamento é típico do comportamento caótico do sistema. Entretanto, aumentando o número de ciclos para 1000 e 3000 iterações, para o  $\alpha$  utilizado, a assimilação das três observações (asterisco) pelo 3D-Var, faz com que a trajetória com a CI ruidosa convirja para a “trajetória verdadeira”.

### 3.3 Experimento 3 (EXP3)

No EXP3 procura-se descobrir o nível de ruído nas CI com o qual o 3D-Var consegue manter-se efetivo como um bom processo de assimilação. Para tal, faz integrações similares as feitas no EXP2, com a diferença de que adiciona-se 20% e 40% de ruído nas CI e explora-se apenas ciclos de 100, 1000 interações na minimização do gradiente da função J.

Na Figuras 3 a e 3 c grafam-se os resultados com 20% de ruído ( $X=1,20$ ;  $Y=3,60$  e  $Z=6,0$ ),

para 100 e 1.000 interações respectivamente e nas Figuras 3 b e 3 d grafam-se os resultados com 40% de ruído ( $X=1,40$ ;  $Y=4,20$  e  $Z=7,0$ ), para 100 e 1.000 interações respectivamente.

Devido a natureza caótica do sistema e ao pouco número de observações assimiladas, com 100 ciclos, a assimilação não faz com que a trajetória seja corrigida com 20% e nem com 40% de ruído nas CI. Com 1000 ciclos, e 20% de ruído o 3D-Var consegue corrigir a trajetória, mas com 40% de ruído o 3D-Var falha.

### 3.4 Experimento 4 (EXP4)

Nas Figuras 4, o experimento acima é repetido para ruído de 10% na CI e a assimilação é realizada em apenas um passo de tempo ( $180^\circ$ ). Sendo que nas Figuras 3 a, 3c e 3 e grafam-se os resultados para observações assimiladas nos vetores X, Y e Z; enquanto nas Figuras 3 b, 3 d e 3 f grafam-se os resultados para observações assimiladas apenas no vetor Y.

A assimilação 3D-Var em apenas um ponto da janela de assimilação falha ao tentar aproximar a trajetória ruidosa do controle, para 100 e 1.000 ciclos de minimização do Gradiente Descendente. Este resultado é similar ao se comparar a assimilação realizada nas três variáveis do modelo com a assimilação realizada apenas na variável Y. Este experimento ilustra o comportamento da assimilação quando poucos dados são assimilados.

