MAUSAM, **66**, 2 (April 2015), 225-236

551.521.31 (510)

Estimation of solar radiation using two-step method in Yangtze River basin in China

JI-LONG CHEN, BEI-BEI XIAO*, CHUN-DI CHEN, ZHAO-FEI WEN, YI JIANG, MING-QUAN LV, SHENG-JUN WU and GUO-SHENG LI**

Chongqing Institute of Green and Intelligent Technology, CAS, Chongqing – 401 122, China

**Chongqing Institute of Surveying and Planning for Land Resource and Housing, Chongqing – 400 020, China*

***Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing – 100 101, China*

(*Received 20 January 2012, Modified 17 November 2014*)

e mail : ligscas@163.com

सार - सौर विकिरण कई भौतिक, रसायनिक और जैविक प्रक्रियाओं की प्रमुख और मूलभूत ऊर्जा है। मापे गए सौर विकिरण के आँकड़े उपलब्ध न होने पर अन्य मापित मौसम वैज्ञानिक परिवर्तिताओं से प्राप्त सौर विकिरण का आकलन एक महत्वपूर्ण विकल्प प्रदान करता है। इस शोध पत्र में हमने चीन में यंगट्ज़ नदी बेसिन के 14 स्थानों के दीर्घावधि आँकड़ों का उपयोग कर विकसित किए गए हार्ग्रीव और समानी, हार्ग्रीव, बिसटॉव और केम्पबेल व चेन मॉडल तथा एक स्थानीय समाश्रयण मॉडल सहित सामान्य रूप से प्रयुक्त चार वायु तापमान आधारित मॉडलों को प्रमाणित कर उनका निर्धारण किया। सामान्यत : मापित वायु तापमान से सौर विकिरण का आकलन करने के लिए द्वि-पदी पद्धति प्रस्तुत की। वर्ग माध्य मूल त्रुटि (RMSE), सापेक्षिक वर्ग माध्य मूल त्रुटि (RRMSE), माध्य निरेपक्ष प्रवृत्ति त्रुटि (MABE) और माध्य निरपेक्ष प्रतिशत त्रुटि (MAPE) का उपयोग करते हुए मॉडल के निष्पादन का मूल्यांकन किया गया। इनसे प्राप्त परिणामों से पता चला है कि द्वि-पदी पद्धति बेहतर निष्पादन देती है और तापमान आधारित मॉडल उल्लेखनीय रूप से अच्छा निष्पादन देता है। द्वि-पदी पद्धति के प्राचल निर्धारण समीकरण उन स्थानों पर प्राचल अंशांकन की कठिन समस्याओं का समाधान करते हैं जहाँ पर सौर विकिरण और धूप की अवधि के दीर्घावधि प्रेक्षण उपलब्ध नहीं होते हैं। यह देखा गया है कि प्रस्तावित समीकरणों द्वारा निर्धारित प्राचलों का उपयोग करने वाली द्वि-पदी पद्धति 0.881 के औसतन R^2, 15.04 के RRMSE और 12.67 प्रतिशत के MAPE के साथ बेहतर निष्पादन देती है। अत: यंगट्ज नदी बेसिन में सौर विकिरण का आकलन करने के लिए प्रस्तावित समीकरणों द्वारा निर्धारित प्राचलों सहित द्वि-पदी पद्धति का उपयोग किया जा सकता है। ऐसा माना जाता है कि जहाँ पर मापित सौर विकिरण और धूप की अवधि के आँकड़े उपलब्ध नही होते हैं उस स्थान के लिए यह उपयोगी होता है जबकि वायु के तापमानों को सामान्य रूप से मापा जाता है।

ABSTRACT. Solar radiation is the principal and fundamental energy for many physical, chemical and biological processes. Estimation of solar radiation from other measured meteorological variables offers an important alternative in the absence of availability of measured solar radiation data. In this paper, we validate and assess four commonly used air temperature-based models including the Hargreaves and Samani, Hargreaves, Bristow & Campbell and Chen model and a local regression model, developed using long-term data from 14 sites in Yangtze River basin in China. We present the two-step method to estimate solar radiation from the commonly measured air temperature. The model performance is evaluated using root mean square error (RMSE), relative root mean square error (RRMSE), mean absolute bias error (MABE) and mean absolute percentage error (MAPE). Results show that the two-step method gives good performance and significantly outperforms the temperature-based models. The parameters determination equations of the two-step method are proposed to solve the difficult problem in parameter calibration at the site where no long-term observations of solar radiation and sunshine duration are available. It is found that the two-step method using the parameters determined by the proposed equations gives good performance, with the averaged R^2 of 0.881, RRMSE of 15.04% and MAPE of 12.67%. Therefore, the two-step method with the parameters determined by the proposed equations could be used to estimate solar radiation in Yangtze River basin. It is believed to be useful for the site where no measured solar radiation and sunshine duration data is available, whereas the air temperatures are common measured.

Key words – Econ Solar radiation, Air temperature, A-P model, Two-step method, Determination equations.

1. Introduction

Solar radiation at the earth's surface is the principal and fundamental energy source for many physical, chemical and biological processes, such as crop growth, plant photosynthesis, and it is also an essential and important variable to many simulation models studies, such as agriculture, environment, hydrology, meteorology

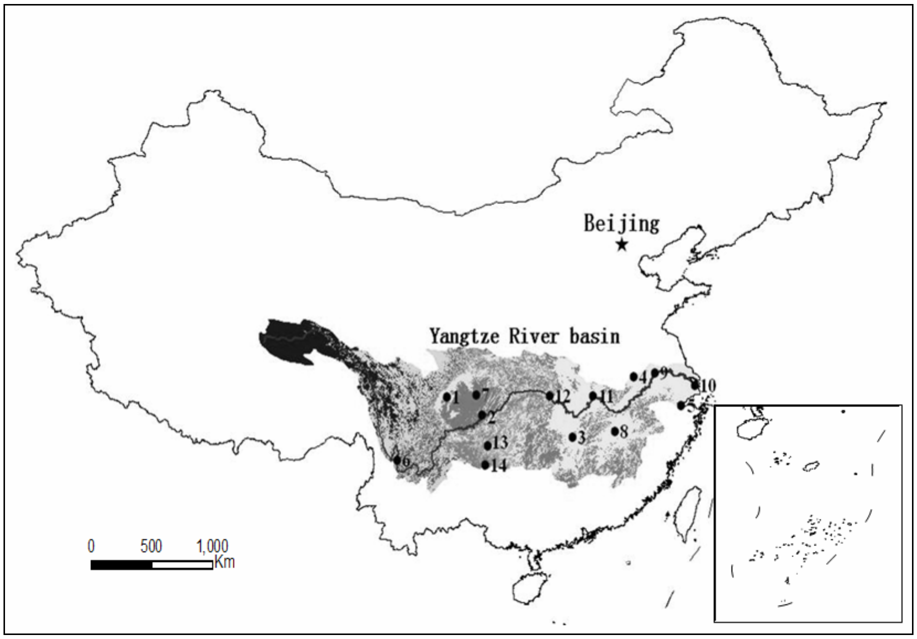

Fig. 1. Location of the study meteorological stations in the Yangtze River basin (stations are numbered in compliance with Table 1)

