User's Guide to the MACC-RAD Services on solar energy radiation resources

Bella Espinar, Carsten Hoyer-Klick, Mireille Lefèvre, Marion Schroedter Homscheidt, Lucien Wald

MACC-II Deliverable D_122.2

USER'S GUIDE to the MACC-RAD Services on solar energy radiation resources

Date: 05/2013
Lead Beneficiary: Armines (#4)
Nature: R
Dissemination level: PU

Monitoring atmospheric composition & climate – II

Grant agreement n°283576

Work-package	122 (RAD, Radiation Support and Service Activities)
Deliverable	D_122.2
Title	USER'S GUIDE to the MACC-RAD Services on solar energy radiation resources
Nature	R
Dissemination	PP
Lead Beneficiary	Armines (#4)
Date	05/2013
Status	Final version
Authors	Bella Espinar, Carsten Hoyer-Klick, Mireille Lefèvre, Marion Schroedter-Homscheidt, Lucien Wald (Armines)
Approved by	Marion Schroedter-Homscheidt
Contact	info@gmes-atmosphere.eu

This document has been produced in the context of the MACC-II project (Monitoring Atmospheric Composition and Climate - Interim Implementation). The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7 THEME [SPA.2011.1.5-02]) under grant agreement n° 283576. All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission has no liability in respect of this document, which is merely representing the authors view.

Executive Summary

The European Earth observation programme GMES (Global Monitoring for Environment and Security), now Copernicus (the European Earth Observation Programme) since December 2012, aims at providing environmental information to support policymakers, public authorities and both public and commercial users. A systematic monitoring and forecasting of the state of the Earth's subsystems is currently under development. Six thematic areas are developed: marine, land, atmosphere, emergency, security and climate change. A land monitoring service, a marine monitoring service and an atmosphere monitoring service will contribute directly to the monitoring of climate change and to the assessment of mitigation and adaptation policies. Additional GMES services will address respectively emergency response and security-related aspects.

The pre-operational atmosphere service of GMES is currently provided through the FP7 projects MACC and MACC-II (Monitoring Atmospheric Composition and Climate). MACC combines state-of-the-art atmospheric modelling with Earth observation data to provide information services covering European Air Quality, Global Atmospheric Composition, Climate, and UV and Solar Energy.

Within the radiation subproject (MACC-RAD) existing historical and daily updated databases for monitoring incoming surface solar irradiance are further developed. The service will meet the needs of European and national policy development and the requirements of (commercial) downstream services (e.g. planning, monitoring, efficiency improvements, integration into energy supply grids). The SOLEMI and the HelioClim-3 databases operated by respectively DLR and ARMINES and its subsidiary Transvalor have been specifically developed in several national, European and ESA projects to fulfil the requirements for long-term databases and NRT services.

On its transition process from the precursor services HelioClim and SOLEMI the following User's Guide intends to summarize existing knowledge, which has been published only in a scattered manner.

Part A 'Users' Expectations' describes the communities of users, their expectations and gives an overview of the compliance of the MACC RAD service with those.

In Part B 'The legacy HelioClim-3 and SOLEMI databases', the current databases HelioClim-3 and SOLEMI as well as the methods used to convert satellite images into solar surface irradiance are presented. The quality of the retrieved irradiances is discussed. An overview of the operations and workflow is presented for the creation, updating and monitoring of these databases.

Part C 'The new HelioClim-4 database' describes the new Heliosat-4 method and the new HelioClim-4 database and provides an overview of the operations and the workflow.

Part D 'Quality control of estimates of irradiance' discusses the means to control the quality of the elaboration of the products and to assess the uncertainty of the estimates of irradiance.

Part E 'Delivering products' is devoted to the supply of HelioClim-4 products. The products are defined. A prototype of a means to access the HelioClim-4 products is presented.

It is intended to update this User's Guide regularly following the realisation of the MACC RAD service line.

Version history

Version	Date	Authors	Description of changes
V2.0	2013-05-30	B. Espinar, C. Hoyer-Klick, M. Lefèvre, A. Oumbe, M. Schroedter-Homscheidt, L. Wald	Title is “User’s guide to the MACC-RAD Services on solar energy radiation resources”. Includes the QC/validation procedures of the HelioClim-4 chain Has been restructured to accommodate future changes in the document to present the new method Heliosat-4 and the HelioClim-4 database
V1.0	2010-06-26	C. Hoyer-Klick, M. Lefèvre, A. Oumbe, M. Schroedter-Homscheidt, L. Wald	Initial version. Title: User’s guide to the SoDa and SOLEMI services. Towards the “solar energy radiation resources MACC-RAD Service”

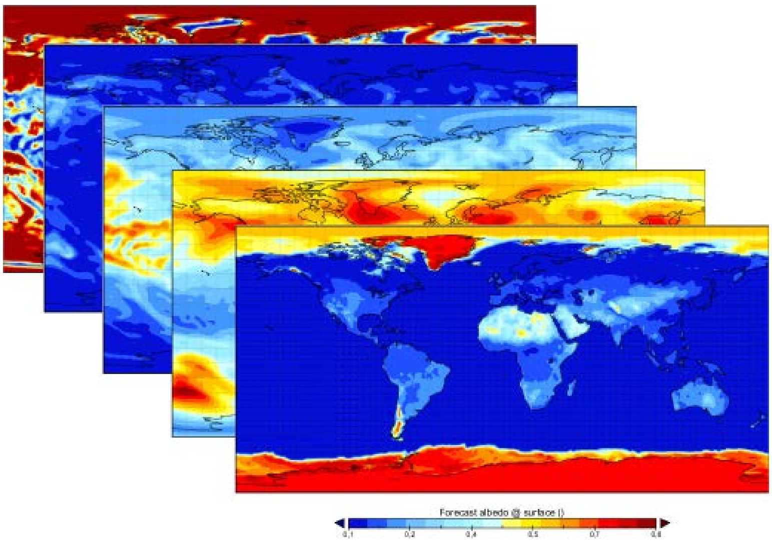
Table of Contents

1. Introduction

The MACC (Monitoring Atmospheric Composition and Climate) project is designed to meet the requirements that have been expressed for the pilot GMES Atmospheric Core Service. The project has been prepared by the consortia of the EC-FP6 project GEMS and the ESA-GSE project PROMOTE, whose core service lines will provide the starting point for MACC. From mid-2009 MACC will continue, improve, extend, integrate and validate these service lines, so that the overall MACC system is ready near the end of 2011 for qualification as the operational GMES Atmospheric Core Service. MACC will prepare the core service in terms of implementation, sustained operation and availability. It will maintain and further develop the efficiency and resilience of the end-to-end pre-operational system, and will refine the scientific basis and quality of the products of the system. It will ensure that its service lines best meet both the requirements of downstream service providers and end users at the European, national and local levels, and the requirements of the global scientific user community.

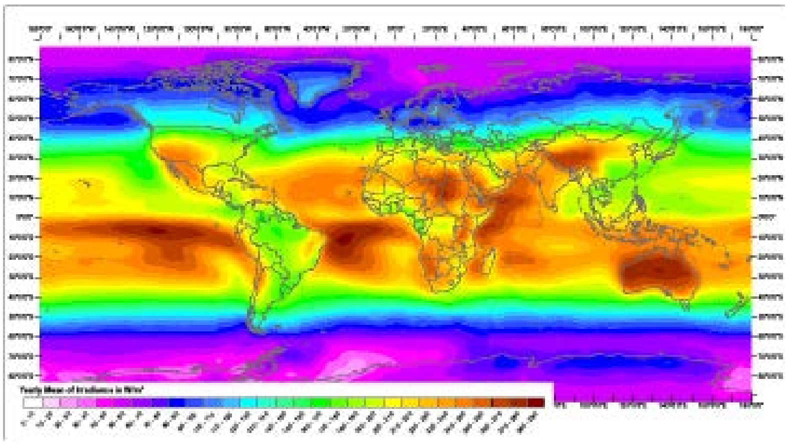
The service lines will cover air quality, climate forcing, stratospheric ozone and solar radiation. MACC will deliver operational products and information that support the establishment and implementation of European policy and wider international programmes. It will acquire and assimilate observational data to provide sustained real-time and retrospective global monitoring of greenhouse gases, aerosols and reactive gases such as troposphere ozone and nitrogen dioxide. It will provide daily global forecasts of atmospheric composition, detailed air-quality forecasts and assessments for Europe, and key information on long range transport of atmospheric pollutants. It will provide comprehensive web-based graphical products and gridded data on which downstream services may be based. Feedback will be given to space agencies and providers of in-situ data on the quality of their data and on future observational requirements.

One of the core services in MACC will provide radiation values at the ground level, which will fulfil the needs in European and national policy developments and the requirements of partly commercial downstream services, e.g., for planning, monitoring, efficiency improvements, and the integration of solar energy systems into energy supply grids.



To do so, several data originating from various sources are assembled. Many of them describe the optical state of the atmosphere, e.g., aerosol optical properties, water vapour and ozone contents over the atmospheric column. Others depict the ground properties, e.g., ground albedo, ground elevation.

These data are inputs to a model that simulates the scattering and absorption phenomena occurring in the atmosphere and affecting the solar radiation in its way downwards to the ground. The outputs of this model are values of the solar radiation available at ground level that can be used to produce energy, either as heat or electricity.



1.1. Objectives of the document

For the sake of simplicity, we denote by "MACC-RAD Service" the future service within the GAS (GMES Atmospheric Service) that will deliver products on solar energy radiation resources.

The objectives of this document are to document and specify this MACC-RAD Service as well as the user-access procedures to the products delivered by this service.

Actually, such a service is an information system composed of two major blocks. One is made of databases, processing chains, daily operations, methods, input data, controlling and monitoring of operations and quality of products. The other deals with the dissemination of the products, the means of access, control of dissemination and quality of service.

This report will document both blocks. As a service aims at satisfying needs of its users, the specifications of this service must be built on the description of the users expectations on products and delivery, and therefore of their profiles. The first part of this document (Part A) is therefore the description of the users and their expectations. Parts B and C deal with the description of the legacy databases and of the new one, of the algorithms, of the operations towards the creation / update of the databases, validation process, monitoring. Part D discusses the means to control the quality of the elaboration of the products and to assess the uncertainty of the estimates of irradiance. Part E deals with the delivery of products.

This document intends to be a living one. It will evolve as the service and its components themselves evolve in the transition process from the precursor services to the MACC-RAD Service.

1.2. Definition of Product Service status

The MACC-II project is preparing the Atmospheric service. It will end in July 2014 and should be replaced by an operational service. In the transition phase, products will experience various states, which have been defined as follows:

- In development: Product is under development and not available in either plot or data format.
- Experimental routine production: Product is routinely produced and available through the Product Catalogue. There is no guarantee on product availability, accuracy is not necessarily monitored, and timeliness is not necessarily in line with user requirements.
- Pre-operational: Product is routinely produced and available through the Product Catalogue. Expected product timeliness, availability, and accuracy are documented

and as much in line with user requirements as possible and are documented in the Service Specification document. Monitoring of input data, timeliness and availability of output data, and quality of output data against independent observations is in place, where possible. Validation of products is available through Validation reports and through monitoring plots available from the Product Catalogue. The production of the product is monitored 24/7 with analyst support during working hours.

- Operational: Same as pre-operational, but with 24/7 analyst support.

1.3. Legacy services

Before proceeding, it is worth mentioning two precursor services –called legacy services- which are the foundations of the MACC-RAD Service: the HelioClim-3 service and SOLEMI service that are presently fully operational and managed by Armines and DLR. There is a variety of documentation, validation and user information available for both services – either from standard operations, from recent project related work or from international benchmarking exercises. Therefore, another objective of this user's guide is to combine and summarize this scattered information.

The HelioClim-3 service is part of the SoDa Service that originates from the EC-funded SoDa project (IST, 200-2003). The SoDa Service turned operational in 2003. Its number of customers increases regularly. From 400 users in 2003, it reached 42 000 users in 2009, and answered approximately 400 000 requests for data in 2009. The SoDa Service is actually a portal and more exactly a collaborative information system. This means that by connecting to the SoDa Service, the user can access several services -including the HelioClim-3 service- offered by several providers in a transparent and homogeneous manner. A service can be a data set, a database or an application. There are presently eleven providers for approximately 60 services. The SoDa Service is formally a means for disseminating knowledge. However, it is becoming so popular that its users often confuse the services offered by the SoDa Service into a single name: SoDa. Behind the SoDa Service, and among the various services offered, a major service is the access to the HelioClim databases of solar radiation at ground level. These HelioClim-1 and -3 databases have been created by Armines and are managed by Armines. A routine exploitation of Meteosat images received by Armines provides 15-min values of radiation that are stored in the databases. Control of operations is performed automatically. Quality of products is checked regularly. In 2006, access to the most recent parts of the HelioClim database became for-pay. There is several tens customers for-pay of the SoDa Service. They are companies having themselves their own customers and own requirements and contingencies. The number of customers is increasing as well as their exigencies regarding the quality of service. Because Armines is a research institute, it does not have the necessary means and knowledge to handle a commercial service and to ensure and secure operations. In 2009, the company Transvalor, a daughter company of Armines, took over the control of the SoDa Service in order to enhance the safety in provision of data and to better handle customers. Presently, the operational processing chain converting Meteosat images into radiation values is implemented at premises of Transvalor to increase reliability in the timely provision of radiation data to customers, while the chain at Armines is now efficiently working as a back-up.

The SOLEMI service has various properties that differ from the HelioClim-3 service. Nevertheless, it is similar in principle to HelioClim-3 in the sense that it is a service: there is one series of actions to construct the database SOLEMI and a series of actions to exploit it for the benefit of customers. The properties of the database SOLEMI are described in the following pages. A series of efforts is underway to ease the access to the database, another series deals with the increase in quality of the irradiation values. The customers of the SOLEMI service request more studies while HelioClim-3 customers require more data.

Therefore, there is no need for the SOLEMI service to offer access via the web like the SoDa Service. However, a service delivering part of the SOLEMI database has been developed by DLR and exposed in the prototype of the second generation of the SoDa Service (project.mesor.net).

One may see the HelioClim-3 and SOLEMI services as concurrent services. This is true in many aspects. However, Armines and DLR share the same objective: performing research to open-up opportunities for companies. Armines and DLR are not competing on market grounds and have signed a legal agreement in 2008, binding Armines, DLR and Transvalor for the sales of next generation radiation products. This agreement is one of the essential pillars of the proposed MACC-RAD Service.

Regarding product service status, the HelioClim-3 service is operated as a 'pre-operational' service. Products are routinely produced and available via the SoDa Service. The SOLEMI service is operated in the 'experimental routine production mode'. Products are routinely produced and available via direct contact to SOLEMI project scientists.

1.4. Acronyms and definitions

DirHI	Direct Irradiance, or Irradiation. Part of the radiation that is received from the direction of the sun on a horizontal plane.
DifHI	Diffuse Irradiance, or Irradiation. Part of the radiation that is received on a horizontal plane from all directions except that of the sun.
DNI	Direct Normal Irradiance, or Irradiation. Part of the radiation that is received from the direction of the sun by a plane facing the sun.
GAS	GMES Atmospheric Service
GHI	Global Horizontal Irradiance, or Irradiation. The radiation that is received by a horizontal plane from all directions.
HelioClim-1	A service and a database containing daily solar irradiation available at ground level, created by Armines from Meteosat images in reduced spatial and temporal resolution. It covers the period 1985-2005.
HelioClim-3	A service and a database containing 15 min solar irradiation available at ground level, created by Armines from Meteosat images. It starts in February 2004 and is updated daily.
HelioClim-4	A service and a database containing 15 min solar irradiation available at ground level, created by Armines within the MACC / MACC-II projects. HelioClim-4 should start in October 2013 and should be updated regularly.
Heliosat	Name of a family of methods to convert images acquired by meteorological geostationary satellites into images of solar radiation available at ground level. For example, the databases HelioClim-1 and -3 are constructed with the method Heliosat-2.
IEA	International Energy Agency
MACC / MACC-II	Monitoring Atmosphere Composition and Climate. Two EC-funded research projects to establish the GMES Atmospheric Service.
MACC-RAD	A sub-project within MACC dealing with radiation.
MACC-RAD Service	The service within the GMES Atmospheric Service that delivers products on solar radiation at ground level.
McClear	A service and a database containing 15 min solar irradiation available at ground

	level, created by Armines within the MACC / MACC-II projects. It starts on 2004-01-01 and is updated regularly.
SoDa Service	A Web portal offering a one-stop access to several services (databases, applications) relating to solar radiation.
SOLEMI	Solar Energy Mining. It corresponds to both method and service providing SSI.
SSI	Surface Solar Irradiance, also called surface downward solar irradiance, or surface downward shortwave irradiance. It can also denotes irradiation, which is the irradiance multiplied by a duration. For example, hourly irradiation is equal to the hourly average of irradiance multiplied by 3600 s.

PART A.

USERS’ EXPECTATIONS

2. Communities of users

Who are the expected users of the MACC-RAD Service? To best design the information system, we need to know who are the users, their typology, how do they structure in communities and domains, and for which purposes they are using the SSI data.

Relationships between users and development of information systems providing SSI are long-standing ones. There are well-established communities of users such as

- planners and managers of solar energy systems and power plants,
- architects and building engineers,
- researchers in renewable energies and building engineering.

There are more and the situation is evolving. New communities of users are attracted by the availability of irradiance data, because they are new in renewable energies such as local authorities and municipalities, or they are in a new domain of activity, e.g., material weathering and ageing.

The managers of the existing information systems delivering SSI such as SOLEMI and SoDa Service and other providers are regularly performing surveys in order to better know their users and establish typologies to improve answers to needs.

Regarding the use of SSI for energy production and building engineering, several recent surveys and outcomes of international initiatives are of interest:

- MESoR (Management and exploitation of solar resource knowledge), funded by the EC DG-TREN, 2007-2009 (www.mesor.org) (Hoyer-Klick et al. 2008, 2009),
- EnviSolar (Environmental information services for solar energy industries), funded by the European Space Agency, 2004-2006 (www.envisolar.com) (Schroedter-Homscheidt et al. 2006),
- GEOSS Energy Community of Practice (www.geoss-ecp.org),
- the Energy Chapter in the 2007 GEO report *The Full Picture* (Schroedter-Homscheidt et al. 2007),
- SWERA (Solar and wind energy resource assessment), a programme of UNEP, 2004-2008 (swera.unep.net),
- IEA SHC 36. International Energy Agency, Solar Heating and Cooling Implementing Agreement, Task 36, 2005-2010 (www.iea-shc.org/task36) (Huld et al. 2008).

A first typology of users can be found from these surveys:

- companies (engineering bureaus, energy producers, investors, plant managers, maintenance services and electricity grid managers),
- experts (engineering bureaus, private R&D),
- public research institutes (engineering bureaus, private R&D),
- public authorities and other organisations supporting policy making, incentives and permit delivery at national, regional or local levels, as well as European policy makers in charge of supporting the implementation of EU policies.

The present work aims at better defining this typology of users, what are their domains of application (or discipline or field), and what are their purposes in using data. It is based on the exploitation of several recent surveys, described in Table 2.1, and of the knowledge gained in the exploitation for several years of services supplying SSI to users or in the above-mentioned projects.

Name	Date	Number of respondents	Comments
IEA SHC 36	2007	111	Respondents were not selected. The questionnaire was on-line on several web sites
SoDa 2005	2005	159	Respondents were not selected. The questionnaire was on-line on the Web SoDa Service

Table 2.1. List of surveys used in the present work

Figure 2.1 exhibits the typology of users and their relative importance as reported by SHC 36 and the SoDa Service. Questions were not exactly the same in both surveys. It is possible that several users are misplaced between "companies" and "engineering bureaus". Nevertheless, both surveys clearly show the importance of the private sector: approximately two-thirds of the users belong to this sector.

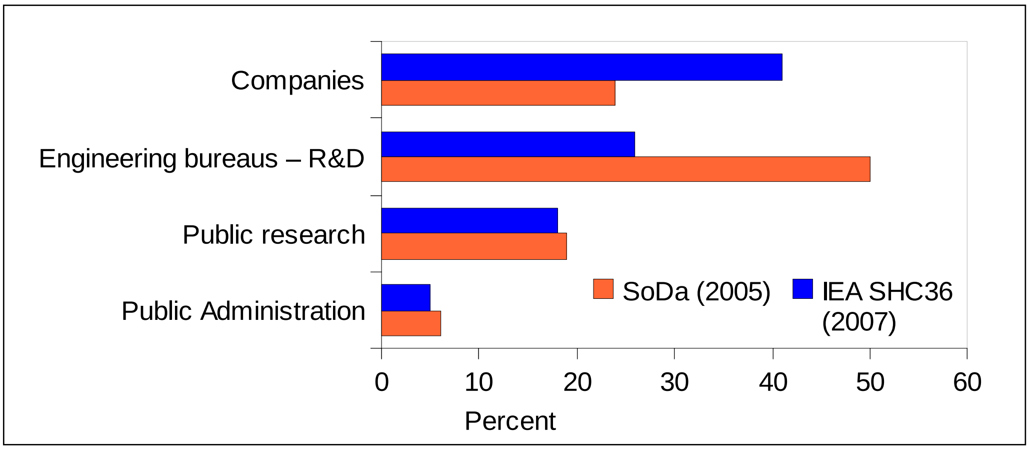


Figure 2.1. Typology of users and their relative importance

There are several purposes in the use of data. Their list and their relative importance are presented in Figure 2.2.

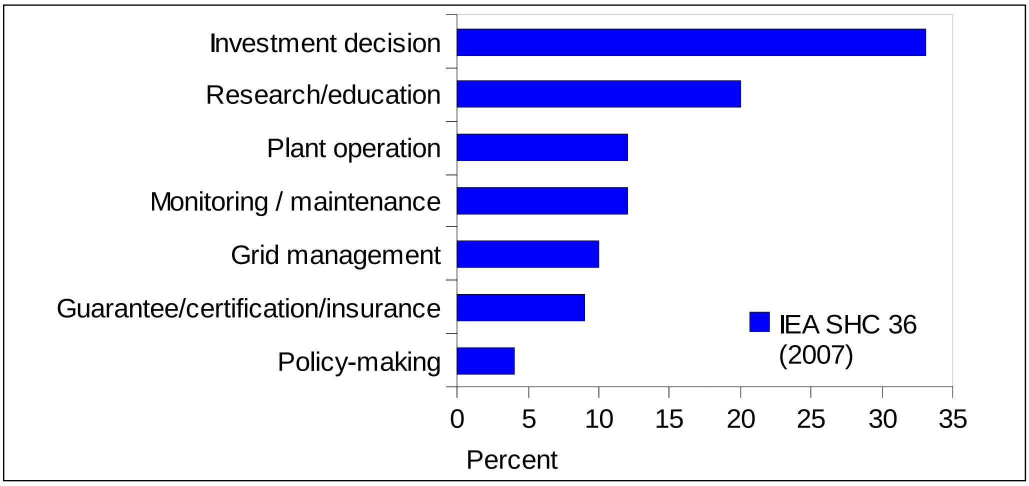


Figure 2.2. List of purposes for use of SSI data and their relative importance

The most important purposes are related to the investment decision: system design, feasibility study, cost assessment, and site selection. SSI data are often used for research and education as already shown in Figure 2.1. Plant operation and monitoring are getting more and more importance as the number of plants is increasing. Smaller solar energy systems need low cost monitoring with a performance check while large solar energy systems need detailed monitoring with automatic fault detection routines. Certification and guarantee are attractive features to customers of the energy suppliers and installers of power systems. Electric power transmission systems collect power from the conventional plants as well as from different renewable sources like solar PV plants. Two major aspects for the management of such a complex grid system to ensure the quality of the supply in electricity is an accurate forecast of the solar power generated and of the expected demand, and a good knowledge of past events and their probabilities.

Finally, policy-making has a limited importance in the request for SSI data. An increase is expected as the local authorities and municipalities are more and more orientated towards sustainability for which exploitation of renewable energy is an important issue.

The survey made by the SoDa Service provides an insight of the various domains of activities of its users and their relative importance in the total number of requests for SSI (Figure 2.3). Several answers were possible for each respondent.

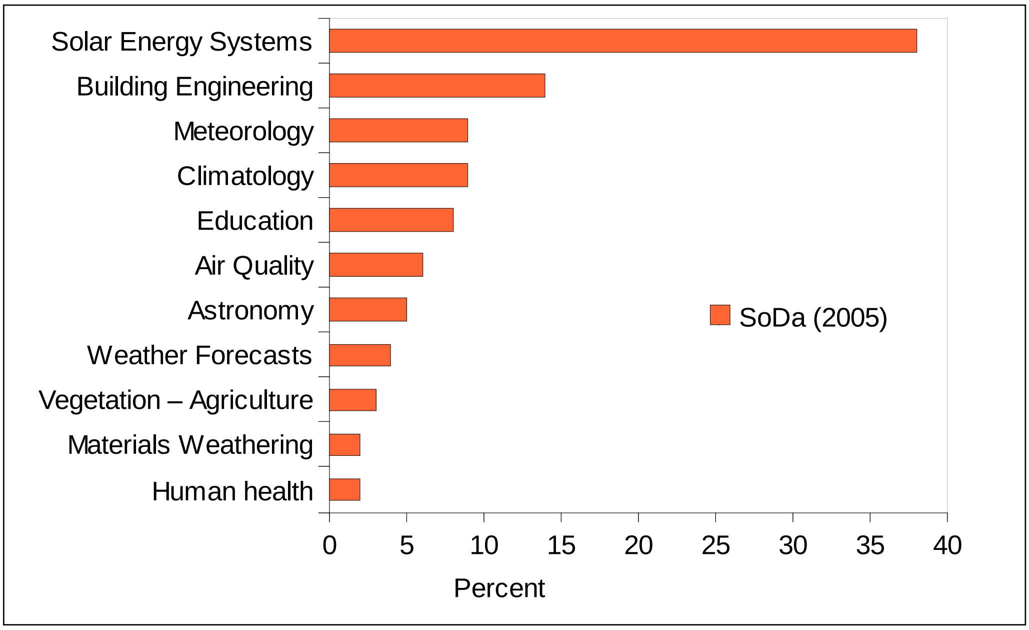


Figure 2.3. Domains of activities of the users of the SoDa Service and their relative importance

Users exploit the SoDa Service in relation with activities in production of energy (38%) and building engineering (15%). These domains represent more than half of the answers (53%). For companies, figures are much larger: production of energy (46%) and building engineering (20%). For companies, these domains represent two-thirds of the answers (66%).

The other half of answers spread themselves in meteorology and climatology (10% each), air quality (urban pollution, radiative budget), astronomy (position of the sun relative to the observer), weather forecasts (SSI, cloud cover), vegetation and agriculture (SSI, PAR), materials weathering and ageing, and human health (skin cancer, eye diseases, multisclerosis). Education is an important activity for the users of the SoDa Service (10%).

One concludes that the SoDa Service is well used by the community in production of energy and building engineering, and also by other communities. It is believed that its usages and the typology of users are representative of what could happen within the MACC-RAD Service.

3. Users' expectations

Users' expectations have been assessed by several surveys and projects with respect to the parameters to be delivered and temporal and spatial resolution. A few dealt with the concept of service as a whole such as the project ENVISOLAR (Schroedter-Homscheidt et al. 2006) and those funded by the IEA (Cros, Wald 2003; Cros et al. 2004; Wald 2006).

We exploit this mass of information together with two recent relevant surveys: IEA SHC 36 and MESoR (Table 3.1). The IEA SHC 36 survey has been presented in the previous Chapter. The MESoR survey differs in several aspects. The sample of 30 respondents was selected by the MESoR consortium according to criteria of importance of data and frequency of usage as well as attitude to scientific cooperation. The surveyed organisations are active in the fields of solar energy systems and building engineering. They are located in various European countries plus the USA. The interviews by phone aimed at collecting information and evaluation about requirements, including needs for data (parameters, format, quality, resolution, etc.), service provision (coverage, availability), as well as about the attitude and interest towards the integration of the existing services into a unique point of access.

In addition, as it has been found that the usages and users of the SoDa Service are representative of what could happened with the MACC-RAD Service, we have exploited several findings on the usage of the SoDa Service in years 2008 and 2009 (Table 3.1).

Name	*Date*	*Number of questionnaires*	*Comments*
IEA SHC 36	2007	111	Users were not selected. The questionnaire was on-line on several web sites
MESoR	2008	30	Users were selected and interviewed on the phone
SoDa	2008 or 2009	Based on several thousands of requests	It is not a survey filled in by users but a survey of the use of data provided by the SoDa Service during periods spanning one or more years

Table 3.1. List of surveys and findings used in the present work

The products of the MACC-RAD Service will be exploited by commercial users to produce their own services dedicated to their own customers (downstream services). We rely on the experience gained by the present SOLEMI and SoDa Services to obtain the business-related users' expectations.

Users expectations are treated in several categories:

- *data*: which data are expected, which geophysical parameters,
- *metadata and ancillary information*: which other data are useful to exploit the main data,
- *access to data*: how can users practically access data, and which format,
- *documentation*: which documentation is expected,
- *quality of service*: operations, help desk, customers desk, data policy,
- *monitoring quality*: scientific validation of products, quality in delivery.

3.1. Data

3.1.1. Geophysical parameters

Figure 3.1 lists the geophysical parameters requested by the users and their relative importance according to the surveys IEA SHC 36 and MESoR.

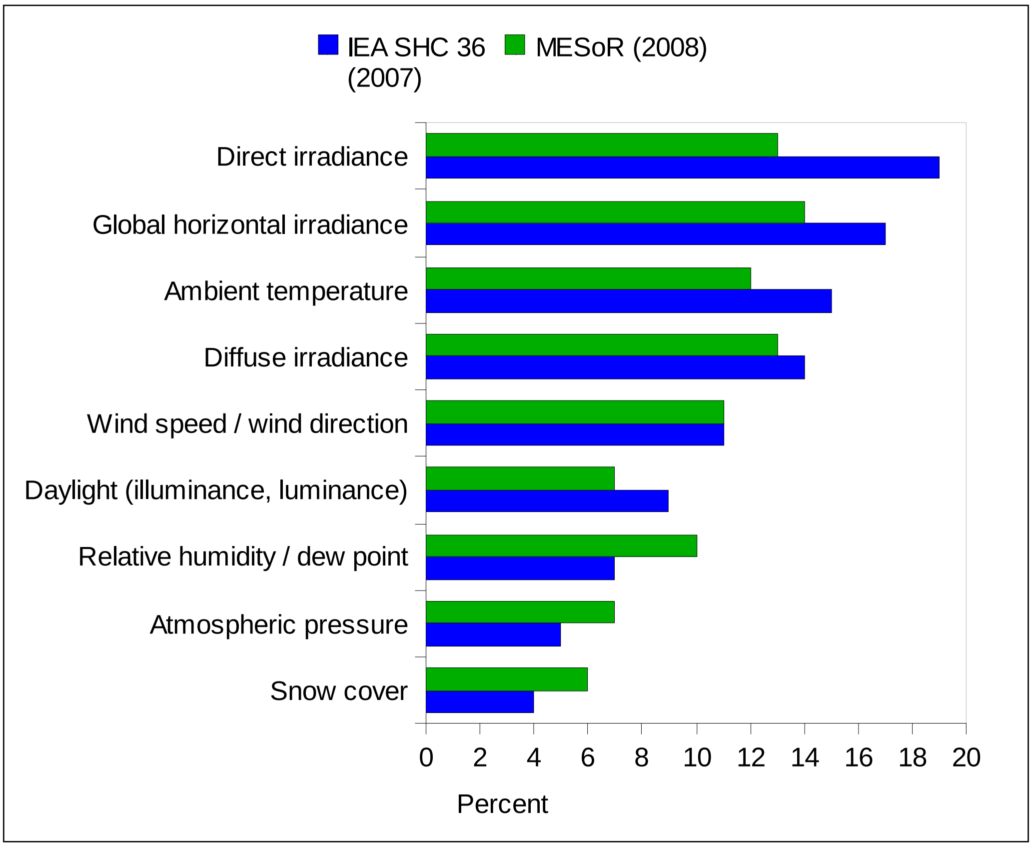


Figure 3.1. Geophysical parameters requested and their relative importance

Results differ slightly by a few percents from one survey to another. This may be explained by the diversity in panels and by the content of the questionnaires. Nevertheless, both agree on the importance of global, direct and diffuse irradiances (GHI, DirHI, DifHI).

