ELSEVIER

Ecological Modelling 98 (1997) 173–186

ECOLOGICAL MODELLING

Prediction of functional characteristics of ecosystems: a comparison of artificial neural networks and regression models

José M. Paruelo [1,a,*], Fernando Tomasel [2,b]

[a] *Depto Ecologia, Facultad de Agronomia, Universidad de Buenos Aires, Av. San Martin 4453, (1417) Buenos Aires, Argentina*
[b] *Department Electrical Engineering, Colorado State University, Fort Collins, CO 80523, USA*

Accepted 7 October 1996

Abstract

We tested the potential of artificial neural networks (ANNs) as predictive tools in ecology. We compared the performance of ANNs and regression models (RM) in predicting ecosystems attributes, with special emphasis on temporal (interannual) predictions of functional attributes of the ecosystem at regional scales. We tested the predictive power of ANNs and RMs using simulated data for six functional traits derived from the seasonal course of the normalized difference vegetation index (NDVI): the annual integral of the NDVI curve (NDVI-I), the maximum (MAX) and minimum (MIN) NDVI, the date of the MAX NDVI (DM) and the date of start (SGS) and end (EGS) of the growing season. For one of these traits (NDVI-I), we also generated a set of data that incorporated the effects of the state of the system in previous years (inertial effects). Even simple non-linearities in the actual functional form of the relationship between environmental variables and ecosystem attributes preclude a precise prediction of these attributes when the rules are not explicit. That was evident for predictions based on both ANNs and RMs under absolutely deterministic conditions (error-free scenario). Non-linearities in the simulated traits of the NDVI curve derive from multiplicative terms in the models. Under the presence of these non-linear terms, a different aggregation of the driving variables (monthly vs. annual or quarterly climatic data) reduce substantially the ability of both RMs and ANNs to predict the independent variable. For the six traits analyzed, the ANNs were able to make better predictions than RMs. The correlation between observed and predicted values of each of the six traits considered was higher for the ANNs than for the RMs. ANNs showed clear advantages to capture inertial effects. The ANN used was able to use previous year information on climate to estimate current year NDVI-I much better than the RM that used the same input information. © 1997 Elsevier Science B.V.

Keywords: Aboveground net primary production; Artificial neural networks; Grasslands; Normalized difference vegetation index; Prediction; Regression models; Remote sensing

* Corresponding author: Fax: + 54 1 5211384; e-mail: paruelo@ifeva.edu.ar

[1] Temporary address: Department of Rangeland Ecosystem Science, Colorado State University, Fort Collins, CO 80523, USA.

[2] On leave from: Depto Física, Facultad de Ingeniería, Universidad Nacional de Mar del Plata, (7600) Mar del Plata, Pcia Buenos Aires, Argentina.

PII S0304-3800(96)01913-8

1. Introduction

A proper evaluation of human impact on terrestrial ecosystems depends on the availability of 'systems of reference' (undisturbed areas). Reference systems need to be characterized not only in terms of averages but also in terms of the probabilistic distribution of extreme values. Given the global nature of many of the human-driven changes (atmospheric composition, climatic change, land-use) and the availability of long term datasets, the generation of these reference systems will depend on models able to reproduce in the future the present relationships between ecological attributes and environmental factors.

Regression or correlative models have been widely used in ecology to explain patterns of ecosystem attributes using environmental variables. The attributes analyzed include plant and animal diversity (Currie and Fritz, 1993), primary production (Le Houerou et al., 1988; Sala et al., 1988; Milchunas and Lauenroth, 1993), secondary production (McNaughton et al., 1989), herbivore biomass (Oesterheld et al., 1992), seasonality of the vegetation (Paruelo and Lauenroth, 1995), distribution of plant functional types (Teeri and Stowe, 1976, Paruelo and Lauenroth, 1996), and soil organic carbon (Burke et al., 1989) among others. Those models provide useful insights on the environmental controls of the ecosystem attributes analyzed. However, they do not always possess enough predictive power.

The lack of predictive power is particularly critical for temporal models, those analyzing how climate controls ecosystem attributes through time. Lauenroth and Sala (1992) showed, for a long-term dataset of aboveground net primary production (ANPP), that the proportion of the variance of ANPP explained by climate among years was substantially lower than for the spatial model (that generated using average values for different sites across a region). A common explanation for the low predictive power of temporal models is the existence of inertial effects. The response of the ecosystem in a given year is a function not only of the current year climatic conditions, but also of the conditions on previous years. Different kinds of non-linearities, and threshold responses are additional explanations for the lack of predictive power of these models.

The global nature of environmental problems requires the use of tools able to capture changes in the structure and function of the ecosystems at large scales. A description of the ecosystem function at broad scales may be derived from the seasonal course of the normalized difference vegetation index (NDVI) (Lloyd, 1990; Soriano and Paruelo, 1992; Running et al., 1994; Paruelo and Lauenroth, 1995). This index, derived from the red and infrared band of the AVHRR sensor on board of the NOAA satellites, shows a high correlation with biophysical rates of the target area, such as transpiration or primary productivity (Sellers et al., 1992). Information derived from remote sensing offers a unique opportunity for monitoring global change issues. However, to evaluate the impact of human-induced changes on the ecosystem, actual spectral data from, for example, agricultural areas needs to be compared with the expected patterns for undisturbed areas. Predictions for the mean and probabilistic distribution of the ecosystem attributes for these reference situations can be derived in different ways. Simulation and correlative models have been the most common tools used to predict ecosystems attributes from environmental variables.

Our objective is to show the potential of artificial neural networks (ANNs) as predictive tools in ecology. We compared the performance of ANNs and regression models (RMs) in predicting functional attributes of the ecosystem using climatic data among years. We tested the predictive power of ANNs and RMs using simulated data for six functional traits derived from the seasonal course of the NDVI: the annual integral of the NDVI curve (NDVI-I), the maximum (MAX) and minimum (MIN) NDVI, the date of the MAX NDVI (DM) and the date of start (SGS) and end (EGS) of the growing season. The use of simulated data allowed us to evaluate these tools in an error-free scenario. For one of these traits (NDVI-I), we also generated a set of data that incorporated the effects of the state of the system in previous years (inertial effects).

