Customer	: ESRIN	Document Ref	: SST_CCI-RRP-UOL-001
WP No	:	Issue Date	: 23 September 2011
		Issue	: 1

Project   **: CCI Phase 1 (SST)**

Title   **: Round Robin Protocol**

Abstract   **:** This document summarises the round robin protocol.

Author   **:** _____

Gary Corlett
University of Leicester

Accepted   **:** _____

Chris Merchant
University of Edinburgh

Accepted   **:** _____

Craig Donlon
ESA

Distribution   **:** SST_cci team members

ESA (Craig Donlon)

TABLE OF CONTENTS

AMENDMENT RECORD

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 FOR CHANGE
Draft A	20 Jun 2011	Initial Issue
Draft B	23 Jun 2011	Passed back to author with comments from Science Leader for consideration.
Draft C	28 Jun 2011	Updated with comments from Science Leader and discussion at GHRSST
Draft D	28 Jun 2011	Cosmetic / editorial changes by Project Manager
Draft E	13 Sep 2011	Updated to action RIDS raised by ESA technical officer
Draft F	20 Sep 2011	Updated to respond to ESA comments on project team's responses to RIDs
Issue 1	23 Sep 2011	Cosmetic changes, and reference to project website in section 1.3

1. INTRODUCTION

The SST_CCI project is part of the ESA Climate Change Initiative, which aims to produce and validate sea surface temperature (SST) SST essential climate variable (ECV) data products.

In order to identify the best performing algorithm or combination of algorithms, the SST_CCI project is holding an open round-robin (RR) algorithm intercomparison and product validation exercise following the protocol defined in this document and using the selection criteria defined in the Product Validation Plan (PVP, RD.216, Section 4). By maximising the number of users participating in the Round Robin exercise, ESA expects to identify the best algorithms for a future operational system.

The chosen algorithm(s) will then be implemented in an end-to-end system to generate the first SST_CCI data records. It is expected that future algorithm selection exercises will be carried out for each subsequent reprocessing to ensure the best performing algorithm is always implemented.

1.1 Purpose and Scope

This document summarises the protocol for the SST_CCI round robin algorithm selection exercise.

1.2 Structure of the Document

After this introduction, the document is divided into a number of major sections that are briefly described below:

2 PARTICIPATION

This section summarises how one participates in the SST_CCI algorithm selection exercise.

3 SCHEDULE

This section summarises the schedule of the SST_CCI round robin algorithm selection exercise.

4 EXPERIMENT DESIGN AND SELECTION CRITERIA

This section details the experiment design and algorithm selection criteria for the SST_CCI round robin exercise.

5 DATA

This section summarises the input and output data of the SST_CCI round robin exercise.

6 IMPORTANT CONTACTS

This section provides contact details of key personnel involved in the running of the SST_CCI round robin exercise.

1.3 Referenced Documents

The following is a list of documents with a direct bearing on the content of this report. Where referenced in the text, these are identified as RD.n, where 'n' is the number in the list below:

RD.216 SST_CCI Product Validation Plan, SST_CCI-PVP-UoL-001 (Oct 2011)

RD.217 SST_CCI Round Robin Data Package Specification, SST_CCI-RRDP-UoL-001 (Sep 2011)

The current version of each reference document is available via the SST CCI web pages at http://www.esa-sst-cci.org/?q=documents#   .

1.4 Definitions of Terms

The following terms have been used in this report with the meanings shown.

Term	Definition
ESA	European Space Agency
GHRSST	Group for High Resolution SST
PVASR	Product Validation and Algorithm Selection Report
PVP	Product Validation Plan
RR	Round Robin
RRDP	Round Robin Data Package
SL	Science Leader
SST	Sea Surface Temperature
SST_CCI	ESA Climate Change Initiative on SST

2.   PARTICIPATION

2.1   Who can participate?

The SST_CCI round robin algorithm selection exercise is open to anyone who can contribute to and/or benefit from the development of better SST algorithms.

