Neural network for data assimilation method applied to shallow water equation

Article · June 2014

5 authors, including:

Helaine Furtado
Universidade Federal do Oeste do Pará
29 PUBLICATIONS **93** CITATIONS

Rosângela Cintra
National Institute for Space Research, Brazil
35 PUBLICATIONS **95** CITATIONS

Haroldo Fraga de Campos Velho
National Institute for Space Research, Brazil
494 PUBLICATIONS **3,517** CITATIONS

Elbert E N Macau
Universidade Federal de São Paulo
322 PUBLICATIONS **2,310** CITATIONS

UNCERTAINTIES 2014
Proceedings of the 2nd International Symposium on Uncertainty Quantification and Stochastic Modeling
July 7th to July 11th, 2014, Rouen, France

Neural network for data assimilation method applied to shallow water equation

Helaine C. M. Furtado[1], **Rosangela S. Cintra**[1], **Haroldo F. de Campos Velho**[1], **Elbert E. N. Macau**[1]

[1] Laboratório Associado de Computação e Matemática Aplicada LAC/CTE/INPE

Abstract. *The description of the most physical phenomena is based on differential equations. But, the modeling error is a permanent feature. One basic uncertainty is to identify the initial condition. For the operational prediction systems, a strategy to deal with such uncertainty is to add some information from the real world into the mathematical model. This additional information consists of observations (measurement values). However, the observed data might be carefully inserted, in order to avoid negative impacto on the prediction. Techniques for data assimilation are tools to produce an effective combination of two sources of data: observation and model, for computing the analysis. The analysis is the initial condition used in the prediction computer model. The goal of the present work is to employ artificialneural networks as a data assimilation method applied to shallow water equation used to represent ocean dynamics.*

Keywords. *data assimilation, artificial neural networking, Kalman Filter, shalow water equation*

1 INTRODUCTION

Many geosciences problems, for exemplo, meteorology, oceanography, and geophysics are modeled by differential equations. This problems may to require the estimation of state variable on time using noisy measurements. However, the mathematical model is always an approximation of reality (Ismail-Zadeh, and Tackley, 2010). Thus, the modeling error is a permanent feature. For the operational prediction systems, a strategy to deal with such uncertainty is to add some information from the real world into the mathematical model. This additional information consists of observations, this is, measurement values from the modeled phenomena. However, the observed data might be carefully inserted, in order to avoid the degradation of the prediction. Data assimilation techniques combine these sources (data model and data observation) in order to obtain the data analysis. The analysis is the initial condition used in the prediction computer model.

An important key point to understanding data assimilation tools is errors concept, in other words, the estimation of the error associated with the observed data and modeling error. Observations has errors from various sources: instrumentation, sampling, and representativeness. The dynamic model is imperfect, because has errors deriving from approach physical (chemical or biological) that govern the explicit evolution of the state variables, the representation of physical processes described by the use of parameters. Therefore, it is to require to deal with uncertainties and limitations inherent in the physical- mathematical model and the uncertainty associated with the observed data. These factors stimulate the use of data assimilation tools.

The basic components for operational forecasting systems are: a network of observation data, the numerical model, and the method of data assimilation. Data assimilation is a science to produce a nice combinations of data assimilation from mathematical model and from observations in order to produce the analysis or initial condition for the numerical predictions systems. The historical evolution of the methods of data assimilation passes through by adjustment functions (verificar se está correto a tradução), successive corrections, analysis correction, optimal Interpolation, Variational Methods, Kalman filter, Monte Carlo techniques, and artitifical neural networking (Daley, 1993, Griffith, 1997, Nowosad, 2000, and Kalnay, 2003). All methods imply different ways of combining forecast with observations. From a mathematical point of view, we have:

(1) forecast step:

$$\mathbf{x}^f_{k+1} = \mathbf{F}(\mathbf{x}^a_k) \tag{1}$$

(2) analysis step, that can be computed of different way, as follow:

(2.1) estimation theory

$$\mathbf{x}^a_k = \mathbf{x}^f_k + \mathbf{K}(\mathbf{y}_k - \mathbf{H}[\mathbf{x}^f_k]) \tag{2}$$

(2.2) variational calculus

$$\mathbf{x}^a_k = \mathrm{Min} J(\mathbf{x}^f) \tag{3}$$

$$\mathbf{x}^a_k = \mathbf{x}^f_k + \sum_{m=1}^{M} \beta_m r_m(x,k) \tag{4}$$

(2.3) artificial neural networking

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

where $\mathbf{x}_k^p$ denotes the vector of model states, k is discrete time, $\mathbf{x}_k^a$ is the value of the analysis (initial conditional), $\mathbf{y}_k$ observations vector. $\mathbf{F}$ is a non-linear operator describing the evolution of the states from time k to time $k+1$ (Equation 1. Equation 2, $\mathbf{H}$ is a linear operator that represents the observation system, $\mathbf{K}$ is the weight matrix (Kalman gain). Equation 4, x is coordenate space, β is representer coefficient, and r is the value representer. $\mathbf{F}_{RNA_w^*}$ is mapping function between input data $(\mathbf{y}_k, \mathbf{x}_k^p)$ and output data $\mathbf{x}_k^a$ in order to find w^* optimal weight (Equation 5).

2 TESTING MODEL: LINEAR SHALOW WATER EQUATION 2D

The model presented in this work is linear in two dimensions, given by:

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

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

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

The spatial domain is $0 \le x \le X$ e $0 \le y \le Y$. We have f is the Coriolis coefficient associated with the Coriolis force; 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 (see Figure1), if $q \equiv q'$, then the ocean is in hydrostatic balance (Bennett, 2002); F_u e F_v are the model forcings H is mean depth the ocean.

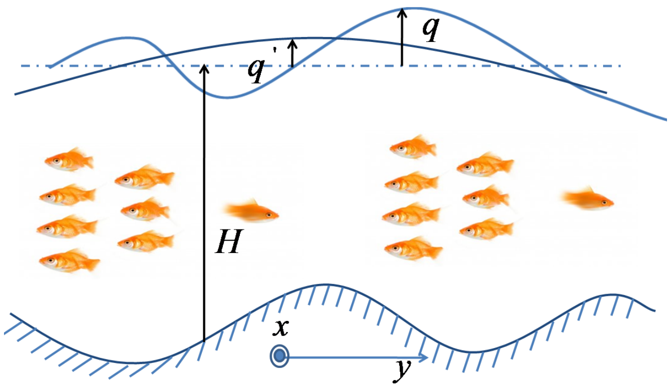

Figure 1: Shallow-water theory. Adapted from (Bennett, 2002) page 127.

