An OGC Web Processing Service for assessing the quality of solar radiation measurements

Bella Espinar, Benoît Gschwind, Lucien Wald, Claire Thomas

An OGC Web Processing Service for assessing the quality of solar radiation measurements

Bella Espinar[1], Benoît Gschwind[1], Lucien Wald[1], Claire Thomas[2]

Abstract

A service is presented that assesses the quality of measurements of daily global irradiation or means of global irradiance acquired by a ground station. Measurements are checked against models resulting in a measure of plausibility. This on-line service obeys the WPS (Web Processing Service) standard of the Open Geospatial Consortium. It is free of use and can be integrated into routine operations and Web portals thanks due its interoperability capability.

1. Introduction

Assessments of the solar radiation reaching the ground are a key element in the domain of electricity production by means of solar-powered systems. These assessments can be derived from a proper processing of satellite images (Blanc et al 2011) or from measurements of solar radiation performed by one or more stations at ground. This communication deals with the latter and more precisely with measurements of the monthly and daily global irradiation or equivalently, means of global irradiance.

Being the basis of decision for large investment or for maintenance of large plants, these data must be accurate and precise. Close examination of data often reveals a lack of quality and, frequently, for extended periods of time. The sources of deficient quality of radiation data are of very different nature. Radiation sensors are typically affected by incorrect sensor levelling, shading caused by near building structures, or complete or partial shade-ring misalignment in the case of measurements of diffuse solar radiation. Other physical agents that affect all type of instruments are dust, snow, dew, water-droplets, bird droppings, electric field in the vicinity of cables, mechanical loading on cables, station shut-down. The maintenance and operation programme may also affect the quality of the measurements: maintenance mishandling, displacement of the instruments during the period under analysis, change in sensors, power shortages, or battery failures (Muneer/Fairooz 2002). A rigorous maintenance program may avoid part of these causes. Nevertheless, the application of quality control is recommended in all stages: acquisition, transmission and processing of data. Even when data originate from a meteorological centre whose quality control practices are acknowledged, data corruption can be introduced in further encoding and transmission (Zarzalejo 2006).

Data quality is a measure of how well data serve the purpose for which they were produced, ensuring that released data is of a known and reasonable quality suitable for scientific research (WMO No.488; Moore et al 2007). The quality of data is deduced from their plausibility, that is considering the consistency of values with respect to physical laws as well as past records. For that, data are submitted to several conditions or checks. Each of these conditions will be named afterwards a quality control procedure (QCP). If the data is flagged by a QCP, it means they do not pass the condition; otherwise, the data are considered as plausible for this specific QCP. The result of the application of all QCPs will be named quality control (QC). Non flagged data after the application of the full QC will be considered as plausible data.

[1] MINES ParisTech, Centre O.I.E., CS 10207, 06904 Sophia Antipolis, France
[2] Transvalor, 694 Avenue du Dr. Maurice Donat, 06255 Mougins, Frances

However, the application of quality control may be a lengthy and tedious work. Several research centres and official institutions which handle data sets of solar radiation have automated these procedures to reduce the time and labour necessary to discover suspicious values in their own data series. Information about the quality results is often available in their web sites. Nevertheless, private institutions and companies have also their own data series whose quality is unknown.

The present work describes a Web service that applies automatically different QCPs to the data uploaded by the user on a standard browser, in a daily time resolution. It has been inspired by earlier experience in web processing service which compared daily solar irradiation observations with expected values based on extraterrestrial irradiation and a simulation of the irradiation for clear and overcast skies (Geiger et al 2002).

2. Quality control procedures

A set of QCPs has been identified from literature. The constraints for selecting them were that they should be applicable for the data user without specific knowledge of any site characteristics which are available only to the site manager, and that they should be robust with respect to the different climates in the world as the QCPs have not been regionally or seasonally optimized.

The QC is guided by basic physical principles and it may also take into account statistical knowledge on the spatial and temporal variability of these variables (ESRA 2000; Geiger et al 2002). Meteorological variables seldom reach these natural limits. In only rare cases it is even possible to pass these limits in very particular and local situation, but only for a short time period.

