Customer	**: ESA/ESRIN**	**Document Ref**	**: IDEAS+**
Contract No	**:**	**Issue Date**	**: 19 May 2016**
WP No	**:**	**Issue**	**: 2.0**

Proba-V Clouds Detection Round Robin Protocol

Abstract : This document presents the Protocols for participation to the Proba-V Clouds Detection Round Robin (PV-CDRR). It includes the aim of the project, the rules for participation, the description of the input and output data as well as the adopted Quality Assessment (QA) approach and metrics.

Author : **Rosario Q. Iannone** **Approval** : **Fabrizio Niro**

Table of Contents

AMENDMENT POLICY

This document shall be amended by releasing a new edition of the document in its entirety. The Amendment Record Sheet below records the history and issue status of this document.

AMENDMENT RECORD SHEET

ISSUE	DATE	REASON
1.0	16 April 2016	Initial version
2.0	19 May 2016	Information regarding the access to the Cloud Toolbox has been added.

1. INTRODUCTION

1.1 Background

Clouds screening represents a critical preprocessing step for the retrieval of surface and atmospheric parameters from satellite observations. In the case of Proba-V [RD-1], clouds detection is particularly challenging, considering the limited number of spectral bands (Blue, Red, NIR and SWIR) and the lack of TIR channels or dedicated cirrus band (as the 1.38 micron band).

Clouds detection for Proba-V is currently based on multiple thresholds method applied to the Blue and the SWIR spectral bands [RD-2]. The drawback of such thresholds method is that its accuracy depends on the amount of contrast in radiometry between the underlying surface and the clouds. Owing to the heterogeneity of the surface reflectance characteristics, the definition of a globally valid threshold is practically impossible. To overcome this limitation, a new algorithm was recently developed and it is now being implemented for next Proba-V reprocessing [RD-3]. The new algorithm uses monthly means of reflectance in the blue band and associated status maps derived from MERIS FR mission. This auxiliary information is used to design a "dynamic threshold" algorithm, with cloud tests customized for each status class (land, water, snow/ice, unknown land cover).

Despite the improvements obtained with this new approach, some issues still remain that need to be further addressed, in particular a dependence on illumination and viewing geometry and some remaining misdetection at the edges of each status class (e.g., land/water). In order to address these issues and define the baseline for the future operational processor, a Round Robin exercise on Proba-V cloud detection (PV-CDRR) was organized by the European Space Agency (ESA) in collaboration with the Belgian Science Policy Office (BELSPO).

The aim of this document is to provide a description of the protocols for participating to this Round Robin; this includes the rules for participation, the required input and output data and the associated quality assessment approach and metrics.

1.2 References

Table 1: List of reference documents.

Number	Reference
RD-1	Dierckx, Wouter, et al. "PROBA-V mission for global vegetation monitoring: standard products and image quality." International Journal of Remote Sensing 35.7 (2014): 2589-2614.
RD-2	Lisens, G., P. Kempeneers, F. Fierens, and J. Van Rensbergen. Development of Cloud, Snow, and Shadow Masking Algorithms for VEGETATION Imagery. Proceedings of Geoscience and Remote Sensing Symposium, IGARSS 2000, Honolulu, HI 2: 834–836.
RD-3	Wolters, E.L.A., Swinnen, E., I. Benhadj, Dierckx, W., PROBA-V cloud detection evaluation and proposed modification, QWG Technical Note, 17/7/2015
RD-4	Brockmann C., Paperin M., Danne O., Kirches, G., Bontemps, S., Stelzer, K., Ruescas, A. Cloud Screening and Pixel Characterization: IdePix Approach and Validation Using PixBox, Sentinel-3 OLCI/SLSTR and MERIS/(A)ATSR workshop, ESA-ESRIN, Frascati, Italy, 15-19 Oct 2012.