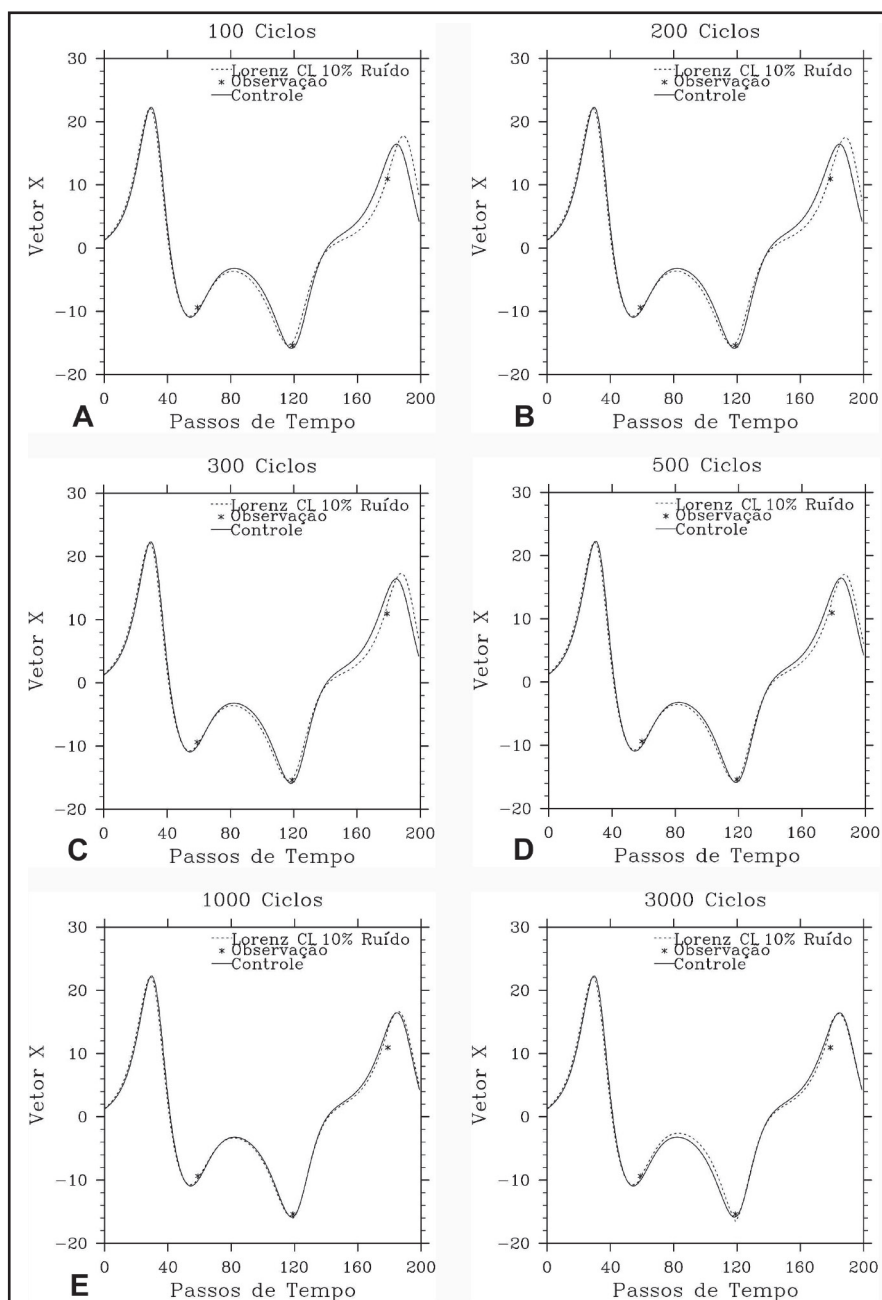


Figura 2 Componente X do modelo de Lorenz (linha contínua), observações (asterisco) e 3D-Var (linha tracejada) com (a) 100, (b) 200, (c) 300, (d) 500, (e) 1000 e (f) 3000 interações do gradiente descendente.

#### 4 Considerações Finais

Neste trabalho, testou-se o método 3D-Var aplicado a assimilação de dados sintéticos no modelo de Lorenz, em regime caótico. Este modelo, apesar de matematicamente simples, possui similaridade com padrões atmosféricos. Através dos experimentos realizados, adquire-se conhecimento sobre a aplicabilidade do método variacional ao problema proposto, bem com suas limitações, o que é difícil num modelo completo de equações primitivas.

Mostrou-se ensaios experimentais em que a adição de 10% de ruído nas CI do modelo, para  $\sigma=32$  (número de Rayleigh associado a regime caótico), leva a perda de previsibilidade após certo período de integração, o que não ocorre para outros regimes de escoamento periódicos e semiperiódicos, de característica linear. Desta forma, evidencia-se numericamente a importância da Assimilação de Dados para a previsão de tempo.

Foi mostrado também, que o método 3D-Var é bastante dependente do número de ciclos do

Figura 3 Componente X do modelo de Lorenz com assimilação por 3D-Var. 3(a) e 3(c) resultados com 20% de ruído ( $X=1,20$ ;  $Y=3,60$  e  $Z=6,0$ ), para 100 e 1.000 interações respectivamente. 3(b) e 3(d) resultados com 40% de ruído ( $X=1,40$ ;  $Y=4,20$  e  $Z=7,0$ ), para 100 e 1.000 interações respectivamente.

