



UNIVERSIDADE FEDERAL DO AMAPÁ  
PRÓ-REITORIA DE ENSINO E GRADUAÇÃO  
CURSO DE LICENCIATURA PLENA EM MATEMÁTICA

*ASSIMILAÇÃO DE DADOS NO MODELO DE LORENZ*

MACAPÁ-AP

2012



MEYCE PEREIRA DA ROCHA

*ASSIMILAÇÃO DE DADOS NO MODELO DE LORENZ*

Trabalho de conclusão de curso apresentado ao colegiado de Matemática da Universidade Federal do Amapá, como parte das exigências para a obtenção do título de Licenciatura Plena em Matemática, sob orientação do Prof. Dr. José Walter Cárdenas Sotil.

MACAPÁ-AP

2012

MEYCE PEREIRA DA ROCHA

*ASSIMILAÇÃO DE DADOS NO MODELO DE LORENZ*

Trabalho de Conclusão de Curso apresentado como pré-requisito para obtenção do título de Licenciatura Plena em Matemática da Universidade Federal do Amapá, submetida à aprovação da banca examinadora composta pelos seguintes membros:

---

Prof. Dr. José Walter Cárdenas Sotil

---

Prof. Dr. Erasmo Senger

---

Prof. Dr. Guzmán Eulálio Isla Chamilco

Avaliado em: \_\_\_\_ / \_\_\_\_ / \_\_\_\_

MACAPÁ-AP

2012

## AGRADECIMENTOS

A Deus pela benção concedida.

Aos meus pais e minha avó materna que contribuíram diretamente para minha formação.

Ao meu esposo pela paciência, carinho e dedicação que teve comigo durante os quatro anos de curso.

As minhas irmãs por me apoiarem sempre.

Aos meus colegas de curso que me ajudaram nos estudos e na vida.

A todos os professores que fizeram parte da minha formação, em especial ao meu orientador o Prof. Dr. José Walter Cárdenas Sotil pela dedicação, paciência e esforço para realização deste trabalho.

*A minha família a todos que contribuíram direta ou indiretamente para realização deste trabalho, ao meu orientador, e especialmente ao meu filho Lucas.*

*A Matemática apresenta invenções tão sutis que poderão servir não só para satisfazer os curiosos como, também para auxiliar as artes e poupar trabalho aos homens.*

(DESCARTES)

# Resumo

Neste trabalho final de curso são abordadas técnicas de assimilação de dados, no qual usando informação de diferente qualidade gera-se uma resposta mais próxima do estado verdadeiro do fenômeno em estudo. Usando o problema das duas temperaturas, o qual é um problema escalar simples, é desenvolvido métodos sequenciais e variacionais, os quais resultam ser equivalentes no sentido de fornecer a mesma solução ótima. Estes métodos são generalizados para problemas de maior complexidade usando assimilação sequencial no tempo. Para testar os métodos de assimilação é usado um modelo de Lorenz com três variáveis desconhecidas, o qual é um modelo caótico devido à não linearidade do modelo. As integrações são codificadas no Matlab usando Runge-Kutta de quarta ordem. Os resultados mostram que a análise ou solução ótima apresenta um erro muito menor que o erro da integração, na qual se perturba ligeiramente a condição inicial, mas sem assimilação. Realizando assimilação de dados periodicamente a análise consegue manter-se próxima da solução exata.

**Palavras-chave:** Assimilação de dados; métodos estatísticos, métodos variacionais, assimilação sequencial, análise, observações.

# Resumen

En este Trabajo Final de Curso se analizan las técnicas de asimilación de datos, en el que la calidad de diferentes mediante el uso de la información que genera una respuesta más cercana a la situación real del fenómeno en estudio. Utilizando el problema de las dos temperaturas, un problema que es escalar simple, se desarrolla y variacionales métodos secuenciales, que resultan ser iguales a fin de proporcionar la solución óptima mismo. Estos métodos se generalizan a los problemas más complejos con el tiempo la asimilación secuencial. Para probar los métodos de asimilación se utiliza un modelo con tres incógnitas de Lorenz, que es un modelo caótico debido a la no linealidad del modelo. Las integraciones se codifican en Matlab usando Runge-Kutta de cuarto orden. Los resultados del análisis muestran que la solución óptima o tiene un error mucho menor integración del error, el cual está ligeramente perturbar la condición inicial, pero no asimilación. Realización de un análisis regular la asimilación de datos puede permanecer cerca de la solución exacta.

**Palabras clave:** La asimilación de datos, métodos estadísticos, métodos variacionales, la asimilación secuencial, análisis, observaciones.



# Sumário

<b>Introdução</b>	<b>5</b>
<b>2 O Problema das duas Temperaturas</b>	<b>8</b>
2.1 Interpolação Estatística com Estimação dos Mínimos Quadrados . . . . .	8
2.1.1 Método Variacional . . . . .	11
2.1.2 Resumo das Equações . . . . .	13
<b>3 Assimilação de Dados Sequencial</b>	<b>15</b>
3.1 Assimilação de Dados para as Equações de Fluxo não Periódico de Lorenz	16
3.1.1 Integração do Modelo . . . . .	17
3.1.2 Perturbação do Modelo . . . . .	19
3.1.3 Assimilação de Dados . . . . .	19
<b>4 Resultados</b>	<b>22</b>
<b>Considerações Finais</b>	<b>31</b>

# Lista de Figuras

4.1	Solução exata do modelo de Lorenz . . . . .	22
4.2	Solução exata para X e Y . . . . .	23
4.3	Solução exata para X . . . . .	23
4.4	Solução exata no tempo(t)para X . . . . .	24
4.5	Solução exata e perturbada para X . . . . .	24
4.6	Solução exata, perturbada e assimilada para X . . . . .	25
4.7	Erro do modelo perturbado, e erro da análise para X . . . . .	25
4.8	Solução exata para Y . . . . .	26
4.9	Solução exata no tempo para Y . . . . .	26
4.10	Solução exata e perturbada para Y . . . . .	27
4.11	Solução exata, perturbada e assimilada para Y . . . . .	27
4.12	Erro da analise e erro do modelo perturbado para Y . . . . .	28
4.13	Solução exata para Z . . . . .	28
4.14	Solução exata no tempo para Z . . . . .	29
4.15	Solução exata e perturbada para Z . . . . .	29
4.16	Solução exata, perturbada e assimilada para Z . . . . .	30
4.17	Erro da analise e erro da solução perturbada para Z . . . . .	30

# Introdução

Assimilação de dados e previsão por conjuntos são de primordial importância nos modelos atuais de prognóstico de tempo e clima em meteorologia e oceanografia e tema de constantes pesquisas para aprimoramento dos resultados de previsão destes modelos nos principais centros de previsão dos Estados Unidos, Europa e Ásia. No Brasil, o Centro de Previsão de Tempo e Clima (CPTEC) é o principal centro de pesquisa em Meteorologia e Oceanografia. O CPTEC já conta com alguns esquemas de assimilação de dados de produtos provenientes de outros países, entretanto os resultados ainda não são totalmente satisfatórios. Assim, são necessárias pesquisas em assimilação de dados que priorizem a melhoria das previsões de tempo e clima no Brasil e América do Sul. O CPTEC também usa com sucesso esquemas de previsão de conjuntos para ampliar o tempo de validade de suas previsões, neste projeto analisamos estes modelos e pesquisamos melhorias que possam ser inserida nestes esquemas.