and ecology. Hence, accurate records of solar radiation are of vital importance. However, it is not widely available due to excessive cost and difficulties in maintenance and calibration of the measurement equipment (Hunt *et al.*, 1998). This parameter is measured only at few meteorological stations. For example, in USA, less than 1% of meteorological stations are recording solar radiation (NCDC, 1995; Thorton and Running, 1999). In China, more than 2000 stations have records of meteorological data; only 98 stations are recording solar radiation (Chen *et al.*, 2004). Therefore, developing method to estimate solar radiation for the site where no solar radiation is readily available has been the focus of many studies.

A number of methods, including satellite-derived (Frulla *et al.*, 1988; Pinker *et al.*, 1995; Olseth and Skartveit, 2001), stochastic algorithm (Richardson, 1981; Hansen, 1999; Wilks and Wilby, 1999) and empirical relationships (Ångström, 1924; Prescott, 1940; Hargreaves and Samani, 1982; Bristow and Campbell, 1984; Hargreaves *et al.*, 1985), interpolation-based (Hay and Suckling, 1979; Rivington et al., 2006), learning machine method (Tymvios *et al.*, 2005; Lam *et al.*, 2008; Chen *et al.*, 2011) have been developed for this purpose. Among them, the empirical relationship method is attractive for its simplicity, efficiency and much lower data requirement. Ångström (1924) first proposed a linear relationship between the ratio of average daily solar radiation to the corresponding value on a completely clear day and the ratio of average daily sunshine duration to the maximum possible sunshine duration. Prescott (1940) modified this method by replacing the clear sky radiation with the extraterrestrial radiation. Subsequently, the well-known sunshine-based Ångström-Prescott (A-P) model was developed. Despite its good performance, it is often limited by the lack of sufficient sunshine duration records. To obviate this difficulty, lots of air temperature-based models have been developed and widely used (Hargreaves and Samani, 1982; Bristow and Campbell, 1984; Hargreaves *et al.*, 1985; Chen *et al.*, 2004). These models allow widespread applications because air temperatures are routinely measured at most meteorological stations. However, a number of studies have shown that the sunshine-based A-P model significantly outperformed the temperature-based models (Iziomon and Mayer, 2002; Podestá *et al.*, 2004; Trnka *et al.*, 2005). Therefore, more accurate solar radiation estimation using air temperature is of vital importance and significance.

The much better performance of the sunshine-based A-P model over the temperature-based model seems to indicate that if the sunshine duration could be estimated from air temperature, then the estimated sunshine duration could be used as inputs for A-P model to estimate solar radiation, and the conversion in this way may be able to improve the estimation accuracy over the temperature-based models. We name this method as two-step estimation method (hereafter referred to two-step method), namely, estimate the sunshine duration using air temperature firstly, and then estimate solar radiation by A-P model using the estimated sunshine duration.

In the present work, four widely used air temperature-based models and a local regression model

TABLE 1

Detailed information of the studied 14 stations in Yangtze River Basin

Station ID	Station Name	Latitude (N)	Longitude (E)	Altitude (m)	Calibration period	Validation period
1.	Chengdu	30.67	104.02	506	1973-1992	1993-2000
2.	Chongqing	29.58	106.47	259	1973-1992	1993-2000
3.	Changsha	28.22	112.92	68	1987-1996	1997-2000
4.	Hefei	31.87	117.23	28	1978-1992	1993-2000
5.	Hangzhou	30.23	120.17	42	1973-1992	1993-2000
6.	Lijiang	26.83	100.47	2394	1977-1992	1993-2000
7.	Nanchong	30.78	106.10	309	1974-1985	1986-1990
8.	Nanchang	28.60	115.92	47	1973-1991	1993-2000
9.	Nanjing	32.00	118.80	9	1973-1992	1993-2000
10.	Shanghai	31.17	121.43	3	1961-1983	1983-1990
11.	Wuhan	30.62	114.13	23	1973-1983 1985-1992	1993-2000
12.	Yichang	30.70	111.30	133	1973-1992	1993-2000
13.	Zunyi	27.7	106.88	844	1973-1984	1985-1990
14.	Guiyang	26.58	106.72	1074	1973-1992	1993-2000

are studied. The main objectives of this study are to (1) calibrate and validate the temperature-based models from data of 14 sites in Yangtze River basin in China; and (2) investigate and evaluate the two-step method.

2. Data and methodology

2.1. *Sites and data set*

The current study focuses on the Yangtze River basin (Fig. 1). The Yangtze River is 6300 km long with a basin area of 180 × 10^4 km^2 with decreasing altitude from west to east. The basin is characterized by abundant water resources, and thus plays significant role in water supply for agriculture industry, because economy of much of the Yangtze River basin is largely dependent on agricultural production. A large part of the Yangtze River basin is of subtropical monsoon climate type. Average annual precipitation in the basin varies from 270-500 mm in the west to 1600-1900 mm in the southeast (Zhang *et al.*, 2008). Average annual sunshine hours ranges from 1000 h to 2500 h (Gong *et al.*, 2006). A total of 14 stations with long-term available records of solar radiation are used in the present study. The mapping of stations roughly range from 26° to 34° latitude North, from 97° to 121° longitude East, and from 3 to 3394 m altitude. Table 1 shows the temporal period of data calibration and validation and the geographical information of the meteorological stations.

