Proceedings of the

World Congress of Computers in Agriculture and Natural Resources

Edited by

Fedro Zazueta
Jiannong Xin

Iguaçu Falls, Brazil
September 19-21, 2001

Rescheduled for
13-15 March, 2002
Use this date for citations

Published by
American Society of Agricultural Engineers
2950 Niles Road, St. Joseph, Michigan 49085-9659 USA

Solar Radiation Prediction Methods Applied to Improve Greenhouse Climate Control

J.P. Coelho[1], J.Boaventura Cunha and P.B. de Moura Oliveira[2]

Abstract

In this paper, deterministic and Artificial Neural Networks (ANNs) based techniques are applied to generate solar radiation forecasts with the purpose of being incorporated within a greenhouse predictive control strategy. These predictions are essential to estimate heat load fluctuations in the greenhouse caused by high frequency solar radiation changes, and so to improve ventilation and heating computation requirements for the greenhouse.

Keywords: Time Series Prediction, Horticulture, Artificial Neural Networks, Linear Regression.

Introduction

In order to achieve optimal climate control in greenhouses it is essential to have dynamic models to compute short and long-term predictions of the outside weather. The most relevant weather factor that must be accurately predicted is the solar radiation, since it is the variable with major influence on the greenhouse thermal load and crop photosynthesis. If a mathematical model describing a studied event is known, prediction becomes an easy task. However, if a model of the phenomenon is either indefinite or incomplete, it is usual to attempt to predict using a model that only takes into account previous results of the event while disregarding any other exterior influence.

Considering the above items, the aim of this work is to investigate modelling techniques that attain the most accurate solar radiation predictions in a given time horizon. This article reports the use of deterministic and ANNs based techniques to solar radiation time series prediction, in order to infer which one provides best results within a prediction horizon of sixty samples ahead.

Time Series Prediction

The observation of past results of an occurrence in order to foresee its future behaviour characterizes the essence of prediction. Generally, time series prediction problems are solved either from a deterministic or ANNs point of view.

[1] J.P. Coelho, ITIDAI – Instituto de Trás-os-Montes para a Investigação e Desenvolvimento Agro-Industrial, 5000 Vila Real , Portugal, jpcoelho@pol.pt
[2] J.Boaventura Cunha and P.B. de Moura Oliveira, Dep. Engenharias, Universidade de Trás-os-Montes e Alto Douro, 5000 Vila Real, Portugal, jboavent@utad.pt, oliveira@utad.pt

Deterministic methods are of limited applicability since they usually employ linear models, whereas ANNs are more powerful due to its intrinsic non-linear nature.

In practice, a time series is a set of discrete data collected sequentially in time at equal spaced instants *T*. In the present work, solar radiation will be modelled by a parametric autoregressive (AR) model based on:

$$y(k)=\sum_{m=1}^{M} a_m \cdot y(k-m)+e(k)$$

where M is the series order, $y(k)$ represent the system output at discrete time k and e is the model error.

When using a deterministic technique, finding an appropriate AR model means choosing an appropriate model order M, and to estimate the coefficients a_m, which is usually done through a least squares optimisation procedure (Ljung 1987). This technique is naturally restricted, since it assumes a linear relationship between sequence elements, moreover it assumes stationarity of the time series, i.e. its statistical distribution does not change across time. This idiosyncrasy makes them very attractive in practice since it implies that time invariant models could represent them. The first step in deterministic modelling is thus an attempt to stationarize the studied time series.

In this study, the performance of the models is analysed by using the computation of the residuals mean μ and determination coefficients r^2.

$$\mu=\frac{1}{n}\cdot\sum_{k=1}^{n}\left(x(k)-\hat{x}(k)\right)$$

$$r^2=1-\frac{\sum_{k=1}^{n}\left[x(k)-\hat{x}(k)\right]^2}{\sum_{k=1}^{n}\left[x(k)-\bar{x}\right]^2}$$

Where x and $\hat{x}$ denote actual and predicted process values respectively, while $\bar{x}$ denotes the mean of actual data.

The error mean is usually used to measure whether the predictor is biased or not. A predictor with error mean close to zero is called unbiased, whereas a predictor with error mean far from zero indicates a biased predictor, and points out that there are some deficiencies with the predictor that need to be solved. The error mean should not be considered in an absolute sense but with respect to the actual process value. The most important parameter for evaluating the prediction accuracy is the coefficient of determination, which is a function of the mean squared error normalized by the variance of the actual data. For a perfect predictor, the coefficient of determination should be equal to one.

Artificial Neural Networks for Prediction

Artificial neural networks are collections of mathematical models that reproduce some of the observed properties of biological nervous systems. The key element of the ANN is the structure of the information processing system. It is composed of a large number of highly interconnected processing elements that are analogous to neurons and are coupled together with weighted connections that are analogous to synapses. Although the original goal of ANNs was to investigate and replicate human information processing tasks such as speech, vision, and knowledge processing, ANNs also demonstrated their capability for classification and function approximation problems.