2. Methodology

2.1. Artificial neural networks

The field of ANNs has been extremely prolific since its resurgence in the early 1980s, and especially in the last few years. Many books are now available on this subject (Hertz et al., 1991; Hecht-Nielsen, 1991; Wasserman, 1993). Of the many different architectures available today, the most popular are the multi-layer, feed-forward networks. A feed-forward neural network can be considered as a transformation which maps a set of input variables into a set of output variables. For a single-layer network, these transformation can be expressed as follows:

$$A_i = f\left(\sum \mathbf{W}_{ij} I_j\right),$$

where I_j is the jth component of the input vector $\boldsymbol{I}$, A_i is the ith component of the output vector $\boldsymbol{A}$, and $\mathbf{W}$ and $f(.)$ are called the weight matrix and the activation function, respectively. Depending on the nature of the system under study, the components of the input vector can represent, for example, the value of different variables of the system measured at a given time, as it would be the case in a typical multivariate regression problem. Or, as in the case of prediction of the time evolution of a system, the component of the input vector can represent the value a given variable of the system for a number of past measurements. The range of problems treated by means of feed-forward multi-layer networks is very vast, including the analysis of a wide variety of environmental problems. Some examples are weather forecasting (McCann, 1992; Derr and Slutz, 1994), prediction of daily solar radiation (Elizondo et al., 1994), classification of remotely sensed data (Liu and Xiao, 1991; Kanellopoulos et al., 1992; Foody et al., 1995), resource management (Gimblett and Ball, 1995), and many others.

In the present paper we used a multi-layer network consisting of two layers of nodes fully interconnected, i.e. with links among all the nodes in adjacent layers. Each of the nodes was characterized by a sigmoidal transfer function on the hidden layer and a linear transfer function on the output layer. The links were characterized by a connection strength or weight, which was adjusted during the training of the network and stored once the network was trained.

The network was trained using the error back-propagation training algorithm, on which the error between the desired result and the result computed by the network is back-propagated through the network to adjust its weights. Training was based on the minimization of the total error by using the gradient descent technique (Hertz et al., 1991). The learning process was stopped when a specified error goal was reached.

Programming of the ANN was based on the neural network toolbox of MATLAB™. The learning speed of the pure backpropagation method was improved by introducing momentum, which makes backpropagation less sensitive to small features of the error surface and an adaptive learning rate, that maximizes the learning rate while maintaining the learning process stable. Appendix A summarizes the basic expressions describing the structure of the ANN (Fig. 1).

2.2. Data generation

To characterize the function of the ecosystem (its exchange of energy and matter) using the seasonal course of the NDVI, we derived six traits that capture the essentials of ecosystem carbon gain dynamics (Lloyd, 1990; Soriano and Paruelo, 1992; Running et al., 1994; Paruelo and Lauenroth, 1995). We analyzed the integral of the NDVI (NDVI-I), the MAX and MIN NDVI, the day of the year of the MAX NDVI (DM), the day of the year of the SGS, and the day of the year of the EGS (Fig. 2). The integral of the NDVI is an accurate estimator of the ANPP in grassland areas (Diallo et al., 1991; Prince, 1991; Tucker et al., 1985; Wylie et al., 1991; Box et al., 1989; Paruelo et al., 1997). We used the inflection points of the ascending and descending portions of the seasonal curve of the NDVI as an estimator of the SGS and EGS (Fischer, 1994) (Fig. 2). The traits MAX, MIN and DM provide additional information on the seasonality of carbon gains.

We generated 13 years of data (1981–93) for the six traits using simple rules on the relation-

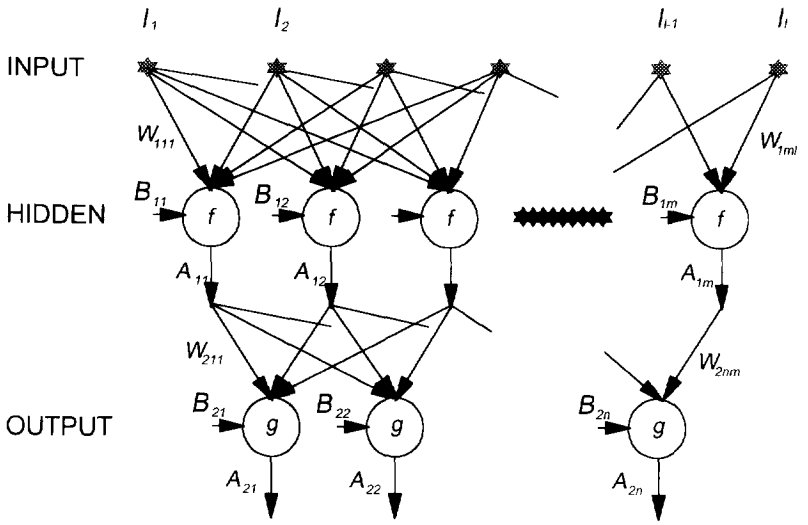

Fig. 1. Schematic representation of the ANN, Appendix A.

ships between each trait and climatic variables. Because we do not know the actual functional form of these relationships ('Nature's laws') we used arbitrary but plausible rules. These rules were based on spatial models of the relationship between the six traits and climate (Paruelo, 1995; Paruelo and Lauenroth, 1995) (Table 1). These models, though quite simple, include several non-linear terms. The spatial models were derived from average values over 3 years of each trait and climate for 43 sites over the central grasslands of North America (Paruelo, 1995; Paruelo and Lauenroth, 1995). To minimize the problems of using a spatial model to make temporal predictions, we selected a site in the central portion of

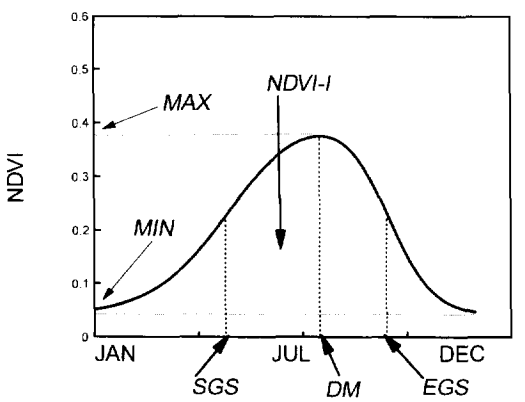

Fig. 2. Example of the seasonal NDVI curve showing the different traits considered: the annual integral of the NDVI curve (NDVI-I), the MAX and MIN NDVI, the date of the MAX NDVI (DM) and the date of SGS and end EGS.