The monthly mean daily solar radiation (MJ m^{-2} d^{-1}), sunshine duration (h), air temperature (°C), including mean maximum temperature and minimum temperature are used in this study. The data were obtained from the National Meteorological Information Center (NMIC), China Meteorological Administration (CMA). The period of records ranges from 6 to 30 years covering the period between 1961 and 2000. Quality control tests were conducted by the suppliers of data. A year with more than 5 days of missing or faulty data in the same month was discarded (*e.g.*, the year of 1992 for Nanchang and the year of 1984 for Wuhan). For each station, two data sets were created. About 70% of the total records were used for calibrating the parameters, and the remaining records were used for evaluating the model (Table 1).

2.2. *Temperature-based solar radiation estimation models*

2.2.1. *H-S model*

Hargreaves and Samani (1982) suggested a solar radiation estimation model that is a function of extraterrestrial radiation, maximum and minimum temperatures as follow:

$$\frac{Rs}{Ra} = a(T\max - T\min)^{0.5} \quad (1)$$

TABLE 2

The empirical parameters of the temperature-based models *

Station	H-S model[a]		H-Sm model[b]			Chen model[c]			Local Rs model[d]		
	a	R^2	a	b	R^2	a	b	R^2	a	b	R^2
Chengdu	0.112	0.500	0.196	-0.227	0.614	0.261	-0.213	0.617	0.036	0.036	0.605
Chongqing	0.105	0.577	0.225	-0.305	0.813	0.267	-0.225	0.787	0.046	-0.037	0.829
Changsha	0.126	0.365	0.298	-0.466	0.550	0.392	-0.437	0.559	0.055	-0.068	0.534
Hefei	0.129	0.463	0.247	-0.337	0.600	0.353	-0.369	0.605	0.043	0.019	0.591
Hangzhou	0.129	0.370	0.249	-0.330	0.483	0.332	-0.314	0.482	0.046	0.005	0.480
Lijiang	0.165	0.612	0.329	-0.559	0.817	0.547	-0.773	0.816	0.049	-0.012	0.813
Nanchong	0.124	0.608	0.233	-0.279	0.783	0.283	-0.209	0.769	0.047	0.006	0.789
Nanchang	0.106	0.423	0.253	-0.405	0.643	0.336	-0.386	0.640	0.047	-0.065	0.640
Nanjing	0.132	0.390	0.169	-0.107	0.403	0.236	-0.121	0.404	0.027	0.152	0.399
Shanghai	0.153	0.360	0.229	-0.210	0.405	0.312	-0.209	0.407	0.042	0.104	0.401
Wuhan	0.126	0.286	0.209	-0.238	0.340	0.306	-0.279	0.349	0.035	0.070	0.328
Yichang	0.120	0.383	0.243	-0.344	0.515	0.328	-0.337	0.511	0.044	-0.011	0.511
Zunyi	0.101	0.549	0.233	-0.360	0.814	0.292	-0.304	0.786	0.045	-0.065	0.831
Guiyang	0.106	0.423	0.253	-0.405	0.643	0.336	-0.386	0.640	0.047	-0.065	0.640

* The metric R^2, varying from 0 to 1, is adopted to measure the degree of success of a fit in explaining data variation, with 0 denoting that model does not explain any variation and 1 denoting that it perfectly explains the observed variation.
[a] Hargreaves and Samani (1982)
[b] Hargreaves *et al*. (1985)
[c] Chen *et al*. (2004)
[d] See Eqn. (7)

where *Rs* is monthly mean daily actual global radiation (MJ m^{-2} d^{-1}), *T*max and *T*min are maximum and minimum temperature, respectively, Ra is monthly mean daily extra-terrestrial solar radiation (MJ m^{-2} d^{-1}), which is a function of latitude and day of the year. The detailed procedure for calculation of extraterrestrial radiation can be found in Allen *et al*. (1998), where 'a' is empirical parameter, which was recommended to be 0.16 for interior regions and 0.19 for coastal regions. In the present work, parameter listed in Table 2 is locally calibrated for each station using the training data set.

2.2.2. *H-Sm model*

Hargreaves *et al*. (1985) further modified the H-S model by adding an empirical correction parameter and obtained the following equation:

$$\frac{Rs}{Ra} = a(T\max - T\min)^{0.5} + b \quad (2)$$

where, a and b are empirical parameters as listed in Table 2.

2.2.3. *B-C model*

Bristow and Campbell (1984) studied the relationship between solar radiation and maximum, minimum temperature and developed the equation given below:

$$Rs = aRa[1 - exp(-b(\Delta Ti)^c)] \quad (3)$$

a, b and c are parameters, values widely used for these coefficients are 0.7 for a, 2.4 for c and the value of b can be calculated from the equation:

$$b = 0.036\,exp(-0.154\overline{\Delta T}) \quad (4)$$

where, $\overline{\Delta T}$ is the monthly average of the ΔTi daily values, ΔTi is the air temperature difference calculated by subtracting the average *T*min of the current and the next day from the *T*max of the current day:

$$\Delta Ti = T\max{,i} - \frac{T\min{,i} + T\min{,i+1}}{2} \quad (5)$$

TABLE 3

The empirical parameters of A-P model and local S model*

Station	A-P model[a]			local S model[b]			
	a	*b*	R^2	*a*	*b*	*c*	R^2
Chengdu	0.550	0.164	0.747	0.052	-0.047	-0.193	0.715
Chongqing	0.585	0.118	0.867	0.060	-0.055	-0.218	0.828
Changsha	0.621	0.125	0.867	0.074	-0.066	-0.314	0.767
Hefei	0.590	0.103	0.773	0.067	-0.064	-0.130	0.678
Hangzhou	0.586	0.127	0.786	0.074	-0.072	-0.187	0.607
Lijiang	0.576	0.225	0.873	0.065	-0.074	-0.116	0.903
Nanchong	0.583	0.155	0.872	0.065	-0.060	-0.234	0.856
Nanchang	0.579	0.120	0.915	0.096	-0.089	-0.345	0.757
Nanjing	0.536	0.147	0.785	0.055	-0.051	-0.104	0.595
Shanghai	0.565	0.158	0.867	0.076	-0.070	-0.192	0.505
Wuhan	0.564	0.110	0.771	0.069	-0.060	-0.243	0.709
Yichang	0.594	0.120	0.810	0.065	-0.060	-0.208	0.712
Zunyi	0.580	0.131	0.895	0.058	-0.053	-0.263	0.817
Guiyang	0.582	0.133	0.870	0.057	-0.049	-0.267	0.747

* The metric R^2, varying from 0 to 1, is adopted to measure the degree of success of a fit in explaining data variation, with 0 denoting that model does not explain any variation and 1 denoting that it perfectly explains the observed variation.
[a] Ångström (1924); Prescott (1940); [b] See Eqn.(9)

where, $T\max,i$, $T\min,i$, $T\min.i+1$ are Tmax, Tmin in *i*th day, Tmin in (i +1)th day, respectively.

