Cumulative frequency distributions of daily clearness index for temperate and high latitudes

S E Frid and N V Lisitskaya

Joint Institute for High Temperatures of the Russian Academy of Sciences, Izhorskaya 13
Bldg 2, Moscow 125412, Russia

E-mail: s_frid@oirtran.ru

Abstract. Cumulative distributions of daily clearness index, which is based on actinometrical observations from meteorological stations and NASA POWER data, have been constructed for the Russian Federation territory, which spans a wide range of latitudes and covers more than half of the high-latitude Northern Hemisphere. Possible applications of NASA POWER data for the simulation of solar power systems at temperate and high latitude areas have been shown. Daily clearness index distributions were approximated by equations derived from data in tropical climate conditions, with modified dimensionless variables. These equations were found to be adequate for high latitudes, which provides greater confidence that they are universal. It has also been shown here that the selection of the extra-terrestrial radiation calculation method has an influence on the calculated clearness index value, especially for high latitudes. The difference between Cooper’s equation for declination and calculations conducted with NASA POWER data at high latitudes is more than 35%.

1. Introduction

One of the properties of solar energy as a renewable resource is its very small energy flux at earth surface (less than 1.4 kW/m²). This flux is 2–3 orders of magnitude less than typical energy fluxes seen in traditional power equipment. Thus, it becomes necessary to spend considerable funds for energy collection, and it is impossible to construct power equipment using traditional energetic techniques, considering its operation at the design point. Therefore, a designer should simulate the work of a solar power plant or water heater operation, and needs input from solar radiation data to do so. Climate databases and handbooks, as a rule, contain long-term monthly or yearly averaged data that can only be used to estimate solar energy resources. Frequency distributions of solar radiation are more useful than such solar radiation data. These distributions can be considered in terms of clearness indices.

Approximately 60 years ago, Liu and Jordan found that cumulative distributions of daily clearness indices were location independent. This notion is the foundation of the utilizability concept. However, some authors doubted the universal quality of Liu and Jordan’s distributions and constructed daily clearness index distributions for tropical climate and temperate storm belts. We tried to expand their results to temperate and high latitude territories using Russian actinometrical data, due to the many climate zones that are present throughout Russian territory. At the same time, we analysed the application of satellite climate databases instead of actinometrical observation data for the simulation of solar power plant operation.
2. Actinometrical data sources

Solar radiation fluxes and sums at the selected location can be estimated analytically [1–3] but with considerable error. Long-term actinometrical observations are the preferable sources of solar radiation data, as they contain several variables, higher accuracy and longer observation periods. The observation results obtained from actinometrical stations all over the world are accumulated in an open-access database from the World Radiation Data Centre (WRDC) [4]. There are several free [5] or commercial [6, 7] databases with similar information.

One major problem is an insufficient number of actinometrical stations to carry out ground-based measurements. For example, the average distance between stations is approximately 500 km in Russia, increasing to 1000 km in Asian parts of the country [8]. Furthermore, the extrapolation and interpolation of actinometrical data yielding necessary accuracy may be used for distances no greater than 100–130 km [3]. Numerical simulation and data reconstruction techniques were widely applied to solve this problem in recent decades. Reanalysis applied to meteorology, climatology, oceanography, etc can be considered one of the techniques for solar radiation data processing [9,10]. For example, NASA SSE (NASA Surface meteorology and Solar Energy) databases have been constructed [11] based on regular gridded data reconstruction.

A lot of actinometrical data sources are accessible [12, 13]. The International Renewable Energy Agency (IRENA) [14] is the most complete source of such information and includes a large set of interactive maps and links to actual data about different renewable energy resources for many industries. The previously mentioned sources, as well as climate handbooks, generally contain long-term averaged data and can be used for solar energy resource evaluation only. Time series of total solar radiation and other climatic parameters that are required for dynamic simulation of solar power units can be generated from these averaged data [1, 15]. However, these potential results are average annual series that are similar to a typical meteorological year [15] and do not reflect extreme climatic data. Artificial meteorological years are useful for the simulation of power units that have auxiliary energy sources from fossil fuels or electricity and provide a solar fraction much less than 100%. The performance of autonomous solar power units providing a solar fraction near 100% is very sensitive to climate parameter deviations from data of a typical year. Thus, simulations using artificial meteorological years poorly reflect the operation of such power units.