Table 1
Rules used to generate the functional traits used to train the ANN and to fit the RM

NDVI-I = 0.1080–0.240. FALL + 0.000413. MAP − 9.30 − 10^{-6} MAP. MAT
MAX = 0.174–0.454. FALL + 0.00102. MAP − 3.84×10^{-5} MAP. MAT
MIN = −0.111 + 0.0575. SUM + 0.000236. MAP + 0.0105. MAT − 1.218×10^{-5}. MAP MAT
DM = 216 + 107. SUM − 287. WIN + 2.57. MAT − 3.42.AMP
SGS = 62 + 182. SUM − 213. WIN
EGS = 201 + 121. SUM − 319. WIN + 0.00443. MAP

MAP, mean annual precipitation; MAT, mean temperature; AMP, annual thermal amplitude; WIN, proportion of precipitation falling in winter (December, January and Febraury); SUM, proportion of precipitation falling in summer (June, July and August); FALL, proportion of precipitation falling in fall (September, October and November).
ANN, artificial neural network; RM, regression model; NDVI, normalized difference vegetation index.
The rules were applied to climatic data for Hays (KS, USA) for the period 1981–1993. The functional traits generated using the rules were: the annual integral of the NDVI curve (NDVI-I), the MAX and MIN NDVI, the date of the MAX NDVI (DM) and the date of SGS and EGS.

the typical precipitation gradient found in the central grassland region (Hays, Kansas, 38.87°N, 99.38°W). Paruelo et al. (1997) proposed that the difference between temporal and spatial models would be minimal for sites with intermediate levels of precipitation.

For the integral of the NDVI we generated an additional set of data using a rule that incorporated inertial effects. The integral of the NDVI calculated using the rule presented in Table 1 was multiplied by the ratio between the NDVI-I of the previous year and the mean NDVI-I over the whole period.

2.3. Generation of predictive models from the simulated data

2.3.1. ANN training

We trained the ANN using 13 years of monthly precipitation and temperature data as inputs and the values of each of the six traits generated using the rules presented in Table 1 as known outputs. Climatic data were obtained from the EarthInfo, 1993 database. For the integral of the NDVI that include inertial effects we also included in the input layer the precipitation and the temperature of the previous year.

The high power shown by ANNs in fitting highly nonlinear data is menaced by the risk of overfitting, which endangers the predictive power of the network. To reduce the overfitting problem, we carried out an optimization on two parameters of the network: the number of nodes of the hidden layer and the error goal used to stop the training process.

We trained cyclically the network by presenting 12 of the 13 years as training samples, and we took the data of the remaining year as validation sample. For each pair of values of the optimization parameters mentioned above, we calculated a figure of merit that we called the total validation error. This quantity is the sum of the squared errors of the network for the 13 different validation samples.

Using this validation procedure, we observed that the optimum value for the number of nodes in the hidden layer was between 12 and 14. The optimum number of nodes in the hidden layer is a parameter of the ANN that depends on both the size of the data set available for training and the specific problem to be solved.

The optimum error goal also depends on the data set, and it is different for each of the traits considered for training. The total validation error is a function of the error goal, and it usually reaches a MINIMUM value after which it grows rapidly. This growth is indicative of overfitting. If the validation error does not present a MINIMUM, then the network is not complex enough to overfit. That could mean that the number of nodes in the hidden layer is not large enough to perform satisfactorily. To optimize the error goal in our study, we trained the network with decreasing values of this quantity until we reached a MINIMUM on the total validation error.

It should be noticed that the span covered by the training data is also an important issue that can endanger the generalization power of the resulting ANN. As the purpose of the network is to generalize to new cases, the input data for training should be as extensive as possible, while still being of a practical size.

2.3.2. Regression analysis

For the same set of data used to train the ANN we performed a step-wise regression analysis (Kleinbaum and Kupper, 1978) for each of the six traits. The independent variables considered were the same used as the input layer of the neural network: monthly precipitation and temperature. For the data set with inertial effect we considered also the monthly temperature and precipitation of the previous year. Regression analysis were performed in SAS (SAS, 1988). The RM fitted are presented in Appendix B.

2.3.3. Comparison of the ANN and RM

To evaluate the predictive power of both the ANN and the RM we used a independent climatic data set (1971–1980) from the one used to generate the RMs or to train the ANN. For each trait we compared the mean absolute errors (ABS ((predicted − observed)/observed) × 100) of the predictions for the 10 year period. We also looked at the correlation coefficient of the observed and predicted values of the six traits. The diagram in

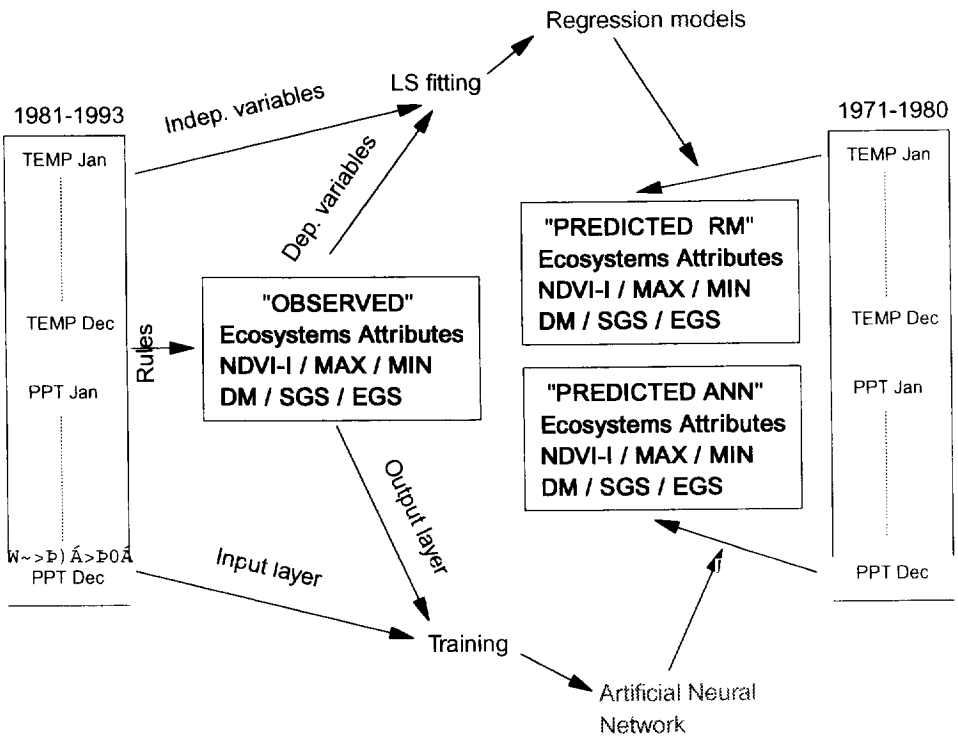

Fig. 3. Scheme showing the procedure to generate the functional traits from the 1981–93 climatic records using the rules presented in Table 1 (observed ecosystem attributes). These attributes were used to fit the RM and to train the ANN. The RM and the ANN were used to predict the ecosystem attributes from the climatic data corresponding to the period 1971–80.