2.2   What do I gain from participating?

By participating in the SST_CCI round robin algorithm selection exercise you will:

- See how results of your algorithm objectively compare with all participating algorithms

- Gain early use of an SST_CCI multi-sensor match-up dataset[1]

- Contribute to a major initiative for a SST climate data record

- Have opportunity to be a contributing co-author of a peer-reviewed paper

- Potentially provide the winning algorithm!

2.3   What am I expected to contribute?

All participants in the SST_CCI round robin algorithm selection exercise are required to contribute:

1. Retrieved satellite SSTs generated by your algorithm(s) based on the data provided by the project, associated uncertainties and sensitivities, for any sensor included in RRDP

   The RRDP contains extracted satellite reflectances and brightness temperatures, data quality masks, cloud masks, NWP fields and RTTOV simulations, for multi-sensor match-ups between ATSR/AVHRR/SEVIRI and in situ data. You can provide retrieved SSTs, associated uncertainties and sensitivities, for one or all of the available sensors.

2. A list of peer-reviewed references describing the algorithm for retrieval and uncertainty estimation, and a <u>very brief</u> technical note summarising the algorithm's

   - theoretical basis

   - degree of dependence on tuning to in situ data

---

[1] *Within the SST CCI project, the multi-sensor matchup dataset (MMD) is a central capability for algorithm development, assessment and validation. MMD capability will be designed into the prototype system being specified and demonstrated for creating SST CDRs, and is intended to be a lasting innovation. The RRDP is a pre-prototype example of this concept.*

- generality

- improvability

- and difficulty of implementation (see below)

2.4   What commitment do I give?

All participants in the SST_CCI round robin algorithm selection exercise are required to:

1. Provide their own resources to cover their participation

2. Register their intention to participate with the SST_CCI team to gain access to the data

3. Agree to the SST_CCI Round Robin (RR) conditions of use[2]

    - To use the RR data package (RRDP) only for RR participation

    - To not redistribute the data to other parties without the permission of ESA or the original data provider, as appropriate

    - To acknowledge the assistance of the ESA CCI programme in any publication that is based upon the use of the SST_cci Round Robin data.

    Agreement to these conditions is implicit upon registration.

4. Download the RR data package and documentation from a dedicated download site

5. Deliver their contributions in the specified data format to a dedicated upload site by the date specified

6. Give permission for the SSTs, uncertainty estimates and calculated SST sensitivities to be made publicly available

In addition, any optional comments you wish to make regarding the RR exercise, the design and content of the RRDP, etc, will be welcomed.

2.5   What happens next?

Once you submit your results the algorithm selection team led by the Science Leader will compare all submitted results on an equal basis using the pre-defined metrics described in the PVP (RD.216). The objective is to determine the preferred algorithms to implement for subsequent processing within the SST_CCI project to create two data records:

1. A long-term (1991-2010) dataset of ATSR and AVHRR to demonstrate a climate data record

---

[2] *For various data within the RRDP, the original data provider has given us permission to include and distribute the data within the RRDP only on these conditions*

2. A short-term dataset (six months within the period October 2010 to June 2011) of AMSR-E, ATSR, AVHRR, TMI and SEVIRI to demonstrate a climate service

All the results for different metrics and the outcome of the algorithm selection will be publicly available.

Information on how to submit your data to the exercise is given in Section 5.

2.6   How will progress and results be reported?

Participants will receive periodic email updates about the progress of the Round Robin exercise.

The results of the SST_CCI algorithm selection exercise will be published in the Product Validation and Algorithm Selection Report (PVASR[3]). This document will be published on the SST_CCI website (http://www.esa-sst-cci.org/).

In addition, results will be submitted for publication in a peer-reviewed journal. We will contact you about your interest in co-authoring this publication.