RD-5	Congalton, R.G.; Green, K. (2009). Assessing the Accuracy of Remotely Sensed Data, 2nd ed.; CRC Press: Boca Raton, FL, USA; p. 193.
RD-6	Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement. Vol. 20, No. 1, pp. 37–40.
RD-7	Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. Communication Methods and Measures, 1, 77-89.
RD-8	Scott, W. (1955). "Reliability of content analysis: The case of nominal scale coding." Public Opinion Quarterly, 19(3), 321-325.
RD-9	Curran, Paul J. "The semivariogram in remote sensing: an introduction." Remote sensing of Environment 24.3 (1988): 493-507.
RD-10	Román, Miguel O., et al. "The MODIS (Collection V005) BRDF/albedo product: Assessment of spatial representativeness over forested landscapes." Remote Sensing of Environment 113.11 (2009): 2476-2498.
RD-11	CMS/Météo-France, 2005, Validation report for PGE01-02-03 of SAF/NWC/MSG. Météo France / Centre de Météorologie Spatiale Report SAF/NWC/IOP/MFL/SCI/VAL/01, version 1.0.

1.2 Acronyms

ATBD	Algorithm Theoretical Basis Document
BELSPO	Belgian Federal Science Policy Office
CCI	Climate Change Initiative
CE	Commission Error
EO	Earth Observation
ESA	European Space Agency
ESRIN	European Space Research Institute
FAR	False Alarm Rate
HDF5	Hierarchical Data Format 5
HR	Hit Rate
IDEAS+	Instrument Data quality Evaluation and Analysis Service
MEP	Mission Exploitation Platform
NetCDF	Network Common Data Form
NIR	Near-InfraRed
NN	Neural Network
OE	Omission Error
PA	Producer's Accuracy
PDF	Portable Document Format
POD	Probability of detection
PROBA-V	Project for on-board Autonomy-Vegetation

PV-CDRR	Proba-V Cloud Detection Round Robin
QA	Quality Assessment
QC	Quality Control
SPPA	Sensor Performance and Product Algorithm
SWIR	Short-Wave InfraRed
TIR	Thermal Infra-Red
TOA	Top Of Atmosphere
UA	User’s Accuracy
VITO	Flemish institute for technological research

2. PARTICIPATION

2.1 Who is invited to participate?

Participation to PV-CDRR is open to all the scientists who:

- are the original developers of the cloud detection algorithm to be tested,
- agree on submitting the PV-CDRR processing results within the required constraints (i.e., deadline, format).

2.2 Benefits of participation

By participating to the PV-CDRR you will have:

- The opportunity to compare the output of your cloud detection algorithm with other leading algorithms, identifying the relative strengths and weaknesses.
- The opportunity to be a co-author of an official ESA report and a peer-reviewed paper summarizing the results of the exercise. Note also that a final Workshop will be organized in ESRIN during Feb 2017 to discuss the results.
- The opportunity to provide potentially the most suited algorithm for the future Proba-V operational chain as well as to be considered for other ESA optical missions.

2.3 What are my responsibilities as a participant?

As participant to the PV-CDRR you are committing to the following:

- To analyse Proba-V input TOA products using your algorithm.
- To submit the results (cloud masks) according to the recommended formats within the deadline.
- To provide a description of your cloud-screening algorithm, in the form of an ATBD document, including reference to peer-reviewed documents, where available.

2.4 How will progress and results be reported?

Participants will be informed through emails about the progress of the Round Robin exercise. Furthermore, any relevant information about this project, including news and documentation will be published in a dedicated website:

https://earth.esa.int/web/sppa/activities/multi-sensors-timeseries/pv-cdrr.

The results of the PV-CDRR will be summarized in a final ESA report and peer-reviewed paper. Co-authorship of the paper will be granted to the Round Robin participants.

Results will be further discussed during a dedicated one-day workshop in ESRIN.

3. ROUND ROBIN DESIGN

3.1 Design

The design of the Round Robin exercise is depicted in Figure 1. There are three types of datasets, which are relevant for the PV-CDRR: the input reference scenes, the validation dataset and the test dataset. The input reference scenes and the test dataset will be provided to the participants, while the validation dataset will be kept “in a vault” and used at the end of the project to perform the quality assessment. Brockmann Consult will be responsible for this final quality assessment.

A detailed description of these datasets is provided hereafter.