In this connection, the NASA Prediction of Worldwide Energy Resource (NASA POWER) databases [16] become very useful sources of actinometrical information. NASA POWER was conceived in 2003 as an extension of NASA SSE and uses satellite observations for simulating atmospheric processes. Within the NASA POWER project, daily data (top atmosphere solar radiation, surface solar radiation, long wave radiation, air and surface average temperatures, wind speed, etc) are included for 1° × 1° and for 0.5° × 0.5° from 1 July 1983 to date and are openly accessible. The problems surrounding NASA POWER data verification are of particular interest. Our previous verification of the NASA SSE monthly averaged parameter database for the Russian territory [8] demonstrated that relative deviations of average solar radiation sums are less than 10–15% in “solar” months. This error is increased in winter months for several locations.

3. Clearness index distribution

As previously noted, a monthly average of daily total radiation on a horizontal surface can only be suitable for initial estimation of solar energy resources and not for the simulation of long-term solar power unit operation. Insolation value frequency distributions are more detailed solar radiation data. These can commonly be summarized in terms of clearness indices.

A monthly average, a daily or an hourly clearness index, $\bar{K}_T$, $K_T$ or $k_T$, is defined as the ratio of monthly average, daily, or hourly radiation on a horizontal surface to the monthly average, daily, or hourly extra-terrestrial radiation, respectively.
The first reference to this representation method was made by Liu and Jordan [17]. They found that cumulative distributions of daily clearness indices, \( f \), for locations having the same values of \( K_T \) are similar (despite that the locations could have different latitudes and elevations). Based on this fact, they constructed generalized frequency distribution curves of \( K_T \) versus \( f \) for different values of monthly averaged clearness indices. Bendt et al [18] obtained approximations for Liu and Jordan’s distributions using daily total radiation data for 90 meteorological stations in the USA over 20 years. Hollands and Huget [19] have suggested another slightly different approximation.

The theory of cumulative distribution independence of daily or hourly clearness index from geographic locations [17, 18] makes it possible to suppose that the utilizability [20], introduced by Whillier [21] to calculate the performance of flat-plate collectors, has a universal quality. Utilizability is defined as a statistical value [1], which is the fraction of total solar radiation that is received at an intensity higher than critical level. By multiplying the average radiation sum for the period by this fraction, one could find the total utilisable energy for practical solar energy application.

However, the results of Saunier et al [22] for tropical climate and of Olseth and Skartveit [23] for temperate storm belts (areas near 60\(^{\circ}\)N and 60\(^{\circ}\)S) have called into question the universal quality of this theory. We obtained similar results for Russian climatic conditions [24]. It is necessary to note that after [22, 23], researchers interested in daily clearness indices have focused on hourly or shorter durations up to the minute clearness indices, based on solar radiation measurement processing results for 1–3 years [25, 26].

A generalized probability density function \( P(K_T, \bar{K}_T) \), obtained from [18, 19], is in the following form [22]:

\[
P(K_T, \bar{K}_T) = C \left( \frac{K_T^{\max} - K_T}{K_T^{\max}} \right)^n e^{\gamma K_T},
\]

where \( K_T^{\max} \) is a maximum value of clearness index, \( n \) is a parameter; and constants \( C \) and \( \gamma \) are defined from the following normalization conditions:

\[
\int_{K_T^{\min}}^{K_T^{\max}} (K_T)^i P(K_T, \bar{K}_T) dK_T = (\bar{K}_T)^i,
\]

where \( i = 1, 2; K_T^{\min} \) is a minimum value of clearness index. In [18, 19] nonlinear equations for parameter \( \gamma \) had been obtained, and an approximate solution had been noted.

The expression for cumulative frequency of occurrence (integral probability) of daily clearness index values that do not exceed the specified value \( K_T \) (\( \bar{K}_T \) is fixed) is the integral of the generalized probability density function:

\[
f(K_T, \bar{K}_T) = \int_{K_T^{\min}}^{K_T} P(K_T, \bar{K}_T) dK_T.
\]

In [18] \( n = 0, K_T^{\min} = 0.05, K_T^{\max} = 0.6113 + 0.267\bar{K}_T - 11.9(\bar{K}_T - 0.75)^8 \), in [19] \( n = 1, K_T^{\min} = 0, K_T^{\max} = 0.864 \).

