A Layered Digital Dynamic Network for Prediction Purposes

SALAZAR-MORENO, Raquel; LÓPEZ-CRUZ, Irineo; FITZ-RODRÍGUEZ, Efren y ROJANO-AGUILA, Abraham

R. Salazar, I. López, E. Fitz y A. Rojano
1Universidad Autónoma Chapingo, Km 38.5 Carr. México-Texcoco Chapingo, Edo. México, C.P 56230. México.
raquels60@hotmail.com

D. Sepúlveda, F. Rérez, D. Sepúlveda, E. Figueroa, R. Salazar, L. Godínez (eds.) Matemáticas Aplicadas. Handbook T-I. -©ECORFAN, Texcoco de Mora-México, 2016.

Abstract

Producers not only harvest products but useful information to develop optimization models to increase efficiency in crop management. The problem we face is the prediction of tomato yields in two greenhouses located at Humboldt University of Berlin, one is a conventional high technology greenhouse and the other is a semiclosed greenhouse, both have the same area and equipment. Yield was recorded for 26 weeks in both greenhouses. Because of the small data pattern, a dynamic artificial neural networks (DANN) was implemented, which is generally more adequate than static networks, because they learn sequential or time varying patterns. In particular, the Layered Digital Dynamic Network (LDDN) was used where solar radiation, transpiration and CO_2 fixation were the input variables for predicting yield. The best LDDN model was chosen on the basis of a suitable architecture (i.e. a minimum number of input neurons connected to a hidden layer) combined with a high R, a low mean absolute error (MAE). The model explained the weekly fluctuations of tomato yields based on external conditions and the stage of the plant. The sensitivity analysis show that the most influencing variable in prediction of tomato yields was CO_2 enrichment. The results from the model, could be a valuable information for making decisions on climate and crop management and in synchronizing crop production with market demands.

12 Introduction

The advances and sofistication of the control systems in agriculture has generated information useful for producers. Most of the variables could be correlated directly or indirectly to yield so regression or correlation analysis cannot be used. According to Ehret et al.(2011) the fisiological information can provide complementary data specially if this data is integrated in computacional models. An efective way for modelling yields of different crops has been the Artificial Neural Networks.

The non-linear modeling approach based on the Artificial Intelligence (AI) techniques has received considerable attention from the hydrologists in the last two decades (Boucher et al., 2010; de Vos and Rientjes, 2005; Toth et al., 2000; Weigend et al., 1995; Xiong et al., 2004, El-Shafie et al. , 2012), as well as in agriculture. Topuz (2010) applied Artificial Neural Networks (ANN) in predicting moisture content of agricultural products. The author mentioned that the recent advances in computer technology and parallel processing have made the use of ANN more economically feasible; ANN is composed of nets of non-linear basis functions, it has the ability to evolve good process models from example data and require little or no a priori knowledge of the task to be performed; ANN has the potential to solve certain types of complex problems that have not been satisfactorily handled by more traditional methods. Also, ANN have been effective in modelling yield of different crops (Masters, 1993).

The design of a neural network and selection of the proper dataset for a given problem, it is not an easy task. First, a proper selection of the input data to explain the phenomenum at hand and the training dataset is crucial and leads to the success of the neural network prediction. Larger networks require large training datasets. However, some problems arise like overfitting. If we want an effective neural network for prediction purposes, the training dataset must be complete enough in such a way that every group must be represented, for the particular case of yield prediction, each group represents the stage of the productive time for tomato crop . Each class is characterized by a statistical variation, so the data presented to a neural network must be the entire range of data with noise included.