A previsão numérica é em boa parte um problema de condições iniciais: dada uma estimativa atual do estado da atmosfera, um modelo numérico simula sua evolução para obter uma previsão no estado futuro. Esta condição inicial se estabelece a partir da interpolação das observações disponíveis sobre os pontos da grade do modelo. Este processo de obtenção da condição inicial a partir das observações se denomina assimilação de dados. O principal problema deste processo é que a quantidade de dados disponíveis não é suficiente para inicializar o modelo em todos seus graus de liberdade (por exemplo, um modelo global com  $1^\circ$  de resolução horizontal e 40 níveis verticais poderá ter  $180 \times 360 \times 40 = 2,6 \times 10^6$  pontos de grade. Sendo que em cada ponto de grade estão definidas 7 variáveis prognósticas, com o que teríamos aproximadamente  $10^7$  graus de liberdade. Para uma janela temporal de  $\pm 3$  horas, existem normalmente  $10^4$  a  $10^5$  observações na atmosfera, ou seja duas ordens de magnitude menor que o número de graus de liberdade do modelo. De outro lado, a distribuição temporal e espacial dos dados não é uniforme,

existindo regiões como Europa e Norte América com quantidade de dados muito maior que outras regiões como a América do Sul. Portanto, é necessário usar informação previa (chamado first-guess, background ou chute inicial) para preparar as condições iniciais usadas pelo modelo. Inicialmente era usada a climatologia como first-guess com resultados pouco satisfatórios. Posteriormente foi usada uma previsão de curto prazo do próprio modelo como first-guest que combinadas de maneira adequada com os dados iniciais geram a condição inicial do modelo, isto é, os valores iniciais das variáveis prognósticas são definidas em cada ponto de grade do modelo.

O first-guess  $x_p$  é interpolada aos pontos de observação mediante um operador  $H(x_p)$  e convertida ao mesmo tipo que as variáveis observadas  $y_0$ . As diferenças entre as observações e o first-guess  $y_0 - H(x_p)$  são os incrementos observacionais. A análise  $x_0$  se obtém adicionando os incrementos observacionais ao first-guess do modelo com uns pesos  $W$  que são determinados com base nas covariâncias dos erros estatísticos da previsão e observação

$$x_0 = x_p + W[y_0 - H(x_p)]$$

Os diferentes esquemas de assimilação estão baseados neste método:

- a) Interpolação ótima: onde a matriz de pesos se determina minimizando os erros em cada ponto da grade.
- b) Métodos variacionais 3D e 4D com funções custo proporcionais ao quadrado da distância entre a análise, first-guess e as observações.
- c) Filtros de Kalman: onde a função peso é função das variâncias dos erros (Cárdenas, 2002; Judd, 2003; Pacheco, 2006).

Assim, o ciclo de assimilação de dados é uma continua integração do modelo que se combinam com as observações de modo que se permaneça o mais perto possível do estado real da atmosfera. A função do modelo é transportar informação de regiões com muitos dados a zonas com poucos dados para oferecer uma melhor estimación do estado da atmosfera. Os erros que se cometem no processo de assimilação impõem uma incerteza no prazo da validade da previsão, devido à não-linearidade da atmosfera, e portanto, dos modelos que aproximam sua dinâmica.

Neste trabalho estudamos métodos de assimilação de dados em um modelo caótico de Lorenz que simula o comportamento da atmosfera e oceano. Os métodos são descritos

completamente usando o problema das duas temperaturas. Simulações numéricas no Matlab mostram a vantagem de usar assimilação de dados em problemas de previsão de tempo.

## Capítulo 2

# O Problema das duas Temperaturas

O problema das duas temperaturas é um modelo simples, que pode ser formulado como segue:

”Dadas duas observações independentes da temperatura  $T_1$  e  $T_2$  de uma sala, determine a melhor estimativa da verdadeira temperatura da sala  $T_t$ ”.

Este modelo simples é importante porque permite formular as aproximações estatísticas e variacionais, sendo que sua metodologia poder ser transferida para problemas complexos e mais realistas.

### 2.1 Interpolação Estatística com Estimação dos Mínimos Quadrados

O problema das duas temperaturas pode-se reformular com a terminologia clássica da assimilação de dados. O problema de estimar a temperatura  $T_t$  de uma sala segue o seguinte procedimento:

- a) Uma primeira medida  $T_o$  da temperatura da sala é realizada com um termômetro.
- b) Assume-se que o termômetro tem um erro de medida  $\epsilon_o$ . Isto é, a medida do termômetro  $T_o$  difere em  $\epsilon_o$  do verdadeiro valor da temperatura da sala  $T_t$  :

$$T_o = T_t + \epsilon_o$$

- c) Assume-se que a medida do termômetro não tem viés, isto é, que a medida não tem erro sistemático ou tendência a apontar para um mesmo valor. Ou seja, a média ou

esperança matemática do erro da medida é zero:

$$\mu_o = E(T_o - T_t) = E(\epsilon_o) = 0$$

d) Denotamos a variância do erro  $\epsilon_o$  por  $\sigma_o^2$ . Em ausência de outra informação a melhor estimativa da temperatura da sala é  $T_o$  com exatidão  $\sigma_o$ .

e) Assume-se que outra informação sobre a temperatura da sala é conhecida, tal como a medida de outro termômetro com diferente exatidão, ou informação desta temperatura em outra data, mas em condições similares do ambiente. Seja  $T_b$  esta medida da temperatura da sala com erro  $\epsilon_b = T_b - T_t$  e variância do erro  $\sigma_b^2$ , a qual não apresenta viés:

$$\sigma_b = E(T_b - T_t) = E(\epsilon_b) = 0$$

f) Assume-se que os dois erros  $\epsilon_o$  e  $\epsilon_b$  não são correlacionados, isto é, o coeficiente de correlação é nulo:

$$\rho(\epsilon_o\epsilon_b) = 0$$

como

$$\rho(\epsilon_o\epsilon_b) = \frac{\text{cov}(\epsilon_o, \epsilon_b)}{\sigma_o\sigma_b} = \frac{E(\epsilon_o\epsilon_b) - \mu_o\mu_b}{\sigma_o\sigma_b} = 0$$

e  $\mu_1 = \mu_2 = 0$  temos:

$$E(\epsilon_o\epsilon_b) = 0$$

g) Procura-se uma melhor estimativa de  $T_t$  a partir dos valores das medidas  $T_o$  e  $T_b$ . Seja  $T_a$  esta estimativa, com erro  $\epsilon_a = T_a - T_t$  e variância do erro  $\sigma_a$ , a qual não apresenta viés:

$$\mu_a = E(T_a - T_t) = E(\epsilon_a) = 0$$

h)  $T_a$  é estimada como combinação linear de  $T_o$  e  $T_b$ :

$$T_a = \alpha T_o + \beta T_b$$

O problema agora é:

*”determinar os valores apropriados dos pesos  $\alpha$  e  $\beta$  para que  $T_a = \alpha T_o + \beta T_b$  seja uma boa aproximação para  $T_t$ ”*

Para resolver este problema, observa-se o fato de  $T_o$ ,  $T_b$  e  $T_a$  não apresentarem viés que

$$E(T_o) = E(T_b) = E(T_a) = E(T_t)$$

logo,

$$E(T_a) = \alpha E(T_0) + \beta E(T_b) \longrightarrow E(T_t) = \alpha E(T_0) + \beta E(T_b),$$

cancelando  $E(T_t)$  temos

$$\alpha + \beta = 1$$

Uma "boa aproximação" é obter  $T_a$  com variância mínima  $\sigma_a^2$ , logo reformulamos o problema como:

"Minimizar

$$\sigma_a^2 = E[(T_a - T_t)^2] = E\{[\alpha(T_0 - T_t) + \beta(T_b - T_t)]^2\} = E\{[\alpha\epsilon_o + \beta\epsilon_b]^2\}$$

sujeito à restrição

$$\alpha + \beta = 1"$$

Observando que

$$\sigma_a^2 = E\{[\alpha\epsilon_o + \beta\epsilon_b]^2\} = \alpha^2 E(\epsilon_o^2) + \beta^2 E(\epsilon_b^2)$$

e substituindo  $\beta = 1 - \alpha$ , temos

$$\sigma_a^2 = \alpha^2 \sigma_o^2 + (1 - \alpha)^2 \sigma_b^2$$

Para minimizar  $\sigma_a^2$  requeremos que

$$\frac{\partial \sigma_a^2}{\partial \alpha} = 0,$$

logo,

$$\frac{\partial \sigma_a^2}{\partial \alpha} = 2\alpha \sigma_o^2 + 2[1 - \alpha](-1)\sigma_b^2 = 2\alpha(\sigma_o^2 + \sigma_b^2) - 2\sigma_b^2 = 0.$$

Colocando  $\alpha$  em evidência, temos:

$$\alpha = \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2}$$

De  $\beta = 1 - \alpha$  resulta,

$$\beta = 1 - \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2} = \frac{\sigma_o^2 + \sigma_b^2 - \sigma_b^2}{\sigma_o^2 + \sigma_b^2} = \frac{\sigma_o^2}{\sigma_o^2 + \sigma_b^2}.$$

Obtidos os pesos  $\alpha$  e  $\beta$ , temos provado a seguinte proposição:

**Proposição 1.** Sejam  $T_0$  e  $T_b$  aproximações da temperatura  $T_t$ . Se

**I**  $\epsilon_o, \epsilon_b$  são respectivamente os erros de  $T_0$  e  $T_b$  ao aproximar  $T_t$



II a média dos erros é nula:  $\mu_o = E(\epsilon_o) = 0$  e  $\mu_b = E(\epsilon_b) = 0$

III  $\sigma_o^2$  e  $\sigma_b^2$  são respectivamente as variâncias dos erros  $\epsilon_o$ ,  $\epsilon_b$

IV o coeficiente de correlação dos erros é nulo:  $\rho(\epsilon_o, \epsilon_b) = 0$

Então, a melhor aproximação de  $T_t$  é dada pela temperatura  $T_a$ :

$$T_a = \left( \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2} \right) T_o + \left( \frac{\sigma_o^2}{\sigma_o^2 + \sigma_b^2} \right) T_b$$

com variância mínima  $\sigma_a^2$ :

$$\sigma_a^2 = \left( \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2} \right)^2 \sigma_o^2 + \left( \frac{\sigma_o^2}{\sigma_o^2 + \sigma_b^2} \right)^2 \sigma_b^2$$

### Prova

Só resta provar que a variância  $\sigma_a^2$  é mínima. Elevando ao quadrado o lado direito temos:

$$\sigma_a^2 = \frac{(\sigma_b^2)^2}{(\sigma_o^2 + \sigma_b^2)^2} \sigma_o^2 + \frac{(\sigma_o^2)^2}{(\sigma_o^2 + \sigma_b^2)^2} \sigma_b^2 = \frac{\sigma_o^2 \sigma_b^2 (\sigma_b^2 + \sigma_o^2)}{(\sigma_o^2 + \sigma_b^2)^2} = \frac{\sigma_o^2 \sigma_b^2}{\sigma_o^2 + \sigma_b^2} = \frac{1}{\sigma_o^{-2} + \sigma_b^{-2}}$$

logo,

$$\sigma_a^2 \leq \frac{1}{\sigma_o^{-2}} \quad e \quad \sigma_a^2 \leq \frac{1}{\sigma_b^{-2}}$$

Por tanto,

$$\sigma_a^2 \leq \sigma_o^2 \quad e \quad \sigma_a^2 \leq \sigma_b^2$$

Em particular, das contas acima se verifica:

$$\frac{1}{\sigma_a^2} = \frac{1}{\sigma_o^2} + \frac{1}{\sigma_b^2}$$

### 2.1.1 Método Variacional

Outro método de obter uma melhor estimativa de  $T_t$  é minimizando uma função custo. A função de custo é definida como a soma dos quadrados das distâncias de  $T$  para as duas observações, ponderadas por sua precisão do erro de observação:

$$J(T) = \frac{1}{2} \left[ \frac{(T - T_o)^2}{\sigma_o^2} + \frac{(T - T_b)^2}{\sigma_b^2} \right]$$

Desejamos obter  $T$  de modo que minimize a função custo  $J(T)$ . O mínimo da função custo  $J$  é obtido exigindo que  $\frac{\partial J}{\partial T} = 0$ :

$$\frac{\partial J}{\partial T} = \frac{1}{2} \left[ \frac{2(T - T_o)(1)}{\sigma_o^2} + \frac{2(T - T_b)(1)}{\sigma_b^2} \right] = \frac{(T - T_o)}{\sigma_o^2} + \frac{(T - T_b)}{\sigma_b^2}$$

Logo,

$$\frac{\partial J}{\partial T} = \frac{(\sigma_o^2 + \sigma_b^2)T - \sigma_b^2 T_o - \sigma_o^2 T_b}{\sigma_o^2 \sigma_b^2} = 0$$

Portanto,

$$T = \left( \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2} \right) T_o + \left( \frac{\sigma_o^2}{\sigma_o^2 + \sigma_b^2} \right) T_b$$

Temos provado a seguinte proposição:

**Proposição 2.** Nas condições da Proposição 1, o valor ótimo da temperatura que minimiza a função custo

$$J(T) = \frac{1}{2} \left[ \frac{(T - T_o)^2}{\sigma_o^2} + \frac{(T - T_b)^2}{\sigma_b^2} \right]$$

é dada por

$$T = \left( \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2} \right) T_o + \left( \frac{\sigma_o^2}{\sigma_o^2 + \sigma_b^2} \right) T_b.$$

Observa-se que ambos os métodos produzem o mesmo valor ótimo:  $T = T_a$ , isto é, os métodos são equivalentes. Entretanto os processos dos métodos são diferentes, enquanto a variável de controle para a minimização de  $J$  (ou seja, a variável em relação à qual estamos minimizando a função custo) é a temperatura, no método da interpolação estatística por mínimos quadrados as variáveis de controle foram os pesos. Esta equivalência também é válida para o caso multidimensional.

Exemplo: Se  $T_b = 2$ ,  $\sigma_b = 2$  e  $T_o = 0$ ,  $\sigma_o = 1$ , obtemos os pesos

$$\alpha = \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2} = \frac{2^2}{1^2 + 2^2} = \frac{4}{5}$$
$$\beta = \frac{\sigma_o^2}{\sigma_o^2 + \sigma_b^2} = \frac{1^2}{1^2 + 2^2} = \frac{1}{5}$$

Logo, a temperatura ótima é dada por

$$T_a = \alpha T_o + \beta T_b = \frac{4}{5} * 0 + \frac{1}{5} * 2 = 0,4$$

e a variância mínima sendo

$$\sigma_a^2 = \frac{\sigma_o^2 \sigma_b^2}{\sigma_o^2 + \sigma_b^2} = \frac{1 \times 4}{1 + 4} = 0,8.$$

## 2.1.2 Resumo das Equações

Em assimilação de dados a medida  $T_o$  é usualmente denominada de observação, sendo obtida por diversos instrumentos de medidas como estações pluviométricas, satélites, radares, etc. A medida  $T_b$  é denominada de chute inicial ou background, obtida usualmente por integrações de um modelo matemático. A medida  $T_a$  é denominada análise, a qual é obtida combinando as observações com o background para obter uma melhor estimativa do valor verdadeiro  $T_t$ . Como as observações são medidas reais obtidas por instrumentos, elas têm maior precisão que as medidas obtidas pelo background, assim a variância do erro das observações  $\sigma_o^2$  é menor que a variância do erro do background  $\sigma_b^2$ . De outro lado, como a análise é uma melhor estimativa que as observações e o background, resulta que a variância do erro da análise  $\sigma_a^2$  é menor que  $\sigma_o^2$  e  $\sigma_b^2$ .

