Controlling the quality of solar irradiation data by means of a web service

Michael Geiger, Lamissa Diabaté, Lionel Ménard, Lucien Wald

Geiger M., Diabaté L., Ménard L., Wald L., Controlling the quality of solar irradiation data by means of a web service. European Geophysical Society, 27th General Assembly, Geophysical Research Abstracts, vol. 4, April 2002, EGS02-A-03419.

CONTROLLING THE QUALITY OF SOLAR IRRADIATION DATA BY MEANS OF A WEB SERVICE

Geiger M. (1), Diabaté L. (2), Ménard L. (1), Wald L. (1)
(1) Groupe Télédétection & Modélisation, Centre d'Energétique, Ecole des Mines de Paris, BP 207, 06904 Sophia Antipolis cedex, France. Tel.: +33 (0) 4.93.95.74.49 Fax: +33 (0) 4.93.95.75.35. Corresponding author: L. Wald, e-mail : lucien.wald@ensmp.fr.
(2) UFAE / GCMI, BP E4018, 410, avenue Van Vollenhoven, Bamako, Mali.

ABSTRACT.

The control of the quality of the irradiation data is often a prerequisite to the further processing of the data. Though data are usually controlled by meteorological offices, the sources are so numerous that the user often faces time-series of measurements containing questionable values. As customers of radiation data, we established our own procedures to screen time-series of measurements. Since this problem of quality control is of concern to many researchers and engineers and since it is often a lengthy and tedious task, we decided to make this screening procedure available to everyone as a web service. This service is the purpose of this paper. The objective is not to perform a precise and fine control, an objective out of reach without details on the site and instruments, but to perform a likelihood control of the data and to check their plausibility. This is achieved by comparing observations with some expectations based upon the extraterrestrial irradiation and a simulation of the irradiation for clear skies. This service is available to everyone on the Web site www.helioclim.net. It offers a very convenient mean to check time-series of irradiation: data are input in a HTML page by a copy and paste procedure and the return is also a HTML page that can be analyzed in details for the data flagged as suspicious.

1. Controlling the quality of irradiation measurements

The measurements of solar global radiation on the ground level are made primarily with pyranometers. These hemispherically sensitive sensors measure the downward solar radiation flux density on a horizontal surface. The properties of pyranometers which are of concern when evaluating the accuracy and quality of radiation measurements are: sensitivity, stability, response time, cosine response, azimuth response, linearity, temperature response, and spectral response (ESRA 2000; ISO-9060 1990). The deviations are primarily caused by the directional sensitivity of the individual receiver surfaces and by effects of the glass dome which covers the upward-facing sensor. Some errors that appear are due to a dust layer or snow or ice on the dome. Other errors may occur from inaccurate horizontal adjustment of the receiver surface, moisture inside the dome, etc. Achievable relative uncertainties at the 95 per cent confidence level are about 3, 8 and 20 per cent for operational pyranometers of high, good and moderate quality in case of hourly sums. In case of daily sums, relative uncertainties of 2, 5 and 10 per cent can be achieved (WMO, 1981).

When taking observations and processing their results, errors may be made which should be detected and, as far as possible, corrected. A reliable estimate of errors is only possible by the final user, when information about the pyranometer is available. This is seldom the case. The data from radiation measurements should be checked in two stages: technical control, and quality control. When observations are acquired from national meteorological services, it is usually assumed that the measurements first went through a technical control by these services. The process of quality control is distinct from technical control. Not only is the correctness of arithmetic calculations checked but also agreement of every radiation parameter with the sun elevation and other parameters. The quality control of radiometric observations is guided by basic physical principles which determine the relations between individual radiation parameters; it may also take into account statistical knowledge on the spatial and temporal variability of these parameters (ESRA 2000).

Quality control may be a lengthy task. Most often, customers do not have the necessary information to perform a precise and fine control. Accordingly, they are trusting the procedures used by the meteorological offices and are ready to use the data in confidence. However, the sources of information are very numerous and do not exhibit the same reliability. It often happens that the user faces time-series of measurements containing questionable values. The reasons are numerous, ranging from non-standard values for the absence of measurements (e.g., 0 instead of -999) to typing errors if the processing chain of measurements is not entirely digital.

As customers of radiation data, we felt the need to establish our own procedures to screen time-series of measurements. The objective is not to perform a precise and fine control, an objective out of reach without details on the site and instruments, but to flag the data that are obviously questionable. Examples are null values when -999 is expected, values much larger than the extraterrestrial irradiation or positive values when the sun is well below the horizon. The aim of these procedures are to perform a likelihood control of the data and to check the plausibility of the measurements. This is performed by comparing observations with those expected. Since this problem of quality control is of concern to many researchers and engineers, we decided to make this screening procedure available to everyone as a web service. This service is the purpose of this paper.

