Executive Summary:
We present the summary of the activities within the third year (Year 3) of SOLID, covering the time from December 1, 2014 to November 30, 2015. The key activities and accomplishments of the third reporting period of the SOLID Project are:

(i) the creation of an observational composite dataset of the SSI from 1951-2014 using 20 quality-checked SSI records. This also includes the creation of a new composite of the TSI (soon to be the official one).
(ii) extensive investigation of the SSI stability, and provision of recommendations on existing SSI dataset and future observations.
(iii) intensive interaction with the other work packages to work on the composite and discuss how to use our quality estimates of the SSI database
(iv) provision of SOHO/CDS irradiance data set to complement irradiance observations.
(v) the completion of the sunspot catalogue
(vi) improvement of the Mg II index and uncertainty estimates
(vii) image segmentation of the PSPT, SOHO/EIT and SOHO/MDI datasets
(viii) detailed uncertainty estimates of the segmentation tools SPoCA and ASAP
(ix) modelling of the solar spectral irradiance based on solar image segmentation and semi-empirical modelling of the solar spectrum
(x) critical discussion and release of solar spectral irradiance reconstructions starting in 1700 based on modelled proxies
(xi) outreach activities such as close interaction with the user communities at international conferences and the provision of an annual newsletter

The scientific outcome of the tasks is published in 25 accepted/published publications, and still about 7 submitted/in preparation. In case all of them get accepted, SOLID has produced at least 32 peer-reviewed publications.

Project Context and Objectives:
The role of solar variability in climate change is a topic of considerable scientific and societal importance, and also a topic of major controversy. The key parameter is the time-varying solar spectral irradiance (SSI), whose highly fragmented observations by over 25 different spacecraft has made its scientific exploitation a major and so far unmet challenge. The First European Comprehensive SOLar Irradiance Data exploitation (SOLID) project will take up this challenge by creating for the first time one single homogenous solar irradiance record out of all these scattered observations. Such a record is essential for climate science, but is also highly relevant for disciplines such as space exploration, space weather and heliospheric science.

The overall objective of SOLID is twofold. First, it will bring together a large variety of different solar spectral irradiance data sets (spectral measurements, solar images, and solar proxies), merge them, enrich them, and deliver one single continuous and high level SSI record that covers mainly the space era (from 1979 onwards) but also before. Secondly, SOLID uses novel methods for reconstructing the SSI, thereby enhancing our capacity to extend the reconstruction farther back in time, but mitigating the risk of losing crucial information when observational gaps occur. To do so, SOLID will mobilise the best expertise in Europe, and also involves experts from the US. The products to be delivered are very timely, as there has recently been a clear community requirement for more comprehensive data products that comply with metrological requirements and a better coordination of existing resources. The main benefits will be an authoritative source of information on SSI, a significant enhancement of present knowledge of solar variability and a long-term structural improvement in the interaction between heliospheric and climate science communities.

The influence of solar spectral irradiance on climate
Although solar electromagnetic radiation is by far the dominant source of energy to the Earth, the actual contribution of solar forcing to recent climate change is hotly debated. The excellent correlation of glacial cycles with Earth’s orbital variations is one among the many pieces of evidence in favour of an impact of our variable star on climate. Thus, the solar irradiance, or more precisely, variation of the SSI, is the most important parameter of natural influence on climate.

On shorter time scales, such as 100 to 10,000 years the identification of the dominant forcing is less
On shorter time scales, such as 100 to 10000 years the identification of the dominant forcing is less obvious. Nevertheless, there is a consensus that irradiance variations have likely again a leading role in the climate forcing but arguably, there are also other natural climate influences such as major volcanic eruptions that have to be considered. In any case, variations of solar irradiance are an important factor and this influence has to be considered and attributed to when investigating climate change.

Even in the present time, when the focus is on anthropogenic forcing, solar irradiance and its variations are still an important factor that has to be included. Surprisingly, very few climate simulations use realistic solar forcings as input. In the absence of better data most use a solar irradiance with a constant energy distribution, or even a constant solar energy input! The first reason for this stems from the fact that none but a few advanced models have the option for detailed spectral input specification. The second is more structural and has much deeper consequences: the widespread ignorance of solar forcing mainly results from the lack of readily available SSI data set that can be used by non-experts. Although space based observations of solar irradiance exist for about the last three decades, these data are highly scattered, inhomogeneous in both time and wavelength coverage, and often are very hard to access. This longstanding problem is exactly what SOLID will solve by delivering one single and well-documented record, together with a sound estimate of the uncertainty level.

The compilation of a good set of solar forcing and therefore presently, only those climate modelers care about the solar irradiance, who are investigating the Sun-Earth relation. For those users semi-empirical reconstructions have been compiled, as detailed below, but all these data sets have known weaknesses. There is a strong need for improving these data sets in several aspects: time resolution, wavelength distribution, and most importantly, reliability in reproducing the real solar irradiance variations on solar cycle timescales and longer. The SOLID consortium comprises the expertise, and with the support of this project also the manpower, to deliver a new improved and simple to use spectral irradiance and related proxy data sets. SOLID has a specific work package (WP6) whose the role is to ensure that the needs of the user community are met. Then, in WP8 the results from this project will be efficiently disseminated to the climate research community and other user communities.

The goal of SOLID is to bring added value to the European spectral irradiance observations by combining all existing observations and at the same time, filling the temporal and spectral gaps through the modelling and reconstruction of spectral irradiance variations. In this way it will enable space researchers to take full advantage of the potential value of the existing irradiance data sets.

The objectives are addressed in the following four work packages, WP2 to WP5:

- Irradiance Data Exploitation (WP2)
- Multi-wavelength solar image processing for novel SSI proxies (WP3)
- Modelling of spectral irradiance based on observed proxies (WP4)
- Solar irradiance reconstructions using simulated magnetograms (WP5)

WP6 and WP8 is then dedicated to the scientific interaction with the climate research community and the dissemination of the added value data product: the uniform Spectral Solar Irradiance data set as a function of time for the space era and beyond.

WP2 - Irradiance Data Exploitation

Solar irradiance data sets acquired from space are sparse and of varying quality. The main reasons for this are the complexity of single instrument calibration, the variety of instruments with different duration and differing spectral specifications, and, more importantly, the optical degradation of the instrument.
and differing spectral specifications, and, more importantly, the optical degradation of the instrument performances caused by the harsh space environment. The calibration and stability (better corrections of degradations) of a single instrument that monitor solar irradiance in a given spectral band is already a difficult task; combining the data of instruments with different characteristics is consequently a true challenge. SOLID will tackle this problem by combining observation expertise in data analysis and empirical as well as physical modelling.

Satellite data now cover nearly 30 years of data, yet they originate from different instruments and satellites; WP2 will in a first step collect and quality check the available datasets, characterizing their strengths and weaknesses, in particular in term of long-term stability. This will be done by means of statistical testing but also by comparison with proxies of solar variability, that are less affected by optical degradation and which have recently revealed new important information (Clette, 2011; Lefevre and Clette, 2011).

In a next step, WP2 will concentrate on the ultraviolet and extreme ultraviolet spectral range and combine data sets in order to build new robust composite. The work of DeLand and Cebula (2008) will be improved and extended; for this, existing formal collaboration (MUSIC) with US colleagues (LASP) will be an asset. Apart from the need to have suitable estimates of SSI, equally important are assessment of errors either in the observed SSI or parameters (proxies, images) that are used to reconstruct SSI. The errors may vary with time as observation techniques and instruments have evolved and changed over time. Particular care will be put in assessing the error on both original and new composite datasets. Datasets from the SOHO (CDS-SEM), ENVISAT (SCIAMACHY), ERS-2 (GOME), PROBA2 (LYRA), and PICARD (PREMOS), all ESA missions, will greatly benefit from this work. Firstly, inter-calibration will be achieved, and, secondly, these datasets will be put into contexts with other spectral irradiance observations in the space era. Understanding solar variability and its effects on Earth climate requires observations on timescales much longer than a single space mission lifetime. WP2 will provide the links between these various datasets and all other irradiance dataset obtained since the beginning of the space era; As an outcome of our analysis, WP2 will also provide recommendations for future observations in order to identify observational gaps and to optimize future long-term monitoring of solar variability.

