Proceeding Series of the Brazilian Society of Applied and Computational Mathematics, Vol. 4, N. 1, 2016.
Trabalho apresentado no DINCON, Natal - RN, 2015.

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Artificial Neural Networks emulating Representer Method at a shallow water model 2D

Helaine C. M. Furtado[1]
Instituto de Engenharia e Geociências (IEG), UFOPA, Santarém, Pará.
Rosângela S. C. Cintra[2]
Haroldo F. de Campos Velho[3]
Laboratório de Computação e Matemática Aplicada (LAC), INPE, São José dos Campos, SP

Abstract. The goal of the present work is to employ artificial neural networks as a data assimilation method applied to shallow water equation. This model is used to represent ocean dynamics. Data assimilation is a computational procedure to combine observation data with model data for identifying the best initial condition (analysis) to an operational prediction system. Here we compare two techniques: representer method and artificial neural network.

Keywords. Data assimilation, differential equations, numerical prediction, artificial neural networks, representer method.

1 Introduction

Many problems in geosciences (meteorology, oceanography, and geophysics) are modeled by differential equations. These problems may require the estimation of time-dependent state variable using noisy measurements. However, the mathematical model is always an approximation of reality [11]. This means that, the modeling error is a permanent feature. For operational prediction systems, a strategy to deal with such uncertainty is to add some real world information into the mathematical model. It consist of observations extracted from the modeled phenomena (data observation). However, observed data should be carefully inserted in order to avoid prediction degrade. Data assimilation techniques combine two sources - data model and observation - to produce a data analysis. The analysis is the initial condition used in the computer model prediction.

The basic components for an operational forecasting system are: a network of observed data, a numerical model and a data assimilation method. Data assimilation is the process responsible for combining a mathematical model and observed data, producing an analysis (or initial condition) for a numerical prediction system. The historical evolution of

[1]helaine.furtado@ufopa.edu.br; helaine.furtado@lac.inpe.br
[2]rosangela.cintra@lac.inpe.br
[3]haroldo@lac.inpe.br

data assimilation methods comes from function adjustment, successive corrections, analysis correction, optimal Interpolation, variational methods, Kalman filter, Monte Carlo techniques and Artificial Neural Networks ([4], [8], [12], [15]). Each method has different ways to combine background values (previous forecasting) with observations. The model presented in this work to test the methodology is linear in two dimensions:

$$\frac{\partial u}{\partial t} - fv + g\frac{\partial q}{\partial x} + r_u u = F_u \tag{1a}$$

$$\frac{\partial v}{\partial t} + fu + g\frac{\partial q}{\partial y} + r_v v = F_v \tag{1b}$$

$$\frac{\partial q}{\partial t} + H\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + r_q q = 0 \tag{1c}$$

The spatial domain is $(x, y) \in \Omega$, where $\Omega \equiv (0, X) \times (0, Y)$. The Coriolis coefficient is denoted by f; g is gravitational acceleration; r_u, r_v, r_q damping coefficients; u is the velocity in the x-direction (or zonal velocity); v is the velocity in the y-direction (or meridional velocity); q is the sea-level diturbance: if $q \equiv q'$, then the ocean is in hydrostatic balance [2]; F_u and F_v are the model forcings; finally H denote ocean depth. The differential equations (1a, 1b, and 1c) are discretized on the Arakawa C-grid with a forward-backward scheme for time-stepping [13]. The boundary conditions we have periodic channel with rigid walls, namely, $u(x, 0, t) = v(x, Y, t) = 0$. The model forcings are: $F_u = -C_d \rho_a u_a^2/(H\rho_w)$ and $F_v = 0$.

2 Metodology

2.1 Variational Method: Representer Technique

The Variational Method is an data assimilation technique based on minimizing a functional. However, it's problem dependent and leads to different variational formulations for different problems. For some cases, it is not obvious to obtain the variational formulation of the problem, or even to find one. The variational formulation of the representer technique requires several steps, as shown below.

