Estimation of monthly-mean daily global solar radiation based on MODIS and TRMM products

Jun Qin\textsuperscript{a,}\textsuperscript{*}, Zhuoqi Chen\textsuperscript{b}, Kun Yang\textsuperscript{a}, Shunlin Liang\textsuperscript{c}, Wenjun Tang\textsuperscript{a}

\textsuperscript{a}Laboratory of Tibetan Environment Changes and Land Surface Processes, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, P.O. Box 2871, Beijing 100085, China
\textsuperscript{b}College of Global Change and Earth System Science, Beijing Normal University, Beijing 1000875, China
\textsuperscript{c}Laboratory of Tibetan Environment Changes and Land Surface Processes, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, P.O. Box 2871, Beijing 100085, China

A R T I C L E   I N F O

Article history:
Received 27 September 2010
Received in revised form 24 November 2010
Accepted 6 January 2011
Available online 5 February 2011

Keywords:
Global solar radiation
Artificial neural network
Remote sensing
MODIS
TRMM

A B S T R A C T

Global solar radiation (GSR) is required in a large number of fields. Many parameterization schemes are developed to estimate it using routinely measured meteorological variables, since GSR is directly measured at a limited number of stations. Even so, meteorological stations are sparse, especially, in remote areas. Satellite signals (radiance at the top of atmosphere in most cases) can be used to estimate continuous GSR in space. However, many existing remote sensing products have a relatively coarse spatial resolution and these inversion algorithms are too complicated to be mastered by experts in other research fields. In this study, the artificial neural network (ANN) is utilized to build the mathematical relationship between measured monthly-mean daily GSR and several high-level remote sensing products available for the public, including Moderate Resolution Imaging Spectroradiometer (MODIS) monthly averaged land surface temperature (LST), the number of days in which the LST retrieval is performed in 1 month, MODIS enhanced vegetation index, Tropical Rainfall Measuring Mission satellite (TRMM) monthly precipitation. After training, GSR estimates from this ANN are verified against ground measurements at 12 radiation stations. Then, comparisons are performed among three GSR estimates, including the one presented in this study, a surface data-based estimate, and a remote sensing product by Japan Aerospace Exploration Agency (JAXA). Validation results indicate that the ANN-based method presented in this study can estimate monthly-mean daily GSR at a spatial resolution of about 5 km with high accuracy.

© 2011 Elsevier Ltd All rights reserved.

1. Introduction

The sun provides energy for almost all land surface processes on the earth’s surface. GSR data are required in a large number of applications. In ecological modeling, the photosynthetically active part of GSR is used in the photosynthesis and stored as the chemical energy for allocation to different living organs of plants such as root, stem, and foliage. GSR is an indispensable component in land surface models and hydrological models, which controls both water and heat exchanges. The knowledge of GSR distribution is also important for the design of the photovoltaic power and solar heating systems in the clean energy area. Especially, GSR plays an irreplaceable part in crop growth modeling due to its tight relation to photosynthesis. For many crop growth models such as WOFOST (Model for WOrld FoOd StUdies) and DSSAT (Decision Support System for Agrotechnology Transfer), monthly-mean GSR values plus other meteorological parameters, such as air temperature and rainfall, are used to generate daily corresponding ones and then imported into the model through the weather generator [1,2].

Monthly-mean GSR values are also important input for plant distribution models [3].

Despite the great importance of GSR, the number of radiation stations is limited, compared to the number of stations where routine meteorological variables are collected, such as air temperature, humidity, sunshine duration, and cloud coverage [4–7]. Thus, many empirical parameterization schemes are designed to indirectly estimate GSR from these routine variables [8]. These methods could be roughly divided into three classes in accordance with inputs: sunshine-based, temperature-based, and cloud-based. Based on mathematical forms, they could be classified as parametric (such as Ångström–Prescott type methods) [9,10] and non-parametric (such as ANN-based methods) [11–14]. If calibrated well, they (parametric and non-parametric) can estimate GSR with a high accuracy. However, since no explicit physics are considered in these estimation schemes, calibrated parameters often vary from one site to the other, limiting their generalization and leading to large uncertainties in GSR estimates at sites where no measured GSR data are available to calibrate these schemes [15].