The objectives of WP2 are thus
1) to provide an overview of the existing irradiance datasets and of their confidence level,
2) to perform an in-depth analysis of particularly relevant datasets with emphasis on ESA missions,
3) to perform an exploration of the SSI-proxy relationships (in collaboration with WP4) to provide best-possible ultraviolet composites with proper error and long-term uncertainty estimates, and
5) to provide recommendations for future space observations.

WP3 - Multi-wavelength solar image processing for novel SSI proxies
SOLID will use state-of-the-art solar image processing techniques in order to derive high level products that will be of direct use for reconstructions and semi-empirical modelling of SSI (WP4,WP5). Magnetogram, photospheric, and coronal images will be used to extract meaningful information to construct new proxies. SOLID will also be unique in the sense that it will be the first project to develop new SSI proxies for more than one solar cycle. The new proxies will be developed in WP3 by processing multi-wavelength space and ground-based images from different observatories (e.g. SOHO, PROBA2, PICARD, PSPT, etc). Advanced and highly automated image processing and machine learning techniques will be applied for large-scale processing of these data sets to detect relevant solar structures.
techniques will be applied for large-scale processing of these data sets to detect relevant solar structures that have specific contributions to the solar spectrum (e.g. sunspots, faculae, coronal holes, solar network, etc). Filling factors will be calculated for these structures by measuring what fraction of the solar disc is covered by each of them and then assigning reference spectra to them. In doing so a range of novel imaging technologies will be developed and will be used to upgrade existing state-of-the-art technologies developed by partners such as the Automated Solar Activity Prediction (ASAP) system developed by UNIBRAD and the Spatial Possibilistic Clustering Algorithm (SPoCA) developed by ROB. Special attention will be paid to the inter-calibration of SSI proxies produced from different observatories to provide higher-level products and to the dissemination of some of these products to the wider solar physics community by integrating them into the SOLID database. WP3 is at the heart of SOLID. It will receive input and feedback from WP2 and particular care will be taken to deliver the proxies and higher-level products in a format that makes them useful to the interested communities, which are represented in WP4, WP5 and WP6.

WP4 - Modelling of spectral irradiance based on observed proxies

Merging different SSI observations, stitching them together and filling their gaps are some of the most challenging tasks SOLID will face. Such tasks require adequate methods, and one of our main objectives will precisely be the introduction of a series of novel and more advanced techniques for deriving proxies, and using them for building a single and homogenous SSI composite record. There are different aspects to this problem. The first one is the technical merging of different data sets collected in WP2, which so far was often done manually. With over 25 different data sets to handle, SOLID will enforce a more rigorous framework that involves techniques developed by some of its members and will also build on the experience gathered in the Earth observation community. WP4 will be in charge of systematically encouraging such a framework, which will be crucial for properly assessing the uncertainty of the data. Let us stress that for many climate modellers, having sound uncertainty estimates is equally important as good SSI values, because the former are needed to organise model runs. This is why SOLID aims at delivering both SSI data and confidence intervals.

The second aspect is the filling in of data gaps. WP5 addresses this aspect by providing semi-empirical models to bridge gaps. This filling in has traditionally been done with various solar proxies such as the sunspot number. Such proxies are of valuable use for reconstructing the SSI before the space age, but they also have serious limitations in their capacity of reconstruction all spectral bands at both short and long time-scales. Nevertheless, such proxies are an indispensable (and often only) tool for probing the solar variability. For that reason, SOLID will heavily rely on them as well. The main novelty is the introduction of new solar proxies such as the often overlooked solar radio flux data going back to the 1950’s (from Toyokawa and Nobeyama observatories) and the recently developed Ca II K index that goes back to 1915. SOLID will not only use these proxies to fill in missing values but also to deepen our understanding of the causes of solar variability.

Another objective of WP4 is to develop better solar proxies, and feed back to WP3 for this. The idea here is to use solar proxies as derived from segmented solar images to improve the SSI reconstruction. This can be done either by multiple regression, as with classical proxies, or by incorporating a spectral synthesis stage, using physically realistic spectra. This apparent redundancy between all these approaches is one of the unique features of SOLID and will guarantee that for each spectral band and for each period in time, the most appropriate solution will be used.

Finally, WP4 also aims at providing a more complete atomic data set, which is essential for a more correct modelling of the solar spectrum. The task mainly involves a comparative study between the EUV/Lyman Alpha.
Modeling of the solar spectrum. The task mainly involves a comparative study between the EUV/UV spectral radiances/irradiances as observed by the various instruments, and as obtained with modelling and the CHIANTI atomic data.

WP5 - Modelling of spectral irradiance based on modelled proxies

Measurements of total and spectral solar irradiance have only been available since the late 1970s and the spectral data set in particular is very incomplete in terms of wavelength as well as time coverage, especially before 1991. In order to obtain a consistent spectral variability data one thus has to rely on irradiance reconstructions. WP5 aims to provide just such a well-characterised set of total and spectral solar irradiance reconstructions going back to 1874.

SATIRE-S (Spectral And Total Irradiance REconstructions for the Satellite era) is one of the most successful models providing such reconstructions, capturing more than 90% of the observed variability (see, e.g. Krivova et al. 2003, 2011b and Wenzler et al. 2006). SATIRE-S uses magnetograms together with solar continuum images to track changes in the distribution of the solar surface flux as traced by sunspots and faculae (predominantly bright, small-scale magnetic flux tubes). No magnetograms are available before 1974, so that the surface flux is inferred from sunspot numbers and/or areas (Wang et al. 2005, Krivova et al. 2007, 2010, Crouch et al. 2008). As sunspot numbers only carry information about the active-region flux emergence rate, but not about its surface distribution, such reconstructions are successful on monthly and yearly timescales, but fare less well on timescales of days and weeks. Variability on longer timescales (decades to centuries) is thought to be due to a build-up of background magnetic flux; it is modelled using an extended sunspot cycle reflecting the appearance of ephemeral-region flux belonging to the new cycle before the preceding cycle has come to an end.

Within the framework of SOLID, we will improve existing total and spectral solar irradiance reconstructions by using modelled magnetograms to recover the solar surface-flux distribution and thus better capture the irradiance variability on timescales of days to weeks. We will use a surface flux transport model to simulate daily full-disc magnetograms since 1874 from the directly measured sunspot areas and positions (Cameron et al. 2010, Jiang et al. 2010a). This allows a more realistic description of the magnetic flux emergence than earlier flux transport models (e.g. Wang et al. 2005) since it does not require assumptions on the varying strengths or the number of active regions in different solar cycles. Reconstructions based on modelled magnetograms will thus describe irradiance changes on timescales of days to weeks more accurately than models employing the sunspot number (or areas) directly (e.g. Krivova et al. 2007, 2010). The spectral irradiance reconstructions from WP5 will be consolidated by contrasting them with those undertaken by WP4 based on observed proxies and images.

Project Results:

WP2 - Irradiance Data Exploitation

Task 2.1 Collection of space era SSI data and relevant solar proxies

The main task here was to collect SSI and solar proxy data that are available to the scientific community. This includes also historical data, some of which were not readable. In the initial database the original data in their native data format were collected. The most important datasets that were also used for further processing (data screening, outlier removal, and updated uncertainty estimates, data merging) were converted into netcdf format, which is a common data format used in the climate modelling community. The initial database includes three different data types of observations: SSI time series, SSI reference data, and solar proxy data. SSI reference data are single spectra that have been recorded once or were assembled from few measurements to form a wavelength composite. Here the focus is on reference data...
assembled from few measurements to form a wavelength composite. Here the focus is on reference data that are or were considered as radiation standards in the past or until today. SSI timeseries data also include composite datasets that merged different satellite data records and are published (de Land et al., 2008). An initial overview of available SSI datasets is given by Ermolli et al. (2013). The initial database was transferred to PMOD and is publically available on the SOLID web page including graphical interfaces for displaying the data.

Task 2.2: SSI data analysis and quality checking
A global study of all solar spectral irradiance datasets has been performed. Historical datasets obtained from pioneer missions (e.g. San Marco or OSO missions) have been disregarded because of the too large uncertainties on the observed irradiance variability, or insufficient documentations. All other datasets have been through the same processing where outliers were detected and all «bad» points (outliers, gaps, data with uncertain time or wavelength, ect.) were replaced using a proxy model fitted to the observations on the short time scale (proxy models have very good agreement with data on rotational time scales). This is described in Scholl et al. (2016).

An independent precision, intended as the uncertainty on the daily repeatability, of observed SSI, has then be estimated. The precision is computed using an adapted Donoho estimator based on wavelets; it provides a good estimation of the 1-sigma dispersion of the data, even for non-white noise. Additionally, the precision is increased when the data are suspected to be corrupted (outliers or filled missing data).