2. Quality control procedures

The quality control procedure adopted in the present paper meets the guidelines of the World Meteorological Organisation (WMO 1981) and implements some of the procedures used in the construction of the European Solar Radiation Atlas (ESRA 2000). The procedure applies to daily or hourly sums of global irradiation. It applies to individual values or time-series. Because of the limited amount of information, the procedure is restricted to the analysis of each measurement with respect to the extraterrestrial irradiation and the extreme that can be expected for this particular instant. The procedure is exactly the same for daily and hourly irradiation.

2.1. Quality control for maximum values of daily sums of global solar irradiation

First of all, each observed daily sum G_d (resp. hourly sum G_h) should be less than the daily (resp. hourly) extraterrestrial irradiation received by a horizontal plane located at the same geographical co-ordinates. Otherwise it is flagged as questionable.

In addition, the daily irradiation should only exceed by a small amount the daily irradiation G_{cd} (resp. hourly irradiation G_{ch}) likely to be observed under exceptionally clear skies with a high atmospheric transparency. The clear sky irradiation G_{cd} (resp. G_{ch}) is assessed by the means of the ESRA model for clear sky. This model is based upon parameterization formula by Kasten (in ESRA 1984 and Kasten 1996) and has been detailed in Rigollier *et al.* (2000). Corrections were brought to this model as proposed by John Page (personal communication). The formula for the Rayleigh optical thickness in Rigollier *et al.* behaves incorrectly with terrain altitude in the original model. This was not evidenced in this paper because the sites used for validation have altitudes less than 500 m. Europe as a whole offers a few areas of altitudes larger than 1000 m. Other continents such as Africa or Asia have large areas of larger altitude and such corrections are necessary. Using the notations of Rigollier *et al.*, the beam transmittance under cloudless skies is given by their Equation 2:

$$exp\ [-0.8662\ T_L(AM2)\ m\ \delta_R(m)] \qquad (1)$$

where $T_L(AM2)$ is the Linke turbidity factor for an air mass equal to 2, m is the relative optical air mass and $\delta_R(m)$ is the integral Rayleigh optical thickness. The air mass m is given by (their Equation # 3):

$$m(\gamma_S) = (p/p_0) / [\ sin\ \gamma_S + 0.50572\ (\gamma_S + 6.07995)^{-1.6364}] \qquad (2)$$

where γ_s is the solar altitude angle, possibly corrected for refraction effects, and the station height correction is given by their Equation # 4: $(p/p_0) = exp\ (-z / z_h)$, where z is the site elevation and z_h is the scale height of the Rayleigh atmosphere near the Earth surface, equal to 8434.5 meters. The parameterization of the Rayleigh optical thickness is given by their Equation # 7:

$$if\ m \leq 20,\ (\gamma_S \geq 1.9°),\ 1/\delta_R(m) = 6.6296 + 1.7513m - 0.1202m^2 + 0.0065m^3 - 0.00013m^4 \qquad (3)$$

$$if\ m > 20,\ (\gamma_S < 1.9°),\ 1/\delta_R(m) = 10.4 + 0.718m$$

In the modified ESRA model, the equations become: A VIRER !!!!!!

$$m(\gamma_S) = 1 / [\ sin\ \gamma_S + 0.50572\ (\gamma_S + 6.07995)^{-1.6364}] \qquad (4)$$

Let $corr\_\delta_R(p/p_0)$ be the correction of the integral Rayleigh optical thickness due to the elevation of the site. John Page has determined this function for two values of (p/p_0):

$$corr\_\delta_R\,(0.75) = 1.248174 - 0.011997\, m(\gamma_S) + 0.00037\, m^2(\gamma_S) \quad (5)$$

$$corr\_\delta_R\,(0.50) = 1.68219 - 0.03059\, m(\gamma_S) + 0.00089\, m^2(\gamma_S)$$

Given that $corr\_\delta_R(1)=1$ and assuming that $corr\_\delta_R(p/p_0) = corr\_\delta_R(0.5)$ for $(p/p_0)>0.5$, $corr\_\delta_R(p/p_0)$ can be determined for any (p/p_0) by piecewise linear interpolation. The integral Rayleigh optical thickness is thus given by:

$$\text{if } m \leq 20,\ (\gamma_S \geq 1.9°)$$

$$1/\delta_R(m) = corr\_\delta_R\,(p/p_0) / [6.625928 + 1.92969m - 0.170073m^2 + 0.011517m^3 - 0.000285m^4]$$

$$\text{if } m > 20,\ (\gamma_S < 1.9°),\ 1/\delta_R(m) = 10.4 + 0.718\, m\, (p/p_0) \quad (6)$$

The observed value G_d (resp. G_h) is compared to this clear sky value G_{cd} (resp. G_{ch}). The second condition is that G_d should be less than $(1.1\ G_{cd})$. Otherwise the observation for this day is flagged as questionable.