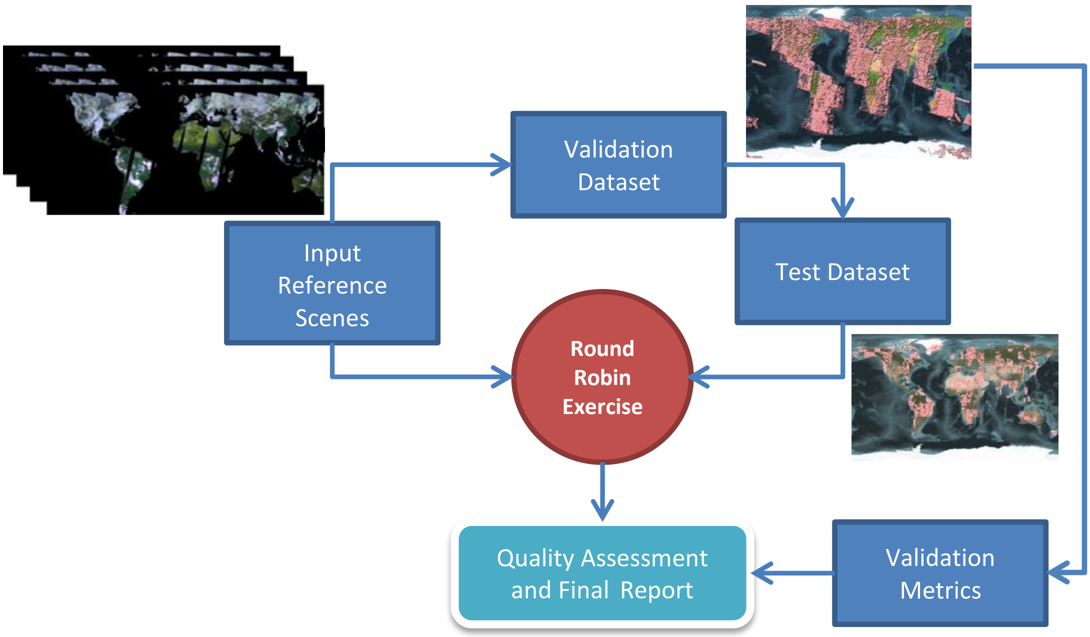

Figure 1 – Proba-v Cloud Detection Round Robin design.

Input Reference Scenes

The input reference scenes that will be the subject of the Round Robin exercise (the one to be processed by the participants) are Proba-V TOA products. More specifically, *Level 2a* products will be provided, consisting of TOA reflectances in the four Proba-V bands, radiometrically and geometrically corrected, projected to Plate Carrée grid, provided as full segments (not tiled) and resampled to the chosen spatial resolution.

Proba-V data are currently provided to users at three spatial resolutions: 1km (for continuity with SPOT-VGT archive), 333m and 100m (100m only over a limited 500km swath with 5 days revisit time).

For this Round Robin exercise we restraint to one single resolution: 333m.

The input products will be blinded (no cloud/ice/shadows mask) in order to avoid that algorithms’ developers “calibrate” or train their algorithms on the provided pixel classification.

The input data will consist ultimately in four days (for the four seasons) of *Level 2a* data covering the full globe (daylight measurements of land and coastal areas). Products format will be HDF5.

Finally, it is noteworthy that *Level 2a* products are scenes generated by *Level 1c* data and therefore they are provided by individual cameras. Depending on the camera, the swath covered by the *Level 2a* products changes; the area covered by the central camera is about 500 km, whereas the area covered by the left and right cameras is of the order of 875 km. They slightly overlap in the across-track direction. The fact that these *Level 2a* products are single scenes has, however, no impact on the validation exercise.

Validation Dataset

The dataset on which the final quality assessment will be made is defined as Validation Dataset. The validation dataset is composed by a collection of pixels (several ten thousand pixels), which were manually selected and classified by visual inspection. The pixel classification and collection is made using a dedicated tool called PixBox [RD-4], developed by Brockmann Consult and successfully used for this purpose in the frame of various ESA CCI projects.

The visually classified pixels of the validation dataset are extracted from the same input reference scenes distributed to the PV-CDRR participants.

The validation dataset will be statistically representative of the different pixel classes, which are relevant for our scope, i.e.: clear sky, thick clouds, and semi-transparent clouds over different surfaces: land, snow/ice and coastal water. The statistical distribution of the classes will be representative of mean global clouds cover condition, i.e.: 60% of cloudy pixels (half of which semi-transparent clouds), 40% of clear pixels. The distribution of surface type will be in-line with the typical observation scenario of the Proba-V mission, acquiring largely over land (70%), with the remaining 30% equally distributed over coastal, inland water and snow/ice.