- The description of the penalty function in terms of residuals (constraints) function is given by:

$$\begin{aligned} J[u, v, q] \;=\; & \sum_{k=1}^{3} \left[W_f^k \int_0^T dt \int_\Omega [f^k(x, y, t)]^2 \, d\Omega + W_i^k \int_\Omega [i^k(x, y)]^2 \, d\Omega \right] \\ & + \sum_{l=1}^{2} \left[W_b^l \int_0^T dt \int_0^X [b_l(x, t)]^2 \, dx \right] + w \sum_{m=1}^{M} \varepsilon_m^2 \end{aligned} \tag{2}$$

  where $k = 1, 2, 3 = u, v, q$, and f^k represents the Eqs. (1a) (1b) (1c). The second and third terms are the initial and boundary condition constraints, and the last term is the square difference between the model and observations.

DOI: 10.5540/03.2016.004.01.0097

- The equation for the Lagrange multipliers λ^k are obtained deriving the Euler-Lagrange equations for extremum of the penalty (constraints):

$$-\frac{\partial \lambda^u}{\partial t} + f\lambda^v - H\frac{\partial \lambda^q}{\partial x} + r_u\lambda^u = 0 \tag{3a}$$

$$-\frac{\partial \lambda^v}{\partial t} - f\lambda^u - H\frac{\partial \lambda^q}{\partial y} + r_v\lambda^v = 0 \tag{3b}$$

$$-\frac{\partial \lambda^q}{\partial t} - g\left(\frac{\partial \lambda^u}{\partial x} + \frac{\partial \lambda^v}{\partial y} + r_q\lambda^q\right) = S \tag{3c}$$

$$S = -w\sum_{m=1}^{M}\left[\hat{q}(x_m, y_m, t_m) - d_m\right]\delta(x - x_m)\delta(y - y_m)\delta(t - t_m) \tag{3d}$$

with null final conditions ($\lambda^k = 0$ for $t = T$), and the same boundary conditions as the forward problem.

- Representer equations is obtained firstly solving the associated Green function α_m^k:

$$-\frac{\partial \alpha_m^u}{\partial t} + f\alpha_m^v - H\frac{\partial \alpha_m^q}{\partial x} + r_u\alpha_m^u = 0\ , \tag{4a}$$

$$-\frac{\partial \alpha_m^v}{\partial t} - f\alpha_m^u - H\frac{\partial \alpha_m^q}{\partial y} + r_v\alpha_m^v = 0\ , \tag{4b}$$

$$-\frac{\partial \alpha_m^q}{\partial t} - g\left(\frac{\partial \alpha_m^u}{\partial x} + \frac{\alpha_m^v}{\partial y}\right) + r_q\alpha_m^q = \delta(x - x_m)\delta(y - y_m)\delta(t - t_m)\ . \tag{4c}$$

The above system is solved with the same final and boundary conditions as adopted to the Lagrange multipliers λ^k.

- Representer solution r_m^k is identified from the Green funcion computed before:

$$\frac{\partial r_m^u}{\partial t} + fr_m^v + g\frac{\partial r_m^q}{\partial x} + r_u r_m^u = [W_f^u]^{-1}\alpha_m^u \tag{5a}$$

$$\frac{\partial r_m^v}{\partial t} + fr_m^u + g\frac{\partial r_m^q}{\partial y} + r_v r_m^v = [W_f^v]^{-1}\alpha_m^v \tag{5b}$$

$$\frac{\partial r_m^q}{\partial t} + H\left(\frac{r_m^u}{\partial x} + \frac{r_m^v}{\partial y}\right) + r_q r_m^q = [W_f^q]^{-1}\alpha_m^q \tag{5c}$$

with the following initial and boundary conditions:

$$r_m^k(X, Y, 0) = [W_i^k]^{-1}\alpha_m^k(x, y, 0) \quad \text{(IC)} \tag{6a}$$

$$r_m^k(x \pm X, y, t) = r_m^k(x, y, t) \tag{6b}$$

$$r_m^k(x, y \pm Y, t) = r_m^k(x, y, t) \quad (k = u, q) \tag{6c}$$

$$r_m^v(x, 0, t) = H[W_b^v]^{-1}\alpha_m^v(x, 0, t) \tag{6d}$$

$$r_m^v(x, Y, t) = H[W_b^v]^{-1}\alpha_m^v(x, Y, t)\ . \tag{6e}$$

DOI: 10.5540/03.2016.004.01.0097

- In the representer approach, the minimum of the functional is determined by: $\hat{\Phi}(x,y,t) = \Phi^F(x,y,t) + \sum_m \beta_m r_m(x,y,t)$. To complete the solution, the expansion coefficients β_m are computed by solving the linear system:

$$\left[\mathbf{R} + w^{-1}\mathbf{I}\right]\,\beta = \sum_{m=1}^{M}\left[\Phi_m^{obs} - \Phi_m^{mod}\right] \tag{7}$$

  where $\Phi = [u \;\; v \;\; q]^T$, $\beta = [\beta_1 \;\ldots\; \beta_M]^T$, $\mathbf{R}$ is the covariance matrix of measurement error, and $\mathbf{I}$ is the identity matrix.

2.2 Artificial Neural Networking (ANN)

Artificial Neural Networks (ANN) became important tools for information processing [9]. Much research has been conducted in pursuing new neural network models and adapting the existing ones to solve real world problems. ANN are made of arrangements of processing elements called neurons. The artificial neuron model basically consists of a linear combiner followed by an activation function. Connected processing units form the ANN. They are characterized by: very simple neuron-like processing elements; weighted connections between the processing elements; highly parallel processing and distributed control; automatic learning of internal representations.

A feedforward network is a non-linear mapping to compute the output vector from an input vector. The connections among the several neurons (1b) have associated weights that are adjusted during the learning process, thus changing the performance of the network. Two distinct phases can be devised while using ANN: the training phase (learning process) and the run phase (activation of the network). The training phase consists of adjusting the weights for the best performance of the network in establishing the mapping of many input/output vector pairs. Once trained, the weights are fixed and the network can be presented to new inputs for which it calculates the corresponding outputs, based on what it has learned.

3 Results

Data assimilation is an essential step for operational forecasting systems by means of a weighted combination between observational data and data from a mathematical model. Artificial neural networks (ANN) have been proposed as a technique for data assimilation ([15], [5], [10], [3]). The Figure 1a) presents two steps of an ANN. The first one, training process, where the ANN learns to emulate the Representer Method. After that, the analysis can be computed. The method is presented with applications on shallow water model 2D (Equation 1). The shallow water equations were integrated at 60 time steps.The training set for the ANN use the first 160 time steps. The remaining time steps were used for the generalization. The performance of the ANN is to be compared with Representer Method (kind Variational Method). The q variable was initialized with sin function, and variables was $u = v = 0$. Figure 3 presents a comparison between the Representer Method (red curve) and the Artificial Neural Network (green curve) and the temporal evolution of point to variables u and q. In this experiment, the ANN was able to follow the dynamics of shallow water flow, showing a good performance of the data

DOI: 10.5540/03.2016.004.01.0097

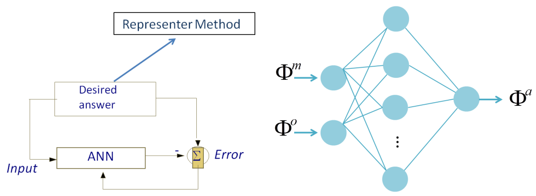

Figure 1: Computing analysis by an Artificial Neural Network. (a) training stage (emulating representer method); (b): Analysis calculated.