The clear-sky model makes use of the Linke turbidity factor, a parameter quantifying the atmospheric visibility (aerosol plus water vapor) under clear skies. Values of this factor are available on climatological basis (Angles *et al.* 1998, 1999; the SoDa web site: http://soda.jrc.it). It varies from site to site and from month to month. Typical values can be found in WMO (1981). A value of 1 stands for very optically clean atmosphere and may be adopted to control the upper limit of the daily irradiation. It is recommended to perform this test twice with two different Linke turbidity factors (e.g., 1 and 3 or 2 and 4) to assess the sensitivity of the results to this factor.

2.2. Quality control for minimum values of daily sums of global solar irradiation

The third condition to be fulfilled is that the daily sum G_d (resp. hourly sum G_h) is greater than a minimum value expected in continuous overcast conditions. This is equivalent to saying that the clearness index KT_d (resp. KT_h) should be greater than a minimum value. This minimum value was set up according to the analysis of the collected data and to the minima of KT_d (resp. KT_h) found in the last and previous editions of the European solar radiation atlas (ESRA 1984, 2000). Accordingly, the smallest acceptable value for KT_d and KT_h is 0.03, which represents a heavily overcast sky.

2.3. The case of very high latitude sites

Special attention was paid to the observed solar irradiation data for the sites at very high latitudes, when the elevation of the sun above horizon at noon is less than 2°. When the extraterrestrial irradiation on the horizontal plane G_{0d} is greater than 1 J cm^{-2} (2.78 Wh m^{-2}), the upper acceptance limit is set to $(2\ G_{cd})$ and the lower acceptance limit is set to $(0.015\ G_{0d})$. When the extraterrestrial irradiation is less than 1 J cm^{-2}, values of G_d are still accepted if they are less than 10 J cm^{-2} (27.78 Wh m^{-2}) and no minimum value is set. The reasons for doing this are that on the one hand measuring instruments have an optimum precision of 5 W m^{-2} and that on the other hand atmospheric phenomena like refraction are present (ESRA 2000).

3. Web service

The quality control procedure is part of an on-going effort of the Group 'Télédétection & Modélisation' of the Ecole des Mines de Paris to provide free of charge valuable tools and information to the solar radiation community through the most know and used media, namely the world wide web.

The web server HelioClim (http://www.helioclim.net) hosts the quality control procedure and other solar radiation oriented sizing tools. It is built from open source software. Apart the fact that those software can be used free of charge, they are rated among the most used, reliable, fast and secure tools available on the market (Netcraft 2001). Installed on a Linux box an Apache web server is delivering result by the mean of CGI (Common Gateway Interface) based scripts (see references on tools Linux and Apache). The interesting feature of the CGI-based quality control procedure is the re-use of source code (mainly C code), procedures and libraries already written, used and validated by the scientific community in various projects like ESRA (2000), Helioserve (Angles *et al.* 1998) and SoDa (Wald 2000). The HelioClim web site use also extensively MySQL a relational database to store, extract and manipulate solar radiation parameters and PHP, a server-side, cross-platform, HTML embedded scripting language (see references on tools MySQL and PHP).

The user interface has been kept simple for a maximum reliability over the network bandwidth. The user is requested to provide information to compute the quality control procedure: geographical co-ordinates, elevation and dates (if single value). Detailed help HTML pages are available to better understand and correctly fill the forms of each quality control procedure. Also in line are a document explaining the algorithm used in the calculation and references to articles, web site of interest and publications on solar radiation topic.

The quality control procedure has been divided in four HTML documents:

1. Daily irradiation for a single day
2. Daily irradiation for several days
3. Hourly irradiation for a single hour
4. Hourly irradiation for several hours

Simple forms are used inside the HTML document allowing a one-click per page result. Figure 1 exhibits a snapshot of the input and output forms for screening a single day. The service was tested by the means of several time-series of hourly and daily irradiations available in the CD-ROM of the European solar radiation atlas (ESRA 2000).

