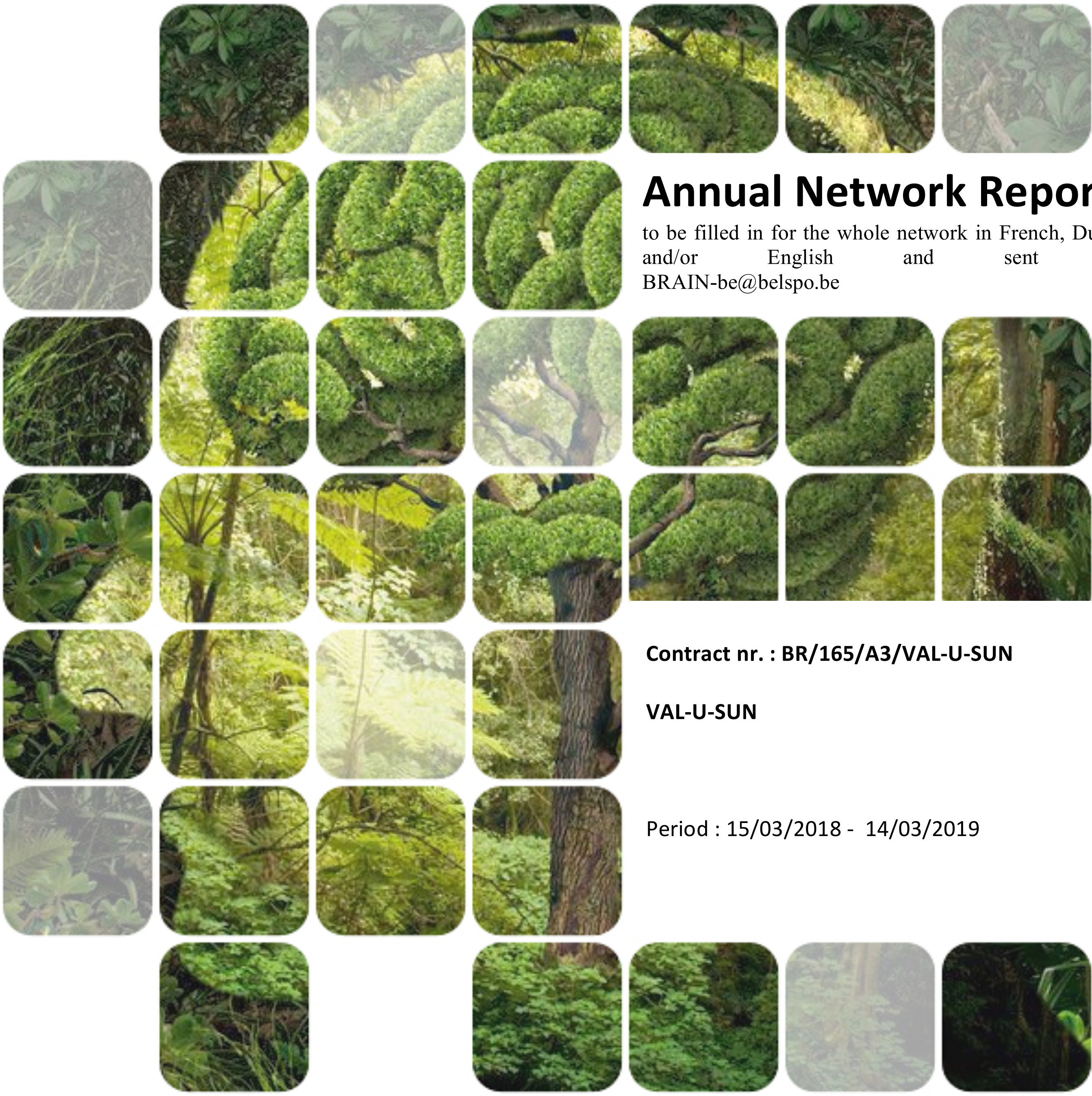
BRAIN-be

BELGIAN RESEARCH ACTION THROUGH INTERDISCIPLINARY NETWORKS



Annual Network Report

to be filled in for the whole network in French, Dutc and/or English and sent t BRAIN-be@belspo.be

Contract nr. : BR/165/A3/VAL-U-SUN

VAL-U-SUN

Period : 15/03/2018 - 14/03/2019

NETWORK

COORDINATOR

LEFEVRE, Laure, Royal Observatory of Belgium, SILSO WDC
Function: Researcher
Address: Avenue Circulaire, 3, 1180 Uccle

Phone: +32 2 790 3923
E-mail: laure.lefevre@oma.be
www: http://sidc.be/valusun/

PARTNERS

1. **von Sachs, Rainer**
   Université Catholique de Louvain - ISBA
   Function: Professor, Researcher
   Phone: +32 10 478806
   E-mail: rvs@uclouvain.be

2. **Sophie Mathieu**
   Université Catholique de Louvain- ISBA
   Function: PhD Student
   Phone: +32 10 478806
   E-mail: soph.mathieu@uclouvain.be

3. **Dudok de Wit, Thierry**
   Université d'Orléans
   Function: Professor, Researcher
   Phone: +33 (0)2 38 25 52 77
   E-mail: ddwit@cnrs-orleans.fr

AUTHORS

1. Lefèvre, Laure
   Function: Researcher
   Phone: +32 2 790 3923
   E-mail: laure.lefevre@oma.be

2. Delouille, Véronique
   Function: Researcher
   Phone: +32 2 790 3938
   E-mail: veronique.delouille@oma.be

PROJECT WEBSITE:

http://www.sidc.be/valusun/

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY OF THE REPORT

Sunspots have been counted since the invention of the telescope in the early 17th century. In 1844 it was discovered that time series of the sunspot count show a periodic pattern of approximately 11 years. This became known as the solar cycle, which deeply affects the entire solar system. The solar magnetic field embedded in sunspots is the driving force behind the solar variability that on a day-to-day basis influences the space environment of the Earth. The **focus of this project** is to open up the ROB Sunspot Data Collections to the international research communities and to contemporary scientific methods by **contextualizing of the collections, updating them with new quality control mechanisms and making them available**. A multidisciplinary consortium has been set-up to achieve these goals, in which solar physicists at the Royal Observatory of Belgium collaborate with statistical experts at the Université de Louvain and with an international partner from the Université d'Orléans (France).