The differential equations (6a, 6b, and 6c) are discretized on the Arakawa C-grid according to Figure 2 with a forward-backward scheme for time-stepping (Mesinger, 1976). The boundary conditions we have periodic channel with rigid walls, namely, $u(x,0,t) = v(x,Y,t) = 0$ in accordance with Figure 3.

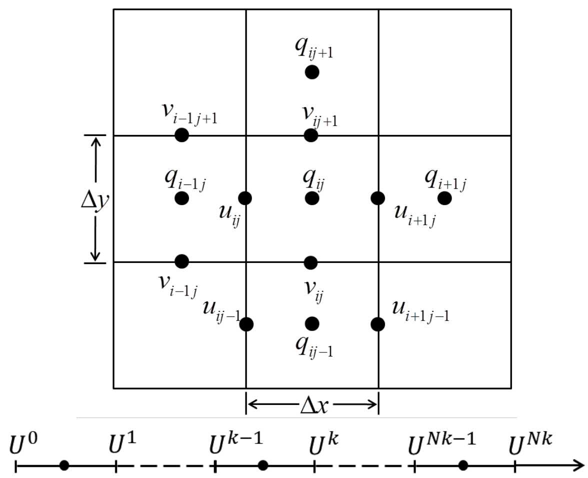

Figure 2: (above) Arakawa C-grid for space differences; (below) sheme for time-stepping.

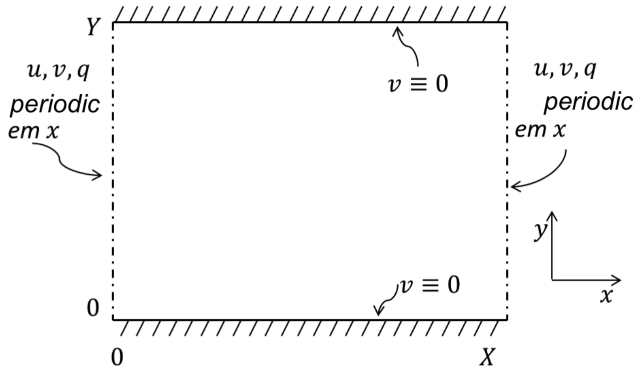

Figure 3: Periodic channel with rigid walls, rotating as the rate $\frac{f}{2}$ about an axis normal to the $xy-$ plane (Adapted (Bennett, 2002) page 196.

Table 1: Model parameters.

parameters	value
H	$5000\ m$
T	$1,8\times10^4\ s$
g	$9,806 m\ s^{-2}$
f	$1,0\times10^{-4} s^{-1}$
r_u	$(1,8\times10^4 s)^{-1}$
r_v	$(1,8\times10^4 s)^{-1}$
r_q	$(1,8\times10^4 s)^{-1}$
C_d	$1,6\times10^{-3}$
ρ_a	$1,275\ kg\ m^{-3}$
ρ_w	$1,0\times10^3\ kg\ m^{-3}$

The model forcings are: $F_u = -C_d\rho_a u_a^2/(H\rho_w)$ and $F_v = 0$. The parameters used in order to integrate the model is showed at Table 1.

Figure 13 and 14 displays the As Figuras 13 e 14 mostram a projeção em três dimensões da variável q, para os passos de tempo 60 e 100, respectivamente. Essa estimativa foi realizada com a RNA.

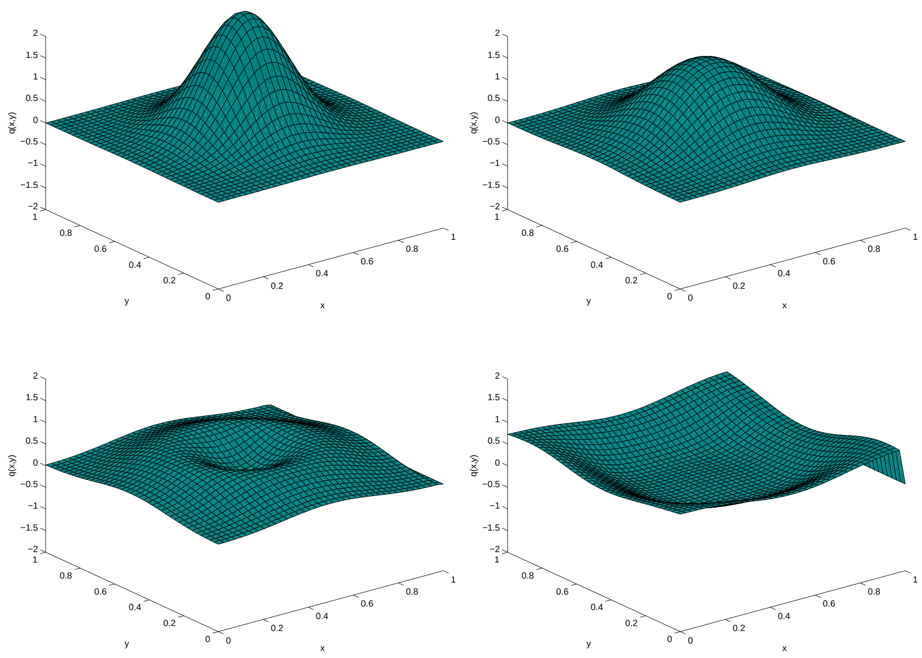

Figure 4: Temporal evolution of the variable q, projections (x,y); initial condition is a gaussian function.

3 METHODOLOGY

In this section we describe the techniques used: the Kalman Filter and Artificial Neural Networking by Multilayer Perceptron in the context to data assimilation.

3.1 Kalman Filter

A Kalman filter provides a recursive solution to the linear optimal filtering problem. It is an optimal estimator, infers parameters of interest from indirect, inaccurate and uncertain observations. It is recursive so that new measurements can be processed as they arrive.