Remote sensing (RS) provides an alternative method to radiation estimation [16–19]. One advantage of RS is its ability to collect continuous signals in space and time at the top of atmosphere. This
could be illustrated in Fig. 1. On the left panel of this figure, the network (MLP) is used to realize the fitting ability of ANN, which units organized in layers. In this study, a multilayer perceptron presents an ANN-based method to directly build the functional relationship.

## 2. Method and data

### 2.1. Method

ANN can be used to fit a function, recognize patterns, and cluster data [21]. It is typically composed of interconnected neuron units organized in layers. In this study, a multilayer perceptron network (MLP) is used to realize the fitting ability of ANN, which could be illustrated in Fig. 1. On the left panel of this figure, the schematic diagram of a neuron unit is displayed and its function can be mathematically abstracted as:

\[ y = g \left( \sum_{i=1}^{n} w_i x_i \right) \]  

where \( i \) denotes the index, \( x_i \) one component of the input vector, \( w_i \) the weight for each \( x_i \), \( g(.) \) the transfer function which can take many function forms such as linear, sigmoid, and hyperbolic tangent. On the right panel of Fig. 1, a MLP with three layers is illustrated and its mathematical formula can be written in a compact form as:

\[ Y = f(X, W) \]  

where \( X \) represents the input vector, \( Y \) the simulated output vector by ANN, \( W \) all the connection weights between network neuron units, and \( f(.) \) denotes functional relationship between the input vector and the output vector.

Its weights \( W \) need to be determined before ANN can be used to estimate or predict. The process of determining the weights according to observations is called training or learning. This process is realized by adjusting the weights in order to minimize the following cost function:

\[ J = \frac{1}{N} \sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2 (Y_i - \hat{Y}_i) \]  

where \( i \) denotes the index, \( Y_i \) the simulated output vector (simulated observations) by ANN, \( \hat{Y}_i \) the observed output vector, \( N \) is the number of input/output pairs, and \( J \) the sum of differences between these pairs. There are many optimization algorithms to realize the minimization of \( J \). The whole training process can be illustrated in Fig. 2. However, the over-fitting problem may be incurred and the ANN generalization quality will weaken if the cost-function form like in Eq. (3) is taken during the learning process.
In the current process of training, the following cost-function form is taken [22]:

\[
J = \gamma J_t + \left(1 - \gamma \right) J_w
\]

\[
J = \frac{1}{N} \sum_{i=1}^{N} (Y_i - Y_i^*)^T (Y_i - Y_i^*)
\]

\[
J_w = \frac{1}{M} W^T W
\]

where \(J_w\) is a regularization term used to make the response of ANN smoother and less likely to overfit, \(\gamma\) is an adjusted parameter, and \(M\) denotes the number of weights. In this study, the Levenberg–Marquardt back-propagation optimization algorithm coupled with Bayesian regularization is used to train the ANN based on the cost function Eq. (4). This scheme can adjust the value of \(\gamma\) automatically. In order to indicate the advantage of this training procedure, another optimization algorithm, the scaled conjugate gradient method, is also utilized to train the same ANN according to the cost function Eq. (3).

It is proven that a multilayer feed-forward network with a single hidden layer has the capability to approximate a function to any accuracy [23]. Thus, this ANN architecture is chosen here. However, both the optimal number of neurons and the type of transfer function in the hidden layer need to be determined through the trial-and-error procedure. In this work, two types of transfer function (i.e., hyperbolic tangent sigmoid function and radial basis function) and different numbers of neurons in the hidden layer are tested. Moreover, the training results are sensitive to the initial weights, which are usually assigned randomly. In this study, the training course is performed twenty times to avoid this effect. All of these ANN tasks are performed using MATLAB neural network toolbox.