The currently proposed QCPs in the web service may be classified in the following categories:

- Range checks: the first developed and the most extensively used algorithm is the range check that compares measurement values against previously defined limit values. Range checks can be divided into checks based on extrema and checks based on rare observations (Long/Dutton 2002). Limits for these ranges have been established on the basis of analyses of distributions, examination of climatological statistics and extreme weather events, as well as using general meteorological theory and experience (Vejen et al 2002).

QCP based on extrema: $0.03\ GHItoa < GHI < 1.2\ I_0$

QCP based on rare observations: $0.03\ GHItoa < GHI < GHItoa$

where GHItoa is the global horizontal irradiance on the top of atmosphere (TOA), GHI is the global horizontal irradiance in Earth surface and I_0 is the solar constant corrected by the eccentricity of the Earth orbit.

The main idea of QCP based on rare observations is to estimate the magnitude of the limit value to fulfil the criterion that only a small quota of all data are allowed to exceed these limits. The reason to use limits based on rare observations is that, very often, erroneous data values do not exceed the limits based on extrema even though they are in error. In this way, more measurements can be flagged as suspicious before dissemination. Unfortunately, this does not prevent true values from being flagged. Values outside of limits based on rare observations are not necessarily non plausible; they must be studied carefully by the station manager or the users to decide if they are actually plausible or not, and therefore, acceptable or not for his purpose.

- Consistency checks: QCPs that verify the relationship between two or more time series. Both, two collocated but different variables or the same variable from two different locations may inform about the plausibility of the data. The former one is named internal consistency check; the latter one is named spatial consistency check. Only the internal consistency check is currently proposed in the web service.

In an internal consistency check an observation is compared with at least one co-located measurement of another variable to see if they are physically or climatologically consistent, either instantly or for time series according to adopted observation procedures (Vejen et al 2002). For irradiance time series, when the global horizontal irradiance (GHI) and the diffuse horizontal irradiance (DHI) measurements are available simultaneously, the consistency check proposed by most of the references is:

$$DHI \leq 1.1\ GHI$$

which takes into account an authorized bias because the measurement instrument may have different responses.

3. Design of the service

The design is based on previous experience of similar service and the own experience of the authors in processing routinely radiation measurements acquired by various ground stations.

The service receives a text file containing the data as sole input. There is one line per day and per location in the file. All values are separated by semicolon, also known as CSV format. The lines are not necessarily ordered by date or location. Several locations may appear in the same file.

The file contains

- Date in ISO format, e.g. 2005-01-02T00:00:00. Here the time following letter T is useless and can be set to 00:00:00 or any other value. Time is UT,
- Latitude and longitude obeying ISO standard, i.e. positive latitude for North, and positive longitude for East, e.g. 12.76, -14.52,
- Elevation above mean sea level, in m,
- Summarization period following ISO standard. P01D means: period (P) is 01 day (D),
- Daily means of irradiance, e.g. 222.3,
- Unit: W/m^2, Wh/m^2 and J/m^2 are accepted

The following lines are an example of the input format:

2005-03-13T00:00:00;12.76;-14.52;38;P01D;349.2;W/m2;

2005-03-15T00:00:00;12.76;-14.52;38;P01D;566.2;W/m2;

2005-03-16T00:00:00;12.76;-14.52;38;P01D;1722.6;W/m2;

The output is a text file that has three new columns in addition to the input data. The first additional column gives the value GHItoa, in W/m^2, and the two others are the flags, one for extrema QCP and one for rare observations QCP. 0 means that the data is not flagged, i.e. is a plausible data, and 1 means that the data is flagged, i.e. is not plausible for this QCP. In the example given above, the output is the following:

2005-03-13T00:00:00;12.76;-14.52;38;P01D;349.2;W/m2;422.017959;0;0

2005-03-15T00:00:00;12.76;-14.52;38;P01D;566.2;W/m2;423.891244;0;1

2005-03-16T00:00:00;12.76;-14.52;38;P01D;1722.6;W/m2;424.796441;1;1

where the daily mean of irradiance for 13th March is plausible for both QCPs, data for 15th March is outside the range of rare observation limits and the data for 16th March is outside the range of extreme limits, and, therefore the rare observations limits.