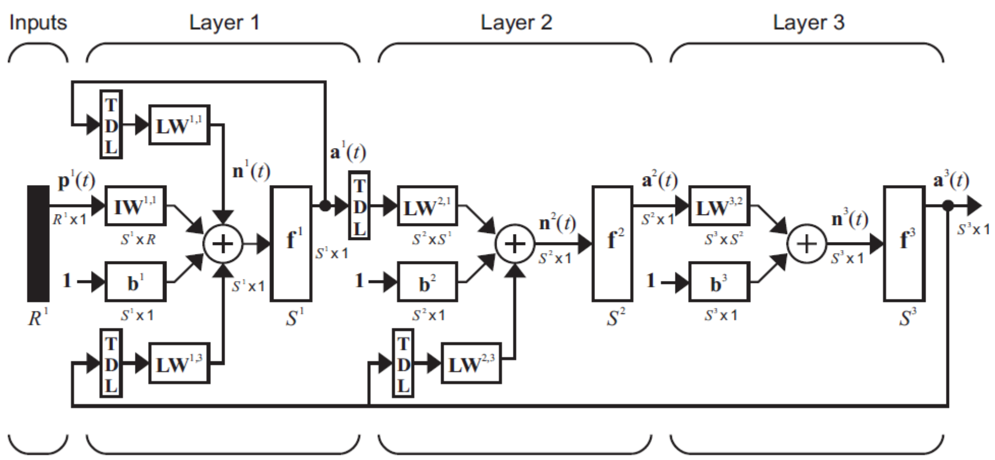
Another important issue is if the data set need any transformations. Besides, the neural network architecture, number of layers , number of hidden nodes, the transfer functions and the training time are some of the factors to be considered in any Neural Network Model.

Transfer functions are very important because they contain adaptive parameters that are optimized, commonly they have limits between -1 to +1. The most widespread transfer functions are the sigmoid, the hyperbolic tangent and pure linear functions. The first two are continuous, non linear function whose domain is the real number set, whose first derivative is always positive, and whose range is bounded. The sigmoid function never reaches its theoretical maximum and minimum; however the hyperbolic tangent function is an ideal transfer function Bardina and Rajkumar (1993).

The process of defining an appropriate neural network architecture can be divided into the following categories: (i) determining the type of neural network; (ii) determining the number of hidden neurons; (iii) selecting the type of transfer functions; (iv) devising a training algorithm; and (v) checking for over and/or under fitting of the results and validation of neural network output. If a function consists of a finite number of points, a three layer neural network is capable of learning that function. This results agree with Bardina and Rajkumar (1993) who conclude that a three layer neural network with a Levenberg-Marquardt training algorithm using pure linear, hyperbolic tangent, and sigmoid as a transfer functions was sufficient for prediction of aerodynamic coefficients.

The type of Artificial Neural Network (ANN) could be static or dynamic. Dynamic networks are generally more powerful than static networks (although somewhat more difficult to train). Because dynamic networks have memory, they can be trained to learn sequential or time-varying patterns. One principal application of dynamic neural networks is in control systems. In order to predict temporal patterns, an ANN requires two distinct components: a memory and an associator. The memory is generated by a time delay unit (or shift register) that constitutes the tapped delay line, it holds the relevant past information, and uses the memory to predict future events. The associator can be a static Multilayer Perceptron Neural Network (MLPNN) is a memoryless network that is effective for complex non-linear static mapping El-Shafie et al. (2012). Figure 1 display the main structure of the dynamic neural network called Layered Digital Dynamic Network (LDDN), which is available in the software Matlab.

Figure 12 Layered Digital Dynamic Network (LDDN)



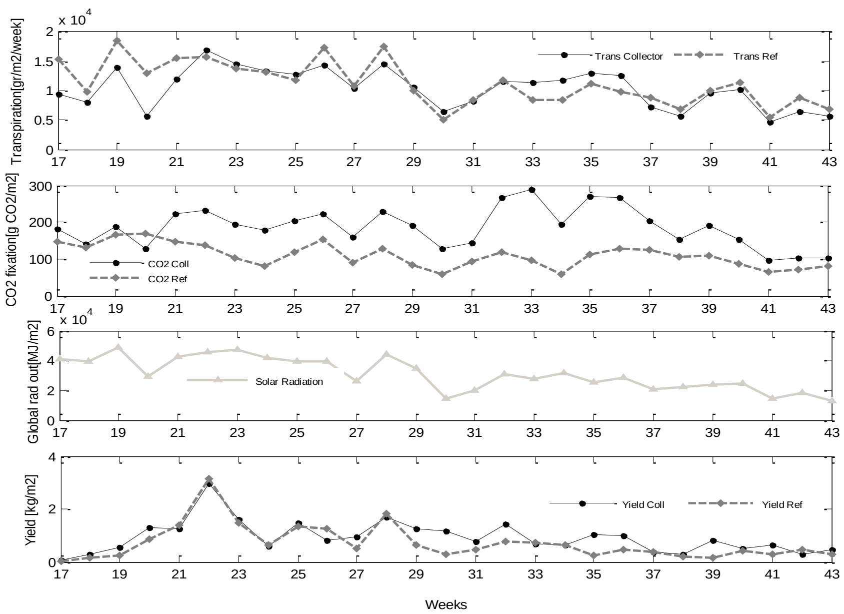
TDL=tapped delay line; LW= weights in the hidden layers; IW= weights in the input layer, b= bias unit; f= transfer functions between layers