In this report we describe the tasks associated to the different Work Packages defined for VAL-U-SUN, that have been achieved and how they have been achieved, as well as the planning for the months ahead. Some tasks are lagging behind schedule, but others have been started ahead of time. All in all, the project is progressing as expected.

WP1 Meetings and reports
For this work package, the tasks were to organize different meetings and report on those meetings and on the evolution of the project. The coordinator plays the interface between the partners, the follow-up committee and BELSPO. T**ask 1.1** mentions this very annual report. For **task 1.2**, we describe the different formal and informal meetings that served the purpose of the VAL-U-SUN project during year 2. Part of the meetings were specifically organized for that purpose, another part of the work consisted in using programmed research meetings to meet VAL-U-SUN partners or follow-up committee members. For these meetings, we provide a description of the location, persons present and a summary of the minutes.

WP2 SILSO network documentation and performance analysis
The SILSO database contains the numbers of sunspots and sunspot groups since 1981. These numbers are provided by individual stations of the worldwide SILSO network and are used to compute the International Sunspot Number. Despite its success, the statistical stability of the International Sunspot Number needs constant concern. The 'averaging' process that combines the individual observations into one official International Sunspot Number was defined in the 1850s. The process depends on the choice of a reference to which an observation is compared. As the reference is provided by a single station called the 'pilot station', the International Sunspot Number is sensitive to any problem that could arise with that station.
During year 1 and year 2, we improved the network documentation, and the associated metadata in the database. This work is almost completed at the beginning of year 3, it is described in **task 2.1 in sections 2 (achieved work), 3 (intermediary results).**
The performance of stations is now checked regularly, but the computed Sunspot Number based on the SILSO database needs constant quality control. Through the alliance with the UCL/IASB and the SMCS, as well as the University of Orléans, we have developed a model for the quantities available from the database (number of groups and number of spots) to be able to monitor closely any solar and non-solar variations. This is described in greater details in **task 2.2 in sections 2, 3 and 5 (future prospects and planning)**.
Task 2.3 is focused on the documentation of the data specific to USET: it is described in sections 3 and 5.
Task 2.4 is focused on the quality control of the data from task 2.3, it is described in section 5.

WP3 USET drawing area measurements
At the end of year 2, near the beginning of year 3, we are in the process of adding contextualization and metadata to the almost 80 years of ROB/USET sunspot drawings acquired since March 1940. A first version of a more modern quality control will be realized during year 2 of the project. As a consequence, the USET drawing collection will be made accessible to the scientific community and should be used in scientific publications well before the end of year 4. This is described in sections 2, 3 and 5 in **task 3.1**.
Task 3.2 focuses on the automatic determination of the quality of the stations and hence the reference stations. As it was started earlier than anticipated, it is described in sections 2, 3 and 5.

WP4 Website, open access
The VAL-U-SUN website is on, it contains a 'news' sections that presents all results as they come in. **Task 4.1** and its future progress are described in sections 2 and 5**. Task 4.2**, focuses on the engagement of the different communities of users, and thus **task 4.3**, the installation of an online consultation platform will directly impact 4.2. Task 4.3 has been realized at the end of year 2. This is described in sections 2 and 5.

2. ACHIEVED WORK

Detailed description of the achieved work and tasks of the past reporting year

WP1:
Task 1 .1: VALUSUN Annual report

The present report is the first task of WP1 due Month 24.

Task 1 .2: VALUSUN meetings

During year two (March 2018-March 2019), a series of organizational and progress meetings were organized to follow-up on the project. In this section, we describe the meetings and a summary of the minutes is provided.

- *March 30th 2018: Progress Meeting.*

Present: L. Lefèvre, V. Delouille, R.von Sachs, S. Mathieu
Location: ROB
Minutes:
- Preparation of the Article for the IEEE meeting in June.
-*Noise Model discussions*.

- *May 8th 2018: Progress meeting.*

Present: L. Lefèvre, C. Ritter, R.von Sachs, S. Mathieu
Location: LLN
Minutes:
This meeting concentrates on the short-term model. We try to correct our model for the presentation of the poster on June 2018 and not spend too much time on new ideas.
Plan for the meeting.
1. Presentation of recent analysis on the conditional distribution (Update by C. Ritter).
2. Discussion on the short-term model (comments by V. Delouille).
3. Discussion on the referee remarks.

- *June 21st 2018: International meeting of the VAL-U-SUN project.*

Present: L. Lefèvre, O. Lemaître, V. Delouille, C. Ritter, R.von Sachs, S. Mathieu, F. Clette, T. Dudok de Wit.
Location: ROB
Minutes:
Choice of the scaling factors to rescale all observers/observing stations to the same observer (discussion with F. Clette). Proxy used for the solar signal (mean, median, more complex?). Choice of the moving average for the mid-term analysis. Discussions with T. Dudok de Wit on the distribution of the values of the different stations when the network gives a signal level s.

Problems raised:
Start thinking about the monitoring procedures.

- *October 3rd 2018: Progress meeting.*

Present: L. Lefèvre, V. Delouille, C. Ritter, R.von Sachs, S. Mathieu

Location: ROB
Minutes:
We decide to write two articles. The first one will contain the model, the distribution of the different terms and the scaling procedure. The second will focus on the monitoring. For the first paper, we aim to publish it in the Astrophysical Journal (https://iopscience.iop.org/journal/0004-637X), the second one will be submitted to a statistical review (to be determined).
Discussions about the choice of the time-scales for the scaling of the different observers to the network.

