Renewable Energy 36 (2011) 3054–3060

Contents lists available at ScienceDirect

Renewable Energy

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

# Prediction of daily maximum temperature using a support vector regression algorithm

A. Paniagua-Tineo [a], S. Salcedo-Sanz [a,*], C. Casanova-Mateo [c], E.G. Ortiz-García [a], M.A. Cony [b], E. Hernández-Martín [c]

[a] Department of Signal Theory and Communications, Universidad de Alcalá, 28871 Alcalá de Henares, Madrid, Spain
[b] Department of Renewable Energy, Center for Energy, Environmental and Technologic Research (CIEMAT), Spain
[c] Department of Physics of the Earth, Astronomy and Astrophysics II, Universidad Complutense de Madrid, Spain

A R T I C L E  I N F O

*Article history:*
Received 22 October 2010
Accepted 23 March 2011
Available online 22 April 2011

*Keywords:*
Daily maximum temperature prediction
Support vector regression algorithms
Neural networks

A B S T R A C T

Daily maximum temperature can be used a good indicator of peak energy consumption, since it can be used to predict the massive use of heating or air conditioning systems. Thus, the prediction of daily maximum temperature is an important problem with interesting applications in the energy field, since it has been proven that electricity demand depends much on weather conditions. This paper presents a novel methodology for daily maximum temperature prediction, based on a Support Vector Regression approach. The paper is focused on different measuring stations in Europe, from which different meteorological variables have been obtained, including temperature, precipitation, relative humidity and air pressure. Two more variables are also included, specifically synoptic situation of the day and monthly cycle. Using this pool of prediction variables, it is shown that the SVMr algorithm is able to give an accurate prediction of the maximum temperature 24 h later. In the paper SVMr technique applied is fully described, including some bounds on the machine hyper-parameters in order to speed up the SVMr training process. The performance of the SVMr has been compared to that of different neural networks in the literature: a Multi-layer perceptron and an Extreme Learning Machine.



## 1. Introduction

Accurate prediction of air temperature is a problem which has attracted the attention of researchers in the last few years, since it has many different applications in areas such as industry, agriculture or energy [1]. Some of these applications are the dimensioning in air conditioning systems in buildings [2], the design of solar energy systems [3], the calculation of temperatures in greenhouses to avoid crop lost [4], the prediction of natural hazards such as wild-fires [5], etc. Specifically, the prediction of daily maximum temperature is an important problem in the energy field, since this figure can be used to predict peak energy consumption due to the massive used of heating systems in winter or air conditioning devices in summer [6–8]. Also, different studies have associated maximum air temperature with direct solar radiation parameters at a given point [9], and also with the prediction of solar radiation [10], what has important consequences for photovoltaic farms and devices.

In the last few years, different soft computing approaches have been applied in different areas to temperature prediction problems. The majority of these approaches have used neural computing techniques, which are fast and provide accurate results. For example, there are several recent works applying Multi-layer perceptrons (MLP) to temperature prediction problems, in different scenarios or countries [4,11–14]. In [1] a different type of neural network, called abductive network was successfully applied to a problem of hourly temperature prediction over data in Seattle, USA. Other types of neural networks such as Radial Basis Function (RBFs), generalized regression networks and also statistical approaches have been applied to temperature prediction in the last few years [14–16].

The regression Support Vector Machine (SVMr) [17,18], however, is a type of robust regression technique which has not been much applied to temperature prediction problems, in spite of the good results it has obtained in the prediction of other atmospheric phenomena, such as wind speed or solar radiation [19,20]. In this paper, it is shown that the SVMr can be successfully applied to a problem of daily maximum temperature prediction from data of measuring stations in Europe, improving the results of other neural-type algorithms. This paper also discuses an important

* Corresponding author. Tel.: +34 91 885 6731; fax: +34 91 885 6699.
*E-mail address:* sancho.salcedo@uah.es (S. Salcedo-Sanz).

0960-1481/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.renene.2011.03.030

point related to the inclusion of synoptic meteorological conditions into the prediction variables, and how it affects to the maximum temperature prediction. A complete comparison with an MLP and an Extreme Learning Machine (ELM), completes the experimental part of the paper.

The rest of the paper is structured in the following way: next section provides a small description of the SVMr algorithm, focusing on the $\epsilon$-SVMr method, the one applied in this paper. Section 3 presents the available data in which the SVMr approach is applied, and also the methodology carried out to show the significance of the obtained results. Section 4 presents the experimental part of the paper. First, the neural algorithms used for comparison are briefly described, and then the results obtained by the different approaches considered in this paper are shown, in the data described in Section 3. Section 5 closes the paper giving some final remarks .