In this work we dealt with the problem of predicting tomato yields, which a problematic area in greenhouse production. The tomato crop data has nonlinear dynamic behaviour and whose response depends not only on several environmental factors but also on the current and previous crop conditions Qaddoum et al. (2013). Therefore, our purpose is to develop a Layered Digital Dynamic Network (LDDN), to explain the weekly fluctuations of greenhouse tomato yields, for two greenhouses one is a high technology conventional greenhouse and the other is a semiclosed greenhouse.

12.1 Methodology

An experiment was done to measure tomato yields in a semiclosed and reference greenhouses, located at Humboldt University of Berlin. The collection yield data is expensive and time consuming, it requires the destruction of vegetative material, so the model development will save money and give additional advantages. The first yield of Pannovy cultivar was achieved in calendar week 17, the last one in calendar week 43 (Dannehl, 2012). First there was a selection of the important input variables to predict yields: Cumulative transpiration, CO_2 fixation and global radiation, the behaviour of such variables are displayed in Figure 2, together with yields for the two greenhouses.

Figure 12.1 Cumulative transpiration, CO_2 fixation, global radiation and yield for the collector and the reference greenhouses



As can be seen in Figure 2, there were only 27 data patterns for training , validation, testing and simulation of the ANN. Many architectures were tested for the Multilayer Perceptrum Artificial Neural Network (MLPNN), but none of them have acceptable performance. Therefore, a dynamic neural network was the best choice. The Layered Digital Neural Network (Figure.1), includes delay lines between the layers. So the output depends also on previous inputs and/or previous states of the network. The dynamic neural networks in which all layers have feedback connections with several time delays mean that the temporal feature could be considered in the model structure.

Each layer in the LDDN have a set of weight matrices that come into that layer (which can connected from other layers or from external inputs), associated weight function rule used to combine the weight matrix with its input (normally standard matrix multiplication, dotprod), and associated tapped delay line. The weights have two different effects on the network output. The first is the direct effect, because a change in the weight causes an immediate change in the output at the current time step (This first effect can be computed using standard backpropagation). The second is an indirect effect, because some of the inputs to the layer, such as a (t −1), are also functions of the weights. To account for this indirect effect, the dynamic backpropagation is used to compute the gradients, which is more computationally intensive, and training is more likely to be trapped in local minima. This suggests that there is a need to train the network several times to achieve an optimal result. The Bias vector is a net input function rule that is used to combine the outputs of the various weight functions with the bias to produce the net input (normally a summing junction, netprod).

Still there are some questions to be answered like number of hidden layers and hidden nodes and the transfer functions. It is unclear how many layers and how many neurons in each layer should be used , usually are chosen empirically (Bardina and Rajkumar, 1993). The additional hidden layers through which errors must be backpropagated makes the gradient more unstable, and the number of false minima increases. Besides, overfitting can arise when training sets are small relative to the number of hidden neurons, the training set size and the hidden layer size are tied together (Moustafa, 2011). Although dynamic networks can be trained using the same gradient-based algorithms that are used for static networks, the performance of the algorithms on dynamic networks can be quite different, and the gradient must be computed in a more complex way. An important issue in the ANN is how to measure the performance and according to Masters (1993), the Mean Square Error (MSE) fails to distinguish between minor and serious errors. Therefore, other two statistical measures were used to examine the goodness of fit, the correlation coefficient "R" and the mean absolute error defined in Marzban (2009)

