

# Assimilação de dados com redes neurais artificiais aplicado ao modelo de água rasa em duas dimensões

Helaine C. M. Furtado<sup>1</sup>, Haroldo F. de Campos Velho<sup>2</sup>, Elbert A. N. Macau<sup>2</sup>,  
Rosângela Cintra<sup>2</sup>

<sup>1</sup>Programa Doutorado em Computação Aplicada – CAP  
Instituto Nacional de Pesquisas Espaciais – INPE

<sup>2</sup>Laboratório Associado de Computação e Matemática Aplicada – LAC  
Instituto Nacional de Pesquisas Espaciais – INPE

{helaine.furtado,haroldo,elbert,rosangela.cintra}@lac.inpe.br

**Abstract.** *Data assimilation techniques are tools to combine data from physico-mathematics model with observational data to obtain a better forecast. Two data assimilation methods are compared here: the Kalman Filter method, and artificial neural network. Artificial neural networks appear as a novel method in the context for data assimilation. The performance of the methods is evaluated under application to shallow water equation 2D. The neural network was able to emulate the Kalman filter for the assimilation, with reduction of the processing time.*

**Resumo.** *Técnicas de assimilação de dados são ferramentas para combinar dados de um modelo físico-matemático com dados de observação para se obter uma melhor previsão do fenômeno em estudo. Dois métodos são comparados neste trabalho: o Filtro de Kalman e a técnica de Redes Neurais Artificiais (RNA). A abordagem de Redes Neurais Artificiais aparecem como uma metodologia recente no contexto de assimilação de dados. A performance dos métodos é avaliada no modelo de água rasa em duas dimensões. A Rede Neural Artificial foi capaz de emular o Filtro de Kalman para realizar a assimilação, com redução do tempo de processamento.*

**Palavras-chave:** *assimilação de dados, Filtro de Kalman, Redes Neurais Artificiais, modelo de água rasa 2D*

## 1. Introdução

Em meteorologia e oceanografia, assimilação de dados é um meio de estimação do estado da atmosfera ou do oceano pela combinação de dados observacionais com uma estimativa *a priori* proveniente do modelo dinâmico [Griffith 1997]. Esta estimativa do estado da atmosfera ou do oceano é frequentemente denominada de análise. No caso da meteorologia, a principal aplicação da assimilação de dados é na Previsão Numérica do Tempo, em que é usada para obter uma boa estimativa do estado atual da atmosfera para iniciar uma previsão.

Um fator importante para a compreensão das ferramentas de assimilação de dados é o conceito de erros, isto é, a estimação do erro associado aos dados observados e ao erro de modelagem. As observações contêm erros oriundos de diversas fontes, por exemplo, erro de instrumentação, amostragem, representatividade. O modelo dinâmico, por sua vez,

é imperfeito, com erros provenientes da aproximação física (química ou biológica) que governam a evolução explícita das variáveis de estado, a aproximação física que parametriza a interação entre as variáveis e a discretização do modelo. Portanto, é necessário considerar as incertezas e limitações inerentes ao modelo físico matemático e com as incertezas associadas aos dados observados, que dar um impulso para o uso das ferramentas de assimilação de dados.

Para a oceanografia, modelar o oceano consiste em resolver as equações do modelo tão exatas quanto possível, e comparar os resultados com as observações. Contudo, o oceano é um fluido turbulento, e pequenas mudanças nas condições iniciais podem originar impactos significantes sobre a evolução subsequente do fluido. No entanto, mesmo que fosse possível resolver completamente a equação diferencial parcial do movimento do fluido, a previsão do estado do oceano será limitada pela exatidão dos dados das condições iniciais e condições de contorno. O modelo numérico, seja oceânico ou atmosférico, requerem discretização e truncamento dos graus de liberdade da equação de continuidade. A parametrização do movimento sobre a escala resolvida é uma fonte significativa de erro. Estes fatores impossibilitam simular o fluxo de fluido com precisão [Zaron 2009], [e Pierre F. J. Lermusiaux 2000].