Fig. 3 shows the different steps of the analysis performed.

3. Results

The mean absolute errors of the predictions were in general lower for the ANN than for the RM (Fig. 4). The differences between both methods were more dramatic for the NDVI-I, MAX and MIN variables. Only for DM the mean absolute error of the prediction of ANN was larger than of RM.

For each individual trait the absolute error varied according to the mean square errors (MSE) specified in the training process of the ANN or to the number of variables included in the regression model (Fig. 4). Problems of overfitting readily appeared in the ANN for SGS and EGS. An increase in the precision of the fitting during the training period translated into a lack of predictive power for the resulting ANN.

Overfitting was also evident for the RMs. For the NDVI-I and DM, the mean absolute error of the predictions increased with the number of variables included in the models (from A to C) (Fig. 4). The step-wise regression procedure fitted models with 6–11 variables depending on the trait (Appendix B). These models explained more than 98% of the variance of the training data, however, they have a limited predictive power.

The correlation coefficient between observed (generated using the rules presented in Table 1) and predicted data was always higher for the ANN than for the RM (Fig. 5). ANN predictions showed a better correspondence with observed values than RM even in the case of DM, for which the absolute error was higher for ANN than for RM. For both methods, the correlation coefficients were lower for the variables that characterize the timing of some features of the NDVI curve than for those characterizing the level of NDVI (Fig. 5).

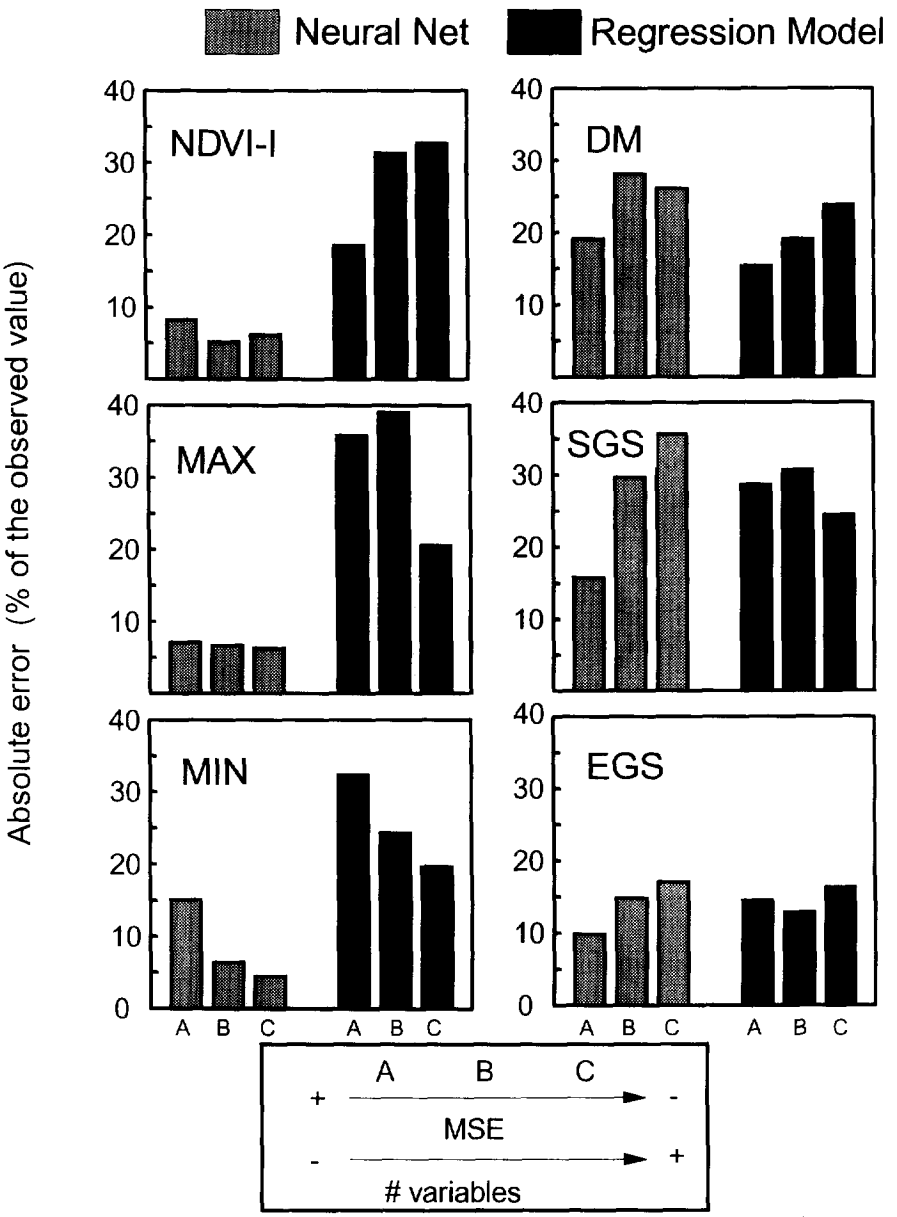

Fig. 4. Absolute error (ABS ((predicted – observed)/observed) × 100) for the six ecosystem attributes analyzed and for predictions derived from RMs and ANNs. Dashed bars correspond to ANNs predictions and closed bars to RMs predictions. Bars A correspond to ANNs trained with a MSE of 0.05 and to RMs including one variable. Bars B correspond to ANNs trained with a MSE of 0.01 and to RMs including two variables. Bars C correspond to ANNs trained with a MSE of 0.005 and to RMs considering all the variables included by the step-wise regresion procedure.

