A support vector machine approach to estimate global solar radiation with the influence of fog and haze

Wanxiang Yao a, b, Chunxiao Zhang c, d, Haodong Hao e, Xiao Wang e, Xianli Li c

Article history:
Received 7 February 2018
Received in revised form 25 April 2018
Accepted 19 May 2018
Available online 21 May 2018

Keywords:
Support vector machines
Global solar radiation
Fog and haze
Air quality index

Abstract

In recent years, fog and haze occurred frequently, due to energy crisis and environmental pollution. Fog and haze have significant scattering-weakening effect on solar radiation, resulting in a severe weaken to solar radiation received on a horizontal surface. In this paper, air quality index (AQI) is taken as an additional input parameter, and some new models for estimating global solar radiation on a horizontal surface are proposed based on a support vector machine (SVM). The accuracy of SVM-1 and SVM-2 models are compared and analyzed, and the results show that the performance of SVM-2 models with an extra input parameter AQI are generally improved, for which the R value is promoted from 0.848 to 0.876, the NSE value is lifted from 0.682 to 0.740, the RMSE value is reduced from 0.114 to 0.102, and the MAPE value is decreased from 9.257 to 8.214. Comparing with existing models, SVM models proposed in this paper can improve the accuracy of global solar radiation models. If AQI is used as an additional input parameter to estimate global solar radiation, the accuracy will be further improved.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

In recent years, population growth has led to the increasingly intensified energy crisis and environmental pollution. All countries have focused their research on renewable energy. As a renewable energy source with no pollution, large reserves and wide distribution, solar energy has been widely used as an energy source. However, the interaction of multiple meteorological parameters (e.g. latitude, sunshine duration, temperature, relative humidity, etc.) on a horizontal surface of solar radiation, and many empirical models have been proposed based on a large number of historical data [4–6]. In addition, some researchers have tried to estimate global solar radiation on a horizontal surface by coupling satellite images with ground observation data [7–9], which is more suitable to estimate solar radiation in a large scale (larger than 10 km), especially for special weather conditions (such as overcast days, cloudy days, rain and snow days, etc.). Solar radiation is influenced by many factors at the same time. Considering the relative simplification of functional relationship, empirical models can only reflect the influence of one or several parameters on solar radiation. However, the interaction of multiple parameters is rarely considered.

With the development of artificial intelligence and bio-anthropophic technology, many researchers apply neural networks to solar energy, especially in the study of solar radiation [10–12]. Because neural network is not restricted by conventional functions and can consider the interaction between multiple parameters at the same time, it has unique advantages to estimate solar radiation. Support vector machine (SVM) has a wide range of applicability compared to other neural networks, which can not
only deal with linear problems, but also map nonlinear problems into high-dimensional linear space by kernel function to ensure the adaptability of this algorithm. In addition, based on statistical learning theory, SVM uses the principle of structural risk minimization to learn from finite samples, and seeks trade-off between approximation accuracy and approximation function complexity of given data, so that this algorithm has strong generalization ability. Therefore, lots of solar radiation models have been established by various scholars based on SVM algorithm in the past decades [13–18]. Chen et al. [13] proposed a series of models which use a combination of seven temperatures as input parameters, and linear, polynomial and radial basis functions as kernel functions. These models use support vector machine to estimate monthly average solar radiation. Results show that the model which uses polynomial as kernel function and $T_{\text{max}}$ and $T_{\text{min}}$ as input parameters is the best. Based on historical data (solar radiation, cloud cover, relative humidity and wind speed, etc.), Zeng et al. [14] used SVM algorithm to predict atmospheric transmissivity and then converted it into solar energy according to latitude and the time of day. Through verification with the National Solar Radiation Database, this model shows a good consistency. If additional meteorological variables are used, especially cloud cover, the accuracy of this model will be further improved. Chen et al. [16] established global solar radiation models based on sunshine duration using support vector machine learning, and compared SVM models of seven different input parameters with five empirical models. The results demonstrate that the performance of SVM model is better than empirical models. Meanwhile, combining sunshine duration with other meteorological parameters can improve the accuracy of models. Z. Ramedani et al. [15] proposed a support vector regression method to predict global solar radiation. Two SVR models are studied, where one is radial basis function model (SVR-rbf) and the other is a polynomial function model (SVR-poly). The estimated performance of SVR-rbf model is better than SVR-poly model, which can effectively improve the accuracy and shorten calculation time. Olatomiwaa et al. [17] used sunshine duration, maximum and minimum temperature as input parameters, and combined support vector machine with firefly algorithm (SVM-FFA) to predict monthly average solar radiation on a horizontal surface. Compared artificial neural network (ANN) and Genetic Programming (GP) models, the SVM-FFA models have higher prediction accuracy. S. Shamshirband et al. [18] combined support vector machine with wavelet analysis to establish a coupled model for estimating solar radiation, and tested with data of Iran (cloud cover and clearness index). As a result, the model has a good agreement with measured data.