A Oceanografia operacional é uma área de crescente relevância para uma melhor caracterização do oceano. Os oceanos exercem uma grande influência no ambiente global, tornando a sua investigação fundamental para compreender os fenômenos que neles ocorrem. O seu conhecimento exige, entre outros, estudos das correntes, temperaturas, salinidades e elevação da superfície do mar, constituindo objetivos da Oceanografia Operacional, com aplicações no âmbito científico, ambiental e econômico.

As técnicas de assimilação de dados tornaram-se uma ferramenta de extrema importância para modelos numéricos, por exemplo, na previsão numérica do tempo [Kalnay 2003, Daley 1993], em circulação oceânica [Bennett 2002], e mais recentemente tem sido utilizada para a previsão de clima espacial [Harter et al. 2008, Scherliess et al. 2004].

Matematicamente, assimilação de dados é um processo de dois passos:

(1) Passo de Previsão:

$$\mathbf{x}_t^p = \mathbf{F}(\mathbf{x}_{t-1}^a) \quad (1)$$

(2) Passo de Análise:

(2.1) teoria da estimação

$$\mathbf{x}_t^a = \mathbf{x}_t^p + \mathbf{K}(\mathbf{y}_t - \mathbf{H}[\mathbf{x}_t^p]) \quad (2)$$

(2.2) cálculo variacional

$$\mathbf{x}_t^a = \text{Min}J(\mathbf{x}^p) \quad (3)$$

$$\mathbf{x}_t^a = \mathbf{x}^p + \sum_{m=1}^M \beta_m r_m(x, t), \text{ método do representante} \quad (4)$$

(2.3) redes neurais artificiais

$$\mathbf{x}_t^a = \mathbf{F}_{RNA_w^*}(\mathbf{y}, \mathbf{x}^p) \quad (5)$$

em que  $\mathbf{x}_t^p$  é o vetor de variáveis de estado prevista do modelo,  $\mathbf{F}$  representa o modelo numérico,  $\mathbf{x}_t^a$  representa o dado de análise (c.i.),  $\mathbf{y}$  o vetor de observações,  $t$  o tempo e  $x$  a coordenada espacial. Na Equação 2 a matriz  $\mathbf{H}$  é o operador linear que representa o sistema de observação,  $\mathbf{K}$  é a matriz de ganho de Kalman. Na Equação 5,  $\mathbf{F}_{RNA^*}$  é uma função que mapeia os dados de entrada ( $\mathbf{y}, \mathbf{x}^p$ ) com os dados da saída desejada  $\mathbf{x}^a$ , obtido com a Equação 2 para a derterminação do  $w^*$  ótimo. Este trabalho concentra-se no cálculo da análise baseado no uso da teoria da estimação, por meio do FK, e baseado em inteligência artificial, por meio das RNAs. O artigo está disposto da seguinte maneira: uma descrição do modelo usado para aplicação, as metodologias abordadas, resultados e discussões.

## 2. Modelo de água rasa 2D

O modelo abordado neste trabalho é o linear em duas dimensões dado a seguir:

$$\frac{\partial u}{\partial t} - fv + g \frac{\partial q}{\partial x} + r_u u = F_u \quad (6a)$$

$$\frac{\partial v}{\partial t} + fu + g \frac{\partial q}{\partial y} + r_v v = F_v \quad (6b)$$

$$\frac{\partial q}{\partial t} + H \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) + r_q q = 0 \quad (6c)$$

sobre o domínio  $0 \leq x \leq X$  e  $0 \leq y \leq Y$ .  $f$  representa o parâmetro de coriolis,  $r_u, r_v, r_q$  são os coeficientes de amortecimento,  $u, v$  componentes de vento e  $q$  é a perturbação de superfície livre,  $H$  é a profundidade média do oceano.