When the NDVI-I incorporated inertial effects, the absolute error of the ANN predictions was similar to the case without the effect of the previous year (Figs. 4 and 6). The predictions based on the RMs showed higher absolute errors than the ANN, and similar to those of the RM for the case of non-inertial effects. The coefficient of correlation of the observed and estimated NDVI-I for

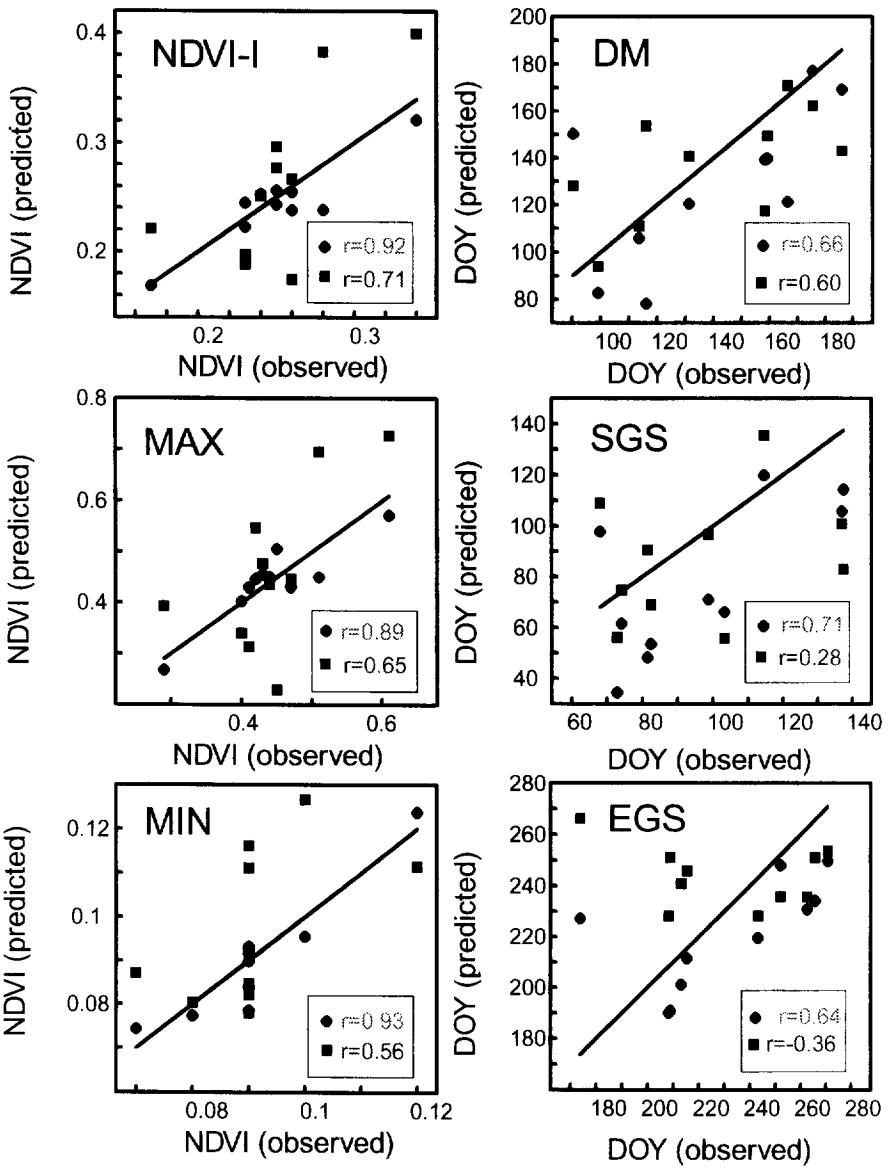

Fig. 5. Comparison of the observed and predicted values of the six ecosystem attributes considered: the annual integral of the NDVI curve (NDVI-I), the MAX and MIN NDVI, the date of the MAX NDVI (DM) and the date of SGS and EGS. Diamonds corresponded to the ANN predictions and squares to the RM predictions. We used for the comparsion the ANN and the RM showing the lowest absolute error in Fig. 4. The line corresponds to the 1:1 line and r is the coefficent of correlation between observed and predicted values.

the best ANN (B) was only 10% lower than in the case of non-inertial effects ($r = 0.82$, Fig. 6, vs. $r = 0.92$, Fig. 5). In the case of the regression model the reduction in the correlation coefficient between observed and predicted NDVI-I was about 30% ($r = 0.50$, Fig. 6, vs. $r = 0.71$, Fig. 5).

4. Discussion and conclusions

To what extent is our ability to describe temporal processes in ecosystems constrained by the analytical tools used? Assuming that the functional form of the relationship between environmental variables and ecosystem attributes is known, the problem of fitting the parameters is relatively easy. If the relationship with the parameters is linear, it reduces to a simple minimization problem. If the functional form depends non-linearly on the parameters (it includes terms that are products, powers or functions of the variables) the problem is slightly more complex and requires an iterative solution. Complexity increases when both the functional form and its dependence on the parameters are unknown. Even simple non-linearities in the actual functional form of the relationship between environmental variables and ecosystem attributes preclude a precise prediction of these attributes when the rules are not explicit. That was evident for predictions based on both ANNs and RMs under absolutely deterministic conditions (error-free scenario). Multiplicative terms in the model are responsible from the non-linearities in the simulated traits of the NDVI curve. Under the presence of these non-linear terms, a different aggregation of the driving variables (monthly vs. annual or quarterly climatic data) substantially reduces the ability of both RMs and ANNs to predict the independent variable.

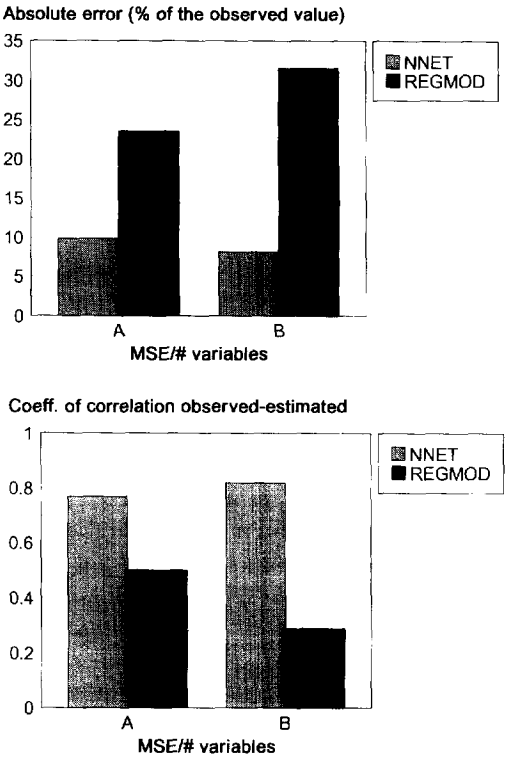

Fig. 6. Absolute error (ABS ((predicted − observed)/observed) × 100) (upper graph) and coefficent of correlation between observed and estimated values (lower graph) of the NDVI intergral (NDVI-I) incorporating inertial effects for predictions derived from RMs (closed bars) and ANNs (dashed bars). Bars A correspond to the ANN trained with a MSE of 0.05 and to the RM including one variable. Bars B corespond to the ANN trained with a MSE of 0.005 and to the RM considering all the variables included by the step-wise procedure.