- *December 3rd 2018: Progress meeting.*

Present: L. Lefèvre, V. Delouille, C. Ritter, R.von Sachs, S. Mathieu
Location: LLN
Minutes:
Discussions about the conditional model of our data. Ordering of the article. Discussions about the estimator of the solar signal: either the median of the network or the filtered median of the network. (solution, create a simulation and check the distribution with, without filtering.). Discrete nature of Ns, Ng and Nc.
Composite Nc:We agree that it is very difficult to retrieve the distribution of the composite Nc = 10Ng + Ns from the linear combination of Ng and Ns as these two variables are dependent and have complex mixture distributions. We will analyze the conditional correlation of Ns and Ng when ISN = s Nc = s.

- *January 25th 2019: Progress meeting.*

Present: L. Lefèvre, V. Delouille, C. Ritter, R.von Sachs, S. Mathieu
Location: Skype
Minutes:
Discussions on the writing of the article to be published in the Astrophysical Journal.

WP2:
Task 2 .1:SILSO network documentation and metadata inventory

The inventory of which information was already collected from our network of observers is now done. We checked the email addresses of all our current observers. We are now looking for missing data in the most recent observers and contacting them to retrieve it, and will be looking into the historical part of the metadata, i.e. information on stations/observers that have stopped observing before the current period. In addition to the work planned for VALUSUN, we also endeavoured to make the SILSO database and especially the metadata, GDPR compliant, which is why this task is still ongoing.

Task 2.2: Performance analysis of SILSO network stations (due month 24)

In order to assess the performance of stations from the SILSO network, we have to determine what is the solar signal and what is “noise” associated to each observing station. Hence the first step was the determination of a valid noise model for the Sunspot Numbers of the different stations. A first description of this noise model was done in an article for the IEEE conference in June 2018. The model was progressively built over the course of year 2 and a publication will be submitted before the end of the first semester 2019 (target date is June 2019).

Task 2.3:USET documentation

The inventory of historical information available at ROB is still in progress at month 24.

WP3:
Task 3.1: Update of the USET database with area measurements (due month 24)

During year 2, we hired 5 student/months for summer jobs to make up for the lag in area measurements due to less hiring of job students in year 1. We have now completed 62% of the data with area measurements.

About 30 years of area measurements still need to be added. Last year's evaluation was a bit optimistic. The USET database should be completed with area measurements by the end of year 2019. To improve our ability to fill in the database, our colleague S. Bechet developed a whole new version of the Digisun software to make up for recurrent and time-consuming problems in the old version. We now have a complete set of 3 solar cycles (22-24) on which to realize a first version of the quality control. After that the quality control will be partially automated while the database keeps on growing.

Task 3.2: Automatic determination of SILSO quality measurements and reference stations (due month 48 – started month 13)

As mentioned in task 2.2, the determination of the model for our data is the first step. As this model is separated in different timescales, the long-term analysis is part of the monitoring of the stations. It is developed in section 3.

In addition to that, during 2018, a first analysis of different monitoring techniques was realized by our PhD student Sophie Mathieu. This analysis was presented at our meeting in October 2018. There are different procedures available for monitoring: from classical control charts (CUSUM, EWMA, Zhang, 1998, Qiu, 2013) to non parametric dynamic curve monitoring (Qui et al, 2011, 2013, 2014 2017).

We have to decide between all reviewed procedures, which are adapted to our specific problem (equally spaced data). We study three data sets: independent and identically distributed data randomly drawn from a normal distribution of zero mean and variance 1, the real data (with gaps filled using our partner T. Dudok de Wit algorithm) and simulated data randomly drawn from a Poisson-lognormal model.
The first step is the standardization. We implement the different methods and compare their performances. We study then the correlation of the residuals. Finally, we apply the CUSUM and EWMA charts on the normal residuals, we adjust the control limits and create artificial shifts to study their detection power. We also propose some solutions to fix the control limit and to evaluate the performance of the charts for real data.

The techniques studied here will be the subject of an upcoming article.

WP4:
Task 4.1: Website.

The VAL-U-SUN website is online. Results are published jointly on the SILSO website and VAL-U-SUN website. The open-access VAL-U-SUN website presents the results of this project whenever they are available. A first version of the network access to their metadata and the quality analysis exists at http://www.sidc.be/valusun/wolf/ We give access to all observers to their own metadata (not the metadata of the other stations for privacy reasons), and our analysis of their quality. The first version of this quality control process is the computation of a monthly k-factor that they can access from the website.

Task 4.3: Online Consultation Platform.

The VAL-U-SUN online consultation platform is now online (Figure 1). The online consultation platform has been developed in 2018, but the results will be coming in late 2019. The results will also be used by the UCL/IASB team in the context of the PhD project to check the validity of the model, as this will help materialize the human component of the observations of spots and groups of spots. Regarding schedule (Fig. 5), we are a few months late but we have time to study the results during years 3 and 4.