An attempt to approximate data for tropical locations has been made by Saunier et al [22], and it has been shown that the use of equations (1)–(3) by [18] does not yield an accurate approximation. Better results could be obtained if the value of \( K_T^{\max} \) were calculated as a parameter of a regression function. Nevertheless, at high \( K_T \) values (near 1), the regression function deviates from experimental values, so authors of [22] have suggested a more complicated expression for \( P(K_T, \bar{K}_T) \):

\[
P(x, \bar{x}) = C x (1 - x) e^{\gamma x},
\]

where \( x = K_T / K_T^{\max} \) (\( 0 \leq x \leq 1 \), because \( K_T^{\min} = 0 \) is assumed, just as in [19]).
Equations for describing $\gamma$ and $C$ are obtained from the normalization conditions and the computation of the mean value of $\bar{K}_T$ (2):

$$\bar{K}_T = \frac{e^{\gamma}(\gamma^2 - 4\gamma + 6) - 2\gamma - 6}{\gamma[e^{\gamma}(\gamma - 2) + \gamma + 2]}, \quad C = \frac{\gamma^3}{e^{\gamma}(\gamma - 2) + \gamma + 2}. \tag{5}$$

This allows for the creation of the density probability distribution function and cumulative clearness index distribution. It should be noted that equation (5) is different from the equation defined by [22] on the left side, of which $\bar{K}_T$ was replaced by $\bar{x}$ mistakenly.

Let us note that numerators and denominators of expressions (5) are equal to 0 when $\bar{K}_T = 0.5$; therefore, it is necessary to redefine them for this value of average clearness index: $\gamma = 0, C = 6$.

4. Formulation of the problem and methods for a solution

This work expands upon the results obtained by [24]. Our goals were as follows:

- to elaborate upon the cumulative distributions of daily clearness indices for temperate and high latitudes;
- to analyze the usage of NASA POWER data for distribution development;
- to fit the distributions with approximation equations.

We used actinometrical data from the Russian Federation because wide latitude ranges are present in this territory, and the country covers more than a half of the high-latitude terrestrial area of North Hemisphere. Our computations were performed using the following:

- surface data from 27 meteorological stations during the 1964–2015 period from WRDC;
- information from the NASA POWER daily time series for $1^\circ \times 1^\circ$ grid squares (3199 squares) during the 1983–2016 period.

Cumulative data has been represented in the database made by the Joint Institute for High Temperatures of the Russian Academy of Sciences.

The literature that discusses clearness index distributions does not determine how extra-terrestrial solar radiation should be calculated. The extra-terrestrial solar radiation is direct, so, mathematically, only geometric calculations are possible [1]. At the same time, the top of the atmosphere daily insolation from NASA POWER data ($H_0^{\text{NASA}}$) is different from the calculated extra-terrestrial solar radiation ($H_0^{\text{calc}}$). Figure 1 shows the ratio of $H_0^{\text{NASA}}$ to $H_0^{\text{calc}}$ after taking into account ellipticity of the orbit of planet Earth, Cooper’s approximate equation for solar declination [1], and more accurate equations used by Spencer [1, 27]. The datasets were compared for the Russian Federation area; NASA POWER data have been averaged for years within the period under review. The significant part of the Russian Federation territory is located nearby and behind the Arctic Circle, which is characterized by an insufficient amount of surface solar radiation, and approaches zero during the Arctic night. Under these conditions, solar power units are out of operation, so it is possible to exclude the days with a solar radiation less than 0.5 kWh from the study.

The computation has demonstrated that, at crude approximation, the difference between calculated daily extra-terrestrial solar radiation and the corresponding NASA POWER data can exceed 35% at a latitude of 60°N. This means declination is heavily influenced by the value of the zenith angle and incidental radiance on a horizontal surface. At the same time, applying Spencer’s equations yields a difference of no more than 8% at the nearby latitude of 70°N. It is possible to obtain an even better result by using the solar position algorithm of Reda and Andreas [28], but we have not investigated this. To ensure that the obtained results could be compared with the results of other researchers, the daily extra-terrestrial radiation was defined in terms of Spencer’s equations.
Figure 1. The ratio of daily extra-terrestrial solar radiation based on NASA POWER data, estimated by taking into account ellipticity of the orbit of planet Earth and Cooper’s approximate equation for solar declination (open circles) [1], and more accurate equations used by Spencer (full circles) [1, 27].

