Computers and Electronics in Agriculture 122 (2016) 94–102

Contents lists available at ScienceDirect

# Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

# Prediction of the temperature in a Chinese solar greenhouse based on LSSVM optimized by improved PSO

Huihui Yu [1], Yingyi Chen [1], Shahbaz Gul Hassan, Daoliang Li *

*College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China*
*Key Laboratory of Agricultural Information Acquisition Technology, Ministry of Agriculture, Beijing 100083, China*
*Beijing Engineering and Technology Research Center for Internet of Things in Agriculture, Beijing 100083, China*

A R T I C L E  I N F O

*Article history:*
Received 20 December 2014
Received in revised form 7 December 2015
Accepted 15 January 2016
Available online 3 February 2016

*Keywords:*
Greenhouse
Temperature
Predictions
Least squares support vector machine
Particle swarm optimization algorithm

A B S T R A C T

Predictions of the Chinese solar greenhouse temperatures are important because they play a vital role in greenhouse cultivation, with solar greenhouse crops susceptible to potential losses because of cold and hot temperatures. Therefore, it is important to set up a precise predictive model of temperature that can predict the occurrence of temperatures several hours before head to reduce financial losses. This paper presents a novel temperature prediction model based on a least squares support vector machine (LSSVM) model with parameters optimized by improved particle swarm optimization (IPSO). The IPSO with probability of mutation was employed to optimize the required hyper parameters of the LSSVM model. The performance of the IPSO–LSSVM model was compared with traditional modeling approaches by applying it to predict solar greenhouse temperatures, and the results showed that its predictions of the maximum and minimum temperature were more accurate than those of the standard support vector machine (SVM) and Back propagation neural network (BPNN). Therefore, it is a suitable and effective method for predicting the Chinese solar greenhouse temperatures.



## 1. Introduction

Typical solar greenhouses in northern China have a plastic film covering the slanted front roof during the day with a thermal blanket added at night to maintain the heat inside. In the winter, northern China experiences sunshine more than 50% of the time, but the temperature is very low, sometimes even falling to below −10 °C. These greenhouses enable the extension of the growing season in the cold climatic conditions of northern China, so that flowers and vegetables can be produced year round (Tong et al., 2009). Such greenhouses must provide the good environment conditions for plant growth, especially maintaining adequate the temperature. Plants exposed to low or high temperatures may lead to mass death or fungal diseases. Therefore, it is important to set up a precise predictive model of temperature in the Chinese solar greenhouses that can predict the occurrence of temperatures by several hours to reduce financial losses.

Accurately predicting temperatures in greenhouses has been a focus of research because it is the most important influencing factor for crop growth and development (Zuo et al., 2010), and suitable temperatures can improve the mortality rate, quality and quantity of plants produced. Specifically, crops may be seriously damaged if extreme temperatures in greenhouses occur without any protective measures, and such temperatures increase the susceptibility to diseases (Li et al., 2013). The resulting decline in the quality and yield of greenhouse cultivation causes significant economic losses to growers. Therefore, farmers cultivating crops in greenhouses must perform preventative protective measures to avoid the occurrence of high and low temperatures (Ou et al., 2014). However, other than the research of Tong et al. (2009), there have been no studies of the temperature variations in the Chinese solar-type greenhouse. The temperature inside a solar greenhouse is nonlinear and dynamic changing and is affected by uncertainties like the human activity, environment changing, and biological factors. The modeling method must be successful enough to accurately forecast the solar greenhouse temperature with incomplete background information and uncertainties.

In recent years, different computational approaches have been applied for various aspects of temperature prediction problems (Patil et al., 2008; Paniagua-Tineo et al., 2011; Van Beveren et al., 2015). The temperature prediction methods can be primarily divided into two groups: the first physical methods based on mathematical theory; and the second black box methods based

* Corresponding author at: P.O. Box 121, China Agricultural University, 17 Tsinghua East Road, Beijing 100083, China. Tel.: +86 10 62736764; fax: +86 1062737741.

*E-mail address:* dliangl@cau.edu.cn (D. Li).

[1] Yingyi Chen and Huihui Yu contributed equally to the work.