2.2.4. *Chen model*

Chen *et al*. (2004) presented a new model as follow:

$$\frac{Rs}{Ra} = a \ln (T\max - T\min) + b \qquad (6)$$

where, a and b are empirical parameters listed in Table 2.

2.2.5. *Local model*

In addition to these well-known models, we propose the following equation based on the investigation of relationship between *Rs/Ra* and maximum, minimum temperature (hereafter referred to local *Rs* model).

$$\frac{Rs}{Ra} = a(T\max - T\min) + b \qquad (7)$$

where, *a* and *b* are empirical parameters as listed in Table 2.

2.3. *Sunshine-based A-P model*

A-P model was proposed by Ångström (1924) and further modified by Prescott (1940). The original form of this model is:

$$\frac{Rs}{Ra} = a\frac{S}{So} + b \qquad (8)$$

where, a and b are empirical parameters which are calibrated from regression analysis between *S/So* and *Rs/Ra*. *S* is monthly mean daily actual duration of sunshine hours (h), *So* is monthly mean daily maximum possible sunshine duration (h) which is calculated using the equations detailed by Allen *et al*. (1998).

2.4. *Sunshine duration estimation model*

In this paper, we propose a formula to estimate monthly mean daily sunshine duration using maximum and minimum temperature in Yangtze River basin (hereafter referred to local *S* model).

$$\frac{S}{So} = aT\max + bT\min + c \qquad (9)$$

TABLE 4

Root mean square error (RMSE in MJ m^{-2}), Relative root mean square error (RRMSE), Mean absolute bias error (MABE in MJ m^{-2}) and mean absolute percentage error (MAPE) of the temperature-based models and the two-step method

Station	H-S model[a]				H-Sm model[b]				B-C model[c]			
	RMSE	RRMSE	MABE	MAPE	RMSE	RRMSE	MABE	MAPE	RMSE	RRMSE	MABE	MAPE
Chengdu	1.488	17.28%	1.294	20.35%	1.362	15.36%	1.131	15.68%	1.361	15.80%	1.203	19.95%
Chongqing	1.963	22.55%	1.425	19.23%	1.238	14.22%	0.906	10.60%	1.504	17.28%	1.210	18.71%
Changsha	2.656	24.80%	2.084	22.52%	2.250	21.01%	1.733	16.46%	2.347	21.92%	1.923	22.14%
Hefei	2.282	18.72%	1.538	11.09%	2.202	18.07%	1.517	11.25%	2.145	17.60%	1.476	11.11%
Hangzhou	2.282	19.54%	1.635	14.39%	2.041	17.47%	1.487	12.40%	2.097	17.95%	1.550	14.23%
Lijiang	2.211	13.15%	1.888	11.56%	1.612	9.59%	1.306	7.76%	1.807	10.75%	1.621	11.29%
Nanchong	1.627	17.56%	1.380	20.80%	1.585	17.11%	1.221	15.55%	1.587	17.13%	1.318	20.44%
Nanchang	2.407	23.78%	1.883	18.96%	2.314	22.86%	1.734	15.78%	2.321	22.92%	1.808	18.74%
Nanjing	1.691	14.19%	1.409	10.40%	1.629	13.67%	1.225	10.24%	1.620	13.59%	1.219	10.39%
Shanghai	2.079	17.35%	1.638	14.69%	2.047	17.08%	1.612	13.82%	2.058	17.18%	1.626	14.62%
Wuhan	2.578	22.60%	1.949	15.52%	2.479	21.73%	1.824	14.97%	2.444	21.43%	1.794	15.49%
Yichang	2.117	19.69%	1.455	22.78%	1.962	18.24%	1.447	21.59%	1.995	18.55%	1.443	22.69%
Zunyi	1.705	24.24%	1.024	16.54%	1.176	18.79%	0.822	11.13%	1.349	21.55%	0.875	16.05%
Guiyang	2.407	23.78%	1.883	18.96%	2.314	22.86%	1.734	15.78%	2.321	22.92%	1.808	18.74%

Station	Chen model[d]				Local Rs model[e]				Two-step method			
	RMSE	RRMSE	MABE	MAPE	RMSE	RRMSE	MABE	MAPE	RMSE	RRMSE	MABE	MAPE
Chengdu	1.404	16.30%	1.130	15.84%	1.336	15.51%	1.138	16.06%	1.320	15.33%	1.128	15.57%
Chongqing	1.183	15.66%	0.894	0.0943	1.157	13.29%	0.839	9.11%	1.101	12.64%	0.759	8.33%
Changsha	2.252	21.03%	1.767	17.59%	2.266	21.16%	1.704	15.40%	1.698	15.85%	1.333	13.47%
Hefei	2.195	18.01%	1.513	11.36%	2.206	18.10%	1.524	11.24%	1.846	15.14%	1.320	10.20%
Hangzhou	2.040	17.47%	1.469	12.29%	2.050	17.55%	1.511	12.60%	1.934	16.56%	1.423	11.98%
Lijiang	1.583	10.01%	1.265	8.15%	1.564	9.30%	1.250	7.40%	1.413	8.40%	1.103	6.55%
Nanchong	1.575	17.00%	1.239	16.12%	1.613	17.41%	1.209	15.27%	1.560	16.84%	1.177	14.41%
Nanchang	2.294	22.66%	1.736	16.07%	2.336	23.08%	1.734	15.55%	1.574	13.25%	1.260	11.17%
Nanjing	1.629	13.67%	1.229	10.27%	1.629	13.67%	1.224	10.29%	1.358	11.39%	1.075	9.03%
Shanghai	2.052	17.13%	1.615	13.82%	2.043	17.05%	1.612	13.87%	1.949	16.27%	1.462	12.63%
Wuhan	2.495	21.88%	1.838	15.10%	2.459	21.56%	1.809	14.85%	1.678	14.71%	1.345	12.87%
Yichang	1.957	18.20%	1.431	21.61%	1.972	18.33%	1.459	21.59%	1.697	15.78%	1.220	19.42%
Zunyi	1.239	19.79%	0.732	10.24%	1.167	18.64%	0.711	8.64%	1.119	13.40%	0.673	9.26%
Guiyang	2.294	22.66%	1.736	16.07%	2.336	23.08%	1.734	15.55%	1.702	16.82%	1.298	12.56%

[a] Hargreaves and Samani (1982); [b] Hargreaves *et al*. (1985); [c] Bristow and Campbell (1984); [d] Chen *et al*. (2004); [e] See Eqn. (7)

where, a, b and c are empirical parameters listed in Table 3 which are locally determined using the calibration data set.