## 3. Filtro de Kalman

A partir de um modelo de previsão e um sistema de observação, em que o subscrito  $n$  denota o tempo discreto, o sobrescrito  $p$  representa o valor previsto, tem-se:

$$\begin{aligned} \mathbf{x}_{n+1}^p &= \mathbf{F}_n \mathbf{x}_n^p + \mu_n \\ \mathbf{y}_n^p &= \mathbf{H}_n \mathbf{x}_n^p + \nu_n \end{aligned} \quad (7)$$

em que  $\mathbf{F}_n$  representa o modelo matemático,  $\mu_n$  é uma forçante estocástica (ruído aleatório do modelo). O sistema de observações é modelado por um operador  $\mathbf{H}_n$  e  $\nu_n$  é o ruído associado as observações. A hipótese para a função densidade de probabilidade Gaussiana com média zero é utilizada para os ruídos de modelagem e de observação,  $\mathbf{y}_n^p$  é o vetor de observações. O vetor de estado  $\mathbf{x}_{n+1}$  é estimado recursivamente no tempo por meio de:

$$\mathbf{x}_{n+1}^a = (\mathbf{I} - \mathbf{G}_{n+1} \mathbf{H}_{n+1}) \mathbf{F}_n \mathbf{x}_n^a + \mathbf{G}_{n+1} \mathbf{y}_{n+1}^p \quad (8)$$

sendo que  $\mathbf{x}_{n+1}^a$  é o valor de análise estimado,  $\mathbf{G}_n$  é a matriz de ganho de Kalman, que é calculada da a partir da minimização variância do erro  $J_{n+1}$  [Jazwinski 1970]:

$$J_{n+1} = E\{(\mathbf{x}_{n+1}^a - \mathbf{x}_{n+1}^p)^T (\mathbf{x}_{n+1}^a - \mathbf{x}_{n+1}^p)\} \quad (9)$$

em que  $E\{\cdot\}$  é o valor esperado. Para o cálculo da matriz de ganho de Kalman ( $\mathbf{G}_n$ ) considera-se três matrizes:  $\mathbf{Q}_n$  (a covariância de  $\mu_n$ ),  $\mathbf{R}_n$  (a covariância de  $\nu_n$ ) e  $\mathbf{P}_n^p$  é a matriz de covariância do erro, dada por:

$$\mathbf{P}_n^p = \mathbf{F}_n \mathbf{P}_n^a \mathbf{F}_n + \mathbf{Q}_n \quad (10)$$

sendo o ganho de Kalman calculado por:

$$\mathbf{G}_{n+1} = \mathbf{P}_{n+1}^p \mathbf{H}_n^T [\mathbf{R}_n + \mathbf{H}_n \mathbf{P}_n^p \mathbf{H}_n^T]^{-1} \quad (11)$$

A assimilação é feita por meio da amostragem:

$$\mathbf{r}_{n+1} \equiv \mathbf{y}_{n+1} - \mathbf{y}_{n+1}^p = \mathbf{y}_{n+1} - \mathbf{H}_n \mathbf{x}_n^p \quad (12)$$

a propagação do erro é expressa pela matriz

$$\mathbf{P}_{n+1}^a = [\mathbf{I} - \mathbf{G}_{n+1} \mathbf{H}_n] \mathbf{P}_{n+1}^p \quad (13)$$

O método de Filtro de Kalman como uma técnica de assimilação de dados pode ser resumida como:

1. calcule o passo de previsão:  $\mathbf{x}_{n+1}^p = \mathbf{F}_n \mathbf{x}_n^p$ ;
2. calcule o valor:  $\mathbf{y}_{n+1}^p = \mathbf{H}_n \mathbf{x}_n^p$ ;
3. calcule a matriz de ganho de Kalman:  $\mathbf{G}_{n+1}$  (Equação 11);
4. execute a assimilação (cálculo da análise):  $\mathbf{x}_{n+1}^a$  (Equação 8);
5. atualiza a matriz de covariância do erro  $\mathbf{P}_{n+1}^a$  (Equação 13).