For the six traits analyzed, the ANNs were able to make better predictions than RMs. The correlation between observed and predicted values of each of the six traits considered was higher for the ANNs than for the RMs (Fig. 4). The better performance shown by ANNs derived mainly from their ability to capture non-linearities. Their intrinsic nonlinear structure makes them particularly suitable as fitting tools. It can be shown that three-layered backpropagation networks can approximate an arbitrary function with the desired degree of accuracy (Hecht-Nielsen, 1991).

The incorporation of non-linear terms in the RMs is of course possible, but it has some practical limits. Including quadratic and logarithmic transformations in the RM of this exercise increases the number of dependent variables from 24 up to 72. Including two-way interactions among the 24 variables means to add $n(n-1)/2$ more variables (276). Prior knowledge on the functional form of the relationship between environmental variable or the level of aggregation of the independent variables poses serious constraints on the possibility of describing these relationship using RMs.

ANNs provide a 'black-box' approach to the description of the relationship between two sets of variables. Even an ANN able to make perfect

predictions would tell us nothing about the functional form of the relationship between the input variables and the output layer. The bias vectors or the weight matrices of the network do not have any direct biological meaning. Numerical experimentation with the ANNs may help, however, to identify the most important input variables and the functional form of the relationship between them and the outputs. ANNs are a useful tool two understand dynamical systems. The behavior of these systems is complicated both in space and time. These problems are particularly hard because they are both large and non-linear. Ecological problems often have these characteristics.

Overfitting is one of the main concerns regarding the use of ANNs. For our example, overfitting problems were similar for both ANNs and RMs. Both the absolute error and the ability to predict (correlation coefficient) changed according to the MSE specified during the training (in the case of the ANN) or according to the number of variables considered (in the case of the RM). The optimum number of variables included in the regression model or the MSE used to train the ANN depended on the trait considered.

ANNs showed clear advantages to capture inertial effects. The existence of historical effects is common in ecological attributes (Pennington, 1986; Malanson et al., 1992; Cole, 1985). In arid and semiarid areas community structure may buffer changes associated to climatic fluctuations (Lauenroth and Sala, 1992). The response of the ecosystems to the current level of resources can be influenced by the availability of resource on previous years (Milchunas and Lauenroth, 1995). For the example analyzed here, the ANN was able to use the previous year's information on climate to estimate current year NDVI-I much better than a RM that use the same input information. The correlation between observed and predicted values of the NDVI integral was only 10% lower in the presence of inertial effects than in the case without inertial effects.

Acknowledgements

This work was funded by the LTER-NSF grant BSR 90-11659. J.M. Paruelo received support from the Consejo Nacional de Investigaciones Cientificas y Tecnicas (CONICET-Argentina) and the University of Buenos Aires (Argentina). Rick Gill provided useful comments on the manuscript.

Appendix A

The basic structure of the ANN used in this paper is shown in Fig. 1. When presented to an input vector $\boldsymbol{I}$, the hidden neuron i receives an input equal to

$$\boldsymbol{H}_{1i} = \boldsymbol{B}_{1i} + \mathbf{W}_{1ik}\boldsymbol{I}_k,$$

where $\boldsymbol{B}_1$ and $\mathbf{W}_1$ are the bias vector and weight matrix of the first neuron layer, respectively, and repeated dummy indices imply a summation over those indices. The neuron i then produces an output

$$\boldsymbol{A}_{1i} = f(H_{1i}) = f(\boldsymbol{B}_{1i} + \mathbf{W}_{1ik}\boldsymbol{I}_k),$$

where f is the transfer function of the neuron. The neuron j at the output layer receives an input

$$\boldsymbol{H}_{2j} = \boldsymbol{B}_{2j} + \mathbf{W}_{2ji}\boldsymbol{A}_{1i} = \boldsymbol{B}_{2j} + \mathbf{W}_{2ji} f(\boldsymbol{B}_{1i} + \mathbf{W}_{1ik}\boldsymbol{I}_k),$$

producing a final output of

$$\begin{aligned}\boldsymbol{A}_{2j} &= g(\boldsymbol{H}_{2j}) = g(\boldsymbol{B}_{2j} + \mathbf{W}_{2ji}\boldsymbol{A}_{1i}) \\ &= g[\boldsymbol{B}_{2j} + \mathbf{W}_{2ji} f(\boldsymbol{B}_{1i} + \mathbf{W}_{1ik}\boldsymbol{I}_k)],\end{aligned}$$

where $\boldsymbol{B}_2$, $\mathbf{W}_2$ and g are the bias vector, weight matrix and transfer function of the output layer, respectively. In order to train the network, the transfer functions must be differentiable. Both the choice and shape of the activation function strongly affect the speed of the network learning and its generalization capabilities. For our calculations we have used a logistic sigmoid as the activation function for the hidden layer, which is given by:

$$f(x) = 1/(1 + e^{-x}).$$

The network is trained using the gradient descent technique to minimize the sum of the square errors of the output values

$$SSE = \sum_{i=1}^{N} (O_i - \bar{O}_i)^2$$

where O_i are the observed values, O_i are the values computed by the ANN, and N is the number of cases available for training. The learning algorithm of backpropagation networks uses the derivative of the transfer function as a multiplicative factor for the adjustment of the weights (Hertz et al., 1991). As the derivative of the sigmoidal functions peaks at zero and vanishes for large absolute values of the argument, the weights of neurons with uncertain responses are more strongly corrected than those of neurons which are turned on or off.

Appendix B

RMs fitted using the step-wise procedure for the six functional attributes analyzed and for the period 1981–93. The independent variables included for each year were monthly precipitation and temperature. The dependent variables (ecosystem attributes) were calculated using the rules presented in Table 1. The models corresponding to one, two, and all the variables selected by the method, are presented.