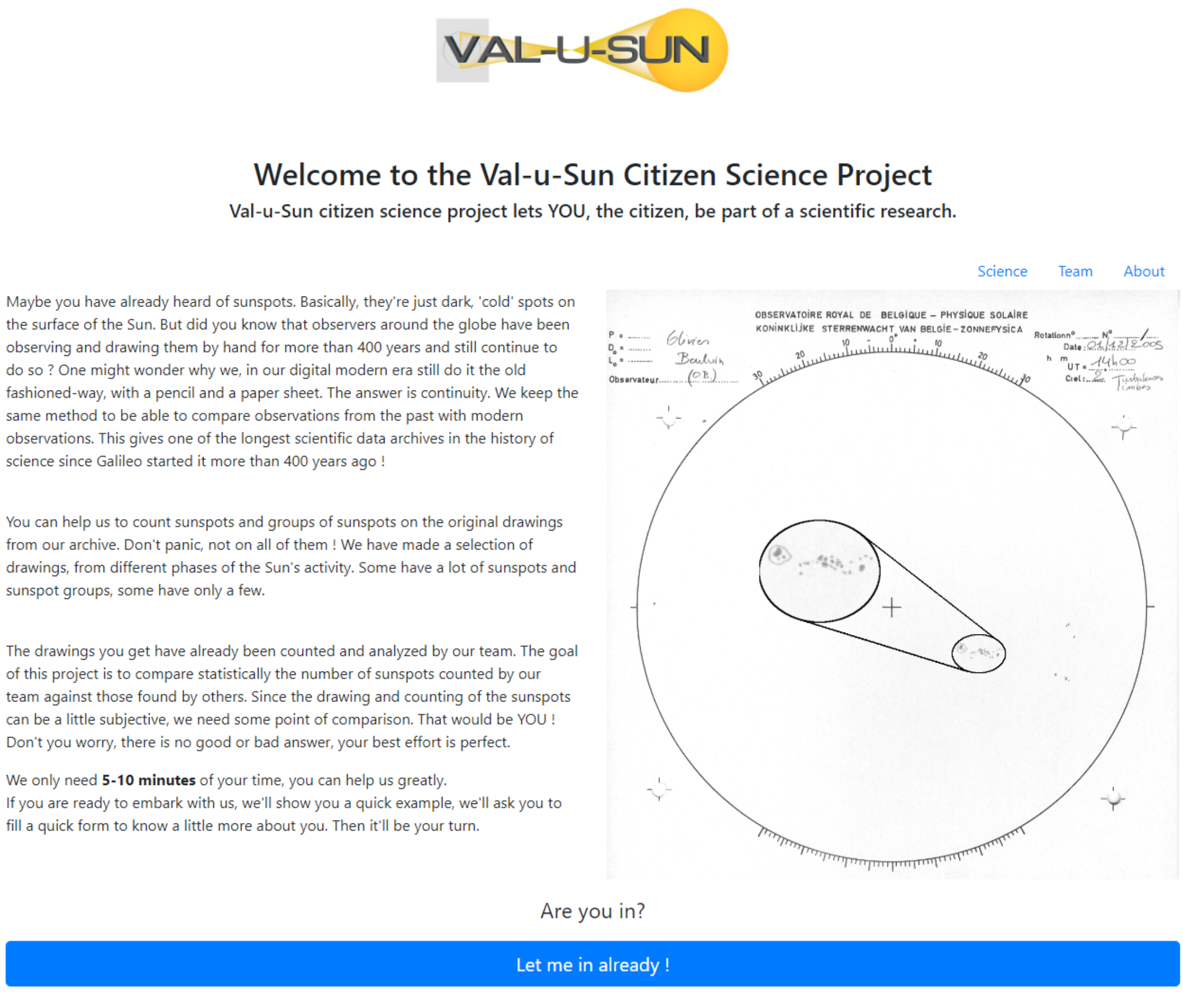


Fig. 1 Screenshot of the VAL-U-SUN online consultation platform developed by O. Lemaître

3. INTERMEDIARY RESULTS

WP2
Task 2.1: Metadata added and updated in the SILSO database

The database has been improved (field comments containing various metadata has been separated in many different usable fields) and results have been added (all email addresses of current observers have been verified). This updated metadata is in the database, but not open access yet.

Task 2.2: Development of a Model of the number of spots and number of groups, for the analysis of the Sunspot Number

In the context of our project there is an **article by Sophie Mathieu, Veronique Delouille and the VAL-U-SUN team** that has been published in the context of the **2018 IEEE Data Science Workshop**. This workshop aims to bring together researchers in academia and industry to share the most recent and exciting advances in data science theory and applications.

In this article, we build upon the work in Dudok de Wit et al. (2016), which presented a first uncertainty analysis of time domain errors and dispersion amongst the stations assuming a Poisson distribution. In the present paper, we propose a more comprehensive error model that accounts for all types of errors known to the experts, and fit them with distributions that comply with essential features of the number of spots (Ns) and the number of groups (Ng), namely their zero-inflated and overdispersed nature. We present the dataset, a proxy for the *true* solar signal, which is further used to estimate the various errors and our results on error estimation.

We have **a noise model:** $Y_i(t)=(\varepsilon_1(t)+\varepsilon_2(i,t))\, s(t)+\varepsilon_3(t)$ **(1)**

Where $Y_i(t)$ represents either the number of spot counts Ns or the number of group counts Ng or Nc the composite (Nc= 10NG+Ns) recorded by station i at time t. The term s(t) is a random variable (r.v.) representing the *true* solar signal, i.e. the actual number of spots or sunspot groups. We denote E (s(t))= $\mu_s(t)$ where E is the expectation sign.

Both ε_1 (t) and ε_2 (i,t) are multiplicative errors, as typically a station will count x% more (or less) sunspots than a reference station. To differentiate ε_1 from ε_2, we use the fact that they live on different time scales: ε_1 will be estimated on short time scale (< 27 days), while the mean of ε_2 will be estimated when short time scales are filtered out.
On longer timescales (typically with a moving average of 81 days) we can estimate the drift component as $\mu_2(t)$ (the mean of $\varepsilon_2(i,t)$).

Finally, an additive error term, identically distributed (i.d.) amongst the stations, is denoted ε3(t). It models essentially errors occurring during minima of solar activity, when there can be extended periods with no or few sunspots. ε3 thus captures effects like short-duration sunspots and non-simultaneity of observations between stations.

Now that we have the noise model, we have to describe it statistically speaking. Thus, instead of defining error bars, we have the complete distribution of the values.