The demand for ambient temperature, or air temperature at 2 m height, is high. Actually, there is a request for having auxiliary meteorological data at the surface: air temperature, wind speed and direction, relative humidity and atmospheric pressure, which are coincident with irradiance data. These data are useful for computing the radiative budget of the solar energy system, thus having a more accurate assessment of the gains and therefore of the energy production. This is also true in building engineering to compute gains and losses in heat of buildings.

Unsurprisingly, users in building engineering request data in illuminance as this is the parameter used in daylighting in building.

A few users ask for snow cover. Snow may cover panels, thus reducing the energy production. It may also be confused with clouds by methods assessing the SSI from satellite images; its knowledge increases the reliability of the SSI data.

3.1.2. Spatial and temporal resolutions

According to these surveys, the spatial resolution of the data ranges clearly from 5 to 10 km, though there is marked interest for higher resolution of 1 km.

Figure 3.2 exhibits the list of temporal resolutions requested by users and their relative importance. Numbers differ slightly between both surveys but are in agreement. Both show that monthly values of SSI attract the greatest attention, followed by hourly values. The demand for daily values is the lowest.

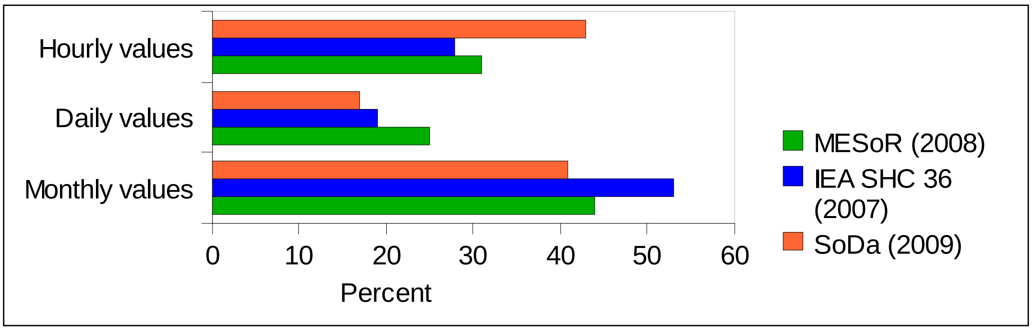


Figure 3.2. Temporal resolutions requested and their relative importance

We add on this graph findings from the usage of the SoDa Service. During 2009, SSI data for the year 2005 were for free independently of the temporal resolutions. Computing the number of requests made for each resolution relative to the total number of requests provides a clear view of the most requested temporal resolutions.

The demand for daily values is small, in agreement with the surveys. For hourly and monthly values, the results are somehow conflicting. There were much more requests to hourly values than expected and slightly less for monthly values. The discrepancy between the expressed expectations and the effective behaviour of the users may be explained by the time lag between the surveys and the findings from the SoDa Service. At the time of the surveys (2007 and 2008), hourly data were not easily accessible. We believe that making them freely available in an easy manner creates a change in the usages of data.

It could be reported here that there has been an explosion of the number of requests for hourly values made to the SoDa Service in 2010 due to the increasing use of such SSI data for monitoring solar power plants.

3.1.3. Time-series or maps

Most data sources deliver data in the form of time-series. Users expressed differently in surveys when asked about their preferences for time-series of data (1 or more sites) or maps (i.e. gridded values). According to IEA SHC 36, both types seem useful to respondents. In the MESoR survey, users prefer clearly time-series to maps.

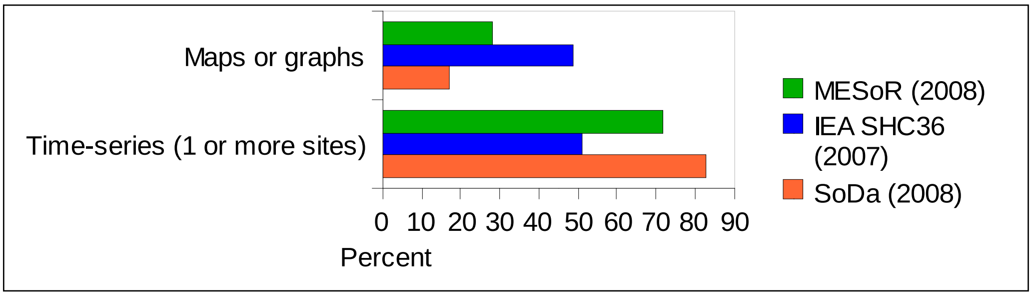


Figure 3.3. Percentage of users preferring time-series or maps

Findings from the usage of the SoDa Service were added on this graph (Figure 3.3). During 2008, a mapping capability was made available for monthly values. Computing the number of requests made for maps relative to the total number of requests provides a view of the relative importance of maps. There are much more requests for time-series than for maps. Clearly, a provider of data should put priority on the capability of delivering time-series.

3.2. Metadata and ancillary information

Several pieces of additional information should be delivered with SSI data, besides the meteorological data mentioned above. These are metadata and ancillary information.

Metadata provide description of the data, e.g., date and time, geographical coverage, etc. They are necessary to exploit the data. There are two types of metadata.

One type is for discovery. It describes the data that could be delivered by a service, the list of products and how to obtain them. This type is exploited by catalogues of services and data. Examples are geographical or temporal coverage, spatial or temporal resolution, parameters.

The second type is for exploitation of data by computers or not. It describes the exact content of the supplied file containing the requested data. Examples are geographical location, elevation, parameters, instant.

It is recommended to use standard metadata such as those proposed by INSPIRE, GEOSS, and ISO.

Figure 3.4 demonstrates the importance of the metadata for exploitation. It displays the percentage of users using the supplied SSI data as inputs to a simulator and more generally to an application, or using SSI data directly in their work. This percentage is large. It means that provided the metadata supplied and those understood by the application share the same format, delivering exploitation metadata should ease the workload of the users.

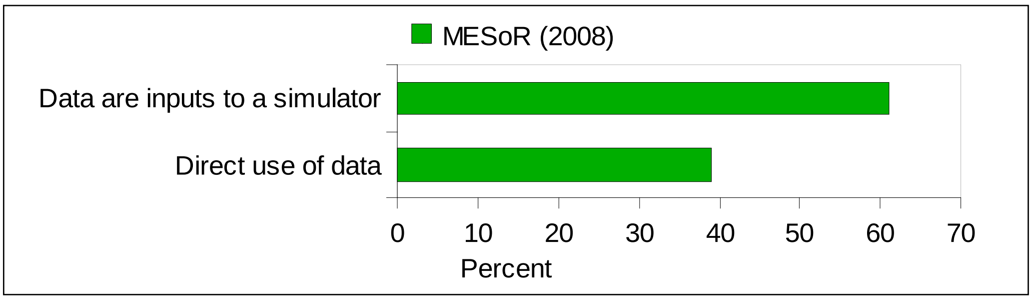


Figure 3.4. Percentage of users using the SSI data as inputs to an application

Users gather many different aspects in the term "ancillary information", such as:

- request for having as outputs in addition to SSI the data input to the method for SSI assessment, such as aerosol properties, water and ozone column content, etc.,
- say whether data are interpolated or not in space or time,
- reliability, i.e., one or more quantities stating whether the supplied data, as a whole or individually, are reliable,
- uncertainty, i.e., one or more quantities about the accuracy of the supplied data, as a whole or individually.

3.3. Access to data

Access to data covers several aspects: where, when, how. As a whole, users desire to access to data for anywhere and anytime; they desire to have an easy and fast access to these data whenever they want, from anywhere.

The geographical coverage requested by users depends strongly of their area of interest. For European users, emphasis is on Europe and Mediterranean basin. Ideally, the whole world should be covered.
The temporal coverage should span over several years, typically 10 or 15 recent years. Figure 3.5 displays the answers of respondents about the age of the SSI data.

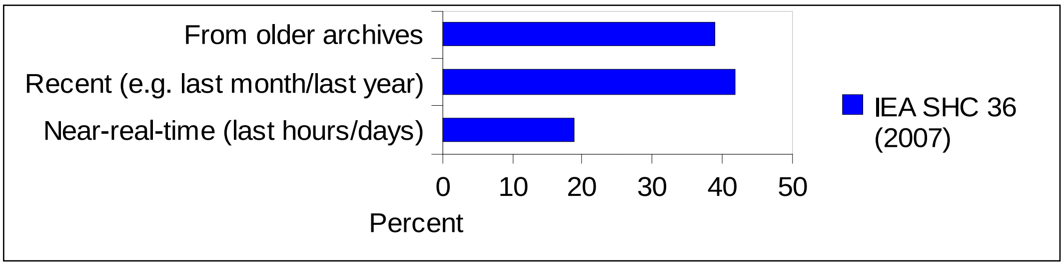


Figure 3.5. How recent should be the data

There is an equal spread of answers between archives and recent data (last month). Requests for very recent data are weak. The latter may have evolved in recent years due to the increase in monitoring activities of solar power plants, as discussed in section 3.1.

In the MESoR survey, all respondents underlined the importance of the Web as a mean to deliver information. They believed that accessing to the data by the Web makes things easier and faster, and because it would not be necessary to store information and applications in own information systems, thus reducing the burden in ICT expertise and hardware.

The Web has several advantages for users:

- it could be a one-stop shop to access various data,
- it can be accessed anytime from anywhere,
- a standard browser (e.g., Internet Explorer, or Firefox) can access the Web,
- the delivery of data can be instantaneous or so.

Finally, there is a large consensus in surveys about format of data. Users do not request for a specific format but do request for a standard type of file, namely text files (e.g., CSV files), or Excel-like spreadsheets. This is in line with the graph (Figure 3.4) showing that products are further ingested into applications. As discussed earlier, these files should contain metadata describing the content of a file.

3.4. Documentation

Surveys reveal a great demand on documentation. Actually, there are several types of documentations requested:

- scientific documentation,
- service documentation,
- list of products and catalogue,
- data policy.

The scientific documentation comprises at best, the articles that support the methods and operations used to supply the products. These articles have been published in international peer-reviewed journals and are freely available on-line.

A report may be available that describes how the contents of these articles have been exploited to produce the SSI and other data. The respondents stressed the need for clarity and transparency about the procedures and methods.

The service documentation comprises the users' guide, describes operations, the workflow, and the various monitoring elements for quality assurance. Several procedures are described such as those used to fill gaps (spatial or temporal), or to flag and possibly remove erroneous data in temporally averaged products. The assessment of the reliability and the uncertainties of the data should be clearly depicted.

The list of products should be available, clear and comprehensive. A catalogue could be established with search capabilities.

Finally, data policy and intellectual property rights should be available and easy to read.

3.5. Quality of service

Users expressed concerns about the quality of service as a whole. By this, they expressed their expectations for a reliable service, working well, with no problem at all, and supplying what is expected.

Several users believe that such a service will work only if it can deliver products to users building their own services on the core products (downstream services). Accordingly, this service should work as a service to customers with customers' desk and helpdesk. The customers' desk deals with the commercial aspects of requests if any. The helpdesk helps in the use of the service.

The service should clearly express the means of delivery of the products and the time schedule. It should supply products in due course and punctually. If the Web is used to deliver, whether through a browser or e-mail, the delivery should be made in the moments following the order.

The SoDa Service offers capability to access data in an automated way by computers at a fixed hour every day. This capability should be offered in the MACC-RAD Service and enhanced. As several commercial applications are based on such capability, especially for monitoring solar power plants, supply of products should be safe. The Service should clearly indicate how it handles the various possible breakdowns and the ways to mitigate their consequences.

3.6. Monitoring quality

Users expressed concerns about the means for controlling the quality of the products. More precisely, questions arise about:

- the scientific quality. How it is established? What is the scientific validation of products? How often such a validation is made? How is it monitored?
- the quality of the workflow, i.e., monitor the smooth running of each operation and of the chaining. How is it made and how often? What are the plans to improve the method / workflow if scientific improvements are available?
- the quality of the service. How is it monitored and how often? How does it include surveys of incidents, concerns of users?

4. Answering the expectations

The MACC-RAD Service is designed in such a way that it satisfies the users' expectations as a whole. Each element is designed in view of its integration in a more complete system and not individually. The use of Web technologies and more generally, of GEOSS standards and best practices in exchange of data and interoperability of applications permits to design an efficient information system in a modular way. The concept of a collaborative information system on the Web adopted in the MACC-RAD Service is discussed at the end of this document. Best designed for its integration in the MACC-RAD Service, each element can be updated and improved without re-designing the whole information system.

The concept of the MACC-RAD Service is based on the following elements:

- operations and procedures to create databases of SSI, monitor the workflow, and control the quality,
- databases of SSI and other parameters. These databases contain basic information from which products are created,
- applications that exploit these databases in order to create MACC-RAD core products. These applications are here Web services, i.e., applications that can be invoked on the Web,
- exposition of these Web services in a community repository (community portal) for their exploitation,
- a dedicated information system that interacts with the users, presents products, exploit the Web services to obtain products that it delivers to the users.

These elements are discussed in the next Chapters.

Prior to detailing these elements, the compliance of the proposed design to users' expectations is discussed. To that goal, the expectations are listed to show whether the future MACC-RAD Service bring an appropriate answer or not to each of them. In order to show improvements with current situation, the same is done for the present services: HelioClim-3 and SOLEMI. In addition, the same comparison is done for the planned transition system that is based on the legacy services. This transition system will be set up in order to bring answers more rapidly, as the creation of the MACC-RAD Service is a lengthy task.

4.1. Data

Table 4.1 shows the compliance of the present services, the transition service, and the first and second versions of the MACC-RAD Service with respect to users' expectations on geophysical parameters.

The present services HelioClim-3 and SOLEMI deliver values for the global irradiance and its direct and diffuse components. These parameters will be delivered in the forthcoming services. Illuminance will be derived from total irradiance in the transition service, and will be directly computed in the MACC-RAD Service.

Request	Irradiance, illuminance	Snow cover	Other meteorological parameters
Present services	Yes - No	No	No
Transition Service	Yes - Yes	Yes, partly	No
MACC-RAD v1	Yes - Yes	Yes	No
MACC-RAD v2	Yes - Yes	Yes	No

Table 4.1. Compliance with the users' expectations regarding geophysical parameters

Snow cover is not a parameter delivered by the present services. It will be considered in the transition service, at least as a flag indicating a possible lack of reliability of the SSI estimate. Further versions will both take into account snow cover in the SSI estimates and provide snow cover parameter.

Other meteorological parameters are not delivered by the present services. Neither the transition service, nor the MACC-RAD Service will provide them because of the data policy in force within the meteorological community.

Table 4.2 shows the compliance of the services with respect to users' expectations on the spatial resolution of the geophysical parameters. The expectations will be met when parameters are available, except for the other parameters in version 1 of the MACC-RAD Service. These parameters will have the resolution of the re-analyses produced by ECMWF, i.e. a few tens of km. Smart spatial resampling procedures will be applied in version 2 to meet users' expectations.

Request	Irradiance, illuminance	Snow cover	Other parameters
Present services	Yes	-	-
Transition Service	Yes	Yes	-
MACC-RAD v1	Yes	Yes	No
MACC-RAD v2	Yes	Yes	No

Table 4.2. Compliance with the users' expectations regarding spatial resolution (5 to 10 km)

Table 4.3 shows the compliance of the services with respect to users' expectations on the temporal resolution. The expectations will be met for irradiance and illuminance. It is not expected to have hourly values of snow cover. As for other parameters, daily and monthly values will be available. Smart temporal resampling procedures will be applied in version 2 to meet users' expectations.

Request	Irradiance, illuminance	Snow cover	Other parameters
Present services	Yes – Yes - Yes	-	-
Transition Service	Yes – Yes - Yes	No – Yes - Yes	-
MACC-RAD v1	Yes – Yes - Yes	No – Yes - Yes	No – No - No
MACC-RAD v2	Yes – Yes - Yes	No – Yes - Yes	Yes – No - No

Table 4.3. Compliance with the users' expectations regarding temporal resolution: hour – day - month

Table 4.4 shows the compliance of the services with respect to users' expectations on the provision of time-series and maps. The present services deliver time-series as a standard product and maps on request. The situation will remain the same till the whole inclusion of the MACC-RAD Service version 2 of a mapping capability to provide maps as standard products.

Request	Time-series	Maps
Present services	Yes	On request
Transition Service	Yes	On request
MACC-RAD v1	Yes	On request
MACC-RAD v2	Yes	Yes

Table 4.4. Compliance with the users' expectations regarding provision of time-series and maps

4.2. Metadata and ancillary information

Table 4.5 shows the compliance of the services with respect to users' expectations on the metadata for discovery and exploitation. The present services use metadata for exploitation only. The transition service will be present in GEOSS catalogues due to the use of metadata for discovery. These metadata will be INSPIRE-compliant. Exploitation metadata in the MACC-RAD Service will be based on the on-going work of the GEOSS Architecture and Data Committee.

Request	Discovery	Exploitation
Present services	No	Yes
Transition Service	Yes (INSPIRE)	Yes
MACC-RAD v1	Yes (INSPIRE)	Yes (GEOSS)
MACC-RAD v2	Yes (INSPIRE)	Yes (GEOSS)

Table 4.5. Compliance with the users' expectations regarding metadata for discovery and exploitation

Table 4.6 shows the compliance of the services with respect to users' expectations on the ancillary information. The present services deliver flags and quantities relating to interpolation, reliability and uncertainty. This will remain in the forthcoming services with enhancements. In addition, data input to the method will be delivered by the MACC-RAD Service.

Request	Data input to method	Flag on interpolation	Reliability	Uncertainty
Present services	No	Yes	Yes	Yes
Transition Service	No	Yes	Yes	Yes
MACC-RAD v1	Yes	Yes	Yes	Yes
MACC-RAD v2	Yes	Yes	Yes	Yes

Table 4.6. Compliance with the users' expectations regarding ancillary information

4.3. Access to data

Table 4.7 shows the compliance of the services with respect to users' expectations on the depth of the archive (10 years at least) and the age of the data. The services meet the demand on this point.

Delivering real-time products, i.e. having less than a few hours, will be technically difficult and is not in the scope of the MACC-RAD Service. The method to be used in the transition service and MACC-RAD Service requests inputs that cannot be provided in real-time. The delay in obtaining these inputs leads to a delay in provision of SSI products.

In addition, data policy and downstream services should be taken into account. Disseminating very recent data is a sensitive topic subject to the data policy of the various providers and that could be harmful to markets generated by downstream services and further, to the companies behind.

Request	Archive (at least 10 years)	Archive (10 recent years)	Recent (last month)	Near-real time (last hours/days)
Present services	Yes	Yes	Yes	Yes
Transition Service	Yes	Yes	Yes	Yes
MACC-RAD v1	Yes	Yes	Yes	Yes
MACC-RAD v2	Yes	Yes	Yes	Yes

Table 4.7. Compliance with the users' expectations regarding the depth of the archive and the age of data

Table 4.8 shows the compliance of the services with respect to users' expectations on the use of the Web and format of data. The present services use the Web for ordering and delivering data. Data are supplied in the form of text files (CSV) or Excel-like spreadsheets. The forthcoming services will meet the expectations.

Request	Use the Web	CSV files or spreadsheets
Present services	Partly	Yes
Transition Service	Yes	Yes
MACC-RAD v1	Yes	Yes
MACC-RAD v2	Yes	Yes

Table 4.8. Compliance with the users' expectations regarding use of the Web and format of data

4.4. Documentation

Table 4.9 shows the compliance of the services with respect to users' expectations on documentation. The present services meet the demand for documentation. Many documents are on-line. The forthcoming services will enhance the documentation, which will be on-line.

Request	Scientific	Users guide	Workflow / procedures	Quality	List / catalogue	Data policy
Present services	Yes, on-line	Partly, on-line	Partly, on-line	Partly, on-line	Yes, on-line	Yes
Transition Service	Yes	Yes	Yes	Yes	Yes	Yes
MACC-RAD v1	Yes	Yes	Yes	Yes	Yes	Yes
MACC-RAD v2	Yes	Yes	Yes	Yes	Yes	Yes

Table 4.9. Compliance with the users' expectations regarding documentation

4.5. Quality of service

Table 4.10 shows the compliance of the services with respect to users' expectations on the quality of service as a whole. The present services have taken steps to meet these expectations. The forthcoming services will follow this policy with improvements.

Request	Delivery procedure	Customers desk	Helpdesk	Automated access
Present services	Yes	Partly	Partly	Partly
Transition Service	Yes	Yes	Yes	Yes
MACC-RAD v1	Yes	Yes	Yes	Yes
MACC-RAD v2	Yes	Yes	Yes	Yes

Table 4.10. Compliance with the users' expectations regarding the quality of service as a whole

4.6. Monitoring quality

Table 4.11 shows the compliance of the services with respect to users' expectations on the quality assurance and control.

The present services perform regularly validation of their products using ground measured data of opportunity. The validation procedure is now standardised thanks to the efforts made in the projects IEA SHC 36 and MESoR. Other manual procedures include a visual analysis of annual patterns for several years and a single geographical site, or visual inspection of images of SSI or other data in a movie.

In the future, it is foreseen to follow the same path for the transition service, and then, to exploit the re-analyses from ECMWF. It is foreseen to be able to report daily to customers. There will be two approaches in validation.

The first one consists in comparing reliable ground data with products. It will be performed as often as possible. It aims at assessing the uncertainty of the products. A model of uncertainty can be established or improved that provides for any instant the plausible uncertainty level of the SSI, given known explanatory variables such as the solar zenithal angle or the temporal variability of the SSI.

The second approach consists in comparing re-analyses with products. It will be performed once a day. Discrepancies are expected between both assessments. However, it is expected that a model can be found that explains these discrepancies as a function of known explanatory variables. Therefore, the distance between products and re-analyses can be predicted. It will be used to monitor the quality of products: large values should indicate suspect products. This approach is only sketched here. Its implementation will necessitate a great deal of effort and care.

The present services monitor the workflow at the following points:

- reception of the Meteosat images,
- missing lines in an image, position of extreme lines and columns in image,
- storage of images of radiance in the temporary archive,
- storage of clear-sky indices in the HelioClim databases,
- smooth running of each piece of software.

This will improve in the forthcoming services. More points will be controlled and reporting to management will improve. The future method Heliosat-4 to compute SSI is more sophisticated than the current method Heliosat-2. Monitoring its smooth running will require more control points.

The present services do not have automated procedure to monitor the quality of the service. Efforts are made on an opportunity basis to warn customers about planned shutdowns. Efforts are planned in the forthcoming services. It is expected to establish a few control points, such as monitoring the number of connected customers and the computers' capability, to report automatically to the management, and store reports in an archive. The MACC-RAD Service will comprise more control points and automated warning to customers. Enhanced reporting will be added in the future and daily controls of the interfaces to Web services (i.e. access to products) will be performed.

Request	Scientific validation	Monitoring workflow	Monitoring service
Present services	Using ground data of opportunity, standardised procedure	Two control points, warning	None
Transition Service	Using ground data of opportunity, standardised procedure	More control points	Reporting to management, archive of incidents, early warning to customers
MACC-RAD v1	Systematic control with re-analyses, warning, ground data of opportunity	Monitor each element, warning	Reporting to customers and management, archive, early warning
MACC-RAD v2	Daily control with re-analyses, ground data of opportunity, reporting to customers	Enhanced control	Enhanced reporting, archive, early warning, daily control of interfaces to Web services

Table 4.11. Compliance with the users' expectations regarding the quality assurance and control

PART B.

THE LEGACY HELIOCLIM-3 AND SOLEMI DATABASES

5. The HelioClim-3 and SOLEMI databases

The legacy HelioClim-3 (Blanc et al. 2011) and SOLEMI (Schillings 2004; Schillings et al. 2004) databases originate from a processing of Meteosat images provided by Eumetsat, a European agency located in Darmstadt, in Germany. The Meteosat satellites are geostationary and located over the Gulf of Guinea, at longitude 0 approximately. The currently operational satellite is called Meteosat Prime. As it is aging, the Meteosat Prime is replaced regularly and since 1998, the "old" satellite is shifted eastwards over the Indian Ocean, at 63°E; it is then called Meteosat East. Figure 5.1 depicts the geographical areas covered by both satellites and by SOLEMI.

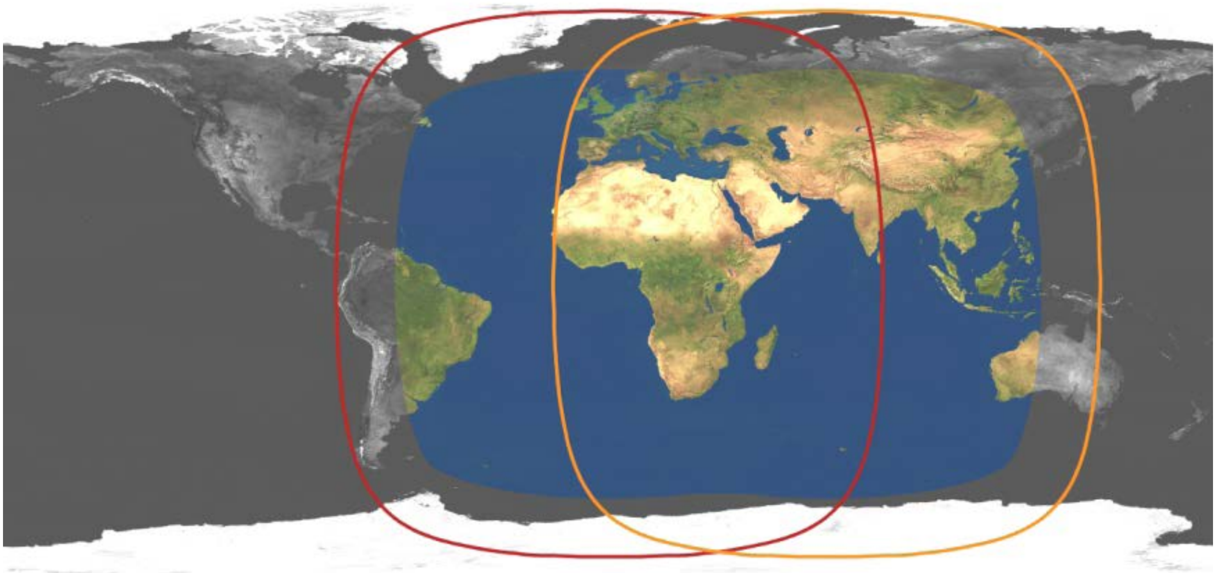


Figure 5.1: Field of view of the Meteosat Prime (in red) and Meteosat East (in orange). The blue area denotes the part of the area where irradiances can be computed; this is the coverage of the database SOLEMI.

Since 2004, a new series of Meteosat satellites is in operation: Meteosat Second Generation covers the same area than the prime satellites. The Meteosat First Generation satellites were operated concurrently till 2006.

DLR has a complete copy of the Eumetsat archives of images acquired by the satellites Meteosat First Generation back to 1984. Raw data from 1991 to June 2006 (Meteosat Prime) and October 1998 to December 2006 (Meteosat East) are easily and operationally accessible in the Data Information and Management System (DIMS) at DLR. These images are exploited to produce the database SOLEMI.

Since 2004, the images of the satellite Meteosat Prime (MSG) are routinely received at MINES ParisTech / Armines and are processed in near-real-time, owing to collaboration with Eumetsat and Meteo-France. These images cover Europe, Africa, the Atlantic Ocean and the Western part of the Indian Ocean (see left half of the blue area in Figure 5.1). These images are exploited to produce the database HelioClim-3.

The main properties of the databases HelioClim-3 and SOLEMI are presented in Table 5.1.

	HelioClim-3	***SOLEMI***
Period	Since 2004	Since 1991
Temporal Resolution	15 min	1 hour
Geographical Coverage	Europe, Africa, Middle East	Europe, Africa, Middle East, Asia (except Eastern Asia)
Spatial Resolution	3 km at satellite nadir, approx. 5 km mid-latitude	3 km at satellite nadir, approx. 5 km mid-latitude
Parameters	Global total irradiance on horizontal plane	Global total irradiance on horizontal plane; direct normal irradiance
Processing	Near-real-time	On request
Update of the database	End of day	End of month
Version	V3	V14, build 98

Table 5.1. Main properties of the databases HelioClim-3 and SOLEMI

6. Brief description of the method converting satellite images into surface solar irradiance

6.1. History of the Heliosat methods

Several studies have demonstrated the feasibility of extracting the global solar surface irradiance (SSI) from geostationary satellites images like Meteosat (Tarpley, 1979; Möser, Raschke, 1984). Indeed, these satellites, which have passive sensors, observe the state of the atmosphere and the cloud cover above the target. These observations can be used to calculate the radiation reaching the ground.

Very early, the European Commission funded research to develop methods for retrieving the SSI from Meteosat images (Grüter et al., 1986). Among those, the Heliosat method was developed at MINES ParisTech (Cano et al., 1986). It became very popular and has been adopted by many researchers. Therefore, it underwent many changes aiming at improvements; the versions bearing major improvements were numbered (Heliosat-1, Heliosat-2, Heliosat-3).

The principles of the Heliosat method are illustrated in Figure 6.1. In most cases, a cloud exhibits a larger reflectance than the ground. Consequently, the appearance of a cloud in the field of view of the satellite sensor should result in an increase of the perceived signal: the cloud (target 2) appear brighter (whiter) than the ground (target 1). The magnitude of the difference between both targets is related to the depletion of the downwards radiation by the atmosphere. Of course, the situation where one can compares a cloudy pixel to a neighbour cloud-free pixel rarely happens. Therefore, Heliosat comprises a modelling of the SSI that should be observed by the sensor if the sky were clear for any pixel.

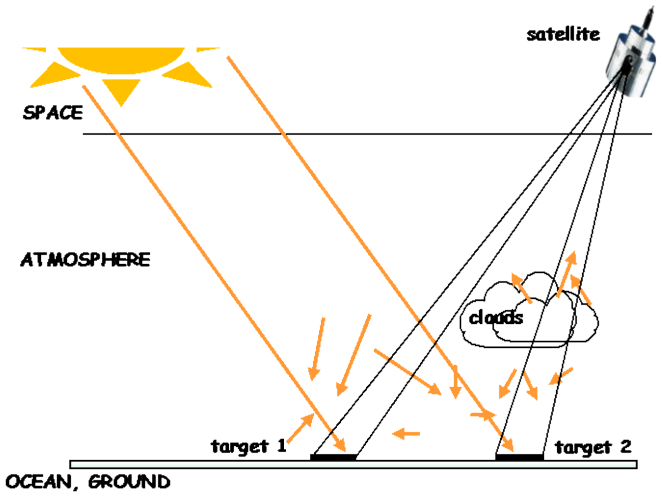


Figure 6.1. Measurement principle as used in cloud-index methods

The versions of Heliosat, which are inverse models, have in common to be divided into two parts regarding the physical modelling: converting the satellite image into a cloud index and converting the cloud index in irradiance. Therefore, they are called cloud index methods. Other methods with similar principles do not use cloud index. In the Tarpley's method (1979), the irradiance is estimated from the cloud fraction above the pixel.

