Energy Conversion and Management 154 (2017) 168–179

Contents lists available at ScienceDirect

Energy Conversion and Management

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

Classification of solar radiation zones and general models for estimating the daily global solar radiation on horizontal surfaces in China

Yanfeng Liu*, Yong Zhou, Dengjia Wang, Yingying Wang, Yong Li, Ying Zhu

School of Environmental and Municipal Engineering, Xi'an University of Architecture and Technology, No. 13 Yanta Road, Xi'an 710055, China

ARTICLE INFO

Keywords:
Two-step solar radiation zoning method
Sunshine duration
Regional global solar radiation estimation
General model

ABSTRACT

Classification of solar radiation zones constitutes the prerequisite for the establishment of regional daily global solar radiation (H) estimation general model. Current zone establishment methods are ordinary based on solar radiation observation stations (SROS) which present a sparse and non-uniform distribution. As a result, the possibility of misclassifications of the stations occurs in cases there are no records of radiation and far away from SROS. Therefore, by using k-means cluster and Support Vector Machine-Genetic Algorithm, a two-step radiation zoning method was proposed in this paper according to: (a) H, sunshine duration, temperature and relative humidity from 98 SROS and (b) sunshine duration, temperature and relative humidity from 562 stations without radiation. The method is capable to combine the SROS and the stations without radiation in the process of classification. Thus, these misclassifications have been effectively reduced and the accuracy of each classification has been significantly improved. Based on the method, five radiation zones have been identified. Concurrently, four sunshine-based models were obtained for each SROS and the analysis of statistical indexes indicated that the cubic models presented the best performances in each zone. According to the best site-specific models and radiation zones, the general models of regional H estimation were developed by introducing the geographical parameters, including latitude and altitude. The comparative results demonstrated that the general models proposed in this paper had better accuracies and can represent the general models for the H estimation of stations without radiation records in China.

1. Introduction

During the last decade, the rapid economic development and steady improvement of living standards increased rapidly China's energy consumption. In 2012, China's total primary energy consumption reached to 3.62 billion coal equivalents, which equals to 1.5 times as much as that in 2000 [1]. Energy has become a significant problem that affects China's economy and environment [2]. Solar energy, representing a safe, economic, environmental and renewable energy form [3], is regarded as the alternative energy to improve the energy structure of China. Solar radiation (H) data correspond to the basic data of various solar energy utilization technologies [4]. Compared to temperature and sunshine duration (S) parameters, the distribution of SROS is generally much sparser and non-uniform, resulting in the lack of H in most of stations [5,6]. Thus, it is difficult to meet the requirement of solar energy utilization data. At present, there are 756 surface meteorological observation stations (SMOS), where the weather data (including daily dry-bulb temperature, relative humidity, wind velocity, pressure and rainfall) are provided by China meteorological stations. Only 122 of them measured and recorded H [7]. Therefore, it is particularly critical and essential to seek for a method to accurately estimate the daily H on horizontal surfaces in China.

The recent research has ordinary adopted indirect statistical methods to estimate the H; that is, according to the correlation between H and other meteorological parameters, such as S, temperature, cloud, rainfall and fog, empirical formulas were established [8], including sunshine [9–13], temperature [14–18] and multi-parameter [19–23] based models. Nevertheless, the majority of these models are site-dependent and it is questionable if an application to other stations without records of H, can occur. Therefore, to estimate the daily H of these stations without H in different areas, it is necessary to establish the general models of regional daily H estimation. Zhou et al. [24], introduced the geographical parameters to establish the general models for estimating the monthly mean daily H, by using the data from mainland city of China. But the models were useless in estimating the daily H. Liu et al. [25], established two general models for estimating the daily H in China, based on temperature and S data. The temperature-based general models proved to be less accurate, while during the

* Corresponding author.
E-mail address: lyfxjd@163.com (Y. Liu).

http://dx.doi.org/10.1016/j.enconman.2017.10.043
Received 24 July 2017; Received in revised form 29 September 2017; Accepted 15 October 2017

Nomenclature

Symbols

b	scalar
C	custom penalty parameter
h	altitude of stations (m)
H	global solar radiation (MJ/m^2 d)
H_0	extraterrestrial horizontal radiation (MJ/m^2 d)
$K(x_i,y_i)$	Kernel function
l	number of training data set
S	sunshine duration (h)
S_0	maximum possible sunshine duration (h)
T_a	average temperature (°C)
T_{max}	maximum temperature (°C)
T_{min}	minimum temperature (°C)
ΔT	temperature diurnal(°C)
X_i	the attribute variables of data set i
$Y_{i,m}$	measured value
$Y_{i,c}$	calculated value