Consider the nonlinear stochastic discrete-time dynamical system 7:

$$\begin{aligned} \mathbf{x}^f_{k+1} &= \mathbf{F}_k \mathbf{x}^f_k + \mu_k \\ \mathbf{y}^f_k &= \mathbf{H}_k \mathbf{x}^f_k + \nu_k \end{aligned} \tag{7}$$

where $\mathbf{F}_k$ is a mathematical model, μ_k is the model error, $\mathbf{H}_k$ is the mesurement function, ν_k is the white-noise sequence associetaded to observations. Under the specified hypotheses the optimal way, in the least squares sense, to assimilate sequentially the observations is given by the Kalman filter algorithm defined below by recurrence over the observation times:

1. State forecast:

$$\mathbf{x}^f_{k+1} = \mathbf{F}\mathbf{x}^a_k \tag{8}$$

2. Error covariance forecast:

$$\mathbf{P}^f_{k+1} = \mathbf{F}_k \mathbf{P}^a_k \mathbf{F}^T_k + \mathbf{Q}_k \tag{9}$$

3. Kalman gain computation:

$$\mathbf{K}_{k+1} = \mathbf{P}^f_{k+1} \mathbf{H}^T_{n+1} [\mathbf{R}_{k+1} + \mathbf{H}_{k+1} \mathbf{P}^f_{k+1} \mathbf{H}^T_{k+1}] \tag{10}$$

4. State analysis:

$$\mathbf{x}^a_{k+1} = \mathbf{x}^f_{k+1} + \mathbf{K}[\mathbf{y}^o_{k+1} - \mathbf{H}_{k+1} \mathbf{x}^f_{k+1}] \tag{11}$$

5. Error covariance of analysis:

$$\mathbf{P}^a_{k+1} = [\mathbf{I} - \mathbf{K}_{k+1} \mathbf{H}_{k+1}] \mathbf{P}^f_{k+1} \tag{12}$$

In order to use the Kalman Filter method is necessary to determine the dynamic matrix model. This matrix is computed as the physical problem. To develop this dynamic matrix on time is a problem. For the model described in section 2 we need to estimate the state variable model $\Psi = [\mathbf{q}, \mathbf{u}, \mathbf{v}]$, see Equation 13, we have:

$$\Psi^{n+1} = \mathbf{F}\Psi^n$$

where Ψ is given by:

$$\Psi = \begin{bmatrix} \mathbf{Q} \\ \mathbf{U} \\ \mathbf{V} \end{bmatrix}$$

and the vectors, $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{Q}$ are variables $u(x,y,t)$, $v(x,y,t)$, and $q(x,y,t)$ after space discretization:

$$\mathbf{Q} = \begin{bmatrix} \begin{bmatrix} q_{11} \\ q_{21} \\ \vdots \\ q_{nx,1} \end{bmatrix} \\ \begin{bmatrix} q_{12} \\ q_{22} \\ \vdots \\ q_{nx,2} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} q_{nx,1} \\ q_{nx,2} \\ \vdots \\ q_{nx,ny} \end{bmatrix} \end{bmatrix} \qquad \mathbf{U} = \begin{bmatrix} \begin{bmatrix} u_{11} \\ u_{21} \\ \vdots \\ u_{nx,1} \end{bmatrix} \\ \begin{bmatrix} u_{12} \\ u_{22} \\ \vdots \\ u_{nx,2} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} u_{nx,1} \\ u_{nx,2} \\ \vdots \\ u_{nx,ny} \end{bmatrix} \end{bmatrix} \qquad \mathbf{V} = \begin{bmatrix} \begin{bmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{nx,1} \end{bmatrix} \\ \begin{bmatrix} v_{12} \\ v_{22} \\ \vdots \\ v_{nx,2} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} v_{nx,1} \\ v_{nx,2} \\ \vdots \\ v_{nx,ny} \end{bmatrix} \end{bmatrix}$$

To shallow water equation the matrix $\mathbf{F}$ is given by:

$$\mathbf{F} = \begin{bmatrix} (1-r_q)\mathbf{I} & \mathbf{A}_1 & -\mathbf{B}_1 \\ -\mathbf{A}_1 & (1+\triangle t r_u)\mathbf{I} & \mathbf{B}_2 \\ \mathbf{B}_1 & -\mathbf{B}_2 & (1+\triangle t r_v)\mathbf{I} \end{bmatrix} \tag{13}$$

where

$$\begin{aligned} (\mathbf{A}_1)_{ii} &= -H\frac{\triangle t}{\triangle x} \\ (\mathbf{A}_1)_{i,i+3} &= H\frac{\triangle t}{\triangle x} \\ (\mathbf{B}_1)_{ii} &= -(\mathbf{B}_1)_{i,i+1} = g\frac{\triangle t}{\triangle y} \\ (\mathbf{B}_2)_{ii} &= -(\mathbf{B}_2)_{i,i+1} = f\frac{\triangle t}{4} \end{aligned}$$

The differential equations 6a, 6b, and 6c are discretized on the Arakawa C-grid. The parameters are described at Section 2. $\triangle x$ is grid spacing (x-direction) and $\triangle y$ is grid spacing (y-direction); $\triangle t$ is temporal spacing.

The dimension of the matrix 13 is related to number of points used for the spatial discretization. It can be seen that the method of assimilation through the Kalman filter is computationally very costly. In addition to, the dynamic model matrix $\mathbf{F}$ we have the error covariance matrices: modeling and observation, which are updated at each time step. The matrix operations, such as, multiplication, inverse matrix calculation carried out to compute the Kalman gain matrix to make this apply to large models heavy tool . These difficulties motivate an investigation of ANN tool as a method of data assimilation. In the next section we describe the Artificial Neural Networking as a data assimilation technique.

3.2 Artificial Neural Networking (ANN)

Artificial Neural Networks (ANN) have became important tools for information processing (Haykin, 2004). Much research has been conducted in pursuing new neural network models and adapting the existing ones to solve real life problems, such as those in engineering (Haykin, 2004). 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, Figure 1 (left side), given by:

$$y_k = \varphi\left(\sum_{j=1}^{n} w_{kj}x_j + b_k\right) \tag{14}$$

where w_{kj} are the connection weights, b_k is a threshold parameter, x_j is the input vector and y_k is the output of the k_{th} neuron.

Arrangements of such units form the ANN that 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.