Cumulative occurrence frequency of daily clearness index values was constructed as described by Bendt et al [18]: monthly average $\bar{K}_T$ values were determined with the specification that all months having more than one day without data were eliminated. Additionally, days with extra-terrestrial solar radiation values less than 0.5 kWh were excluded. We then selected months for which $\bar{K}_T$ was within the limits of ±0.01 from values of 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, or 0.7. Then, for each of these months and $\bar{K}_T$ values, the daily distributions of $K_T$ were added to the final histogram (with an interval of 0.02). The histograms were integrated to obtain cumulative frequency and were normalized for the total quantity of values. The results of these calculations are discussed below.

5. Discussion of results
The results of plotting cumulative distribution curves of days with various values of $\bar{K}_T$ are shown in figure 2 (distributions with values of $\bar{K}_T = 0.35, 0.45, 0.55, 0.65$ are similar and are not shown in the figure). It is necessary to note the reasonably good agreement between NASA POWER data and WRDC measured radiation; therefore, this makes it possible to use NASA POWER data in simulations of solar power systems.

An attempted use of equations (1)–(3) for obtained cumulative distribution with fitted values of $K_{T_{\text{max}}}$, as was done by Saunier et al [22], did not result in a close approximation. This is illustrated by dotted curves in figure 2.

In contrast to the results obtained by [22], we did not obtain the expected good agreement between regression curves and initial data. There is a significant difference at $f = 1$. More
Figure 2. Cumulative distributions of daily clearness index for various values of $\bar{K}_T$: squares—WRDC data; solid lines—NASA POWER data; dotted lines—regression using (1)–(3); dashed lines—regression curves for WRDC data; dash-dot lines—regression curves for NASA POWER data.

desirable results have nonetheless been shown by the methods used in this study, as described by [22] with some modifications.

For the data regression, we used equation (4) which was integrated with $K_T$, as directed by equation (3) and taking into account the redefinition of equation (5) at $\bar{K}_T = 0.5$:

$$f(K_T, \bar{K}_T) = \begin{cases} 
\frac{\gamma + 2 - e^{\gamma x} [(\gamma x)^2 + (1 - \gamma x)(\gamma + 2)]}{e^{\gamma (\gamma - 2)} + (\gamma + 2)} & \text{for } \bar{K}_T \neq 0.5, \\
x^2(3 - 2x) & \text{for } \bar{K}_T = 0.5,
\end{cases}$$

(6)

where $K_T^\text{min} \neq 0$, unlike the study done by [22], and $x = (K_T - K_T^\text{min})/(K_T^\text{max} - K_T^\text{min})$, as in the study done by [23].

Parameter $\gamma$ was defined by the numerical solution of equation (5). The resulting deviation from the $\bar{K}_T$ calculation, using an obtained $\gamma$ with equation (5), is not more than $6 \times 10^{-4}\%$. Results can be fitted with a cubic polynomial (the coefficient of determination [29] $R^2 = 0.9999$):

$$\gamma = 103.4K_T^\beta - 155.0K_T^\alpha + 96.40K_T - 22.36.$$  

Values of $K_T^\text{min}$ and $K_T^\text{max}$, obtained by the regression as well as by an evaluation of their standard deviation ($\sigma$) and coefficients of determination ($R^2$), are shown in table 1. The values of $R^2$ near 1 indicate a close approximation of WRDC data (dashed line in figure 2) and NASA POWER data (dash-dot line in figure 2), as obtained from equation (6). Similar calculations have been performed for cumulative distribution, which were constructed using daily extra-terrestrial radiation of NASA POWER data (results are shown in table 2). WRDC data and NASA POWER data regression curves (see figure 2) fit more closely than data interpolation curves.
Table 1. Values of $K_{T}^{\text{min}}$ and $K_{T}^{\text{max}}$ obtained by the regression, evaluation for standard deviation $\sigma$ and coefficient of determination $R^2$ (for calculated extra-terrestrial radiation).