## 2. Support vector machines for regression problems

One of the most important statistic models for prediction are the Support Vector Regression algorithms (SVMr) [18]. The SVMrs are appealing algorithms for a large variety of regression problems [19,20], since they do not only take into account the error approximation to the data, but also the generalization of the model, i.e., their capability to improve the prediction of the model when a new dataset is evaluated by it. Although there are several versions of SVMr, in this case the classic model presented in [18] is considered and described:

The $\epsilon$-SVMr method for regression consists of, given a set of training vectors $C = \{(\mathbf{x_i}, y_i), i = 1, \ldots, l\}$, training a model of the form $y(\mathbf{x}) = f(\mathbf{x}) + b = \mathbf{w}^T\phi(\mathbf{x}) + b$, to minimize a general risk function of the form

$$R[f] = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{l} L(y_i, f(\mathbf{x})) \tag{1}$$

where $\mathbf{w}$ controls the smoothness of the model, $\phi(\mathbf{x})$ is a function of projection of the input space to the feature space, $b$ is a parameter of bias, $\mathbf{x_i}$ is a feature vector of the input space with dimension $N$, $y_i$ is the output value to be estimated and $L(y_i, f(\mathbf{x}))$ is the loss function selected. In this paper, the L1-SVMr (L1 support vector regression) is used. It is characterized by an $\epsilon$-insensitive loss function [17]

$$L(y_i, f(\mathbf{x})) = |y_i - f(\mathbf{x_i})|_\epsilon \tag{2}$$

In order to train this model, it is necessary to solve the following optimization problem [17]:

$$\min\left(\frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{l}\left(\xi_i + \xi_i^*\right)\right) \tag{3}$$

subject to

$$y_i - \mathbf{w}^T\phi(\mathbf{x_i}) - b \le \epsilon + \xi_i, i = 1, \ldots, l \tag{4}$$

$$-y_i + \mathbf{w}^T\phi(\mathbf{x_i}) + b \le \epsilon + \xi_i^*, i = 1, \ldots, l \tag{5}$$

$$\xi_i, \xi_i^* \ge 0, \quad i = 1, \ldots, l \tag{6}$$

The dual form of this optimization problem is usually obtained through the minimization of the Lagrange function, constructed from the objective function and the problem constraints. In this case, the dual form of the optimization problem is the following:

$$\max\left(-\frac{1}{2}\sum_{i,j=1}^{l}\left(\alpha_i - \alpha_i^*\right)\left(\alpha_j - \alpha_j^*\right)K(\mathbf{x_i}, \mathbf{x_j}) - \epsilon\sum_{i=1}^{l}\left(\alpha_i + \alpha_i^*\right) + \sum_{i=1}^{l} y_i\left(\alpha_i - \alpha_i^*\right)\right) \tag{7}$$

subject to

$$\sum_{i=1}^{l}\left(\alpha_i - \alpha_i^*\right) = 0 \tag{8}$$

$$\alpha_i, \alpha_i^* \in [0, C] \tag{9}$$

In addition to these constraints, the KarushKuhn–Tucker conditions must be fulfilled, and also the bias variable, $b$, must be obtained. Details about this process for simplicity, the interested reader can consult [17] for reference. In the dual formulation of the problem the function $K(\mathbf{x_i}, \mathbf{x_j})$ is the kernel matrix, which is formed by the evaluation of a kernel function, equivalent to the dot product $\langle\phi(\mathbf{x_i}), \phi(\mathbf{x_j})\rangle$. A usual election for this kernel function is a Gaussian function, as follows:

$$K(\mathbf{x_i}, \mathbf{x_j}) = exp\left(-\gamma\cdot\|\mathbf{x_i} - \mathbf{x_j}\|^2\right). \tag{10}$$

The final form of function $f(\mathbf{x})$ depends on the Lagrange multipliers $\alpha_i, \alpha_i^*$, as follows:

$$f(\mathbf{x}) = \sum_{i=1}^{l}\left(\alpha_i - \alpha_i^*\right)K(\mathbf{x_i}, \mathbf{x}) \tag{11}$$

In this way it is possible to obtain a SVMr model by means of the training of a quadratic problem for a given hyper-parameters $C$, $\epsilon$ and $\gamma$. However, obtaining these parameters is not a simple procedure, being necessary the implementation of search algorithms to obtain the optimal ones or the estimation of them [21].

## 3. SVMr in daily temperature prediction: data available

The presented study is focused on a 10-year period of data, from 1989 to 1998, in Europe. In this study, information extracted from the ECA (European Climate Assessment) dataset [22] has been used. This dataset has been chosen because its temporal resolution, meteorological variables stored and spatial coverage are pretty good for this kind of studies, since it contains series of about 2000 stations and nine variables: minimum, mean and maximum temperature, precipitation, sea level air pressure, snow depth, sunshine duration, relative humidity, and cloud cover [23]. A selected group of five predictor meteorological variables were chosen from the ECA dataset:

1. Maximum temperature.
2. Minimum temperature.
3. Precipitation.
4. Sea level air pressure.
5. Relative humidity.

In order to choose the stations from which data would be extracted, two simply test were applied:

a) Though it is possible to fill in missing data through different spatial interpolation techniques, only series without missing data in the 10-year period of study have been selected.