ANN aim to explore the massively parallel network of simple elements in order to yield a result in a very short time slice and, at the same time, with insensitivity to loss and failure of some of the elements of the network. These properties make artificial neural networks appropriate for application in pattern recognition, signal processing, image processing, financing, computer vision, engineering, etc.

There are different architectures of ANN that are dependent upon the learning strategy adopted. This paper briefly describes the Multilayer Perceptron (MLP) with error backpropagation learning. Detailed introductions on ANN can be found in (Haykin, 2004) and (Nadler, 1993). MLP with backpropagation learning algorithm, are feedforward networks composed of an input layer, an output layer, and a number of hidden layers, whose aim is to extract high order statistics from the input data (Haykin, 2004). Figure 5 (right side) describe a multilayer neural network with a hidden layer. Functions $\varphi(.)$ provide the activation for the neuron. Neural networks will solve nonlinear problems, if nonlinear activation functions are used for the hidden and/or the output layers. From several activation functions, the sigmoid are commonly used:

$$\text{bipolar function } \varphi(v) = \frac{1-\mathtt{exp}(-av)}{1+\mathtt{exp}(-av)} \tag{15}$$

A feedforward network is a non-linear mapping to compute the output vector from an input vector. The connections among the several neurons (Figure 5 (right side)) 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.

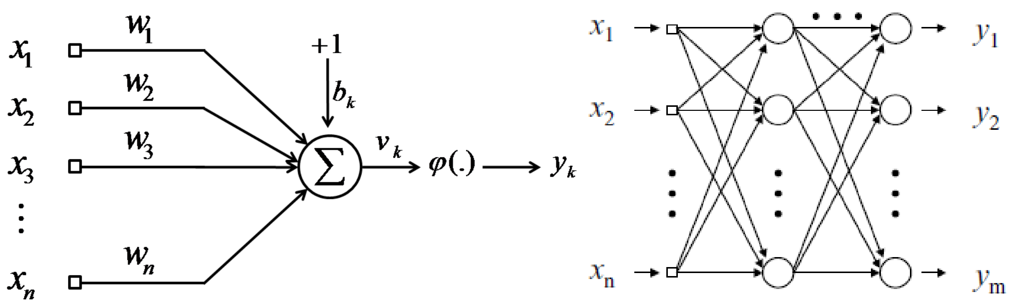

Figure 5: (left side) Single Neuron, (right side) Multilayer Neural Network

4 RESULTS

In order to evaluate the techniques used in this work was made two experiments that are described to follow.

4.1 Experiment 1

The first experiment was characterized by the integration of the model in 60 time steps and the data assimilation process was made each 10 time step. Table 2 shows the numerical model parameters used in this experiment. The q variable was initialized with gaussian function. Initially, the others variable u and v were zero, that is, on time initial the variable were resting state. The spatial distribution of the observations were made in the grid according to illustrated at Figure 8, 25 observations were assimilated in 10 time steps. In order to test the MLP-ANN for emulating an KF in data assimilation scenario were needed a architecture to each variable as shown at Figure 6. The input data are: data model $[u^m, v^m, q^m]^T$ and data observation $[u^o, v^o, q^o]^T$; output data is the data analysis $[u^a, v^a, q^a]^T$. The parameters used were:

- hyperbolic tangent activation function

$$\phi(v_j) = \tanh\left(\frac{\mathrm{av_j}}{2}\right) \tag{16}$$

- $\eta^u = 0,2$ and $\eta^v = \eta^q = 0,7$
- Number of hidden layer neurons (hln): $hln^u = hln^v = 15$ e $hln^q = 10$

The shallow water equations were integrated at 60 time steps. In order to training the MLP-ANN the data set was made until 40 time steps. The remaining time steps were used for the generalization according to 7.

Figure 9 shows a comparison between Kalman Filter (red curve) and artificial neural networking (green curve). The Figure display the evolution to variables u, v, and q variable. In the Table 3, we have the error value to each variable, to all cases the result obtained with ANN was better than the result by Kalman Filter. In this experiment just q variable was assimilated.

Table 2: Numerical model parameters.

parameters	value
number of grid points in $x-$direction	$NI = 40$
number of grid points in $y-$direction	$NJ = 40$
number of time steps	$NK = 60$ and $NK = 100$
grid spacing (x-direction)	$\triangle x = 100\ km$
grid spacing (x-direction)	$\triangle y = 100\ km$
time step	$\triangle t = 180\ s$

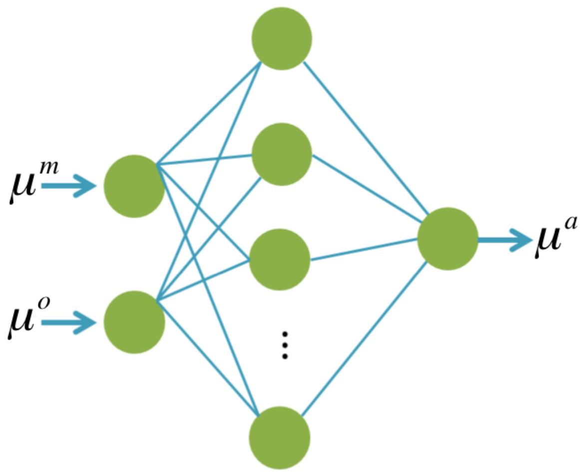

Figure 6: Architecture of ANN for each variable to shallow water model with $\mu = u, v, q$.

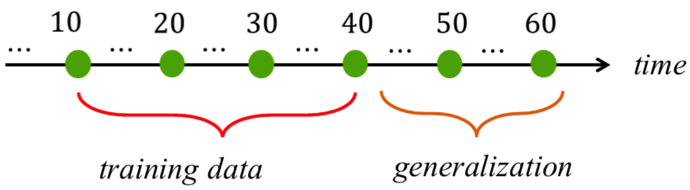

Figure 7: Composition of the training set for the experiment 1.