2.5. *Performance criteria*

To assess the performance of models, root mean square error (RMSE), relative root mean square error (RRMSE) (%), mean absolute bias error (MABE), mean absolute percentage error (MAPE) (%) and coefficient of determination (R^2) are determined. The metric R^2 is adopted to measure the fit of the model on calibration data, and the correlation between the estimated and observed values. The former four indicators are calculated by the following equations:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{10}$$

TABLE 5

The relative improvement of Root mean square error (RIrmse) and mean absolute bias error (RImabe) of the two-step method over the temperature-based models

Station	H-S model[a]		H-Sm model[b]		B-C model[c]		Chen model[d]		Local Rs model[e]	
	RIrmse	RImabe	RIrmse	RImabe	RIrmse	RImabe	RIrmse	RImabe	RIrmse	RImabe
Chengdu	11.30%	12.83%	3.10%	0.27%	2.96%	6.26%	5.97%	0.18%	1.15%	0.93%
Chongqing	43.94%	46.73%	11.11%	16.21%	26.84%	37.26%	6.97%	15.03%	4.89%	9.48%
Changsha	36.06%	36.05%	24.54%	23.10%	27.66%	30.68%	24.61%	24.59%	25.08%	21.79%
Hefei	19.11%	14.16%	16.17%	13.00%	13.94%	10.58%	15.91%	12.76%	16.32%	13.40%
Hangzhou	15.23%	12.95%	5.21%	4.28%	7.76%	8.22%	5.19%	3.16%	5.64%	5.81%
Lijiang	36.12%	41.58%	12.36%	15.51%	21.83%	31.96%	10.75%	12.78%	9.67%	11.74%
Nanchong	4.08%	14.71%	1.58%	3.56%	1.68%	10.69%	0.92%	4.94%	3.27%	2.57%
Nanchang	34.62%	33.09%	31.99%	27.33%	32.18%	30.32%	31.39%	27.43%	32.64%	27.36%
Nanjing	14.68%	11.09%	16.64%	12.26%	16.19%	11.86%	16.64%	12.59%	16.64%	12.21%
Shanghai	6.27%	10.69%	4.78%	9.29%	5.32%	10.06%	5.02%	9.44%	4.59%	9.29%
Wuhan	26.32%	23.10%	32.30%	26.28%	31.34%	25.06%	32.75%	26.84%	31.76%	25.66%
Yichang	19.85%	16.12%	13.52%	15.64%	14.94%	15.42%	13.30%	14.73%	13.93%	16.32%
Zunyi	34.40%	16.85%	4.91%	21.00%	17.07%	23.13%	9.73%	8.10%	4.13%	5.34%
Guiyang	29.29%	31.07%	26.44%	25.15%	26.64%	28.22%	25.79%	25.24%	27.14%	25.17%

[a] Hargreaves and Samani (1982); [b] Hargreaves *et al*. (1985); [c] Bristow and Campbell (1984); [d] Chen *et al*. (2004); [e] See Eqn. (7)

$$\text{RRMSE} = \frac{100}{\bar{y}}\sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \qquad (11)$$

$$\text{MABE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \qquad (12)$$

$$\text{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \qquad (13)$$

where, n, y, $\hat{y}$ and $\bar{y}$ represent the number of testing data, the observed value, the estimated value and the average value of the observation, respectively. Lower values of RMSE, RRMSE, MABE, and MAPE indicate a better estimation accuracy of the model.

The relative improvement of RMSE (RIrmse) and MABE (RImabe) were used to measure the improvement of the evaluated model accuracy over the reference model accuracy:

$$\text{RIrmse} = \frac{\text{RMSEr} - \text{RMSEe}}{\text{RMSEr}} \qquad (14)$$

$$\text{RImabe} = \frac{\text{MABEr} - \text{MABEe}}{\text{MABEr}} \qquad (15)$$

where, RMSEr and MABEr are the root mean square error and mean absolute bias error for the reference model, respectively. RMSEe and MABEe are the root mean square error and mean absolute bias error for the evaluated model, respectively. Higher values of RIrmse and RImabe indicate a better improvement over the reference model.

3. Results and discussion

3.1. *Performances of the temperature-based models*

The performances of the five temperature-based models are presented in Table 4. Overall, all the temperature-based models give good performances with the RRMSE < 25% (averaged 18.28%) and MAPE < 23% (average 14.94%). The H-Sm, Chen and local *Rs* models give quite similar values of RMSE (averaged 1.872, 1.871, 1.867 MJ m^{-2}, respectively), RRMSE (averaged 17.72%, 17.96%, 17.69%, respectively), MABE (averaged 1.409, 1.399, 1.386 MJ m^{-2}, respectively), and MAPE (averaged 13.79%, 13.80%, 13.40%, respectively). These three models have similar equation expressions, they differ in the form of the term *T*max-*T*min, namely, the square root of *T*max-*T*min, logarithm of *T*max-*T*min, and *T*max-*T*min for the H-Sm, Chen model, and local Rs model, respectively. However, they give nearly identical performances according to the indicators of RMSE, RRMSE, MABE and MAPE, indicating that the variations of term *T*max-*T*min are generally not very effective and gave no significant improvement. Evidently, the modified H-S model (H-Sm model) by adding a empirical parameter outperforms the original H-S model, in terms of RMSE, the accuracy could be on average 11.41% higher, and at some sites (Chongqing, Lijiang and Zunyi), the accuracy could be 27-36% higher.

3.2. *Performance of the two-step method*

Despite the simplicity and significant performance of A-P model, it is often limited by the lack of available sunshine duration records. Therefore, in the present work, an attempt has been made to estimate sunshine duration by local *S* model [Eqn. (9)] using air temperature. Consequently, the estimated sunshine duration is used as input for A-P model to estimate solar radiation, and the performance is presented in Table 4. We have named this method as two-step method. Such a method has not been studied before, which may be contributed to the lack of study on the relationship between the air temperature and sunshine duration. Another reason may be that there are two estimation processes which may largely decrease the performance accuracy. However, in the present work, the two-step method gives good performance with the RMSE < 2 MJ m^{-2} (averaged 1.568 MJ m^{-2}), RRMSE < 20% (averaged 14.46%), MABE < 1.5 MJ m^{-2} (averaged 1.184 MJ m^{-2}), and MAPE < 20% (averaged 11.96%) as shown in Table 4. Therefore, it could be used to estimate solar radiation in Yangtze River basin.