Finally, a method has been developed to estimate the stability, intended as the uncertainty on the long time scale repeatability, of each time series. It is based on two approaches which rely on proxies in different ways. The first approach determines the link between irradiance variations and proxies at the rotational time scale, where proxies allow to reproduce irradiance variations with a good precision. This link is then extrapolated to the cycle time scale, and the residuals between the observed and simulated cycle variations are fitted using long-term variations of the proxies. The difference in the slope of the residuals and their fits provide the stability in %/year for each data points. The second approach consists in fitting the data with a multiscales multi-proxies data. This approach is more permissive as the combination of the long-term variations of several proxies can also reproduced artefacts in the data. A mix of the two approaches is made to determine the final stability.

Although these uncertainties are not determined from the instrumentation, they allow a measure to intercompare the SSI time series and as such should be useful for the users not familiarized with all the details of the datasets. They are also used as weights when making the bayesian composite.

The final database contains two data products: 1) the original (untouched) observed solar spectral irradiance with flags and estimated precision and stability, and 2) the original observed SSI corrected for outliers and missing values, together with precision and stability.

Additionally to this global approach on all datasets, specific work was done in order to improve the calibration of specific datasets. This was in particular the case for the ISS/SOLAR (Thuiller et al., 2013, 2014a,b), PICARD/PREMOS (Cessateur et al., 2016), and the SOHO/CDS (Del Zanna and Andretta 2015) datasets.

Task 2.3: Ground-based proxies for SSI
The sunspot number is one of the most used solar proxies; however, it takes only into account the number of spots and groups of spots while other information are available from these historic observations. For this project, we extended a very detailed sunspot catalogue developed in the context of a previous
For this project, we extended a very detailed sunspot catalogue developed in the context of a previous FP7 projects (SOTERIA, COMESEP), after a thorough survey of available sunspot data since the end of the 19th century. For the early 20th century, we use the Brussels local USET (Uccle Solar Equatorial Table) catalogue that stems from drawings starting in 1940 and, in parallel, the well-known Royal Greenwich Observatory data (RGO: Willis et al., 2013a,b). This work gave birth to the first version of a merged catalogue with detailed information on active regions from 1874 to the present, with an increased level of detail as we near the present time, allowing us to obtain extended information on the active regions, their positions, evolution and complexity. In parallel between SOLID and the TOSCA cost ES1005 action, a collaboration with J. Vaquero, V. Carrasco, Maricruz Gallego and A. Aparicio, (UNEX, Spain) enabled the acquisition and exploitation of a Spanish sunspot dataset covering a part of our catalogues that facilitated the inter-comparison between catalogues. In this context, two articles have been written with the Spanish team: Carrasco et al. (2015) and Lefèvre et al., (Solar Physics, accepted in Mar. 2016).

A quality assessment of these datasets has been realized in 2015 and a more extended intercomparison work of the different catalogues has been realized to determine the error bars associated to the different measurements. This assessment on overlapping periods enables us to determine a confidence level on the measurements.

Extensive analysis was also done on the widely used Mg II index and lead to two major improvements: on the one hand, a disagreement between the two available Mg II composite (« LASP » and « Bremen ») was solved (Snow et al., 2014) and, on the other hand, the Bremen Mg II index is now daily provided with its uncertainty (http://www.iup.uni-bremen.de/UVSAT/Datasets/mgii).

Task 2.4: EUV spectral irradiances
Extensive work has been done to provide irradiance time series in the extreme ultraviolet range from 1998 to 2014, which represents the longest EUV time series after the SOHO/SEM data. The SOHO/CDS spectrometer data were used to built mosaic of the solar disk and careful analysis lead to well calibrated irradiance time series in several transition region and coronal wavelengths. The analysis required a new calibration for the local degradation of the spectral lines, and the overall degradation of the instrument, as detailed in Del Zanna and Andretta (2015). The CDS irradiance data were then compared with other EUV experiments (e.g. sounding rockets, SOHO/SEM, TIMED/SEE, SDO/EVE). These comparison lead to significant improvements in our overall knowledge of the cycle variability at these wavelengths. Previously, discrepancies of factors of two to three in the measured irradiances of even the strongest lines were common. With the latest revisions of the various instruments calibrations, good agreement (to within an overall 20-30%) has been achieved. In particular, the new CDS and a revised calibration for the SOHO/SEM first-order band now show agreement to within 10% in the past few years. During solar minimum, this band is dominated by the strongest EUV line, the He II resonance line at 30.4 nm (detailed of this comparison are shown in Del Zanna et al. 2015). Good agreement with most of the SDO/EVE data (except a few strong lines and the last few years) was found, whilst the TIMED/SEE EGS data suffered a strong degradation. Overall, the SOHO/CDS has performed exceptionally well with only a factor of two degradation during its lifetime, and provided the most accurate EUV irradiances. The transition region lines were found to vary very little over the solar cycle, while hotter coronal lines formed around 3 MK showed variations of about 40. Unfortunately, the short-wavelength channel of the SDO/EVE spectrometer failed in 2014, and the SOHO/CDS instrument was switched off because of lack of funding, so accurate measurements of EUV irradiances will not be available in the future.
Measurements of EUV irradiances will not be available in the future.

Task 2.5: UV spectral irradiance composite

Several new composites for solar proxies and SSI from the existing database were created during this project. Updated proxy data are the sunspot numbers (see Task 2.3 above), several radio fluxes (Dudok de Wit et al., 2014), and the Bremen composite Mg II index (Snow et al., 2014). The Bremen composite Mg II index now provides uncertainty estimates that accounts for the uncertainty of the individual time series (noise level) and uncertainties due to the scaling and merging of the of the various datasets to form the composite (manuscript in preparation).

These three solar proxies are the most useful ones and correlate best with SSI variations on rotational (~27d) and solar cycle (~11y) time scales. While in the EUV the faculae contribution (e.g. Mg II index or radio fluxes) dominates, in the near UV both sunspot darkening and faculae contribute significantly (see also Task 4.6). For solar proxy based reconstruction of SSI generally a two component model (faculae and sunspots) is needed as a minimum as done, for instance, for the NRLSSI model (Lean et al.).

From the available SSI timeseries datasets, two types of SSI composites were created. The first is based on merging the available individual SSI timeseries by matching and scaling the various datasets in overlapping periods. Some degradation effects were corrected with the aid of solar proxies. It can be considered as an update on and improvement of the de Land et al. (2008) composite. Detailed comparisons show some significant improvement, but due to limited availability of data in certain periods, not all issues with regard to optical degradation (drifts) affecting to at least some degree all the UV satellite data could be completely resolved. The second SSI composite is based upon an empirical reconstruction and is described in Task 4.1.

Task 2.6: Recommendation for space-based SSI and proxy time series

Here we concentrate on the UV and Visible bands, which are of prime importance for climate, although deliverable D2.6 addresses other bands as well. In the EUV band, radiometric stability and calibration are major issues. All instruments heavily suffer from degradation (e.g. TIMED/SEE) and contamination (e.g. PICARD/PREMOS). SoHO/SEM stands out, thanks to periodic rocket calibration flights. In spite of this, several of the data sets are sufficiently good to enable quantitative tests with EUV models. In the UV band, instruments suffer less from degradation except for the strong Lyman-alpha line at 121.6 nm. However, the signal-to-noise ratio is also lower because of the lower solar variability. Generally speaking, SSI observations are well understood below approximately 200nm-240nm. Above these wavelengths, the uncertainty on the stability of the measurements prevents to accurately measure the low amplitude of the solar cycle. This is true for basically all instruments. Several considerations can be done, however:

The actual SORCE/SIM measurements, showing a larger UV variations compensated by an out-of-phase variations in the visible for Solar Cycle 23, appear very unlikely for several reasons: 1) this behavior is not reproduced in Solar Cycle 24 in the same observations, 2) it contradicts other observations (e.g. Wehrli et al., 2013; Marchenko et al., 2014), and 3) Models are unable to reproduce this behavior without braking the agreement at other times. These observations are still being studied by the instrument teams, in particular to reach a better understanding of long-term instrumental effects; their good performances on short time scale let us hope that reliable measurements of the solar cycle in the NUV and visible can still be reached.
Several SSI datasets have been corrected using solar proxies (NOAA/SBUV instrument series, at least one version of the SOLSPEC data) or by doing some assumptions on the variations of the SSI (for ex. constant SSI during solar minimum for AURA/OMI). Although highly valuable, it is important to remind that these datasets are not purely independent measurements, especially when long-term variations are concerned.