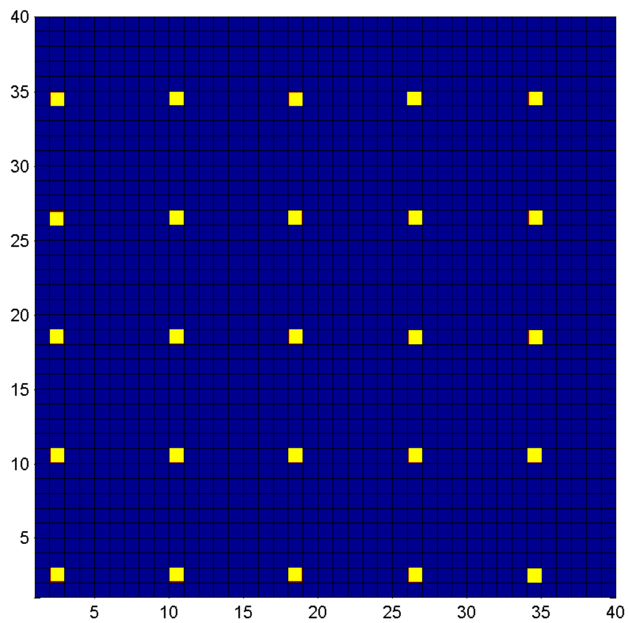

Figure 8: Spatial distribution of observations for the experiment 1. The yellow squares represent the observations at each grid point, 25 observation assimilated.

The estimation error is given by:

$$\varepsilon = \int_0^{T_{last}} \int_\Omega \left| \phi_i^{est} - \phi_i^{true} \right|$$

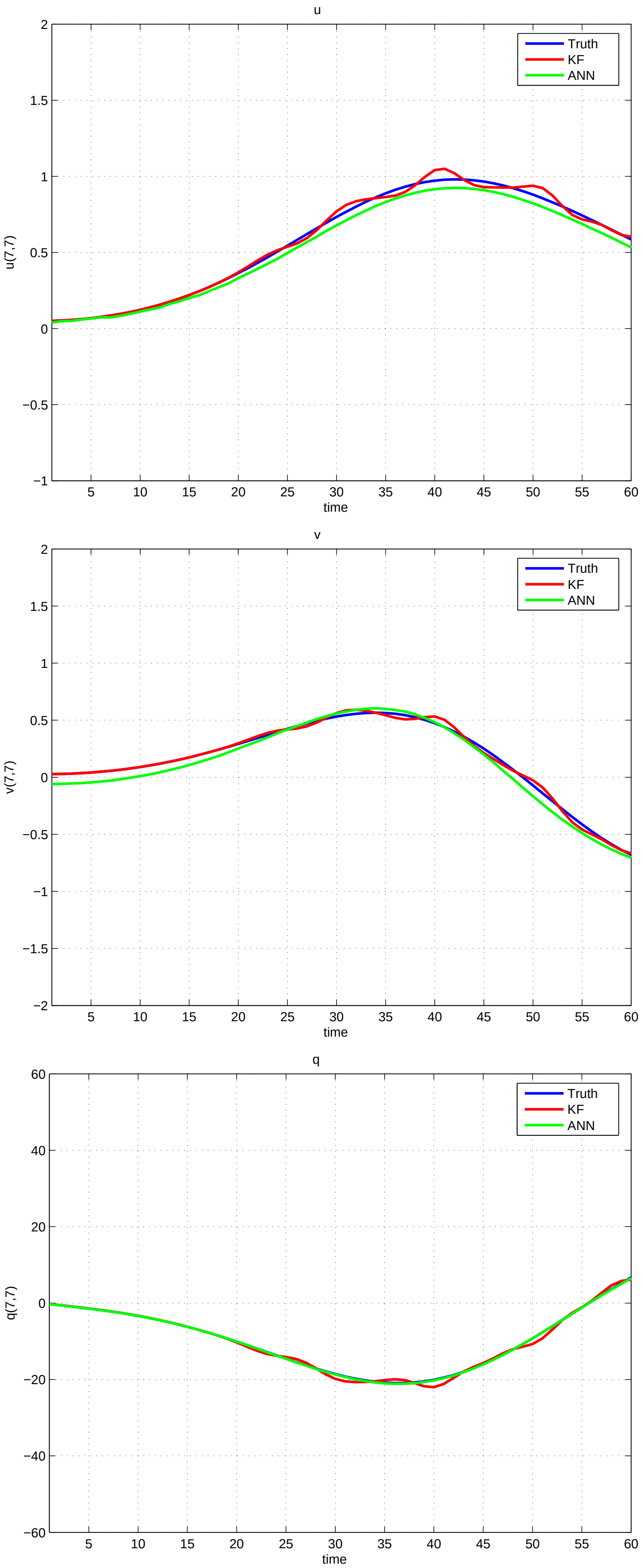

Figure 9: Experiment 1: temporal evolution of point $u, v, q(7,7)$.

where: $\Omega \in (0, L_x) \times (0, L_y)$. ϕ_i^{est} is a estimated value and ϕ_i^{true} is reference truth.

Table 3: Error results of assimilation to KF and ANN (experiment 1).

variable	ANN-MLP	KF
u	0,0199	0,0296
v	0,7794	0,8554
q	0,1460	0,5969

4.2 Experiment 2

In this experiment all variable of the model are assimilated. The variable u, v, and q were initialized with gaussian function. The model were integrated until 100 time step; 100 observations were assimilated the each 10 time step represented by green points at Figure 10.

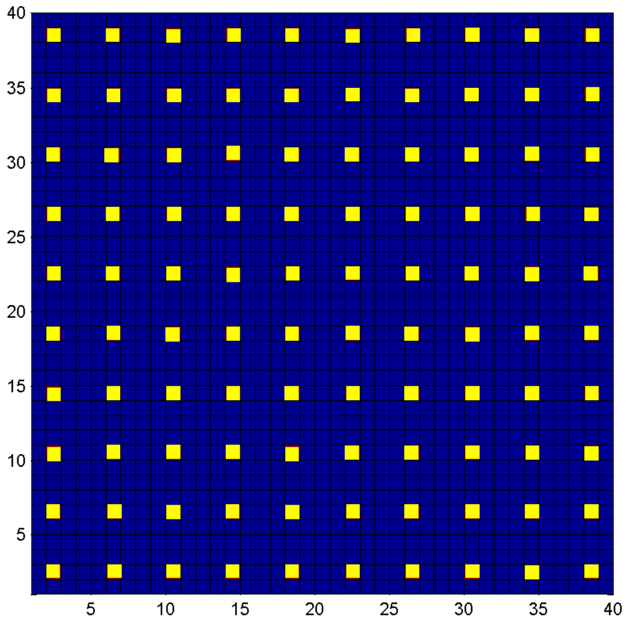

Figure 10: Spatial distribution of observations for the experiment 2. The yellow squares represent the observations at each grid point, 100 observation assimilated.