3.3. *Comparison of the two-step method between temperature-based models*

Performance comparisons of the two-step method and temperature-based models given in Table 4 shows that the two-step method gives the lowest RMSE of 1.525 ± 0.424 MJ m^{-2}(averaged 1.568 MJ m^{-2}), RRMSE of 12.62% ± 4.22% (averaged 14.46%), MABE of 1.068 ± 0.395 MJ m^{-2} (averaged 1.184 MJ m^{-2}) and MAPE of 12.98% ± 6.43% (average 11.96%) within the same station. The results of studies on relative improvement of the two-step method over the temperature-based models are presented in Table 5. Evidently, the two-step method significantly outperforms the temperature-based models, with the averaged RIrmse of 23.66%, 14.62%, 17.60%, 14.64%, and 14.06%, RImabe of 22.93%, 15.21%, 19.98%, 14.13%, and 13.36% over H-S, H-Sm, B-C, Chen and Local *Rs* model, respectively. These results further confirm that the two-step method can significantly improve the estimation accuracy over the temperature-based models which directly estimate solar radiation using air temperature only.

3.4. *Determination of model parameters*

The two-step method has substantial potential for application in solar radiation estimation due to the greater availability of air temperature data and the significant performance, but it requires the long-term observations of solar radiation and sunshine duration to calibrate the parameters, including the parameters of the A-P model and local *S* model. Generally, the parameters of two-step method (a, b of A-P model and a, b, c of local S model in Table 3) vary from station to station as shown in Table 3, namely, they are site dependent. And it is therefore open to question how transferable these calibration values are to other locations without measured solar radiation and sunshine duration data for calibration. No literature has reported the use of uniform set of determination equations of the parameters in a large area. Therefore, in the present work, based on the investigation of the relation between the parameters and geographical information (longitude, latitude and altitude in Table 1), mean daily extraterrestrial radiation and solar radiation, we propose the determination equations of parameters of the two-step method in Yangtze River basin as follows:

For the parameters of A-P model, the determination equations are:

$$a = -0.3494\sin(\varphi) - 1.1321\sin(\lambda) - 0.1311\frac{\sin(\varphi)}{\cos(\lambda)} + 3.306\times10^{-6}\beta + 1.1636 \quad (16)$$

$$b = -0.5082\sin(\varphi) + 0.1458\sin(\lambda) - 0.0106\frac{\sin(\varphi)}{\cos(\lambda)} - 1.6108\times10^{-5}\beta + 0.7168 \quad (17)$$

For the parameters of local *S* model, the determination equations are:

$$a = -0.0279\varphi + 5.8570\times10^{-4}\lambda - 0.0771Ra + 0.5057So - 3.3985\times10^{-5}\beta - 2.7931 \quad (18)$$

$$b = 0.0323\varphi - 6.3304\times10^{-4}\lambda + 0.0956Ra - 0.6195So + 3.2338\times10^{-5}\beta + 3.4581 \quad (19)$$

$$c = 0.0767\varphi + 4.0649\times10^{-3}\lambda - 0.0916Ra - 0.7524So + 1.7810\times10^{-4}\beta + 3.1221 \quad (20)$$

TABLE 6

Empirical parameters of the two-step method determined by the corresponding determination equation

Station	A-P model[a]		Local S model[b]		
	a[c]	b[d]	a[e]	b[f]	c[g]
Chengdu	0.569	0.165	0.055	-0.051	-0.207
Chongqing	0.583	0.135	0.064	-0.058	-0.261
Changsha	0.597	0.115	0.084	-0.076	-0.339
Hefei	0.565	0.124	0.066	-0.062	-0.144
Hangzhou	0.576	0.140	0.076	-0.071	-0.211
Lijiang	0.566	0.226	0.065	-0.074	-0.117
Nanchong	0.572	0.140	0.056	-0.052	-0.194
Nanchang	0.592	0.122	0.083	-0.076	-0.310
Nanjing	0.563	0.131	0.061	-0.056	-0.111
Shanghai	0.568	0.147	0.073	-0.068	-0.166
Wuhan	0.577	0.116	0.071	-0.065	-0.219
Yichang	0.576	0.115	0.065	-0.060	-0.207
Zunyi	0.590	0.131	0.057	-0.051	-0.252
Guiyang	0.595	0.131	0.058	-0.051	-0.271
RMSE	0.015	0.010	0.006	0.006	0.022
RRMSE	2.59%	7.06%	8.33%	2.64%	9.98%

[a] Ångström (1924); Prescott (1940); [b] See Eqn. (9); [c] See Eqn. (16)
[d] See Eqn. (17); [e] See Eqn. (18); [f] See Eqn. (19); [g] See Eqn. (20)

where, φ, λ, β, *Ra* and *So* are the latitude (rad), longitude (rad), altitude (m), mean daily extra-terrestrial solar radiation and maximum possible sunshine duration (h), respectively.

The determined parameters by the corresponding equation are presented in Table 6. It is found that, in general, the parameters determined by these equations are relatively close to the corresponding values (Table 3) calibrated from the data set, indicating that these determination equations give good performances, with the RMSE of 0.015, and 0.010, RRMSE of 2.59% and 7.06% for parameter a, and b of the A-P model; RMSE of 0.006, 0.006, and 0.022, RRMSE of 8.33%, 2.64%, and 9.98% for parameter a, b, and c of local *S* model, respectively. Therefore, they could be used to determine the parameters of the two-step method.

3.5. *Solar radiation estimation using the estimated parameters*

The local *S* model with the parameter a determined by Eqn. (18), b by Eqn. (19) and c by Eqn. (20) is consequently used to estimate *S/So*, which is later used as input to estimate solar radiation by A-P model using the parameter a determined by Eqn. (16) and b by Eqn. (17), and the performances are presented in Table 7 and Fig. 2. Good performance of the two-step method is found, with the RMSE < 2.5 MJ m^{-2} (averaged 1.623 MJ m^{-2}), RRMSE < 20% (averaged 15.04%), MABE < 1.8 MJ m^{-2} (averaged 1.253 MJ m^{-2}), MAPE < 20% (averaged 12.67%) and R^2 > 0.8(averaged 0.881). And there is a substantially good agreement between the estimated and observed values as shown in Fig. 2, where the points tend to line up around the 1:1 line, further indicating that the estimated solar radiation are close to the observed. Therefore, the two-step method with the parameters determined by equations could be used to estimate solar radiation in Yangtze River basin.