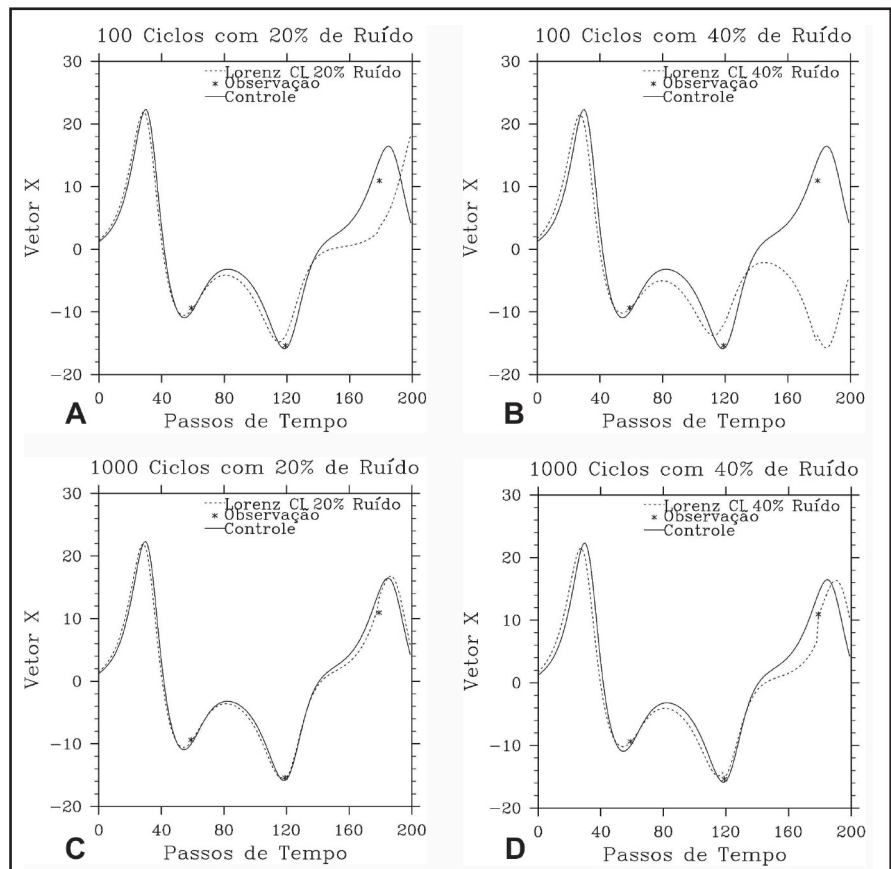
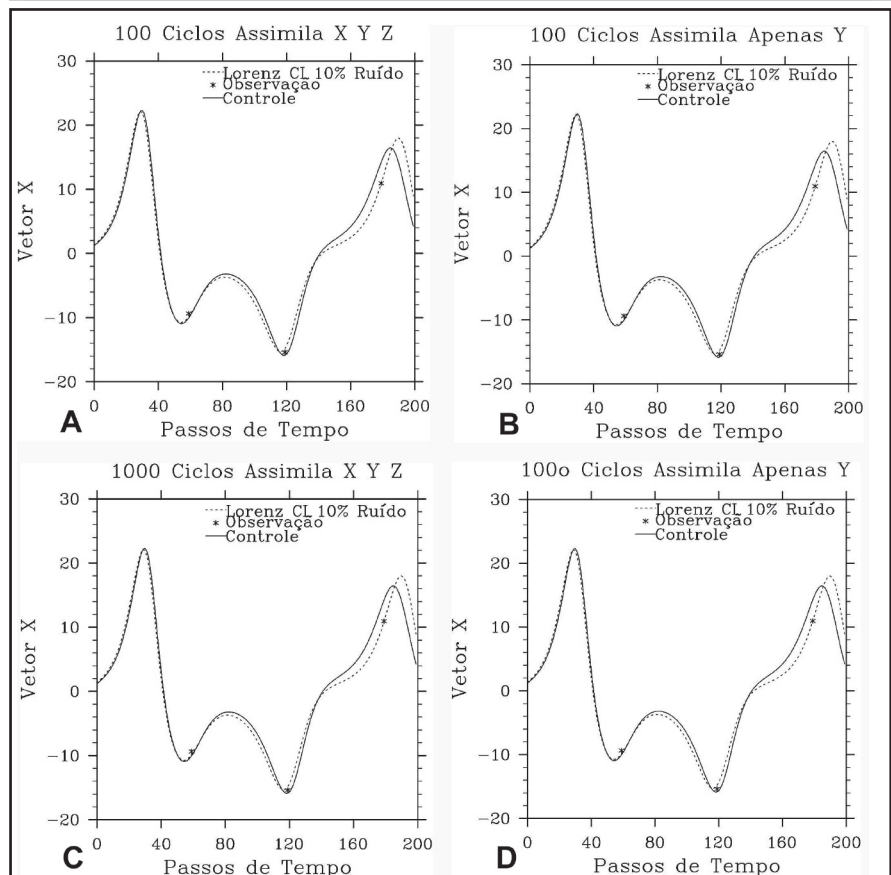