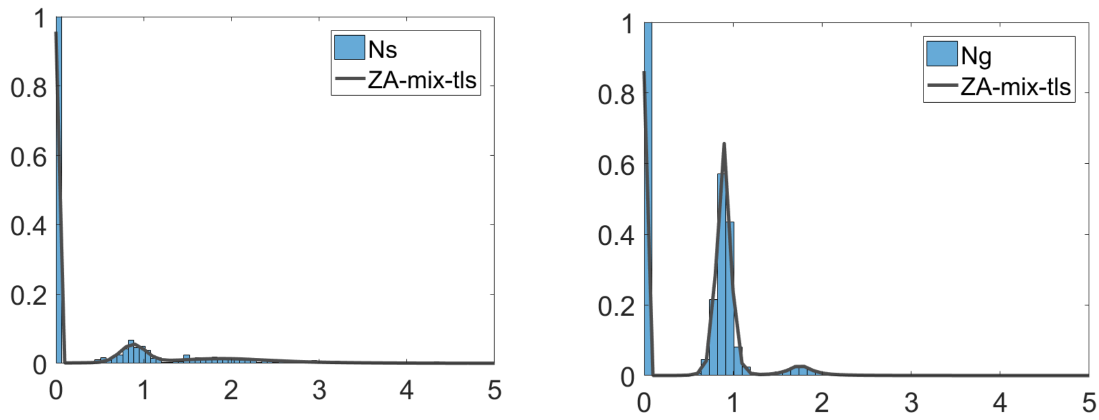


Figure 2: Histogram of $\varepsilon_3(t)$ for the counts of spots Ns (left) and groups Ng (right). The continuous line shows the fit using a ZA-mixture of tLS distributions. The binning (bw= 0.0917) is the same for both figures and is computed using the Scott procedure for Ns.

The probability distribution function (PDF) of $\varepsilon_3(t)$ may be described by a ZA-mixture of t-Location scale (t-LS), for Ns and Ng. They appear in Fig. 2. This distribution is a generalization of the Student distribution: it allows the modeling of a random variable with asymmetry and heavier tails than the normal distribution (Taylor & Verbyla, 2004; Evans, Hastings and Peacock, 2000s). The visual closeness between the histogram and the fitted distribution was used as a criteria to select the best PDF among few intuitive candidates. In Fig. 2, we immediately notice the mportant excess of 'true' zeros, together with dominant peaks around one and two for both Ns and Ng. These small excesses at zero represent the most represented value at ow regime: no spots. Our proxy for s(t) might be equal to zero even if some spots appear several minutes on the sun, due to the non-simultaneity of the observations between stations. It is indeed possible that only part of the network (observing the sun at a different time than other stations) has actually seen these spots. The exact correspondence between Ns and Ng is very interesting: it means that what we observe at that regime is simply groups of only one spot. Even if there are two groups, there are only two spots, i. e. one spot per group. This is exactly what we would expect at low regime of solar activity.

Task 2.3 USET Documentation

The inventory of all the historical information available in ROB logbooks, databases and annual reports has started during year 2. Extraction of all context information that can be of eventual use for the end data user is in progress.

WP3
Task 3.1: DIGISUN area computation progression

During 2018, we mostly focused on digitzing data between 1983 and the present to have 3 complete solar cycles on which to realize a first quality control at the beginning of year 3.

Year	Comment	Year	Comment	Year	Comment
1940			**TO BE DONE**	**1950**	**11 years**
1951	DONE	**1962**	DONE	**1973**	DONE
1952	programmed– 6 months	**1963**	DONE	**1974**	programmed– 6 months
1953	DONE	**1964**	DONE	**1975**	programmed– 6 months
1954	programmed– 6 months	**1965**	DONE	**1976**	programmed– 6 months
1955	DONE	**1966**	programmed– 6 months	**1977**	programmed– 6 months
1956	programmed– 6 months	**1967**	programmed– 6 months	**1978**	DONE
1957	DONE	**1968**	programmed– 6 months	**1979**	DONE
1958	DONE	**1969**	programmed– 6 months	**1980**	DONE
1959	programmed– 6 months	**1970**	programmed– 6 months	**1981**	programmed– 6 months
1960	DONE	**1971**	DONE	**1982**	programmed– 6 months
1961	DONE	**1972**	DONE		
1983	DONE	**1995**	DONE	**2007**	DONE
1984	DONE	**1996**	DONE	**2008**	DONE
1985	DONE	**1997**	DONE	**2009**	DONE
1986	DONE	**1998**	DONE	**2010**	DONE
1987	DONE	**1999**	DONE	**2011**	DONE
1988	DONE	**2000**	DONE	**2012**	Nearing completion
1989	DONE	**2001**	DONE	**2013**	Nearing completion
1990	DONE	**2002**	DONE	**2014**	DONE
1991	DONE	**2003**	DONE	**2015**	Started
1992	DONE	**2004**	DONE	**2016**	DONE
1993	DONE	**2005**	DONE	**2017**	day-to-day
1994	DONE	**2006**	Nearing completion	**2018**	day-to-day

Table 1: Area computation progression table. In red are the years that still have to be programmed for computation (no personnel assigned yet). In orange the years for which personnel has been assigned and which should be computed within 6 months. In green, years already done or in progress.

Overall 48 years out of a total of 78 have been added area measurements, 33 have been done in 2018. On average, a job student can compute 10 years of area measurements during a full month of work. This relatively slow progression for year 1 was compensated for during year 2 by the hiring of a maximum of summer students. We still have 30 years to go, which will be facilitated by a much more user-friendly DIGISUN developed by our colleague S. Bechet.