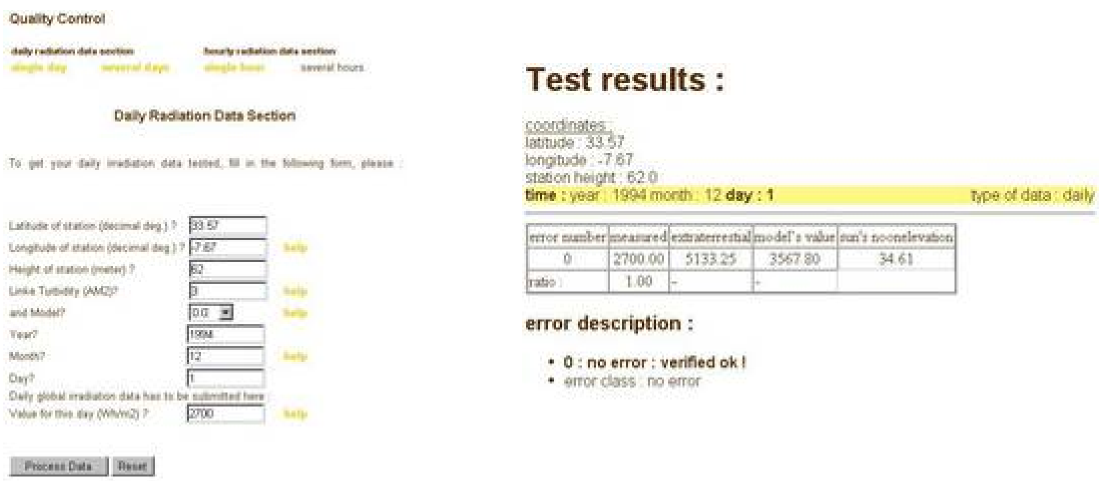

Figure 1. Screening daily irradiation. a) the form for screening a single day. b) the form of the screening results.

4. Example

This web service was used to perform a quality screening of time-series of daily irradiation for the period 1994-1998 and for sites in Africa. Data were provided by the networks belonging to the World Meteorological Organization. These data are prominent in constructing global maps of the Linke turbidity factor within the SoDa project (Wald 2000) as well as in assessing the accuracy of the method Heliosat-II, aiming at retrieving solar irradiations from an appropriate processing of satellite data (Rigollier, Wald 1999). Hence, it is important that these data are properly screened.

Figure 2 exhibits an extract of the HTML page containing the results of the test. The site under concern is Casablanca (Senegal). Here are only reported three months of 1994 and nine of 1998 for illustration. Each line represents a month. For each line, there are 31 cells, one per day. The character "V" stands for "verified" and means that the value is OK with respect to the procedure. The days not passing the test are flagged with a number taking values 10, 11, 12 or 21 to 24.

The number 10 means that the daily irradiation is greater than the extraterrestrial value. An example is the day 24 November 1994. When clicking on a cell flagged as incorrect, another page appears, similar to Figure 1b. This new page gives the observed irradiation for this instant (day or hour), the extraterrestrial irradiation, the irradiation for clear sky and the solar elevation at noon. It permits to see how large is the discrepancy between the observed value and the threshold under concern. Values may be rehabilitated or discarded.

In August 1998, two days (13 and 21) exhibit values that are well exceeding the expected value for very clear skies. Here, the Linke turbidity factor was set to 2, a value well below the typical values for this site. It follows that the irradiations for these two days are overestimated and should be rejected. They are flagged with the number 11.

Missing values are set to 0. Such days (e.g., 31 November 1994, 29, 30 and 31 February 1998) are flagged with the number 12.

The other flagging values 21 to 24 are used for the tests for low solar elevations at noon.

5. Conclusion

This web service is available to everyone on a free basis. Data are input in the HTML page by a copy and paste procedure. The return is also a HTML page that can be saved and further processed if necessary. This service calls upon up-to-date algorithms that have proven quality. They may be changed as progresses are made. This service offers a very convenient mean to users to check time-series of irradiation without writing and testing themselves the procedures.

6. Acknowledgements

The content of the present paper was influenced by the chapter II of the ESRA handbook. Fruitful discussions were held with John Page (Sheffield, United Kingdom) and Gerhard Czeplak (Hamburg, Germany). The program Leonardo of the Commission of the European Communities offered a grant for the stay of Michael Geiger in France. This work was also partly supported by the Commission under Contract Number IST-1999-12245.