Renomeando o peso  $\alpha$  como  $W$ , descrevemos um resumo das equações obtidas no problema das duas temperaturas:

i.  $T_a = T_b + W(T_o - T_b)$

a análise é obtida pelo background acrescentada da diferença entre a observação e o background (denominada inovação) ponderada pelo peso  $W$ .

ii.  $W = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2} = \frac{1}{1 + \left(\frac{\sigma_o^2}{\sigma_b^2}\right)}$

o peso ideal é a variância do erro do background pelo inverso da variância total (a soma da variância do erro do background e da variância do erro de observação). O peso ideal depende da razão entre a variância do erro da observação e a variância do erro do background, quanto maior a variância do erro do background, maior a correção para a estimativa inicial. Se a variância do erro da observação é nula, então o peso é igual a um, e a análise coincide com a observação.

iii. A variação da análise é

$$\sigma_a^2 = \frac{\sigma_b^2 \sigma_0^2}{\sigma_b^2 + \sigma_0^2}$$

Isso também pode ser escrito

$$\frac{1}{\sigma_a^2} = \frac{1}{\sigma_b^2} + \frac{1}{\sigma_0^2}$$

A precisão da análise (inverso da variância do erro de análise) é a soma da precisão do background e da observação.

iv.  $\sigma_a^2 = (1 - W)\sigma_b^2$

A variância do erro da análise é a variância do background, reduzido por um fator igual a 1 menos o peso ideal, o que também pode ser escrito como:

$$\sigma_a^2 = W\sigma_0^2$$

Este conjunto de equações se aplica a problemas multidimensionais mais complexos, onde são usados assimilação de dados sequenciais, como interpolação ótima e Filtro de Kalman. Nestes problemas, em que  $T_b$  e  $T_a$  são campos tridimensionais de tamanho muito maior que o conjunto de observações  $T_o$ , substituímos os escalares por matrizes. Em particular substitui-se a variância do erro pela covariância do erro e o peso ideal pela matriz de ganho ótima.

## Capítulo 3

# Assimilação de Dados Sequencial

No caso de realizar uma previsão, podem-se assimilar dados cada certo intervalo de tempo. Esta assimilação de dados no tempo recebe o nome de assimilação de dados sequencial. Este ciclo de análise tem duas fases, uma fase de previsão para atualizar o background  $T_b$  e sua variância do erro  $\sigma_b^2$ , e uma fase de análise, para atualizar a análise  $T_a$  e sua variância de erro  $\sigma_a^2$ .

### Fase de Previsão

Na fase de previsão do ciclo de análise, o background no tempo  $t_{i+1}$  é obtido por:

$$T_b(t_{i+1}) = M[T_a(t_i)]$$

onde, o modelo matemático de previsão  $M$  é alimentado pela análise no tempo  $t_i$ .

Na Interpolação Ótima, é feita a suposição que a integração do modelo aumenta a variância do erro inicial em uma razão fixa um pouco maior que 1:

$$\sigma_b^2(t_{i+1}) = a\sigma_a^2(t_i)$$

Isso permite que o novo peso  $W(t_{i+1})$  possa ser estimado como,

$$W(t_{i+1}) = \frac{\sigma_b^2(t_{i+1})}{\sigma_o^2(t_{i+1}) + \sigma_b^2(t_{i+1})}$$

**Fase de Análise** Na fase do ciclo de análise são coletadas as observações  $T_0(t_{i+1})$  no tempo  $t_{i+1}$ . A análise  $T_a(t_{i+1})$  é calculada por

$$T_a = T_b + W(T_0 - T_b)$$

A variância do erro da nova análise é determinada pela relação:

$$\sigma_a^2(t_{i+1}) = (1 - W)\sigma_b^2(t_{i+1})$$

Após a análise, se o ciclo de tempo  $t_{i+1}$  não estiver concluído, fazemos um novo ciclo de assimilação.

### 3.1 Assimilação de Dados para as Equações de Fluxo não Periódico de Lorenz

Para simular o ciclo de assimilação de dados sequencial, usamos o modelo de Lorenz para um "Fluxo Determinístico Não-periódico" (Lorenz, 1993):

$$\begin{aligned}\frac{dx}{dt} &= a(y - x) \\ \frac{dy}{dt} &= rx - y - xz \\ \frac{dz}{dt} &= xy - bz\end{aligned}$$

o qual é um modelo simplificado do comportamento da atmosfera que simula o comportamento de um fluido em um plano retangular, cuja temperatura do lado inferior é maior que a do superior, gerando correntes de convecção. Onde, "x" representa o fluxo convectivo; "y" a distribuição de temperaturas horizontal e "z" a distribuição de temperaturas vertical. Os três parâmetros que intervêm nas equações são: "a" a relação entre a viscosidade e a condutividade térmica, ou número de Prandtl; "r" proporcional à diferença de temperaturas entre os lados inferior e superior, ou número de Rayleigh reduzido e "b" relação entre a altura e a largura do retângulo.

Este sistema de equações é não linear, o que origina o comportamento caótico, pequenas perturbações na condição inicial se propagam e amplificam com sucessivas integrações no tempo. A não linearidade propaga os erros rapidamente produzindo futuras previsões com pouca confiabilidade, a assimilação de dados permite que as integrações fiquem sempre próximas da solução real.

### 3.1.1 Integração do Modelo

O modelo de Lorenz pode ser escrito em forma matricial:

$$\frac{d}{dt} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -a & a & 0 \\ r & -1 & 0 \\ 0 & 0 & -b \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 0 \\ -xz \\ xy \end{bmatrix}$$

denotando

$$Y = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad L = \begin{bmatrix} -a & a & 0 \\ r & -1 & 0 \\ 0 & 0 & -b \end{bmatrix}, \quad V = \begin{bmatrix} 0 \\ -xz \\ xy \end{bmatrix},$$

e  $f(t,y) = LY+V$ .

onde,  $Y$  é o vetor que contém as variáveis desconhecidas,  $L$  é a matriz que contém a parte linear do sistema e  $V$  é o vetor que contém a parte não linear do sistema. O sistema não linear de Lorenz é escrito em notação matricial como:

$$\frac{dY}{dt} = f(t, Y)$$

Para integrar o modelo no tempo foi utilizado o método de Runge - Kutta de quarta ordem, a qual tem ordem de convergência de ordem 4. Este esquema de integração é descrito pelo seguinte procedimento:

$$Y_{n+1} = Y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + 2k_4)$$

$$k_1 = hf(t_n, Y_n)$$

$$k_2 = hf\left(t_n + \frac{h}{2}, Y_n + \frac{k_1}{2}\right)$$

$$k_3 = hf\left(t_n + \frac{h}{2}, Y_n + \frac{k_2}{2}\right)$$

$$k_4 = hf(t_n + h, Y_n + k_3)$$

onde,  $h$  é o passo do tempo e  $Y_n = Y(t_n)$  e  $t_{n+1} = t_n+h$ , com  $h = 0,01$ .