There is one thing worthy of being noticed. It depends on two points whether or not ANN performs satisfactorily when being used to estimate or predict. On the one hand, the physical relationship must exist between input parameters and corresponding output ones; on the other hand, the chosen structure of ANN must have the ability to reflect the physical relationship between inputs and outputs. If these two requirements are not met, ANN cannot work well.

2.2. Study area and data

The Tibetan Plateau (TP) and its surroundings are chosen as the study area. The TP is the highest and largest plateau in the world, with an average elevation of 4000 m above sea level (ASL) and an area of about 2.5 million square kilometers. As shown in Fig. 3, the study area extends from 25°N to 45°N and 70°E to 105°E. There, the elevation range is drastic and the terrain is complex. Moreover, the amounts of cloud and precipitation vary significantly from the east to the west and from the south to the north. All of above factors have great impacts on GSR. The possibility of generalizing it to other areas, therefore, is high, if some GSR estimation scheme can work well in this region [24].
Due to the harsh climate in this region, merely 34 radiation stations are deployed by China Meteorological Administration (CMA) to directly measure GSR, which are illustrated in Fig. 3. The GSR measuring activity at Chengdu station was moved to Wenjiang station in 2004 and the one at Lanzhou station was moved to Yuzhong station in 2005. As Shi et al. [25] pointed out, the errors in radiation measuring instruments used at CMA weather stations do not exceed 5% after the year of 1990. However, GSR measurements may contain uncertainties larger than 5% due to many problems other than instruments such as inappropriate measuring manipulations and occasional voltage instability. Thus, quality control of these measurement data has been made according to steps presented by Tang et al. [26]. These steps could be roughly divided into two main procedures. In the first place, daily GSR measurements, which do not meet some hard physical conditions such as exceeding the theoretical clear-sky GSR values, need removing. In the second place, at each radiation site, the functional relationship is constructed by ANN between measured GSR values and six routinely measured meteorological parameters, which include daily air temperature range, daily mean temperature, relative humidity, sunshine duration, precipitation, and relative optical air mass. After this, three sub-steps are performed: first, the scatter-plot of the estimated versus the observed is created; second, the fitting line and the 95% confidence interval lines are calculated; third, the GSR data is considered as the suspected data and then is excluded if its value is outside of the 95% confidence interval line. More details on GSR data quality control can be found in Tang’s article.

In this study, monthly-mean daily GSR data from 2001 to 2007 at 22 stations are used in the training process as input vectors and 5 years of GSR data from 2003 to 2007 at the remaining 12 stations are applied in the process of validation. Both the training and validation stations distribute over wide ranges of latitude, longitude, and altitude in order to ensure the representativeness of the estimation algorithm presented in this study.

As mentioned above, RS products are used as inputs to train the ANN for building their mathematical relationship with measured GSR. Thus, these RS products must directly or indirectly contain information that may affect GSR. Since both diurnal temperature range (DTR) and precipitation at weather stations are correlated with the amount of cloud and thus have a fairly large impact on GSR at the earth’s surface, they are often chosen as independent variables in order to estimate GSR in many previous studies [6,27,28]. Thus, MODIS monthly averaged daytime/nighttime LSTs [29] and TRMM monthly accumulated precipitation [30] are chosen as one part of inputs to train ANN. However, the spatial resolutions of MODIS and TRMM products are 0.05° and 0.25°, respectively. Thus, TRMM data are spatially interpolated to 0.05° in order to match with MODIS products. In this study, MODIS daytime and nighttime LSTs need to be transformed as:

\[ \Delta T = T_d - T_n \]

and

\[ T_m = \frac{T_d + T_n}{2} \]

where \( T_d [K] \) denotes the daytime LST, \( T_n [K] \) the nighttime LST, \( \Delta T [K] \) the difference between the daytime and nighttime LSTs, \( T_m [K] \) the mean LST.