The original Heliosat method makes use of the clearness index K_T. K_T is defined as the ratio of the SSI to the irradiance received at the top of the atmosphere. It characterizes the depletion of the solar radiation by the atmosphere. The cloud index n is converted into the clearness index by an empirical affine function $K_T = a\, n + b$, whose parameters a, and b, should be derived empirically by comparison with coincident ground measurements. These parameters can be computed for each location of ground station and then spatially interpolated to produce maps of parameters (Cano et al. 1986). They can also be averaged; the mean values are considered valid for a given region, e.g., Europe (Diabaté et al., 1988). Diabaté et al. (1989) observed that for Europe, three sets of parameters were needed: one for morning, one at noon, and one in the afternoon. A delicate part in the Heliosat method is the determination of the cloud-free instants. As Heliosat uses only one channel in visible range, the cloud-free instants should be detected by exploiting the time-series. A cloud-free instant should correspond to a minimum in the time-series, provided all other conditions are equivalent, which is not the case; for example, the sun position is changing within the day and also from day to day for the same hour. Espinar et al. (2009b) or Lefèvre et al. (2007) found that a relative error in the ground albedo leads to a relative error of the same magnitude in SSI under clear-sky, i.e., a relative error of order 10 % of the SSI in cloudy cases. Another delicate part in cloud-index based methods is the determination of the albedo of the brightest clouds. The error due to an error in this albedo increases as the sky is becoming cloudy; consequently, the relative error in the SSI can be very large, e.g., 60 % (Espinar et al., 2009b; Lefèvre et al., 2007).

Beyer et al. (1996) at the University of Oldenburg (Germany) produced a version called later Heliosat-1. It enhanced the original Heliosat method in several aspects. The major one is the adoption of the clear-sky index Kc instead of the clearness index K_T. The clear-sky index is defined as the ratio of the actual SSI to the SSI that would be received if the sky were clear. The great advantage of the substitution is that the relationship between Kc and n is universal and is now: $Kc = 1-n$. It has been found by these authors and confirmed by others that little was lost in quality by adopting this relationship for any part of the world and any time. Further work was done to remove partly the dependence of the received radiance with the viewing angle, thus leading to a more spatially-homogeneous cloud-index. In addition, work was performed on the determination of the ground and cloud albedo. Several empirical parameters used in this determination, e.g., the allowed change in time of the ground albedo or the threshold to detect cloud-free instants were revisited and new values were proposed to better account for actual measurements of SSI made by European ground stations.

To improve the accuracy and the reliability of the estimation and to facilitate the implementation of the method, Rigollier et al. (2004) designed the Heliosat-2 version at MINES ParisTech. It exploits the advances proposed by Heliosat-1 and seeks at removing empirical parameters. This is done by adopting several models that have been published independently of Heliosat or Meteosat. This requests a calibration of the Meteosat images to convert gray values into radiances and then reflectances. The clear-sky model proposed in the European Solar Radiation Atlas (ESRA) was adopted (Rigollier et al., 2000). The albedo of the brightest clouds is given by the model of Taylor and Stowe (1984a, b). This Heliosat-2 version is presented hereafter.

Zarzalejo et al. (2009) combined the Heliosat-2 method and the statistical approach proposed in the original version to determine the relationship between the clear-sky index Kc and the cloud-index. Actually, they searched a relationship between ground measurements, the cloud index n and several statistical moments of n, e.g., median, first and third quartiles. It is a means to account for the local climate effects. They found unbiased results for 28 stations in Spain and the relative error decreases compared to the Heliosat-2 version.

The Heliosat-3 version has been designed in a collaborative EU-funded project led by University of Oldenburg, and comprising MINES ParisTech and DLR among others. It is characterized by a clear-sky model, called SolIS, which is an approximation of radiative transfer equations for fast implementation (Mueller et al. 2004).

6.2. Overview of the Heliosat-2 method

Both HelioClim-3 and SOLEMI are constructed by the same Heliosat-2 method. They differ in the implementation. The concept of the Heliosat-2 method is as follows. The irradiance *I* for an instant *t* and location (*x, y*) is equal to

$$I(t, x, y) = I_c\,(t, x, y)\;K_c(t, x, y) \qquad (6.1)$$

where $I_c(t, x, y)$ is the irradiance for the clear-sky case; $K_c(t, x, y)$ is called the clear-sky index, is positive, and quantifies the depletion of I_c due to clouds. Thus, the method is based on *1)* a model of irradiance for clear-sky whose results are more or less depleted as a function of the cloud properties to yield actual irradiance. This concept is the basis of many published models outside Heliosat-2 (Rigollier et al., 2004).

The clear-sky index $K_c(t, x, y)$ is computed from the analysis of the Meteosat image at instant *t* and from the time-series of images prior to the current one. A cloud-index *n(t, x, y)* is defined:

$$n(t, x, y) = [\rho_{cloud}(t, x, y) - \rho(t, x, y)] \,/\, [\rho_{cloud}(t, x, y) - \rho_g(t, x, y)] \qquad (6.2)$$

where ρ, ρ_{cloud}, and ρ_g are the reflectances respectively observed by satellite for the pixel under concern, the brightest clouds, and the ground. The cloud index is close to 0 when the observed reflectance is close to the ground reflectance, i.e., when the sky is clear. It can be negative if the sky is very clean, in which case ρ is smaller than ρ_g. The cloud index increases as the clouds are appearing. It can be greater than 1 for clouds that are optically very thick.

An empirical relationship was derived from coincident ground measurements and Heliosat-2 results that links *n* to K_c (Figure 6.2):

$$\begin{array}{ll} n < -0.2 & K_c = 1.2 \\ -0.2 < n < 0.8 & K_c = 1 - n \\ 0.8 < n < 1.1 & K_c = 2.0667 - 3.6667\,n + 1.6667\,n^2 \\ n > 1.1 & K_c = 0.05 \end{array} \qquad (6.3)$$

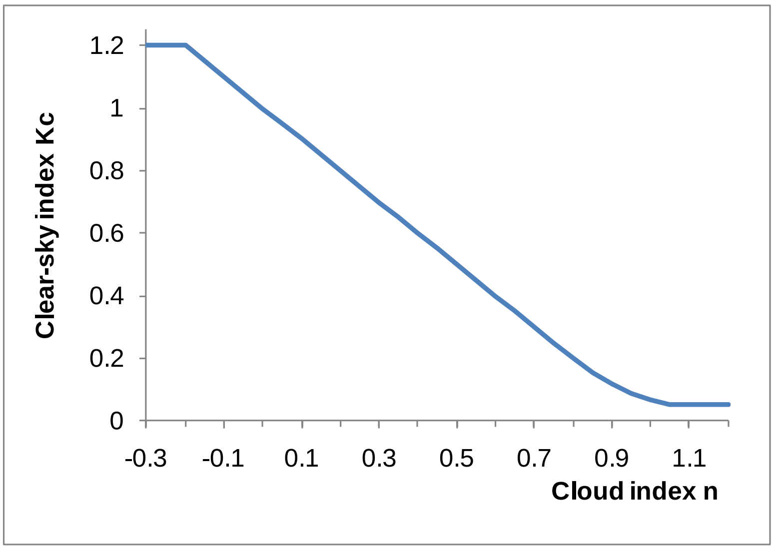


Figure 6.2. Relationship between the cloud index n and the clear-sky index Kc

For the computation of the DNI (direct normal irradiance) in the SOLEMI database, the following equation is used:

$$DNI = DNI_{clear} * \exp(a\ n) \tag{6.4}$$

where *a* is a number which depends on the viewing geometry, the brightness temperatures in thermal infra-red of the pixel, and the spatial variability in the cloud index.

The model of irradiance for clear-sky used for HelioClim is that of the European Solar Radiation Atlas (ESRA); the clear-sky model for SOLEMI is the model of Bird. These models and their inputs are discussed in further sections.

There are limitations in the implementation of this concept. One major limitation is that $I_c(t, x, y)$ is unknown. Knowledge on aerosols and other influencing atmospheric parameters is too poor to permit to retrieve on an operational basis the irradiance $I_c(t, x, y)$ for any time and any location. Therefore, the best that can be provided is a typical value of $I_c(t, x, y)$ for this instant and location. In order to cope with that uncertainty, the clear-sky index $K_c(t, x, y)$ is allowed to be greater than 1 while it should not in principle.

6.3. Inputs to the clear-sky models

The clear-sky models require inputs that are now discussed.

6.3.1. Aerosols

Aerosols have the strongest influence on clear-sky irradiances through absorption and scattering processes. Since aerosol particles are much larger than the solar irradiance wavelength, the scattering processes follow the Mie scattering theory. Unfortunately scattering and absorption processes cannot be well discriminated from each other. Ångstrøm introduced a formula covering both processes that provides the optical thickness as a function of the wavelength λ:

$$k_\lambda = \beta\ \lambda^{-\alpha} \tag{6.5}$$

β is the Ångstrøm turbidity coefficient indicating the aerosol content integrated in a vertical column of the atmosphere. The values are usually between 0 and 0.5. α is the wavelength exponent related to the size distribution of the aerosol particles. α usually is between 0.25 and 2.5 with an average of 1.3. Extreme values up to -0.5 or 3.0 are possible.

Modelling aerosols in the atmosphere is very difficult and is one of the major current tasks in atmospheric and climate research. Liu and Pinker (2005) give an overview of the current state of the art. The sources of aerosols are highly variable in space and in time. The interaction of the aerosol particles with the atmospheric trace gases and clouds is complex; the life time is approximately one week and is rather short. Models are making good progress in capturing aerosol evolution, but the characterization of the sources is still difficult (Tanré et al., 2005). Current state-of-the-art data sets currently include satellite observations of aerosols, precursor trace gases, clouds and precipitation and networks of surface-based instruments assimilated into a chemical transport model.

Chemical transport models are off-line models driven by meteorological data or from global circulation models which take aerosol processes as an integral part within the simulation scheme. The available global data sets have made use of satellite data. Many of them have been developed for the use in climate models, which analyze direct and indirect effects of aerosols in global warming. Kinne et al. (2001, 2003) give a comparative overview to the different available aerosol data sets.

SOLEMI currently is an on request processing, so different aerosol data sets are selected based on comparisons with local ground data (if available) or on a regional basis (e.g. GACP and AeroCom tend to match better in European areas whereas NCAR-MATCH showed best results on the Arabian Peninsula).

GACP Data Set
The GACP (Global Aerosol Climatology Project) data set has been prepared by the NASA Goddard Institute for Space Studies (GISS, Tegen et al. (1997). They used a chemical transport model to calculate optical thickness for sea salt, soil dust, sulfate, carbonaceous aerosols and black carbon. It has a vertical resolution of 4° and a horizontal resolution of 5°. Figure 6.3 exhibits the annual average of the aerosol optical thickness in this data set.

Aerocom Data Set
The AeroCom project is an open initiative to compare different aerosol data sets and models (Kinne et al., 2005). Within this project aerosol fields of the model median were created: All AeroCom models were regridded to a 1° by 1° horizontal resolution data set and then the center value was picked at each model grid separately for each month. Outliers were eliminated, which otherwise would have affected the average. The data set covers aerosol optical thickness, single scattering albedo and the Angstrøm parameter. Figure 6.4 exhibits the annual average of the aerosol optical thickness in this data set.

MATCH Data Set
MATCH (Model of Atmospheric Transport and Chemistry, Collins et al., 2001; Zender et al., 2003) is a chemistry transport model with an irregular grid of about 1.9°x1.9°. Inputs to the model are surface properties, emission data bases and wind and rain fields from NCAR/NCEP reanalysis data. The model calculates aerosol uptake, transport, the chemistry and change with time and the deposition of aerosols. The direct output of the models is mass concentrations which can be converted into optical thicknesses. In general the model is capable of calculating aerosol data for actual days. A dataset covering the full Meteosat period back to 1984 is in preparation within MACC. Up to now, based on data from the years 2000 – 2005, a climatology of daily values has been created by averaging the respective day within each year. This gives 366 different values for each grid cell. Figure 6.5 exhibits the annual average of the aerosol optical thickness in this data set.

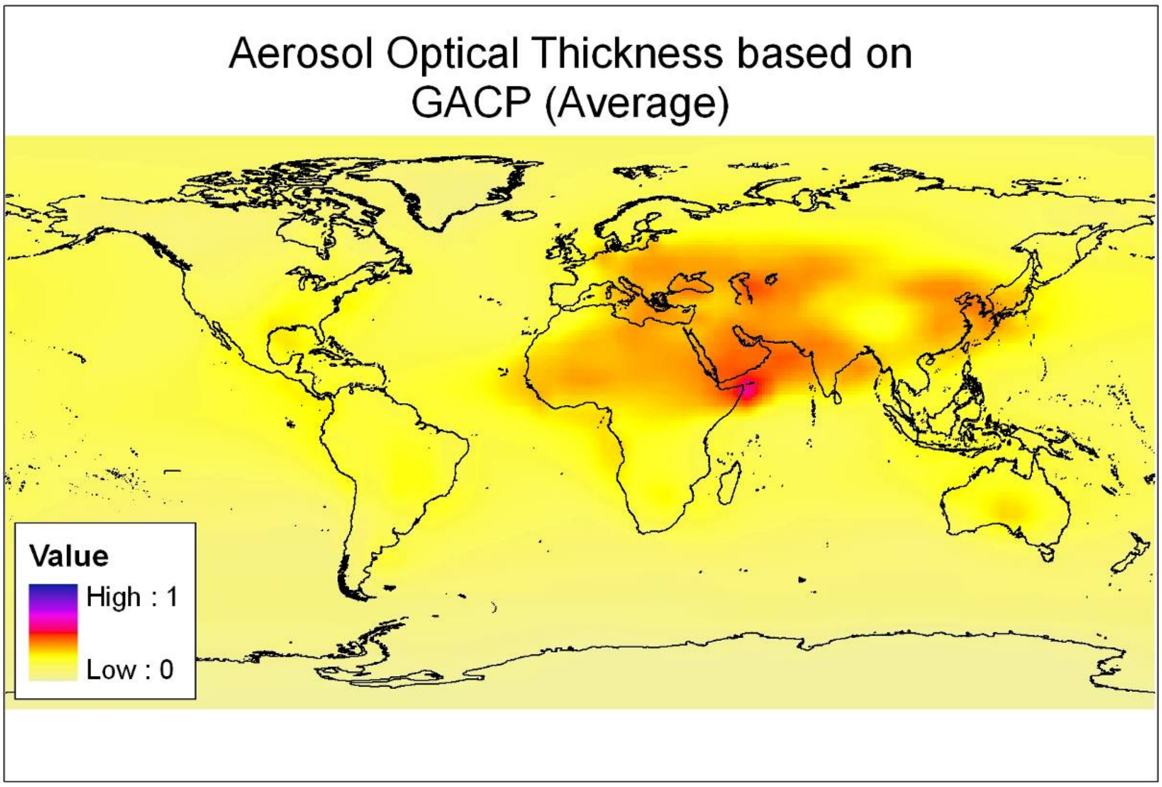


Figure 6.3. Annual average of the aerosol optical thickness in the GACP data set.

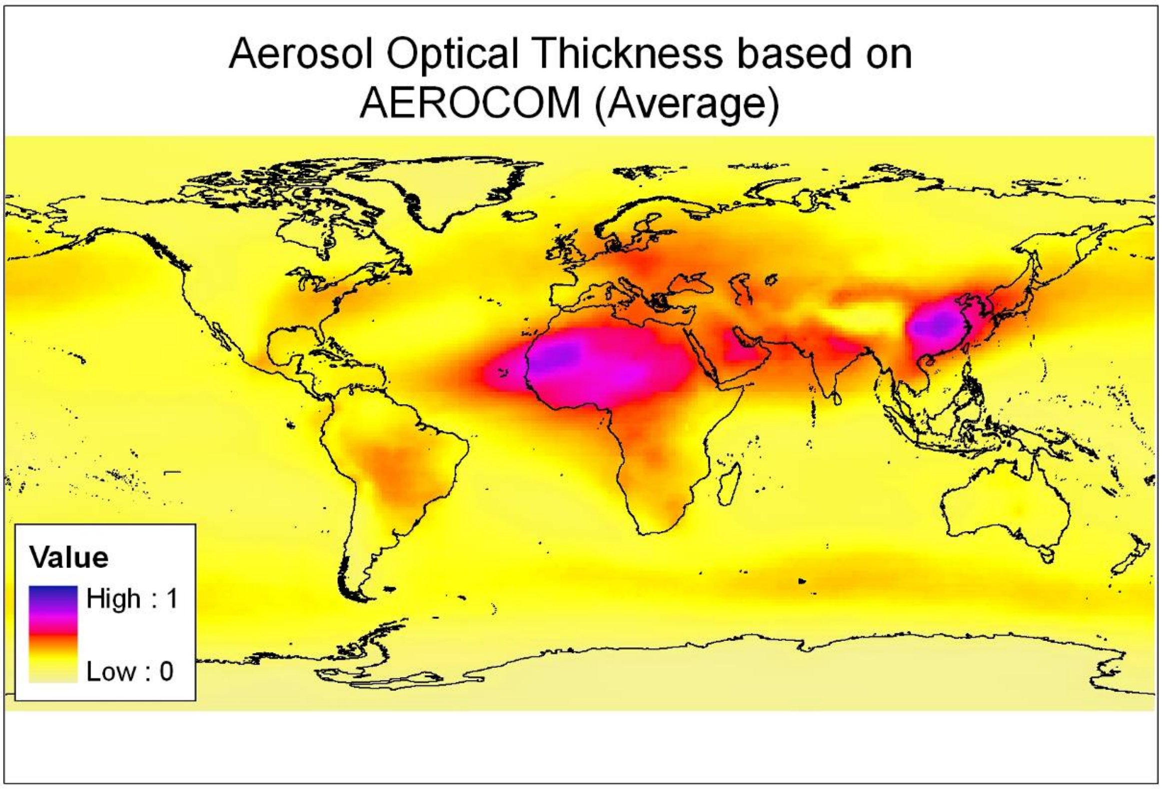


Figure 6.4. Annual average of the aerosol optical thickness for the Aerocom data set

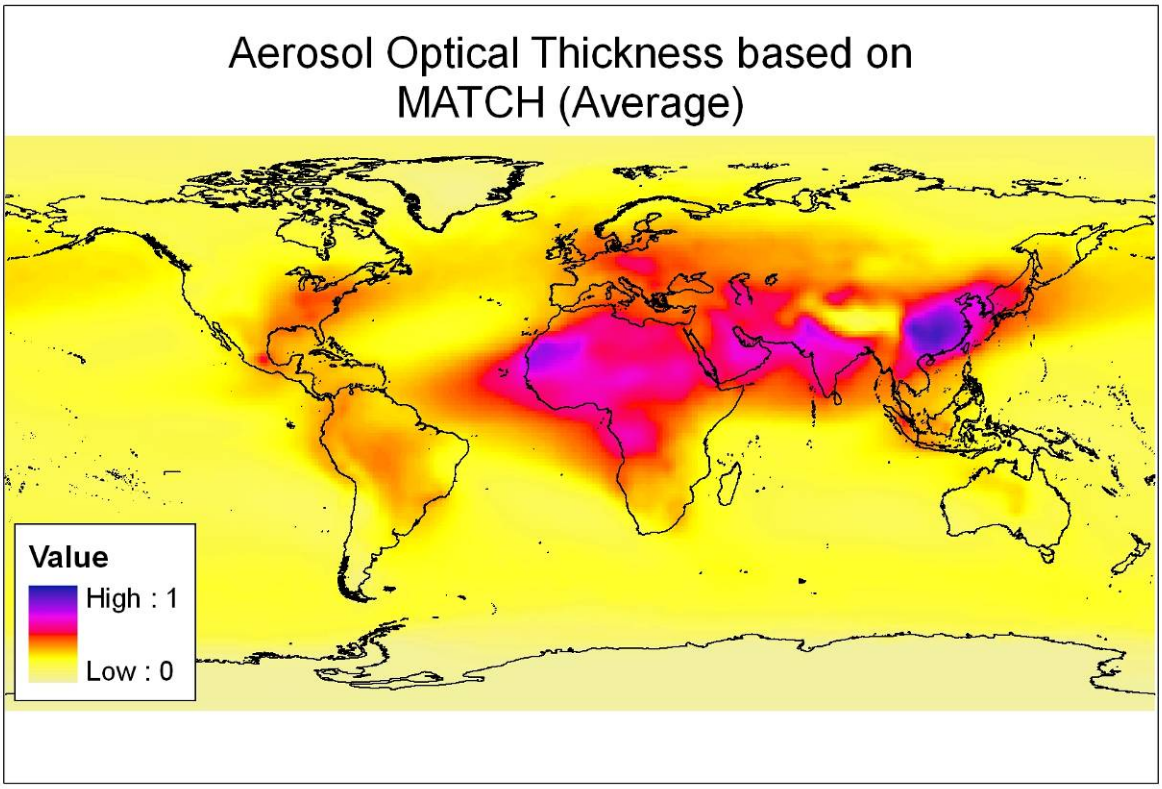


Figure 6.5. MATCH annual average (2000-2005) of the aerosol optical thickness.

6.3.2. Water vapour

Water vapour mainly absorbs the solar irradiance in the thermal spectrum and has a larger influence than ozone. Therefore a data set of the NCEP/NCAR-Reanalysis of the Climate Diagnostic Center (CDC-NOAA) with a spatial resolution of 2.5° x 2.5° is used. This data set is a joint effort of the National Centers for Environment Prediction (NCEP) and the National Center for Atmospheric Research (NCAR) in the United States providing historic analysis from 1948 onwards. The reanalysis continues with the Climate Data Assimilation System (CDAS) so that the products are available to the present. A detailed description of the data set can be found in Kalnay et al. (1996); Kistler et al. (2001).

Trenberth and Guillemot (1998) evaluated the NCEP/NCAR data set and found a correlation of above 0.9 for most regions of the world for precipitable water from the NVAP (NASA Water Vapor Project) dataset (Randeletal, 1996).The comparison has been done for the period 1988 to 1992 when NVAP data exist. Large differences occurred in the tropics and the Sahara region.

6.3.3. Ozone

Ozone absorbs the irradiance predominantly at wavelength lower than 0.3 µm. Therefore the extinction of ozone is fairly low for the complete solar spectrum. The variability of ozone depends mainly on geographical latitude and time of the year. In the solar belt the ozone concentration is between 0.2 and 0.4 cm[NTP][1],. Since the effect of ozone is very small, a data set from the Total Ozone Mapping Spectrometer (TOMS) sensor is used (McPeters et al., 1998). The TOMS sensor uses a six-band sensor measuring backscattered earth

1 The unit cm[NTP] refers to the thickness under normal temperature and pressure.

radiances. It was originally flown in a 500 km high orbit starting in July 1996 but was raised to a 750 km orbit in December 1997 to increase coverage. In this orbit a 90 % daily coverage is achieved. The derived ozone columns from the level 2 product are gridded into a 1° latitude by 1.25° longitude level product. As the daily products only have 90 % coverage and the influence of ozone is relatively small, a monthly data set is used. Compared to ground data, the derived ozone columns have a bias of about 1% (McPeters et al., 1998).

6.3.4. Linke Turbidity

The Linke turbidity factor (TL, for an air mass equal to 2) is a very convenient approximation to model the atmospheric absorption and scattering of the solar radiation under clear skies. It describes the optical thickness of the atmosphere due to both the absorption by the water vapour and the absorption and scattering by the aerosol particles relative to a dry and clean atmosphere. It summarizes the turbidity of the atmosphere, and hence the extinction of the direct solar radiation (WMO, 1981; Kasten, 1996). The larger TL, the larger the extinction of the radiation by the clear atmosphere.

The Linke turbidity factor denotes the transparency of the cloudless atmosphere. If the sky were dry and clean, TL would be equal to 1. When the sky is deep blue, TL is small. In summer, in Europe, the water vapour is often large and the blue sky is close to white. TL is larger than 3. In turbid atmosphere, e.g. in polluted cities, TL is close to 6 - 7.

A typical value of TL for Europe is 3. However, this value exhibits strong fluctuations in space and time as did the aerosols optical properties and the column-integrated amount of water.

A worldwide database for TL has been proposed by Remund et al. (2003). It has the form of gridded values, whose cells are squared and have a size of 5' of arc angle. There is one grid per month. For a given cell, the value of TL could be considered as representative of the monthly mean value averaged over several years. For a given day, one may interpolate the TL values for the month of this day and the closest month.

6.3.5. Conclusions

This section clearly demonstrates one of the current limitations of the methods for assessing the irradiance. The inputs used so far for the construction of the databases SOLEMI or HelioClim-3 are climatological values, whether they are aerosol properties and water contents or Linke turbidity factors. These climatologies cannot account for the daily variability nor the day-to-day variability of these parameters.

These inputs have a major influence on the value of the clear-sky SSI. The lack of variability impacts the quality of the retrieval of the SSI at daily and smaller time-scales. The impact is enhanced in the case of the DNI (direct irradiance at normal incidence).

The poor availability of accurate inputs is a recurrent problem, and not only in the retrieval of the SSI. The MACC Service will provide assessment of these inputs by the means of the re-analyses in the ECMWF model suite or the DLR/MATCH model version. In the transition step, we will incorporate more and more inputs from MACC. It is expected to have one value per day for each cell of say, 1° in size. The spatial description is coarse but a step forward in quality is expected by taking into account the day-to-day variability of the optical state of the atmosphere.

6.4. Brief description of the clear-sky models

6.4.1. The Bird model

The currently used Bird model has originally been proposed by Bird and Hulstrom (1981) and has later been modified by Iqbal (1983) as Model C. The broadband direct normal irradiance G_{dn} is calculated as:

$$G_{dn} = 0.9751\, G_{ext}\ \tau_r \tau_O \tau_g \tau_w \tau_a \qquad (6.6)$$

The factor 0.9751 is a conversion faction, since the model used for the development (SOLTRAN) considered the spectral interval of 0.3 µm - 3.0 µm. G_{ext} is the irradiance for a given instant on a plane normal to the sun rays at the top of atmosphere. The τ_x are the individual transmittances of the different atmospheric constituents. They are described below.

Air mass *m* – as used to describe water vapour and ozone transmission - is defined as:

$$m = \frac{1}{\cos\theta_z + 0.15(93.885 - \theta_z)^{-1.253}} \qquad (6.7)$$

For Rayleigh and aerosol transmission the SOLEMI model uses the air mass of Kasten and Young (1989) with a correction for the height *h* above sea level:

$$m_a = \frac{\dfrac{1-h}{10000}}{\cos\theta_z + 0.50572 * (96.07995 - \theta_z)^{-1.6364}} \qquad (6.8)$$

where θz is the solar zenith angle. The transmittance due to the molecular scattering τ_r is calculated by:

$$\tau_r = e^{-0.0903 m_a^{0.54} (1.0 + m_a - m_a^{1.01})} \qquad (6.9)$$

The transmittance τ_o due to ozone is:

$$\tau_c = 1 - 0.1611 U_3 (1.0 + 139.48 U_3)^{-0.3035} - \frac{0.002715 U_3}{(1.0 + 0.044 U_3 + 0.0003 U_3^2)} \qquad (6.10)$$

with

$$U_3 = l \cdot m \qquad (6.11)$$

where *l* is the vertical ozone layer thickness in cm and *m* is the uncorrected air mass. The transmittance τ_g due to uniformly mixed gases is given by:

$$\tau_g = e^{-0.0127 m_a^{0.26}} \qquad (6.12)$$

The transmittance τ_w due to water vapour is calculated from:

$$\tau_w = 1 - \frac{2.4959 U_1}{(1.0 + 79.034 U_1)^{0.6828} + 6.385 U_1} \qquad (6.13)$$

with

$$U_1 = w \cdot m \qquad (6.14)$$

where *w* is the precipitable water in cm and *m* the uncorrected air mass. The aerosol transmittance τ_a at 550 nm is given by:

$$\tau_a = e^{-k_a^{0.873}(1.0 + k_a - k_a^{0.7088}) m_a^{0.9108}} \qquad (6.15)$$

where the aerosol optical thickness k_a at 550 nm is either calculated by

$$k_a = 0.2758 k_{a,\lambda=0.38\mu m} + 0.35 k_{a,\lambda=0.5\mu m} \qquad (6.16)$$

or if α and β from the Angstrom law are known, the following equation from Mächler (1983) can be used:

$$\tau_a = (0.1245\,\alpha - 0.0162) + (1.003 - 0.125\,\alpha)\, e^{-\beta\, m_a (1.069\alpha + 0.5123)} \qquad (6.17)$$

β can be calculated from the aerosol optical thickness k_a at 550 nm:

$$\beta = \frac{k_{550}}{0.55^{-\alpha}} \qquad (6.18)$$

If unknown α is set to 1.3.

The clear-sky diffuse irradiance G_{diff} is calculated of three components, where $G_{diff,r}$ is the diffuse irradiance due to Rayleigh scattering, $G_{diff,a}$ due to aerosol scattering and $G_{diff,m}$ due to multiple reflections between surface and atmosphere. The diffuse irradiance due to Rayleigh scattering is:

$$G_{diff,a} = 0.79\, G_{ext} \cos\theta_z\; \tau_o \tau_g \tau_w \tau_{aa}\; F_c \frac{1 - \tau_{as}}{1 - m_a + m_a^{1.02}} \qquad (6.19)$$

where τ_{aa} is the is the transmittance of direct irradiance due to aerosol absorption:

$$\tau_{aa} = 1 - (1 - \omega_0) \cdot (1 - m_a + m_1^{1.06}) \cdot (1 - \tau_a) \qquad (6.20)$$

where ω_0 is the single scattering albedo of the aerosols. It is the relation of the radiation scattered by the aerosol to the total aerosol extinction. Bird and Hulstrom (1981) recommend to use 0.9 for ω_0 as long as no better value is available.

The aerosol-scattered diffuse irradiance after the first pass through the atmosphere is:

$$G_{diff,a} = 0.79 G_{ext} \cos\theta_z \tau_o \tau_g \tau_w \tau_{aa} F_c \frac{1 - \tau_{as}}{1 - m_a + m_a^{1.02}} \qquad (6.21)$$

where τ_{as} is the transmittance of direct irradiance due to aerosol scattering:

$$\tau_{as} = \frac{\tau_a}{\tau_{aa}} \qquad (6.22)$$

and *Fc* is the ratio of forward scattering to total scattering for which a value of 0.84 is recommended.

To calculate the irradiance by multiple reflections, the following expression for the atmospheric albedo ρ_a is suggested:

$$\rho_a = 0.0685 + (1 - F_C)(1 - \tau_{as}). \qquad (6.23)$$

The irradiance due to multiple reflections is then given by

$$G_{diff,m} = \frac{(G_{dn}\cos\theta_z + G_{diff,r} + G_{diff,a})\rho_g\rho_a}{1 - \rho_g\rho_a} \qquad (6.24)$$

where ρ_g is the ground albedo, which is set to 0.2. The diffuse irradiance is the sum of its components:

$$G_{diff} = G_{diff,r} + G_{diff,a} + G_{diff,m} \qquad (6.25)$$

Finally the total clear sky irradiance G_{clear} on the horizontal surface is the sum of the direct and diffuse components:

$$G_{clear} = G_{dn} \cdot \cos\theta_z + G_{diff} \qquad (6.26)$$

6.4.2. The ESRA model

The database HelioClim-3 is constructed by the means of the clear-sky model ESRA. This model was developed in the construction of the European Solar Radiation Atlas. It is based on Kasten's (1996) Rayleigh optical depth parameterization and the Linke turbidity factor at air mass 2. In this model, the global horizontal irradiance for clear sky, *G*, is split into two parts: the direct component, *B*, and the diffuse component, *D*. Each component is determined separately. This model is presented here very briefly. Details can be found in Rigollier et al. (2000) with revision proposed by Page and Remund and reported in Geiger et al. (2002).