Moreover, the validation dataset will be globally spread and representative of different seasons and different geometry of observation (sun and viewing geometries).

Test Dataset

The test dataset represents a small subset of the validation dataset. This dataset will be representative of different pixel classes and it will include metadata information on pixel classification derived by visual inspection using PixBox.

This dataset will be delivered to the participants together with the input reference scenes in order to provide example of our pixel classification criteria and nomenclature.

3.2 Quality Assessment

The adopted Quality Assessment (QA) approach is primarily based on the pixel-by-pixel comparison of each output cloud mask with the corresponding visually classified PixBox validation dataset, which represents our "truth" data.

PixBox has the advantage of being based on visual inspection, which is the most effective approach for clouds identification, and implements, in addition, a set of tools for verifying the representativeness (both statistical and geographical) and the correctness of the collected validation dataset, this includes for instance a posteriori QC of the visual classification in order to detect potential human errors.

The quality metrics will be based on confusion matrices formalism, with the two axes of the matrices representing respectively the PixBox visual classification and the algorithm cloud masks output. Standard quality metrics will be used to quantify the performances of each method with respect the "truth" data, this includes in particular the computation of the following statistical quantities: Producer, User and Overall Accuracy (PA, UA

OAA), Omission and Commission Error, Scott's pi, Cohen's Kappa and Krippendorf Alpha coefficients, see the following papers for additional details [RD-5], [RD-6], [RD-7], [RD-8].

In addition to the pixel-by-pixel validation, visual comparison of selected images will be used to investigate the performances of the various algorithms in clouds structure delineation and for assessing critical cases (thin, patchy, or cirrus clouds detection). The visual inspection, in fact, enables to better assess the contextual information on the images, which is lost within the PixBox validation dataset.

Alternative methods for clouds validation are deemed at the moment not sufficiently reliable to be part of the baseline for this project. However, two optional methods will be explored during this project, i.e.:

- Investigation of multi-days composite images. This method is particularly useful in order to evaluate the trade-off between cloud detection and clear pixels' availability; in addition, it allows detecting misclassification and under-detection of clouds, which appear as artefacts in the resulting composites. The QA of the composites will be based on visual inspection, and on the usage of relevant quality metrics, such as measure the semivariogram [RD-9] and of the Relative Coefficient of Variation [RD-10].
- Cross-comparison with in-situ synoptic observations from meteorological stations. This method is the closest one to a typical "geo-physical validation", since independent measurements are used. On the other hand, the accuracy of such approach is limited, mainly because the up scaling of point-measurements to the sensor pixel data will introduce systematic bias. In addition, total sky cameras allow only verification of clouds fraction over the sky area imaged by the ground camera, and the size of this area depends on the altitude of the clouds, meaning that a coincident active measurement of cloud height should be made to improve the accuracy of the validation. Nevertheless, the possibility of using the SYNOP dataset for cloud detection validation will be investigated, such as here [RD-11]. For such method, standard quality indices, such as Probability of Detection (POD), Hit Rate (HR) and False Alarm Rate (FAR) will be used to quantify algorithm classification accuracy.

It should be stressed that if any of these optional QA methods will be adopted, additional input reference scenes will need to be provided to the participants and additional cloud masks will need to be generated. This will be decided during the course of the project and this additional effort will be required only after the processing of the first set of input reference scenes (when algorithms are assumed to be consolidated and tested). We will inform the participant in due time if this additional effort will be required.

3.3 Schedule

The Proba-V Cloud Detection Round Robin will be performed according to this plan:

Table 2: Schedule of the PV-CDRR exercise.

Activity	Deadline
Delivery of Round Robin input data (Input Reference Images)	30 Jun 2016
Delivery of the Test Dataset	30 Jun 2016
Delivery of Round Robin output data (Cloud masks, ATBD)	1 Nov 2016
Quality Assessment Report	15 Jan 2017
Final Workshop in ESRIN	Feb 2017

3.4 Deadline for Submissions

The deadline for the submissions for the algorithm output is the **1st November 2016**. Due to the schedule of this exercise, any submissions after this date will not be considered in the selection process. The description of your algorithm and statement about potential future development must also be submitted by this deadline.