As well known, the ground surface conditions have an effect on GSR through the multiple-scattering between surface and the atmosphere, so MODIS monthly EVI [31] is selected as one input to ANN since the high value of EVI normally corresponds to small surface albedo and then decreases the magnitude of multiple-scattering between the ground and the atmosphere, and vice versa. Besides, the number of days (NOD), in which the LST retrieval is
performed in 1 month, is normalized by the number of days in the corresponding month and then taken as one input. This quantity can reflect information on the mean amount of cloud cover in 1 month because MODIS LST is not retrieved at a cloudy pixel. Moreover, the relative optical air mass has a great influence on GSR at the ground surface. Where the altitude is higher and thus the relative optical air mass is smaller, the radiation from sun is less attenuated, and vice versa. This effect has to be taken into account by using the relation between the atmospheric pressure and the altitude and can be represented as:

$$\tau = 10^{\left(-\frac{z}{18400} - \frac{t_m}{273}\right)}$$

where $\tau$ denotes the ratio of the local air pressure and the one at sea level, $z$ [m] the altitude ASL, and $t_m$ [K] the mean LST. The entire input vector to ANN can be represented as $X = [d, t_m, p, v_i, g, s]^T$ in which $p$ [mm] denotes precipitation, $v_i$ the EVI, and $g$ the NOD.

The ground GSR depends not only on the extraterrestrial GSR, but also on the atmospheric status. Since the extraterrestrial GSR can be analytically calculated with high accuracy, the ground GSR could be completely determined once the atmospheric condition is known. Therefore, the clearness index, which represents the atmospheric status, is selected as the output vector and defined as:

$$\kappa = \frac{H}{H_0}$$

where $H$ [MJ day$^{-1}$] denotes GSR at the ground surface and $H_0$ [MJ day$^{-1}$] is the extraterrestrial GSR. So, the output vector can be represented as $Y = [\kappa]^T$. Once $\kappa$ is determined, GSR can be obtained easily according to Eq. (7).

---

Table 1

<table>
<thead>
<tr>
<th>Number of neurons in hidden layer</th>
<th>Bayesian regularization</th>
<th></th>
<th></th>
<th>Scaled conjugate gradient</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Best</td>
<td>Worst</td>
<td>Best</td>
<td>Worst</td>
<td>Best</td>
<td>Worst</td>
</tr>
<tr>
<td>2 neurons</td>
<td>1.30</td>
<td>1.35</td>
<td>1.30</td>
<td>1.36</td>
<td>1.34</td>
<td>1.75</td>
</tr>
<tr>
<td>4 neurons</td>
<td>1.13</td>
<td>1.20</td>
<td>1.13</td>
<td>1.29</td>
<td>1.23</td>
<td>1.47</td>
</tr>
<tr>
<td>6 neurons</td>
<td>1.09</td>
<td>1.14</td>
<td>1.09</td>
<td>1.22</td>
<td>1.19</td>
<td>1.43</td>
</tr>
<tr>
<td>8 neurons</td>
<td>1.08</td>
<td>1.12</td>
<td>1.07</td>
<td>1.11</td>
<td>1.14</td>
<td>1.42</td>
</tr>
<tr>
<td>10 neurons</td>
<td>1.03</td>
<td>1.09</td>
<td>1.05</td>
<td>1.10</td>
<td>1.14</td>
<td>1.30</td>
</tr>
<tr>
<td>12 neurons</td>
<td>1.04</td>
<td>1.15</td>
<td>1.01</td>
<td>1.08</td>
<td>1.13</td>
<td>1.30</td>
</tr>
<tr>
<td>14 neurons</td>
<td>0.99</td>
<td>1.10</td>
<td>1.03</td>
<td>1.05</td>
<td>1.12</td>
<td>1.24</td>
</tr>
<tr>
<td>16 neurons</td>
<td>1.01</td>
<td>1.14</td>
<td>1.01</td>
<td>1.13</td>
<td>1.07</td>
<td>1.28</td>
</tr>
<tr>
<td>18 neurons</td>
<td>1.01</td>
<td>1.14</td>
<td>1.00</td>
<td>1.14</td>
<td>1.07</td>
<td>1.27</td>
</tr>
<tr>
<td>20 neurons</td>
<td>1.01</td>
<td>1.12</td>
<td>1.01</td>
<td>1.12</td>
<td>1.07</td>
<td>1.17</td>
</tr>
</tbody>
</table>