The direct irradiance on a horizontal surface (or beam horizontal irradiance) for clear sky, *B*, is given by:

$$B = G_{ext} \cos\theta_z \exp(-0.8662\ TL(AM2)\ m\ \delta_R(m)) \qquad (6.27)$$

where

- G_{ext} is the irradiance normal to the solar beam at the top of atmosphere,
- θ_z is the solar zenith angle. θ_z is 90° at sunrise and sunset,
- *TL(AM2)* is the Linke turbidity factor for an air mass equal to 2,
- *m* is the relative optical air mass (note that *m* is computed slightly differently from Eq. 6.7),

- $\delta_R(m)$ is the integral Rayleigh optical thickness.

The quantity $\exp(-0.8662\, TL(AM2)\, m\, \delta_R(m))$ represents the beam transmittance of the beam radiation under cloudless skies. The relative optical air mass m expresses the ratio of the optical path length of the solar beam through the atmosphere to the optical path through a standard atmosphere at sea level with the sun at the zenith.

The diffuse irradiance falling on a horizontal surface for clear sky (or diffuse horizontal irradiance), D, also depends on $TL(AM2)$ at any solar elevation. When the turbidity increases, the beam irradiance falls, the proportion of the scattered energy in the atmosphere increases and the diffuse irradiance normally rises. At very low solar altitudes and high turbidity, however, the diffuse irradiance may fall with turbidity increase due to large overall radiative energy loss in the atmosphere associated with long path length. Thus, the diffuse horizontal irradiance, D, is determined by:

$$D = G_{ext}\, Trd(TL(AM2))\, Fd(\theta_z, TL(AM2)) \qquad (6.28)$$

In this equation, the diffuse radiation is expressed as the product of the diffuse transmission function at zenith (i.e. sun zenith angle is 0°), Trd, and a diffuse angular function, Fd.

7. Known problems in the retrieval of the SSI

Radiative transfer in the atmosphere is a complex phenomenon. In an operational method for the assessment of the SSI, methods should run fast at the expenses of the complexity of the models and therefore of the accuracy of the retrieved SSI. For example, several interactions between radiation and ground (e.g., reflections on the surrounding slopes) or clouds (reflections on the sides of clouds, multi-layered cloud ...) are currently not taken into account. Another problem is the low level of availability of data characterising the optical state of the atmosphere under clear skies. Climatological data are used by default at the expenses of a misrepresentation of the temporal variability of the SSI.

7.1. Sub-pixel phenomena

There will be an error when the condition differs from the average state. The frequency (15 min) of observation by satellite is very satisfactory to describe the transitional phenomena such as convection. But the size of the pixel is not adapted to the micro-meteorology. There is a spatial integration which smoothes the phenomena. Meteosat pixels have actually an elliptic shape and their average diameter ranges from 1 km to 8 km depending on the viewing geometry of the satellite. Figure 7.1 shows a typical Meteosat First Generation pixel in Central Europe with N-S extent of about 4 km and E-W extent of about 3 km. The resolution of this satellite at sub satellite point is 2.5 km at ground.

Figure 7.1. Typical Meteosat pixel in Europe with sizes of about 3x4 km².

In this example, one can note the presence of clouds in the ellipse, whose size is much lower than the pixel size.

7.2. Empirical parameters in method

Another limitation of the current methods is that are still included parameters determined empirically that significantly influence the quality of the results. Lefèvre et al. (2007) has shown that an error of 0.05 on the estimated ground albedo causes an error of 9 % on the estimated SSI. Similarly, an estimation error of 25 % on cloud albedo can cause up to 60 % error on SSI.

7.3. Change in terrain elevation within a grid cell in databases

The computation of the SSI from satellite images calls upon a digital terrain model (DTM) whose cell size fits that of the pixel. For example, the ESRA atlas (ESRA 2000) or the HelioClim-1 database (Lefèvre et al. 2007) exploits the DTM TerrainBase (TerrainBase 1995) whose cell size is 5' of arc angle, i.e. approximately 10 km at mid-latitude. The size of the cell is even larger for the NASA-SSE database: 1° of arc angle (Chandler et al. 2004) or for the ISIS database: 280 km (Lohmann 2006).

These sizes are too large to describe changes in altitude with a sufficient accuracy in areas of steep relief; large discrepancies can be found between the mean altitude of a cell and the altitude of a particular site within this cell (Figure 7.2). For example, in Switzerland, the altitude of the measuring station at Saentis Mountain is 2490 m while that provided by the DTM TerrainBase is 1126 m, i.e. an underestimation of 1364 m. Wahab et al. (2009) report that a difference in altitude of 300 m may induce a relative difference greater than 1 % on the monthly mean value of the SSI. This means that change in elevation must be accounted for.

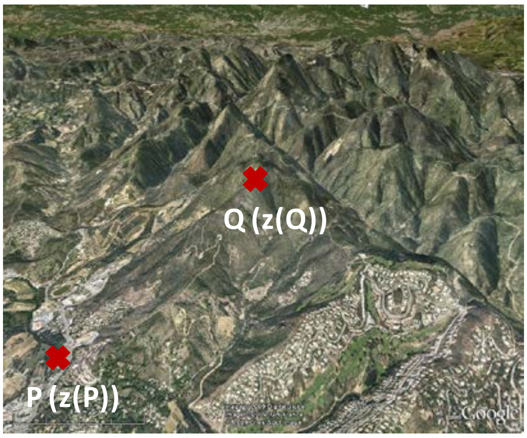


Figure 7.2. Sample of cell in a SSI database. Change in altitude is very large between site P and site Q

In very steep relief, irradiance depends upon shadows cast on the sites by surrounding obstacles (Figure 7.3) and not only on change in altitude.

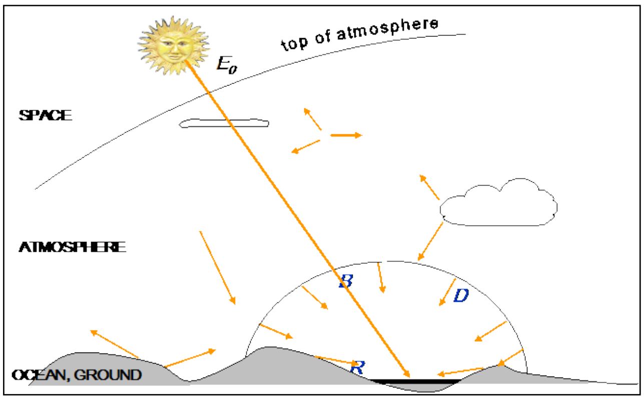


Figure 7.3. Incident irradiance in a complex terrain. B: direct irradiance; D: diffuse irradiance; R: irradiance reflected by nearby terrain

Figure 7.3 shows that the SSI on a horizontal surface is composed of the direct irradiance, diffuse irradiance partly masked by the surrounding mountains, and a reflected part R due to the reflexion on the surrounding slopes. Most often, if not always in operational methods, R is not accounted for; the irradiance calculations are done under the assumption of a flat terrain within the pixel. In that case, the tilt angle α and azimuth angle β of the element receiving the radiation are set to 0 and the cosine of the local incident angle θ is:

$$\cos\theta(0,0) = \cos\omega \cos\delta \cos\phi + \sin\delta \sin\phi \quad (7.1)$$

where ω is the hour angle, δ is the solar declination, and ϕ is the latitude of the site.

In case of non-flat pixel, for each element (*dx, dy*) within a pixel, the cosine of the local incident angle θ is:

$$\cos\theta(\alpha, \beta) = (\cos\omega \cos\delta \cos\phi + \sin\delta \sin\phi) \cos\beta + \cos\omega \cos\delta \sin\phi \cos\alpha \sin\beta + \sin\omega \cos\delta \sin\alpha \sin\beta - \sin\delta \cos\phi \cos\alpha \sin\beta \quad (7.2)$$

where α and β correspond to the direction of the local slope, respectively in azimuth and tilt. Thus, the SSI of the pixel should be modified by the ratio *R'*:

$$R' = \iint_{pixel} \cos\theta(\alpha(x, y)), \beta(x, y)) \, dx \, dy \,/\, \cos\theta(0,0) \quad (7.3)$$

7.4. Bidirectional reflectance and albedo

Reflexion properties of the ground are a function of the incident and viewing angles. Up to now, these parameters are not operational available. Figure 7.4 shows the angular variation of the reflectance of a coniferous forest in the near infrared (Oumbe 2009).

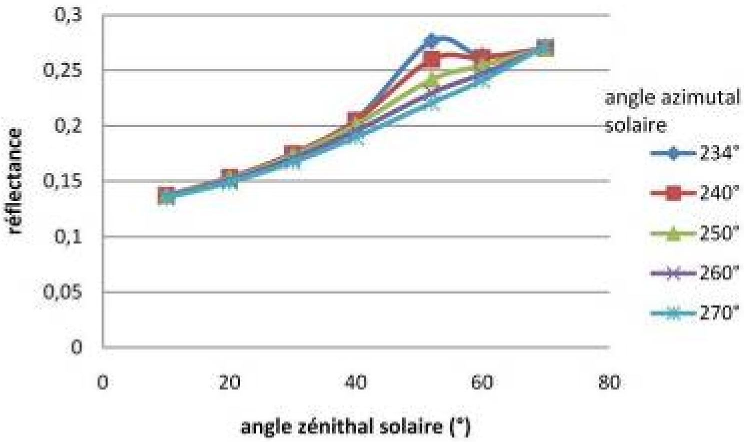


Figure 7.4. Example of variation in ground reflectance (vertical axis) with the solar zenithal angle (horizontal axis). Viewing zenithal and azimuthal angles are respectively 52° and 234°.

The change in reflectance with the angles can be large; it is greater than 0.1 in Figure 7.4. It can be much larger in the specific case of oceans, where the reflectance depends also on wind speed and varies from values close to zero to values greater than cloud reflectance (Lefèvre et al. 2007). By considering the hemispherical albedo instead of the bidirectional reflectance, one commits a significant error on the part of irradiance reflected by the ground then backscattered by the atmosphere, thus contributing to the diffuse fraction of the SSI. This omission is very often made for operational reasons because of the lack of data describing the ground.

A similar effect can be found in case of clouds because cloud reflectance changes with illuminating and viewing angles. The method Heliosat-2 accounts for such changes by adopting the models of Taylor and Stowe (1984a, b).

7.5. Cloud vertical position – Snow cover

Contrary to objects at ground, clouds are located at different altitudes in the atmosphere. If the viewing angle is large, i.e. far from the satellite nadir, the parallax effect becomes noticeable for clouds at high altitude and not too thick. This results in an erroneous assessment of the geographical position of the cloud. The cloud will be assigned to a pixel farther from the nadir of the sensor than the actual one. The pixel over which the cloud is actually located will be seen as a cloud-free pixel and the SSI assessment will be inaccurate.

The current method Heliosat-2 does not account for the sudden appearance of snow; large errors may appear in the presence of snow cover in cloud-free atmosphere. The satellite derived irradiance values indicate a cloudy sky, since the snow covered pixels appear bright as clouds in the visible channels of the satellite images, as illustrated in Figure 7.5.

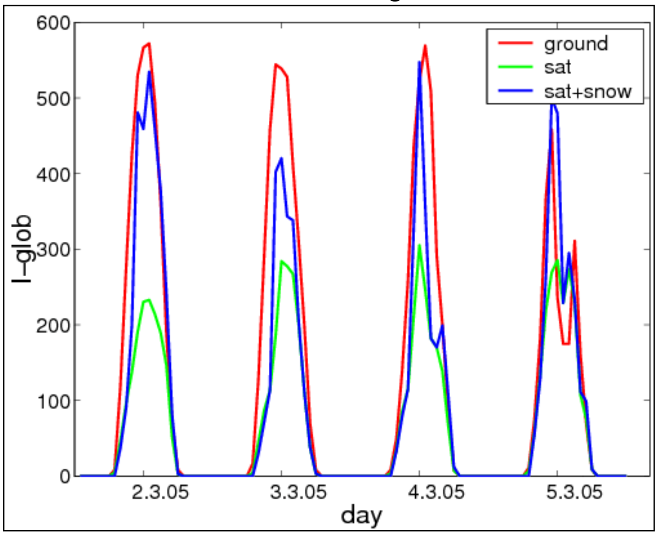


Figure 7.5. Example for satellite derived global irradiance with (blue) and without (green) snow detection in comparison versus ground measured values for a German station in March 2005 (source Univ. Oldenburg).

This figure stands for a German measuring station in March 2005. During this month, the sky was clear for many days and the snow was covering the ground. The ground measurements in red show large values in irradiance, while the SSI retrieved by the Oldenburg method (similar to Heliosat-2) are much too low (in green). An attempt to correct this drawback was successful in this case (blue line).

7.6. Input data

Aerosol loading and water vapour amount are difficult to measure with remote sensing methods over land. The retrieval of aerosols is handicapped by the small aerosol reflectance and the perturbation of the weak signal by clouds and surface reflection. It can also be difficult to distinguish between water vapour and clouds. Thus, aerosols and water vapour data are usually taken from numerical model reanalyses and the accuracy and resolution are limited. These data are often available on a daily or monthly basis and with a resolution close to 1° or coarser. Furthermore, there are many different data sets available. It may be difficult to select a specific one. It may happen that the data set matching most of the validation sites may not be the best for a specific site. Figure 7.6 shows three different aerosol data sets. It can be seen that the absolute values and distributions are very different even if only the annual average is compared.

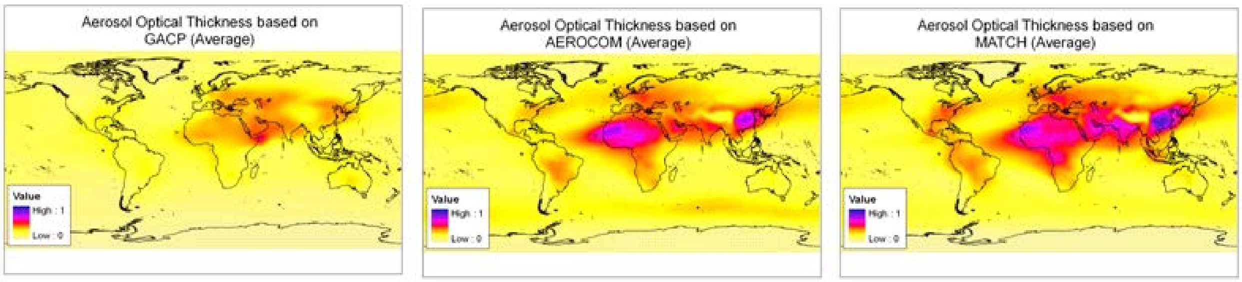


Figure 7.6. Different global annual average aerosol data sets. The colour scaling is the same on all three images.

Nonetheless, aerosol loading and water vapour amount are very variable in time and space and have a great influence on SSI, particularly in clear skies. Wald and Baleynaud (1999) demonstrate noticeable variation in atmospheric transmittance due to local pollution in cities at scale of 100 m. The error induced on the SSI depends on the variation of these parameters within the pixel. The clearer the sky, the greater the error. Comparisons between ground measurements of hourly means of SSI made at sites in Europe less than 50 km apart for clear skies show that the spatial variation in SSI, expressed as the relative root mean square difference, can be greater than 10%.

The main influence on direct irradiance comes from clouds, determined from pictures taken at a distance of 36 000 km. The distinction between different cloud types is very difficult. Observations made by satellite bear a spatial average over pixels; a satellite-derived information of 50% cloud cover for a pixel can result from a 50% semitransparent homogenous cloud or a broken cloud field with 50% cloud coverage within this pixel. These two cloudy conditions attenuate the radiation in a very different way.

Therefore, models have to make simplifying assumptions to transform the cloud information derived by the satellite into an effective cloud transmission. E.g. the transmission of direct normal irradiance is estimated with a simple exponential function depending on the cloud observations in the visible and infrared channels of the satellite. Figure 7.7 gives such a sample based on the visible cloud index. The applied function may not be the best for all sites and climatic conditions.

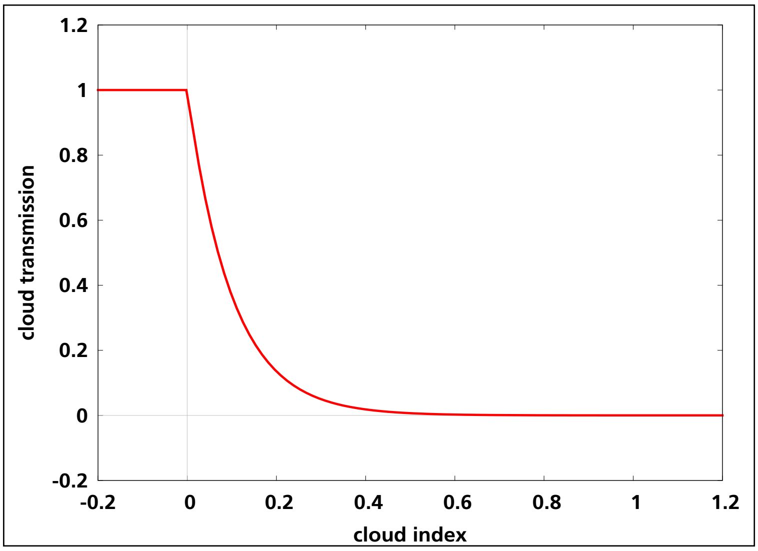


Figure 7.7. Example of transfer function from cloud observation (cloud index) to cloud transmission.

The variability in space and time has been reported in the scientific literature. For example, Wielicki and Welch (1986), Rossow and Schiffer (1999) note typical scale of variation of 30 m and 10 min. This is smaller than the spatial and temporal sampling in the Meteosat images: 3 km and 15 min.

8. Overview of the operation chains (workflow) in the HelioClim-3 and SOLEMI legacy services

The present HelioClim-3 and SOLEMI services share the same method Heliosat-2 for converting satellite images into SSI. They differ in workflow for computing irradiance products, i.e. the inputs, operations and their chaining.

The major reason for difference lies in their concept. The HelioClim-3 service aims at delivering products in an automated manner with a large capability of dissemination (1000+ requests per day), while the SOLEMI service has been conceived to deliver products with a lower dissemination level.

Nevertheless, both workflows for computing irradiance products can be represented by the same schematic representation in Figure 8.1.

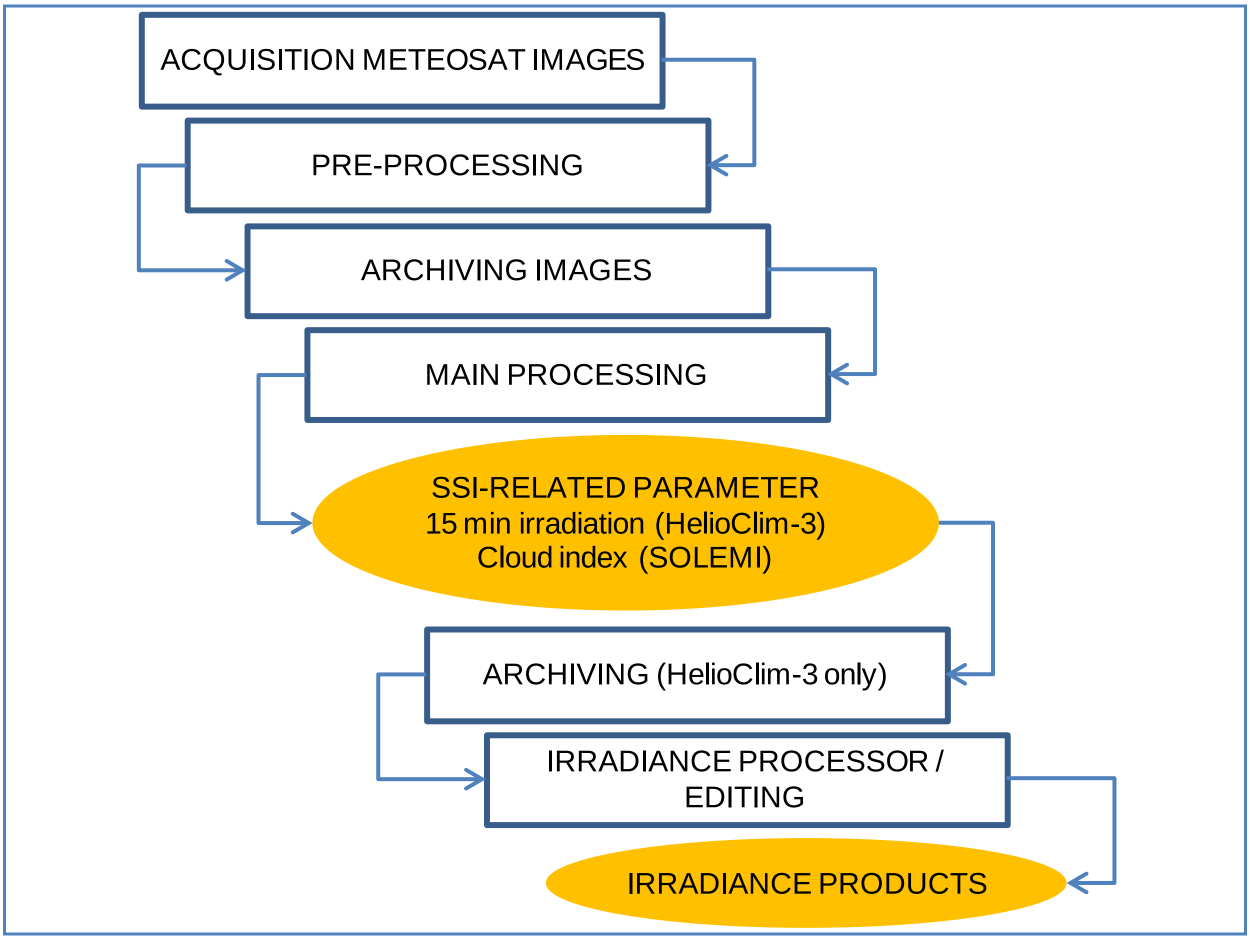


Figure 8.1. Schematic representation of the workflow of the HelioClim-3 and SOLEMI services for computing irradiance products

Meteosat images are acquired by the means of a receiving station. A pre-processing is performed that includes the control of the quality of images and calibration. These images are stored into a database.

The main processing converts Meteosat images into images of a parameter that is related to the SSI in the Heliosat-2 method: the 15-min irradiation for HelioClim-3 and the cloud index *n* for SOLEMI. This method has been described earlier in the User's Guide (Chapter 0) as well as several scientific articles.

In the HelioClim-3 service, the main processing is triggered automatically at the end of the day once all images have been received; the 15 min irradiations are stored into the HelioClim-3 database. The editing part is triggered at the request of a user. It performs a post-processing of the requested time-series extracted for the HelioClim-3 database and a formatting of the outputs to fit the user requirements. This post-processing corrects the irradiation to take into account the difference between the ground elevation of the requested site and that used to compute the irradiation for the given Meteosat pixel. It also performs an empirical correction of the irradiation to palliate observed defects and computes the uncertainty.

In the SOLEMI service, a request triggers the main processing that produces the cloud index and then the irradiance processor that produces the requested irradiance.

8.1. Overview of the workflow in the HelioClim-3 service

Figure 8.2 displays a schematic representation of the workflow for producing irradiance products in the HelioClim-3 service.

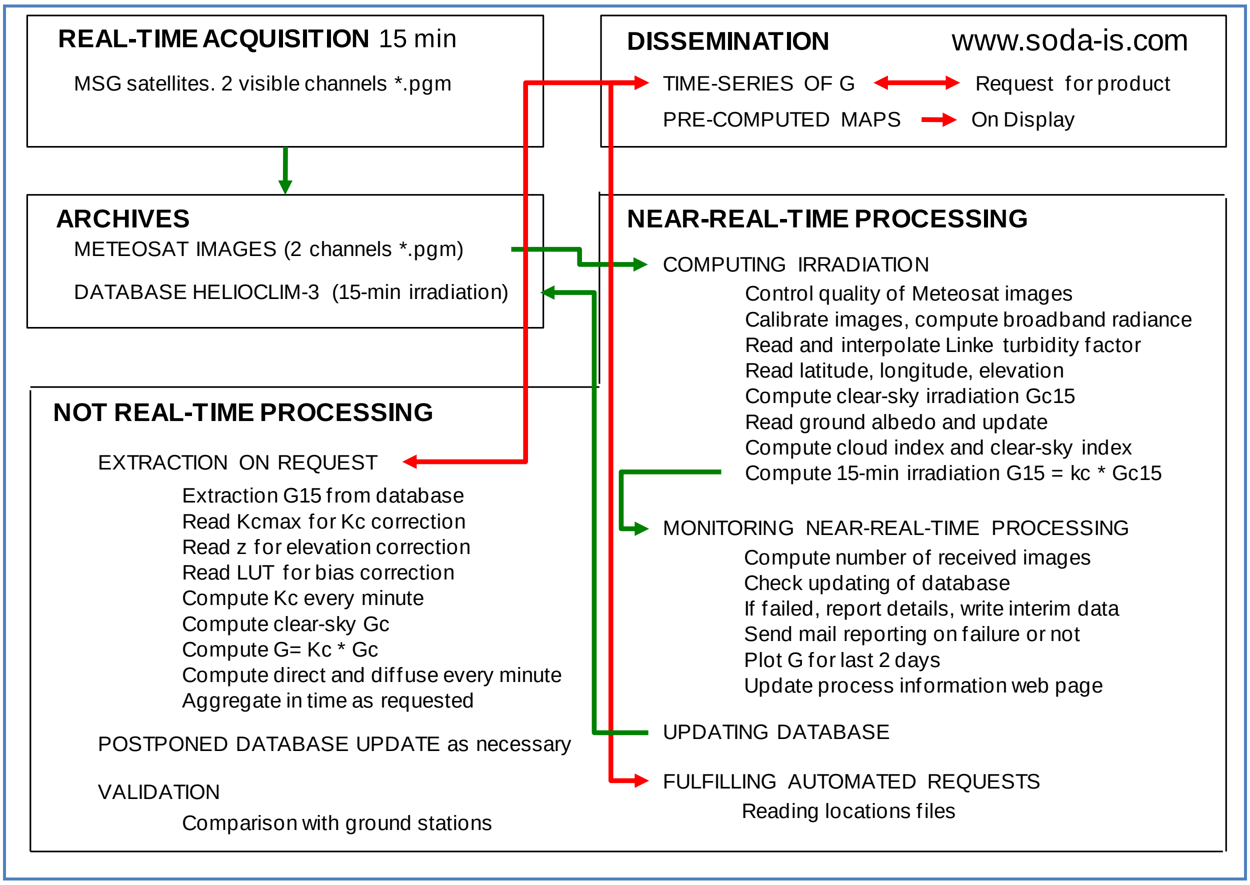


Figure 8.2. Overview of the workflow in the HelioClim-3 service for computing irradiance products

Meteosat images are received every 15 min by a PC connected to an antenna. Only the two channels in the visible band are kept. The images are stored into a database.

The near-real-time processing is triggered at the end of the day by the reception of the last images at 21:00 UT. The first operation is the control of the quality of Meteosat images. It consists in looking for the missing lines if any and in the check of the position of extreme lines and columns in each image. If a failure is observed, the processing stops.

Meteosat images are calibrated using the calibration coefficients supplied in the flow emitted by Meteosat, resulting into two images of radiance every 15 min. These two images are then combined to create an image of broadband radiance (Cros et al. 2006).

Inputs to the method Heliosat-2 are the latitude, longitude, and elevation of the current pixel. Another input is the Linke turbidity factor characterising the optical state of the clear atmosphere. It is read from a database containing monthly values and interpolated for the day under concern. The clear-sky irradiation for the 15 min period is computed by the ESRA model, as explained in Chapter 0. Then, using a database of ground albedo, the cloud-index, then the clear-sky index, and finally the irradiation G15 for the 15 min period are computed. The irradiation is stored in the HelioClim-3 database.

There are a number of control points in the workflow for monitoring the smooth running of the near-real-time processing. In case of failure, detailed reporting is made to the management, written in a log file and sent by e-mail as well. A visual monitoring is in place to perform visually a gross check of the computed irradiance. It comprises a graph of the irradiances for the last two days for three selected sites: Sophia Antipolis where the chain is operated, and two extreme locations in the East and the West on the Equator. A Web page is updated that provides access to the archive of log files and graphs. Figure 8.3 provides examples of graphs produced by the workflow monitoring during the day 2013-05-22 at 09:30 CET time.

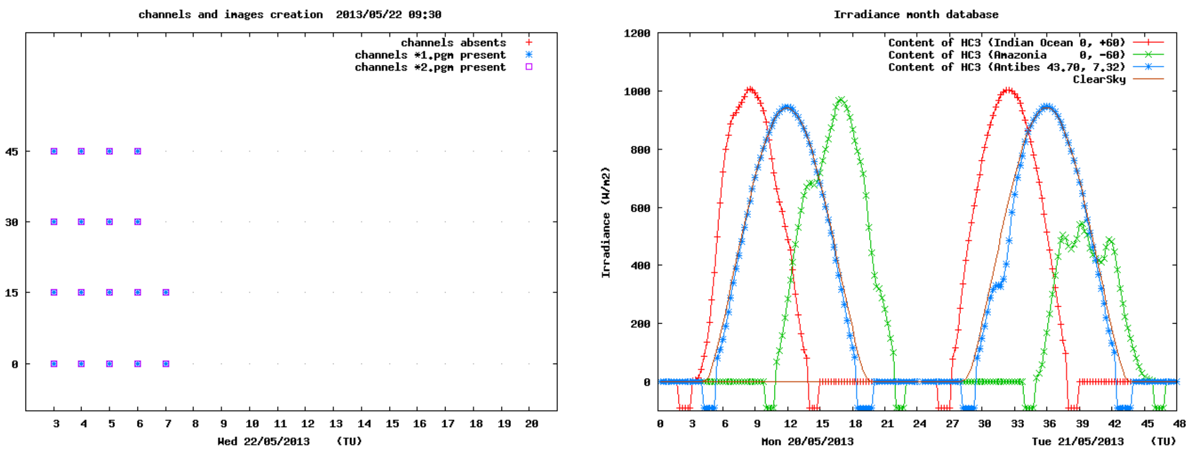


Figure 8.3. Examples of graphs produced by the monitoring of the workflow. Left: images acquired and controlled during a day. Right: irradiance computed for three sites for the past two days.

Radiation products are time-series of irradiation for a given site and a given period. Request for products are made through the Web site of the SoDa Service (www.soda-is.com). There are two types of requests: manual and automated. Manual requests are made by the means of a browser. Automated requests are emitted by computers. The latter are typically requests for hourly irradiation every day for large numbers of geographical sites. Automated requests are processed in the same way than the manual requests, except that their processing is part of the near-real-time processing and takes part immediately after the updating of the HelioClim-3 database.

Maps are computed on an *ad-hoc* basis and are set on display in the SoDa Service.