O modelo de Lorenz foi codificado no software Matlab, que descreve o esquema de Runge-Kutta para o modelo da seguinte forma:

---

$x = \text{funcarg};$

```

p = parameters;
f = feval(func, x, p);

c1 = stepsize .* f;
x = funcarg + c1 /2;
f = feval(func, x, p);

c2 = stepsize .* f;
x = funcarg + c2 /2;
f = feval(func, x, p);

c3 = stepsize .* f;
x = funcarg + c3;
f = feval(func, x, p);

c4 = stepsize .* f;
funcarg = funcarg + (c1 + 2.*c2 + 2.*c3 + c4)./6;
test = funcarg;

```

---

O método de Runge-Kutta de ordem 4 resolve o sistema por recorrência, dado um estado inicial:

$$Y_0 = Y(0) = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} 8 \\ 0 \\ 30 \end{pmatrix}$$

pode-se avaliar  $Y_1$ , conhecido  $Y_1$  pode-se calcular  $Y_2$ . Em geral, conhecido  $Y_n$  pode-se calcular  $Y_{n+1}$ .

Dado que o problema é não linear, a integração leva um tempo para se estabilizar, para resolver esta instabilidade inicial integramos a partir da condição inicial  $Y_0$  por 600 iterações, o valor final substitui  $Y_0$ . Isto permite obter um background estável que será uma componente do processo de assimilação.

Consideramos os seguintes parâmetros para a integração do modelo: a=10, r=28, e b=8, os quais são parâmetros característicos do modelo de Lorenz.



### 3.1.2 Perturbação do Modelo

O modelo de Lorenz é perturbado na condição inicial e devido à não linearidade a trajetória se afastará das trajetórias do modelo não perturbado. Consideramos as trajetórias do modelo não perturbado como o estado real, e as trajetórias do modelo perturbado como o background. As observações serão combinadas com o background para obter a análise pela assimilação de dados.

Usamos a seguinte perturbação da condição inicial:

$$dY_o = \begin{pmatrix} 0, 1 \\ 0, 1 \\ 0, 1 \end{pmatrix}$$

logo, o valor inicial perturbado é  $Y_{p,0} = Y_0 + dY_0$ . O modelo perturbado é integrado por 1000 iterações, com assimilação das observações cada 8 iterações.

A codificação do modelo de Lorenz descreve a perturbação do modelo da seguinte forma:

---

```
x = x0;  
xp = x + dx0;  
for i = 1:numstep  
  for j = 1:bst  
    x = stepit('L63eqs',x,par,h2);  
    xp = stepit('L63eqs',xp,par,h2);  
  end  
  state(i,:) = x;  
  statep(i,:) = xp;  
end
```

---

### 3.1.3 Assimilação de Dados

Denominamos de Interpolação Ótima o método sequencial de interpolação estatística pelo método dos mínimos quadrados descrito no capítulo 2. Como agora estamos num problema vetorial, as variâncias dos erros são substituídas por covariâncias dos erros.

Descrevemos a seguir cada termo do esquema de interpolação ótima:

$$Y_a = Y_b + W(Y_{obs} - HY_b)$$

onde,

**i.**  $Y_a$  é o vetor da análise

**ii.**  $Y_{obs}$  é o vetor das observações

**iii.**  $H$  é a matriz de interpolação dos pontos espaciais do background para os pontos onde existem observações. No caso geral o vetor background é de dimensão muito maior que o vetor das observações, assim  $HY_b$  é da mesma dimensão que  $Y_{obs}$ . Neste modelo de Lorenz a dimensão espacial para cada variável é um, logo  $H$  é a matriz identidade de ordem três.

**iv.** A variância das observações é substituída pela covariância das observações  $R$ . Assumimos que  $R$  é da forma:

$$R = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

**v.** A variância do background é substituída pela covariância do background  $B$ :

$$B = E((Y_b - Y_t) \cdot (Y_b - Y_t)^t)$$

Consideramos na integração a seguinte matriz  $B$ :

$$B = \begin{bmatrix} 0,5124 & 0,7191 & -0,0067 \\ 0,7191 & 1,3460 & 0,0335 \\ -0,0067 & 0,0335 & 1,1116 \end{bmatrix}$$

a qual é obtida dos resultados de várias integrações.

**vi.** O peso ótimo é agora substituído pela matriz de pesos:

$$W = BH^t(R + HBH^t)^{-1}$$

isto é, a matriz de pesos é o produto da covariância dos erros do background pela inversa da soma das covariâncias dos erros do background e das observações.

O código do modelo de Lorenz que realiza a assimilação de dados é descrito a seguir:

---

```
x = x0 + dx0;  
for i = 1:numstep  
  for j = 1:bst  
    x = stepit('L63eqs',x,par,h2);  
  end  
  time(i) = i*bst*h2;  
  stateb(i,:) = x;  
  
  x = RUNIE(stateo(i,:),x,H,R,B);  
  statea(i,:) = x;  
end
```

---

# Capítulo 4

## Resultados

O software Matlab o qual codifica a solução numérica do modelo de Lorenz, facilita a manipulação de quantidades vetoriais e matriciais, além de fornecer gráficos de alta qualidade. Na figura 4.1 se apresenta a solução exata em 3D, para as tres variáveis (X,Y,Z). Esta figura lembra as asas de uma borboleta.

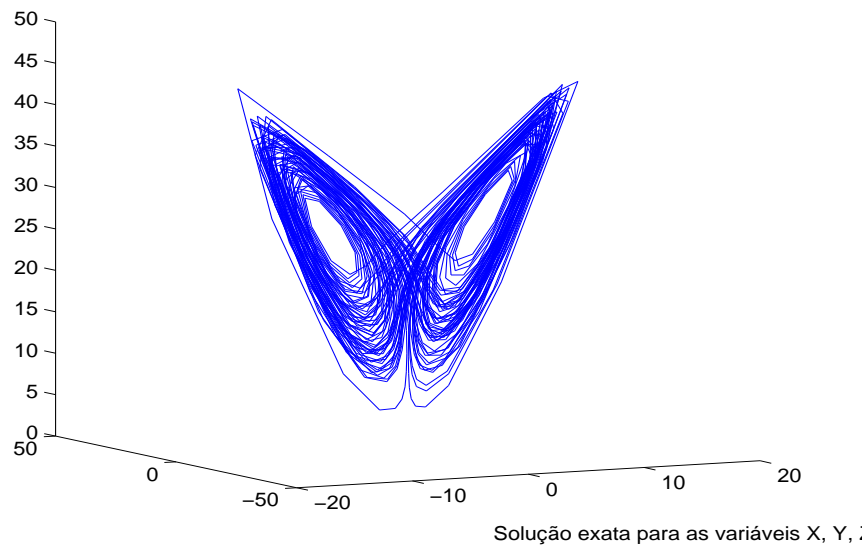


Figura 4.1: Solução exata do modelo de Lorenz

Na figura 4.2 se apresenta a solução exata do modelo de Lorenz para as variáveis X e Y.