In addition, environmental pressure is increasing with the process of industrialization. There are different degrees of fog and haze in many regions of China, in which Beijing-Tianjin-Hebei region is particularly serious [19,20]. The number of air pollution days in Beijing was 143 days in 2017, accounting for 39.2% of the whole year, with severe and serious pollution for 27 days. On sunny days, solar radiation is absorbed and scattered by the upper atmosphere; when it is overcast or cloudy days, especially heavy fog and hazy weather, solar radiation is not only passing through atmosphere, but also low-level clouds and particles in air near the ground. Solar radiation is scattered, reflected and refracted, and it is severely weakened when reaching the ground. Therefore, near-surface atmosphere seriously affects the acceptable solar radiation on a horizontal surface. The near-surface atmospheric conditions can be characterized by air quality index (AQI), which is almost directly proportional to the concentration of particulate matter (PM) in the air [21,22]. Some researchers considered AQI as an additional input parameter to establish solar radiation models. For example, Yao et al. [23] established empirical model of daily diffuse solar radiation based on the data of nearly 55 years in Beijing, China, and modified it with AQI. Result demonstrates that the accuracy of this model is improved. Based on meteorological data of Tehran in Iran for one year, Masoud Vakili et al. [24] established daily global solar radiation models by using an artificial neural network which input parameters include temperature, relative humidity, wind speed and air particulate matter (PM).

The main purpose of this paper is using SVM algorithm to realize the learning mechanism of finite samples based on the principle of minimizing structural risk, and investigate the influence of different input parameters on output performance of these models, especially the contribution of AQI to output performance. In this paper, based on the influence of fog and haze, SVM neural network is applied to establish daily global solar radiation models of Beijing. The influence of different input parameters (surface temperature difference, air temperature difference, relative humidity, sunshine duration and AQI) on the accuracy of SVM models are compared and analyzed. In addition, these new SVM models and existing models are compared and analyzed based on statistical parameters. The results of this paper can provide a reference for estimating solar radiation in the regions which are lacking of detection equipment, and provides a basis for the evaluation, conversion, and utilization of solar energy resources, especially for regions where environmental pollution is more serious.

2. Method and data

2.1. Support vector machine

Support vector machines (SVM) were originally proposed by V. Vapnik [25,26] for classification, non-linear regression and other related fields. The theoretical basis of SVM is statistical learning theory, more precisely, SVM is an approximate realization of structural risk minimization. According to statistical learning theory, the actual risk of regression estimation is related to empirical risk and confidence range. The empirical risk is related to training sample, and the confidence range is related to Vapnik-Chervonenkis dimension (VC dimension) of learning machine and number of training samples. When these two factors are both small, that is, not only empirical risk is small, but also VC dimension is
calculated as Eq.(1).

$$|y - f(x)| = \begin{cases} 0 & |y - f(x)| \leq \varepsilon \\ |y - f(x)| - \varepsilon & |y - f(x)| > \varepsilon \end{cases}$$

(1)

Where $f(x)$ is the regression estimation function constructed by learning sample dataset, and $y$ is corresponding target value of $x$. The purpose of SVM training for samples is to establish regression function $f(x)$, so the distance between estimated value and expected value of regression function is less than $\varepsilon$. When VC dimension of regression function is the smallest, regression function has the best generalization ability, which perfectly solves the problem of sample over fitting.