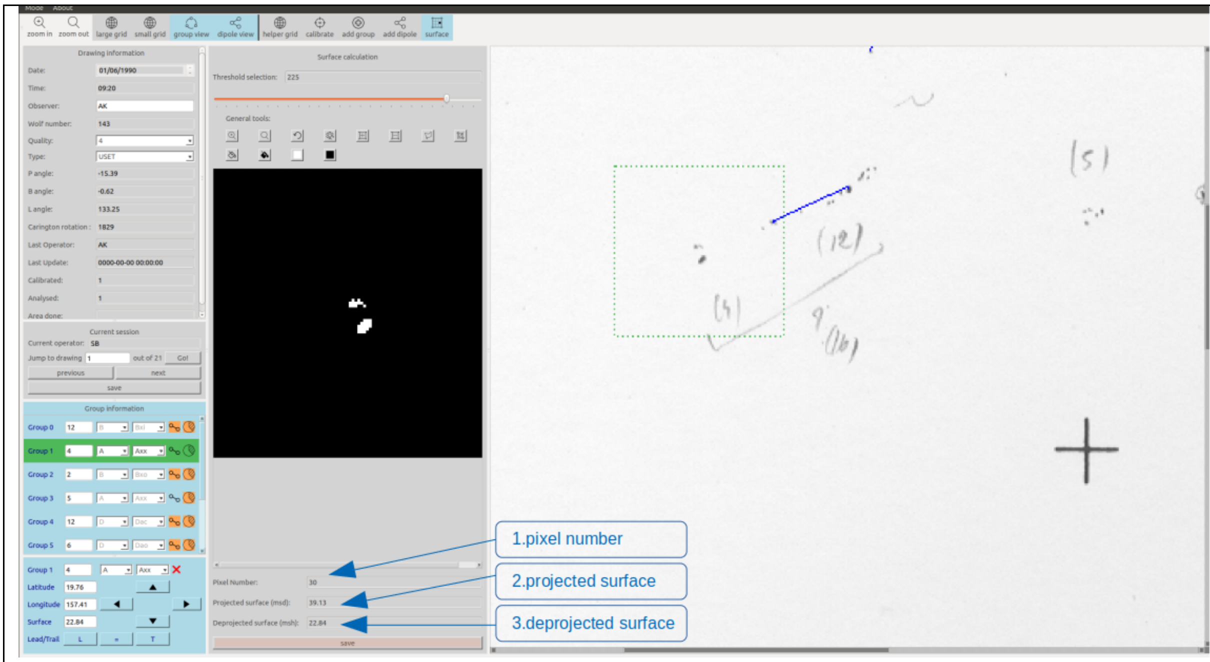


Fig. 3 Screenshot of the new DIGISUN developed by S. Bechet

Figure 3 above presents a screenshot of the new DIGISUN interface. It is in testing by the team. Its advantage compared to the old version is that it is much more user-friendly and thus much faster to use, for the computation of areas and the definition of regions on historical periods, but also for the USET observer on a day-to-day basis.
It is also developed in Python, which enables exportation to every platform in different environments (whether for amateur or professionals). It does not require any extensive coding knowledge to be installed and used.
This software is for all intents and purposes a valorization of the VAL-U-SUN project. Although we were not awarded the valorization grant by BRAIN, we have added value to our project.

Task 3.2 Automatic determination of SILSO quality measurements and reference stations

The long-term analysis of the work as described in task 2.2 will be used in the monitoring of the WDC-SILSO stations in the near future (i.e. 2020-2021).

Historically, the observers were rescaled to each other using what we call the k-factors. They were computed yearly and used to fill the gaps in the main observer's values (i.e. the director of the Zurich Observatory). In this study, we need a rescaling for the short-term analysis, of the stations. Thus, we decided to compute an improved version of the k-factors, using ordinary or total least squares or a weighted version of the k-factors. We also use a Kruskal-Wallis test to test on which length of time these rescaling factors have to be computed. This study leads us to an optimized time interval of about 11 months, which is very close to the 1 year used historically.
Of course, for the long term analysis, we want to assess a quantity similar to these k-factors, so the rescaling is not necessary.

We have defined a generic estimator for the long-term error:

$$\widehat{\mu_2}(i,t) = \left(\frac{Y_i(t)}{\hat{\mu}_s(t)}\right)^{\star} \quad \text{when } \hat{\mu}_s(t) > 0$$

where the * operator denotes the smoothing process, namely a moving average of at least 81 days, the ^ refers to a median. We choose $\mu_s(t)$ as the median of the network (all Yi(t) for N stations).

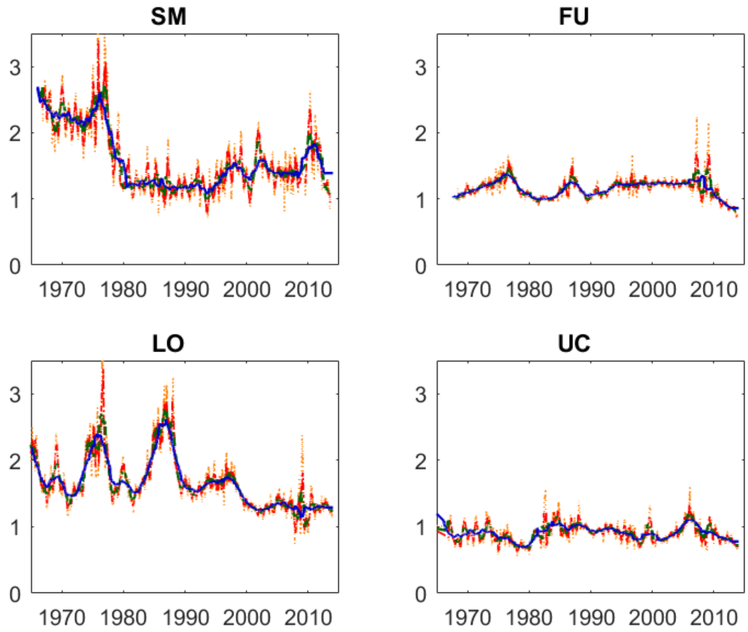


Figure 4: Estimation of the long-term drifts $\mu_2(i,t)$ of four stations (FU, LO, SM and UC). $\mu_2(i,t)$ is shown with different MA window lengths: 81 days (orange dotted line), 162 days (red dash-dot line), 1 year (green dashed line) and 2.5years (blue plain line)

Fig. 4 presents the long-term drifts of 4 different stations including our own USET station (code UC). These figures actually show the drift of the pilot station of the network, Locarno (LO) and the stability of USET. This method will be used in year 3 of the project to implement the monitoring of the stations.