| $\tilde{K}_{T}$ | $K_{T}^{\text{min}}$ | $\sigma$ | $K_{T}^{\text{max}}$ | $\sigma$ | $R^2$ | $K_{T}^{\text{min}}$ | $\sigma$ | $K_{T}^{\text{max}}$ | $\sigma$ | $R^2$ |
|----------------|---------------------|--------|---------------------|--------|------|---------------------|--------|---------------------|--------|------|
| 0.30           | 0.086               | 0.002  | 0.741               | 0.007  | 0.9993 | 0.104               | 0.001  | 0.703               | 0.005  | 0.9997 |
| 0.35           | 0.069               | 0.001  | 0.821               | 0.004  | 0.9998 | 0.113               | 0.001  | 0.744               | 0.004  | 0.9998 |
| 0.40           | 0.068               | 0.003  | 0.858               | 0.008  | 0.9991 | 0.118               | 0.002  | 0.781               | 0.004  | 0.9997 |
| 0.45           | 0.091               | 0.006  | 0.856               | 0.013  | 0.9971 | 0.128               | 0.003  | 0.808               | 0.008  | 0.9990 |
| 0.50           | 0.127               | 0.007  | 0.847               | 0.007  | 0.9964 | 0.157               | 0.006  | 0.812               | 0.006  | 0.9977 |
| 0.55           | 0.194               | 0.007  | 0.817               | 0.014  | 0.9972 | 0.212               | 0.006  | 0.798               | 0.013  | 0.9973 |
| 0.60           | 0.272               | 0.007  | 0.796               | 0.013  | 0.9971 | 0.288               | 0.005  | 0.781               | 0.009  | 0.9984 |
| 0.65           | 0.310               | 0.006  | 0.806               | 0.010  | 0.9988 | 0.359               | 0.005  | 0.780               | 0.008  | 0.9991 |
| 0.70           | 0.332               | 0.006  | 0.832               | 0.010  | 0.9988 | 0.327               | 0.018  | 0.827               | 0.029  | 0.9908 |

Table 2. Values of $K_{T}^{\text{min}}$ and $K_{T}^{\text{max}}$ obtained by the regression, evaluation for standard deviation $\sigma$ and coefficient of determination $R^2$ (for extra-terrestrial radiation from NASA POWER data).

| $\tilde{K}_{T}$ | $K_{T}^{\text{min}}$ | $\sigma$ | $K_{T}^{\text{max}}$ | $\sigma$ | $R^2$ | $K_{T}^{\text{min}}$ | $\sigma$ | $K_{T}^{\text{max}}$ | $\sigma$ | $R^2$ |
|----------------|---------------------|--------|---------------------|--------|------|---------------------|--------|---------------------|--------|------|
| 0.3            | 0.065               | 0.002  | 0.799               | 0.008  | 0.9992 | 0.094               | 0.001  | 0.720               | 0.004  | 0.9998 |
| 0.4            | 0.062               | 0.004  | 0.871               | 0.012  | 0.9978 | 0.112               | 0.002  | 0.790               | 0.006  | 0.9995 |
| 0.5            | 0.134               | 0.009  | 0.842               | 0.009  | 0.9942 | 0.166               | 0.006  | 0.803               | 0.006  | 0.9976 |
| 0.6            | 0.256               | 0.008  | 0.808               | 0.015  | 0.9968 | 0.293               | 0.006  | 0.777               | 0.010  | 0.9981 |
| 0.7            | 0.301               | 0.008  | 0.817               | 0.013  | 0.9980 | 0.428               | 0.008  | 0.795               | 0.014  | 0.9973 |

enabling the construction of generalized curves by averaging $K_{T}^{\text{min}}$ and $K_{T}^{\text{max}}$ with identical $\tilde{K}_{T}$. The results of such averaging, shown in table 3, fit the WRDC data and NASA POWER data reasonably well.

The noticeable differences between regression curves and data occur only for $K_{T}$ near $K_{T}^{\text{min}}$ ($f < 0.1$, corresponding to heavily overcast days and days near Arctic night) and $K_{T}^{\text{max}}$ ($f > 0.8$, corresponding to very clear days). Both these areas are situated at the boundaries of $f$ and are characterized by small values of $P(K_{T}, \tilde{K}_{T})$ (at the boundaries, $P(K_{T}, \tilde{K}_{T})$ reaches 0 [18]); therefore, they have no great influence on the estimation of solar power unit performance. Nevertheless, the regression accuracy might be increased upon investigation of distribution for separated latitude limits, climate zones and annual seasons.