TABLE 7

Root mean square error (RMSE in MJ m^{-2}), Relative root mean square error (RRMSE), Mean absolute bias error (MABE in MJ m^{-2}) and mean absolute percentage error (MAPE) of the two-step method with the parameters determined by the corresponding determination equation

Station	RMSE	RRMSE	MABE	MAPE
Chengdu	1.593	18.50%	1.371	18.70%
Chongqing	1.261	14.48%	0.981	11.67%
Changsha	1.723	16.09%	1.357	13.81%
Hefei	1.597	13.10%	1.131	9.08%
Hangzhou	2.282	19.54%	1.741	15.79%
Lijiang	1.431	8.51%	1.124	6.66%
Nanchong	1.034	11.16%	0.722	8.39%
Nanchang	1.770	14.90%	1.424	12.54%
Nanjing	1.417	11.89%	1.106	9.35%
Shanghai	1.742	14.54%	1.338	11.14%
Wuhan	1.858	16.29%	1.436	12.88%
Yichang	1.665	15.48%	1.207	19.22%
Zunyi	1.430	17.14%	1.119	14.22%
Guiyang	1.921	18.97%	1.481	13.98%

4. Conclusion

Solar radiation is an essential and important variable for use in many simulation models. Estimation of solar radiation from other commonly measured meteorological variables is generally done when direct measurement are not readily available. In this work, we present the two-step method to estimate solar radiation from the commonly measured air temperature, namely, estimate sunshine

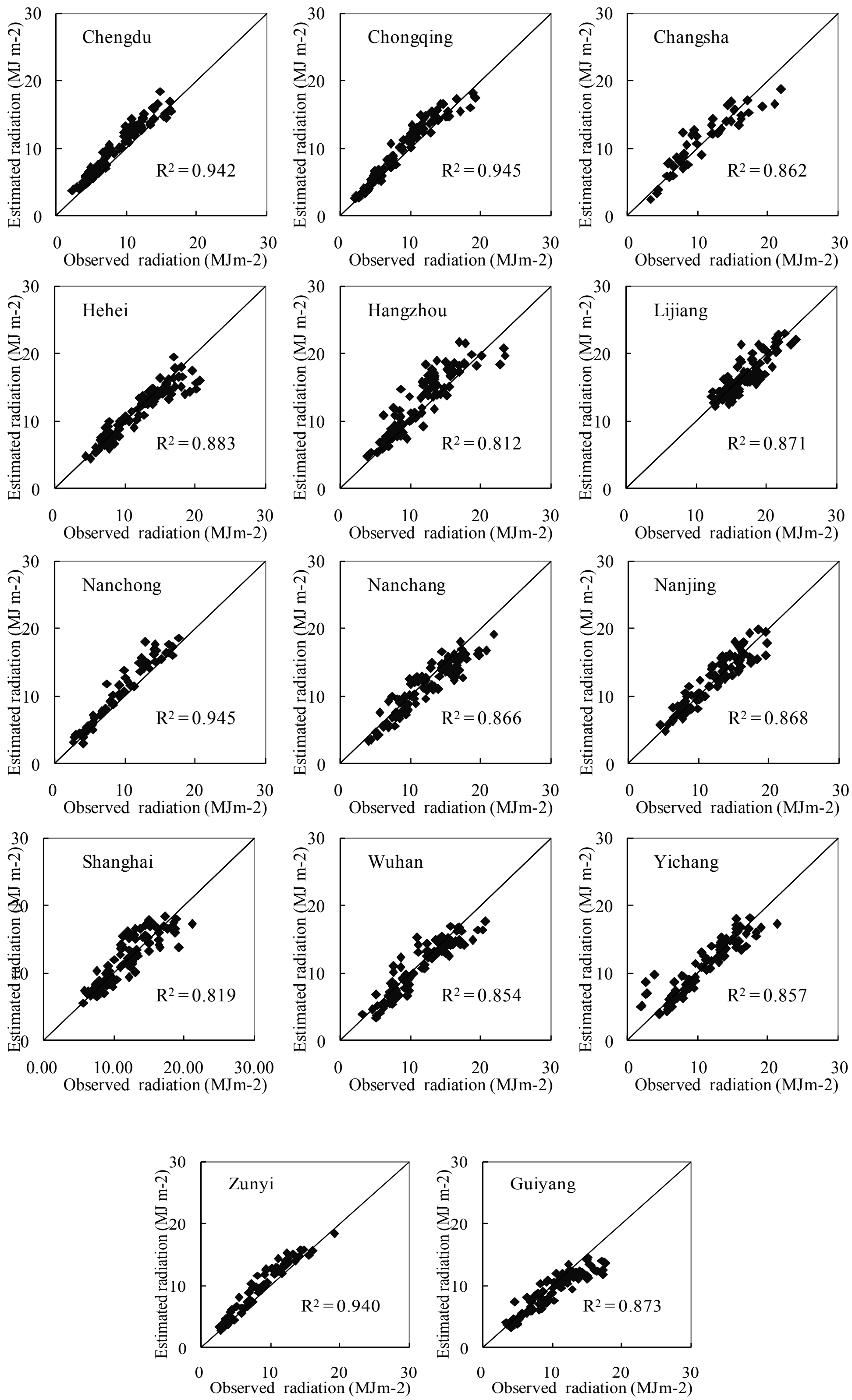

Fig. 2. Scatter plots of the observed *vs* estimated solar radiation at 14 stations in Yangtze River basin

duration by local *S* model [Eqn. (9)] using air temperature firstly, consequently, the estimated sunshine duration is used as inputs for A-P model to estimate solar radiation. The model performance indicators of RMSE, RRMSE, MABE, and MAPE illustrate that the two-step method gives good results. Further comparisons show that the two-step method significantly outperforms the temperature-based models. Therefore, it could be used to estimate solar radiation in Yangtze River basin, and it can significantly improve the estimation accuracy over the temperature-based models which directly estimate solar radiation using air temperature only.