Once extracted from the HelioClim-3 database, the time-series experiences a post-processing before delivery. The post-processing comprises several steps. As several benchmarking activities demonstrated a few biases in the raw irradiation data contained in the HelioClim-3 database, *a posteriori* corrections to the raw irradiation are brought to the raw values in step #1 and #3:

1. The raw data are converted into clear-sky indices. These indices are corrected to take into account the non-lambertian properties of the ground for a given pixel and given day.
2. The following step corrects the resulting irradiation to take into account the difference between the ground elevation of the requested site and that used to compute the irradiation for the given Meteosat pixel.
3. Step #3 is an empirical correction of the irradiation as a function of the irradiation itself. The empirical correction has been defined through the benchmarking activities. Briefly speaking, low values are lowered and high values are amplified.
4. In order to cope with gaps and to compute direct and diffuse components, the clear-sky index is interpolated every minute. If requested, the algorithms of Ruiz-Arias et al. (2010) are used to compute direct and diffuse components on horizontal surface from the global irradiance every minute.
5. The 1-min values are summed up to yield the requested aggregation period, e.g. hour, day, or month.
6. The uncertainty model is performed in order to associate lower and upper bounds to each data for global irradiation on horizontal surface. There is a 68% chance that the actual value is comprised between the lower and upper bounds. This model is described in Wald et al. (2011).

Comparison between irradiation products and measurements at ground by well-calibrated instruments is made as often as possible. This is not systematic as the access to such data is limited. The protocol and performance indicators are described in Chapter 0 of the User's Guide. The results of such comparisons are used:

- to detect possible flaws in the method or processing workflow, and correct them,
- to verify the consistency in quality of the delivered products,
- to document the uncertainty in the retrieval and to display it in the Web site,
- to improve *a posteriori* correction procedures if possible in order to improve the quality,
- to improve the model of uncertainty allowing to allocate uncertainty values to each irradiance value.

Several versions of HelioClim-3 have been made available that reflect changes in the post-processing part. All these versions and the improvements are documented in the SoDa Service. On 1st January 2011, i.e. in the course of the project MACC, the version 3 (HC-3v3) was set in operation.

8.2. Improvements and changes in the current HelioClim-3 workflow

An empirical function is applied to the data extracted from the HelioClim-3 database in step#3 of the post-processing (editing part of the workflow). The only input to this function is the HelioClim-3 irradiation itself.

A new correction function has been devised in the course of the MACC-II project. The methodology is similar to the previous one. It is based on the deviations between the HelioClim-3 data, corrected from altitude, and ground measurements aggregated over 15 min. However, now 15 BSRN stations are used located from Estonia to Brazil from 2005 to 2010, instead of 32 stations mostly located in Western Europe with various periods of measurements. Here, 15 min values are used instead of hourly values. Finally, the deviations are put into bins of clearness index and solar elevation angles instead of irradiation as previously. There is a total of 768 946 samples for the analysis.

Figure 8.4 illustrates the main defect of HC-3v3. This figure (left) exhibits a 2-D histogram of the deviation between HC-3v3 and corresponding ground measurements as a function of the solar elevation angle (the complement to 90° of the solar zenithal angle). The black line is the mean deviation for each angle. One may observe an underestimation of the irradiance (negative deviation) for solar elevation angles less than 25° (solar zenithal angle greater than 65°). For angles greater than 30° (solar zenithal angles less than 60°), the overestimation increases with the elevation angle.

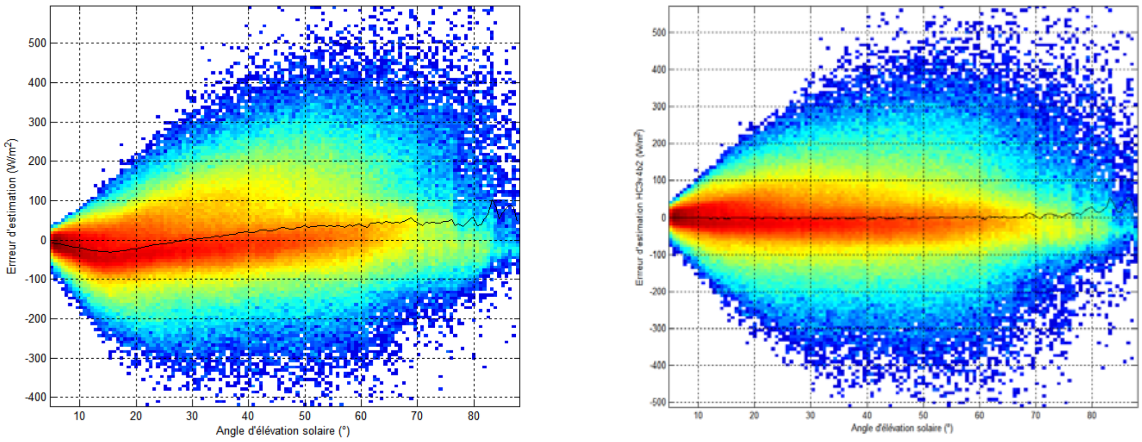


Figure 8.4. 2-D histogram of the deviation between HC-3v3 (left), or HC-3v4 (right), and corresponding ground measurements as a function of the solar elevation angle (horizontal axis). Vertical axis is the deviation. Colours represent the number of samples, increasing from dark blue to dark red. The black line is the mean deviation for each angle.

The new correction function, *i.e.* the HC-3v4 version, removes the influence of the solar elevation angle, as shown in the figure (right). The mean deviation (black line) is always equal to 0, except for the very large elevation angles greater than 80°.

Table 8.1 shows the improvements of performances between v3 and v4. The four indicators are improved in a noticeable way.

	Mean irradiance (W m^{-2})	Bias (W m^{-2})	RMSE (W m^{-2})	Correlation coefficient	KSI (W m^{-2})
HC-3v3	394.6	5.0	90.6	0.953	8.3
HC-3v4	394.6	-0.8	84.4	0.958	6.9

Table 8.1. Performances of HC-3v3 and HC-3v4 for all stations merged.

The version 4 (HC-3v4) was set in operation on 8th April 2013.

8.3. Overview of the workflow in the SOLEMI service

Figure 8.5 presents a schematic representation of the workflow in the SOLEMI service for producing irradiance products. There are several similarities in the operations with the HelioClim-3 service.

Meteosat images are received by a receiving station at DLR. This station is external to the SOLEMI service and acts as a provider of images. Images are quality-controlled for all channels. They are then cut into overlapping tiles, which are stored into a database. This database is part of the DIMS system of DLR (Data Information and Management System). The SOLEMI service exploits data every 30 min for the Meteosat and every 15 min for the Meteosat Second Generation satellites.

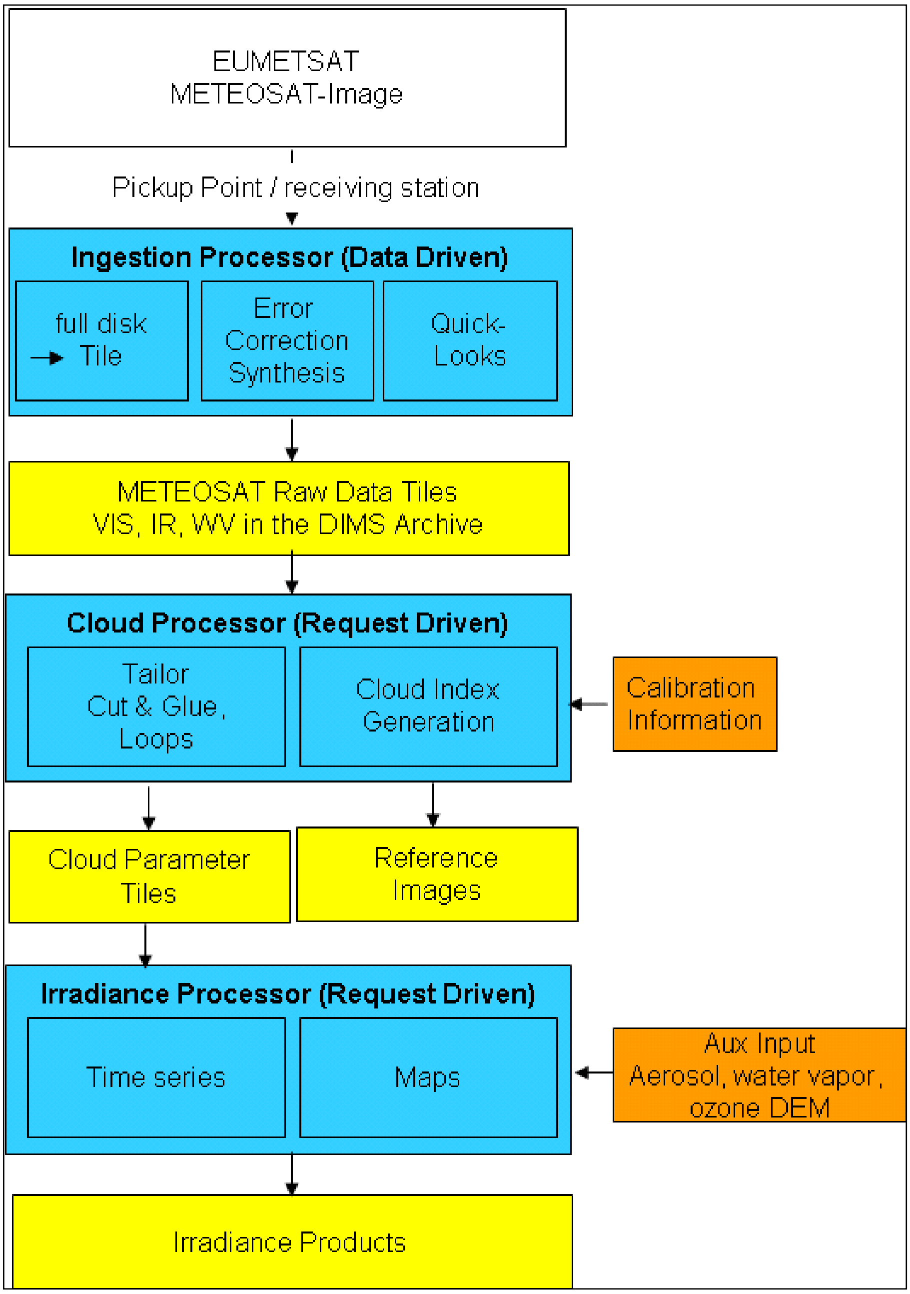


Figure 8.5. Overview of the workflow in the SOLEMI service for computing irradiance products

The processing to produce an irradiance product is made only on-request. Standard requests are for time-series for a given site, or for maps. As for the HelioClim-3 service, the processing for maps is the same than for a single site. Inputs are latitude, longitude, and elevation of the site of interest.

Meteosat images are extracted from DIMS and calibrated. The two images in the visible band are combined to create an image of broadband radiance (Cros et al. 2006).

Other inputs to the method Heliosat-2 in SOLEMI are the aerosol properties, the total water column, and the total column ozone. These parameters are read from databases containing monthly or daily values and interpolated for the day under concern if needed. The clear-sky irradiance is computed by the Bird model. Then, using a database of ground albedo, the cloud-index, then the clear-sky index, and finally the irradiance are computed.

Products are created by aggregating irradiance values over the requested aggregation period, e.g., hour, day, month.

There are a number of control points in the workflow for monitoring the smooth running of the processing. Each time a failure is observed, reporting is made.

A manual inspection of the product is made based on scientific experience, taking into account the history available in the archive. Such procedures include comparing monthly means, e.g. all months of January in a period spanning several years, or comparing annual patterns and frequency distributions of single years against each other. If larger areas are processed a visual inspection of satellite raw data in movies is done.

Comparison between irradiance products and measurements at ground by well-calibrated instruments is made as often as possible. This is not systematic as the access to such data is limited. The results of such comparisons are used:

- to document the uncertainty in the retrieval displayed on the Web site,
- to detect possible flaws in the method or processing workflow, and correct them.

8.4. Improvements and changes in the current SOLEMI workflow

Recent initiatives from NCAR (USA, MATCH scheme) and the EU-funded project GEMS yielded to accurate assessments of optical properties of the aerosols with capability of producing daily maps. These recent results have been ingested in the SOLEMI workflow instead of the climatological maps described in Chapter 6. This results in a better accounting of the day-to-day changes in aerosol properties (Schroedter-Homscheidt et al., 2013). Comparison with ground measurements of irradiance will confirm the improvement. This is a the first step in the transition process. The change in workflows is small.

PART C.

THE NEW HELIOSAT-4 METHOD

AND THE NEW HELIOCLIM-4 DATABASE

9. Overview of the new Heliosat-4 method

The new Heliosat-4 method is based on the principle that the SSI *I* for a cloudy atmosphere can be approximated in an accurate manner by the product of the irradiance *Ic* obtained under a clear sky for a null ground albedo and a function *f(cloud, albedo)* modelling the extinction of the radiation by clouds and the contribution due to the ground albedo:

I = Ic(albedo=0) f(cloud, albedo) (9.1)

This relation is very interesting for practical reasons, in particular for rapid calculations. Each part of the equation can be processed following the available spatial and temporal resolutions of their inputs. This relation also allows for the modular development of methods for computing SSI. Several options can be adopted for each model. A few of them were tested by Oumbe et al. (2010) in a preliminary study.

In case of clear-sky, the equation is

Ic = Ic(albedo=0) f(albedo) (9.2)

The possibility of considering several spatial and temporal resolutions is of practical importance in MACC. MACC outcomes should be daily values or every 3 h but by no means every hour. The expected spatial resolution should be in the range 50 km - 200 km. The cloud properties can be derived from the processing of Meteosat images as done by DLR with the adapted APOLLO chain. Such products will be available every 15 min for each Meteosat pixel. Therefore, having two different modules: clear-sky and cloudy atmospheres, eases the burden of coping with these differences in resolution and availability.

A new fast clear-sky model called McClear was developed to estimate the downwelling shortwave direct and global irradiances received at ground level under clear skies. McClear implements a fully physical modelling replacing empirical relations or simpler models used before. It exploits the recent results on aerosol properties, and total column content in water vapour and ozone produced by the MACC project. It accurately reproduces the irradiance computed by the libRadtran reference radiative transfer model with a computational speed approximately 10^5 times greater by adopting the abaci, or look-up tables, approach combined with interpolation functions. It is therefore suited for geostationary satellite retrievals or numerical weather prediction schemes with many pixels or grid points, respectively. McClear irradiances were compared to 1 min measurements made in clear-sky conditions in several stations within the Baseline Surface Radiation Network in various climates. For global, respectively direct, irradiance, the correlation coefficient ranges between 0.95 and 0.99, resp. 0.86 and 0.99. The bias is comprised between -14 and 25 W m^{-2}, resp. -49 and +33 W m^{-2}. The RMSE ranges between 20 W m^{-2} (3% of the mean observed irradiance) and 36 W m^{-2} (5%), resp. 33 W m^{-2} (5%) and 64 W m^{-2}2 (10%). These results are better than those from the clear-sky model used in the HelioClim-3 service. The study demonstrates the quality of the McClear model combined with MACC products, and indirectly the quality of the aerosol properties modelled by the MACC reanalysis

Figure 9.1 displays a schematic view of the Heliosat-4 method. The method will exploit the McClear model. Radiative transfer in clouds is modelled by the Delta-Eddington approximation.

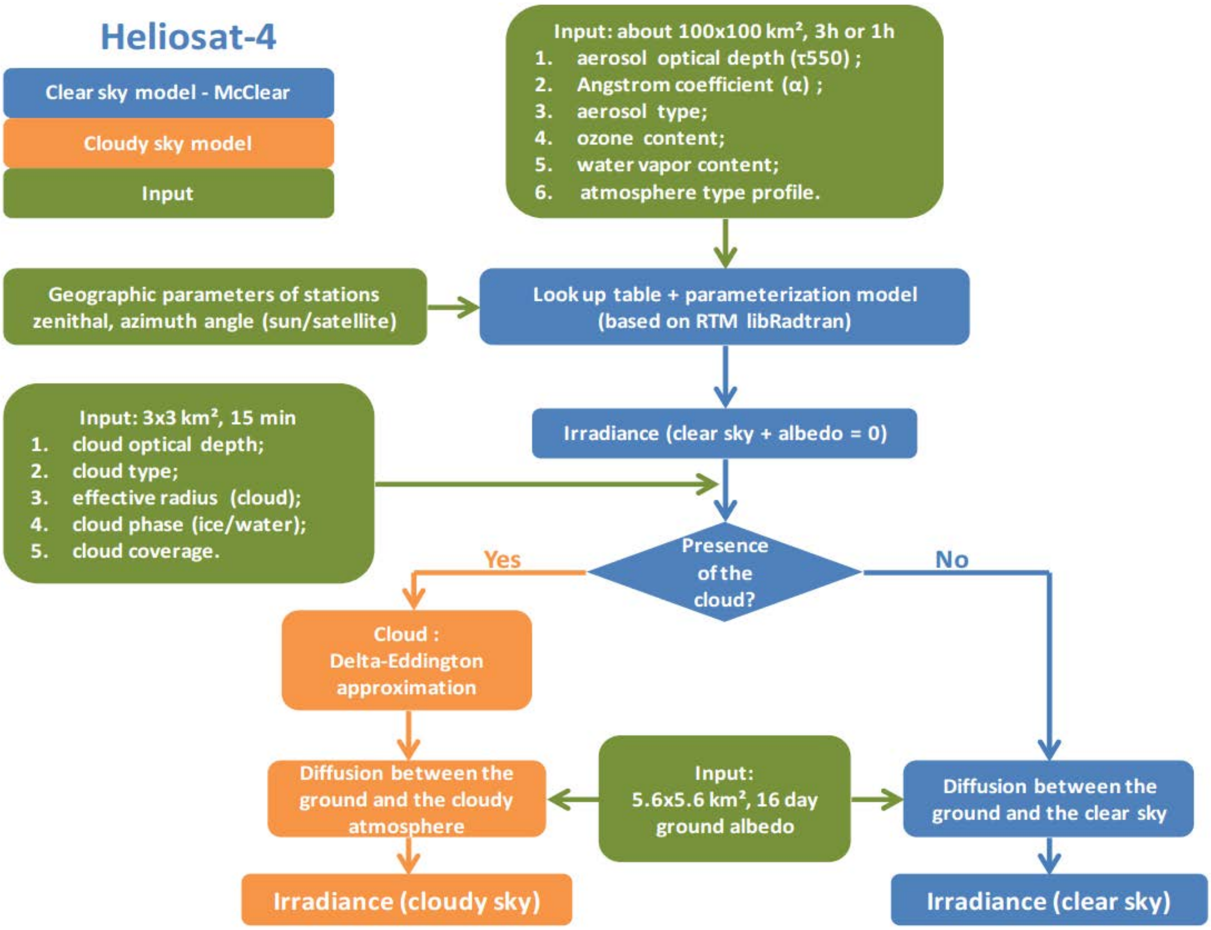


Figure 9.1. Schematic view of the Heliosat-4 method

10. Overview of the workflow in the HelioClim-4 chain

The adoption of the Heliosat-4 method leads to an entirely new workflow for producing irradiances. This workflow could be sketched by Figure 10.1 though it is too early to provide details. This overview does not comprise all elements as in Figures 8.2 and 8.3 for the HelioClim-3 and SOLEMI services.

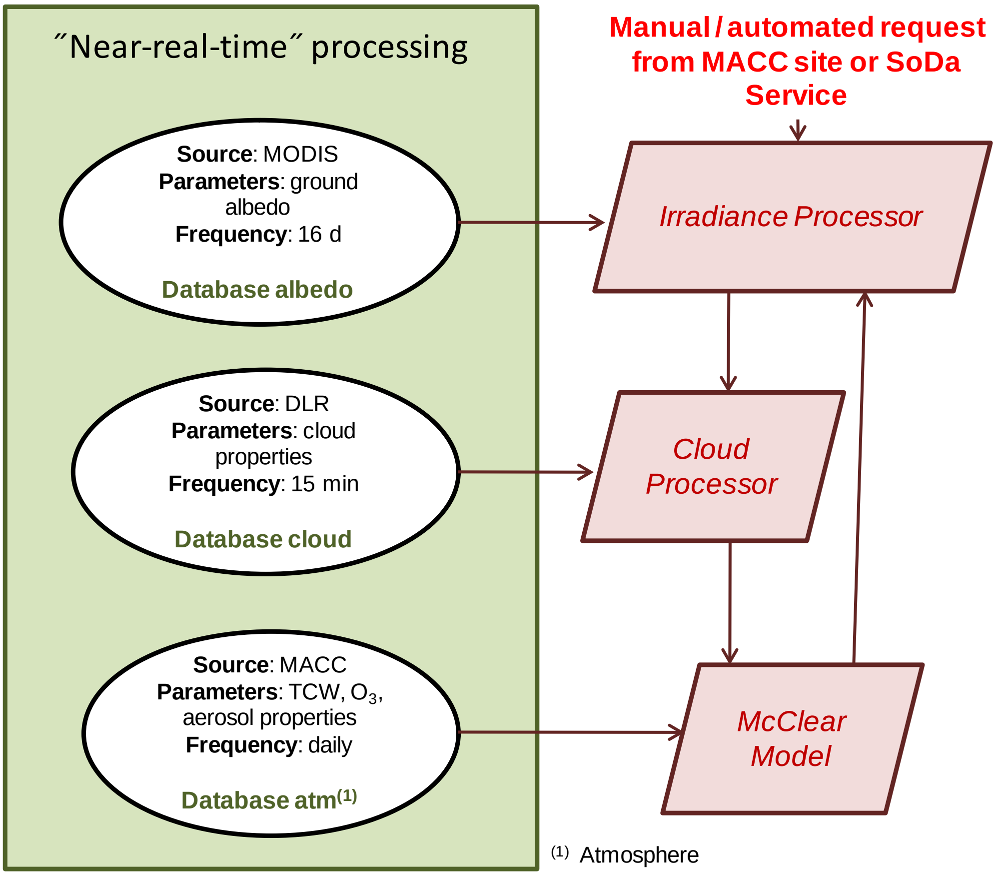


Figure 10.1. Schematic representation of the workflow of the MACC-RAD Service for computing irradiance products. The three databases are part of the HelioClim-4 database.

Several data will be received from various sources: MACC, DLR and a NASA. These data will be pre-processed and stored in several databases. These intermediate databases (or tables) will constitute the HelioClim-4 database.

The APOLLO (AVHRR Processing scheme over cLouds, Land and Ocean) method is used by DLR to produce the necessary cloud properties. It is adapted to the images of the SEVIRI instrument of Meteosat Second Generation. Ozone and water vapour quantities are retrieved from MACC products. MODIS (MODerate-resolution Imaging Spectroradiometer) BRDF/Albedo Snow-free Quality product (MCD43C2) is used for the ground albedo. The daily mean values of atmospheric column content in ozone and water vapour and aerosol properties are provided by MACC and used in the McClear clear sky model. Information about the input data is summarized in Table 10.1.

Variable	Data sources	Temporal resolution	Spatial resolution
Aerosols properties and type	MACC (EU FP7) / MATCH (DLR)	3 h / 1 h	1.125° / 1.9°
Cloud properties and type	APOLLO (DLR)	15 min	3 to 10 km
Total column content in ozone	MACC (EU FP7)	3 h	1.125°
Total column content in water vapour	MACC (EU FP7)	3 h	1.125°
Ground albedo	MODIS (NASA)	16 d	5.6 km

Table 10.1. Input data to the Heliosat-4 method

The delivery deadline from each source is still unknown. It could amount to several days. Therefore, the term 'near-real-time' has to interpreted as within 24 hours (preferably) or longer. Operations are triggered by successfully received input data sets.

Currently, the concept to produce products is that the processing is made on-request (on-the-fly). When a request is made for a time-series for a given site and a given period, the irradiance processor will invoke the McClear model with inputs extracted from the database "atm". Extinction coefficients are then obtained by the cloud processor using inputs from the database "cloud". Finally, the contribution to the diffuse irradiance due to the ground albedo is computed.

Tests have been made that demonstrate that such an on-the-fly processing is sustainable in the perspective of the MACC-RAD Service. Nevertheless, it is likely that databases of irradiance will be pre-computed, at least to best answer to request for maps of daily or monthly values of irradiance.

PART D.

QUALITY CONTROL AND VALIDATION OF PRODUCTS

11. Overview of the quality control / validation procedures

Quality control / validation procedures span over many aspects of the whole system.

Several types of quality control will be performed in the HelioClim-4 workflow:

- checking the smooth running of the product generation workflow in compliance with specifications,
- checking the quality of the inputs to the Heliosat-4 method,
- benchmarking the HelioClim-4 irradiance products against high quality measurements made in ground stations, in order to provide estimates of the quality of the retrieved irradiance,
- monitoring the consistency of this quality of product, and detecting possible trends
- checking the smooth running of the service.

Figure 11.1 is a schematic overview of the quality control (QC) / validation procedures. It is based on the graph depicting the Heliosat-4 method. The QC / validation procedures are in red in this graph. These procedures are not yet implemented and are still under development. Each procedure is described in the following Chapters, if not stated otherwise.

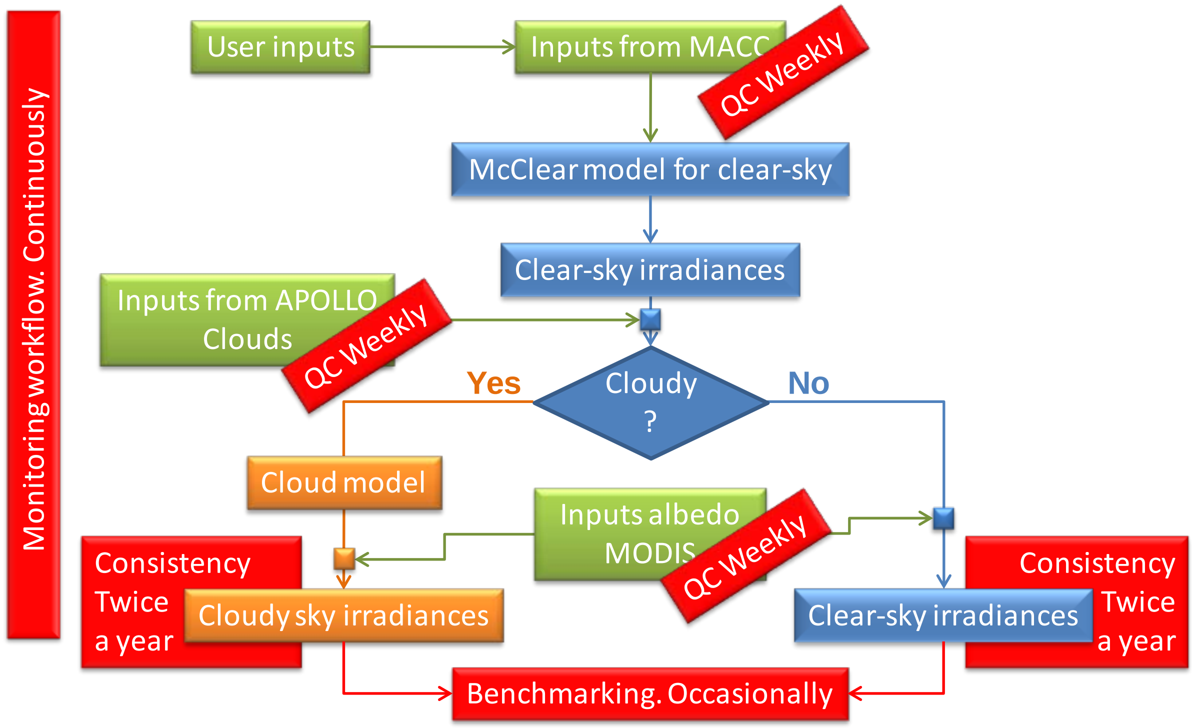


Figure 11.1. Schematic overview of the quality control (QC) / validation procedures

On the left in the graph, is the box for a continuous monitoring of the workflow. Similarly to the HelioClim-3 and SOLEMI workflows, control points will be set up in the workflow for a constant monitoring the smooth running of the processing chain. For example, one will check that inputs are available, that the database is updated as planned... In case of failure, detailed reporting will be made to the management, written in a log file and sent by e-mail as well. A kind of dashboard will be set up to perform a visual monitoring of several elements. This may include a gross check of the computed irradiance. This point is not detailed here.

Inputs to the Heliosat-4 method originate from several sources and have been quality-controlled by their providers. The role of the QC procedures linked to inputs (red boxes on green boxes) is to detect any gross problem due to transfer and archiving of data and to offer support to the analysis in case of the detection of a possible drawback of the Heliosat-4 method or the HelioClim-4 workflow. The procedures may be automated and may be invoked weekly.

Comparisons between irradiance products and measurements performed at ground by well-calibrated instruments will be made occasionally (lower red box in the graph). This will not be systematic as access to such data is limited. The benchmarking procedure will be performed manually by skilled researchers. It will consist in the comparison of HelioClim-4 retrievals and McClear estimates to measurements for several sites in the field-of-view of the Meteosat satellites, for the HelioClim-4 retrievals, and worldwide for McClear estimates. The measurements serve as a reference. The protocol for benchmarking and the performance indicators are described in Chapter 13.

The results of such comparisons are used in order to:

- detect possible flaws in the method or processing workflow, and correct them,
- document the uncertainty in the retrieval and to display it in the Web site,
- improve *a posteriori* correction procedures if possible in order to improve the quality,
- establish then improve the model of uncertainty that provides for any instant the plausible uncertainty level of the SSI, given known explanatory variables.

Meanwhile, automatic quality control procedures can be established that monitor the consistency of the quality of the retrieved irradiance (red box labelled "consistency"). The same protocol and performance indicators than for the benchmarking procedure can be used but with a different aim. The aim is now to check that the retrieved irradiance is in line with expectations. The control of the consistency will be performed twice a year.

This control of the consistency can be illustrated by an example. Assume that for a given site, or for a much larger area, one has computed the monthly means of irradiance for the month of January using a period spanning several years. This could be done by using either a long term time-series of high quality measurements, or the HelioClim-3 or SOLEMI databases for areas of proven quality. The HelioClim-4 irradiances can be averaged for a given month of January, e.g. January 2014, and compared to the climatology. If the discrepancy is greater than what is expected given the inter-annual variability and the known quality of HelioClim-4, then a warning will be issued. A manual examination will then be performed by a researcher or a trained engineer to confirm the alarm and the magnitude of the discrepancy with respect to the expected range of variation. Further actions will be decided to make the necessary and appropriate corrections.

12. Quality control procedure of the inputs to Heliosat-4

Inputs to the Heliosat-4 method have been described previously. They originate from several sources. Their producers have themselves established quality control procedures. The scope of the procedure presented here is not to question the intrinsic quality of the inputs but to offer support to the analysis in case of the detection of a possible drawback of the Heliosat-4 method or the HelioClim-4 workflow.

The position of the Sun in the sky is an important factor in the retrieval of the SSI. The ESRA algorithm in place in the HelioClim-3 and SOLEMI workflows for computing this position has been replaced in HelioClim-4 by the SG2 algorithm (Blanc, Wald 2012) which is much more accurate. Given the number of tests done on this algorithm and its proven accuracy, no defect is expected from the computation of the Sun position, and no automatic control will be set up.

12.1. Inputs from MACC for clear-sky

The inputs to the McClear clear sky model are the MACC products: aerosol optical depth, Angstrom coefficient, aerosol type, ozone content, water vapour content and the atmosphere type profile.

The MACC products spans from 1st January 2004 up to now. Maps of the statistical distribution of the values of each parameter: observed maximum, observed minimum, P95 and P5, will be computed for the period 2004-2012 (9 years). Such maps will be constructed for each month with the daily values for ozone and water vapour, and 3 h values for the aerosols.

An automatic routine will be launched say every few days that will compare the newly collected MACC values to these reference maps and will result in binary maps: alarm / no alarm values. Warning will be issued to the management. Additional analyses will be possibly performed in case of alarm. In addition, these maps will support the analyses in case of detection of a possible defect in the retrieved SSI.

The benefit of using this kind of maps is the possibility of detecting extended areas with possible malfunction of the workflow, in addition to the temporal information of this malfunction.

12.2. Albedo from MODIS

Ground albedo is also an input and originates from the MODIS database. Control of the quality of this input is not yet established as the decision to take the MODIS product every week or to use averaged values has not been taken.