#### 4. Redes neurais artificiais

Redes neurais artificiais (RNA) tem se tornado uma importante ferramenta em processamento de sinais. Muitas pesquisas têm sido realizadas na busca de novos modelos de RNA adaptando o modelo existente para resolver problemas da vida real, tais como os problemas de engenharia [Haykin 2004]. Uma rede neural artificial é um arranjo de unidades de processamento denominado de neurônios que são caracterizado por [Haykin 2004]:

- um conjunto de unidades de processamento (neurônios).
- um conjunto de pesos, que são as conexões entre as unidades de processamento, onde o conhecimento da rede é armazenado.
- altamente paralelo, controle distribuído.

O modelo de um neurônio artificial consiste basicamente de uma combinação linear seguida por uma função de ativação, dado por:

$$y_k = \varphi \left( \sum_{j=1}^m \theta_{kj} x_j + b_k \right) \quad (14)$$

em que  $x_1, \dots, x_m$  são sinais de entrada;  $\theta_{k1}, \dots, \theta_{km}$  são os pesos sinápticos do neurônio  $k$ ,  $b_k$  é o viés e  $y_k$  é o sinal de saída do neurônio,  $\varphi(\cdot)$  representa a função de ativação, que restringe a amplitude do sinal na saída de um neurônio. O papel da função de ativação é simular características não lineares de um neurônio biológico.

Existem várias arquiteturas de redes neurais artificiais (RNA), que geralmente dependem da estratégia de aprendizagem utilizada. O funcionamento das RNA é realizado por meio de duas fases: a fase de treinamento<sup>1</sup> e a fase de ativação<sup>2</sup>. O algoritmo usado para o

<sup>1</sup>Um conjunto de exemplos são apresentados à rede, a qual extrai as características necessárias para representar a informação fornecida.

<sup>2</sup>Utiliza as características armazenadas nos pesos após a fase de treinamento.

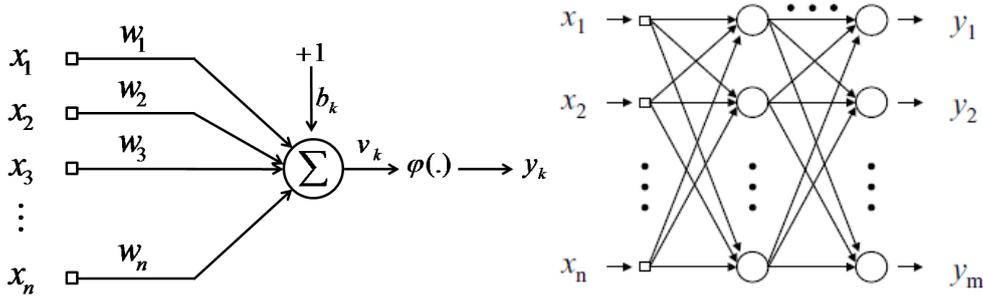


Figura 1. (lado esquerdo) neurônio artificial, (lado direito) Rede Neural Artificial Multicamadas.

treinamento do PMC - Perceptron de Múltiplas Camadas foi o retropropagação do erro, a qual é uma rede alimentada adiante composta por uma camada de entrada, uma camada de saída e camadas escondidas, cujo o objetivo é extrair propriedades estatísticas da entrada dos dados [Haykin 2004]. A Figura 1 ilustra o neurônio artificial e a rede PMC.

A função de ativação implementada na camada oculta do PMC foi a *função tangente hiperbólica*, que é uma função sigmóide que varia entre  $[-1, 1]$  dada pela equação 15.

$$\varphi(v_j) = \tanh\left(\frac{av_j}{2}\right) \quad (15)$$

O método de assimilação de dados baseado em RNA, por exemplo, o Perceptron Multicamadas pode ser escrito como:

$$\mathbf{x}_n^a = F_{RNA}(\mathbf{x}_n^p, \mathbf{y}_n^o, \mathbf{x}_n^{FK}) \quad (16)$$

em que  $\mathbf{x}_n^a$  é a análise,  $F_{RNA}$  representa o processo de assimilação por RNA,  $\mathbf{x}_n^p$  são os dados do modelo,  $\mathbf{y}_n^o$  os dados de observação e  $\mathbf{x}_n^{FK}$  a análise obtida com o Filtro de Kalman descrito na seção 3, que é usado na fase de treinamento da rede.

