Energy Conversion and Management 89 (2015) 318–329

Contents lists available at ScienceDirect

Energy Conversion and Management

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

Assessing the transferability of support vector machine model for estimation of global solar radiation from air temperature

Ji-Long Chen [a,b,c,d], Guo-Sheng Li [c,d], Bei-Bei Xiao [e,f,g], Zhao-Fei Wen [a,b], Ming-Quan Lv [a,b], Chun-Di Chen [a,b], Yi Jiang [a,b], Xiao-Xiao Wang [a,b], Sheng-Jun Wu [a,b,*]

[a] Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing 401122, China
[b] Key Laboratory on Water Environment of Reservoir Watershed, Chinese Academy of Sciences, Chongqing 401122, China
[c] Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
[d] Laboratory of Land Surface Pattern and Simulation, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, China
[e] Chongqing Institute of Surveying and Planning for Land Resources and Housing, Chongqing 400020, China
[f] National Engineering Technology Research Center for Remote Sensing Applications, Chongqing Research Center, Chongqing 400020, China
[g] Chongqing Xinrong Institute of Surveying Technology for Land and Housing, Chongqing 400020, China

ARTICLE INFO

Article history:
Received 3 July 2014
Accepted 1 October 2014

Keywords:
Global solar radiation
Support vector machine
Air temperatures
Distance
Altitude

ABSTRACT

Exploring novel methods for estimation of global solar radiation from air temperature has been being a focus in many studies. This paper evaluates the transferability of support vector machines (SVM) for estimation of solar radiation in subtropical zone in China. Results suggest that solar radiation at one site (estimation site) could be well estimated by SVM model developed at another site (source site). The accuracy of estimation is affected by the distance and temperature difference between two sites, and altitude of source site. Higher correlations between RMSE of SVM and distance, and temperature differences are observed in northeastern region, increasing the reliability and confidence of SVM model developed at nearby stations. While lower correlations between RMSE and distance, and temperature differences are observed in southwest plateau region. When the altitude of estimation site is lower than 1200 m, RMSE show logarithm relationship with altitude of source sites where the altitude are lower than that of estimation site. Otherwise, RMSE show linearly relationship with altitude of source sites where the altitude are higher than 200 m but lower than that of the estimation site. This result suggests that solar radiation could be also estimated using SVM model developed at the site with similar but lower altitude. Based on these results, a strategy that takes into account the climatic conditions, topography, distance, and altitude for selecting a suitable source site is presented. The findings can guide and ease the appropriate choice of source sites for estimation of solar radiation at estimation site.

1. Introduction

The importance of global solar radiation in ecology, agriculture, environment and the associated researches has been well documented [1–3]. However, due to the cost of measuring equipment and its difficult maintenance and calibration, lack of sufficient solar radiation data has been reported in many countries like USA [4,5], United Kingdom [6], Egypt [7], India [8], France [9], Greece [10], Italy [11] and China [12–14]. On the contrary, air temperature is routinely measured at most meteorological stations. In this context, great efforts have been made to estimate solar radiation using air temperature [15,16]. Hargreaves and Samani [15] proposed a model (H–S model) using air temperature range, several modified versions were developed and validated in many places around the world [17–20]. Although some authors claimed that their new versions outperformed the original model, this may not always be the cases in many studies even when other commonly measured meteorological variables are taken into account [4,13,21,22]. Bristow and Campbell [16] developed a model (B–C model) as exponential function of temperature range. Numerous modifications centered on tuning the parameters of B–C model have been made [23–25]. However, many evaluations suggested that such modifications were generally not effective and yielded little or no improvement [26–28].

Accurate estimation of solar radiation has been being a major goal for solar energy practitioners, climatologists and all concerned scientists [29]. It seems that H–S, B–C models and their revised

* Corresponding author at: Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing 401122, China. Tel.: +86 023 6593 5918.
E-mail address: chenjilong@cigit.ac.cn (S.-J. Wu).

http://dx.doi.org/10.1016/j.enconman.2014.10.004
0196-8904/© 2014 Elsevier Ltd. All rights reserved.