In the latter case, there will be no control of this input.

12.3. Cloud parameters

Cloud parameters are retrieved from APOLLO - cloud optical depth, cloud type, and cloud coverage are used.

Quality control for these inputs has not yet been established. Efforts are being made by the DLR to create cloud products for the past years. If the number of years with available data is sufficient, the quality control will be based on the same statistical approach than for the MACC products.

13. Procedures and measures for benchmarking the retrieved irradiances with reference ground measurements

Comparisons between irradiance products and measurements performed at ground by well-calibrated instruments will be made as often as possible. This Chapter describes the procedures and the measures used in these comparisons.

Measurements from well-known high quality operation and management organisation are not numerous. It is foreseen to use the BSRN network measurements and GAW (Global Atmospheric Watch) system, established by the WMO. Measurements are available in an irregular schedule but enough to use them several times per year.

The benchmarking procedure is manual and consist in the comparison of HelioClim-4 retrievals and McClear estimates to measurements for several sites in the field-of-view of the Meteosat satellites, for the HelioClim-4 retrievals, and worldwide for McClear estimates.

These systematic comparisons will help in the detection of trends due to the operational chain and not attributable to natural changes.

High quality measurements of irradiance are not easily available in number, in due time and in an automated manner. Therefore, such comparisons cannot be performed automatically every month or so. For example, within the same network of stations, e.g. BSRN or WMO, data may be available within a few weeks for several stations and within months or years for others. The comparison with ground stations is therefore a manual exercise to be performed by trained researchers or engineers.

13.1. Principles. Comparisons. Limitations

The usual way of assessing the quality of retrievals of SSI derived from satellite images is to compare these SSI to coincident measurements performed at ground level. The typical relative accuracy of SSI measured in the global meteorological network is 3 to 5% in terms of root mean square error. Therefore, the ground measurements can be seen as an accurate reference against which one may compare the SSI derived from satellite. The comparison is made by computing the difference between the two sets of measurements and analysing statistical quantities such as the bias or the root mean square error.

However, the actual situation is not that simple. Several limitations exist that make the assessment of the quality of retrieved irradiances a very difficult task.

The first limitation is the quality of the ground measurements. Well-maintained stations are rare. Data are often questionable. They should undergo extensive procedures for checking quality. Such procedures are often not enough and a final check must be performed by a trained meteorologist to discard suspicious data. For example, a series of ground-measured data sets was produced during the MESoR project[2], to serve as a reference for the benchmarking of any solar radiation product device from satellite data (Hoyer-Klick et al., 2008).

Even in case of accurate measurements, one often encounters a problem of time system. The time system for acquisition may be universal time, mean solar time, true solar time or local time. However, when stored in a database, there is a conversion in another time system, e.g., universal time. There is consequently a change of original values due to a

2 Project MESoR (Management and Exploitation of Solar Resource Knowledge) funded by the European Commission, from 2007 to 2009 , http://www.mesor.org/

resampling in time. This resampling can be done using various techniques, usually unspecified. In any case, it is not possible to return to the original values and there is a systematic shift of a fraction of an hour between the two sets of measurements. This leads to an additional difference.

Moreover, the networks do not always follow the existing standard for defining hourly data. This standard is defined by the WMO (WMO 1981): the time assigned to a data corresponds to the end of the measurement period. For example, a hourly data assigned to 11:00 UT has been measured between 10:00 and 11:00 UT. In several cases, the time associated to a measurement represents the beginning of the period, or the middle of the period, or any instant within the period. Again, the comparison between the two sets requires the resampling of one of the sets at the expenses of decay in quality.

The limitations expressed above due to the time do not hold if one deals with daily, monthly or yearly averages or sums of SSI.

A severe limitation is due to the large differences in principles of measurements. Single point and temporally integrated data (ground measurements) are compared to spatially integrated and instantaneous data (satellite estimates). An assumption of ergodicity (e.g. here equivalence between the temporal and spatial averages) is usually made. This assumption is correct only if the field is spatially homogeneous over an area much larger than a pixel. This is generally false when a significant physiographic feature is present. Other local effects such as reflections on the surrounding slopes or the shadows of clouds may add to the difficulty in comparison.

Perez et al. (1997), Zelenka et al. (1999) have observed the local variability of the SSI using measurements made by well-calibrated ground stations close to each other. They found that the variability itself is highly variable from one region to another. Nevertheless, they demonstrate that this variability cannot be ignored. Expressing this variability as the ratio of the variance relative to the mean value over the area, they found typical variability in hourly irradiances of 17 % for an area of 10 km in radius. This means that within a 10 x 10 km² area, irradiances measured by a series of similar inter-calibrated sensors would exhibit the same mean value but would differ from hour-to-hour, with a relative variance equal to 17 %. Therefore, observing a relative difference hour-to-hour of 17% between a single pyranometer located in a pixel cannot mean that the satellite-derived irradiances are of bad quality.

The relative variability increases as the surface of the area increases. For example, it typically reaches 25% for a radius of 30 km. It decreases as the time integration increases. For example, it is down to 10% for daily values and a radius of 10 km.

Zelenka et al. (1999) analyzed the actual accuracy of satellite estimations of hourly SSI. They suggest that for a relative deviation of 23% (root mean square error) between ground measurements and satellite estimations, only half of it is due to the estimation method itself. The difference comes from:

- error on the measurements provided by the pyranometer (3 to 5%);
- error due to the spatial variability of solar radiation within the pixel (5 to 8%)
- error due to spatial and temporal heterogeneity of the compared data, e.g. assuming ergodicity (3 to 5%) as discussed above.

13.2. Measures of performance

Guidelines of the benchmarking procedures of solar radiation products derived from satellite data, have been proposed by the MESoR project, in order to measure the quality of these products with a common scheme (Beyer et al., 2008). The quality of a product is defined by various statistical quantities that measure the discrepancies between products and ground data considered as a reference.

These different measures are described below. They are sensitive to different properties in the products and by this, allow for an assessment of the uncertainty specific to a given application, and a ranking of the quality of different products. We present here the benchmarking procedure for SSI data.

13.2.1. Usual measures

Following the ISO standard (1995), the MESoR project recommends to compute the difference: modelled - measured, for each pair of values, and to summarize these differences by the bias (also improperly termed mean bias, MB), the root mean squared difference (RMSD) and the correlation coefficient. The bias *MB* is defined by:

$$MB = \frac{1}{n}\sum_{i=1}^{n}\left(x_e(i) - x_m(i)\right) = \overline{x_e(i)} - \overline{x_m(i)} \tag{12.1}$$

n : number of data pairs
$x_e(i)$ = modeled data
$x_m(i)$ = measured data

and the RMSD by:

$$RMSD = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_e(i) - x_m(i)\right)^2} \; . \tag{12.2}$$

The corresponding relative differences *rMB* and *rRMSD* are computed by dividing the bias and the RMSD by the mean of the measurements:

$$rMB = \frac{MB}{\overline{x_m}} \tag{12.3}$$

$$rRMSD = \frac{RMSD}{\overline{x_m}} \; . \tag{12.4}$$

The standard deviation σ can be calculated from *MB* and *RMSD*:

$$\sigma = \sqrt{RMSD^2 - MB^2} \; . \tag{12.5}$$

The correlation coefficient *CC* is computed as follows:

$$CC = \frac{\sum_{i=1}^{n}\left(x_e(i) - \overline{x_e}\right)\cdot\left(x_m(i) - \overline{x_m}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_e(i) - \overline{x_e}\right)^2 \cdot \sum_{i=1}^{n}\left(x_m(i) - \overline{x_m}\right)^2}} \; . \tag{12.6}$$

13.2.2. Quality measures based on the Kolmogorov-Smirnov Test

However, these parameters are often insufficient to establish a complete, coherent comparison for benchmarking. Additional measures can be introduced that quantify the discrepancies between the cumulative distribution functions (CDFs). A comprehensive approach to an analysis of the deviations of measured and modelled CDFs is described by Espinar et al. (2008). It is based on the Kolmogorov-Smirnov (KS) test and defines new parameters to quantify the similarity of the two CDFs. Although there are several statistical tests and ways of evaluating the goodness of a model, the KS test has the advantage of making no assumption about the data distribution, and is thus a non-parametric, distribution-free test.

The KS test tries to determine if two data sets differ significantly. The test consists of comparing the distribution of a dataset to a reference distribution. This can be done by converting the list of the *N* data points to an unbiased estimator $S(x_i)$ of the CDF. The KS statistic *D* is defined as the maximum value of the absolute difference between the two CDFs:

$$D = \max |S(x_i) - R(x_i)| \tag{12.7}$$

where $R(x_i)$ is the CDF of the reference data set. If the *D* statistic is lower than the threshold value V_C, the null hypothesis that the two data sets come from the same distribution cannot be rejected. The critical value depends on *N* and is calculated for a 99% level of confidence as:

$$V_c = \frac{1.63}{\sqrt{N}}, \quad N \geq 35 \tag{12.8}$$

This test detects smaller deviations in cumulative distributions than the χ^2 test does. However, instead of using the original one, an extended KS test is used, in which the distances between the CDFs are calculated over the whole range of the variable *x*, i.e. the solar radiation. A discretisation in *m* levels is applied here. In the project MESoR, setting *m* to 100 was found a reasonable choice. Greater orders of magnitude for *m* are not recommended since it implies more computational cost for no improvement in the accuracy of the result. The interval distance *p* is defined as:

$$p = \frac{x_{\max} - x_{\min}}{m} \;,\; m=100 \tag{12.9}$$

where x_{max} and x_{min} are the extreme values of the independent variable. Then, the distances between the CDFs are defined, for each interval, as:

$$D_n = |S(x_j) - R(x_j)| \;,\; x_j \in [x_{\min} + (n-1)p, x_{\min} + np] \tag{12.10}$$

The representation of the values D_n, along with the critical value, gives the complete testing behaviour of the CDF with respect to the reference over the whole range. Thus, the extended KS test is very useful for model response assessments. However, although application of the KS test contributes valuable information, it only materializes in the acceptance or rejection of the null hypothesis. In the next sections new parameters are proposed, which, based on the estimation of the distance between the two CDFs for the sets compared, define quantitative measures that can be used to rank models.

13.2.3. Kolmogorov – Smirnov test Integral, parameter KSI

The KSI parameter (Kolmogorov-Smirnov test Integral) is defined as the integral of the area between the CDFs for the two sets. The unit of this index is the same for the corresponding magnitude, the value of which depends on it. As an example, the leftmost graph in Figure 12.1 shows the CDF (measured) for the data set of daily irradiation measured at the station in Arkona (Germany) over a period of 9 years (1995-2003) and the CDF of a corresponding satellite-derived data. In the rightmost plot the distances D_n between the two CDFs are displayed. The black dotted line represents the critical limit, V_c, calculated for the number of available data.

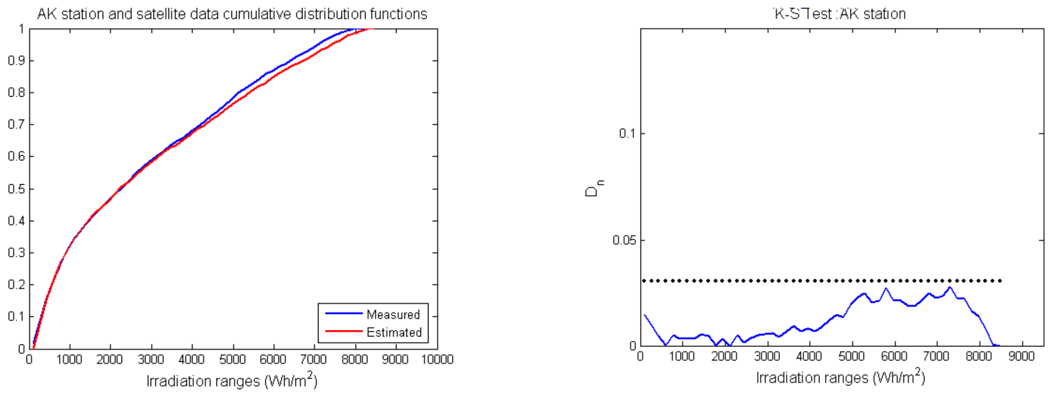


Figure 12.1. Plot of cumulative distribution functions for measured and modelled sets of daily irradiation and the distances between them, for the Arkona station (Germany). Courtesy of Annette Hammer, University of Oldenburg

The KSI is defined as the integral:

$$KSI = \int_{x\,\min}^{x\,\max} D_n d(x) \qquad (12.11)$$

As D_n is a discrete variable and the number of integration intervals is identical in all cases, trapezoidal integration is possible over the whole range of the independent variable *x.* A percentage of KSI is obtained by normalizing the critical area, $a_{critical}$:

$$KSI\% = \frac{\int_{x\,\min}^{x\,\max} D_n dx}{a_{critical}} * 100 \qquad (12.12)$$

where $a_{critical}$ is calculated as:

$$a_{critical} = V_c * (x_{\max} - x_{\min}) \qquad (12.13)$$

where V_c is the critical value for the level of confidence selected and (x_{max}, x_{min}) are the extreme values of the independent variable. Normalization to the critical area enables the comparison of different KSI values from different tests. The minimum value of the KSI index is zero, in which case, it can be said that the CDFs of the two sets are the same.

13.2.4. Parameter OVER_99

In the definition of this new parameter, the critical limit V_c is applied from the original KS test, calculated according to the number of data *N* in the set. For its estimation, only the areas of those distances between the CDF that exceed the critical limit are calculated.

OVER_99 refers to the fact that V_c is based on a 99% confidence level. Basically, other confidence levels can be chosen. To calculate the OVER_99, the auxiliary vector for the values that exceed the critical value *aux* is generated. If any of the components does not exceed the critical value, its corresponding component in the auxiliary vector is zero.

$$aux = \begin{cases} D_n - V_c & if\ D_n > V_c \\ 0 & if\ D_n \leq V_c \end{cases} \qquad (12.14)$$

The OVER_99 and OVER%_99 parameters are then calculated as the trapezoidal integral of that auxiliary vector and its corresponding normalization to the critical area:

$$Over\_99 = \int_{x\,\min}^{x\,\max} aux\ dx \qquad (12.15)$$

$$Over\%\_99 = \frac{\int_{x\,\min}^{x\,\max} aux\ dx}{a_{critical}} * 100 \qquad (12.16)$$

where $a_{critical}$ is the same as that calculated in the *KSI* and *KSI%*.

Like the KSI index, the OVER_99 has the same unit as the variable that is evaluated. This parameter enables a value to be marked that generates two types of results: the sets that behave statistically the same and those that do not. This is an advantage compared to the results supplied by the classical parameters RMSD and MB, which do not provide the possibility of making this differentiation.

13.2.5. Comparability of results

A yet problematic feature of the proposed tests is the dependency of the critical value on the number of samples. The more the samples used in the analysis, the lower the critical value. It can be explained by an example. Assume that the CDFs of two hourly data sets are compared for one year (about 4500 values with the sun above the horizon) and that they are judged similar. If two years are taken, the critical value is $1/\sqrt{2}$ times smaller than before. If the difference between the CDFs is still the same, they might now be judged as different because of this change in critical value. It has to be evaluated if constant values of the critical value should be used for certain time steps (e.g. hourly, daily, monthly data).

13.2.6. Summary on the measures proposed

The basis for the proposal of the new parameters is the KS test, but the aim is its quantification to gain numerically comparable results from the different sets. Therefore, the parameters KSI and OVER are introduced. The KSI estimates the area between the two CDFs. The OVER is also an estimate of the area between the CDFs, but only for the parts where the critical value distance V_c is exceeded. In addition, it provides a value comparable between different data sets.

The OVER% is the only one that enables to observe a significant difference in the behaviour of the comparison. It shows whether the critical value is passed or not, so the classification has to account for this value first, classifying those stations as best that have an OVER equal to zero. That means the sets compared are statistically so similar that they could be the same one.

The differences found in the comparison of the data may in turn have different explanations. Thus, they may originate both in the behaviour of the measured data at a specific station, such as recording errors, and in the behaviour of the estimates made from the satellite images in the pixels for complex topographies.

13.3. Selection of valid data pairs for benchmarking

An important factor to achieve comparable benchmarking results is the selection of valid data pairs (modelled and measured) which are taken into account. The project MESoR recommends the use of data pairs only for which:

- the ground data has passed a QC procedure (e.g. Hoyer-Klick et al. 2008);
- measured global irradiance is greater than zero (this excludes night values and missing measurements);
- the modelled value is available.

Averages are calculated from all valid data pairs. If a subset is selected (e.g. a sun zenith angle interval) averages are calculated from all valid data pairs within this subset.

There are two extreme ways to perform the benchmarking when data from multiple stations are available. One of these ways is to concatenate all data pairs into a single series to analyse. In this way, each data pair has the same weight in the analysis. The other way is to perform the analysis for each station independently and then to compute the mean values of the different measures. In this way, each station has the same weight irrespective of its number of samples in the analysis. Both ways have their merits and drawbacks. If all stations belong to the same "radiation climate", then, it would be justified to adopt the first way. On the contrary, if the stations belong to different climates, e.g., one station in Sahel, and three stations in Germany, then, concatenating all data will give more weight on the German stations, with the risk that discrepancies specific to the Sahel station may disappear in the analysis; in this case, the second approach should be preferred. Actually, it is highly preferable to perform the analysis for each station individually and to discuss the possible discrepancies in quality measures with respect to the properties of these stations.

The same discussion may apply to time aggregation. For example, assume that daily values are available for a year but with gaps, e.g., 15 days available per month in January, February and March, all days available for the other months. Obviously, if one performs an average of all days for assessing the performance over a year, this assessment will be biased by the low number of days in winter. There are several possible solutions. It is recommended to analyse the influence of the choice of a solution.

13.4. Typical performances of the existing HelioClim-3 and SOLEMI services

The quality of the products provided by the existing HelioClim-3 and SOLEMI services has been assessed at several occasions using the MESoR – IEA procedure described previously. The results are published in articles, communications and on the Web site of the SoDa Service.

A sample of the many comparisons made for the products supplied by HelioClim 3 and SOLEMI is presented here. The example is drawn from the work made by Pierre Ineichen, University of Geneva, who made several comparisons between ground measurements and the products from HelioClim-3 and SOLEMI. This work was presented at the meeting of the MESoR project in Geneva in December 2008.

Table 12.1 presents the results for the products HelioClim-3v3 and Table 12.2 those for the products SOLEMI. These tables deal with hourly means of global irradiance, expressed in W/m². The period is the full year 2005.

HelioClim-3v3		Mean value (W m^{-2})	N	R^2	rMB (%)	rRMSD (%)	KSI%	Over%_99
Toravere	Estonia	256	3789	0.937	3	32	89	14
Nantes	France	294	4113	0.933	2	30	40	1
Vaulx-en-Velin	France	310	4066	0.966	5	22	65	1
Thessaloniki	Greece	376	3820	0.976	-4	17	54	4
Geneva	Switzerland	292	4283	0.969	2	23	39	0
Payerne	Switzerland	294	4268	0.955	-5	26	67	8
Cabauw	UK	261	3803	0.962	-1	23	70	17
Camborne	UK	249	4235	0.953	5	29	72	12
Lerwick	UK	199	3624	0.906	10	43	105	40

Table 12.1. Benchmarking of the hourly averaged global irradiance from HelioClim-3v3 for 2005. "Mean value" is the average of the ground measurements; "N" is the number of data pairs; "R²" is the squared correlation coefficient; other quantities are defined in the text.

SOLEMI		Mean value (W m^{-2})	N	R^2	rMB (%)	rRMSD (%)	KSI%	Over%_99
Toravere	Estonia	256	3789	0.948	-3	27	38	0
Nantes	France	294	4113	0.937	2	29	27	3
Vaulx-en-Velin	France	310	4066	0.950	5	26	60	17
Thessaloniki	Greece	376	3820	0.971	3	18	43	1
Geneva	Switzerland	292	4283	0.954	4	27	77	12
Payerne	Switzerland	294	4268	0.957	1	25	33	5
Cabauw	UK	261	3803	0.950	0	26	40	5
Camborne	UK	249	4235	0.950	2	28	32	8
Lerwick	UK	199	3624	0.935	3	33	70	20

Table 12.2. Benchmarking of the hourly averaged global irradiance from SOLEMI for 2005. "Mean value" is the average of the ground measurements; "N" is the number of data pairs; "R²" is the squared correlation coefficient; other quantities are defined in the text.

It is not the scope of this document to discuss specifically these tables and to compare the results. Nevertheless, the reader may note that the performances are very good in view of what is available currently. These performances are typical of those observed for the HelioClim-3v3 and SOLEMI products in other occasions.

The performances increase as the aggregation period increases, except for the bias of course.

The reader may take note of the variability of the quality parameters for the same product for sites belonging to the same geographical area. For example, in Table 12.1, one finds a relative bias (*rMB*) of 2% for Nantes and 5% for Vaulx-en-Velin. On the contrary, *rRMSD* is larger for Nantes than for Vaulx-en-Velin: 30% versus 22%. This can also be observed in UK stations. Even in geographical areas *a priori* fairly similar with respect to radiation climate, one cannot transpose accurately the performances observed at one site to another one. The same observation can be made when comparing benchmarks for the same site but different years. This does not help customers unfortunately but it is better to be aware of it.

14. Monitoring the consistency of the quality of retrievals

14.1. Concept

Once the quality of the retrieval is known, one may establish expectations for the estimates to come and therefore establish a monitoring of the consistency of the quality of the retrieved irradiance. The same protocol and performance indicators than for the benchmarking procedure can be used but with a different aim. The aim is now to check that the retrieved irradiance is in line with expectations. The control of the consistency will be performed twice a year. This would help in detecting unwanted trends in the SSI retrieval.

The control of the consistency of the quality consists in a series of preparatory steps:

- identifying sites or areas and associated periods during which these sites / areas do not exhibit complex features in irradiance in both space and time such as those created by a complex orography,
- then collecting estimates of irradiance that can be collected automatically, and which may originate from measurements of medium quality or derived from satellite data or meteorological re-analyses,
- computing the discrepancies between the HelioClim-4 retrievals and these estimates,
- setting up a site-specific model of discrepancy as a function of known explanatory variables such as the solar zenithal angle, season or the temporal variability of the SSI. This model should be established with a sufficient great number of data.

In the automatic mode, the control procedure collects the estimates, computes the discrepancies between them and the HelioClim-4 retrieval, and compares these discrepancies to the model of discrepancy. If the distance between the estimates and the HelioClim-4 products does not obey the model, there is a suspicion that the HelioClim-4 products may present a drawback and an alarm is triggered. A manual examination will then be performed by a researcher or a trained engineer to confirm the alarm and the magnitude of the discrepancy with respect to the expected range of variation. Further actions will be decided to make the necessary and appropriate corrections.

To perform the control of the consistency of the quality, operations should be established to collect in an automatic way re-analyses values from e.g. MERRA and ground measurements from selected sites or from the World Radiation Data Center. Quality control procedures must be also implemented that apply onto these measurements and re-analyses.

The analysis of a possible drawback will be supported by an automatic analysis of the inputs from MACC, averaged over the same period, say January 2014, with respect to the mean values of these inputs averaged for the month of January over several years. A large difference may explain the discrepancy detected by the control of consistency and may rule out any problem in the Heliosat-4 method and the HelioClim-4 workflow.

Several quality control procedures are being studied for the time being. They will be tested with the existing HelioClim-3 database. Those yielding satisfactory results will be retained. Then, they will be tested with actual HelioClim-4 data and improved during the several months of experimental operations of HelioClim-4 planned from October 2013 to July 2014. Those yielding satisfactory results will be implemented in the HelioClim-4 workflow once the MACC-II project completed.

14.2. Oceanic areas with low variability in SSI

A region exhibiting a low variability in SSI, i.e. fairly homogeneous in space and time, is a perfect element to establish a quality control procedure. If existent and identified, the behaviour of HelioClim-4 retrieval averaged say over the last two weeks with respect to a re-analysis or HelioClim-3 should be comparable to what was observed previously.

Such regions may exist in the ocean. A loop movie of the SSI estimated from the HelioClim-1 database spanning 21 years (1985-2005) has been used to detect oceanic areas exhibiting spatially homogeneous SSI, that is, the area is all clear, or all overcast, or all covered by broken cloud fields.

Six areas were manually selected and studied in terms of their annual variability in daily means of the SSI. They are shown in Figure 14.1. The areas are named according to their geographical position or proximity to a country: Atlantic, Brazil-North, Brazil-Rio, Greece, Somalia and Angola. Each of these areas is roughly 4° x 4° in size, except Somalia which is a quarter of this size.

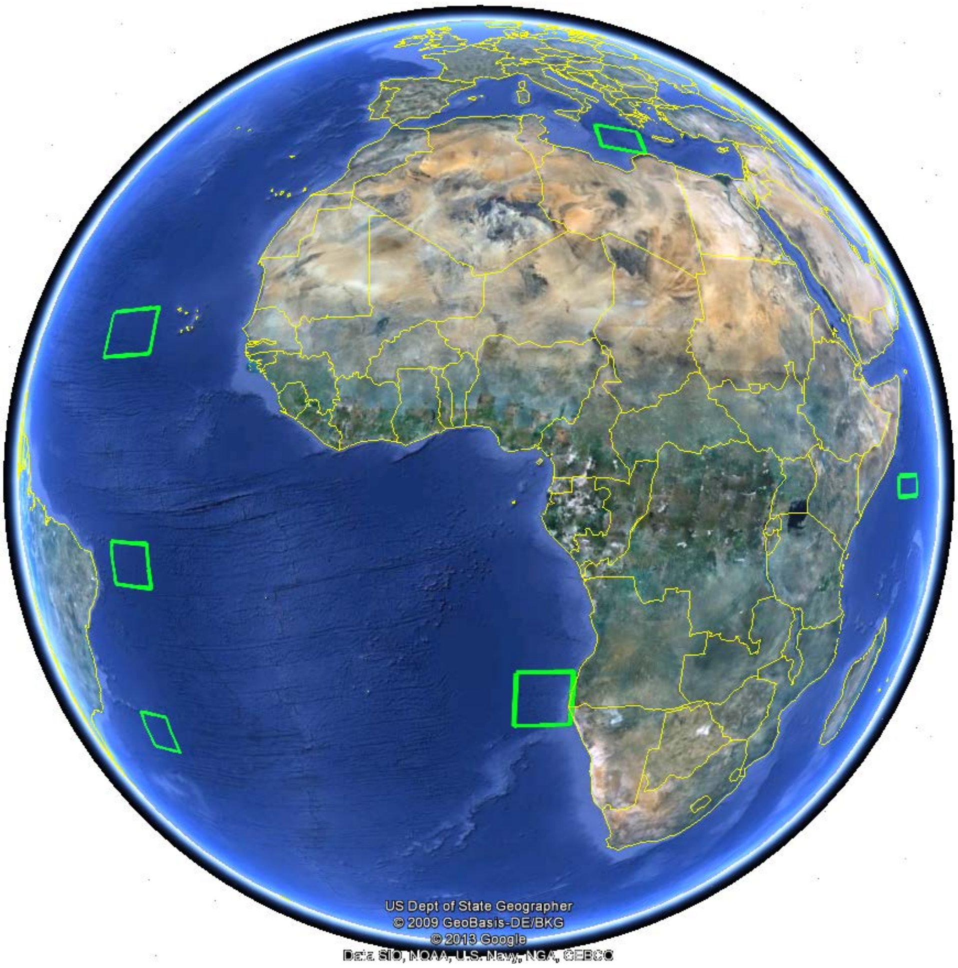


Figure 14.1. Map showing the six oceanic areas selected for the variability study. They are named according to their location or proximity to a country: Atlantic, Brazil-North, Brazil-Rio, Greece, Somalia and Angola. From Google Earth.

The NASA Global Modeling and Assimilation Office produces the Modern Era-Retrospective Analysis for Research and Applications (MERRA) re-analysis. The MERRA data set has a resolution of 0.5° x 0.65° and comprises estimates of the SSI from which the daily means of the SSI can be computed.

For each area, the daily values of SSI have been computed from MERRA for the nearly hundred grid points inside the area and for a period of 28 years, from 1985 to 2012. For each day, the spatial variability of SSI has been represented by the standard deviation for the whole area.

The statistical distribution of the standard deviation for each year observed in the area "Atlantic" is represented in Figure 14.2, in the form of boxplot as a function of the year. In the boxplot, the circle denotes the mean value and the middle line, the median. The limits of the box are the percentiles 25 (*P25*) and 75 (*P75*). The difference *P75-P25* is called the interquartile range (*Q*). The extreme lines denote the outliers, i.e. the values that are outside the range [*P25*-1.5 *Q*; *P75*+1.5 *Q*]. The smaller the interquartile range, the less variable the SSI for a year.

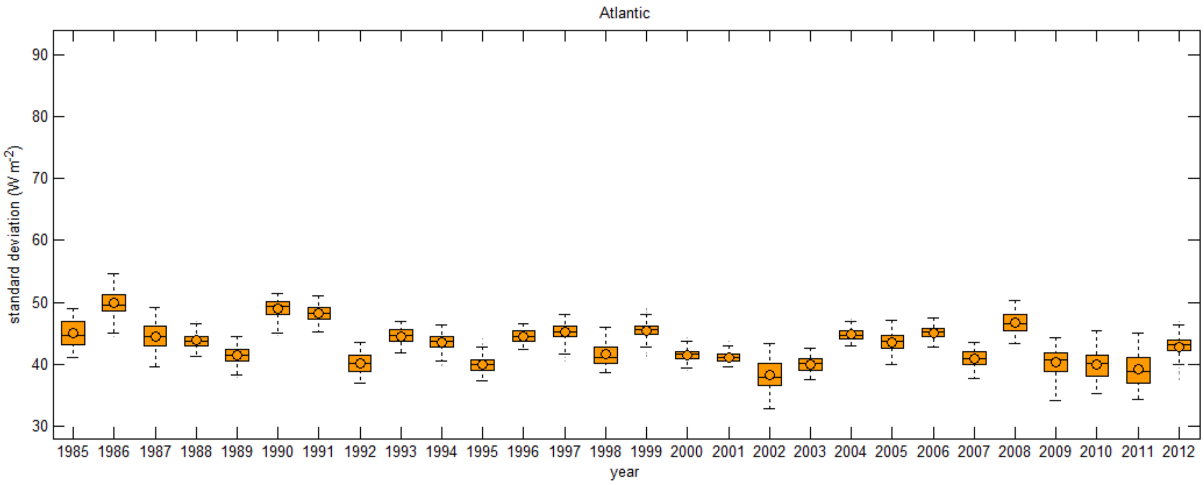


Figure 14.2. Boxplots of the statistical distribution of the spatial standard deviation of the daily SSI, for all the grid-points in the area "Atlantic". Circle: mean value; middle line: median; box limits: percentiles 0.25 and 0.75; extreme lines denote the outliers.

This graph shows that the mean value of the standard-deviation is changing from year-to-year but with a limited range. The interquartile range is fairly constant along the years and is small, approximately equal to 4 W m^{-2}.

The following graphs exhibit the boxplots for the other sites.