The objective of the present Web service 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 means to check time-series of irradiation

4. Implementing the service as a OGC Web Processing Service

The service is implemented as a Web service, i.e. an application that can be invoked on the Web. The Web Processing Service (WPS) standard of the OGC (Open Geospatial Consortium) has been adopted. OGC WPS offer a standard and interoperable approach to access, combine and process remote and spread resources to obtain value-added information. The QC service takes benefit from the interoperability promoted by the GEOSS (Global Earth Observation System of Systems) (Ménard et al 2012) and is a means to invoke the QCPs by a computer without human interaction (Percivall et al 2011).

This WPS is hosted in the Energy Community Portal (www.webservice-energy.org) of the GEOSS. This portal is an effort carried out by the Center Observation, Impacts, Energy (O.I.E.) of MINES ParisTech / ARMINES towards the Energy and Environmental Community. It allows end-users to access a collection of Web services, data and applications in the field of renewable energy, environment and environmental impact assessment. It is registered as a GEOSS Energy Community Portal in the GEOSS registry since 2009. In addition, MINES ParisTech has deployed an OGC Catalog Service for the Web (CSW) which offers a single Internet access point to users seeking data, datasets, services, maps, images, algorithms… related to energy and environment relevant to all parts of the globe. The WPS is described in this catalogue by the means of INSPIRE-compliant metadata.

This web service is available to everyone on a free basis. It calls upon up-to-date algorithms that have proven performance. They may be changed as progresses are made. This service offers a very convenient means to users to check the plausibility of time series of solar radiation measurements without writing and testing the procedures themselves.

5. Limits of the service

QCP programmes are designed to reveal all gross and substantive non plausible data which recur regularly. The detection of rare and irrelevant measurement events is not worthwhile, as the reliability of tests varies in inverse proportion to the magnitude of errors they are designed to detect (WMO No.488). Since there is always the risk of missing a non plausible data or mistaking a correct value for a doubtful one, it must be searched for detection and false alarm probabilities trade-off. Several articles have pointed out the important role of human understanding after the application of QC. The expert user will be in charge of the decision about the final plausibility of data in order to avoid, for example, exclusion of extreme weather phenomena (Vejen et al 2002).

Finally, a few limitations of automatic QC are analyzed. For example, a drift in the time registered by the data logger remains undetectable with the proposed QCPs. A consequence would be that the automatic QC would yield that data are not plausible, since they will not fit the comparison with data from other databases, such as derived from satellite or from reanalysis.

Users may pay attention when handling the input data to fit the input format requirements since this is a source of potential errors. Any manual intervention of data by copy / paste procedure can introduce errors that were not originally present when coming out from the data logger. According to our own experience in the processing of the measurements, data series may suffer from the following mistakes:

- The user can enter the wrong latitude and longitude of the site: a lack of the sign minus, a confusion of the coordinates of a city due to the existence of a homonymous, an error in the conversion from degrees, minutes, and seconds into decimal degrees…

- Data loggers can use either a particular number or string to designate a missing value. Generally, the missing value is -999. Nevertheless, since causes for missing values are numerous, few data loggers use different missing values depending on the cause. If an unknown string is found into the input file, the automatic process of the data may fail.
- The user may enter the wrong unit of measurements from the wealth of possibilities: for example, for solar irradiation time series, confusing J/m^2 and Wh/m^2.

This list is not exhaustive. An automatic solution to cover the full range of possibilities is very unlikely. The interface has been built to drive the user by supplying recommendations for the prevention of this kind of inadvertent mistakes which may distort the QC result.

6. Conclusion and perspectives

A series of quality control procedures (QCPs) has been described for controlling the plausibility of measurements of daily radiation. A Web service implements these QCPs. It allows an on-line and automatic control of the quality of a series of measurements of daily irradiation or daily means of irradiance. Flags are issued for suspicious data.