## 5. Resultados e discussões

Nesta seção são apresentados os resultados. Para testar as metodologias foram realizados dois experimentos. Para a integração do modelo apresentado na seção 2 foi usado o método de *Arakawa C-Grid* para a discretização espacial e para a discretização temporal o método *forward-backward*. No experimento 1 a variável  $q$  foi inicializada com uma função seno, as variáveis  $u$  e  $v$  tiveram valor inicial zero. Para o experimento 2 as variáveis  $u$ ,  $v$  e  $q$  foram inicializadas com uma função seno.

Para a implementação do Filtro de Kalman, foram usadas as seguintes matrizes:  $\mathbf{Q}_t = 0.1\mathbf{I}$ ;  $\mathbf{R}_t = 0,5\mathbf{I}$ ;  $\mathbf{H} = \mathbf{I}$ ; matriz de covariância do erro de modelagem, matriz de covariância do erro de observação e o operador que representa o sistema de observação, respectivamente. A matriz de covariância do erro de previsão é inicializada por:

$$\mathbf{P}_0^f = \begin{cases} 10(x_0^f)_i^2 & \text{para } i = j, \\ 0 & \text{para } i \neq j. \end{cases} \quad (17)$$

As observações assimiladas neste trabalho foram geradas artificialmente. Os dados foram gerados a partir da integração do sistema, adicionando-se um ruído aleatório de variância

0,05. A referência de verdade para o método de assimilação é a curva obtida a partir da integração dos modelos sem ruído, o que significa que, quanto mais próximo as estimativas estiverem da dinâmica do modelo, melhor será a estimativa obtida pelos métodos de assimilação.

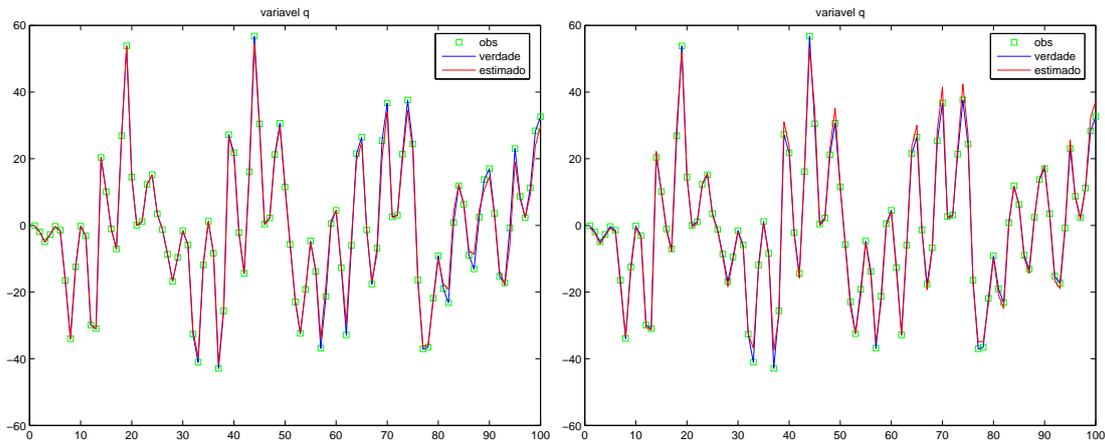
A rede Perceptron de Múltiplas Camadas, descrita na seção 4, é uma rede que exige o treinamento supervisionado, isto é, que seja apresentado a ela a saída desejada, ou o denominado “alvo” da rede. A saída desejada da rede é a estimativa obtida com o Filtro de Kalman, descrito da seção 3. A fase de treinamento da rede, consiste em determinar os melhores pesos que mapeia os dados de entrada aos dados da saída desejada (“alvo” da rede). A rede implementada possui uma camada escondida com 30 neurônios; uma camada de entrada com dois neurônios e um neurônio na camada de saída. A taxa de aprendizagem utilizada foi  $\eta = 0,7$ , em ambos os experimentos.