NDVI-I

	Coefficient	F	$P>F$
One variable ($r^2 = 0.65$)			
INTERCEP	0.16108722	44.82	0.0001
PPT-August	0.00134924	20.21	0.0009
Two variables ($r^2 = 0.76$)			
INTERCEP	0.51999316	9.4	0.0119
PPT-August	0.00125631	22.5	0.0008
TEM-September	−0.01754499	4.5	0.0587
All variables ($r^2 = 0.99$)			
INTERCEP	0.02763296	152.8	0.0065
PPT-January	0.00073670	244.8	0.0041
PPT-March	0.00047173	1327.9	0.0008
PPT-April	0.00009613	25.9	0.0365
PPT-May	0.00037383	391.2	0.0025
PPT-July	0.00044819	843.5	0.0012
PPT-August	0.00074108	1054.5	0.0009
PPT-October	0.00106020	2254.2	0.0004
PPT-November	0.00020286	44.0	0.0219
TEM-January	−0.00191243	68.7	0.0142
TEM-Febraury	−0.00091627	6.2	0.1301

MAX

	Coefficient	F	$P>F$
One variable ($r^2 = 0.63$)			
INTERCEP	0.25939467	28.1	0.0003
PPT-August	0.00266121	19.0	0.0011
Two variables ($r^2 = 0.77$)			
INTERCEP	1.05408810	10.2	0.0095
PPT-August	0.00245545	22.7	0.0008
TEM-October	−0.0388483	5.9	0.0355
All variable ($r^2 = 0.99$)			
INTERCEP	0.00289053	0.24	0.6473
PPT-January	0.00141142	138.7	0.0003
PPT-March	0.00088634	796.0	0.0001
PPT-May	0.00078959	297.5	0.0001
PPT-July	0.00094964	1057.4	0.0001
PPT-August	0.00144167	926.0	0.0001
PPT-October	0.00206903	1206.1	0.0001
TEM-January	−0.00691907	89.4	0.0007
TEM-December	−0.00332483	36.5	0.0038

MIN

	Coefficient	F	$P>F$
One variable ($r^2 = 0.75$)			
INTERCEP	0.06652437	117.9	0.0001
PPT-August	0.00043496	32.3	0.0001
Two variables ($r^2 = 0.92$)			
INTERCEP	0.06078804	262.4	0.0001
PPT-July	0.00014137	22.6	0.0008
PPT-August	0.00031927	39.8	0.0001
All variables ($r^2 = 0.99$)			
INTERCEP	0.05367334	127.5	0.0015
PPT-April	0.00003149	62.2	0.0042
PPT-June	0.00013422	1529.6	0.0001
PPT-July	0.00013058	4525.9	0.0001
PPT-August	0.00026429	4521.3	0.0001
TEM-April	0.00167548	436.4	0.0002
TEM-June	−0.00021012	8.4	0.0626
TEM-August	0.00113402	91.8	0.0024
TEM-October	−0.00351703	226.8	0.0006
TEM-December	−0.00043771	87.19	0.0026

DOYMAX

	Coefficient	F	$P>F$
One variable ($r^2 = 0.64$)			
INTERCEP	170.85180226	784.33	0.0001
TEM-January	7.69378938	19.92	0.0010
Two variable ($r^2 = 0.86$)			
INTERCEP	178.52491153	1548.66	0.0001
TEM-January	9.24654198	57.48	0.0001
TEM-December	4.45031447	14.70	0.0033
All variables ($r^2 = 0.99$)			
INTERCEP	62.22683741	12229.4	0.0058
PPT-August	0.12547250	56156.9	0.0027
PPT-November	−0.30590180	125441	0.0020
TEM-January	7.18357027	525347	0.0020
TEM-Febraury	0.91560916	15229.0	0.0052
TEM-March	−0.79504176	2541.1	0.0126
TEM-April	4.33868142	96548.2	0.0020
TEM-May	−0.43426043	1368.5	0.0172
TEM-July	−2.22517341	5696.5	0.0084
TEM-August	−0.20962879	55.1	0.0852
TEM-October	10.65572552	181151	0.0020
TEM-December	3.78174085	263310	0.0020

SGS

	Coefficient	F	$P>F$
One variable ($r^2 = 0.45$)			
INTERCEP	313.12314559	20.2	0.0009
TEM-August	−8.36301708	9.0	0.0119
Two variables ($r^2 = 0.58$)			
INTERCEP	333.71816484	26.77	0.0004
PPT-November	−0.33819664	3.20	0.1037
TEM-August	−8.79689857	11.92	0.0062
All variables ($r^2 = 0.99$)			
INTERCEP	193.20511601	1896.7	0.0146

PPT-January	−0.39376188	2809.1	0.0120
PPT-March	−0.02991613	158.1	0.0505
PPT-July	0.18179736	4787.3	0.0092
PPT-October	−0.16572134	733.0	0.0235
PPT-November	−0.51597865	6477.7	0.0079
TEM-January	2.83194904	1336.2	0.0174
TEM-Febraury	1.66451740	940.0	0.0208
TEM-May	0.61412766	34.170	0.1079
TEM-August	−4.58509910	1734.5	0.0153
TEM-October	3.15208142	313.6	0.0359
TEM-November	−1.00044555	150.8	0.0517

EGS

	Coefficient	F	$P>F$
One variable ($r^2 = 0.45$)			
INTERCEP	469.71682089	37.1	0.0001
TEM-August	−9.15780142	8.8	0.0126
Two variables ($r^2 = 0.67$)			
INTERCEP	491.10520008	60.3	0.0001
PPT-January	−0.55131453	6.6	0.0277
TEM-August	−9.39111922	14.04	0.0038
All variables ($r^2 = 0.98$)			
INTERCEP	635.81598123	244.8	0.0001
PPT-January	−0.99393678	140.0	0.0001
PPT-August	0.21237644	25.0	0.0024
PPT-October	−0.23090247	12.3	0.0125
PPT-November	−0.54986285	56.0	0.0003
TEM-May	−6.14060120	28.5	0.0018
TEM-October	−10.05940298	109.1	0.0001

References

Box, E.O., Holben, B.N. and Kalb, V., 1989. Accuracy of the AVHRR Vegetation Index as a predictor of biomass, primary productivity and net CO_2 flux. Vegetation, 80: 71–89.

Burke, I.C., Yonker, C.M., Parton, W.J., Cole, C.V., Flach, K. and Schimel, D.S., 1989. Texture, climate, and cultivation effects on soil organic content in US grassland soils. Soil Sci. Soc. Am. J., 53: 800–805.

Cole, K., 1985. Past rates of changes, species richness, and a model of vegetational inertia in the Grand Canyon, Arizona. Am. Nat., 125: 289–303.