Because regression estimation of sample dataset cannot be achieved by linear functions in a low-dimensional space, the sample dataset are mapped to a high-dimensional linear feature space in which a linear regression can be implemented. If mapping function of sample dataset is a nonlinear function $\phi(x)$, the regression estimation function $f(x)$ can be expressed as Eq.(2).

$$f(x) = w \cdot \phi(x) + b$$

(2)

Where dimensions of $w$ is the dimension of high-dimensional linear feature space after sample dataset is mapped. In order to analyze different emphasis degree of data, slack variables $\xi_i$ and $\xi_i^*$ are introduced, respectively representing relaxation factors at the upper and lower bounds. The penalty factor $C$ is an empirical value, which determines recognition degree of this algorithm on the loss of different data. Hence, Eq.(2) is rewritten as Eq.(3).

$$\min_{w,b}\frac{1}{2}||w||^2 + C \sum_{i=1}^{l} (\xi_i + \xi_i^*)$$

s.t. $y_i - w \cdot \phi(x_i) - b \leq \varepsilon + \xi_i$

$w \cdot \phi(x_i) + b - y_i \leq \varepsilon + \xi_i^*$

$\xi_i \geq 0$

$\xi_i^* \geq 0$, $i = 1, 2, \ldots, l$

(3)

The lagrange multiplier method is utilized to solve the optimization problem, and regression estimation function is shown as Eq.(4).

$$f(x) = \sum_{x_i \in SV} (a_i - a_i^*) K(x_i; x_j) + b$$

(4)

Where $a_i$ and $a_i^*$ are lagrange multipliers, and $K(x_i; x_j) = \phi(x_i) \cdot \phi(x_j)$ is kernel function. Although SVM algorithm needs to map a sample set to the high-dimensional feature space by a nonlinear function, it only needs to calculate the kernel function when calculating the regression function. In order to avoid dimensionality disaster caused by high-dimensional feature space, the kernel function must satisfy Mercer condition.

2.2. Data

Considering the present situation of fog and haze in China, Beijing with more serious fog and haze is selected as the research object. Annual mean daily global solar radiation in Beijing is 14.71 MJ/m² and average daily sunshine duration is 6.83 h. The data of Beijing (39°48'N, 116°28'E) [29] are used to establish SVM models and to discuss the coupling effects of fog and haze with other meteorological parameters on solar radiation. The data from December, 2013 to August, 2016 is used to train SVM neural network models, a total of 1000 groups of data. The data from September, 2016 to March, 2017 is employed to verify the accuracy of SVM models, a total of 210 groups of data. Data used in this paper includes daily global solar radiation, sunshine duration, temperature, relative humidity and air quality index (AQI).

2.3. Statistical parameters

Correlation coefficient (R), Nash-Sutcliffe Equation (NSE), Root mean square error (RMSE) and Mean absolute percentage error (MAPE) are used to characterize the error between measured value and calculated value of different models, which reflect the performance of different models [23,30–32]. These four statistical parameters are calculated as follows:

$$R = \sqrt{\frac{\sum_{i=1}^{n} (c_i - c_a) (m_i - m_a)}{\left[ \sum_{i=1}^{n} (c_i - c_a)^2 \right] \left[ \sum_{i=1}^{n} (m_i - m_a)^2 \right]}}$$

(5)

$$NSE = 1 - \frac{\sum_{i=1}^{n} (m_i - c_i)^2}{\sum_{i=1}^{n} (m_i - m_a)^2}$$

(6)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (c_i - m_i)^2}$$

(7)

$$MAPE = \frac{\sum_{i=1}^{n} |c_i - m_i| \cdot 100}{n}$$

(8)

Where $c_i$ is calculated value; $c_a$ is average calculated value; $m_i$ is measured value; and $m_a$ is average measured value. When R and NSE are closer to 1, and MAPE and RMSE are closer to 0, the model is more accurate.