No processo de assimilação os dois experimentos se diferenciam pela a quantidade de observações utilizadas. No experimento 1 foram usados 100 observações. Enquanto que, no experimento 2 foram usados 400 observações. Somente a variável  $q$  foi assimilada. Na Figura 4 os quadrados vermelhos representam os pontos de observações. No experimento 1 foram usados 25 observações em cada passo de tempo e para o experimento 2 foram utilizados 100 observações em cada passo de tempo.

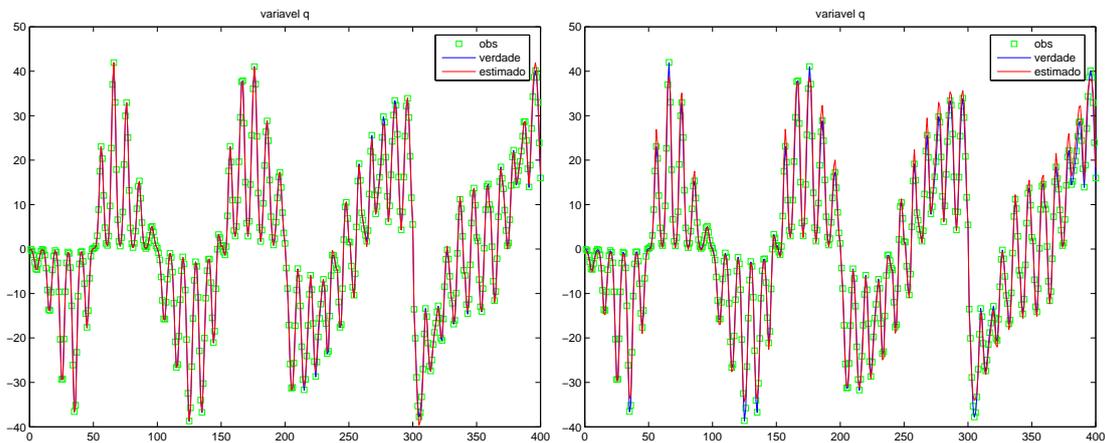
Na Figura 2 apresenta-se o resultado da assimilação realizada com o FK (lado esquerdo) e com a RNa (lado direito). Na Figura 3 tem-se o resultado para o experimento 2. Para uma avaliação quantitativa apresenta-se na Tabela 1 o resultado do erro de assimilação para os dois experimentos. Na Tabela 2 o resultado do custo computacional. Em relação ao erro de assimilação o Filtro de Kalman foi o que apresentou o menor erro nos dois experimentos. A curva azul representa a referência de verdade, a curva vermelha o estimado e os quadrados verdes representam as observações.

Um fato importante neste tipo de aplicação é o custo computacional. A inovação tecnológica permitindo o surgimento de supercomputadores, a evolução dos modelos numéricos de previsão e da rede de observação, vem proporcionando maior confiabilidade na previsão numérica operacional. O aumento na quantidade de satélites lançados disponibiliza uma grande quantidade de informações meteorológicas e oceanográficas. O aperfeiçoamento da resolução de grade dos modelos numéricos de previsão também é fato. Isso significa que, a ordem de grandeza de dados observacionais disponíveis é da ordem de  $10^5$  a  $10^7$ , enquanto que, o número de variáveis do modelo é da ordem de  $10^6$  a  $10^8$  [Kalnay 2003].