WP3 - Multi-wavelength solar image processing for novel SSI proxies

Task 3.1: Investigation Phase

A first investigation phase at the beginning of the project was dedicated to collect all available knowledge on the feature identification schemes and tools available in WP3. The requirements of other work packages that require the results of this WP were also investigated to ensure cohesion, usability and usefulness of the delivered products.

Task 3.2 Developing automated imaging tools for analysing SOHO magnetogram and continuum images

This Task is concerned using image data recorded by the SOHO/MDI satellite in white light intensitygrams and magnetograms to provide information for use in estimating the amount of radiation emitted by the sun over the period of the data. This will supplement other methods of estimating the emitted radiation, which are not necessarily all available at the same time, and lead to more reliable combined estimates. The approach of our Work Package is to start with pairs of intensitygrams and magnetograms, taken where possible one pair per day over the lifetime of the SOHO/MDI programme. Each of almost 2500 such pairs from April 1996 to April 2011 is used to estimate the fraction of the area of the solar disc occupied by each of seven different characteristic features, such as sunspots and network. These feature area fractions are calculated individually in eleven different concentric regions between the centre of the solar disc and the limb, known as filling factors. These filling factors, calculated for each available day, are then to be used as inputs to a model which calculates solar radiation.

The framework for calculating the solar features has been provided by the Automated Solar Activity Prediction (ASAP) tool, an existing system designed to estimate solar features for the purpose of predicting solar flares. This has been extensively modified with new capabilities to produce a version suiting the requirements of the Solid work package.

The estimation of features requires simultaneous intensitygram and magnetograms, but these are generally captured at different times. It is necessary therefore to calculate synchronised versions of the closest available images in time, for input to ASAP. This process was verified by synchronising intensitygrams to later intensitygrams and comparing the synchronised result with reality.

This work package has required ASAP to be extended with a several enhanced segmentation algorithms developed to detect and calculate the area coverage of several magnetic features on the solar disc by making use of and combining information extracted from synchronised intensitygrams and magnetograms. ASAP based automated segmentation algorithms were also developed to segment PICARD/SODISM images in five narrow pass bands from 215 nm to 782 nm. However, only data from for the 535 nm pass band, from August 2010 to January 2014, were available for detailed processing.

The features are provided as binary images of the solar disc showing feature and background. The number of feature pixels within each individual radial region is summed and to increase the accuracy of the filling factors, the calculations are done to sub-pixel accuracy. Feature pixels lying on a boundary between radial regions are split appropriately between the two regions. Multilinear regression fits between VIRGO data and MDI feature data showed that there is a good correlation (0.88) between the two. Comparisons
data and MDI feature data showed that there is a good correlation (0.88) between the two. Comparisons between sunspot features extracted from the MDI and the SODISM images showed 0.98 correlation. Similar comparisons with PSPT archived sunspot data also showed 0.98 correlation with the MDI features.

Task 3.3 Developing imaging tools for analysing SOHO EUV images
As described in Haberreiter et al. (2014), the SPoCA tool automatically segments an image into 6 solar features, or classes, that refer to: coronal features Coronal Holes (CH), Quiet corona (QS1), Coronal quiet network (QS2), Coronal active network (AR1), Active Regions (AR2), Bright active regions (AR3). In Haberreiter et al (2014), SPoCA provides the value for the QS2 class, and the five other class centers are determined using the relative contrast of the images as determined by the synthetic spectra calculated with the SolMod code (Haberreiter, 2011). For each of these features the relative area coverage is determined as a function of time.

To determine the uncertainty of the image segmentation, we made a statistical analysis on the class centers as determined by the SPoCA tool. We used about 11000 EIT images in each wavelength 171Å and 195Å, covering the time period January 1998 to December 2008. First, we consider the 6-dimensional vectors of class centers, and cluster those 11000 vectors using hierarchical trees. Four groups of class centers (and thus of images) can be distinguished: class centers corresponding to the minimum of solar activity (1), class centers for ascending and descending phase (2), class centers for solar maxima (3) and outliers (4), ie image featuring mainly proton storms.

For the subsequent analysis we kept only groups #2 and #3. For each class separately, we then fitted a Gaussian Mixture Model (GMM) of its centers. The fixed SPoCA class center are chosen as the main mode of the GMM, and the corresponding standard deviation is used to determine a 1-sigma error bar on the class center. We then compare these SPoCA class centers with those obtained with the SolMod ratio using the QS2 value as benchmark. We find that SPoCA and SolMod class centers agree pretty well (largely within the 1-sigma confidence interval). For the QS1 class however, SPoCA finds a larger value for the class centers as compared to SolMod. Finally, using the main-mode class centers and the standard deviation class centers, the complete EIT archive in 171Å and 195Å was processed using SPoCA, and the area coverage of each class is computed. The SSI reconstruction based on the area coverage still needs to be done, and compared with previous reconstruction using SolMod-based class centers values.

Task 3.4 Developing imaging tools for analysing ground-based solar images
The main results derived from the study of the segmentation of full-disk ground-based solar observations can be summarized as follows. We developed and then applied image processing tools to PSPT (1996-present) observations, in order to measure the photometric properties of seven classes of solar disk features, including sunspot umbra and penumbra regions, plage and facular region, enhanced and quiet network features, internetwork. We computed maps showing the location and extent of seven classes of solar disk features that were identified on the archival PSPT observations. We delivered time series of the filling factors of solar features from the segmentation of the Rome-PSPT archive, in order to provide semi-empirical models with input data for reconstruction of SSI variations at various bands, including EUV. The image segmentation and determination of filling factors required the development of dedicated software tools that allowed us to reliably estimate various quantities for the identified features in processing the full-disk ground-based solar observations. Results were provided into FITS format files for the segmentation maps, and ASCII files for the times series of filling factors in the eleven rings. The adequacy of the image
maps, and ASCII files for the times series of filling factors in the eleven rings. The adequacy of the image segmentation applied to the archival data for the purpose of SSI modelling rests on the homogeneity of the quality of the analysed observations, which turned out to be important for accurate reconstructions of SSI variations measured from space. We tested the accuracy of time series obtained against findings from different image processing methods applied to same observations.

All the results obtained (Chatzistergos et al., in preparation), besides than providing new information and new tools for the processing of full-disk ground-based solar highlight the importance of continuing the efforts to improve the results of the data processing observations to the purpose of improving SSI models.

Task 3.5 Developing imaging tools for analyzing PROBA2 EUV images

SWAP has a spectral band-pass centered on 174Å and provides images of the low solar corona over a 54x54 arcmin field of view with 3.2 arcsec pixels. For the purpose of EUV irradiance reconstruction, we select a dataset containing one image every two hours from 1 Jan 2010 till 16 September 2014. In order to provide a consistent dataset, we selected Sun-centered images with an exposure time of 10s. Moreover, blurred and dark images are discarded by statistical analysis. Finally, the images are corrected for various instrumental effects, including the Point Spread Function, using the algorithm described in (Seaton et al., 2013), and a correction for limb brightness enhancement is performed. We use the SPoCA algorithm (Verbeeck et al., 2014) to classify each image from our dataset into six classes: Coronal holes (CH), Quiet corona (QS1), Coronal quiet network (QS2), Coronal active network (AR1), Active Regions (AR2) and Bright active regions (AR3). Similarly to what was done in (Haberreiter et al., 2014), we use the median of the therefore found classification center values for (QS2), limited to the year 2010, still a period of quiet solar activity, and use it as our reference value. Next, the SolMod ratio, computed using SDO/AIA bandpasses, are applied to the reference class to fix the value for the other five class centers and to segment our dataset using SPoCA. Fillings factors for each of the six class were computed using the segmented images in view of feeding models for reconstruction of EUV solar spectral irradiance. But as it turns out during the course of the project, due to lack of photometric calibration of the PROBA2/SWAP images, the class centers computed using SolMod ratio did not correspond to the actual solar features as seen by the SWAP telescope, and therefore PROBA2/SWAP images are not well suited for irradiance reconstruction.