2.7   Will the results and data be made public?

Yes, results will be made publicly available as follows:

- All results for algorithm selection metrics will be published in the PVASR on the SST_CCI website (http://www.esa-sst-cci.org/)

- Results will be prepared for peer-reviewed publication in consultation with participants

- The complete algorithm selection dataset (including the submitted SSTs, validation values, etc) will be freely available online

2.8   What if my sensor is not in the round robin data package?

The SST_CCI algorithm selection exercise is directed towards a set of specific sensors chosen for this initial project. If you wish to contribute data from a sensor not in the initial list then you may provide matched SSTs to the in situ locations provided in the RRDP.

Of course any additional sensors will not be included in the algorithm selection exercise but you will be able to compare your data against the sensors used in SST_CCI once the algorithm selection is complete and the dataset is made publically available.

---

[3] *PVASR is the mandated report title, but it is a misnomer: at this stage, the algorithms will not have been used to create products and no product validation will have taken place. The report will describe the results of algorithm selection metrics and the decision process for selecting algorithms on the basis of these results.*

3. SCHEDULE

3.1 What are the time scales?

Seven milestones are defined for the SST_CCI round robin algorithm selection exercise. These are:

1. Launch of RRDP:

   30th June 2011 during the GHRSST-12 meeting (Edinburgh)

2. Release of training and test data:

   23rd September 2011

3. Release of selection data:

   15th December 2011

4. Submission of participant contributions:

   31st January 2012

5. Start of algorithm selection:

   1st March 2011

6. End of algorithm selection:

   30th April 2012

7. Publication of results and data release:

   1st July 2012

A Gantt chart summarising the SST_CCI round robin algorithm selection exercise schedule is shown below.
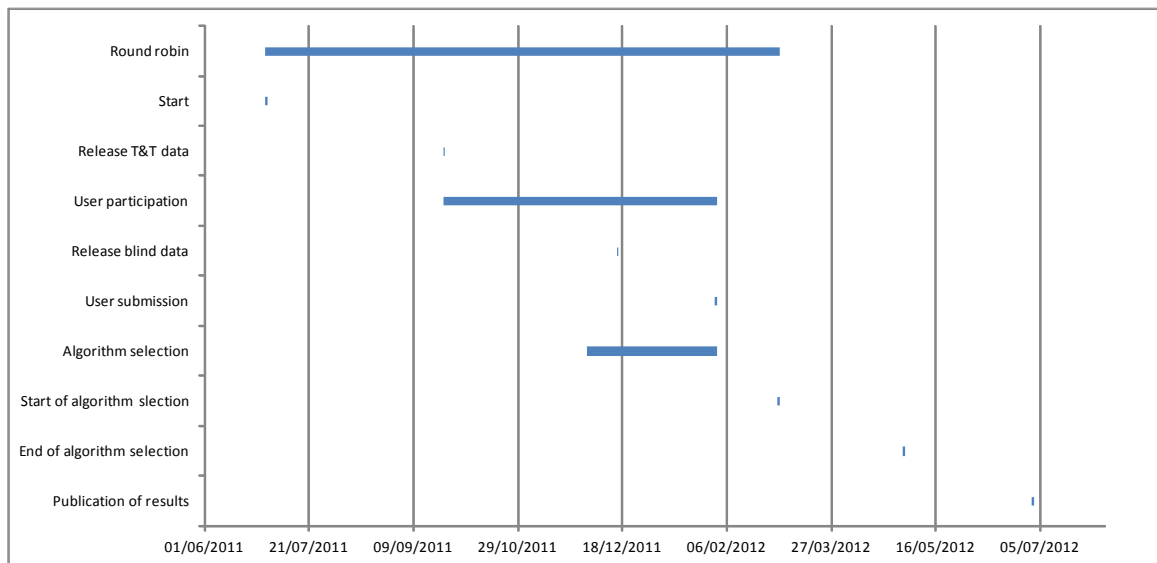
4. EXPERIMENT DESIGN AND SELECTION CRITERIA

The experiment design and selection criteria for the Round Robin are explained in detail in the PVP, and are summarised below.