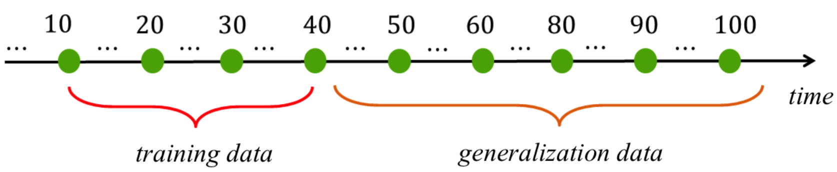

Figure 11: Training data set and generalization data to experiment 2.

Figure 11 shows how to select training data set and generalization data for experiment 2. The shallow water equations was integrated at 100 time steps. The generalization was done after time step 41. In this experiment the result obtained with ANN had agreement with true until 100 time steps.

Figure 12 shows the evolution of the state variables. We have that green curve is result obtained with ANN, and red curve obtained with KF method, and the blue line is the reference of truth. The Table 4 presents the average error (given by Equation 17) to variable u, v, and q, again the result by ANN was better than KF method.

Table 4: Error results of assimilation to KF and ANN (experiment 2).

variable	ANN-MLP	KF
u	0,0211	0,0290
v	0,8441	0,8515
q	0,1971	0,8951

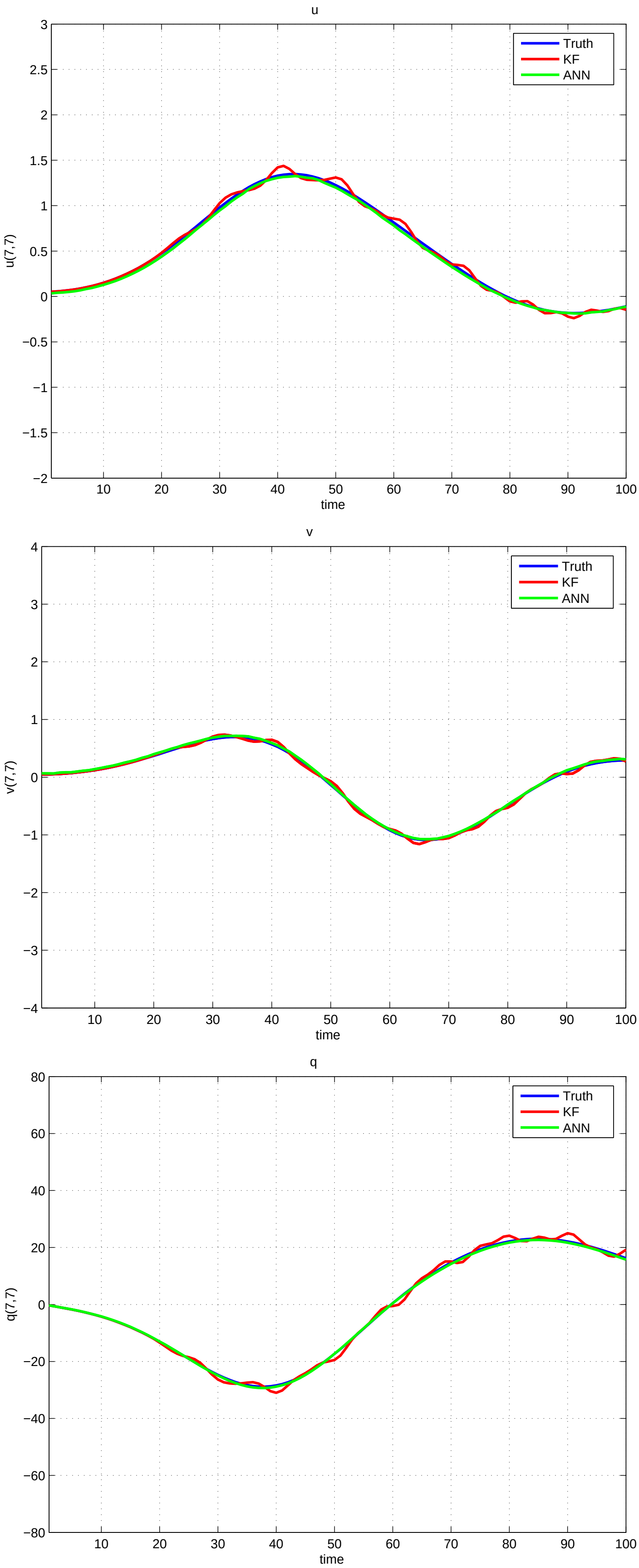

Figure 12: Experiment 2: temporal evolution of point $u, v, q(7,7)$.