3. Results and discussion

In this paper, the SVM algorithm is used to modeling and evaluate global solar radiation with MATLAB. The detailed steps are as follows:

(1) All data are preprocessed, and unreasonable data is removed. Then global solar radiation on a horizontal surface and sunshine duration are normalized to get clearness index and sunshine percentage.

(2) The data is imported into MATLAB software, and divided into training dataset (1000 groups of data) and verification dataset (210 groups of data).

(3) Learning the sample data through the training of data sets, and a linear kernel function is used to establish global solar radiation model (SVR-1 model). The calculated value and measured value are comparatively analyzed to evaluate the accuracy of different models.
(4) To analyze the internal relationship between different influence factors and explore its impact on solar radiation, AQI is imported to establish SVM-2 models based on SVM-1 global solar radiation models.

(5) The SVM-1 models, SVM-2 models and the existing models are comparatively analyzed to evaluate the accuracy of SVM models.

According to the influence of different input parameters on solar radiation, the SVM models are divided into two categories (SVM-1 and SVM-2 models), as shown in Fig. 1. The first category is SVM-1 models, and their input parameters include surface temperature difference, air temperature difference, relative humidity and sunshine duration. In addition, air quality index (AQI) is added to establish SVM-2 models based on SVM-1 models, and the input parameters of SVM-2 models include surface temperature difference, air temperature difference, relative humidity, sunshine duration and AQI.

Table 1 and Fig. 2 show the accuracy of these models based on SVM algorithm is improved in different degrees. For training dataset, compared with SVM-1 models, the accuracy of SVM-2 models after adding AQI parameters is improved accordingly. Among average value of these statistical parameters, the R value is promoted from 0.864 to 0.894, the NSE value is lifted from 0.712 to 0.764, the RMSE value is reduced from 0.091 to 0.088, and the MAPE value is decreased from 8.203 to 7.692. Similarly, for validation dataset, the R value is promoted from 0.848 to 0.876, the NSE value is lifted from 0.682 to 0.740, the RMSE value is reduced from 0.114 to 0.102, and the MAPE value is decreased from 9.257 to 8.214.

Among the first type of models (SVM-1 models), SVM-1-8 model based on 4 input parameters is the most accurate, because the SVM model takes into account the coupling effect among different factors, and the meteorological parameters can be repaired with each other to improve the accuracy of global solar radiation. What’s more, sunshine duration represents the time duration with a certain intensity (greater than or equal to 120 W/m²), which plays a most significant impact on solar radiation. Therefore, SVM-1-4 model based on sunshine duration is the most accurate, which can be verified from Table 1. Meanwhile, sunshine duration and conventional meteorological parameters are used as input parameters, and SVM-1-5 model, SVM-1-6 model and SVM-1-7 model are obtained, in which SVM-1-5 model (its input parameters are sunshine duration and surface temperature difference) is the best. Its R value is 0.940, NSE value is 0.851, and RMSE value is 0.083.