To calculate cumulative distribution of clearness index for various values of $\tilde{K}_{T}$, it is necessary to define the dependence of $K_{T}^{\text{min}}$ and $K_{T}^{\text{max}}$ on it. We obtained the dependence by processing the data in table 3 as follows. $K_{T}^{\text{min}}(\tilde{K}_{T})$ is closely fitted by polynomial equation of 4 degree, and $K_{T}^{\text{max}}(\tilde{K}_{T})$ by a cubic equation:

$$K_{T}^{\text{min}} = a_0 + a_1\tilde{K}_{T} + a_2\tilde{K}_{T}^2 + a_3\tilde{K}_{T}^3 + a_4\tilde{K}_{T}^4, \quad (7)$$

$$K_{T}^{\text{max}} = b_0 + b_1\tilde{K}_{T} + b_2\tilde{K}_{T}^2 + b_3\tilde{K}_{T}^3. \quad (8)$$

The regression curves are shown in figure 3, as well as coefficients $a_i$, $b_i$, the estimate of their standard deviations and their coefficients of determination ($R^2$) are presented in table 4.
Table 3. Average values of $K_T^{\min}$ and $K_T^{\max}$ for approximation by the equation (6).

| $\bar{K}_T$ | Average in table 1 | Average in tables 1 and 2 |
|-------------|---------------------|---------------------------|
|             | $K_T^{\min}$       | $K_T^{\max}$             | $K_T^{\min}$       | $K_T^{\max}$             |
| 0.30        | 0.095               | 0.013                     | 0.722             | 0.027                     | 0.087               | 0.016                     | 0.741             | 0.042                     |
| 0.35        | 0.091               | 0.031                     | 0.782             | 0.055                     | —                   | —                         | —                 | —                         |
| 0.40        | 0.093               | 0.036                     | 0.820             | 0.054                     | 0.090              | 0.029                     | 0.825             | 0.046                     |
| 0.45        | 0.110               | 0.026                     | 0.832             | 0.034                     | —                   | —                         | —                 | —                         |
| 0.50        | 0.142               | 0.021                     | 0.830             | 0.025                     | 0.146              | 0.019                     | 0.826             | 0.022                     |
| 0.55        | 0.203               | 0.013                     | 0.808             | 0.013                     | —                   | —                         | —                 | —                         |
| 0.60        | 0.280               | 0.012                     | 0.788             | 0.011                     | 0.277              | 0.017                     | 0.790             | 0.014                     |
| 0.65        | 0.335               | 0.035                     | 0.793             | 0.019                     | —                   | —                         | —                 | —                         |
| 0.70        | 0.329               | 0.004                     | 0.830             | 0.004                     | 0.347              | 0.056                     | 0.818             | 0.016                     |

Figure 3. $K_T^{\min}$ and $K_T^{\max}$ dependence on $\bar{K}_T$: full squares—values of $K_T^{\min}$ and $K_T^{\max}$; solid lines—regression curves for $K_T^{\min}$ and $K_T^{\max}$ averaged in table 1; dashed lines—regression curves for $K_T^{\min}$ and $K_T^{\max}$ averaged in tables 1 and 2.

Thus, we can plot the cumulative distribution of the daily clearness index for various values of monthly averaged clearness index. The fact that equation (6) does not differ from that used by Saunier et al [22] under completely different climate conditions, and that the definition of a non-dimensional variable $x$ is a more general concept than the notion used by Saunier et al [22], offers hope that the functional dependence (6) is universal in quality.
Table 4. Regression coefficients $a_i$ and $b_i$, evaluation for standard deviations $\sigma$ and coefficients of determination $R^2$ for $K_{T_{min}}(\overline{K}_T)$ and $K_{T_{max}}(\overline{K}_T)$ averaged from table 1 (T1) and both tables 1 and 2 (T2).

|   |   |   |   | 1 | 2 | 3 | 4 |   |   |
|---|---|---|---|---|---|---|---|---|---|
| T1 | $a_i$ | -3.1 | 0.5 | 30.8 | 4.0 | -108.5 | 12.6 | 163.2 | 17.1 | -87.4 | 8.4 | 0.9996 |
|   | $b_i$ | -1.3 | 0.2 | 13.1 | 0.9 | -25.8 | 1.9 | 16.3 | 1.2 | — | — | 0.9826 |
| T2 | $a_i^a$ | -2.2 | — | 22.8 | — | -81.5 | — | 124.2 | — | -66.9 | — | — |
|   | $b_i$ | -1.0 | 0.2 | 11.0 | 1.2 | -21.6 | 2.5 | 13.6 | 1.6 | — | — | 0.9569 |

In this case, the number of points equals the number of the polynomial coefficients, and the approximation is reduced to interpolation. Thus, the standard deviation of coefficients and $R^2$ are not estimated.