$^a$ Tansig denotes hyperbolic tangent sigmoid function. $^b$ Radbas denotes radial basis function.
Fig. 6. Individual validation results for all 12 validation sites. They are Yining, Kashi, Ruoqiang, Hami, Dunhuang, Golmud, Shiquanhe, Yushu, Mianyang, Lijiang, Panzhihua, and Kunming from (a) to (l).
The elevation range over the TP region and its surroundings is rather large as shown in Fig. 3. Since the highest radiation station (Nagqu) is located at about 4500 m ASL, the altitudinal extension of the trained ANN may be questionable even though Nagqu site is included into the training dataset, as there is a large fraction of the Plateau where elevation is more than 4500 m ASL. So, some virtual input/output pairs are generated in order to increase the applicability of the ANN-based method in the vertical dimension. First, pixel points, which have high elevations (above 6800 m ASL), little precipitation (less than 10 mm month\(^{-1}\)), and high NOD (greater than 28 days per month), are extracted from remote sensing images. These pixels correspond to clear skies and thus large clearness index values. Second, GSR values are generated at these pixels by running Yang's hybrid model for clear-skies. Finally, RS data and the corresponding simulated clear-sky GSR values at these pixels are added into the training dataset.

### 3. Results and discussion

There are a total of 1678 input/output pairs at 22 stations from 2001 to 2007. These pairs are used to train ANN to build the relationship between the clearness index (the output vector with only one component) \(T = [\epsilon]^T\) and the input vector \(X = [d_t, \theta_{m}, p, v_t, \eta, \tau]^T\). After training, a comparison is performed between simulated GSR and observed ones at the training stations in order to examine the training performance. Correlation coefficient \(R\), root mean squared error (RMSE) and relative root mean square error (RRMSE) are used as error metrics for comparison. As mentioned in Section 2.1, the different configurations and learning algorithms are tested in order to select the relatively optimal neural network architecture. Both the best and the worst performances in the twenty training runs, which are measured according to RMSE, are shown for each ANN configuration in Table 1. As can be seen, the Bayesian regularization training algorithm, in which the Levenberg–Marquardt method is used, is always superior to the scaled conjugate gradient algorithm in all configurations. It is also indicated that the ANN configurations behave similarly no matter which transfer function is used. Moreover, the training results are hardly improved further when the number of neurons in the hidden layer is greater than ten. Based on the previous discussion, the ANN configuration with 12 neurons and hyperbolic tangent sigmoid function in the hidden layer is applied in the following discussion.

As shown in Fig. 4, error metrics show that ANN can grasp the functional relationship between GSR and its independent variables. The above examination does not guarantee that the built ANN can be applied to estimate GSR on the regional scale. In order to achieve this goal, the ANN needs to be evaluated at sites where GSR observations are not used in the training process. Moreover, the ANN-based GSR estimate also needs comparing with those from the existing satellite products and other estimation methods to indicate that this ANN-based method can estimate GSR both effectively and efficiently. So JAXA monthly-mean daily GSR products with a spatial resolution of 0.25° based on MODIS sensor and GSR estimates derived from Yang's hybrid method are chosen in this study. A total of 12 validation sites, where GSR observations are left for validation, are illustrated in Fig. 3. As shown in Fig. 5, the ANN-based method presented in this study and Yang’s method both perform better than the retrieval algorithm by JAXA. It is also shown in Fig. 5 that Yang’s hybrid model estimates GSR better than the ANN-based method does at small GSR values, but the reverse situation occurs at large GSR values. Moreover, the hybrid model can only work at stations where sunshine data are available.