Task 3.6 Developing imaging tools for processing PICARD images and data

For this task images taken with the SODISM instrument onboard the PICARD mission were analyzed. PICARD was commissioned in October 2010 and decommissioned in April 2014. The fact that the PICARD’s operational phase would end already in 2014 was not known at the start of the SOLID project. Of key interest for the SOLID project were the images taken at 215nm, 393nm and 535 nm. The UV channels are important to identify bright regions on the solar disk, while the visible channel is important for the sunspot detection. However, at the time of the SOLID project phase, only the 535nm images were made available at Level 1B, i.e. the data product that includes the important correction for instrumental issues. Overall, the main challenge for this task was that quality of SODISM raw images revealed an optical aberration that blurs the images in a non-static way. This was a major challenge for the SOLID team when processing these images. The outcomes of this task shows that only sunspot filling factors are suitable for future analysis. Due to the limitation of only having sunspot filling factors a subsequent SSI reconstruction based only on the SODISM analysis was not possible.
Task 3.7 Providing higher level products by inter-calibrating the proxies produced in previous tasks
In order to understand in detail differences in the results of the segmentation approaches we performed a comparison of the filling factors obtained within WP3. Three filling factors datasets, which were produced as part of the SOLID activity, have been compared. The filling factors were obtained by processing SOHO/MDI, PICARD/SODISM 535, and Rome-PSPT/Ca II K observations. An excellent agreement with a high correlation was obtained for the sunspot filling factors obtained by all the datasets.
It is also important to generate error estimates for the generated filling factors. We do not have multiple sets of equivalent solar images which could be used to repeat the calculations and so generate error estimates from variation in results. However, we do have measures of the uncertainties in the intensitygrams and magnetograms estimated from multiple datagrams. Following ideas presented in the Bootstrap Method, suggested by [Efron and Tibshirani, 1994], we have generated multiple (50) sets of intensitygram and magnetogram pairs from each pair of available data with different realisations of the known uncertainties. The variations in resulting feature estimates have been used to estimate the uncertainties in the features and filling factors generated by ASAP for 30 different intensitygam and datagram pairs taken at different times in the data record.

WP4 - Modelling of spectral irradiance based on observed proxies

Task 4.1: Reconstruction of SSI from empirical models
The main output of SOLID is a composite SSI dataset that incorporates all instrumental records from WP2; for that reason, Task 4.1 is at the very heart of the project. This observational composite is now ready, and can be downloaded from the SOLID website (http://projects.pmodwrc.ch/solid/). This dataset has daily values from 1 November 1951 to 31 December 2014, with 915 wavelength bins from 0.5 to 2925.5 nm; the bin size is 1 nm up to 500 nm, and larger above. A detailed description of the input data is given in [Schöll et al., 2015], while the composite itself will be discussed in Haberreiter et al., (2016).
The main issue in making such a composite is the sparse coverage of the time-wavelength space, which has numerous data gaps. In addition, when several instruments happen to be observing simultaneously, they frequently disagree on the long-term evolution of the SSI. Other difficulties are the variable spectral resolution (e.g. broadband radiometers versus spectrometers), incomplete temporal coverage (especially for the early missions), and observations that lack radiometric accuracy and stability. For that reason, we used solar proxies to bridge data gaps, and constrain the variability. These proxies are the solar radio flux at 3.2 8.0 10.7 15 and 30 cm wavelength, and the daily sunspot area. They were chosen for their stability, and their high correlation with specific bands of the SSI.
The composite is built as follows
1. The input data are pre-processed and corrected for outliers (see WP2).
2. The precision (i.e. the random error caused by short-term fluctuations) is estimated by wavelet transform.
3. Missing data are filled in by expectation maximization (Dudok de Wit, 2011).
4. The interpolated data sets (and their standard error) are then decomposed into different time scales by using a translation-invariant wavelet transform. We thus obtain for each wavelength, as many records as there are instruments observing, multiplied by the number of time scales (typically between 8 and 12).
5. For each time scale, and each wavelength, the composite is an average of all existing observations, weighted by their inverse variance. This estimator of the composite is optimal from a maximum likelihood perspective. Its uncertainty is obtained by error propagation. Missing observations that were interpolated in step 3, are not used in this averaging, except when no observations are available, in which case proxies
in step 3, are not used in this averaging, except when no observations are available, in which case proxies are all we have left.

6. The scale-dependent composites are finally merged into a single composite, and its standard error. Large differences in the absolute values of the SSI do not preclude a small uncertainty in the composite, because we decouple precision (relative error) from accuracy (absolute error). Thanks to the multi-scale approach, observations that differ by an offset only, or some trend, can easily be merged into a high-quality product.

On average, the composite agrees better with the observations in the UV band because of the higher solar cycle, and solar rotational variability. In the visible and near-infrared bands, this variability is considerably lower (and actually often buried in instrumental noise), and there are less observations available. For that reason, the composite in the visible and near-infrared relies more heavily on proxies to reproduce solar variability. This gives satisfactory results for sub-decadal time scales. For multi-decadal time scales, however, the composite is constrained by the lack of secular variation in any of the proxies, so that no long-term changes can be generated. The most sensible way for allowing the composite to include secular changes, would be to include in it SSI models. In SOLID, we decided not to do so, by lack of clear framework for assigning weights to the observations and to the modelled data. Cleary, assimilating both modelled and observed SSI in the composite should come next on the list of tasks to do (see Outlook Section).

We performed extended comparisons between the composite and SSI models (SATIRE-TS and NRLSSI2), and also with individual observations. When the input data are have stability problems (e.g. SIM/SORCE in the MUV band) and when no alternative observations are available, then the composite can only rely on the inputs it has, and will thus provide a sub-optimal SSI record. This is the price to pay for having a dataset that relies strictly on the observations only, with additional user input. A cure to this is to enhance manually the uncertainty on the input data, based on external information.

TSI observations are formally part of the SSI. However, because they have received considerably more attention, and expectations are high regarding their weak long-term variability, we processed TSI data separately, by merging them without relying on any proxy data, or SSI model inputs. In addition, we developed a more elaborate merging method that properly excludes the influence of interpolated values in the merging. All of these are detailed in [Dudok de Wit et al., 2016], which focuses on the TSI only.

There exist today three TSI composites, which show marked differences in their long-term behavior. All three use one single instrument at a time to build the composite, which is far from optimal. Subtle as they may be, the differences between these composites have a direct impact on the long-term variation of solar radiative forcing, and on models such as NRLSSI2, which are trained on these data. The new SOLID composite will eventually replace the three existing ones. However, the official version is not ready yet as depends on corrections that ought to be made in some of the original TSI datasets. This is a rather political decision, which is being investigated now by team at the International Space Science Institute (ISSI, Bern). Interestingly, for the first time we have been able to estimate the time-scale dependent error on the TSI without making any model assumptions. The results reveal a $1/f$ (where $f$ stands for frequency) scaling of the noise power, which definitely departs from the usual white noise assumption underlying most studies.

Task 4.2 Comparative Study and improvement of the atomic data used in SolMoD3D and CHIANTI

A comparative study between the EUV/UV spectral radiances/irradiances as observed by the various instruments and as obtained with modelling and the CHIANTI atomic data has been carried out. This has allowed us to assess which atomic data needed improvement and which were missing. New atomic data...
allowed us to assess which atomic data needed improvement and which were missing. New atomic data for a range of ions have been produced, and have been included in CHIANTI version 8 (Del Zanna et al., 2015). The new version 8 is a major update with a lot of new datasets and a new format, although the data have also been produced in the older format, for easy inclusion into SOLMOD. This study was based on the quiet-Sun spectrum obtained by a SDO-EVE prototype during the solar minimum of 2008. All the main spectral lines have been identified and measured, and a Differential Emission Measure (DEM) modelling carried out, exploring a range of parameters. Excellent agreement between observed and predicted irradiances is found. The completeness of the atomic data has been checked against higher-resolution datasets, from SOHO CDS, SUMER and Hinode EIS QS observations, where many of the stronger lines are not blended as in the EVE spectra. Many weaker lines await identifications and atomic data, but for the irradiance, the stronger lines are all covered.

Task 4.3 Improvement of atmosphere models for the chromosphere, transition region and corona

The Lyman-alpha line is the strongest line in the solar spectrum. Therefore, it is important to understand its formation to a high degree. In the case of Lyman-alpha, the microturbulent velocities influence the shape, i.e. the steepness and the width, of the central line profile, thereby also determining the distance of the double peak from the line center. Here we address fitting of the Lyman-alpha profile to observations with as little change to the FAL99 A atmosphere structure as possible. In order to improve the formation of the wings of the Lyman alpha profile we adopted a change of the chromospheric temperature profile, in particular the temperature gradient has been reduced to 3/4 and 1/2. Using the updated velocities along with the changed chromospheric temperature structures leads to a significant improvement of the Lyman-alpha profile.