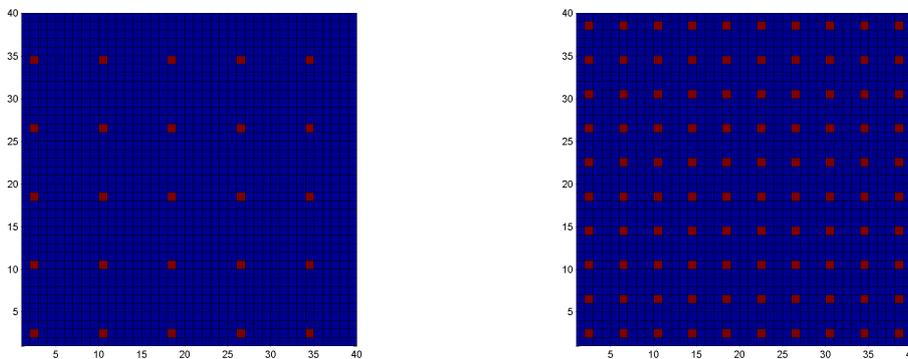
O desafio atual da assimilação de dados consiste em: *“conceber formas para utilizar os dados observados disponíveis em modelos com número de grades crescentes e produzir uma análise dentro do tempo disponível na prática operacional.”* Neste contexto, assimilação de dados por meio de Redes Neurais Artificiais pode ser uma alternativa viável. Na Tabela 2 mostra-se que essa ferramenta é aproximadamente 42 vezes mais rápida que o método do Filtro de Kalman, e com um resultado de assimilação eficiente. As técnicas de assimilação foram implementadas em matlab. Os programas foram executados em um processador Intel(R) core(TM) i5-2310 CPU@ 2.9 GHz com 8 GB de memória RAM.



**Figura 2. Experimento 1: (lado esquerdo) Estimado com o FK; (lado direito) estimado com a Rede Neural Artificial Multicamadas. (curva azul) verdade; (curva vermelha) estimado; (quadrados verdes) observações.**



**Figura 3. Experimento 2: (lado esquerdo) Estimado com o FK; (lado direito) estimado com a Rede Neural Artificial Multicamadas. (curva azul) verdade; (curva vermelha) estimado; (quadrados verdes) observações.**



**Figura 4. (lado esquerdo) experimento 1; (lado direito) experimento 2. Os quadrados vermelhos representam as observações no ponto de grade.**

**Tabela 1. Resultados do erro de assimilação para FK e o PMC.**

|                     | FK     | RNA-PMC |
|---------------------|--------|---------|
| <i>experimento1</i> | 0,6257 | 1,3290  |
| <i>experimento2</i> | 0,2021 | 1,1629  |

**Tabela 2. Custo computacional para FK e o PMC.**

|                     | FK         | RNA-PMC |
|---------------------|------------|---------|
| <i>experimento1</i> | 42 min     | 33 s    |
| <i>experimento2</i> | 1 h e 19 s | 45 s    |

## Referências

- Bennett, A. F. (2002). *Inverse Modeling of the Ocean and Atmosphere*. Cambridge University Press, New York, 1 edition.
- Daley, R. (1993). *Atmospheric data analysis*, volume 3. Cambridge University Press, New York.
- e Pierre F. J. Lermusiaux, A. R. R. (2000). Overview od data assimilation. Harvard Reports in Physical/Interdisciplinary.
- Griffith, A. K. (1997). *Data assimilation for numerical weather prediction using control theory*. PhD thesis, University of Reading, Reading, Inglaterra.
- Harter, F. P., Rempel, E. L., de Campos Velho, H. F., and Chian, A. (2008). Application of artificial neural networks in auroral data assimilation. *Journal of Atmospheric and Solar - Terrestrial Physics*, 70(10):1243–1250.
- Haykin, S. (2004). *Redes neurais princípios e prática*, volume 2. Editora Bookman, Porto Alegre.
- Jazwinski, A. H. (1970). *Stochastic processes and filtering theory*. Academic Press, New York and London.
- Kalnay, E. (2003). *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, New York, 2 edition.
- Scherliess, L., Schunk, R. W., Sojka, J. J., and Thompson, D. C. (2004). Development of a physics-based reduced state kalman filter for the ionosphere. *Radio Science*, 39(RS1S04). doi:10.1029/2002RS002797.
- Zaron, E. D. (2009). Introduction to ocean data assimilation.