Figura 4 Componente X do modelo de Lorenz com assimilação 3D-Var de observações com 10% de ruído na CI em apenas um passo de tempo ( $180^\circ$ ). Figuras 3 a e 3 c observações assimiladas nos vetores X, Y e Z. Figuras 3 b e 3 d e 3 f observações assimiladas apenas no vetor Y.



processo de minimização da função custo. Em poucas interações, no caso mostrado 100, não foi o suficiente para que a trajetória ruidosa em regime caótico fosse corrigida pelo processo de assimilação. Todavia, com 1.000 interações a técnica foi eficiente. O mesmo não acontece para CI muito ruidosa, no caso avaliado 40% de ruído.

Por fim, avaliou-se a capacidade do 3D-Var em situações com escassez de dados observacionais. Simulou-se a situação na qual há dado para assimilação em apenas um instante da janela de assimilação. Neste caso, tanto com assimilação das três variáveis prognósticas como com a assimilação de apenas a variável  $Y$ , o 3D-Var falha na correção da trajetória, frente a expectativa.

## 5 Referências

- Härter, F. & Campos Velho, H. 2008. New Approach to Applying Neural Network in Nonlinear Dynamic Model. *Applied Mathematical Modelling*, 12: 1-25.
- Hayden, K; Olson, E.& Titi, E. 2011. Discrete data assimilation in the Lorenz and 2D Navier–Stokes equations. *Physica D*, 240: 1416-1425
- Huang, X.-Y & Yang, X. 1996. Variational Data Assimilation with Lorenz Model. Hirlam Technical Report, 26, S-601 76 SMHI Norrköping, Sweden.
- Kalnay, E. 2003. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge. Cambridge University Press. 369 p.
- Law, K.; Shukla, A. & Stuart, A. 2013. Analysis of the 3DVAR Filter for the Partially Observed Lorenz '63 Model. *Discrete and Continuous Dynamical Systems*, 3: 1061-1078.
- Lorenz, E. 1963. Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, 20 (2): 130 – 141.
- Miller, R.; Guil, M. & Gauthiez, F. 1994. Advanced Data Assimilation in Strongly Nonlinear System. *Journal of the Atmospheric Sciences*, 51(8): 1037–1056.
- Miyoshi, T. 2005. *Ensemble Kalman Filter Experiments with a Primitive Equation Global Model*. Programa de Pós-graduação em Meteorologia, University of Maryland, Tese de Doutorado, 197p.
- Mogensen, K. & Huang, X.-Y. 1998. Variational Parameter Estimation with the Lorenz Model. Danish Meteorological Institute, Copenhagen, Denmark. Technical Report, DK2100.
- Saltzman, B. 1962. Finite Amplitude Free Convection as an Initial Value Problem. *Journal of the Atmospheric Sciences*, 19: 329 – 341.