The second type of models (SVM-2 models) is similar to the first type of models (SVM-1 models), in which SVM-2-8 model based on five input parameters is the most accurate. When AQI is added into the input parameters, the accuracy of SVM-2 models is improved to a certain extent compared with SVM-1 models. The performance improvement of the SVM-1-2 model is the most significant (see Fig. 2). Its R is increased 18.41%, NSE is increased 55.58%, and RMSE is decreased 13.32%. In Fig. 3, when clearness index is larger (greater than 0.7), it is a sunny day, and its air quality is good. At this time, the effect of fog and haze on solar radiation is small, and the corrective effect of AQI is not significant, which the difference between the calculated value of SVM-1-8 model and SVM-2-8 model is small. When clearness index is small (less than 0.2), it is characterized by overcast or cloudy days. The solar radiation is obscured by clouds in the upper air, and the space further modified by fog and haze near the surface is smaller, as a result, the corrective effect by AQI is poor, or even the accuracy is decreased. Based on the sample data, SVM algorithm seeks optimal solution between the best fitting ability and the best generalization ability, and it has a strong generality. In this paper, although the performance of SVM-1-8 is slightly worse than that of SVM-2-8, it is no larger fluctuations Fig. 4.

SVM models proposed in this paper are compared with existing models (including empirical models and neural network models), as shown in Table 2. In the existing models, The RMSE value is varied from 0.1944MJ·m⁻²·day⁻¹ for said model (relative sunshine duration based model) to 3.8 MJ·m⁻²·day⁻¹ for Ramedani model (temperature based model). The MAPE is distributed between 0.8 and 78.28 with mean value of 10.301, and The R is distributed between 0.7727 and 0.9835 with mean value of 0.907. The accuracy of SVM-1-8 and SVM-2-8 models recommended in this paper has been improved relative to existing models. Relative to the empirical models, the RMSE of SVM-2-8 is decreased from 2.028MJ·m⁻²·day⁻¹ to 1.18MJ·m⁻²·day⁻¹, R value is increased from 0.907 to 0.952, and MAPE is decreased from 0.907 to 0.992. In general, the SVM-2-8 model based on the SVM algorithm and adding AQI input parameters has a good applicability in Beijing.

Considering the time scale, daily value reflects the variation of solar radiation within a day, and it is more random than the monthly average daily value. So the accuracy of monthly average daily solar radiation models is generally higher than the daily models. From the aspect of input parameters, the performance of models which used sunshine hours as the main input parameters is
better than models which use other input parameter such as temperature, relative humidity, etc. From Table 1, SVM-1-4 model based on sunshine duration has good performance, and addition of other parameters (surface temperature difference, air temperature difference and relative humidity) can improve the accuracy of these models. If the effect of fog and haze on solar radiation can be considered, and AQI is used as an extra input parameter, the accuracy will be further improved.

### 4. Conclusion

The horizontal solar radiation plays a crucial role in the utilization of solar energy. In this paper, the effect of near-surface fog and haze on solar radiation is considered. Based on support vector machine, air quality index (AQI) is employed as an additional input parameter to establish daily global solar radiation model and compared with existing models. The conclusions are as follows:

#### Table 1

<table>
<thead>
<tr>
<th>Models</th>
<th>Input parameters</th>
<th>Train</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>NSE</td>
</tr>
<tr>
<td>SVM-1-1</td>
<td>Surface temperature difference</td>
<td>0.797</td>
<td>0.492</td>
</tr>
<tr>
<td></td>
<td>Air temperature difference</td>
<td>0.652</td>
<td>0.351</td>
</tr>
<tr>
<td></td>
<td>Relative humidity</td>
<td>0.770</td>
<td>0.463</td>
</tr>
<tr>
<td></td>
<td>Sunshine duration</td>
<td>0.932</td>
<td>0.871</td>
</tr>
<tr>
<td></td>
<td>Sunshine duration, surface temperature difference</td>
<td>0.942</td>
<td>0.891</td>
</tr>
<tr>
<td></td>
<td>Sunshine duration, air temperature difference</td>
<td>0.931</td>
<td>0.873</td>
</tr>
<tr>
<td></td>
<td>Sunshine duration, relative humidity</td>
<td>0.935</td>
<td>0.876</td>
</tr>
<tr>
<td>SVM-2-1</td>
<td>Surface temperature difference, AQI</td>
<td>0.845</td>
<td>0.555</td>
</tr>
<tr>
<td></td>
<td>Air temperature difference, AQI</td>
<td>0.782</td>
<td>0.458</td>
</tr>
<tr>
<td></td>
<td>Relative humidity, AQI</td>
<td>0.776</td>
<td>0.655</td>
</tr>
<tr>
<td></td>
<td>Sunshine duration, AQI</td>
<td>0.948</td>
<td>0.893</td>
</tr>
<tr>
<td></td>
<td>Sunshine duration, surface temperature difference, AQI</td>
<td>0.942</td>
<td>0.891</td>
</tr>
<tr>
<td></td>
<td>Sunshine duration, air temperature difference, AQI</td>
<td>0.952</td>
<td>0.899</td>
</tr>
<tr>
<td></td>
<td>Sunshine duration, relative humidity, AQI</td>
<td>0.943</td>
<td>0.871</td>
</tr>
<tr>
<td>SVM-2-2</td>
<td>Surface temperature difference, air temperature difference, relative humidity, sunshine duration</td>
<td>0.949</td>
<td>0.875</td>
</tr>
</tbody>
</table>

Fig. 2. The performance of different SVM models.
(1) To study the influence of different meteorological parameters on solar radiation, SVM algorithm is introduced to establish SVM-1 models. Among SVM-1 models, the SVM-1-8 model which considers sunshine duration, surface temperature difference, air temperature difference and relative humidity is the most accurate.