### Table 2 Performance of three estimation methods at the validation sites.

<table>
<thead>
<tr>
<th>Site</th>
<th>ANN RMSE (R)</th>
<th>Yang RMSE (R)</th>
<th>JAXA RMSE (R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yining</td>
<td>1.24 0.98</td>
<td>1.87 0.99</td>
<td>1.86 0.96</td>
</tr>
<tr>
<td>Kashi</td>
<td>1.24 0.98</td>
<td>2.55 0.98</td>
<td>1.85 0.98</td>
</tr>
<tr>
<td>Ruqiang</td>
<td>1.61 0.98</td>
<td>1.63 0.99</td>
<td>1.55 0.99</td>
</tr>
<tr>
<td>Hami</td>
<td>1.34 0.99</td>
<td>0.93 0.99</td>
<td>1.55 0.99</td>
</tr>
<tr>
<td>Dunhu</td>
<td>0.86 0.99</td>
<td>0.89 0.99</td>
<td>0.76 0.99</td>
</tr>
<tr>
<td>Shiquan</td>
<td>1.85 0.97</td>
<td>0.90 0.99</td>
<td>0.73 0.99</td>
</tr>
<tr>
<td>Yushu</td>
<td>1.39 0.96</td>
<td>0.91 0.98</td>
<td>1.74 0.94</td>
</tr>
<tr>
<td>Manyang</td>
<td>1.21 0.96</td>
<td>1.01 0.99</td>
<td>1.48 0.97</td>
</tr>
<tr>
<td>Lijiang</td>
<td>1.37 0.91</td>
<td>1.24 0.94</td>
<td>2.28 0.77</td>
</tr>
<tr>
<td>Panzhihu</td>
<td>1.73 0.90</td>
<td>2.50 0.97</td>
<td>1.47 0.94</td>
</tr>
<tr>
<td>Kuming</td>
<td>1.28 0.91</td>
<td>1.26 0.97</td>
<td>0.82 0.96</td>
</tr>
<tr>
<td>Hotan</td>
<td>1.11 0.98</td>
<td>1.51 0.99</td>
<td>2.41 0.98</td>
</tr>
<tr>
<td>Emeishan</td>
<td>1.33 0.89</td>
<td>1.39 0.90</td>
<td>4.95 0.83</td>
</tr>
<tr>
<td>Garze</td>
<td>1.54 0.94</td>
<td>1.60 0.97</td>
<td>5.94 0.79</td>
</tr>
<tr>
<td>Lhasa</td>
<td>1.14 0.97</td>
<td>0.97 0.98</td>
<td>3.38 0.88</td>
</tr>
</tbody>
</table>
In Fig. 6, GSR time series for each validation site are plotted and error metrics are provided in order to investigate the retrieval ability of ANN method presented in this study. As seen in Fig. 6, the performance of the three estimation methods is different at these 12 validation sites, although ANN-based method exhibits overall stronger retrieval ability than JAXA GSR algorithms does (Fig. 5). At seven sites (Yining, Kashi, Hami Shiquanhe, Yushu, Mianyang and Lijiang), ANN-based method retrieves GSR more accurately. At other sites (Rouqiang, Dunhuang, Golmud, Panzhihua and Kunming), JAXA products perform better. As mentioned above, Yang’s method generally performs better than both the ANN-based method and JAXA algorithm. However, its retrievals at Kashi and Panzhihua have relatively large errors.