6. Conclusions

Our inferences are as follows:

- Cumulative distribution curves of daily clearness index for Russian actinometrical data from ground meteorological station measurements and for NASA POWER data are almost identical. A possibility of NASA POWER data applicability for the simulation or solar power units at temperate and high latitude areas has been shown.

- Calculated cumulative distribution of clearness index is closely fitted by equation, which is obtained from the equation suggested by Saunier et al [22]. Calculated values for regression coefficients allow us to plot the distribution for a wide range of climate conditions. An adequacy of such an equation for high latitudes, similar to that used by Saunier et al [22] for tropical climate conditions, allows for the conclusion that functional dependencies are universal in quality.

- The selection of the extra-terrestrial radiation calculation method has great influence on the calculated clearness index value, especially at high latitudes. The difference between using Cooper’s equation for declination and calculations with NASA POWER data at high latitudes is significant (more than 35%). More precise Spencer’s equations provide much higher accuracy (approximately 8% difference).

Acknowledgments

This work is supported by the Russian Science Foundation, grant No. 16-19-10659. We thank Dr Oleg Popel, Dr Sophia Kiseleva and Alexey Tarasenko for critical reading the manuscript and helpful discussions.

References

[1] Duffie J A and Beckman W A 2013 Solar Engineering of Thermal Processes 4th ed (Hoboken, NJ: Wiley)
[2] Kondratiev K Ya 1969 Radiation in the Atmosphere (New York: Academic Press)
[3] Pivovarova Z I 1977 Radiatsionnyie Harakteristikii Klimata SSSR (Leningrad: Gidrometeoizdat)
[4] WRDC World radiation data centre URL http://wrdc.mgo.rssi.ru
[5] RETScreen URL http://www.nrcan.gc.ca/energy/software-tools/7465
[6] Meteonorm URL http://meteonorm.com/
[7] 3TIER Renewable energy assessment and forecasting URL http://www.3tier.com/en/
[8] Popel O S, Frid S E, Kolomiets Yu G, Kiseleva S V and Terekhova E N 2010 Atlas Resursov Solnechnoj Energi na Territorii Rossii (Moscow: JIHT RAS)
[9] NCEP/NCAR Reanalysis project URL https://www.esrl.noaa.gov/psd/data/reanalysis/
[10] ERA Interim reanalysis system URL http://apps.ecmwf.int/datasets/data/interim-full-daily/
[11] NASA Surface meteorology and solar energy URL https://eosweb.larc.nasa.gov/sse/
[12] SoDa Solar energy services for professionals URL http://www.soda-is.com/eng/index.html
[13] SWERA Solar and wind energy resource assessment URL http://maps.nrel.gov/SWERA
[14] IRENA Global atlas for renewable energy URL http://irena.masdar.ac.ae/
[15] Knight K M, Klein S A and Duffie J A 1991 Sol. Energy 46 109–20
[16] NASA Prediction of worldwide energy resource URL http://power.larc.nasa.gov/
[17] Liu B Y H and Jordan R C 1960 Sol. Energy 4 1–19
[18] Bendt P, Collares-Pereira M and Rabl A 1981 Sol. Energy 27 1–5
[19] Hollands K G T and Huget R G 1983 Sol. Energy 30 195–209
[20] Liu B Y H and Jordan R C 1963 Sol. Energy 7 53–74
[21] Whillier A 1953 Solar Energy Collection and Its Utilization for House Heating Ph.D. thesis MIT Cambridge, MA
[22] Saunier G Y, Reddy T A and Kumar S 1987 Sol. Energy 38 169–77
[23] Olseth J A and Skartveit A 1984 Sol. Energy 33 533–42
[24] Frid S E, Popel O S, Lisitskaya N V and Kiseleva S V 2017 Dokl. Phys. 62 278–80
[25] Hollands K G T and Suehrcke H 2013 Sol. Energy 96 103–12
[26] Pavanello D, Zaaiman W, Colli A, Heiser J and Smith S 2015 J. Atmos. Sol.-Terr. Phys. 130–1 142–50
[27] Spencer J W 1971 Search 2 172
[28] Reda I and Andreas A 2004 Sol. Energy 76 577–89
[29] Reddy T A 2011 Applied Data Analysis and Modeling for Energy Engineers and Scientists (Springer)