We have completed the analysis of the EUV spectral irradiance for the quiet Sun with the updated CHIANTI v.8 atomic data. During the WP2 work, significant problems were found in what was thought to be the best quiet Sun EUV irradiance spectrum, obtained with a well-calibrated clone of the SDO/EVE spectrometer flown in 2008 on a sounding rocket. The advantage of this spectrum is that it observed the Sun during minimum across a wide spectral range, from 70 to 1040 Angstroms. Nearly all the lines in this range are formed in the transition-region and corona and are optically thin, so simple emission measure modelling has shown very good agreement (20%) between observed and predicted EUV irradiances.

Task 4.4 Reconstruction of the SSI spectrum based on filling factors from solar image analysis

The key achievement of Task 4.4 was the reconstruction of the EUV spectral wavelength range. For the reconstruction of the solar EUV irradiance we use the semi-empirical approach as described in Fontenla et al. (2009) and Haberreiter (2011, 2012). Synthetic solar intensity spectra are calculated with the Solar Modeling (SOLMOD) code for five components of the intensity distribution of the solar corona: quiet corona (Quiet Sun 1), coronal quiet network (Quiet Sun 2), coronal active network (Active Region 1), active regions (Active Region 2), and bright active regions (Active Region 3). The coronal hole intensity has been calculated with a fixed contrast of c=0.5 with respect to the quiet corona (QS1). Additionally, we also calculate the intensity as emitted from the chromosphere and transition region (TR) based on the atmosphere structures by Fontenla et al. (2009). The level populations for the ions with charges up to two are calculated with radiative transfer calculations accounting for non-local thermodynamic equilibrium (NLTE). The higher ionization stages are calculated based on the coronal models and the optically thin approach, which means that the radiation field is neglected in the calculation of the statistical equation. For the corona the emergent intensity spectra are calculated in spherical symmetry, which allows us to realistically calculate the emission beyond the solar limb. This is particularly important for the extended...
reasonably calculate the emission beyond the solar limb. This is particularly important for the extended corona, in particular for the sharp increase of the coronal emission at the limb. The optically thin coronal spectrum is then added to the spectrum from the chromosphere and transition region. This leads to the total emission in the EUV. The segmentation analysis of the SOHO/EIT images has been provided by WP3 from Partner ROB, for details see Task 3.3. SOHO/EIT images taken in the 171 Å and 195 Å channels over the time span of over 16 years, i.e. from 4 March 1996 to 29 February 2012. The comparison of the reconstructed SSI time series with SOHO/SEM data shows that the overall solar cycle variation is captured very well, the results have been presented by Haberreiter et al. (2014).

The next achievement of Task 4.4 was the reconstruction based on the analysis of PROBA2/SWAP images. The image segmentation employed for the reconstruction presented here has already been described in Deliverable D3.4 Filling factors catalogue for PROBA2 images, provided by ROB as part of WP3. In particular, the images have been segmented using the SPoCA tool (Verbeeck et al, 2014). For the analysis we use images obtained from 1 Jan 2010 to 16 September 2014. We have carried out the reconstruction of the EUV between 6 and 20 nm for the time frame from from 1 Jan 2010 to 16 September 2014. In the earlier part of the time series, i.e. up to July 2013, the reconstruction matches very well the daily minimum value of LYRA match. After that time the reconstruction seems to overestimate the LYRA observations. One reason for this deviation could be that the degradation correction does not fully account for the loss of sensitivity in the EUV part of the channel. The results are currently being prepared for publication by Haberreiter et al. (2016). As it turns out, the SWAP images are difficult to calibrate. In particular, we encountered the problem that the image segmentation and the subsequent reconstruction is very noisy. Therefore, more understanding of the SWAP detector is needed in order to use the dataset for a robust irradiance reconstruction.

The third point addressed in Task 4.4 was the irradiance reconstruction based on the PICARD SODISM instruments. However, some issues were encountered regarding the SODISM data. First, due to the shutdown of the PICARD satellite in April 2014, the length of the SODISM dataset turned out to be very limited. Second, the available image are not routinely calibrated and processed. One of the big challenges of the SODISM images were ghosts are on the image which need to be corrected. A lot of effort has been put into correcting these ghosts, mainly for the 535-nm channel, which is accounted for in the Level 1B data for that wavelength. However, all other channels are only available as Level 1A. In summary, the combination of the early shutdown of the PICARD satellite, along with a very intermittet and sparce processing of the raw SODISM images lead us to conclude that these data were not of great value for irradiance reconstructions.

Task 4.5 Intercomparison of the results with reconstructions from WP5
This task aims at inter-comparing the SSI reconstructions from WP4 with the ones from WP5 with a focus on the UV, visible and IR part of the spectrum. The reconstructions presented in this task are based on the one hand on the image segmentation of the PSPT and MDI data as provided by WP3. The synthetic spectra used here are calculated with the COde for Solar Irradiance (COSI, Haberreiter, et al., 2008; Shapiro et al., 2010, Tagirov et al., 2016). We carried out a detailed comparison of the SSI reconstruction using the COSI spectra and the PSPT and MDI filling factors, respectively, with the SOLID composite and the SATIRE-S reconstruction. We conclude that the reconstructions presented here are very promising. However, some discrepancies were identified for the longer wavelength range above 250 nm, which need to be further investigated. In particular, a study that links the feature identification from intensity images at different wavelength and the identification from magnetograms would be of great value in order to further understand solar spectral irradiance variation. It is exactly for this purpose that the analysis carried out...
understand solar spectral irradiance variation. It is exactly for this purpose that the analysis carried out here will be of value for future investigations.

Task 4.6: Definition of new solar proxies for irradiance modelling

The paucity of direct and long-term SSI observations has stimulated a continuing search for proxies of solar activity. Best known for the UV band are the radio flux at 10.7 cm (F10.7 index), and the Mg II core-to-wing index. The reconstruction of solar variability in the visible and near-infrared bands is considerably more challenging.

In SOLID, we first performed detailed comparisons between EUV observations from SOHO/CDS and SDO/EVE, and various solar proxies, using the correlation coefficient as a metric. The MgII and the radio flux at 30 cm (F30 index) came out as better proxies. This work was then extended to the full spectrum, using either instrumental records (e.g. from SORCE) or the composite SSI dataset from 1990 till today. No single proxy can adequately describe a spectral band on both long and short time scales. Nevertheless, there are three outstanding candidates:

The Mg II index, which remains the best proxy for the UV below 300 nm, with good performance both for solar rotational variability (i.e. ~27 days) and solar cycle variability.

The 30 cm radio flux (F30) closely follows the Mg II index in terms of performance, with the advantage of being available back to 1957.

The F10.7 index remains a good all-purpose UV proxy, with good performance for the EUV band (10-121 nm).

On decadal time scales, the Mg II index and the radio fluxes agree for all but the last solar minimum (in 2009, between cycles 23 and 24), for which the Mg II index shows a markedly deeper dip. This difference highlights the extreme care by which proxy reconstructions can be extrapolated to multi-decadal time scales.

None of the 9 proxies we considered adequately reproduces the SSI above 300 nm. However, empirical models involving a linear combination of proxies describing facular brightening (such as the MgII index) and photometric sunspot darkening (such as the daily sunspot number), are known to work well. These are further addressed in WP5. Likewise, proxies that are derived from segmented solar images offer a high potential for accurately describing specific bands of the SSI. This approach is already used in models such as SOLMOD [Haberreiter et al., 2014] and SATIRE-S [Yeo et al., 2014], and is further addressed in WP3.

WP5 - Solar irradiance reconstructions using simulated magnetograms

The main objective of WP5 was to reconstruct time series of daily total and spectral irradiance variations that span more than a century and include pre-industrial times. This was achieved by combining reconstructions from different versions of the SATIRE (Spectral And Total Irradiance REconstruction) model that use the best input data available for each time period: sunspot numbers (1700 to 1878), sunspot areas (1878 to 1974) and full-disc images (since 1974). The main new development carried out within WP5 of the SOLID framework was the development of SATIRE-T2 (described below) that uses simulated magnetograms as input to reconstruct irradiance at times before 1974. The simulated magnetograms provide the spatial distribution of active regions (sunspot and faculae) and their decay products (network) on the solar surface. They thus enable a better description of the irradiance variations on timescales of weeks to months than older reconstructions that use only sunspot numbers. The new irradiance dataset is described in Dasi-Espuig et al. (2016) and is available through the website of the Sun-Climate group at the Max-Planck Institute for Solar System Research (www2.mps.mpg.de/projects/sun-climate).
The steps towards the final data product can be split into four main tasks that are the (i) development of SATIRE-T2, (ii) generation of simulated magnetograms, (iii) testing of the irradiance modelling to uncertainties in the model parameters and (iv) reconstruction of the solar irradiance.