4.1 Experiment Design

Algorithms will be compared on a fair basis by standardisation of the approach:

- Competing algorithms will be developed using identified training data within the RRDP and will be compared by looking at their results when applied to test data within the RRDP. For fair comparison, test data must not be used at all in algorithm development; the test data will be reserved for use only after the algorithm is finalized. All participants for a given sensor and category must use the same training and test subsets in order to be considered within the exercise.

- Common metrics describing the results will be used for each type of algorithm to facilitate comparison of performance.

- Where objective/independent external data are available for validation, these will be used to compare performance, but in other situations, more subjective expert evaluation must be relied upon.

Algorithm selection requires joint assessment of a range of metrics and wider considerations. Not all properties of interest are quantifiable as metrics. Among measures that are quantifiable in principle, it may not always be feasible to undertake proper quantification within the scope of the project, and thus a qualitative approach may still be necessary.

For algorithm selection purposes, the validation data to be used are the matches flagged as "test data" in the RRDP. This will include at least two types of validation data (drifting buoys and moored buoys) for which results should be prepared separately. Certain test data will also be flagged as "high latitude" and "coastal" cases, and some metrics will be evaluated for performance separately using these subsets.

4.2 Selection Criteria

The selection criteria for the SST_CCI round robin algorithm selection exercise have been pre-defined before the start of the activity and are summarised below. All assessments will be carried out with reference to drifting buoys. Further details on each criterion can be found in Section 4 of the PVP (RD.216).

For SST:

- Bias – the systematic difference from the truth, and is assessed via systematic differences from validation data

- Precision - observations generally differ from the truth according to a distribution that has a spread (or "dispersion"). The concept of precision is to characterise that dispersion, with a precise observation having a narrow spread

- Stability - constancy of bias in time. Stability of observation is critical when looking at differences between observations (i.e., changes of SST over time).

- Degree of independence - SST retrievals can be based on either empirical correlations to in situ observations, or on radiative transfer modelling. For

applications where satellite SSTs are required to complement, enhance or test in situ observations, independence from in situ SSTs is an advantage (and in some cases a necessity).

- SST sensitivity - for the ideal SST estimate, changes in true SST are (on average) wholly reflected in changes in the estimated SST. Thus, $\partial \hat{x} / \partial x$ needs to be evaluated, where x is true SST and $\hat{x}$ is the estimated SST. We refer to $\partial \hat{x} / \partial x$ as "SST sensitivity" for the SST estimate.

- Generality - the degree to which an algorithm is adaptable to other sensors and/or channel combinations, including future missions

- Improvability - the degree to which an algorithm can be further refined

- Difficulty of implementation – high, medium or low, based upon factors such as the use of external models and the size and nature of any required static or dynamic auxiliary files (e.g., look up tables, NWP, etc.)

For SST uncertainty, the following criteria will be used (definitions are as above):

- Bias

- Degree of independence

- Generality

- Improvability

- Difficulty of implementation

5. DATA

5.1 What is in the round robin data package?

The SST_CCI RRDP contains the necessary satellite and auxiliary data to carry out SST_CCI round robin algorithm selection exercise. Three different subsets of data will be included:

1. 'Training': This is the subset that you should use for determining coefficients in empirically derived algorithms, for any other form of algorithm turning, and/or as a training dataset in a supervised neural net optimisation (or similar). The training set is made available with validation (in situ) values. This subset of data is released at RRP milestone 2 (from 15th September 2011).

2. 'Test': This is the subset that you should use to get a (statistically) independent assessment of your algorithm developed using the training set -- it is "reserved data" for algorithm development, with validation values included. This subset of data is released at RRP milestone 2 (from 15th September 2011).

   Ideally, you should use this type once any algorithm tuning is done; although if the test set performance is perceived as poor, it is recognised that this might prompt another cycle of training/test.