Since almost all measured phenomena are non-linear, the linear modelling methods are often inappropriate to describe their behaviour. Non-linear autoregressive models are potentially more powerful than linear ones because they can model more complex underlying characteristics of the series and they theoretically do not have to assume stationarity. Among non-linear methods, ANNs represent an attractive approach for time series prediction problems (Farmer *et al.* 1987).

There are a broad number of ANNs topologies. Among the most widespread are feedforward networks. In this work a multilayer perceptrons (MLP) network with a hyperbolic tangent (*tanh*) activation functions is used. These types of structure have proved to be a universal aproximator (Hornik *et al.* 1989). This means that they can approximate any reasonable function f with a subjective accuracy by:

$$f(y)=\left(\sum_{j=1}^{k} v_{jl}\tau\left(\sum_{i=1}^{n} w_{ij}\cdot y_i-\theta_j\right)-\theta_l\right), l=1...m$$

where τ is the *tanh* function, k is the number of hidden units, v_{jl} and w_{ij} are weights, θ_i are biases and y the data vector. MLP offer a straightforward extension to a well-known classical way of modelling time series: linear autoregressive models.

Learning in biological systems involves adjustments on the synaptic connections that exist between neurons. This is also true on ANNs.in which learning is typically done by expose the network to a set of input/output data with a training algorithm iteratively adjusting the connection weights. These weights store the knowledge necessary to solve specific problems. The objective of training is to determine a mapping from the set of training data to the set of possible weights so that the network will produce predictions, which in some sense are close to the true outputs. In this research, the non-linear function f is estimated based on samples from the series using the Lavenberg-Marquardt optimisation technique. The

Lavenberg-Marquardt is the standard method for minimization of mean square error criteria, due to its rapid convergence properties and robustness (Marquardt 1963).

ANNs have some major drawbacks. They require a large numbers of data samples due to their large number of degrees of freedom. Problems such as overfitting and sub-optimal minima, which are more severe than in the linear case may occur. In the linear case, overfitting can happen when choosing a value too high for the parameter M.

Problem Description

In order to apply optimal control to agricultural buildings it is essential to have a dynamic model that describes the evolution of the variables required for crop development. The use of Model Predictive Controllers for greenhouse indoor environment control has the advantage of providing system with the ability to react before any deviations in the controlled variable take place, avoiding delays in the system's response, (Boaventura Cunha *et al.* 2000). This class of control algorithms must employ models to predict solar radiation fluctuations over a specified time horizon. The way to obtain those models using linear and non-linear models is the subject of the present paper.

In this section, an example of a time-series identification using both deterministic and ANN techniques is presented. The aim is to develop a time-series model to describe the evolution of the outside solar radiation in time. This model will be useful to predict heat load fluctuations in the greenhouse caused by high frequency solar radiation changes, and to improve ventilation and heating computation requirements.

Despite of the large estimation and validation period usually used, on the present paper, for clarity sake, estimation and validation was done using only a one day period in different months. The data used for model estimation and validation was acquired for the period of one day with one minute sampling time (T_s).

Linear Model

It is assumed that a fourth order autoregressive parametric model is suitable to describe the time-series dynamics.

$$y(k)=a_1\cdot y(k-1)+a_2\cdot y(k-2)+a_3\cdot y(k-3)+a_4\cdot y(k-4)$$

in which $y(k-i)$ with $0\le i\le 4$, denotes the values of solar radiation at time $(k-i)\cdot T_s$, and a_i the coefficients estimated using a least squares algorithm. Before feeding the data into the estimator, it was pre-processed in order to remove high frequency noise. This was done using a second order low-pass filter with cut frequency equal to 40% of the Nyquist frequency.

Table 1 present the statistical results obtained when using a linear model to perform one and sixty minute prediction of solar radiation. Figure 1, illustrates the shapes of the predicted signal for the mentioned prediction horizons.

Non-Linear Model

To perform non-linear modelling of a time series descriptive of the solar radiation, a feed-forward artificial neural network was used. The structure chosed obeys the multilayer perceptrons strategy. The non-linear model used to represent the solar radiation is characterized by

$$y(k) = F_{MLP}\left[y(k-1), y(k-2), y(k-3), y(k-4)\right]$$

where the non-linear function F_{MLP} was estimated based on samples from the series, using the Lavenberg-Marquardt optimisation technique. Making F_{MLP} dependent on four prior sequence elements is equal to use four input units being supplied with four adjacent sequence elements. Initially fully connected network architecture with four inputs units (time window), ten hidden units and one output linear unit was selected. In order to remove redundant weights from the network, the network was pruned with an Optimal Brain Surgeon (OBS) strategy (Pederson et al. 1996).

Table 2 present the statistical results obtained when using a non-linear model to perform one and sixty minute prediction of solar radiation. Figure 2, illustrates the shapes of the predicted signal for the mentioned prediction horizons.

Conclusion

This work has drawn some comparison between deterministic and Artificial Neural Networks techniques applied to generate solar radiation predictions. It was observed from the current research that ANNs improve the mean square error between the predicted and observed series when compared with linear models, particularly when the predictions move forward in time.

Acknowledge

This work was sponsored by the INTERREG II European Program.