7. References

Angles J., Ménard L., Bauer O., Wald L. (1998) A Web server for accessing a database on solar radiation parameters. In *Proceedings of the Earth Observation & Geo-Spatial Web and Internet Workshop '98,* Josef Strobl & Clive Best (Eds), Salzburger Geographische Materialien, Universität Salzburg, Salzburg, Austria, Heft 27, pp. 33-34. http://www.sbg.ac.at/geo/eogeo/authors/angles/angles.html

Angles J., Ménard L., Bauer O., Rigollier C., Wald L. (1999) A climatological database of the Linke turbidity factor. In Proceedings of the ISES Solar World Congress 1999, Jerusalem, Israel, July 4-9, 1999, volume I, pp 432-434.

Apache (2001) The Apache Software Foundation: http://www.apache.org.

ESRA (1984) *European Solar Radiation Atlas. Second Improved and Extended Edition,* Vols. I and II. (1984) W. Palz (Ed.). Commission of the European Communities, DG Science, Research and Development, Report No. EUR 9344, Brussels.

ESRA (2000) *European Solar Radiation Atlas*, includ. CD-ROM. Edited by J. Greif, K. Scharmer. Scientific advisors: R. Dogniaux, J. K. Page. Authors : L. Wald, M. Albuisson, G. Czeplak, B. Bourges, R. Aguiar, H. Lund, A. Joukoff, U. Terzenbach, H. G. Beyer, E. P. Borisenko. Published for the Commission of the European Communities by Presses de l'Ecole, Ecole des Mines de Paris, France. http://www.helioclim.net/esra/

ISO 9060 (1990) Solar energy - specification and classification of instruments for measuring hemispherical solar and direct solar radiation. Instrumental standard ISO 9060. International Organisation for Standardization (ISO), Geneva, Switzerland.

Kasten F. (1996) The Linke turbidity factor based on improved values of the integral Rayleigh optical thickness. *Solar Energy*, **56**, 239-244.

Linux (2001) Linux On line: http://www.linux.org.

MySQL: http://www.mysql.org

Netcraft (2001) Netcraft web server survey: http://www.netcraft.com/survey/ Internet Research Reports: http://www.securityspace.com/s_survey/data/index.html.

PHP Hypertext Preprocessor: http://www.php.net.

Rigollier C., Bauer O., Wald L. (2000) On the clear sky model of the 4th European Solar Atlas with respect to Heliosat method. *Solar Energy*, 68 (1), 33-48.

Rigollier C., Wald L. (1999) The HelioClim Project: from satellite images to solar radiation maps. In Proceedings of the ISES Solar World Congress 1999, Jerusalem, Israel, July 4-9, 1999, volume I, pp 427-431.

Wald L. (2000) SoDa: a project for the integration and exploitation of networked solar radiation databases. In Proceedings of the European Geophysical Society Meeting, XXV General Assembly, Nice, France, 25-29 April 2000.

World Meteorological Organization, WMO (1981). Meteorological aspects of the utilization of solar radiation as an energy source. Annex: World maps of relative global radiation. Technical Note No. 172, WMO-No. 557, Geneva, Switzerland, 298 pp.

test for a station's daily solar radiation data :

coordinates :
latitude : 33.57
longitude : -7.67
station height : 62.0

1994:10	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V
1994:11	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	**10**	V	V	V	V	V	V	12
1994:12	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V
1998:1	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V
1998:2	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	12	12	12
1998:3	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V
1998:4	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	12
1998:5	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
1998:6	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	12
1998:7	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V
1998:8	V	V	V	V	V	V	V	V	V	V	V	V	**11**	V	V	V	V	V	V	V	**11**	V	V	V	V	V	V	V	V	V	V
1998:9	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	12

number of days passed : 331
number of data input errors : 0
number of processing errors occurred : 0
number of test failed errors occurred : 41
number of days processed : 372

Error number description :
V : verified ok
processing errors :
-3 : error in getnoonelev4day while calculating noonelevation
-2 : error in getGC0day while calculating extraterrestrial value
-1 : error in getGC4day while calculating models value
data input errors :
1 : sort of failed : no value was given for this day
test failed errors :
10 : failed : exceeds calculated extraterrestrial value
11 : failed : exceeds model's Gc value by more than 110 %
12 : failed : too small against extraterrestrial value
21 : sol. elev. <2[deg] & extraterr. >= 2.78, failed : exceeds model's Gc value by more than 200 %
22 : sol. elev. <2[deg] & extraterr. >= 2.78, failed : too small against extraterrestrial value
23 : sol. elev. <2[deg] & extraterr. ><2.78, failed : exceeds 27.78 >
24 : sol. elev. <2[deg] & extraterr. ><2.78, failed : below zero >

Figure 2. An example of the HTML page containing the results (see text for more explanations).