Greek symbols

α	Lagrange multiplier
λ	latitude of stations (°)
ξ	slack variables
ω	normal vector
σ	Gaussian noise level of standard deviation
$\varphi(x)$	attribute matrix of the input space vector x

Abbreviations

GA	genetic algorithm
GMOS	general meteorological observation stations
MABE	mean absolute bias error
RBF	radial basis function
RMSE	root mean square error
rRMSE	relative root mean square error (%)
RH	relative humidity (%)
SMOS	surface meteorological observation stations
SROS	solar radiation observation stations
SVM	support vector machine
SVM-C	support vector machine for classification

establishment of the sunshine-based general models, only Angstrom-Prescott model has been analyzed. Mecibah [26], proposed two general models for Algeria based on the sunshine-based models in quadric and cubic forms. However, the models were established based on the observation data of Algeria and could not be adopted in the case of China. Therefore, it is necessary to establish general models of regional daily H estimation suited for China including high accuracy to estimate the daily H for the stations where no records of H exist.

Zoning solar radiation, comprehension of the solar radiation distribution and classification of the regions with similar solar radiation change trends are the prerequisites for the establishment of regional daily H estimation general models. At present, the research on the classification of solar radiation zones focused on the selection of different zoning criteria [25,27–30]. As an example, based on the data of 1966–1975 from 340 SROS, Commission of the European Communities developed the map of European solar radiation using monthly mean daily H and S in 1984 [31]. Lau et al. [32] divided China into five zones based on monthly mean daily S, by using data from 123 SROS in 2007. Joseph et al. [33] proposed to take into account the outdoor temperature, radiation and air velocity as zoning criteria, and divided Madagascar into three radiation zones in 2009. These zoning methods were based on the observation data from SROS. Initially, the SROS are classified into groups. Subsequently, the areas near to the SROS are divided into the same groups. The above radiation zoning methodology is based on the First Law of Geography, namely, “all attribute values on a geographic surface are related to each other, but closer values are more strongly related than are more distant ones” [34]. The method is simple and easy to operate. Nevertheless, a significant limitation is that the accuracy cannot be guaranteed.

At present, the distribution of SROS is ordinarily sparse and non-uniform. Based on the observation data from SROS, may lead to misclassification problems for the areas far away from SROS, due to the reduction of the spatial similarities of solar radiation. The problem becomes more intense as it concerns the areas around the junctions of two zones, resulting in high errors in the estimation of the daily H for the stations without H. However, past research on the reduction of these misclassifications is poor. Therefore, it is essential to propose a method which could effectively reduce these misclassifications and obtain accurate classification results. Nowadays, the distribution of the general meteorological observation stations (GMOS) (since no records of H exist) is much denser than that of SROS. The misclassification issues could be effectively improved by the introduction of GMOS in the process of classification of radiation zones. According to Mohammadi et al. [35] and Olatomiwa et al. [36], the daily H can be predicted by adopting Support Vector Machine (SVM) using daily sunshine rate, daily average temperature (T_a), daily temperature diurnal (ΔT). The results proved to be satisfactory. This pronounces that the SVM could effectively describe the correlation between H and other meteorological parameters. Thus, the SVM could be an effective method, which could introduce the GMOS in the process of classification, to achieve the misclassification reduction.

Therefore, the main objectives of this paper contain the following parts. Firstly, a solar radiation zoning method based on the SVM should be proposed. This method could combine the SROS and GMOS to reduce the misclassifications of the areas far from SROS, especially at the areas adjacent to the junctions of two zones. Furthermore, it should identify different solar radiation zones by the monthly mean daily H. Secondly, the site-specific daily H estimation models should be established and validated by the S and H from the 98 meteorological stations in China and the best models for each solar radiation zone should be obtained. Finally, the general models of regional H estimation should be developed based on the best models to estimate the daily H for different solar radiation zones in China.

2. Methods of solar radiation zoning

Solar radiation zoning is a prerequisite for the establishment of general models of regional daily H estimation. Based on the k-means clustering and Support Vector Machine-Genetic Algorithm (SVM-GA), a two-step solar radiation zoning method is proposed. To distinguish the change trend of solar radiation in different zones, the monthly mean daily H was selected as the zoning criteria. The procedure of two-step method is as follows:

(a) Based on the monthly mean daily H of 98 SROS, five clusters of SROS were obtained by k-means clustering.
(b) By using the SVM-GA, the monthly mean daily S, monthly mean daily extraterrestrial radiation (H_0), monthly mean daily T_a, monthly mean daily ΔT and monthly mean daily relative humidity (RH) of SROS were adopted as the training data set, while the