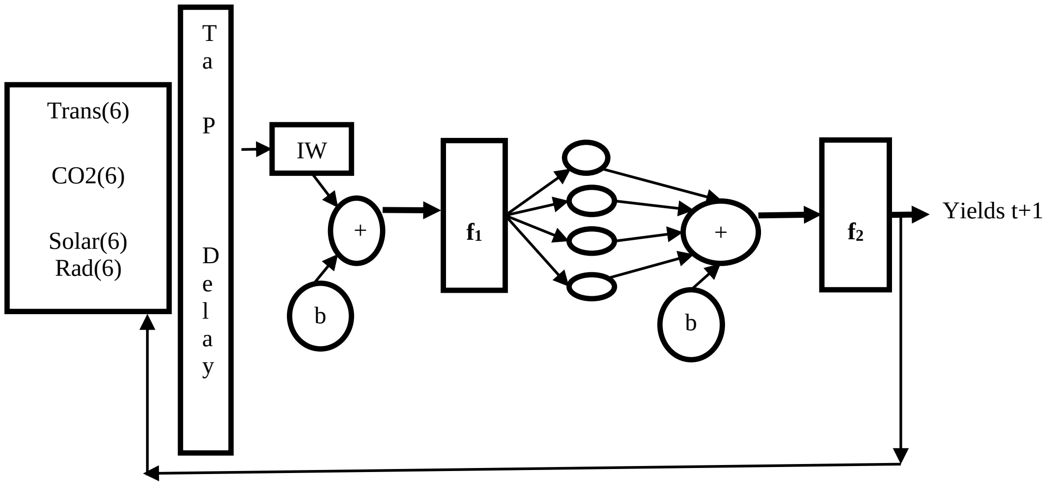
$$MAE = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i| \tag{12}$$

A sensitivity analysis was performed varying individual inputs while all other are fixed to find out what was the most important variable for yield prediction.

12.2 Results

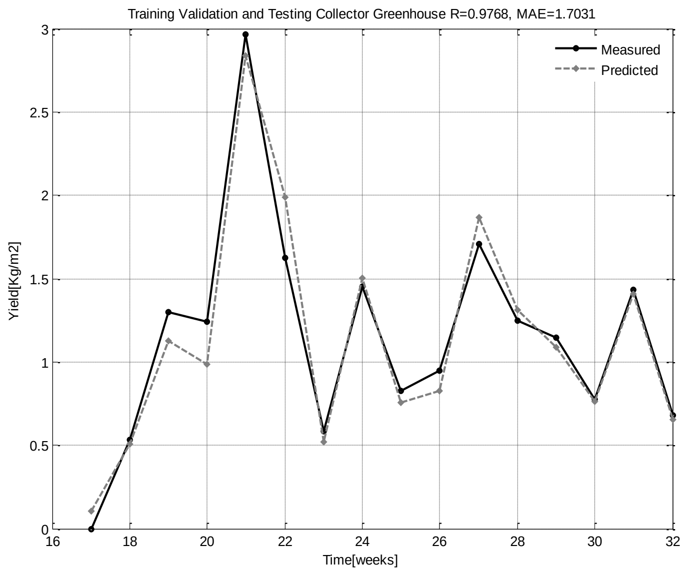
For the input and output variables two delays were included, so we lost two data patterns. The weeks from 17-32 were used for training validation and testing the LDNN, weeks 33-40 were used for simulation purposes. We test different architectures and the best LDNN model was chosen on the basis of minimum number of hidden layer and hidden nodes with a high R, and a low absolute error. The best LDDN has one hidden layer with four nodes, six delays for the input and four delays for the output variables, as shown in Figure 3.

Figure 12.2 Dynamic Neural Network Architecture



The transfer function for the input and hidden layer was the sigmoid tangent function and for the output the purelin function, the training function used the backpropagation algorithm. Figure 4 display the results for training, validation and testing of the LDNN for the collector greenhouse, with R=0.9768 and MAE=1.7031.