Currie, D.J. and Fritz, J.T., 1993. Global patterns of animal abundance and species energy use. Oikos, 67: 56–68.

Derr, V.E. and Slutz, R.J., 1994. Prediction of El Nino events in the Pacific by means of neural networks. AI Appl., 8: 51–63.

Diallo, O., Diouf, A., Hanan, N.P., Ndiaye, A. and Prevost, Y., 1991. AVHRR monitoring of savanna primary production in Senegal, West Africa: 1987–1988. Int. J. Remote Sensing, 12: 1259–1279.

EarthInfo, 1993. NCDC Summary of the Day. Boulder, CO.

Elizondo, D., Hoogenboom, G. and McClendon, R.W., 1994. Development of a neural network model to predict daily solar radiation. Agric. For. Meteorol., 71: 115–132.

Fischer, A., 1994. A model for the seasonal variations of vegetation indices in coarse resolution data and its inversion to extract crop parameters. Remote Sensing Environ., 48: 220–230.

Foody, G.M., McCulloch, M.B. and Yates, W.B., 1995. Classification of remotely sensed data by an artificial neural network: issues related to training data characteristics. Photogramm. Eng. Remote Sensing, 61: 391–401.

Gimblett, R.H. and Ball, G.L., 1995. Neural Network architectures for monitoring and simulating changes in forest resource management. AI Appl., 9: 103–123.

Hecht-Nielsen, R., 1991. Neurocomputing. Addison-Wesley, Reading, MA.

Hertz, J., Krogh, A. and Palmer, R.G., 1991. Introduction to the Theory of Neural Computation. Addison-Wesley, Reading, MA.

Kanellopoulos, I., Varfis, A., Wilkinson, G.G. and Megier, J., 1992. Land-cover discrimination in SPOT HRV imagery using an artificial neural network-a 20-class experiment. Int. J. Remote Sensing, 13: 917–924.

Kleinbaum, D.G. and Kupper, L.L., 1978. Applied Regression Analysis and other Multivariate Methods. Duxbury Press, North Scituate, MA.

Lauenroth, W.K. and Sala, O.E., 1992. Long-Term forage production of North American shortgrass steppe. Ecol. Appl., 2: 397–403.

Le Houerou, H.N., Bingham, R.L. and Skerbek, W., 1988. Relationship between the variability of primary production and the variability of annual precipitation in world arid lands. J. Arid Environ., 15: 1–18.

Liu, Z.K. and Xiao, J.Y., 1991. Classification of remotely sensed image data using artificial neural networks. Int. J. Remote Sensing, 12: 2433–2438.

Lloyd, D., 1990. A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery. Int. J. Remote Sensing, 11: 2269–2279.

Malanson, G.P., Westman, W.E. and Yan, Y.L., 1992. Realized versus fundamental niche functions in a model of chaparral response to climatic change. Ecol. Model., 64: 261–277.

McCann, D.W., 1992. A neural network short-term forecast of significant thunderstorms. Weather Forecasting, 7: 525–534.

McNaughton, S.J., Oesterheld, M., Frank, D.A. and Williams, K.J., 1989. Ecosystem-level patterns of primary productivity and herbivory in terrestrial habitats. Nature, 341: 142–144.

Milchunas, D.G. and Lauenroth, W.K., 1993. Quantitative effects of grazing on vegetation and soils over a global range of environments. Ecol. Monogr., 63: 327–366.

Milchunas, D.G. and Lauenroth, W.K., 1995. Inertia in plant community structure: State changes after cessation of nutrient enrichment stress. Ecol. Appl., 5: 1995.

Oesterheld, M., Sala, O.E. and McNaughton, S.J., 1992. Effect of animal husbandry on herbivore-carrying capacity at a regional scale. Nature, 356: 234–236.

Paruelo, J.M., 1995. Regional patterns and climatic controls of the structure and function of North American graslands and shrublands. Ph.D. Dissertation. Colorado State University, Fort Collins, CO.

Paruelo, J.M., Epstein, H.E., Lauenroth, W.K. and Burke, I.C., 1997. ANPP estimates from NDVI for the Central Grassland region of the US. Ecol. in press.

Paruelo, J.M. and Lauenroth, W.K., 1995. Regional patterns of NDVI in North American shrublands and grasslands. Ecology, 76: 1888–1898.

Paruelo, J.M. and Lauenroth, W.K., 1996. Climatic controls of the distribution of functional types in Grassland and shrubland of North America. Ecol. Appl., 6: 1212–1224.

Pennington, W., 1986. Lags in adjustment of vegetation to climate caused by the pace of soil development: evidence from Britain. Vegetation, 67: 105–118.

Prince, S.D., 1991. Satellite remote sensing of primary production: comparison of results for Sahelian grasslands 1981–1988. Int. J. Remote Sensing, 12: 1301–1311.

Running, S.W., Loveland, T.R. and Pierce, L.L., 1994. A vegetation classification logic based on remote sensing for use in global biogeochemical models. Ambio, 23: 77–81.

Sala, O.E., Parton, W.J., Joyce, L.A. and Lauenroth, W.K., 1988. Primary Production of the central grassland region of the United States. Ecology, 69: 40–45.

SAS, 1988. SAS/STAT User's guide. Release 6.03 Edition. SAS Institute, Cary, NC.

Sellers, P.J., Berry, J.A., Collatz, G.J., Field, C.B. and Hall, F.G., 1992. Canopy reflectance, photosynthesis, and transpiration. III. A reanalysis using improved leaf models and a new canopy integration scheme. Remote Sensing Environ., 42: 187–216.

Soriano, A. and Paruelo, J.M., 1992. Biozones: vegetation units defined by functional characters identifiable with the aid of satellite sensor images. Global Ecol. Biogeogr. Lett., 2: 82–89.

Teeri, J.A. and Stowe, L.G., 1976. Climatic patterns and the distribution of C4 grasses in North America. Oecologia, 23: 1–12.

Tucker, C.J., Vanpraet, C.V., Sharman, M.J. and Vanlttersum, G., 1985. Satellite remote sensing total herbaceous biomass production in the Senegalese Sahel: 1980–1984. Remote Sensing Environ., 17: 233–249.

Wasserman, P.D., 1993. Advanced Methods in Neural Computing. Van Nostrand Reinhold, NY.

Wylie, B.K., Harrington, J.A., Prince, S.D. and Denda, I. 1991. Satellite and ground-based pasture production assessment in Niger: 1986–1988. Int. J. Remote Sensing, 12: 1281–1300.