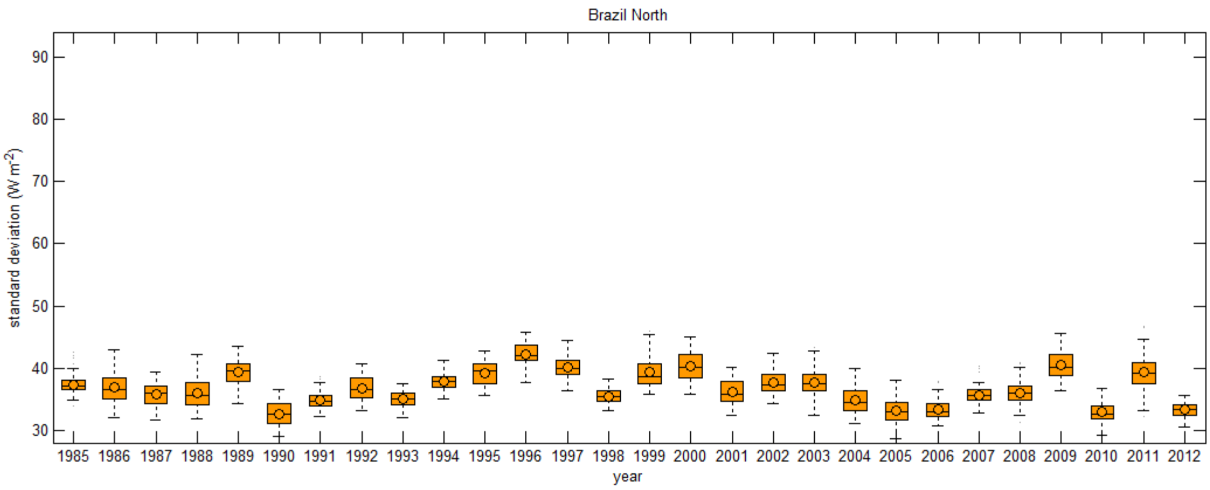


Figure 14.3. As Figure 14.2 but for "Brazil North".

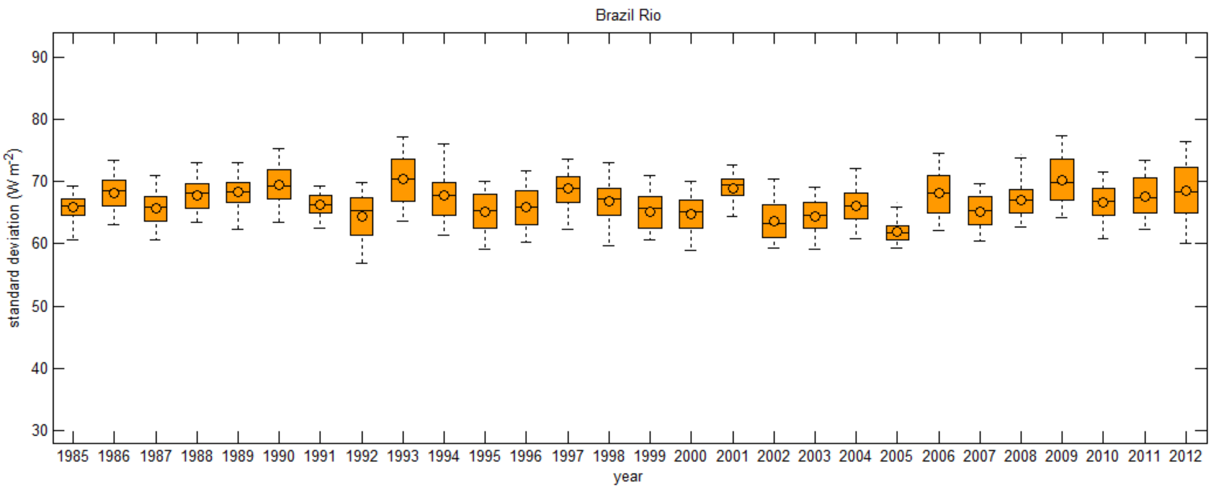


Figure 14.4. As Figure 14.2 but for "Brazil Rio".

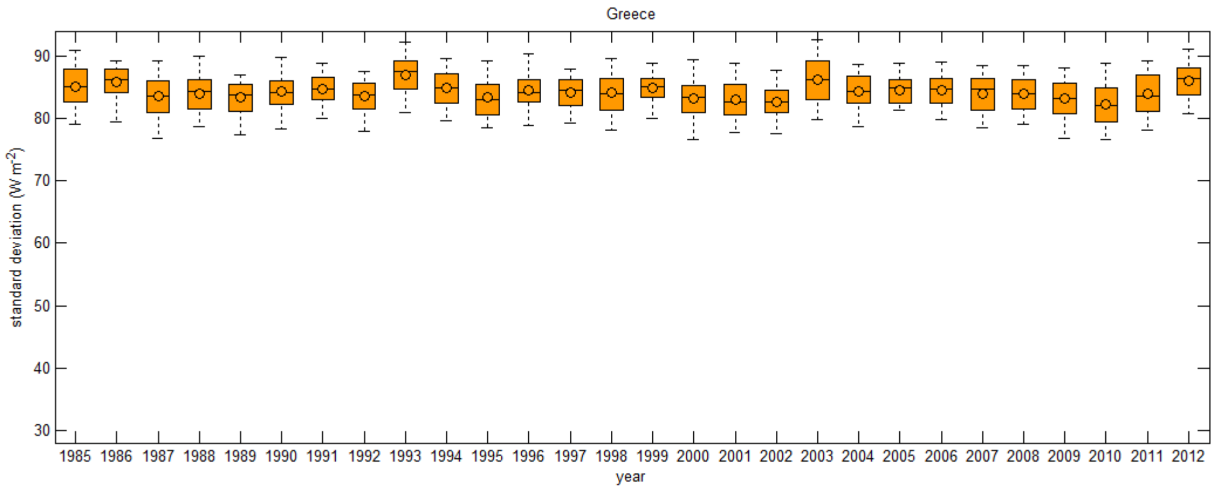


Figure 14.5. As Figure 14.2 but for "Greece".

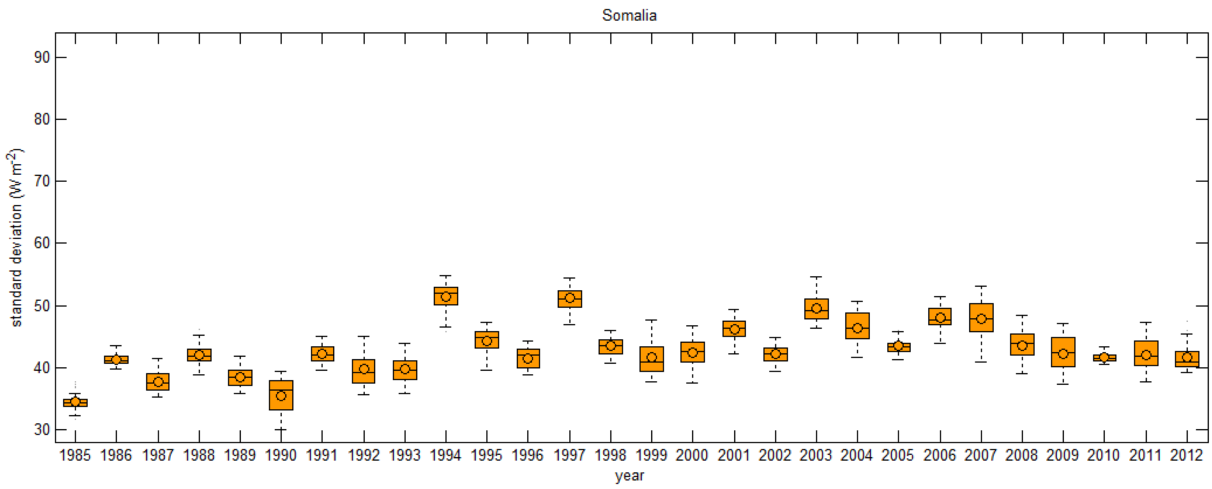


Figure 14.6. As Figure 14.2 but for "Somalia".

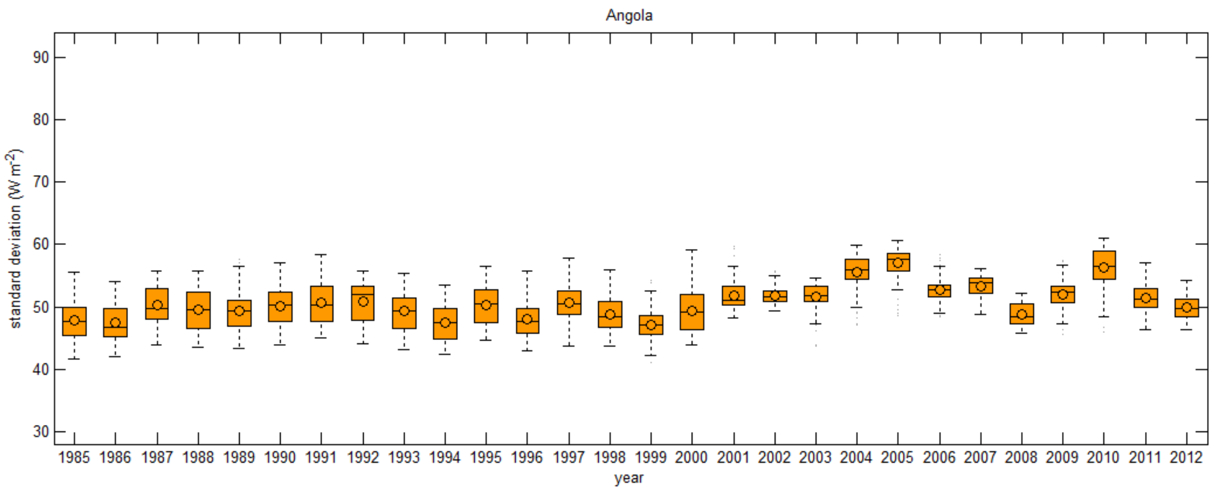


Figure 14.7. As Figure 14.2 but for "Angola".

According to these graphs, three zones: Atlantic, Brazil-North and Somalia, exhibit small interquartile ranges close to 4 W m^{-2} and constant over the 28-year period. They may be selected as homogeneous zones.

Given this homogeneity, the differences between the HelioClim-4 SSI averaged over the past month and the MERRA re-analysis or HelioClim-3 SSI should be comparable to what was observed previously.

For each zone, the daily values of SSI will be extracted from HelioClim-4 for the past month. They will be averaged over the week and the zone, yielding to the mean value m_{HC4}. The standard deviation of these daily values will be computed yielding the interquartile range: σ_{HC4}. Meanwhile, the same computations will be made with the SSI from the MERRA re-analysis and the HelioClim-3 service, yielding m_{MERRA}, m_{HC3}, σ_{MERRA} and σ_{HC3}. σ_{HC4} will be compared to σ_{MERRA} and σ_{HC3}, with the expectation that the differences between them should be small, less than a few W m^{-2}. The differences between m_{HC4} and m_{MERRA}, and m_{HC4} and m_{HC3}, will be compared to the differences computed for previous months, with the expectation that these differences should be fairly constant.

It is possible that a more robust control could be achieved by the means of the clearness index, or the clear-sky index, instead of the SSI, because the influence of the sun position will be limited. In that case, the same procedure may be repeated but using 2, 3, and 4 months for MERRA and HelioClim-3. This multi-scale comparison will help in detection of trends whose time scale is greater than 1 month.

14.3. Controlling the consistency in quality of the McClear estimates in homogeneous zones

The above-mentioned approach may be applied to controlling the consistency in quality of the McClear estimates. In that case, homogeneous zones should be identified all over the world. The comparison will be made between space and time averaged SSI, or clearness indices, or clear-sky indices provided by McClear for the past month, and the same values but averaged for this month over several years.

PART E.

DELIVERING PRODUCTS

15. Core products

The definition of the core products may evolve for various reasons. A value-added product may one day become a core product, depending on advances in science and technology. The data policies of MACC and of the GMES Atmospheric Service should be taken into account in the definition of the products. These policies include the important issue of financing and therefore of the cost of the products to customers. Combined together with others, these elements cast shades on the definition of the products. The proposed description is based on the data policy issued by the MACC project. It may be refined depending upon the GMES / Copernicus data policy.

Presently, there are three series of products in the MACC-RAD Service:

- ARCH products: the first series is an archive: 1984-2005, based on Meteosat First Generation (MFG) data. It is denoted by the letters ARCH (for Archive),
- MSG products: the second series is based on Meteosat Second Generation (MSG) series of satellites, starting from 2004 and on-going. It is denoted by the letters MSG,
- McClear products: this series describes the clear-sky SSI for the whole world, starting from 2004 and on-going. This series is the result of a model and is not based on the processing of Meteosat data.

A new series will be incorporated later as Meteosat Third Generation series of satellites will become available.

For the series currently available, core products in the MACC-RAD Service are time-series of solar radiation available at ground surface, also called surface solar downward irradiance (SSI).

There are three irradiance parameters identified as necessary by users. Table 15.1 provides a definition of these parameters that are contained in the core products.

Acronym	Definition
GHI	Total global irradiance, i.e., the SSI integrated over the whole spectrum available at ground level, on a horizontal surface
DifHI	Total diffuse irradiance, i.e., the diffuse part of the SSI integrated over the whole spectrum available at ground level, on a horizontal surface
DirHI	Total direct (beam) irradiance, i.e., the direct part of the SSI integrated over the whole spectrum available at ground level, on a horizontal surface

Table 15.1. Definition of the surface solar downward irradiance (SSI) parameters contained in the core products

15.1. Main features of the core products. Geographical and temporal coverage

Erreur ! Source du renvoi introuvable. summarizes the geographical and temporal coverage of the core products.

	Europe / Africa / Middle East / Atlantic Ocean	Central Asia / Indian Ocean	World
ARCH	1984 – 2005	1995 - 2005	-
MSG	Since 2004	No	-
McClear	-	-	Since 2004

Table 15.2. Geographical and temporal coverage of the core products

The main features of the core products are presented in Table 15.3.

	ARCH	***MSG***	***McClear***
Type of product	Time-series		
Best temporal resolution	1 h	15 min	Not applicable
Spatial resolution	Spatial resolution is the original pixel of the Meteosat image. Approx. 3 km at satellite nadir, and 5 km at mid-latitude		Not applicable
Parameters	GHI, DifHI, DirHI		
Processing	On request	Near-real-time	On request
Update of the database	End of month	End of day	2 days late

Table 15.3. Main features of the core products of the MACC-RAD Service

15.2. List of products

Within each series of products ARCH, MSG or McClear, there are several products delivered by the MACC-RAD Service. They differ by the temporal aggregation. The main features remain similar.

The period of integration is defined as the time during which the solar radiation is integrated to yield period-averaged irradiance. For example, an integration period of 1 h means that the delivered irradiance is the hourly mean of irradiance during one hour. Following meteorological standards, the time given for an irradiance value is the end of the integration period. For example, the value given for 11:00 means an integration period from 10:00 to 11:00 if the period is 1 h, or 10:30 to 11:00 if the period is 30 min.

This list of products results from request by users and is presented in Table 15.4. It may evolve as the needs of users evolve. A product contains the three irradiance parameters: GHI, DifHI, andDirHI, listed in Table 15.1.

Integration period	Product name	Product name	Product name
1 min	-	-	McClear.1min
15 min	-	MSG.15min	McClear.15min
1 h	ARCH.1h	MSG.1h	McClear.1h
1 d (1 day)	ARCH.1d	MSG.1d	McClear.1d
1 month	ARCH.1m	MSG.1m	McClear.1m
1 y (1 year)	ARCH.1y	MSG.1y	McClear.1y

Table 15.4. List of the core products

15.3. How to make a request for a MACC product

Requests will be made through an information system which is described in the next Chapter. For the time-being, there is the possibility of using the Web SoDa Service which is well-known in the solar radiation community which uses it intensively.

The Web SoDa Service is actually a precursor of the future MACC-RAD Service as discussed in Chapter 11. It already offers an access to the MSG-derived HelioClim-3 database of irradiance, discussed in Chapter 5. There is a major difference between the HelioClim-3 products and the foreseen MSG products from MACC: the access to HelioClim-3 products is made on commercial basis, while the MACC-RAD products obey the MACC data policy explained below.

From a practical point of view, and beyond the various issues related to being a recognised customer of the information system, the inputs by users needed to trigger a request for a selected MACC product will be:

- the geographical coordinates of the site of interest, or the name of this site,
- the elevation of this site above sea-level. By default, the application uses well-known digital elevation models, such as NASA-SRTM,
- the period of time: begin data, end date.

15.4. Delivery deadline

Delivery deadline is the time lag between the moment the request is made and the instant of delivery. The delivery deadline depends on the series of products: ARCH, MSG or McClear.

The ARCH products are computed on request and are not available as such on the shelf. Once the request for a product is made, DLR, the author of the products ARCH, selects the best suitable auxiliary data set for the site among global atmospheric data sets (aerosol, water vapour, ozone) from different earth observation sources and climate models. The atmospheric data is gridded to a resolution of 1°x1° and the cloud data from Meteosat has a nominal resolution of 2.5 km x 2.5 km at the sub satellite point. The geographical coordinates (latitude, longitude and height above sea level) in decimal degrees and meters have to be delivered to DLR. The product is delivered per e-mail within two weeks.

The MSG products are computed on-the-fly from the HelioClim-3 database. The request is made through the Web site of the SoDa Service (www.soda-is.com) and the answer is provided within a few minutes.

The McClear products are computed on-the-fly from the various databases of inputs to the McClear model. The request is made through the Product Catalogue of the MACC Web site or the SoDa Service (www.soda-pro.com) and the answer is provided within a few minutes.

The description above stands for the current situation. It will evolve as the situation evolves from interim to final solutions.

15.5. Data policy. Conditions of use

The data policy of MACC products is an open and free access to data, in compliance with the GEOSS (Global Earth Observation System of Systems) Data Sharing Principles. In addition, the data policy of Eumetsat shall govern the use of Meteosat data to produce the MACC-RAD products. A licence agreement comprising the licence terms will be presented to users at their first registration to the information system.

The principles governing the conditions of use are that the customer assumes all the responsibilities for the use of the provided data, especially concerning:

- the fitness-for-purpose of the data to the needs of the customer,
- the use of the data,
- the qualification and expertise of the staff of the customer regarding such data.

The MACC-RAD product is supplied “as is”. The author and provider disclaim all warranties, expressed or implied, including, without limitation, the warranties or merchantability and of fitness for any purpose. The author and provider shall in no event be liable for indirect damages such as loss of profits, loss of markets, loss of clientele, commercial injury or disturbances, brand-image impairment or unfair-competition suits, or other consequential, incidental or special damages in connection with the MACC-RAD products.

The customer acquires a license for the use of the MACC-RAD product. It does not bear any rights upon the products themselves.

For the time-being, the data policy for both ARCH and MSG products is on a commercial basis. The French company Transvalor, a subsidiary of Armines, is taking care of the commercialisation of the database HelioClim-3 through the SoDa Service.

15.6. Format of products

15.6.1. Data formats

Two formats are provided for the currently available products in ARCH or MSG: CSV (comma-separated values) and HTML.

The HTML format can easily be ingested by copy-and-paste tools into spreadsheet or text files. Text files are easy to handle by tools such as Microsoft or OpenOffice suites, Matlab or proprietary applications written in any language (e.g., C, PHP, Python…).

The CSV format is a human-readable format, more exactly an ASCII format. Each value is separated from the others by a comma or a semi-column. It can be easily ingested by tools such as Microsoft or OpenOffice suites, Matlab or proprietary applications written in any language (e.g., C, PHP, Python…).

In order to follow changes in technology, the NetCDF format has been adopted. NetCDF is a set of software libraries and self-describing, machine-independent data formats that support the creation, access, and sharing of array-oriented scientific data. NetCDF was developed and is maintained at Unidata, part of the University Corporation for Atmospheric Research (UCAR) Community Programs (UCP) in USA. Unidata is funded primarily by the National Science Foundation of USA. It is one of the formats recommended by the GEOSS (Global Earth Observation System of Systems) programme to which is contributing GMES / Copernicus. Several tools are available in the NetCDF Web site to create, handle and exploit NetCDF files (www.unidata.ucar.edu/software/netcdf/).

The McClear products are currently available in CSV and NetCDF formats.

15.6.2. Metadata

Metadata are available to describe the MACC-RAD products in GEOSS-compliant portals. These are *discovery* metadata that permit the cataloguing of these products and therefore to users to discover the products. These metadata obey to INSPIRE implementation rules (Ménard et al. 2009).

The MACC-RAD products are organised as lines of values (columns). Before the first line of data, there is a set of metadata for exploitation of the data. These metadata are written as text in the delivered file. They obey the ISO standard when available. They describe what are the various elements contained in a product. Currently these metadata are:

- *title*: title of the time-series, e.g., "Clear-sky irradiation at ground level. McClear_v1",
- *content*: a short description of the content of the product, e.g., "A time-series of solar radiation that should be received on a horizontal plane at ground level if the sky were clear",
- *provider*: name of the provider, e.g., "MINES ParisTech - Armines (France)",
- *date begin, date end*: dates of the beginning and end of the period. The date follows the ISO 8601 standard, e.g. 2004-01-01T01:00:00.000,
- *site latitude and longitude*: geographical coordinates of the site. The ISO 19115 standard is used,
- *elevation*: elevation above sea level in m,
- *time reference*: time system used: Universal Time (UT), or True Solar Time (TST). This metadata is present only when the period of integration is less than 1 d,
- *summarization*: period during which the energy is summed up or during which the power is averaged, e.g., 15 min,
- *sampling rate*: period with which the resulting irradiation is sampled. Usually, the sampling rate is set equal to *summarization*,
- *noValue*: the code denoting the absence of value, e.g. NaN (not a number).

The content of each column is described as a free text.

Each line corresponds to an instant of observation. The typical content of a line is:

- instant of observation,
- irradiation at the top of the atmosphere,
- irradiation value at ground level, e.g., GHI,
- reliability code, ranging from 0 to 1.

The exact content of a line depends on the type of product and this is why the metadata are included in the file. For example, there is a column "day" for products containing daily values, which does not exist in products containing monthly values.

The meaning of the reliability code depends on the type of products. This code may denote the proximity of the instant under concern with an estimate of the aerosol optical depth provided by the MACC project, or may indicate whether the value results from a time interpolation (case of 15-min or 1-h values), or may provide the number of valid 15-min values used to compute a daily average, or the number of valid daily values used to compute a monthly average. This is fully described in the exploitation metadata conveyed by the product.

When available, the uncertainty can be expressed as a single value, e.g., standard-deviation, or as lower and upper bounds. These bounds are such that there is a 68% chance that the actual value is comprised between lower and upper bounds.

15.6.3. More standard metadata for exploitation in next steps

In forthcoming steps depending on the progress of technologies, the format of the products will contain metadata for their automatic exploitation by other applications. These metadata for exploitation are conveyed in the products themselves. The supply of a product thus provides both the data and the explanations (metadata) to exploit the data in an automatic manner.

These metadata will be INSPIRE and GEOSS-compliant. A first application schema was developed by the international community on solar resource knowledge (Gschwind et al. 2007b). This schema is an example of the future technologies that will help developing web services that are interoperable, i.e., that can exchange information with various portals and other web services.

16. Description of the future MACC-RAD Service

The following is a preliminary version of the description of the future MACC-RAD Service. There are a large number of unknowns in the future of the GMES Atmospheric Service. Given these uncertainties, only a tentative description is made here.

The MACC-RAD Service comprises the ensemble of the databases, means and operations to construct these databases, and means and operations to construct and disseminate the products. The current databases, their construction, the products and their construction have been already described. This Chapter focuses on the dissemination part of the products.

Given the experience gained by several precursor services delivering databases and applications relating to solar radiation, the MACC-RAD Service has been designed as a Web service that disseminate the products. This Web service will follow the GEOSS (Global Earth Observation System of Systems) standards for interoperability and therefore can be exploited by any GEOSS-compliant portal. This will ensure a wide dissemination, more efficient than establishing a specific Web site.

In more details, the system is based on the technologies recommended by the EU-funded MESoR project, the International Energy Agency Task SHC 36, and the GEOSS AIP projects, in order to reduce the amount of technological development and to benefit from the adopted standards. The basic concept of the MACC-RAD Service is the following:

- products from MACC-RAD can be accessed through the MACC-RAD Web services,
- each MACC-RAD Web service is an application that can be invoked through the Web using GEOSS standards,
- the MACC-RAD Web services are deployed on the energy community portal (webservice-energy.org) from which they can be executed,
- the MACC-RAD Web services are described in existing catalogues to increase their visibility,
- clients, i.e. applications including an user-interface that can execute one or more MACC-RAD Web service, can be developed by anyone provided the client bear clear references and links to MACC and the GMES Atmospheric Service as a whole. The look should recall the identity of the MACC pages. A legal agreement must be signed with the owner of the client in line with the MACC policies on data dissemination and intellectual property rights,
- connect to the collaborative SoDa Service, built on the most advanced Web technologies and free software. One of the advantages is to benefit from the large penetration rate of the current SoDa Service in the solar energy community;
- once available, each MACC-RAD Web service is exposed in both the MACC Web site and the collaborative SoDa Service. One of the advantages of the latter is its large penetration rate in the solar energy community.

The commercial aspects are not discussed here as the MACC data policy and likely the GMES / Copernicus data policy aim at an open and free access. Nevertheless, several of these aspects have been considered in the design of the system, several of them not necessary relating to the sales of data but to safety of the information system. The selected tools offer means to manage communities of users with different access rights and tools for the monitoring of the system. It is possible to operate on a discriminatory basis for access to products. It may be possible that access may be restricted to a few products or depending of the requested volume of data, and means are available to manage such restrictions efficiently and in a transparent way.

In the following, the concepts of collaborative information systems and web services are introduced. Then, the concept of the MACC-RAD information system and its integration as a GEOSS component is presented.

16.1. Introduction to collaborative information systems and web services

The collaborative information system aims at providing an easy access to various services (databases, data sets, applications) in a homogeneous and seamless way. It is a one-stop shop where users find the requested information. Whatever the service, the user interface is more or less the same and consequently, very familiar. The delivery means is the same, i.e., by means of the Web. The format of the outputs obeys also some standardization.

It is not exactly a portal which displays links to providers of services. Here, the providers are "hidden" in the sense that though they are known to users of course, users do not access directly to their own web sites. Users only interact with the information system, i.e., one web site, and the information system plays the role of a broker.

Figure 16.1 is a sketch of the structure of the collaborative information system. It describes the actors: customers, broker, providers of services, and their relationships. The broker is offering a unique access to an ensemble of services supplied by providers. The broker is taking care of the relationship with its customers on the one hand, and with its providers on the other hand.

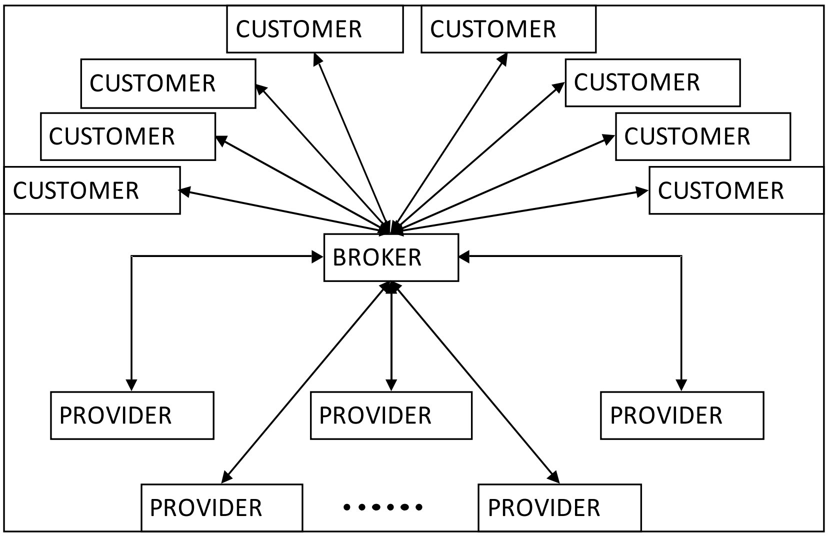


Figure 16.1. Sketch of the structure of the collaborative information system

The benefit of these collaborative systems is that they collect data and call upon other applications to answer as best as possible to the request. In doing so, the broker does not centralize the data into a warehouse nor write the software for applications. It concentrates on the best ways to satisfy the request of the customers and to select the most appropriate services. From the point of view of providers, the benefit is to keep hand on their data and applications. The provider does not need to send its data to a centralized database or data silos nor install its applications on a remote computer belonging to the broker.

The collaborative information system offers several advantages, ranging from immediate availability of new data or latest version of an application to the customers, to less restrictive agreements on intellectual property rights between the broker and a provider. From the point of view of customers, it offers a one-stop shop where to find information necessary to the accomplishment of their tasks and easy to use and reach because it is based on the Web.

According to Cros and Wald (2003), a collaborative information system, or co-operative information system, benefits from the advanced information technologies. Such an information system is essentially based on interactivity between different web servers with an information-delivery system calling on other servers that provide the necessary weather data and other information (e. g., topography, hardware properties, costs...).

An example of a mature collaborative system in the field of solar energy is the SoDa Service (www.soda-is.com). It realizes an integration of information sources of different natures and located in various institutes and companies in the world within a smart network on the Web (Gschwind et al., 2006). These sources include databases containing solar radiation parameters. Other relevant information is also available, such as other meteorological parameters: surface air temperature, relative humidity, forecasts, Linke turbidity factor, astronomical parameters, or digital elevation models. Several of these databases originate from an advanced processing of remote sensing images, especially those about surface solar irradiance or daylighting.

All these databases are maintained by different institutes and companies. Before the advent of the SoDa Service, several of them were already available on the Web but separately, while others were not. The SoDa Intelligent System builds smart Internet co-operation between sources (Wald et al., 2002, 2004). By this means, all databases are available within the same Web site: one-stop shop, and in a standardized manner. In addition to databases, the SoDa Service offers access to application-specific user-oriented numerical models and advanced algorithms. These applications are hosted by various institutes and companies, and here again, the SoDa Intelligent System builds link between one or more applications and the databases that are necessary as inputs (Gschwind et al., 2007a).

In collaborative systems, the various sources of information: databases, data sets, maps, applications…, are called services. In the SoDa information system, these services are actually Web services, in the sense of the World Wide Web Consortium, meaning that they are applications that can be invoked on the Web and that the basis of the network is the Web. As far as databases and data sets are concerned, the service is the application that retrieves data from the database.

Figure 16.1 gives a clue about responsibilities. The broker has the responsibilities for making customers respect the terms and conditions under which a provider is willing to supply information to the broker. The provider has the responsibilities of the quality and ownership of the information it supplies. The customer is granted a license of usage of information and should respect it. In any case, the property rights belong to the creator of the information.

Figure 16.1 also shows the middleware infrastructure, schematized by the arrows, that enables the exchange of information. Besides it, are hidden the metadata that are a key element for facilitate access to information in a seamless way as well as standards for exchange on the Web (e.g., HTTP, SOAP).

16.2. Concept of the MACC-RAD information system and its integration as GEOSS component

The infrastructure adopted for the MACC-RAD information system is fully aligned with GEOSS recommendations on interoperability. It is designed in such a way that it can be integrated as a GEOSS component. This ensures a wide dissemination of the results of the MACC-RAD operational system and a possible use of these results in many clients.

Figure 16.2 depicts the infrastructure and its integration as a GEOSS component. The MACC-RAD operational system and the corresponding radiation database are represented by a box on the right of the graph. Users may request for time-series of radiation values from a web site, shown on the left. In this example, the SoDa Service is shown and is one of many possible clients. It hosts a client: the MACC-RAD Client, which is the interface to the users, collects inputs for time-series, such as location and period, and delivers time-series. The SoDa Service does not connect directly to the MACC-RAD database. An intermediate has been set-up in order to ensure interoperability and facilitate exploitation of MACC-RAD results by other clients. This intermediate is a Web service, more exactly a Web Processing Service (middle of the figure). It receives the request from the MACC-RAD Client, makes itself queries to the MACC-RAD operational system, and flows the time-series from the database to the client.