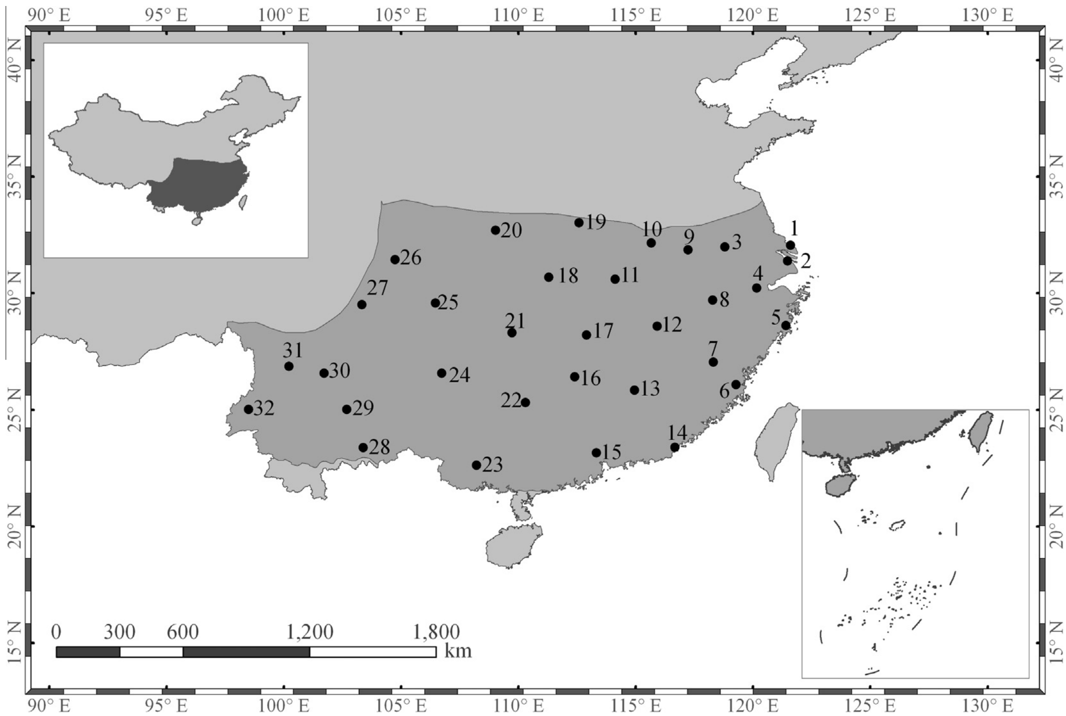

Fig. 1. Location of the studied meteorological stations in China's subtropical zone (stations are numbered in compliance with Table 1).

versions, as well as other empirical models, have far overreached their predictive limits. Therefore, exploring new methods has became the focus of many studies in recent years [30–39].

Recently, as the development of computational technology and sophisticated statistical methods, a novel machine learning method, support vector machine (SVM), has been widely applied in computer [40], environment [41,42] and hydrology researches [43], and proven to be a robust algorithm in classification [44], regression [45,46] and forecasting [47]. Many studies proved that SVM is superior to Neural Network and traditional statistical models [48–50]. Despite successes in many fields, there is not any application in estimation of solar radiation until our work investigated the feasibility of SVM using air temperatures at Chongqing, China [3]. While it is still needed to research further to solve the problem how to apply SVM model to the site where no solar radiation data is available. One possible alternative is to substitute training data from nearby representative meteorological stations. It was proposed by Allen et al. [51] as a possible way of replacing missing data when calculating daily evapotranspiration. However, there is lack of explicit guidance on appropriate choice of the representative stations, which may result in arbitrariness in choosing a source site that provides data for constructing SVM model for solar radiation estimation at the estimation site. Thus, the transferability of SVM model is investigated.

The present study is carried out in China's subtropical zone which plays an important role in ecological, agricultural and climate researches in the world. However, there are only 32 meteorological stations measuring global solar radiation. On the contrary, more than 700 stations have records of air temperatures, offering an important alternative to deriving solar radiation due to the widely availability of temperature data. Some works validated the temperature-based models for this area [52–54], and a few revised models were developed [13,52]. However, these new versions were found to give similar performances with the original ones at many sites [13,52,55]. Apart from model accuracy, one major limitation is that calibration of those empirical models is constrained by unavailability of solar radiation data, limiting applications of those models to other site where solar radiation data is not available. Therefore, exploring model developed at source site for estimating solar radiation at estimation site is of vital importance and significance, not only for China's subtropical zone but also for other regions. The main objectives of this study

Table 1
Detail information of the studied meteorological stations.

Site ID	Site name	Longitude (E)	Latitude (N)	Altitude (m)
1	Lvshi	121.60	32.07	5.5
2	Shanghai	121.48	31.40	6.0
3	Nanjing	118.80	32.00	7.1
4	Hangzhou	120.17	30.23	41.7
5	Hongjia	121.42	28.62	1.3
6	Fuzhou	119.28	26.08	84.0
7	Jianou	118.32	27.05	154.9
8	Tunxi	118.28	29.72	142.7
9	Hefei	117.23	31.87	27.9
10	Gushi	115.67	32.17	57.1
11	Wuhang	114.13	30.62	23.1
12	Nanchang	115.92	28.60	46.7
13	Ganzhou	114.95	25.85	123.8
14	Shangtou	116.68	23.40	2.9
15	Guangzhou	113.33	23.17	41.0
16	Changling	112.40	26.42	116.6
17	Changsha	112.92	28.22	68.0
18	Yichang	111.30	30.70	133.1
19	Nanyang	112.58	33.03	129.2
20	Ankang	109.03	32.72	290.8
21	Jishou	109.73	28.32	208.4
22	Guiling	110.30	25.32	164.4
23	Nanling	108.22	22.63	121.6
24	Guiyang	106.73	26.58	1223.8
25	Chongqing	106.47	29.58	259.1
26	Mianyang	104.75	31.45	486.3
27	Emeishan	103.33	29.52	3047.4
28	Mengzhi	103.38	23.38	1300.7
29	Kunming	102.68	25.02	1892.4
30	Panzhihua	101.72	26.58	1190.1
31	Lijiang	100.22	26.87	2392.4
32	Tengchong	98.50	25.02	1654.6