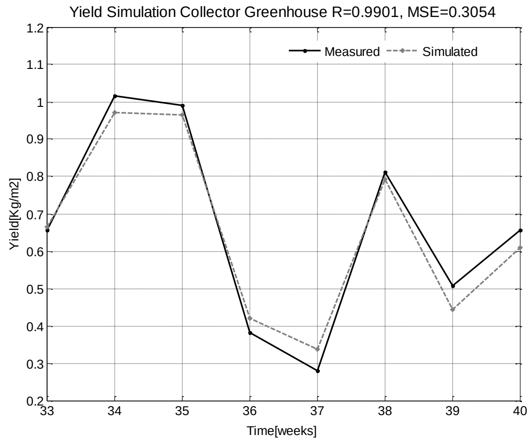
Figure 12.3 Training, validation and testing of the Dynamic Neural Network Collecto Greenhouse with all variables



The Layered Digital Neural Network was used for simulation purposes for the weeks 33-40, the table and graph are shown in Figure 5, this results show the power of the LDNN, the simulation was made using the inputs without providing the outputs, only the delay in yields with an R=0.9901, MAE=0.3054.

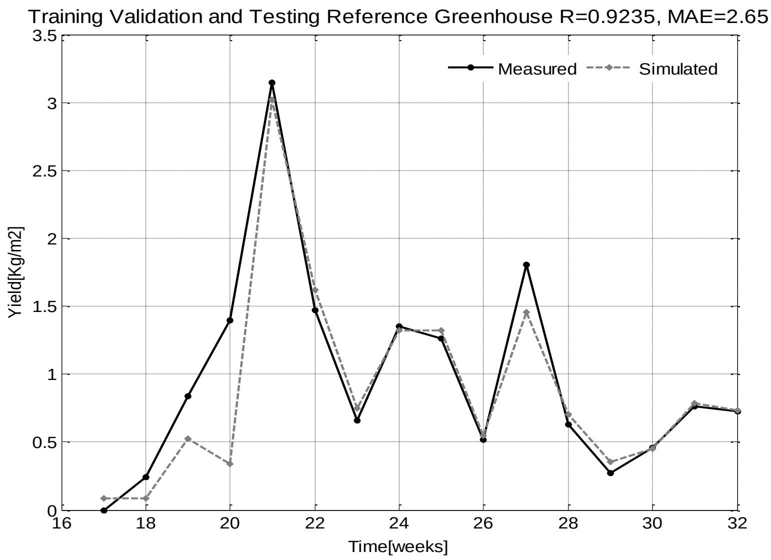
Figure 12.4 Yield simulation for the collector greenhouse

Yields Kg/m^2		
Week	estimated	predicted
33	0.66	0.66
34	1.02	0.97
35	0.99	0.97
36	0.38	0.42
37	0.28	0.34
38	0.81	0.79
39	0.51	0.44
40	0.66	0.61



The results obtained for the reference greenhouse which is a conventional one , were not as good as for the collector greenhouse, the R=0.9235 and MAE=2.65 (Figure 6) our assumption is that the collector greenhouse has less disturbances from the outside environment than the reference greenhouse.

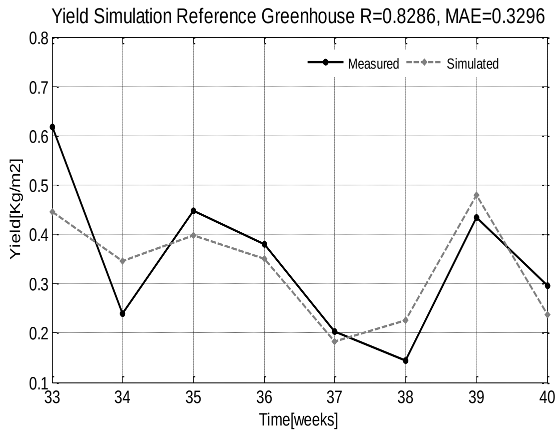
Figure 12.5 Training validation and testing and simulation of the Dynamic Neural Network Reference greenhouse with all variables



For this case, the simulation is shown in Figure 7, with an R=0.8286 and MAE=0.3296.

Figure 12.7 Yield simulation for the reference greenhouse

Yields Kg/m^2		
Week	Measured	Predicted
33	0.62	0.45
34	0.24	0.35
35	0.45	0.40
36	0.38	0.35
37	0.20	0.18
38	0.14	0.23
39	0.43	0.48
40	0.30	0.24



A sensitivity analysis was performed to detect what is the most important variable for yield prediction, results are display in Table 1.