Table 1: Parameters ANN.

	forcing			i.c.		
	u	v	q	u	v	q
η	0,005	0,003	0,005	0,009	0,009	0,009
nhl	35	35	35	45	40	40

assimilation process during all time-integration period. Figure 4 present the error. Error is the absolute difference between estimated by ANN and true reference, we have the errors to q, u and v variable at time $t = 39$, respectively. of the entire field for all model variables. The Table 1 presents the ANN parameters used: learning rate (η) and number of neurons in the hidden layer(nhl), the parameters were determined empirically. This paper estimated the initial conditions, forcings and boundary condition. Developed up a network for each estimated parameter. The Figure 2 presents the model of the integration area, the green dots represent the assimilated observations.

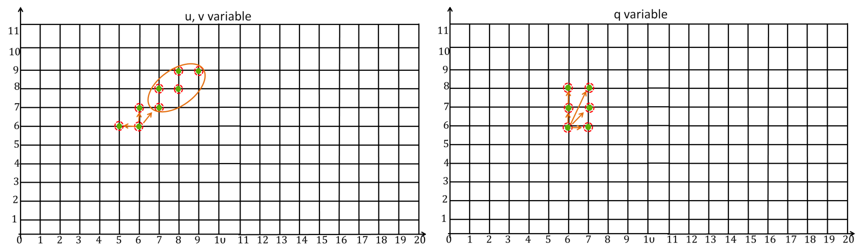

Figure 2: Interpolated observations at the grid point: a) u and v variable; b) q variable.

4 Conclusion

Numerical prediction system is an initial value problem. A better representation of the initial conditial will produce a better forecast. Many methods have been developed for data assimilation. These methods have different strategies to combine the forecasting *(background)* and observations. They differ in the quality of its matching with the real dynamics and the computational cost. In this work, two techniques for data assimilation were compared. These methods were tested to shallow water 2D model. These methods were tested on the shallow water 2D modeling problem. A model commonly used to test new schemes in meteorology and oceanography. According to the results obtained in this work, ANN is a competitive method for data assimilation.

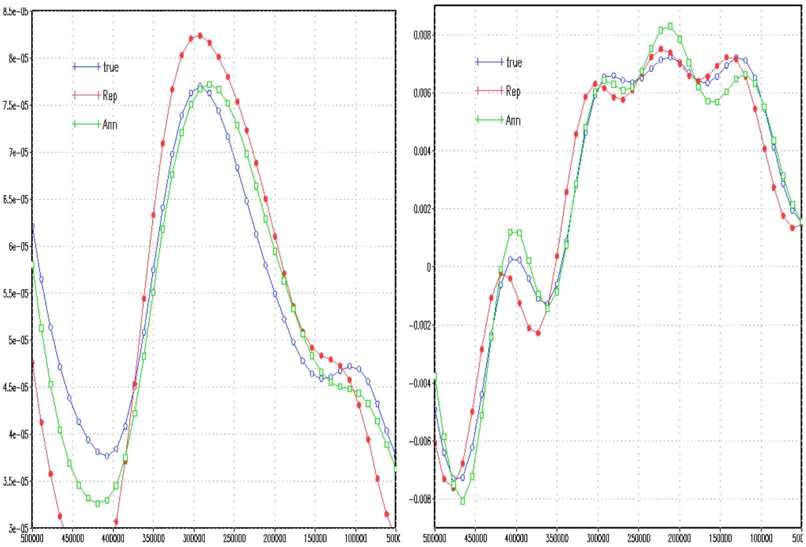

Figure 3: Temporal evolution of point v(6,6) and q(6,6), respectively. Representer Method (red curve), ANN (green curve), True (blue curve).