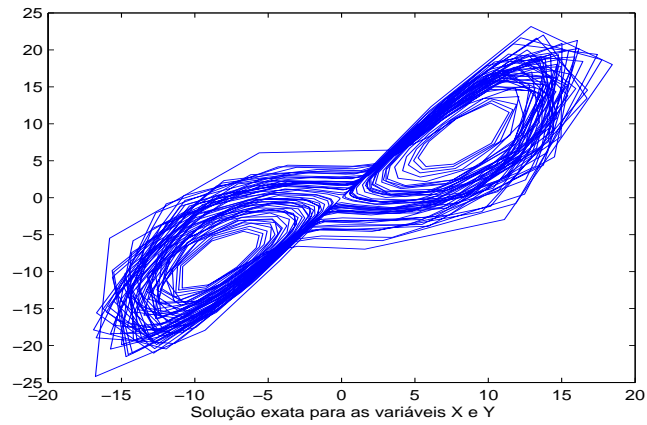


Figura 4.2: Solução exata para X e Y

No modelo de Lorenz consideramos a solução exata como a integração de controle ou solução não perturbada. Na Figura 4.3 se apresenta a evolução da solução exata para a variável x.

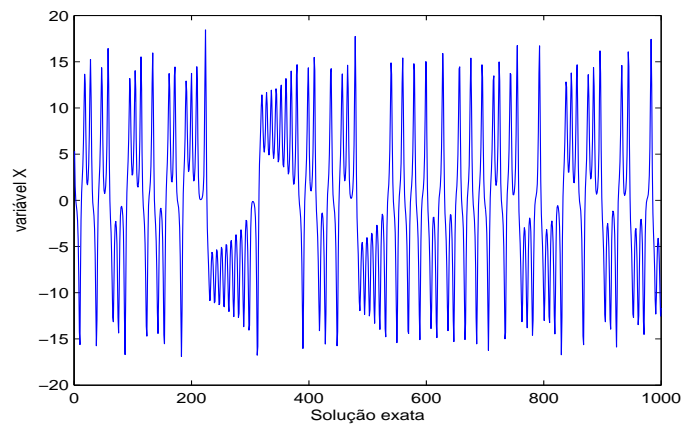


Figura 4.3: Solução exata para X

Na figura 4.4 se apresenta a evolução da solução exata no tempo para a variável  $x$ .

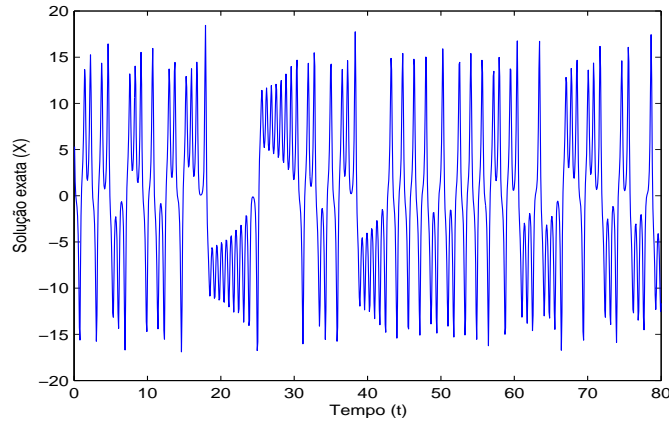


Figura 4.4: Solução exata no tempo(t)para X

Na Figura 4.5 se apresenta a evolução da solução exata (linha azul) e a solução perturbada (tracejado vermelho) para a variável  $x$ . Esta Figura mostra os efeitos da não linearidade do modelo, as condições iniciais de ambas as integrações diferem em 0,1 e são indistinguíveis no gráfico. Com o transcorrer do tempo as trajetórias se afastam até ter trajetórias diferentes.

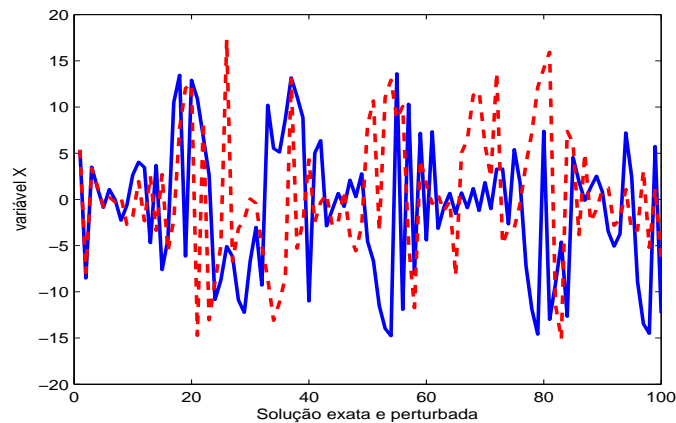


Figura 4.5: Solução exata e perturbada para X

NA figura 4.6 se apresentam as trajetórias da solução exata (linha azul), solução perturbada (linha vermelha) e solução assimilada (traçado verde) para variável  $x$ . Observa-se como a assimilação de dados por interpolação ótima combina o background co as observações para produzir a análise, a solução assimilada fica bem próxima da solução

exata. A solução perturbada devido a não linearidade se afasta da solução exata, enquanto o processo de assimilação, assimilando periodicamente as observações (a cada 8 iterações) conseguem manter a solução ótima próxima da solução exata.

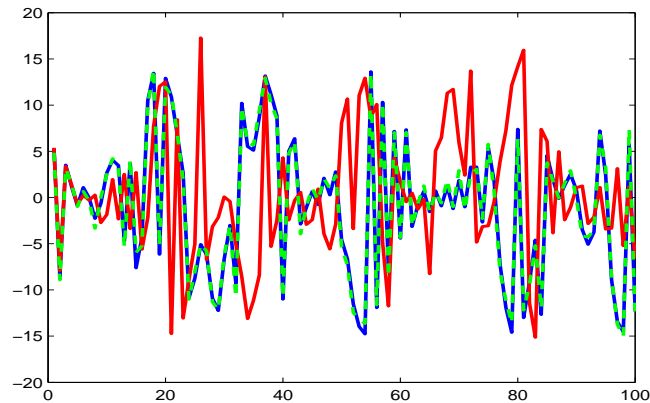


Figura 4.6: Solução exata, perturbada e assimilada para X

Na Figura 4.7 se apresentam o erro da análise e o erro do modelo perturbado para a componente x. Observa-se que o erro da análise (linha azul) é muito menor que o erro do modelo perturbado (linha vermelha), o método de assimilação consegue aproximar corretamente a solução ótima. O erro da análise é até 29 vezes menor que o erro do modelo perturbado.

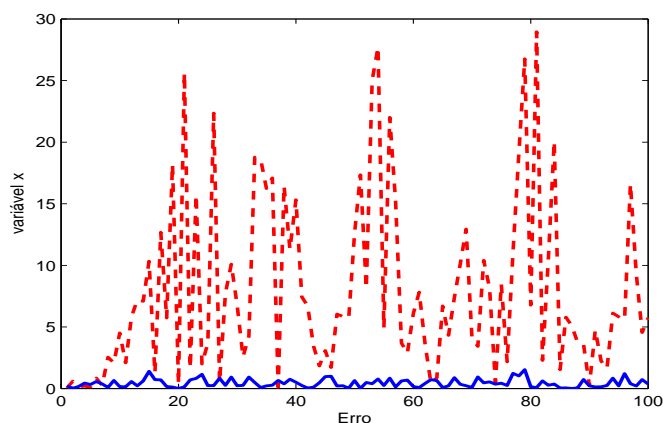


Figura 4.7: Erro do modelo perturbado, e erro da análise para X

Na figura 4.8 se apresenta a evolução da solução exata para a variável Y. Observa-se que o gráfico para a variável Y também nos mostra a não-linearidade do modelo.