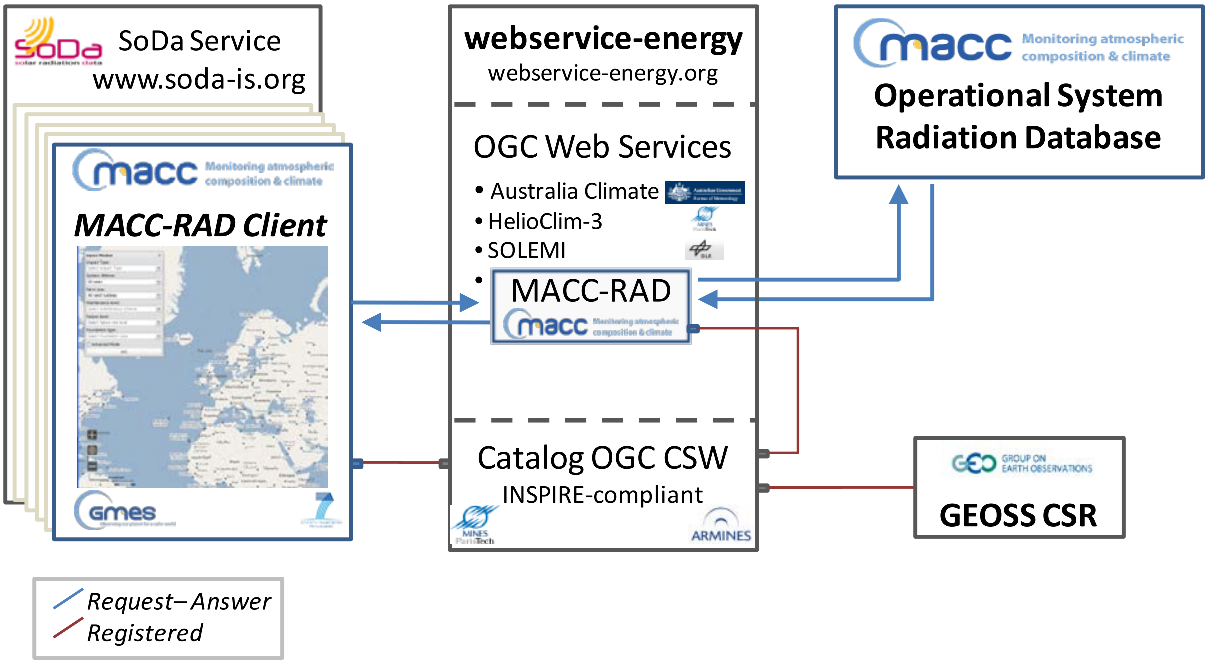


Figure 16.2. Sketch of the MACC-RAD information system and its integration as a GEOSS component

To ensure a wide dissemination, the MACC-RAD information system is declared as a GEOSS component by registering the component Web Service in a thematic catalogue on energy, itself registered in the GEOSS Components and Services Registry (GEOSS CSR box in the lower right part in Figure). By this means, anyone may discover the MACC-RAD Web Service and will be supplied with details on how to exploit them as discussed in the next section.

Invokation of the MACC-RAD Web Service is done by the means of a client, called MACC-RAD client in the Figure. This client should be developed by the broker.

The website webservice-energy.org is a GEOSS community portal. It is an initiative of MINES ParisTech / Armines to host web services dedicated to solar radiation and more generally, energy. The website has several web services: Web Map Services, Web Processing Services, obeying OGC (Open Geospatial Consortium) or W3C (World Wide Web Consortium) standards. The W3C standard is abandoned in favour of the OGC standard. Several services have been developed by MINES ParisTech, others by other providers such as DLR.

The website comprises also a catalogue obeying the OGC CSW (Catalogue Service for the Web) standard. The open-source Geonetwork (geonetwork-opensource.org) has been adopted as a CSW tool. Resources are described via metadata that are INSPIRE-compliant (INSPIRE: Infrastructure for Spatial Information in Europe). The catalogue will describe the MACC-RAD service, thus enabling its discovery and further exploitation. The catalogue is registered in the GEOSS Components and Services Registry; this ensures that the catalogue is fully aligned with GEOSS standard and increases the chances of discovery of the components registered in the catalogue.

16.3. Web Processing Service

Accessing the MACC-RAD database is done by the means of a Web Processing Service which obeys the standard set up by the OGC and recommended by GEOSS.

The OpenGIS® Web Processing Service (WPS) Interface Standard provides rules for standardizing inputs and outputs (requests and responses) for services aimed at processing the geospatial information extracted from the geographic databases. Here, the radiation database is defined for the Meteosat field-of-view and the "processing" is the extraction of a time-series for a given location, given period, and summarization.

The standard also defines how a client can request the execution of a process, and how the output from the process is handled. It defines an interface that facilitates the publishing of geospatial processes, their discovery by clients and the linking to those remotely hosted processes. The data required by the WPS can be delivered across a network or can be available through the local server.

WPS defines three operations, two describing service metadata and Input/Output (I/O) characteristics, and a third to execute/run the process:

- getCapabilities: generic WPS instance metadata, list of services in instance,
- describeProcess: full description I/O of service,
- execute: process execution with provided inputs and returned a formatted output, i.e. a netCDF file.

There are several frameworks available to develop and deploy OGC Web Processing Services. The Toolbox from the company INTECS, developed in the progress of the Integrated Project (2008-2011) GENESIS funded by the European Commission DG-INFSO under the 7th Framework Programme has been selected. The Toolbox is a configurable application released under GNU General Public Licence (GPL) that facilitates the conversion of applications into a Web Processing Service (WPS) in OGC standards. On the front end, the Toolbox encodes the outputs of the WPS using SOAP and transfers it using HTTP. On the back end, it can be connected via shell scripts to the application, here the radiation database.

16.4. The MACC-RAD Client

The MACC-RAD client is the interface between the user and the WPS. Several frameworks can be used to develop such a client. The development of a client is the responsibility of the broker.

In the course of the MACC-II project, a client is being developed in close collaboration with Transvalor for the SoDa Service.

16.5. Status of work

A WPS has been developed for the McClear clear-sky model to serve as a precursor of HelioClim-4 service and demonstrator of technology. It delivers time-series of irradiation under clear-sky. It is available since 1st November 2012.

Figure 16.3 is a screenshot of the McClear client developed by Transvalor in the SoDa Service. This client permits to invoke the McClear WPS. It can also be launched from the MACC Web site.

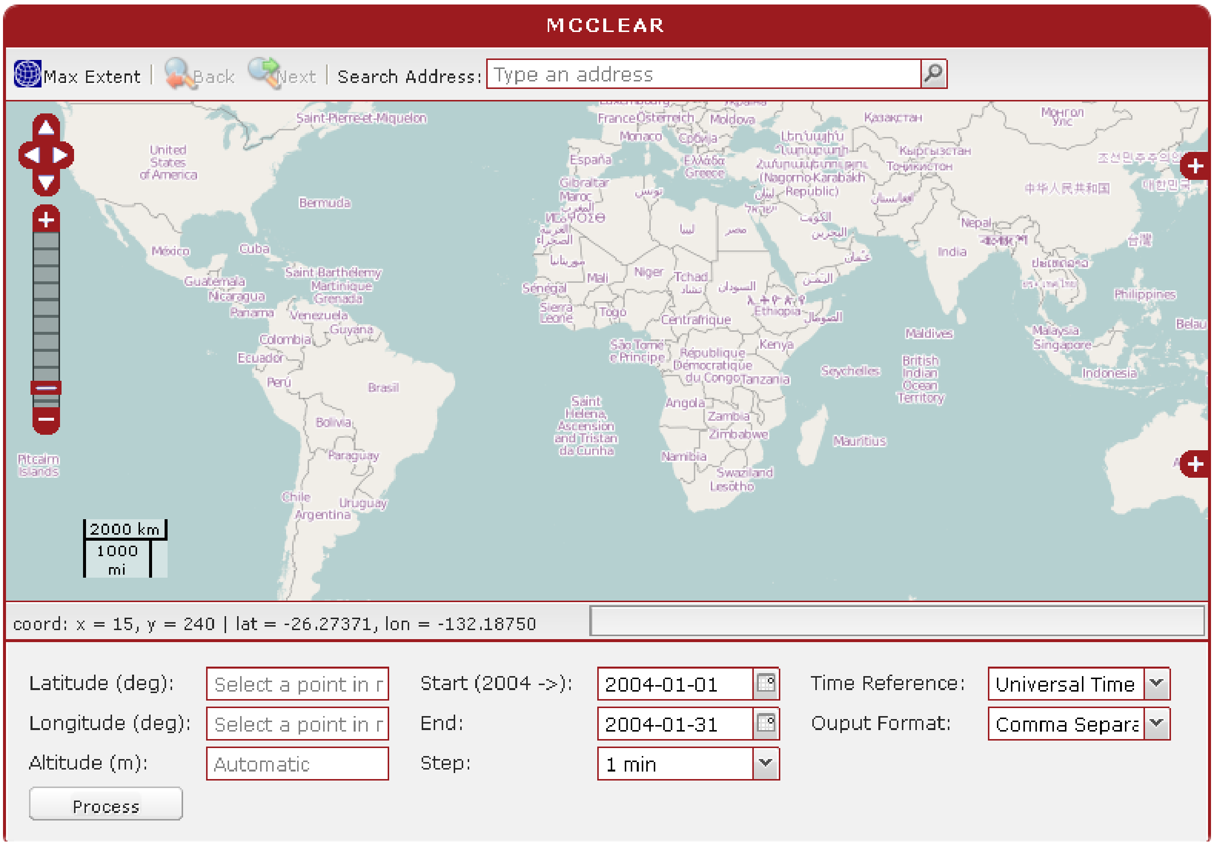


Figure 16.3. Interface of the McClear client in the SoDa Service or MACC Web site.

The result is a file that can be saved or displayed. The format is either CSV (comma-separated values) or NetCDF.

The file contains irradiation, expressed in Wh m^{-2}:

- on horizontal plane at the top of the atmosphere
- global, direct and diffuse at ground level
- direct on a plane normal to the sun rays at ground level,
- a reliability parameter, an experimental parameter aiming at describing the confidence one may have in data.

Means should be established to monitor the use of the MACC-RAD Service. At least, this would show how much the products are used, where and when. Actually, it would permit to establish strategy for a more efficient service. In particular, this would help in future versions of the service by, e.g., permitting a faster access to the most used products, establishing typical profiles of users and thus presenting products in a more relevant way, or in defining new products, e.g., a combination of products or an archive containing several products in one click instead of several.

McClear is currently in the 'MACC II - experimental routine production mode'. This includes that the product is routinely produced and available through the Product catalogue. There is no guarantee on product availability, accuracy is not necessarily monitored, and timeliness is not necessarily in line with user requirements.
The remaining part of the HelioClim-4 service is in the 'MACC-II - under development mode' and not available in either plot or data format.

17. References

Abdel Wahab M., El Metwally M., Hassan R., Lefèvre M., Oumbe A., Wald L., 2009. Assessing surface solar irradiance in Northern Africa desert climate and its long-term variations from Meteosat images. International Journal of Remote Sensing, 31(01), 261-280.

Beyer H.-G., Costanzo C., and Heinemann D., 1996. Modifications of the Heliosat procedure for irradiance estimates from satellite images. Solar Energy, 56, 207-212.

Beyer H. G., Martinez J. P., Suri M., Torres J. L., Lorenz E., Hoyer-Klick C., Ineichen P., 2008. D 1.1.1 Handbook on Benchmarking, Management and Exploitation of Solar Resource Knowledge, CA – Contract No. 038665.

Bird, R., Hulstrom, R., 1981. Review, evaluatoin and improvement of direct irradiance models. Journal of Solar Energy Engineering 103, 182–192.

Blanc Ph., Gschwind B., Lefèvre M., Wald L., 2011. The HelioClim project: Surface solar irradiance data for climate applications. Remote Sensing, 3, 343-361, doi:10.3390/rs3020343.

Blanc P., Wald L., 2012. The SG2 algorithm for a fast and accurate computation of the position of the Sun. Solar Energy, 86, 3072-3083, doi: 10.1016/j.solener.2012.07.018.

Cano, D., Monget, J., Albuisson, M., Guillard, H., Regas, N., and Wald, L., 1986. A method for the determination of the global solar radiation from meteorological satellite data. Solar Energy, 37, 31-39.

Collins, W.D., P.J. Rasch, B.E. Eaton, B.V.Khattatov, J.-F. Lamarque, 2001. Simulating aerosols using a chemical transport model with assimilation of satellite aerosol retrievals: Methodology for INDOEX., J. Geophys. Res., Vol 106, No D7, 7313-7336.

Cros S., Wald L., 2003. Survey of the main databases providing solar radiation data at ground level. In Proceedings of the 23rd EARSeL Annual Symposium "Remote Sensing in Transition", 2-4 June 2003, Ghent, Belgium, Rudi Goossens editor, Milpress, Rotterdam, Netherlands, pp. 491-497.

Cros S., Mayer D., Wald L., 2004. The Availability of Irradiation Data. Report IEA-PVPS T2-04: 2004, International Energy Agency, Vienna, Austria, 29 p.

Cros S., Albuisson M., Wald L., 2006. Simulating Meteosat-7 broadband radiances at high temporal resolution using two visible channels of Meteosat-8. Solar Energy, 80, 3, 361-367, doi: 10.1016/j.solener.2005.01.012.

Diabaté L., Demarcq H., Michaud-Regas N., Wald L., 1988. Estimating incident solar radiation at the surface from images of the Earth transmitted by geostationary satellites: the Heliosat Project. International Journal of Solar Energy, 5, 261-278.

Diabaté L., Moussu G., Wald L., 1989. Description of an operational tool for determining global solar radiation at ground using geostationary satellite images. Solar Energy, 42(3), 201-207.

Espinar B., Ramirez L., Drews A., Beyer H.G., Zarzalejo L.F., Polo J., Martin L., 2009a. Analysis of different error parameters applied to solar irradiation data from satellite and German radiometric stations, Solar Energy, 83(1), 118-125.

Espinar B., Ramírez L., Polo J., Zarzalejo L.F., Wald L., 2009b. Analysis of the influences of uncertainties in input variables on the outcomes of the Heliosat-2 method. Solar Energy, 83, 1731-1741, doi:10.1016/j.solener.2009.06.010.

Geiger, M., Diabate, L., Menard, L., Wald, L., 2002. A web service for controlling the quality of measurements of global solar irradiation. Solar Energy 73, 475–480.

GEOSS, 2005. The Global Earth Observation System of Systems (GEOSS) 10-Year Implementation Plan (As adopted 16 February 2005), http://www.earthobservations.org/documents/10-Year%20Implementation%20Plan.pdf

Grüter W., Guillard H., Möser W., Monget J.M., Palz W., Raschke E., Reinhardt R.E., Schwarzmann P., Wald L., 1986. Solar radiation data from satellite images. Solar Energy R&D in the European Communities, series F: Solar Radiation Data, vol. 4, D. Reidel Publishing Co., 100 p.

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. Environmental Modelling and Software, 21, 1555-1561, doi:10.1016/j.envsoft.2006.05.002.

Gschwind B., Wald L., Mahl R., Irigoin F., Ménard L., 2007a. Test of several approaches for the composition of web services in meteorology. In Proceedings of EnviroInfo 2007, Environmental Informatics and systems research, the 21st International Conference on "Informatics for Environmental Protection", Warsaw, Poland, 2007, vol. 1 "Plenary and session papers", p. 127-133, ISBN 978-3-8322-6397-3.

Gschwind B., Ménard L., Ranchin T., Wald L., Stackhouse P., 2007b. A proposal for a thesaurus for web services in solar radiation. In Proceedings of EnviroInfo 2007, Environmental Informatics and systems research, the 21st International Conference on "Informatics for Environmental Protection", Warsaw, Poland, 2007, vol. 1 "Plenary and session papers", p. 135-142, ISBN 978-3-8322-6397-3.

Hansen, J., Sato, M., Ruedy, R., Nazarenko, L., Lacis, A., Schmidt, G., Russel, G., Aleinov,I., Bauer, M., Bauer, S., Bell, N., Cairns, B., Canuto, V., Chandler, M., Cheng, Y., Genio, A. D., Faluvegi, G., Fleming, E., Friend, A., Hall, T., Jackman, C., Kelley, M., Kiang, N., Koch, D., Lean, J., Lerner, J., Lo, K., Menon, S., Miller, R., Minnis, P., Novakov, T., Oinas, V., Perwitz, J., Perlwitz, J., Rind, d., Romanou, A., Shindell, D., Stone, P., Sun, S., Tausnew, N., Thresher, D., Wielicki, B., Wong, T., Yao, M., Zhang, S., 2005. Efficacy of climate forcings. Journal of Geophysical Research 110, D18104.

Hoyer-Klick C., Dumortier D., Tsvetkov A., Polo J., Torres J. L., Kurz C., Ineichen P., 2008. D 1.1.2 Existing Ground Data Sets, Management and Exploitation of Solar Resource Knowledge, CA – Contract No. 038665.

Hoyer-Klick C., Schillings C., Schroedter Homscheidt M., Beyer H.-G., Dumortier D., Wald L., Menard L., Gschwind B., Martinoli M., Gaboardi E., Ramirez L., Polo J., Huld T., Suri M., Cebecauer T., De Blas M., Lorenz E., Pfatischer R., Remund J., Ineichen P., Tsvetkov A., Hofierka J., 2008. Management and exploitation of solar resource knowledge. Proceedings, EUROSUN 2008, 1st International Congress on Heating, Cooling and Buildings, Lisbon, Portugal (2008).

Hoyer-Klick C., Beyer H.-G., Dumortier D., Schroedter Homscheidt M., Wald L., Martinoli M., Schillings C., Gschwind B., Menard L., Gaboardi E., Polo J., Cebecauer T., Huld T., Suri M., De Blas M., Lorenz E., Kurz C., Remund J., Ineichen P., Tsvetkov A., Hofierka J., 2009. MESoR - Management and exploitation of solar resource knowledge. SolarPACES 2009, Berlin, Germany.

Huld T., Šúri M., Meyer R., Perez R., Wald L., Renne D., Stackhouse P. W., Dumortier D., Schroedter-Homscheidt M., Polo J., Schillings C., Remund J., Lorenz E., Ineichen P., Beyer H.-G., 2008. Customers' Requirements of Solar Energy Resource Information - Results of the IEA SHC Task 36 on-line survey. Available on-line at http://www.geoss-ecp.org/sections/solar/iea-shc-task36-survey-on.

Iqbal, M., 1983. An Introduction to Solar Radiation. Academic Press, New York.

ISO, 1995. Guide to the Expression of Uncertainty in Measurement, first ed. International Organization for Standardization, Geneva, Switzerland.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, G., White, J., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgin, W., Janowiak, J., Mo, K., Rpelewski, C., Wang, J., Jenne, R., Joseph, D., 1996. The nmc/ncar 40-year reanalysis project 77, 437–472.

Kasten, F., Young, A. 1989. Revised optical air mass tables and approximations formula. Applied Optics, 28, 4735-4738.

Kasten F., 1996. The Linke turbidity factor based on improved values of the integral Rayleigh optical thickness. Solar Energy, 56 (3), 239-244.

Kinne, S., Holben, B., Eck, T., Smirnov, A., Dubovik, O., Dlutsker, I., Tanre, D., Zibozdi, G., Lohmann, U., Gahn, S., Easter, R., Chin, M., Ginoux, P., Takemaura, T., Tegen, I., Koch, D., Kahn, R., Vermote, E., Stowe, L., Torres, O., Mishchenko, M., Geogdzhayev, I., Hiragushi, A., 2001. How well do aerosol retrievals from satellites and representation in global circulation models match ground based aeronet statistics. In: Beniston, M.,

Verstraete, M. (Eds.), Remote Sensing and Climate Modeling: Synergies and Limitations (Advances in Global Change Research. 7). Kluwer Academic Publishers, Dordrecht, pp. 103–158.
Kinne, S., Lohmann, U., Gahn, S., Easter, R., Chin, M., Ginoux, P., Takemura, T., Tegen, I., Koch, D., Herzog, M., Penner, J., Pitari, G., Holben, B., Eck, T., Smirnow, A., Dubovik, O., Slutsker, I., Tanre, D., Torres, O., Mshchenk, M., Geoghzhayev, I., 2003. Monthly averages of aerosol properties: A global comparision among models, satellite data and aeronet ground data. Journal of Geophyical Research 108, D20.
Kinne, S., Schulz, M., Textor, C., Guibert, S., Balkanski, Y., Bauer, S., Bernsten, T., Berglen, T., Boucher, O., Chin, M., Collins, W., Dentener, F., Diehl, T., Easter, R., Feichter, J., Fillmore, D., Ghan, S., Ginoux, P., Gong, S., Grini, A., Hendricks, J., Herzog, M., Horowitz, L., Isaksen, I., Iversen, T., Kikevag, A., Kloster, S., Koch, D., Kristjansson, J., Krol, M., Lauer, A., Lamaraque, J., Lesins, G., Liux, X., Lohmann, U., Montanaro, V., Myhre, G., Penner, J., Pitari, G., Reddy, S., Seland, O., Stier, P., Takemura, T., Tie, X., 2005. An AeroCom initial assessment - optical properties in aerosol component modules of global models. Atmospheric Chemistry and Physics Discussions 5, 8285–8330.
Kistler, R., Kalnay, E., Collin, W., Saha, S., White, G., Wollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kusky, V., van den Dool, H., Jenne, R., Fiorina, M., 2001. The NCEP NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. Bulletin of the American Meteorological Society 82, 247–267.
Lefèvre M., Diabaté L., Wald L., 2007. Using reduced data sets ISCCP-B2 from the Meteosat satellites to assess surface solar irradiance. Solar Energy, 81, 240-253, doi:10.1016/j.solener.2006.03.008.
Liu, H., Pinker, R.T., 2005. A global view of aerosols from merged transport models, satellite and ground observations. Journal of Geophysical Research 110, D10S15.
McPeters, R., Bhartia, P., Krueger, A., Herman, C., Wellmeyer, C., Seftor, C., Jaross, G., Torres, O., Moy, L., Labow, G., Byerly, W., Taylor, S., Swissler, T., Cebula, R., 1998. Earth probe total ozone mapping spectrometer (TOMS) data products user's guide. Tech. rep., NASA Technical Publication 1998-206895.
Ménard L., Wald L., Blanc Ph., Ranchin T., 2009. Sitting of a solar power plant: development of Web service based on GEOSS data and guidance. In: Proceedings, 33rd International Symposium on Remote Sensing of Environment, ISRSE 33, Stresa, Italy, May 4-8, 2009, paper 789, [USB key].
MESOR, 2009a. D.1.3.1 Future research objectives and priorities in the field of solar resources, Ed. by M. Schroedter-Homscheidt, EU CA contract 038665 on 'Management and Exploitation of Solar Resource Knowledge'.
MESOR, 2009b. D.1.3.2 Needs for new solar radiation services to faster deploy the market for solar energy applications and optimized grid integration, Ed. by M. Schroedter-Homscheidt, EU CA contract 038665 on 'Management and Exploitation of Solar Resource Knowledge'.
MESOR, 2009c. D.1.3.3 Recommendations on an Improved Earth Observation System to better support solar energy, Ed. by M. Schroedter-Homscheidt, EU CA contract 038665 on 'Management and Exploitation of Solar Resource Knowledge'.
Möser, W. and Raschke, E., 1984. Incident solar radiation over Europe estimated from Meteosat data. Journal of Applied Meteorology, 23, 166-170.
Mueller, R., Dagestad, K., Ineichen, P., Schroedter, M., Cros, S., Dumortier, D., Kuhlemann, R., Olseth, J., Piernavieja, G., Reise, C., Wald, L., and Heinnemann, D., 2004. Rethinking satellite based solar irradiance modelling - the SOLIS clear sky module. Remote Sensing of Environment, 91, 160-174.
Oumbe A., 2009. Exploitation des nouvelles capacités d'observation de la terre pour évaluer le rayonnement solaire incident au sol (Assessment of solar surface radiation using new earth observation capabilities). Thèse de Doctorat, MINES ParisTech, Paris, France, 128 pages).
Oumbe A., Blanc Ph., Schroedter-Homscheidt M., Wald L., 2010. Solar surface irradiance from new meteorological satellite data. In Proceedings of the 29th Symposium of the

European Association of Remote Sensing Laboratories, Chania, Greece, 15-18 June 2009. I. Manakos and C. Kalaitzidis (Eds.). Published by IOS Press, 320-328, doi:10.3233/978-1-60750-494-8-320.

Perez R., Seals R., Zelenka A., 1997. Comparing satellite remote sensing and ground network measurements for the production of site/time specific irradiance data. Solar Energy, 60, 89–96.

Remund J., Wald L., Lefèvre M., Ranchin T., Page J., 2003. Worldwide Linke turbidity information. In Proceedings of ISES Solar World Congress, 16-19 June, Göteborg, Sweden, CD-ROM published by International Solar Energy Society.

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

Rigollier C., Lefèvre M., Wald L., 2004. The method Heliosat-2 for deriving shortwave solar radiation from satellite images. Solar Energy, 77(2), 159-169.

Ruiz-Arias J., H. Alsamamra, J. Tovar-Pescador, D. Pozo-Vázquez, 2010. Proposal of a regressive model for the hourly diffuse solar radiation under all sky conditions. Energy Conversion and Management. 51, 881–893.

Schillings C. 2004. Bestimmung langjähriger stündlicher Zeitreihen und räumlich hochaufgelöster Karten der Direkt-Normal-Strahlung auf der Basis von Meteosat-Daten und Atmosphärenparametern für die Nutzung in konzentrierenden Solarkraftwerken. Dissertation, Uni-Marburg. (in German).

Schillings C., Mannstein H., Meyer R. 2004. Operational method for deriving high resolution direct normal irradiance from satellite data. Solar Energy, 76 (4), S. 475 – 484.

Schroedter-Homscheidt M., Delamare C., Heilscher G., Heinemann D., Hoyer C., Meyer R., Toggweiler P., Wald L., Zelenka A., 2006. The ESA - ENVISOLAR project : experience on the commercial use of Earth observation based solar surface irradiance measurements for energy business purposes. Chapter 10, In: Solar Energy Resources management for Electricity Generation, p. 111-124, ISBN 1-59454-919-2, © 2006 Nova Science Publishers Inc.

Schroedter-Homscheidt M., Stackhouse P., Ranchin T., 2007. Space for Solar energy. In: *The Full Picture*, published by Tudor Rose, 2007, p. 198-201, 278 p. ISBN 978-92-990047-0-8.

Schroedter-Homscheidt, M., Oumbe, A., 2013. Validation of an hourly resolved global aerosol model in answer to solar electricity generation information needs, Atmos. Chem. Phys., 13, 3777-3791, 2013, doi:10.5194/acp-13-3777-2013

Schmidt, G. A., Ruedy, R., Hansen, J. E., Aleinov, I., Bell, N., Bauer, M., Bauer, S., Cairns, B., Canuto, V., Cheng, Y., Del Genio, A., Faluvegi, G., Friend, A. D., Hall, T. M., Hu, Y., Kelley, M., Kiang, N. Y., Koch, D., Lacis, A. A., Lerner, J., Lo, K. K., Miller, R. L., Nazarenko, L., Oinas, V., Perlwitz, J., Perlwitz, J., Rind, D., Romanou, A., Russell, G. L., Sato, M., Shindell, D. T., Stone, P. H., Sun, S., Tausnev, N., Thresher, D., Yao, M.-S., 2006. Present-day atmospheric simulations using giss modele: Comparison to in situ, satellite, and reanalysis data. Journal of Climate 19, 153–192.

Stackhouse P., Renne D., Perez R., Meyer R., Wald L., Suri M., 2006. Towards designing an integrated Earth observation system for the provision of solar energy resource and assessment. In Proceedings of the IEEE IGARSS Symposium 2006, Denver, Colorado, USA.

Tanre´, D., Kaufman, Y., Makajima, T., Ramanathan, V., 2005. Preface to special section on global aerosol system. Journal of Geophysical Research 110, D10S01.

Tarpley, J., 1979. Estimating incident solar radiation at the surface from geostationary satellite data. Journal of Applied Meteorology, 18, 1172-1181.

Taylor V. R., Stowe L. L., 1984a. Reflectance characteristics of uniform Earth and cloud surfaces derived from Nimbus 7 ERB. Journal of Geophysical Research 89, 4987–4996.

Taylor V.R., Stowe L.L., 1984b. Atlas of reflectance patterns for uniform Earth and cloud surfaces (Nimbus 7 ERB—61 days), NOAA Technical Report NESDIS 10, July 1984, Washington, DC, USA.

Tegen, I., Hollring, P., Chin, I., Fung, D., D., J., Penner, J., 1997. Contribution of different aerosol species to the global aerosol extinction optical thickness: Estimates from model results. Journal of Geophysical Research 102, 23895–23915.

Trenberth, K., Guillemot, C., 1998. Evaluation of the atmospheric moisture and hydrological cycle in the NCEP/NCAR reanalysis. Climate Dynamics 14, 213–231.

Wald L., Albuisson M., Best C., Delamare C., Dumortier D., Gaboardi E., Hammer A., Heinemann D., Kift R., Kunz S., Lefèvre M., Leroy S., Martinoli M., Ménard L., Page J., Prager T., Ratto C., Reise C., Remund J., Rimoczi-Paal A., Van der Goot E., Vanroy F., and Webb A., 2002. SoDa: a project for the integration and exploitation of networked solar radiation databases. In: Environmental Communication in the Information Society, W. Pillmann, K. Tochtermann Eds, Part 2, pp. 713-720. Published by the International Society for Environmental Protection, Vienna, Austria.

Wald L., Albuisson M., Best C., Delamare C., Dumortier D., Gaboardi E., Hammer A., Heinemann D., Kift R., Kunz S., Lefèvre M., Leroy S., Martinoli M., Ménard L., Page J., Prager T., Ratto C., Reise C., Remund J., Rimoczi-Paal A., Van der Goot E., Vanroy F., and Webb A., 2004. SoDa: a Web service on solar radiation. In Proceedings of Eurosun 2004, published by PSE GmbH, Freiburg, Germany, pp. (3)921-927, ISBN 3-9809656-4-3.

Wald L., Blanc Ph., Espinar B., Gschwind B., Lefevre M., Menard L., Wald F., Hoyer-Klick C., Schroedter-Homscheidt M., Thomas C., 2011. Early achievements towards an automatic assessment of the uncertainty in solar irradiation using web services. In Proceedings of the 25th EnviroInfo Conference 'Environmental Informatics', 5-7 October 2011, Ispra (VA), Italy. Pillman W., Schade S., Smits P. (Eds). Part 1, pp. 309-318. Publisher Shaker Verlag, Aachen, Germany.

Wielicki B., Welch R., 1986. Cumulus cloud properties derived using Landsat satellite data. Journal of Climate and Applied Meteorology, 25, 261-276.

WMO, 1981. Technical Note N° 172, WMO-No. 554, World Meteorological Organization, Geneva, Switzerland, pp. 121-123.

Zarzalejo L. F., Polo J., Martin L., Ramirez L., Espinar B., 2009. A new statistical approach for deriving global solar radiation from satellite images. Solar Energy, 83(4), 480-484.

Zelenka A., Perez R., Seals R., Renne´ D., 1999. Effective accuracy of satellite-derived hourly irradiances. Theoretical and Applied Climatology, 62, 199–207.

Zender, C. S., H. Bian, and D. Newman, 2003. Mineral Dust Entrainment And Deposition (DEAD) model: Description and 1990s dust climatology. J. Geophys. Res., 108(D14), 4416, doi:10.1029/2002JD002775.