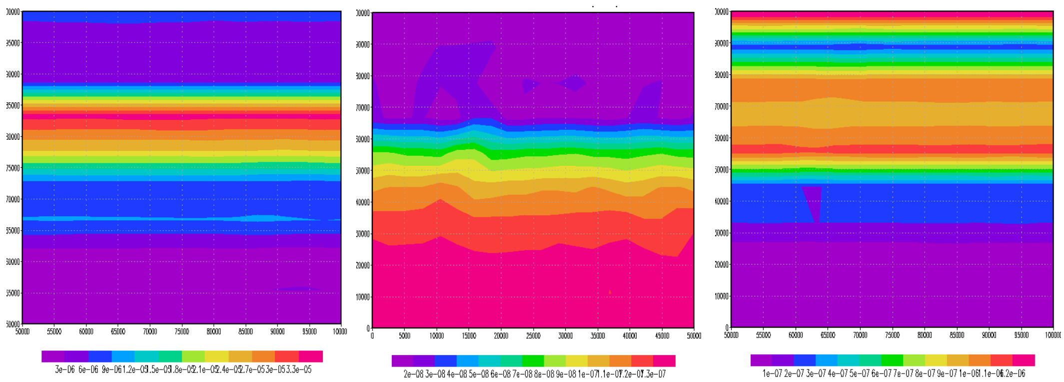

Figure 4: Data assimilation error. Variables: (a) q, (b) u and (c) v.

References

[1] K. P. Belyaev, C. A. S. Tanajura and J. J. O'Brien, A data assimilation method used with an ocean circulation model and its application to the tropical Atlantic. Applied Mathematical Modelling, v. 25 - 655-670), (2000).

[2] A. F. Bennett, Inverse Modeling of the Ocean and Atmosphere, Cambridge University Press. (2002).

[3] R. S. C. Cintra, H. F. Campos Velho and H. C. M. Furtado. Neural Network for performance Improvement in atmospheric prediction systems: Data Assimilation. In: BRICS Countries Congress (BRICS-CCI) and 11th Brazilian Congress (CBIC) on

Computational Intelligence, Porto de Galinhas. Congresso Brasileiro de Inteligência Computacional, (2013).

[4] D. Daley, Atmospheric Data Analysis. Cambridge University Press, (1993).

[5] H. C. M. Furtado, H. F Campos Velho, E. E. N. Macau and F. P. Härter, Diferent data assimilation methods applied to the Lorenz dynamical system, 6th Brazilian Conference on Dynamics, Control and Their Applications (DINCON-2007) May 22-25 (UNESP São José do Rio Preto (SP) Brazil) Proceedings in CD-Rom 1030-1035 - in Portuguese, (2007).

[6] T. W Garner, R. A. Wolf, R. W. Spiro and M. F. Thomsen, First attempt at assimilation data to constrain a magnetospheric model. Journal of Geophysical Research 104, 25145-25152, (1999).

[7] C. Grebogi, E. Ott, J. Yorke, Chaos, strange attractors, and fractal basin boundaries in nonlinear dynamics. Science, 238, 585-718, (1987).

[8] A. K. Griffith, Data assimilation for numerical weather prediction using control theory. University of Reading, Reading, Inglaterra, 1997-04, (1997).

[9] S. Haykin, Neural Networks: A Comprehensive Foundation, New York: Prentice Hall, (1993).

[10] F. P. Härter and H. F. de Campos Velho, New approach to applying neural network in nonlinear dynamic model. Applied Mathematical Modelling - Science Direct, 32, 2621-2633, (2008).

[11] A. Ismail-Zadeh and P. J. Tackley, Computational Methods for Geodynamics, Cambridge University Press, First edition. United States of America: Cambridge University Press, (2010).

[12] E. Kalnay, Atmospheric modeling, data assimilation and predictability, Cambridge, (2003).

[13] F. Mesinger and A. Arakawa, A Numerical methods used in atmospheric models, GARP Publication, WMO/ICSU Joint Organizing Committee, Geneva, Switzerland, 64, (2002).

[14] M. Nadler and E. P. Smith, Pattern Recognition Engineering, New York: John Wiley & Sons, (1993).

[15] A. G. Nowosad, A. Rios Neto and H. F. de Campos Velho, Data Assimilation in Chaotic Dynamics Using Neural Networks, Third International Conference on Nonlinear Dynamics, Chaos, Control and Their Applications in Engineering Sciences, July 31 - August 4, Campos do Jordão (SP), Brasil, Vol. 6, Chapter 6 - Control, 212-221, (2000).

DOI: 10.5540/03.2016.004.01.0097