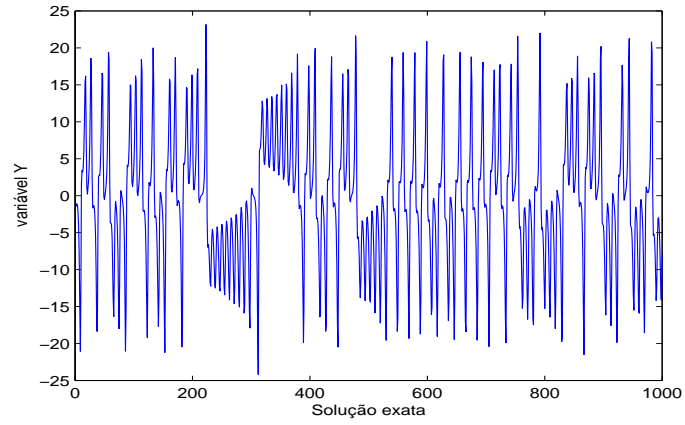


Figura 4.8: Solução exata para Y

Na figura 4.9 se apresenta a evolução da solução exata para variável Y no tempo.

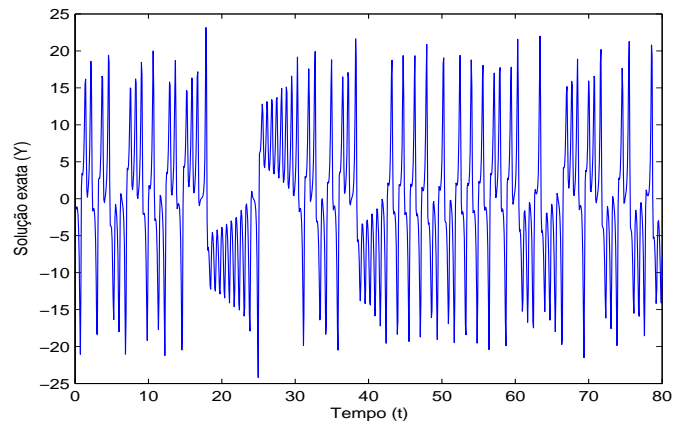


Figura 4.9: Solução exata no tempo para Y



Na figura 4.10 se apresenta a solução exata (linha azul) e solução perturbada (traçado vermelho) para variável  $Y$ . Vemos que para a variável  $Y$  as duas soluções com o decorrer do tempo também tomam trajetórias diferentes.

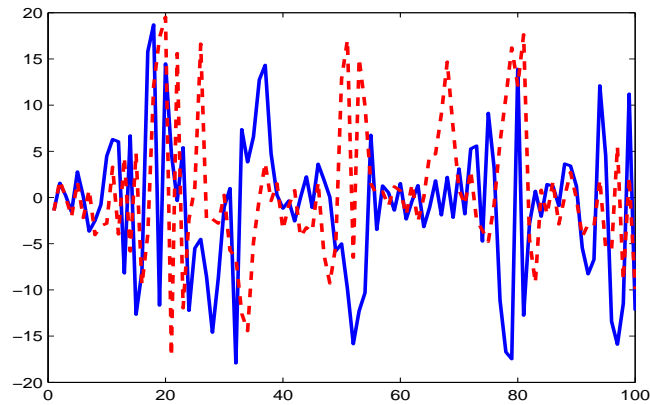


Figura 4.10: Solução exata e perturbada para  $Y$

Na figura 4.11 se apresenta a evolução da solução exata (linha azul), solução perturbada (traçado vermelho) e solução assimilada (traçado verde) para variável  $Y$ . Observa-se que para a variável  $Y$  o processo de assimilação é eficaz, a solução assimilada fica bem próxima da solução exata, enquanto que a solução perturbada se afasta totalmente da solução exata.

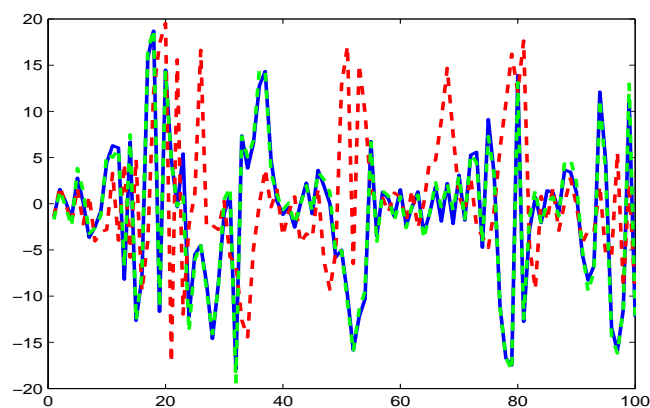


Figura 4.11: Solução exata, perturbada e assimilada para  $Y$

Na figura 4.12 se apresentam o erro da análise (linha azul) e o erro do modelo perturbado (traçado vermelho) para a componente Y. Note que para a variável Y o erro da solução assimilada é muito menor do que o erro da solução perturbada.

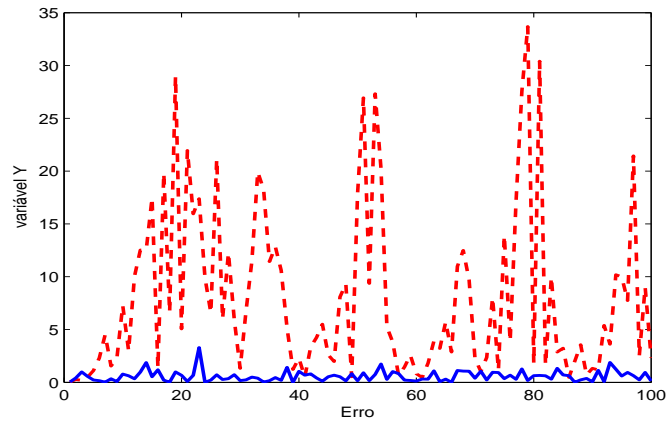


Figura 4.12: Erro da análise e erro do modelo perturbado para Y

Na figura 4.13 se apresenta a evolução da solução exata para variável Z. Nesta figura observamos também a não-linearidade do modelo.

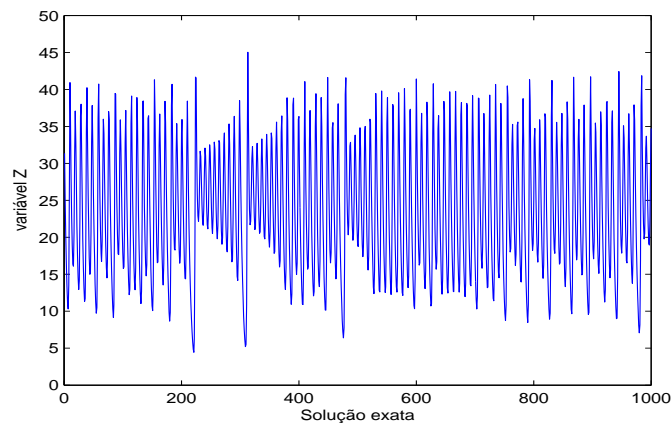


Figura 4.13: Solução exata para Z

Na figura 4.14 se apresenta a evolução da solução exata no tempo para variável  $Z$ .