(2) In order to study the impact of fog and haze on solar radiation, SVM-2 models are established based on SVM-1 models with AQI. The comparison of these models shows that the
Table 2
Statistical analysis of proposed models in contrast with existing models.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Method</th>
<th>Station (location)</th>
<th>Time scale</th>
<th>RMSE (MJ/m²)</th>
<th>MAPE (%)</th>
<th>R</th>
<th>MPE (%)</th>
<th>Input parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>R. Said et al. [31]</td>
<td>Empirical</td>
<td>Tripoli, Libya</td>
<td>Monthly</td>
<td>0.1944</td>
<td>0.8</td>
<td>–</td>
<td>–</td>
<td>S, So</td>
</tr>
<tr>
<td>Muyyiu S. Adaramola [34]</td>
<td>Empirical</td>
<td>Akure, Nigeria</td>
<td>Monthly</td>
<td>0.9252</td>
<td>3.619</td>
<td>–</td>
<td>0.260</td>
<td>S, So</td>
</tr>
<tr>
<td>Empirical model</td>
<td>Akure, Nigeria</td>
<td>Monthly</td>
<td>1.3248</td>
<td>6.626</td>
<td>–</td>
<td>0.510</td>
<td>ΔTair</td>
<td></td>
</tr>
<tr>
<td>Empirical model</td>
<td>Akure, Nigeria</td>
<td>Monthly</td>
<td>1.6092</td>
<td>8.253</td>
<td>–</td>
<td>–0.986</td>
<td>RH</td>
<td></td>
</tr>
<tr>
<td>Empirical model</td>
<td>Akure, Nigeria</td>
<td>Monthly</td>
<td>1.1628</td>
<td>4.798</td>
<td>–</td>
<td>0.459</td>
<td>Tave</td>
<td></td>
</tr>
<tr>
<td>Yao et al. [35]</td>
<td>Empirical</td>
<td>Shanghai, China</td>
<td>Daily</td>
<td>3.704</td>
<td>78.28</td>
<td>–</td>
<td>55.221</td>
<td>S, So</td>
</tr>
<tr>
<td>Chen et al. [36]</td>
<td>SVM</td>
<td>ChaoYang, China</td>
<td>Daily</td>
<td>2.302</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>S, So</td>
</tr>
<tr>
<td>SVM</td>
<td>Dalian, China</td>
<td>Daily</td>
<td>1.801</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>SVM</td>
<td>Shenyang, China</td>
<td>Daily</td>
<td>2.003</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Ramedani et al. [37]</td>
<td>SVR_poly</td>
<td>Tehran province, Iran</td>
<td>Daily</td>
<td>3.44</td>
<td>–</td>
<td>0.887</td>
<td>T, N, S, So</td>
<td></td>
</tr>
<tr>
<td>R. Meenal et al. [38]</td>
<td>SVM</td>
<td>India</td>
<td>Monthly</td>
<td>1.3002</td>
<td>0.9443</td>
<td>–</td>
<td>–</td>
<td>S, So</td>
</tr>
<tr>
<td>NAD</td>
<td>India</td>
<td>Monthly</td>
<td>1.3846</td>
<td>–</td>
<td>0.9472</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>SVM India</td>
<td>India</td>
<td>Monthly</td>
<td>1.1617</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>ANN India</td>
<td>India</td>
<td>Monthly</td>
<td>1.4468</td>
<td>2.325</td>
<td>0.9739</td>
<td>2.325</td>
<td>S, So</td>
<td></td>
</tr>
<tr>
<td>ANN India</td>
<td>India</td>
<td>Monthly</td>
<td>2.2197</td>
<td>2.5079</td>
<td>0.7727</td>
<td>2.5079</td>
<td>T, S, So</td>
<td></td>
</tr>
<tr>
<td>Victor H. Quej et al. [39]</td>
<td>SVM</td>
<td>Yucatan Peninsula, Mexico</td>
<td>Daily</td>
<td>2.786</td>
<td>–</td>
<td>0.8142</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Masoud Vakili et al. [40]</td>
<td>ANN-1</td>
<td>Tehran, Iran</td>
<td>Daily</td>
<td>–</td>
<td>1.5</td>
<td>–</td>
<td>–</td>
<td>T, RH, WS, PM2.5</td>
</tr>
<tr>
<td>ANN-2 Tehran, Iran</td>
<td>Daily</td>
<td>–</td>
<td>3.04</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Current paper</td>
<td>SVM-1-8</td>
<td>Beijing, China</td>
<td>Daily</td>
<td>1.19</td>
<td>4.296</td>
<td>0.945</td>
<td>–</td>
<td>ΔTair, Tave, RH, S, So</td>
</tr>
<tr>
<td>SVM-2-8 Beijing, China</td>
<td>Daily</td>
<td>1.18</td>
<td>3.992</td>
<td>0.952</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

Accuracy of SVM-2 models is improved. Among them, the R value is promoted from 0.848 to 0.876, the NSE value is lifted from 0.682 to 0.740, the RMSE value is reduced from 0.114 to 0.102, and the MAPE value is decreased from 9.257 to 8.214. Compared with existing models, the RMSE value of SVM-2 is decreased from 2.028MJ m⁻² day⁻¹ to 1.18MJ m⁻² day⁻¹, R value is increased from 0.907 to 0.952, and MAPE value is decreased from 10.301 to 3.992.

Prospects: To study the influence of fog and haze on solar radiation, especially overcast or cloudy days, different neural network algorithms should be considered to explore the coupling influence of fog and different meteorological parameters on solar radiation.

Acknowledgements

The authors gratefully acknowledge the support of the National Natural Science Foundation of China (Grant Nos. 51508372 and 51506141), Natural Science Foundation of Tianjin (Grant Nos. 17JCYBJC22100 and 17JCTPJC52700), State Key Lab of Subtropical Building Science, South China University Of Technology (Grant No. 2017ZB16), Tianjin urban & rural construction commission science and technology project (Grant No. 2017-9).

References

daily diffuse solar radiation models modified by air quality index (AQI) in the region with heavy fog and haze, Energy Convers. Manag. 139 (2017) 140–150.