OGC standards were used to design the Web service in order to benefit from the interoperability features. The Web service can be easily searched for in OGC-compliant catalogs and invoked by appropriate tools and portals. At the time of writing, the SoDa Service (www.soda-is.com, Gschwind et al. 2006), a portal dedicated to professionals in solar radiation, is offering a simple interface to the WPS. The user is requested to upload an ASCII file (text file) containing the necessary inputs. Help pages will be available that describe the quality control procedures and the file content.

The work presented here is a first step towards the realization of a more comprehensive tool that will handle other integration steps: from 1 min to 1 month. In addition, an automated output QC report is planned to be given. This report will supply some graphs, histograms and percentages of values which have been flagged by each of the QCP. This output report is aimed at helping the user to interpret the results after the quality check.

Acknowledgements

The research leading to those results has partly receive funding from the European Union’s Seventh Framework Programme (FP7/2007-2013) under Grant Agreement no. 262892 (ENDORSE project). The authors thank all ground station operators of the BSRN network for their valuable measurements and the Alfred-Wegener Institute for hosting the BSRN website.

Bibliography

Blanc, P., Gschwind, B., Lefèvre, M., Wald, L. (2011): The HelioClim Project: Surface solar irradiance data for climate applications, in: Remote Sensing, vol. 3, pp. 343-361.

ESRA (2000): European Solar Radiation Atlas, Scharmer, K., Greif, J. (Ed.), Les Presses de l’Ecole des Mines de Paris, Paris, France.

Geiger, M., Diabaté, L., Ménard, L., Wald, L. (2002): A web service controlling the quality of measurements of global solar irradiation, in: Solar Energy, vol. 73, pp. 475-480.

Gschwind, B., Ménard, L., Albuisson, M., Wald, L. (2006): Converting a successful research project into a sustainable service: the case of the SoDa Web service, in: Env. Modelling Software, vol. 21, pp. 1555-1561.

Long, C., Dutton, E. (2002): BSRN Global Network recommended QC tests, V2.0 Baseline Surface Radiation Network (BSRN).

Menard, L. (2011): Energy Scenario, in: Engineering Report, GEO Architecture Implementation Pilot, Phase 3, 2011 [Online]. Available in: www.ogcnetwork.net/pub/ogcnetwork/GEOSS/AIP3/documents/AIP-3-Energy-Scenario-ER-FINAL.pdf. (last access: 2013-04-23).

Ménard, L., Blanc, I., Beloin-Saint-Pierre, D., Gschwind, B., Wald, L., Blanc, P., Ranchin, T., Hischier, R., Gianfranceschi, S., Smolders, S., Gilles, M., Grassin, C. (2012): Benefit of GEOSS interoperability in assessment of environmental impacts illustrated by the case of photovoltaic systems, in: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 5, pp. 1722 - 1728, doi: 10.1109/JSTARS.2012.2196024.

Moore, S., Peppler, R., Kehoe, K., Sonntag, K. (2007): Analysis of historical ARM measurements to detect trends and assess typical behavior, in: 16th Conference on Applied Climatology held in January 2007, San Antonio, Texas, Sponsored by the American Meteorological Society, Boston, Massachusetts.

Muneer, T., Fairooz, F. (2002): Quality control of solar radiation and sunshine measurements - lessons learnt from processing worldwide databases Building Services, in: Engineering Research & Technology, vol. 23, pp. 151-166.

Percivall, G., Menard, L., Chung, L. K., Nativi, S., Pearlman, J. (2011): Geo-processing in cyber-infrastructure: Making the Web an easy to use geospatial computational platform, in: 34th International Symposium for Remote Sensing of the Environment, Sydney, Australia, 2011.

Vejen, F., Jacobsson, C., Fredriksson, U., Moe, M., Andresen, L., Hellsten, E., Rissanen, P., Palsdóttir, T., Arason, T. (2002): Quality control of meteorological observations. Automatic methods used in the Nordic countries. Norwegian Meteorological Institute.

WMO, No.488. (2007): Guide to the global observing system. World Meteorological Organization, Geneva, Switzerland.

Zarzalejo, L.F. (2006): Irradiancia solar global horaria a partir de imagenes de satéllite. Editorial Ciemat, Madrid, Spain.