We believe that the proposed two-step method has substantial potential for direct application in solar radiation estimation due to the greater availability of air temperature data and it significant improved performance over the temperature-based models. Therefore, the determination equations of parameters of the two-step method are proposed for Yangtze River basin. They give good performances in terms of RMSE and RRMSE. Consequently, the two-step method using the estimated parameters is further evaluated, and it is found that the estimated values agree well with the measured solar radiation. Therefore, the two-step method with the parameters determined by the equations as proposed in the paper could be used in Yangtze River basin, and it is believed to be particularly useful for the sites where no solar radiation and sunshine duration data is available, whereas the air temperatures are commonly measured.

Acknowledgment

The work was supported by National Key Technology Research and Development Program (2012BAC21B01), the Geological Survey program of China Geological Survey (GZH201200503) and Special foundation for scientific research on public Interest (1212010611402, 201111023) and Chongqing Science and Technology Key project (cstc2012ggB20001). We thank the National Meteorological Information Center, China Meteorological Administration for providing the long-term data records. Many thanks go to the anonymous reviewers for the comments on the manuscript.

References

Allen, R. G., Pereira, L. S., Raes, D. and Smith, M., 1998, "Crop evapotranspiration. Guidelines for computing crop water requirements", FAO Irrigation and Drainage paper, **56**, Rome.

Ångström, A., 1924, "Solar and terrestrial radiation. Quarterly Journal of the Royal Meteorological Society", **50**, 121-126.

Bristow, K. L. and Campbell, G. S., 1984, "On the relationship between incoming solar radiation and daily maximum and minimum temperature", *Agricultural and Forest Meteorology*, **31**, 159-166.

Chen, J. L., Liu, H. B., Wu, W. and Xie, D. T., 2011, "Estimation of monthly solar radiation from measured temperatures using support vector machines - A case study", *Renewable Energy*, **36**, 413-420.

Chen, R. S., Ersi, K., Yang, J. P., Lu, S. H. and Zhao, W. Z., 2004, "Validation of five global radiation models with measured daily data in China", *Energy Conversion and Management*, **45**, 1759-1769.

Frulla, L. A., Gagliardini, D. A., Grossi, G. H., Lopardo, R. and Tarpley, J. D., 1988, "Incident solar radiation on Argentina from the geostationary satellite GOES: Comparison with ground measurements", *Solar Energy*, **41**, 61-69.

Gong, L., Xu, C. Y., Chen, D., Halldin, S. and Chen, Y. D., 2006, "Sensitivity of the Penman-Monteith reference evapotranspiration to key climatic variables in the Changjiang (Yangtze River) basin", *Journal of Hydrology*, **329**, 3, 620-629.

Hansen, J. W., 1999, "Stochastic daily solar irradiance for biological modeling applications", *Agricultural and Forest Meteorology*, **94**, 53-63.

Hargreaves, G. H. and Samani, Z. A., 1982, "Estimating potential evapotranspiration", *Journal of Irrigation and Drainage Engineering*, **108**, 225-230.

Hargreaves, G. L., Hargreaves, G. H. and Riley, J. P., 1985, "Irrigation water requirement for Senegal River Basin", *Journal of Irrigation and Drainage Engineering*, **111**, 265-275.

Hay, J. E. and Suckling, P. W., 1979, "An assessment of the net-works for measuring and modelling solar radiation in British Columbia and adjacent areas of western Canada", *Canadian Geographic*, **23**, 222-238.

Hunt, L. A., Kuchar, L. and Swanton, C. J., 1998, "Estimation of solar radiation for use in crop modeling", *Agricultural and Forest Meteorology*, **91**, 293-300.

Iziomon, M. G. and Mayer, H., 2002, "Assessment of some global solar radiation parameterizations", *Journal of Atmospheric and Solar-Terrestrial Physics*, **64**, 1631-1643.

Lam, J. C., Wan, K. K. W. and Yang, L., 2008, "Solar radiation modeling using ANNs for different climates in China", *Energy Conversion and Management*, **49**, 1080-1090.

NCDC (National Climatic Data Center), 1995, "Cooperative summary of the day, dataset TD 3200", U. S. Department of Commerce, National Oceanographic and Atmospheric Administration, National Climatic Data Center, Asheville, NC.

Olseth, A. and Skartveit, A., 2001, "Solar irradiance, sunshine duration and daylight illuminance derived from METEOSAT data for some European sites", *Theoretical and Applied Climatology*, **69**, 239-252.

Pinker, R. T., Frouin, R. and Li, Z., 1995, "A review of satellite methods to derive shortwave irradiance", *Remote Sensing of Environment*, **51**, 108-124.

Podestá, G. P., Núñez, L., Villanueva, C. A. and Skansi, M. A., 2004, "Estimating daily solar radiation in the Argentine Pampas", *Agricultural and Forest Meteorology*, **123**, 41-53.

Prescott, J. A., 1940, "Evaporation from a water surface in relation to solar radiation", *Transactions of the Royal Society of South Australia*, **64**, 114-118.

Richardson, C. W., 1981, "Stochastic simulation of daily precipitation, temperature and solar radiation", *Water Resources Research*, **17**, 182-190.

Rivington, M., Matthews, K. B., Bellocchi, G. and Buchan, K., 2006, "Evaluating uncertainty introduced to process-based simulation model estimates by alternative sources of meteorological data", *Agricultural Systems*, **88**, 451-471.

Thorton, P. E. and Running, S. W., 1999, "An improved algorithm for estimating daily solar radiation from measurements of temperature, humidity, and precipitation", *Agricultural and Forest Meteorology*, **93**, 211-228.

Trnka, M., Zalud, Z., Eitzinger, J. and Dubrovský, M., 2005, "Global solar radiation in Central European lowlands estimated by various empirical formulae", *Agricultural and Forest Meteorology*, **131**, 54-76.

Tymvios, F. S., Jacovides, C. P., Michaelides, S. C. and Scouteli, C., 2005, "Comparative study of Ångström's and artificial neural network's methodologies in estimating global solar radiation", *Solar Energy*, **78**, 752-762.

Wilks, D. S. and Wilby, R. L., 1999, "The weather generation game: A review of stochastic weather models", *Progress in Physical Geography*, **23**, 329-357.

Zhang, Q., Xu, C. Y., Zhang, Z., Chen, Y. D., Liu, C. L. and Lin, H., 2008, "Spatial and temporal variability of precipitation maxima during 1960-2005 in the Yangtze River basin and possible association with large-scale circulation", *Journal of Hydrology*, **353**, 3, 215-227.