References:

- J. Taylor and A. Verbyla. Joint modelling of location and scale parameters of the t distribution.Statistical Modelling, 4(2):91–112, 2004.
- M. Evans, N. Hastings, and B. Peacock, Statistical Distributions, John Wiley and Sons, 3 edition, 2000
- Zhang N. A Statistical Control Chart for Stationary Process Data. Technometrics, 40(1):24–38, 1998. 6.3.4
- Qiu. P. Introduction to Statistical Process Control. Texts in Statistical Science, CRC Press, 1 edition, 2013. 6.2.1, 6.2.2
- Qiu. P and Li Z. On Nonparametric Statistical Process Control of Univariate Processes. Technometrics, 53(4):390–405, 2011. 6.3.4
- P. Qiu and D. Xiang. Univariate Dynamic Screening System : An Approach For Identifying Individuals With Irregular Longitudinal Behavior. Technometrics, 56(2):248–260, 2014. 1, 6.3, 6.3.1, 6.3.4, 6.5.1
- P. Qiu, X. Zi, and C. Zou. Nonparametric Dynamic Curve Monitoring. Technometrics, 2017. 6.3, 6.4

4. PRELIMINARY CONCLUSIONS AND RECOMMENDATIONS

Overall the project is progressing as planned, although we had to adjust the timing for some tasks, we are still on course for the completion of the project on time. The management seems to be going as planned also.

WP2 and WP3 will undoubtedly progress even faster in the next years. The recommendation here is to monitor the progress of the PhD student closely, and organize progress meetings between V. Delouille, L. Lefèvre, R. von Sachs and Sophie Mathieu, as well as a few meetings with C. Ritter, T. Dudok de Wit and F. Clette.

For the DIGISUN area computation, we encountered some trouble recruiting job students during year 1, but recruited more during year 2 and plan 3 more during year 3. We are now more proactive and start looking for students that would be willing to apply to ROB summer jobs from our side before the deadline approaches. Also, due to some problems with the unfriendliness of the older version of DIGISUN, progression was slower. This is why we developed a new version of the software that is more user-friendly and exportable to multiple platforms.

The results are published on the VAL-U-SUN website regularly and more will be published as soon as they are available. During year 3, we will gather the results from the online consultation platform and interpret them for the different audiences: general public, network observers, scientists.

5. FUTURE PROSPECTS AND PLANNING

Overview of the foreseen activities and planning for next reporting year, taking into account the actual state of the work and the intermediary results

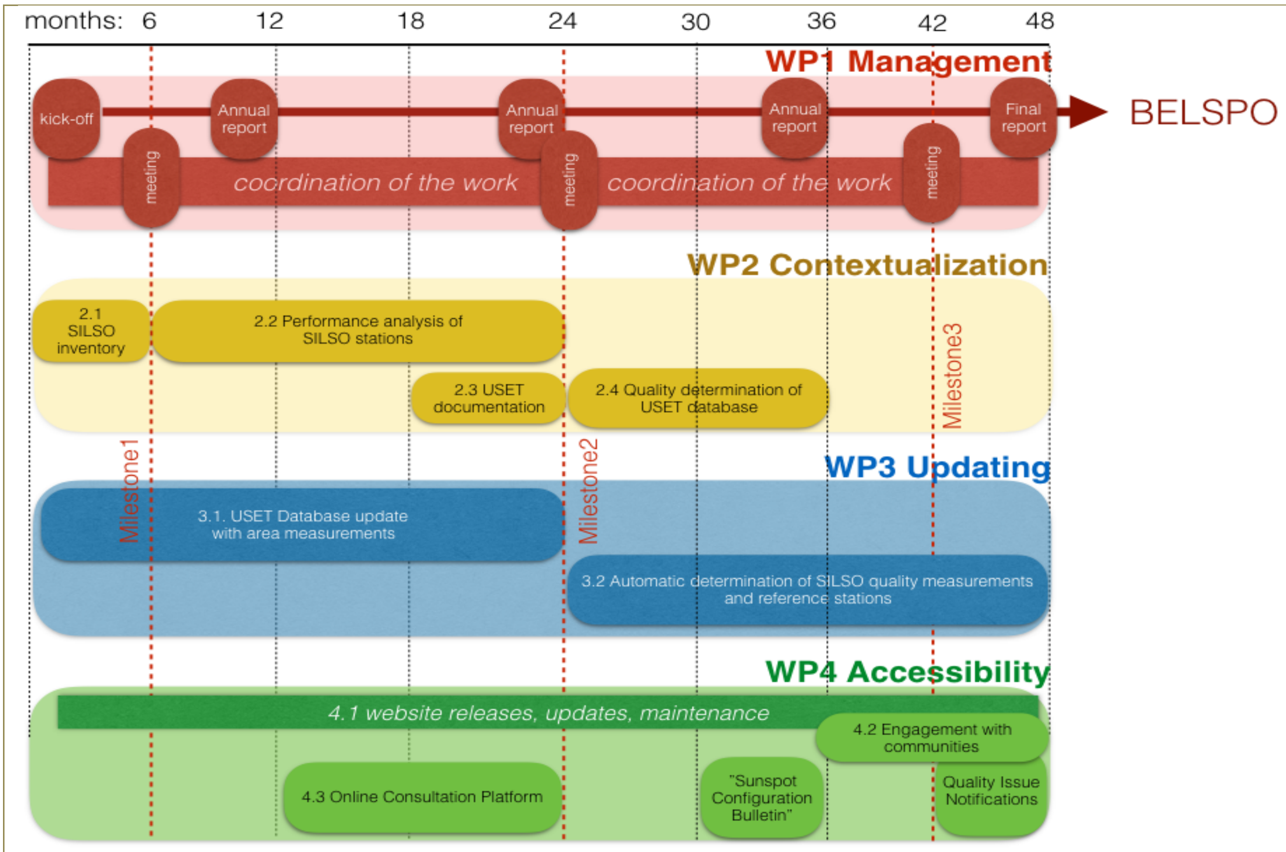


Figure 5: Work breakdown structure and overall planning of VAL-U-SUN

Figure 5 illustrates clearly what has been achieved and what needs to be done during years 3 and 4 We focus here on the tasks that have to be achieved or well under way at the end of year 2.