3. 'Selection': This subset will be distributed to participants (including the SST CCI developers) without validation values, and with fields sufficient only to derive SSTs, SST uncertainties and SST sensitivity for each "blind" matchup.

   These derived quantities are the minimum set of data that you must submit for each sensor of interest.

   This subset will be used to carry out the algorithm selection process and is released at RRP milestone 3 (from 15th December 2011).

In addition a fourth type of data, the 'validation' subset, is being retained to be used exclusively for product validation after system prototyping and product generation. (In ESA's SoW, it is the set referred to as "reference data".) No participants, including those responsible for algorithm selection, will have access to this subset of data prior to product generation.

The product validation will be done by team members not involved in algorithm specification. Thus the product validation will be fully independent, in terms of data and personnel.

5.2 How do I get the round robin data package?

The RRDP will be available for download from a secure FTP site. A detailed content specification of the RRDP can be found in RD.217.

To obtain login details for the RRDP download site you must email a request Gary Corlett (contact details given in Section 6).

5.3 What data do I have to deliver?

All participants are required to deliver

1. Documentation: A brief technical note summarising the retrieval algorithm and uncertainty estimation, giving appropriate references.

   The technical note should be submitted to the Science Leader (see Section 6 for contact details) and can be submitted any time between the start of the RRDP (milestone 1) and the start of the algorithm selection exercise (milestone 5).

2. Data: Retrieved SSTs, uncertainties and SST sensitivities according to the specification given in Section 5.5 below.

   Retrieved SST: This is the estimate of SST arising from your algorithm given the satellite observations provided

   SST uncertainty: This is your estimate of the standard deviation of the distribution of error you expect for your retrievals. This may be a single estimate, or may vary between matches if you have a model for the variations in retrieval uncertainty.

   SST sensitivity: This is an estimate of the responsiveness of your algorithm to a true change in SST, other factors being constant. Partial derivatives of brightness temperature with respect to SST are provided in the RRDP to facilitate consistent calculation of this estimate between participants.

5.4 How do I submit my data?

You will be required to submit your data to a secure FTP site. Further details will be provided as the exercise proceeds.

5.5 Format specification of participant contributions

All participant contributions must be in NetCDF version 3. As a minimum they must contain:

Variable name	Description
matchup.id	Unique MMD record number
<sensor_name>.<sen_id>.sea_surface_temperature	Retrieved SST

Optional fields are:

Variable name	Description
<sensor_name>.<sen_id>.sea_surface_temperature_uncertainty	Total uncertainty of retrieved SST[#]
<sensor_name>.<sen_id>.sea_surface_temperature_dSST_SST	Retrieved SST sensitivity to SST, i.e. dSST/SST

[#] It is encouraged to provide uncorrelated (random), synoptically correlated (pseudo-random) and large-scale correlated (systematic) components of the SST uncertainty budget separately if they are known. In this case the fields should be named:

- sea_surface_temperature_uncorrelated_uncertainty

- sea_surface_temperature_synoptically_correlated uncertainty

- sea_surface_temperature_large_scale_correlated_uncertainty

An example submission file will be provided to all participants.

Resource and time constraints will compel the SST_CCI round robin algorithm selection team to adopt a zero-tolerance policy to submissions with an incorrect format specification. Please contact the RR manager (see Section 6 for contact details) if you have any questions regarding submission of data.

6. IMPORTANT CONTACTS

The SST_CCI round-robin algorithm selection exercise is managed by the RR manger, Gary Corlett (gkc1@le.ac.uk). All technical enquiries should be directed to the RR manager in the first instance and copied to the project manager Paul Spinks (project.manager@esa-sst-cci.org).

The algorithm selection process will be led by the SST_CCI Science Leader, Chris Merchant (science.leader@esa-sst-cci.org). All scientific enquires should be directed to the SL in the first instance and copied to the Project Manager, Paul Spinks (project.manager@esa-sst-cci.org).

The Project Manager maintains a website (http://www.esa-sst-cci.org/) on which project documents are published.