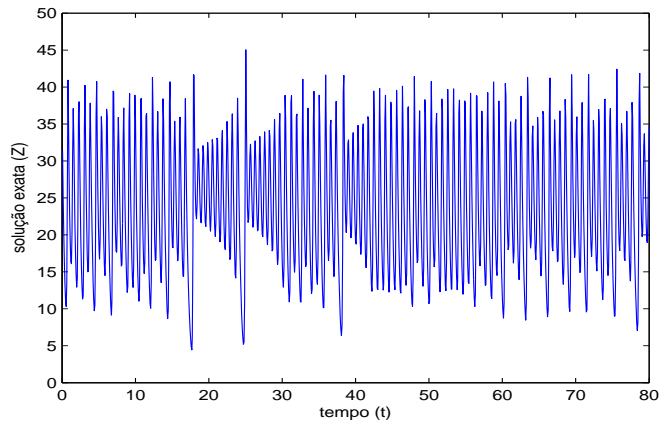


Figura 4.14: Solução exata no tempo para  $Z$

Na figura 4.15 se apresenta a solução exata (linha azul) e a solução perturbada (traçado vermelho) para variável  $Z$ . Podemos observar que para a variável  $Z$  as duas soluções com o decorrer do tempo tomam trajetórias completamente diferentes.

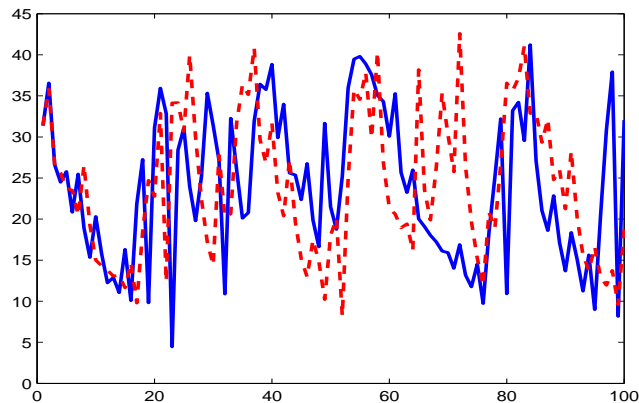


Figura 4.15: Solução exata e perturbada para  $Z$

Na figura 4.16 se apresenta a solução exata (linha azul), solução perturbada (traçado vermelho) e solução assimilada (traçado verde) para variável  $Z$ . Vemos que para a variável  $Z$  o processo de assimilação também é bastante eficaz, a solução assimilada fica bem próxima da solução exata, enquanto que a solução perturbada se distancia da solução exata.

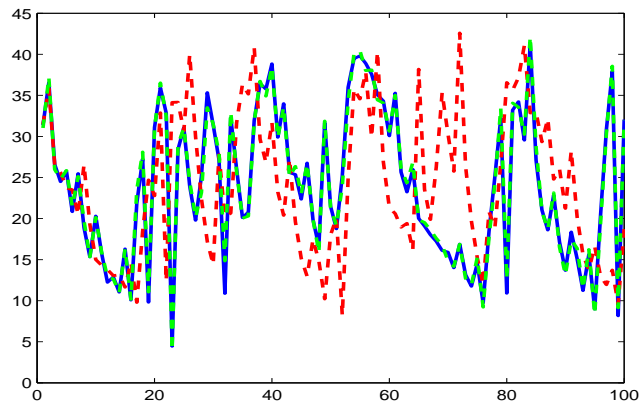


Figura 4.16: Solução exata, perturbada e assimilada para Z

Na figura 4.17 se apresenta o erro da análise e o erro da solução perturbada para componente Z. Esta figura mostra a grande diferença do erro da solução assimilada para o erro da solução perturbada, o erro da solução perturbada chega a ser 29 vezes maior que o erro da solução assimilada.

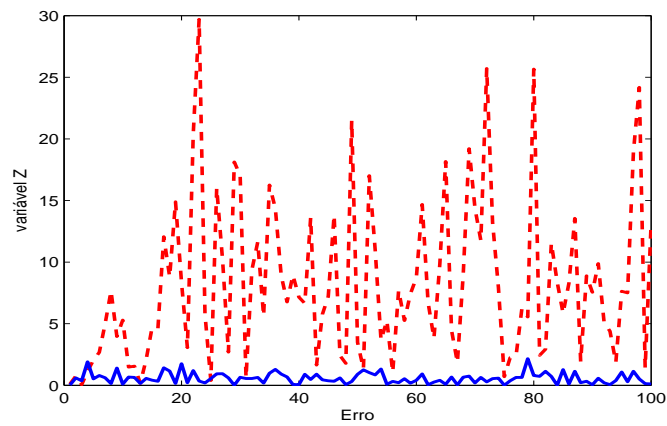


Figura 4.17: Erro da análise e erro da solução perturbada para Z

# Considerações Finais

Os resultados teóricos mostram que para o problema das duas temperaturas o método estatístico é equivalente ao método variacional. No primeiro método se minimiza o peso por mínimos quadrados, enquanto o método variacional minimiza uma função custo. Mostramos também que o problema simples das duas temperaturas serve como paradigma para o problema da análise objetiva do estado atmosférico, e identificamos e explicamos as principais equações relacionadas ao processo de assimilação.. Aplicamos os métodos de assimilação de dados em um modelo não-linear, que foi o sistema de equações de fluxo não periódico de Lorenz, esse sistema é não linear e por esse fato produz previsões do tempo pouco confiáveis, para obter previsões é feito sucessivas integrações do modelo para estabilizá-lo, vimos então como é feito o processo de integração do modelo no qual é usado o Método de Runge-Kutta para fazer tal processo, inserimos uma pequena perturbação na condição inicial do modelo e então integramos novamente o modelo para mostrar que pelo fato do modelo ser caótico uma pequena variação no seu ponto inicial muda toda sua trajetória no decorrer das integrações, conseguimos identificar a diferença que há entre a evolução da solução exata e a evolução da solução perturbada através de um gráfico feito no Matlab, feita a perturbação no modelo aplicamos a assimilação de dados, onde mostramos como é feito a assimilação para as equações de Lorenz, e visualizamos através de outro gráfico a eficácia da assimilação de dados, como realmente consegue-se aproximar muito bem a solução com assimilação à solução exata, vimos que esta solução contém erros, mas que são mínimos em comparação ao erro da solução com perturbação. Podemos concluir então que os métodos de aproximação são muito eficazes, e a assimilação de dados é muito importante para determinar a previsão do estado atmosférico.

# Referências Bibliográficas

- [1] CÁRDENAS, J.W.; TANAJURA, C.A.S.; BELYAEV, K.P.; MULLER, D.P.; CUBASH, U. On the oceanic impact of a data assimilation method on the numerical simulations over the tropics. XII CBMET. In CD-Room, 2002. 172; 592-598, 2001.
- [2] CÁRDENAS J.W. and THOMPSON, M. Error estimates and existence of solutions for the Lorenz atmospheric model. *Nonlinear Analysis*, v. 54, 123-142; 2003.
- [3] DALEY, R. Atmospheric data analysis. Cambridge University Press, 2000.
- [4] KALNAY, E. Atmospheric Modeling, Data Assimilation and Predictability. Cambridge University Press, Cambridge, 2003
- [5] LORENZ, E. N. On the existence of a slow manifold. *Journal of Atmospheric Science*, 43:1547-1557. 1986.
- [6] PACHECO, P.S.; CÁRDENAS, J.W. and TANAJURA, C.A.S. Temperature and salinity data assimilation experiments in the tropical atlantical with MOM3. Proceedings of the 8<sup>a</sup> International Conference on Southern Hemisphere Meteorology and Oceanography, v. 1; 589-594; 2006.
- [7] WILKS, D.S. Statistical methods in the atmospheric sciences. An introduction. Academic Press, San Diego, 2008