Table 12 Sensitivity analysis of the input variables for the collector greenhouse (simulation)

	R	**Absolute error**
All variables	0.9901	0.0697
Without Transpiration	0.9348	0.0836
Without CO_2 enrichment	0.7382	0.1341
Without Rad	0.8905	0.2252

According to this result, CO_2 has a prominent importance in yield prediction, when this variable is not considered in the LDNN, the correlation coefficient is small (0.7382), and MAE increase to 0.1341.

12.3 Conclusions

Choosing the proper type of neural network for a certain problem can be a critical issue. So it is very important to define the following in order to achieve a good model: The input sets, the target sets, the network architecture, the activation functions, the training function, the training rate, the goal and the number of iterations.

For the data at hand, we suggest a certain network structure that yielded optimal results for this particular case , Layered Digital Dynamic Network (LDDN) improves by far the standard MLP forecast accuracy. This work shows the power of the dynamic NN compared to static, specially when there are few data patterns. The Artificial Neural Networks predictions are more precise in closed greenhouse systems, because there are less disturbances from the outside environment as in open greenhouses, besides the lost of humidity from ventilation closes the stomata and photosynthesis decreases. The results from the model could be a valuable information for making decisions on climate and crop management and in synchronizing crop production with market demands.

12.4 References

Bardina, J. and Rajkumar, T. (1993). Training data requirement for a neural network to predict aerodynamic coefficients. Proc SPIE—Int Soc Opt Eng, 5102, pp. 92–103.

Boucher, M.-A., Lalibert´e, J.-P., and Anctil, F. (2010). An experiment on the evolution of an ensemble of neural networks for streamflow forecasting. Hydrol. Earth Syst. Sci., 14, 603–612, doi:10.5194/hess-14-603-2010.

de Vos, N. J. and Rientjes, T. H. M. (2005). Constraints of artificial neural networks for rainfall runoff modelling: trade-offs in hydrological state representation and model evaluation. Hydrol. Earth Syst. Sci., 9, 111–126, doi:10.5194/hess-9-111-2005.

Dannehl, D., Rocksch,T., Schimidt, W. (2014). Modelling to estimate the specific leaf area of tomato leaves (cv. Pannovy). Acta Hort: in press.

Ehret , D,L., Hill, B.D., Helmer, T., Edwards, D.R. (2011). Neural network modeling of greenhouse tomato yield, growth and water use from automated crop monitoring data.Computers and Electronics in Agriculture 79, p 82–89.

El-Shafie, A., Noureldin, A., Taha, M., Hussain, A. and Mukhlisin, M. (2012). Dynamic versus static neural network model for rainfall forecasting at Klang River Basin. Malaysia Hydrol. Earth Syst. Sci., 16, 1151–1169. www.hydrol-earth-syst-sci.net/16/1151/2012/ doi:10.5194/hess-16-1151-2012.

Masters, Timothy. (1993). Practical Neural Network Recipes in C++ . Academic Press Inc. London.

Moustafa A.A. (2011). Performance Evaluation of Artificial Neural Networks for Spatial Data Analysis. Contemporary Engineering Sciences. Vol. 4, 2011, no. 4, 149 – 163.

Marzban, C. (2009). Artificial Intelligence Methods in the Environmental Sciences (Performance Measures and Uncertainty, Chapter 3, Springler Verlag Berlin.

Qaddoum, K., Hines, E. L. Iliescu, D. D. (2013). Yield Prediction for Tomato Greenhouse Using EFuNN. ISRN Artificial Intelligence. Article ID 430986. doi:10.1155/2013/430986.

Topuz, A. (2010). Predicting moisture content of agricultural products using artificial neural networks. Advances in Engineering Software 41, 464–470.

Toth, E., Montanari, A., and Brath, A. (2000). Comparison of short-term rainfall prediction model for real-time flood forecasting. J. Hydrol., 239, 132–147.

Weigend, S., Mangeas, M., and Srivastava, N. (1995). Nonlinear gated experts for time series: discovering regimes and avoiding overfitting. Int. J. Neural Syst., 6, 373–99.

Xiong, L., O’Connor, K. M., and Guo, S. (2004). Comparison of three updating schemes using artificial neural network in flow forecasting. Hydrol. Earth Syst. Sci., 8, 247–255, doi:10.5194/hess-8-247-2004.