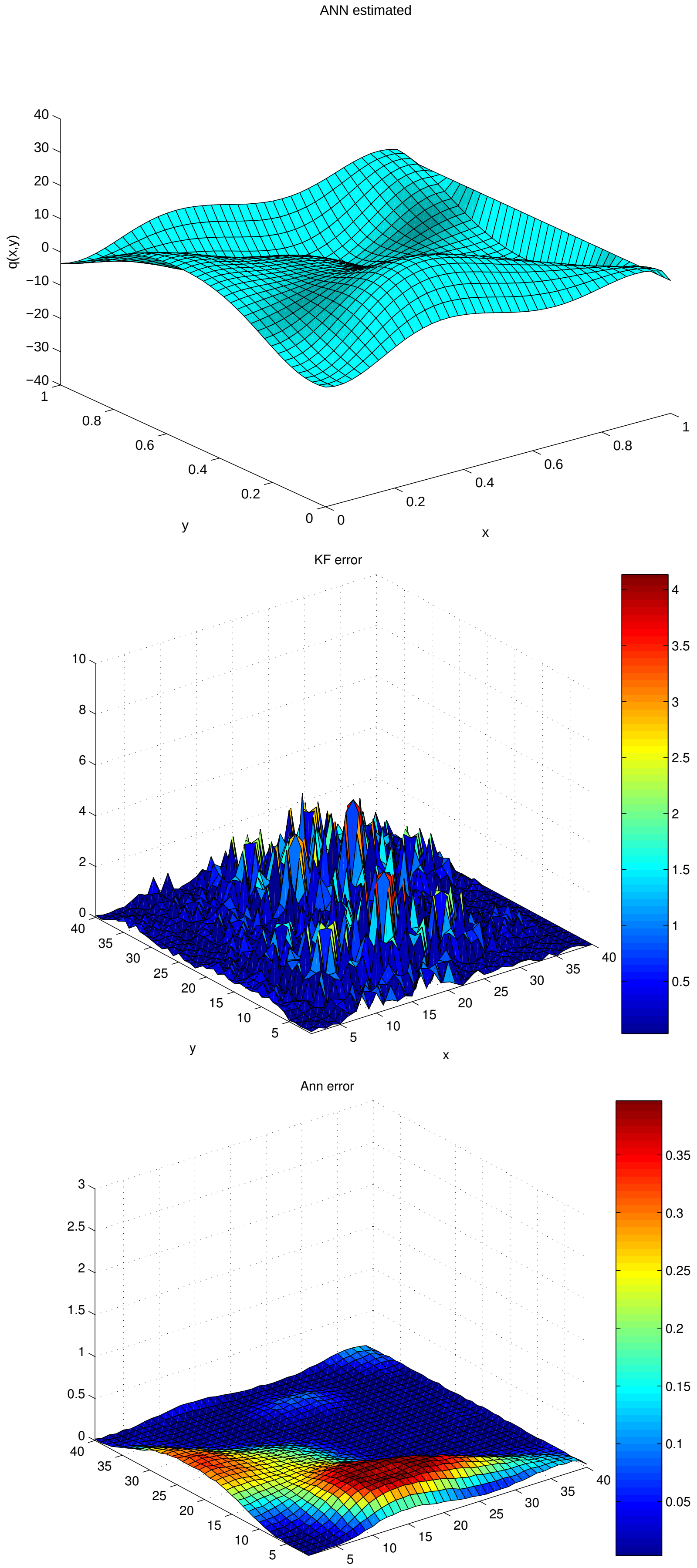

Figure 13: Projection of the variable q in three dimensions in time step 60. Error surface: KF method (figure middle) and Ann.

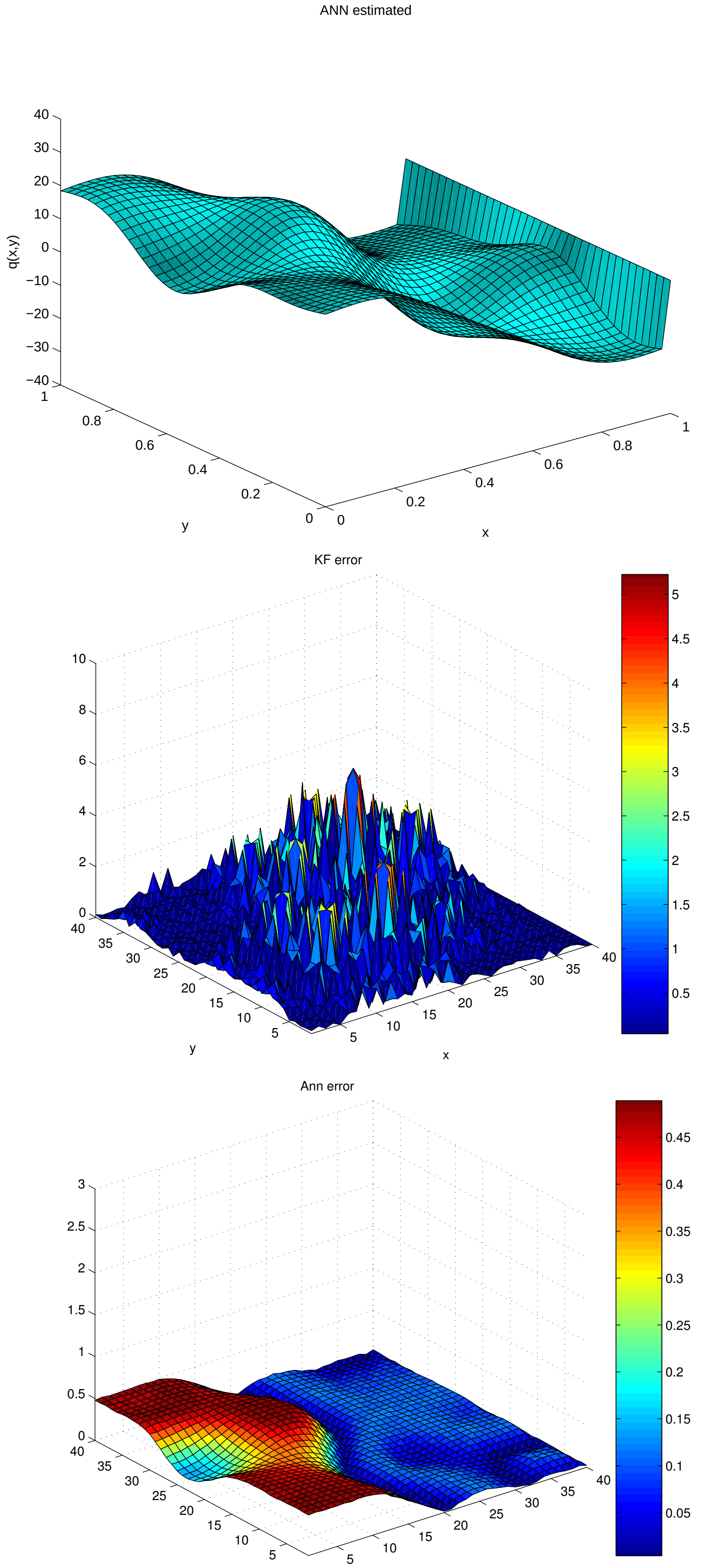

Figure 14: Projection of the variable q in three dimensions in time step 100. Error surface: KF method (figure middle) and Ann.