(i) SATIRE-T2 was developed in the first year of SOLID. In common with all SATIRE-type models, the underlying assumption is that irradiance variations on time scales longer than a day are caused by changes in the distribution of the surface magnetic flux. SATIRE-T2 takes simulated magnetograms as input to account for magnetic flux emerging in active regions and models separately the flux that appears in ephemeral regions (see Dasi-Espuig et al. 2014).

(ii) The simulated magnetograms were obtained using a surface-flux transport model that describes the evolution of magnetic surface flux due to differential rotation, meridional flow, and surface diffusivity (Cameron et al. 2010, Jiang et al. 2011b). The sources of magnetic flux are the active regions. Their emergence on the solar surface is described by the observed record of daily sunspot group positions and areas for times after 1874. For times before 1874, we use a statistical model that reflects the observed relationships between the strength of a cycle (measured using the sunspot number) and sunspot characteristics at the time of emergence such as position, area and tilt angle (Jiang et al. 2011a). The simulated magnetograms were validated by comparisons to measurements of the total surface flux (available for the past 35 years) and to empirical reconstructions of the open flux that go back to the beginning of the 20th century (Jiang et al. 2011b, Dasi-Espuig et al. 2014).

(iii) Different sources of uncertainty enter the irradiance reconstructions. These include uncertainties in the input data (e.g. sunspot areas and numbers), in the comparison data (e.g. in the total solar irradiance or in the magnetic flux) and in the fitting process that is used to determine the free parameters of the model. SATIRE-T2 has 4 free parameters that are fixed by minimising the discrepancy between the observed and modelled total solar irradiance and total magnetic flux from solar cycles 21 onwards, when comparison data become available. Accounting for the uncertainties in the fitting process and exploring different input and comparison data sets, we are able to provide an uncertainty range for our reconstructions, provided that there was no change in the way the solar dynamo has operated (see Dasi-Espuig et al. 2016).

(iv) The final irradiance reconstructions comprise daily irradiance values for wavelengths between 120-nm and 160-μm at a wavelength resolution of 1nm in the UV, 2 nm in the visible and increasingly wider bins at longer wavelengths. Irradiance values from 1974 to 2015 are derived with SATIRE-S using observed magnetograms and disc images from Kitt Peak, MDI and SDO as discussed in Yeo et al. (2014). This dataset has been tested extensively (see also Yeo et al. 2015) and provides a consistent description of the solar total and spectral irradiance changes. For periods between 1878 and 1974, solar irradiance values were obtained with SATIRE-T2 using the sunspot area and location as input sources to simulate solar magnetograms. At even earlier times, magnetograms are derived using semi-synthetic records of sunspot area and latitude (Jiang et al. 2011a). These magnetograms represent hypothetical surface maps that are statistically consistent with what one would expect to see given the observed sunspot number. The resulting irradiances are therefore not intended to be accurate daily irradiances, but rather their monthly average is expected to be reliable. In addition, the variability pattern is typical of what is expected for the Sun for the given phase in its activity cycle.
Comparisons between observed spectral irradiances and those modelled either using observed or simulated magnetograms during the last three solar cycles show that the simulated magnetograms give an accurate statistical representation of solar variability. In addition, we find an excellent match between rotational and cyclic variability amplitudes calculated with SATIRE-S and SATIRE-T2 and good agreement with spectral irradiance measurements where reliable comparison data are available.

Task 6.2: Statistical analysis of irradiance time series
The most important SOLID recommendation is that SOLID observational composite should be used to constrain SSI models. The novel approach taken in our work is a probabilistic one, in which, at each time step, we consider a weighted average of the different observations. These weights are set by the uncertainty of the observations, for which we thus need a metric that can be consistently applied to all the observations; for details see Haberreiter et al. (2016). The SOLID composite is based on 20 instruments and one reference spectrum (due to the high uncertainty in the other available reference spectra). In addition to SSI observations, 6 solar proxies are also considered. While these proxies do not strictly correspond to SSI observations, they are known to properly reproduce SSI variations in given bands, and for specific time scales. The 10.7cm radio flux (or F10.7 index), for example, has been widely used so far as a proxy for the EUV band, whereas the 30cm flux is better suited for reproducing the solar rotational variability at longer UV wavelengths. Furthermore, we have investigated the variation of the observational composite and compared it to the variation of the SATIRE-S, NRLSSI1 and NRLSSI2 models for the annual means of 2003 to 2008 as well as for 1984 to 1996; for details see Haberreiter et al. (2016).

To demonstrate the effects of SOLID spectra on the atmospheric energy balance, we performed calculations with the high resolution model LibRadtran, which is a library of radiative transfer equation solvers widely used for UV and heating rate calculations. These set of calculations are presented in Haberreiter et al. (2016). The solar cycle signature in heating rates is characterized by two distinctive maxima in the stratopause (50 Km) and upper mesosphere (above 70 Km), resulting from the larger solar fluxes available for absorption of O3 and O2. The SOLID dataset introduces a solar-cycle related heating of about 0.12 K/day at the stratopause, which is considerably stronger than NRLSSI2 and SATIRE-S but weaker than SORCE. The difference in heating rates between SOLID and SORCE, however, is not statistically significant, if the 2 standard deviation in the SOLID irradiances is taken into account. The ±2 sigma uncertainty of heating rate anomalies ranges from 0.08 to 0.2 K/day.

References:
Ashamari, O., Qahwaji, R. Ipson, S., Schöll, M., Nibouche, O., and Haberreiter, M., 2015, Identification of photospheric activity features from SOHO/MDI data using the ASAP tool, Journal of Space Weather and Space Climate, 5, A15
Potential Impact:
Impact of the SOLID Project
Thanks to strong interactions with the user communities, part of the SOLID team is now involved in a publication (Matthes et al., 2016) on the recommendation of the CMIP6 model intercomparison for the forthcoming assessment report of the IPCC. The SOLID observational SSI composite is a key contribution to constrain the CMIP6 SSI recommendation. As such, the SOLID project can be considered a full success. In addition, the SOLID community was involved in a number of outreach activities:
1) increased activity of presentations on SSI variability and its effect on the Earth’s atmosphere and climate at international conferences, such as EGU, AGU, AOGS.
2) Strengthening of links, collaboration, and involvement of EU scientists to programmes of the International scientific community dedicated to the study of solar influence on climate, such as the SOLARIS/HEPPA, key programme of SPARC/WRCP under the auspices of WMO/UNEP, allowing thus for greater visibility of the project results.
3) part of the SOLID consortium was involved in the Handbook The Earth’s response to a changing Sun.
3) Part of the SOLID consortium was involved in the Handbook The Earth's response to a changing Sun written within the COST Action TOSCA, and the TOSCA Training Schools.

4) Involvement in International teams at the International Space Science Institute:
   (i) Scenarios of Future Solar Activity for Climate Modelling (lead T. Dudok de Wit)
   (ii) Solar Heliospheric Lyman Alpha Profile Effects SHAPE (lead M. Snow)
   (iii) Towards New Models of Solar Spectral Irradiance based on 3D MHD Simulations (lead S. Criscuoli)

SOLID has also been instrumental in improving the scientific exploitation of a large ensemble of disparate SSI records for use in climate and in space science, in a context that had been almost devoid of coordinated efforts. As an example, during the course of the project, the NASA-funded MUSSIC project as well as several smaller initiatives dealing with SSI reconstructions were initiated.

A less visible, but strategically important impact is the experience gained by the SOLID team. We started as a team in which most partners had hardly interacted before. Three years later, the team has become an essential partner in the community of SSI science.

The making of the composite SSI dataset, however, is not finished. This is continuing process, as new data are being produced, physical models are improving, and the needs of the user community become more specific. The experience gathered by our team is invaluable for defining more durable SSI products that can be used by broader user community, and put us in a strong position for shaping a European strategy. For that reason, the team is determined to pursue this coordinated effort, and maximize the return from the expertise, and from the SSI data.