http://dx.doi.org/10.1016/j.compag.2016.01.019
0168-1699/

on modern computational technology. Over the last decades, numerous physical greenhouse temperature methods have been presented (Du et al., 2012; Van Beveren et al., 2015). Generally, the physical methods have a high degree of complexity with lots of parameters that have to be determined (He and Ma, 2010). And, it's often a difficult task to measure the parameters or get them from the sub-models. In contrast to physical methods, the 'black box' prediction methods don't need to determine every parameter value. A variety 'black box' methods have been used in the greenhouse environment predicting such as time series model, auto regressive exogenous model (ARX model) (Patil et al., 2008; Wakaura and Ogata, 2007). Patil et al. (2008) built linear regression models by auto regressive method (AR), an auto regressive moving average method (ARMAX) and a neural network auto regressive method (NNARX) for a greenhouse in Thailand, and the NNARX model performed better than other models. However, the model has the drawback of noise due to the sequential effect of rain. Wakaura and Ogata (2007) utilized the auto regressive method as modeling tool for the long-term and linear forecasting of the surface air temperature, and the method used many samples. Thus, it is not suitable for nonlinear and small samples greenhouse temperature prediction. Overall, the ARX and time series model provide simple and fast predictions, they require numerous samples, have poor accuracy, and are unable to solve nonlinear prediction. Therefore, different methods based on modern computational technology have been studied to construct quantitative prediction models that do not rely on physical mechanisms or simple structures. Moreover, computer-based modeling is fast and provides accurate results. Thus far, grey fuzzy approach models (Lu et al., 2014; Su et al., 2012) and artificial neural networks (ANN) models have been used for greenhouse inside temperature predication (Kok et al., 1994; Dariouchy et al., 2009; Ferreira et al., 2002). Neural networks methods (Ruano et al., 2006; Baboo and Shereef, 2010), support vector regression algorithm methods (SVM) (Wang et al., 2010; Lins et al., 2013; Radhika and Shashi, 2009) and other combination forecasting models (Kayacan et al., 2010) have been explored for temperature prediction. Lu et al. (2014) and Su et al. (2012) proposed a fuzzy logic approach which enables accurate temperature calculation in an area monitored by a wireless sensor network. Although the fuzzy approach requires less historical data, its ability to provide accurate forecasts is affected by fluctuations in the data, which can produce large errors. Dariouchy et al. (2009) developed an artificial neural network (ANN) and used it to predict the internal temperature and the internal moisture inside the greenhouse starting from the external climatic data. The model can perform better than the time series prediction models. Ruano et al. (2006) used neural networks for air temperature prediction inside buildings. It was shown that the neural models can achieve better results than state-of-the-art physical models. However, all of the methods have numerous drawbacks, including low convergence, over-fitting, poor stability and cannot solve complex system prediction well. Support vector regression machine learning algorithms (SVM) are attracting more attention for the forecast models. Wang et al. (2010) and Lins et al. (2013) built the temperature prediction models, and the model can solve problems that are non-linear and have multiple variables and small sample sizes. Although function estimation and classification can be handled by SVM but it is not suitable for higher dimensional computation.

LSSVM has the better ability to solve the higher computational problem than SVM. Moreover, large scale problem can be handled more easily, more efficiently using LSSVM (Wang and Hu, 2005). The least squares support vector machine (LSSVM) proposed by Suykens et al. (2002), is a modified, simpler version of an SVM (Çalişir and Dogantekin, 2011) that can address linear and nonlinear multivariable calibration and solve multivariable calibration problems relatively quickly. Therefore, it has been successfully applied to many fields of function approximation and pattern recognition with high accuracy and an increased capacity for generalization (Liao et al., 2011). The LSSVM attempts to minimize an upper bound of the generalization error instead of the empirical error, and it can provide more reliable and improved generalization performances under the same training conditions compared with the artificial neural networks method (Long et al., 2014). Since the solar greenhouse environment is a dynamic, complex and nonlinear system, the modeling of the temperature inside the solar greenhouse includes many uncertainty factors, such as the human activity, environment changing, and biological factors. And the time series method cannot solve nonlinear problem, the neural networks method cannot solve too complex system forecasting, and the support vector machine is not suitable for higher computation problem. Thus, in this paper, the least square support vector machine is adopted to develop the solar greenhouse temperature forecasting model. However, LSSVM performance is heavily impacted by the selection of several hyper parameters, which are necessary to define the optimization problem and final LSSVM model. It is important to choose a kernel function, set the kernel parameters and determine a soft margin constant C and e-insensitive loss parameter. The parameters that should be optimized include the regularization parameter C and kernel function parameters, such as the gamma ($\sigma$) parameter for the radial basis function (RBF) kernel. To date, an exact method of obtaining the optimal set of LSSVM hyper parameters has not been determined.

Recently, particle swarm optimization (PSO) and differential evolution (DE) have been introduced and particularly PSO has received increased interest from the evolutionary computation community (Zhang et al., 2009; Yu et al., 2014). Both PSO and DE algorithms have shown great promise in several field applications. In this study, the particle swarm optimization algorithm is adopted to optimize the parameters for the LSSVM method. The particle swarm optimization (PSO) is heuristic global optimization method introduced originally by Kennedy and Eberhart in 1995 (Kennedy and Eberhart, 1995). And it's widely used in the fields such as function optimization, parameters training, and model classification, due to its many advantages including its simplicity and easy implementation (Bai, 2010; Chang, 2015). However, the basic particle swarm optimization does not ensure the convergence to an optimal solution and also easily suffers from the partial optimization (Chang, 2015). Own to its comparatively poor efficiency, this paper presents an improved particle swarm optimization algorithm as a technique to simultaneously optimize the LSSVM parameters. And, a comparative result of the improved particle swarm optimization algorithm and the particle swarm optimization's parameter inspired by differential evolution algorithm has been made. In this paper, traditional SVM and ANN models are used for comparative purposes.

The paper is organized as follows. In Section 2, we describe the study area, sources of data and proposed model. In Section 3, we report the results of numerical experiments, and in Section 4, we provide conclusions from this research and suggest directions for future investigations.

## 2. Materials and methods

### 2.1. Data acquisition

The data used in this study were acquired by the horticultural monitoring and management system installed at the zone of technology application and demonstration of the China Agricultural University –Shouguang Vegetable Industry Group in Shandong Province, China. Up to now, the system has been applied in more