4. DATA

4.1 Input Data

The input data consists of:

- **Input Reference Scenes:** *Level 2a* Proba-V products in HDF5 format, consisting of TOA reflectance in the four bands, provided as full segments and covering the full globe (land and coastal areas). The data set will consist of 331 products. The following dates will be considered for the exercise in order to cover the four seasons:
    - 21/03/2014
    - 21/06/2014
    - 21/09/2014
    - 21/12/2014

- **Test Dataset:** this dataset is intended to provide guidance on how the results will be assessed during the QA phase. The test dataset will be a representative subset of the validation dataset; it will be a collection of visually classified pixels with metadata information on visual classification and relevant auxiliary information, such as time and geographical location of the pixel.

4.2 How do I get the round robin input data package?

The PV-CDRR participants will have access to a remote virtual machine with all the needed input reference scenes, and several relevant tools such as SNAP toolbox, HDF Viewer, R/Rstudio and Python. Additional SW can be easily installed on request, including commercial SW, for which the participants are required to have their own license.

The RSS Cloud toolbox will be initially equipped with 2CPUs x 8 GB RAM and 100GB HD with CentOS 6, but it can be easily tailored on the user needs in terms of available SW, as well as CPU and RAM.

As a first time user, please make sure you complete your registration at following website: http://eogrid.esrin.esa.int/.

After your registration, you can login at the ESA RSS Cloud Toolbox service WebPortal: http://eogrid.esrin.esa.int/cloudtoolbox/.

When you will be logged in, please notify your access by sending an e-mail to the following address: eo-gpod@esa.int. You can find the Virtual Machine in your My VMs page. In order to have access to your personal virtual machine, use the credentials, which are present in the machine notes (visible always from the My VMs page).

Note that this platform is the basis used to build the Proba-V Mission Exploitation Platform (MEP); hence this project will provide the opportunity for the participants to gain experience with such advanced tool. Additional details about the Proba-V MEP can be found here: https://proba-v-mep.esa.int.

4.3 Format specification of participant contributions

We recommend that all participant contributions will be submitted in HDF5 format. It is also requested that submission filename should provide a name that is indicative of the corresponding *Level 2a* product, and it should include the participant acronym. The name of each file submitted should have therefore the following format:

[Level 2a filename]_[participant_acronym].hdf

Spaces and special characters should be avoided.

Although the attention is focused on a binary flag (clouds/clear) to ease inter-comparison of the different algorithm, it is accepted also the development of algorithms able to provide “ambiguous” cases, where ambiguous corresponds to sub-visible or semi-transparent clouds (in the visible and NIR spectrum). In this context, semi-transparent means that clouds optical thickness in the visible is such that surface properties and spatial characteristics are still “visible” in the satellite images.

The table below summarizes the fields that must be included in the submissions:

Table 3: Mandatory fields in the submission files.

Cloud Mask Flag	Conditions
0	clear
1	cloud
2 (optional)	semi-transparent

Please contact the PV-CDRR manager (see section 5 for contact details) if you have any questions regarding submission of data.

4.4 Technical note and ATBD submission

A brief technical note describing the algorithm must be included in the results submission of each processor. The participants should provide information covering the following points:

- A description of the algorithm providing references to peer-reviewed papers where available.
- The channel combinations required by the algorithm.
- The static and dynamic auxiliary data required in order to implement the cloud mask algorithm.
- Description of the main steps of the algorithm and indication of any model used.

5. IMPORTANT CONTACTS

5.1 Project support and troubleshooting

The PV-CDRR is managed by Fabrizio Niro (ESA/ESRIN). In case of technical enquires related to data access and submission, please contact Rosario Q. Iannone (RosarioQuirino.Iannone@serco.com), who will provide support to the participants.

For troubleshooting when using the RSS Cloud Toolbox please sent e-mail to: eo-gpod@esa.int.

If necessary, the queries can also be directed to the rest of the PV-CRDD coordinators: Fabrizio Niro (fabrizio.niro@esa.int) and Carsten Brockmann (carsten.brockmann@brockmann-consult.de).

The project Manager will maintain the following webpage for project's news, documents and information: https://earth.esa.int/web/sppa/activities/multi-sensors-timeseries/pv-cdrr

End of Document