Main Dissemination Activities within the SOLID Project

(i) To set up and maintain the Project Webpage, in order to provide the user communities latest updates of the project, providing introduction to a better understanding of Solar Spectral Irradiance, listing all activities of SOLID, conferences and meetings, positions available within SOLID, publications, news and recent developments

(ii) This web page hosts the Projects database, created continuously update so that of irradiance products are provided and widely distributed, taking into account the scientific community needs as identified via web-based online surveys.

(iii) With this in mind, a survey to extended user community, serving both the need to communicate the Project's Aims and Scopes to the wider scientific community as well as getting their feedback was conducted. This latter task is closely related to the survey conducted under WP6, on the needs of the Climate and Chemistry-Climate modelling communities, who are the first users of SOLID's results

(iv) An Annual Newsletter was promptly issued, available via mailing lists and the Project's webpage, communicating recent developments and the Projects latest news to a larger audience

(v) A crucial task was the participation at International conferences where the aims, scopes and developments on the Project were introduced to the scientific communities. In this respect, a number of talks were given at a wide range of conferences, such as the European Space Weather Week (in 2013, 2014 and 2015), the International CAWSES-II Symposium 2013, DACA2013 Atmosphere and Cryosphere Assembly, the 5th HEPPA/SOLARIS Workshop 2014, the SCC2015 Conference on Sun Climate...
Assembly, the 5th HEPPA/SOLARIS Workshop 2014, the SCC2015 Conference on Sun-Climate Connections 2015, IUGG 2015, and EGU2014 and 2015, where members of the SOLID consortium presented their work, gave keynote invited talks and convened sessions with a view to communicate further SOLIDs scientific results.

(vi) A session on Solar Irradiance Variability in EGU 2016, convened by members of the SOLID consortium will further disseminate the final project results. In addition, results will be presented in 6th International HEPPA-SOLARIS Workshop.

(vii) Central focus in our dissemination planned tasks holds the outreach towards students. To fulfill it, SOLID as a consortium was heavily involved in teaching and co-organizing a COST training school on “Solar variability and climate response”, held in ICTP, Trieste in 2014. The main organizer of this training school was the EU COST Action Project TOSCA, the strong ties to which were mentioned above, together with ICTP.

(viii) One of the major tasks was dissemination via the organization of an open workshop, where SOLID results were presented to a wider audience of end-users, with a view of getting their reaction, comments and feedbacks. This was achieved by the organization of the Workshop on “Long-term variations of solar activity and their impacts: From the Maunder Minimum to the 21st century”, held in autumn 2014, jointly with the COST Action TOSCA. One of the main aims was to discuss lessons learned from past initiatives (such as the Solar Model Inter-comparison Project (SolarMIP), a SPARC/SOLARIS initiative to compare the coupled ocean-atmosphere model response to variability in solar irradiance in the CMIP5 (Coupled Model Intercomparison Project phase 5), but, foremost, give recommendations on the solar forcing to be used in the new simulations in support to the Climate Model Intercomparison Project phase 6 (CMIP6) in support to the next planned IPCC Report. Therefore, it was a unique opportunity to communicate SOLID’s products and discuss its potential role as a solar forcing dataset for use in climate models. The key messages from the participants, implications and suggestions for the SOLID data sets were valuable inputs to the final preparation of the irradiance data sets. One element of crucial importance to SOLID results in order to ensure the attention of the scientific community were the requirements presented and discussed for an irradiance time series to be used as the solar forcing in CMIP6 (detailed description in D8.3).

(ix) And, last but not least, one key point of the dissemination was the participation of SOLID scientists to the Working Group meeting of SOLARIS-HEPPA in autumn 2015, to present the finalized time series, discuss and promote visibility of the project’s results. In this way, along with the other actions mentioned above, we were able to strengthen the links, collaboration, and involvement of EU scientists to programmes of the International scientific community dedicated to the study of solar influence on climate.

Exploitation of Results:

1. Survey of the climate community
A survey of the climate community users’ needs, relevant to the project outcomes was the main objective of this task, starting with a compilation of ongoing and past studies where total and spectral solar irradiance were used in various modelling projects (e.g. CCMVal and CMIP5), followed by a survey on the needs of the scientific communities.
The first survey of chemistry and climate-chemistry was conducted in the period May-mid July 2013. The survey collected responses from a representative sample of modelling groups and end-users that exploit general circulation models for climate research.

The main conclusions of this survey are summarized here:
(i) A spectral resolution of 1 nm is sufficient for broad-band parameterizations but a considerably finer resolution is required for photolysis heating rates.
(ii) Increased spectral resolution in the Schumann-Runge band and near Lyman-a line would certainly be an asset.
(iii) The climate modelling community is highly interested in a new reference solar spectrum.
(iv) The climate modelling community could make use of reference solar spectra for the solar minimum and maximum conditions of the 11-yr solar cycle.
(v) Finally, it would be best to deliver SOLID products in both NetCDF and ASCII formats.

A follow-up survey of a similar nature addressing the needs of scientific communities including, but not limited to, thermospheric, planetary and lunar research, was conducted during August and September 2013. In short, it can be concluded that survey participants are highly interested in reference solar spectra representative of the 11-yr solar maximum and minimum conditions and in a recommendation of a reference solar spectrum. Near-real time spectral solar irradiance products will be an asset to the project. This is consistent to the needs of the climate model community as expressed in the first survey.

2. A review of recent model studies that considered total solar irradiance (TSI) and spectral solar irradiance (SSI) variability was conducted during the first months of the project. It revealed that while early studies simulated climate response due only to TSI changes, in recent years climate models include sophisticated SSI and ozone variability. The upward extension of the model’s domain to include a detailed representation of the stratosphere and mesosphere is considered now as a prerequisite to simulate realistic solar signals at the surface. Solar cycle signals in the stratosphere, however, strongly depend on the radiation scheme implemented in every GCM. The comparison of the radiation schemes used within the Chemistry Climate Model Validation initiative (CCMVal) showed that not all chemistry climate models were able to capture the amplitude of the heating rates between solar maximum and minimum as calculated by line-by-line radiative transfer models. The success depends mainly on the number of radiation bands used in the spectral range where ozone absorption dominates. For instance, the FUB-EMAC radiation scheme with 44 spectral intervals between 121 and 683 nm reproduced the reference heating profile over the whole stratosphere whereas the radiation scheme of ECHAM4 with 2 only bands simulated negligible heating changes in the stratosphere. Most of past studies used TSI and SSI variability according to the NLRSS model, which was also recommended for the Chemistry Climate Model Initiative (CCMI).

3. SOLID webpage and database:
The SOLID database can be found under [http://projects.pmodwrc.ch/solid/index.php/main-database](http://projects.pmodwrc.ch/solid/index.php/main-database). The goal of the visualization tool of the SOLID Database is to disseminate the final reconstruction time series covering the full spectral range, from the XUF to the radio wavelength. Furthermore, the data sets of being used within SOLID will also be made available.
Within SOLID, the following data sets are being processed:

i) spectral irradiance time series
ii) reference spectra
iii) proxy data sets
iv) segmentation maps

These data sets are also reflected in the set-up of the SOLID Database. The idea is that the database will be continuously updated. With the acceptance of Haberreiter et al. (2016), the final product of SOLID, i.e. a recommended SSI time series that covers the full wavelength range and does not contain any temporal and spectral gaps will also be provided through the database.

4. Above all, the observational SSI composite dataset that has been produced by SOLID will be extremely valuable for understanding how to combine disparate observations, and test SSI models, as well as solar proxies. The goal is that the observational SSI data set produced in this project is incorporated and used extensively in a number of atmosphere and ocean models, including also models with coupled chemistry processes. A thorough evaluation of the results and comparisons to observations will help not only to assess differences, weaknesses and strengths of the models, but also to advance knowledge and further understanding of the physical mechanisms that govern the influence of the Sun to the Earth’s atmosphere and climate. A first step was taken in this project, with the assessment of atmospheric effects via a high resolution radiative transfer model. Our future goals are to proceed further with coordinated modelling experiments within the scientific user community.

List of Websites:
http://projects.pmodwrc.ch/solid/

Contact:
Margit Haberreiter,
PMOD/WRC
Dorfstrasse 33
margit.haberreiter@pmodwrc.ch

Last update: 14 July 2016
Record number: 186940