References

Boaventura Cunha, J., Couto, C. and Ruano, A.E.B. 2000. A greenhouse climate multivariable predictive controller, Acta Horticulture N.534, ISHS, 2000, pp: 269-276.

Farmer D. and Sidorowich J. 1987. Predicting chaotic time series. Physical Review Letter 59:845-848.

Girosi F. and Poggio T. 1990. Networks and the Best Approximation Property, Biological Cybernetics 63, 169-176.

Hornik K., Stinchcombe M. and White H. 1989. Multi-Layer Feedforward Networks are Universal Aproximators, Neural Networks 2, 359-366.

Ljung, L. 1987. System Identification – Theory for the User, Prentice-Hall.

Marquardt, D. 1963. An Algorithm for Least-Squares Estimation of Nonlinear Parameters, SIAM J. Appl. Math. 11, pp. 164-168.

Pederson, M. W., Hansen, L.K. and Larsen, J. 1996.Pruning with Generalization Based Weight Saliencies: gammaOBD, gammaOBS, Proceedings of NIPS'8, pp. 521-528.

Tables

Table 1. Statistical results for solar radiation prediction using a linear model.

	1 Step Ahead	60 Steps Ahead
μ	3.3774e-005	0.003847
r^2	0.99953	0.83127

Table 2. Statistical results for solar radiation prediction using a non-linear model.

	1 Step Ahead	60 Steps Ahead
μ	1.9296e-007	0.0091661
r^2	0.9888	0.94979

Figures

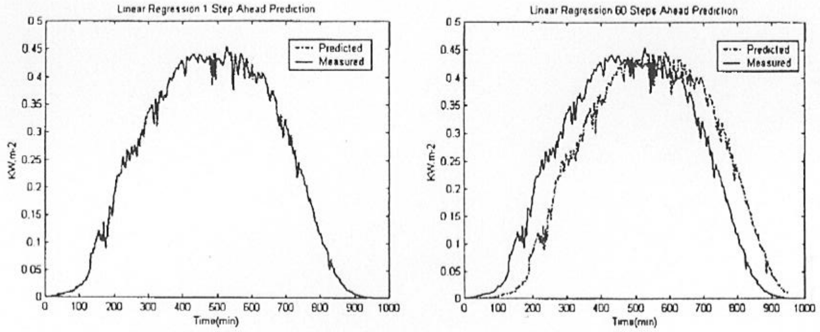

Figure 1. Predictions obtained when using the linear model.

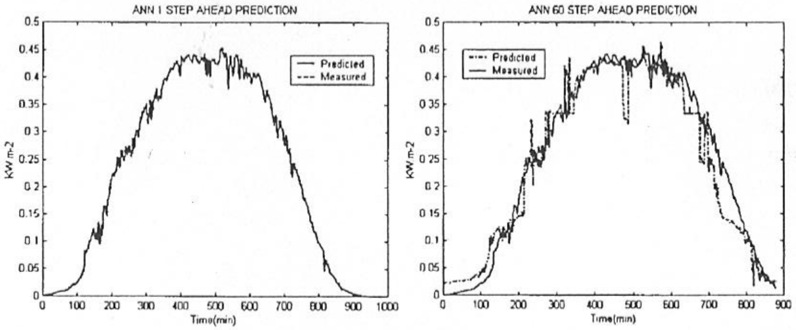

Figure 2. Predictions obtained when using the non-linear model.

Strategies of Climatic Control in Greenhouses with Programmed Automatons

Ingeniero Industrial Antonio Nolasco and José Basilio Nolasco

Abstract

The intensive agriculture in greenhouse is characterized by an agricultural production which controls to the greatest extend the biological processes involve in vegetal production. By gathering conditions, it is possible to achieve a complete control of the cultivation. This is an efficient and modern agriculture that makes use of natural resources in an intelligent an rational fashion with the objective of reaching high production, quality while being full respect with the environment.

Climatic control objective is to achieve that temperatures, relative humidity, light and CO2 is as close as possible to optimal conditions for the cultivation development. The equipments that exist in the market can control the temperature (aerometters and evaporation screens), relative humidity (nebulisators), ventilation (zenithals) and solar radiation (thermic screens) among others. The difficulty lies in being able to manipulate all these elements in a precise way, in order to achieve the desired climate inside the greenhouse, with a minimum waste of energy and water.

A programmable automaton is an electronic device capable to control in real time sequential processes and to carry out a wide variety of logical functions such as: regulations, drips, computations, timing, etc. They have input terminals or peripherals to which sensors are connected, switches and buttons and output terminals to which it can be connected relays, electrovalves, motors, lamps, etc. As a result of this, a programmable automaton will receive an information through the sensors and will activate over the outputs a type of action in one way or another, depending of each program and the control strategy design, achieving a precise climate control and closed to the conditions that have been considered as optimal for the cultivations in each of its phases.

Keywords: Equipment automation; Instrumentation and Control; Greenhouses technology; Climatic Control

Introduction

The intensive agriculture in greenhouse is characterized by an agricultural production which controls to the greatest extend the biological processes involve in vegetal production. By gathering conditions, it is possible to achieve a complete control of the cultivation.

This is an efficient and modern agriculture that makes use of natural resources in an