WP1: Management
Within Work Package 1, we still have to organize a yearly meeting of all participants and a few progress meetings with concerned parties, as well as meeting(s) with members of the follow-up committee. The coordinator will continue producing minutes of these meetings and annual reports at the end of each year of project. The next steps are the organization of a yearly meeting with all participants in June 2019 as well as formal and informal meetings with the follow-up committee at a meeting in July (SC7) and an ISSI meeting in August.

WP2: Contextualization of the ROB Sunspot Data Collections

Task 2.2 Performance analysis of SILSO network stations

The analysis realized in 2017-2019 (years 1 and 2) of the numbers for each station present in the database will be the subject of a scientific publication that will be submitted before the end of the first semester of 2019. The method developed on a subsample of stations still has to be applied to the whole database, but the methods have been developed.

Task 2.3 USET observatory documentation

The inventory and extraction has started as mentioned in the previous section. We still have to analyse the data (check + quality control) and publish it once it has been checked. This will be done in 2019.

Task 2.3 USET database quality control

The quality control of the USET database, and comparisons with catalogues from different observatories will be done in the course of 2019.

WP3: Updating of the ROB Sunspot Data Collections

Task 3.1 Updated USET database with area measurements

During year 3, we are planning to have at least 3x1month summer job students to make up for a lag in area measurements due to the unfriendliness of the old version of the software and a lack of recruitment in 2017. We can expect the USET database to be completed with area measurements in September 2019, and a first quality analysis soon after.

Task 3.2 Automatic determination of SILSO quality measurement and reference stations

As mentioned in section 2, the different techniques for the monitoring of the network will be analysed during year 3. An article will be written during year 3 and should be submitted at the beginning of year 4.

WP4:
Task 4.1 Open access data and documentation website
The open-access VAL-U-SUN website will present the results of this project whenever they are available. A new version of the SILSO network access to their metadata and the quality analysis will be introduced at http://www.sidc.be/valusun/. We give access to all observers to all non-privacy protected metadata from the network and our most recent analysis of their quality. The first version of this quality control process is the computation of a monthly k-factor that they can access from the website.

Task 4.3 Online consultation platform

The online consultation platform has been developed during year 2, and added to the VAL-U-SUN website. During year 3, a report centralizing the results will be written. The results will also be used by the UCL/IASB team in the context of the PhD project to check the validity of the model, as this will help materialize the human component of the observations of spots and groups of spots.

6. FOLLOW-UP COMMITTEE

Dates of the meetings and overview of the concrete contributions of the follow-up committee

Meetings with Follow-up committee members: Ed Cliver (SSN Workshops), Alexei Pevtsov (International Astronomical Union)

- *Informal: July 14-22d: COSPAR Pasadena*

Present: L. Lefèvre, A. Pevtsov
Location: Pasadena, CA
Summary: Study of the Noise model developed by Sophie Mathieu and the LLN team. Alexei Pevtsov mentions the fact that the composite model could be created from the models of the number of groups and the number of spots. This is explained in an article that will be submitted in 2019.

- *Informal: November 2d 2018: ESWW*

Present: L. Lefèvre, E. Cliver
Location: ROB
Summary: Study of the different techniques for monitoring. Ed Cliver gives references to monitoring of observatories.

7. VALORISATION ACTIVITIES

7.1 PUBLICATIONS

S. Mathieu, R. von Sachs, V. Delouille and L. Lefèvre, "UNCERTAINTY QUANTIFICATION IN SUNSPOT COUNTS," 2018 IEEE Data Science Workshop (DSW), Lausanne, 2018, pp. 1-5.
doi: 10.1109/DSW.2018.8439893

7.2 PARTICIPATION/ORGANISATION OF SEMINARS (NATIONAL/INTERNATIONAL)

Oral presentation, poster and/or organisation of workshops, symposia etc.

March 2018, VAL-U-SUN meeting (National), organization, oral presentations
April 2018, Laure Lefèvre, EGU, Vienna, Austria, oral presentation
May 2018 , VAL-U-SUN meeting (National), organization, oral presentations
June 2018, VAL-U-SUN meeting (International), organization, oral presentations
June 2018, Sophie Mathieu, Veronique Delouille, IEEE DSW, Lausanne, Poster
July 2018, Laure Lefèvre**,** COSPAR, Pasadena, CA, oral presentation
October 2018, Sophie Mathieu, RSSB meeting (Royal Statistical Society of Belgium), Domaine des Hautes Fagnes, Poster
October 2018, VAL-U-SUN meeting (National), organization, oral presentations

October-November 2018, Veronique Delouille, Laure Lefèvre, SDO Science Workshop, Ghent, poster presentation
December 2018, VAL-U-SUN meeting (National), organization, oral presentations

7.3 SUPPORT TO DECISION MAKING (IF APPLICABLE)

Solar activity and space weather have an increasing influence on our society as it becomes more dependent on space infrastructure for communication, position and other services. The solar cycle as shown by the International Sunspot Number is the only available forecast on a timescale of years that is relevant for political decision-making.

7.4 OTHER

The coordinator L. Lefèvre introduced the VAL-U-SUN project and its associated collections to every seminar/workshop/conference she attended during year two (section 7.2).

8. ENCOUNTERED PROBLEMS AND SOLUTIONS

Encountered problems/obstacles, adopted and/or envisaged solutions, unsolved problems

I list two items here for which we found efficient solutions:

1. At the end of 2017, we had to go through the accreditation of the SILSO World data center to the World Data System (https://www.icsu-wds.org). For a center of our size (in terms of data and personnel), this was a time-consuming process. To remedy that, we combined the VAL-U-SUN metadata inventory and filling with this accreditation process and the GDPR (https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations_en?pk_source=google_ads&pk_medium=paid&pk_campaign=gdpr2019) compliance exercise.
2. We experienced some delays with the digitization of the areas of the sunspot groups through the DIGISUN program for diverse reasons. The main reason identified was the unfriendliness of the graphical interface and an area module that was not properly linked to the main program. This was remedied with the development of a new DIGISUN. As this had to be done in the context of the day-to-day operations of the USET stations, I did not add (unpaid) personnel to the VAL-U-SUN project.