In addition, measured GSR data at the 22 training sites are also used to validate JAXA GSR products. It is found that the retrieval errors are large at four sites (Hotan, Emeishan, Garze, and Lhasa) and RMSEs even reach roughly 5 MJ day\(^{-1}\) at Emeishan and Garze sites. So it is speculated that there perhaps exist some special conditions at these sites, which makes JAXA GSR estimates having large errors and the similar situations may happen to ANN-based
method presented in this study. However, GSR observations at these four sites are used to train ANN as aforementioned. Therefore, an experiment has to be made to check whether the ANN-based method can obtain satisfactory GSR estimates at these four sites. In the following, these four sites are excluded from the training sites, measurements only at the left 18 sites are used to train ANN, and then this ANN is applied to retrieve GSR estimates at these four stations. Fig. 7 indicates that ANN-based algorithm, after excluding the four sites, performs much better than JAXA retrieval scheme at these sites. This proves that the bad performance of the JAXA algorithm at these four sites is due to its own failure; ANN scheme has the fairly good robustness. Yang’s method also performs well at these four sites as shown in Fig. 7. The error statistics in Figs. 6 and 7 are also summarized in Table 2.

As shown above, ground-data-based methods such as Yang’s method may retrieve GSR with high accuracy. However, it is impossible to estimate GSR on a regional scale since the areal inputs required by them are not available. So the trained ANN is applied to retrieve GSR over the TP and its surroundings in order to show the advantage of the method presented in this study. The estimated monthly-mean daily GSR images for 12 months in 2004 are illustrated in Fig. 8. As seen, these 12 images thoroughly exhibit the spatial–temporal patterns of GSR over the TP region. GSR values in western TP are much larger than ones in eastern TP since the western TP is dry and the eastern TP is relatively wet. Following the northerly propagation of the Plateau summer monsoon in the summer, the dividing line between high and low GSR values moves westward, as shown in the June–August panels in Fig. 8. It is also seen that GSR estimates exceed 30 MJ day$^{-1}$ at many points in the western TP since the elevation is high there, the optical air mass is rather small, and the amount of cloud is low due to lack of water vapor.

4. Concluding remarks

Many algorithms have been developed to estimate solar radiation either based on surface meteorological observations or based on satellite data. In many regions over the world, there are no dense meteorological stations. Therefore, no spatially continuous GSR values can be estimated from routinely observed meteorological parameters. In this study, an efficient and effective ANN-based algorithm is presented to estimate monthly-mean daily GSR at a high spatial resolution in order to overcome some weaknesses, especially low spatial resolution, in existing methods. ANN is trained to build the functional relationship between observed GSR and exiting RS products such as LST and EVI. It is demonstrated that this ANN-based method can retrieve monthly-mean daily GSR with high accuracy and has a good stability in comparison with JAXA GSR products. Moreover, the ANN-based method is used to map the GSR distribution in space and time over the TP and its surroundings. Resultant images illustrate that the reasonable spatial and temporal patterns of GSR could be obtained by the method presented in this study, although this region has the very complex terrain and climate, suggesting that the estimation scheme presented in this study should be applied in other places.

In summary, the following findings are acquired in this work. First, ANN can be used to explore the mathematical relationship between GSR and satellite products. Second, this type of relationship is stable since it can be validated well at the different sites. Finally, this ANN-based method can be applied to map the regional GSR distribution with both high efficiency and effectiveness. In the future, the estimation method presented in this study can be used to obtain the GSR in China and coupled with other datasets to provide the distribution of solar energy potential with a high spatial resolution.

Acknowledgements

This work was supported by National Basic Research Program of China under contract 2005CB723901, the “100-Talent” Program of the Chinese Academy of Sciences, the Natural Science Foundation of China under contract 40701129, 41071220, and China “863” project under Contract 2006AA01A120 and 2009AA122100. The sample GLI data was produced and supplied by the Earth Observation Research and Application Center, Japan Aerospace Exploration Agency.

References