5 CONCLUSION

Numerical prediction system is an initial value problem, and this implies that a better representation of the initial conditial will produce a better forecast. Many methods have been developed for data assimilation (Kalnay, 2003). These methods have different strategies to combine the forecasting *(background)* and observations (see Equation 2, 3, 4, and 5) and differ in their numerical cost, and their suitability for real-time applications. In this work was compared two techniques for data assimilation. These methods was tested at shallow water model (described at Section 2) that are commonly used to test problems in meteorology and oceanography, because they describe characteristics flow present in the atmosphere and the ocean.

Two experiments were managed. The first experiment was assimilated 25 observations in 10 time steps and the model was integrated to 60 time steps, only the variable q was assimilated. The second experiment were assimilated 100 observations in 10 time steps, in this case, the model was integrated into 100 time steps, and all model variables were assimilated. In both experiments, the result assimilation performed with the ANN had the lowest error than Kalman Filter. However, it is important to emphasize that the main advantage of ANNs is the computational performance. Once done training the network, the processing is faster than the Kalman filter methods, for example, as shown in Table 5. The disadvantage of the Kalman filter method is the handling of arrays. The algorithm requires the computation of inverses, which makes the method costly. For the application of operational problems it is an important factor to be considered. According to the results obtained in this work, it leads to believe that the tool ANN is a competitive method as a technique of data assimilation.

Table 5: Computational cost.

	KF	ANN-MPL
experiment 1	42 *min*	1.39 *min*
experiment 2	79 *min*	5.00 *min*

6 REFERENCES

Daley, D. 1993 “Atmospheric Data Analysis” (Cambridge University Press).

Kalnay, E. 2003 “Atmospheric modeling, data assimilation and predictability” (Cambridge University Press).

Bennett, A. F. 2002 “Inverse Modeling of the Ocean and Atmosphere” (Cambridge University Press).

Garner, T. W; Wolf, R. A; Spiro, R. W; Thomsen, M. F. 1999 “First attempt at assimilation data to constrain a magnetospheric model” Journal of Geophysical Research 104, 25145-25152.

GRIFFITH, A. K. Data assimilation for numerical weather prediction using control theory. University of Reading, Reading, Inglaterra, 1997-04- 1997.

Schunk R W, Scherliess L, Sojka J J, Thompson D C, Anderson D N, Codrescu M, Minter C, Fuller-Rowell T J, Heelis R A, Hairston M, Howe B M 2004 “Global assimilation of ionospheric measurements” (GAIM00). Radio Science, 1, RS1S02.

Belyaev, K. P; Tanajura, C. A. S. and O’Brien J. J. 200 “A data assimilation method used with an ocean circulation model and its application to the tropical Atlantic”. Applied Mathematical Modelling, v. 25 - 655-670)

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

F. P. Härter, H. F. de Campos Velho “New approach to applying neural network in nonlinear dynamic model”. Applied Mathematical Modellin - Science Direct 32 (2008) 2621-2633).

Nakano S, Ueno G, Higuchi T 2007 “Merging particle filter for sequential data assimilation Nonlinear” (Processesin Geophysics 14 395-408).

Chua, B. S.; Bennett, A. F. ”An inverse ocean modeling system” (Ocean modeling, v.3, p. 137-165, March 2001).

Jazwinski, A. H. 1970 “Stochastic Processes and Filtering Theory” (Academic Press).

Evensen, G. 2007 “The Ensemble Kalman Filter” (Springer Verlag).

Belyaev, K; Meyers, S. O’Brien J J 2000 “Fokker-Planck equation application to data assimilation into hydrodynamic models” J. Math. Sci., 99(4) 1393-1403.

Belyaev, K. P; Tanajura, C. A; S, O’Brien J J 2001 “A data assimilation method used with an ocean circulation model and its application to the tropical Atlantic Appl. Math. Modelling 25 655-670.

Campos Velho H F, Cintra R S, Furtado H C M. “Introduction to Data Assimilation”. Short-course, INPE, São José dos Campos, Brazil (In Portuguese). Available in the internet (Accessed: September 22, 2010):
`www.lac.inpe.br/~haroldo/Curso-DataAssimilation/Curso_Assi_Dados-INPE-2.pdf`.

Furtado H C M, Campos Velho H F, Macau E E N “Redes neurais e diferentes métodos de assimilação de dados em dinâmica não linear” (125 p). (INPE-15235-TDI/1322). MSc thesis on Applied Computing – Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2008. Available in the internet:
`http://urlib.net/sid.inpe.br/mtc-m17@80/2008/02.07.10.49`

Nowosad A G, Rios Neto A, Campos Velho H F 2000 “Data assimilation in chaotic dynamics using neural networks”. In: International Conference on Nonlinear Dynamics, Chaos, Control and Their Applications in Engineering Sciences 6 212-221 (Associação Brasileira de Ciências Mecânicas).

Härter F P, Campos Velho H F 2005 “Recurrent and feedforward neural networks trained with cross correlation applied to the data assimilation in chaotic dynamic”. Brazilian Journal of Meteorology 20 411-420.

Härter F P, Campos Velho H F 2008 “New approach to applying neural network in nonlinear dynamic model” Applied Mathematical Modelling 32 2621-2633 - DOI 10.1016/j.apm.2007.09.006 - ISSN: 0307-904X.

Härter F P, Rempel E L, Campos Velho H F, Chian A 2008 “Application of artificial neural networks in auroral data assimilation”. Journal of Atmospheric and Solar – Terrestrial Physics, 70 1243-1250, - DOI 10.1016/j.jastp.2008.03.018.

Lorenz, E. N. 1963 “Deterministic Nonperiodic Flow” Journal of the Atmospheric Sciences 20(2) 130-141.

Lorenz, E. N. 1965 “A study oh the predictability of a 28-variable model” Journal Tellus 17

Haykin, S. 1993 “Neural Networks: A Comprehensive Foundation” (New York: Prentice Hall).

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

Furtado, H. C. M. and Campos Velho, H. F. and Macau, E. E. N. and Härter, F. P. 2007 “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.

Ismail-Zadeh, A. and Tackley, P. J., 2010,“Computational Methods for Geodynamics”, Cambridge University Press, First. United States of America: Cambridge University Press, 2010. 332 p. ISBN 978-0-521-86767-2.

MESINGER, F.; ARAKAWA, A. Numerical methods used in atmospheric models. GARP Publication, WMO/ICSU Joint Organizing Committee, Geneva, Switzerland, p. 64 pp, 2002.