CCI

BIOMASS

SYSTEM VERIFICATION REPORT

YEAR 1

VERSION 1.0

DOCUMENT REF:	CCI_BIOMASS_SVR_V1
DELIVERABLE REF:	D3.3-SVR
VERSION:	1.0
CREATION DATE:	2019-05-29
LAST MODIFIED	2019-05-31

Document Authorship

	NAME	FUNCTION	ORGANISATION	SIGNATURE	DATE
PREPARED	M. Santoro				
PREPARED	O. Cartus				
PREPARED					
PREPARED					
PREPARED					
PREPARED					
PREPARED					
PREPARED					
PREPARED					
PREPARED					
VERIFIED	S. Quegan	Science Leader	Sheffield University		
APPROVED					

Document Distribution

ORGANISATION	NAME	QUANTITY
ESA	Frank Seifert	

Document History

VERSION	DATE	DESCRIPTION	APPROVED
0.1	2019-05-29	First draft version	
1.0	2019-05-31	Finalised version	

Document Change Record (from Year 1 to Year 2)

VERSION	DATE	DESCRIPTION	APPROVED

Table of Contents

SYMBOLS AND ACRONYMS

CCI	Climate Change Initiative
ESA	European Space Agency
GSV	Growing Stock Volume
MD5	Message Digest Algorithm 5
OS	Operating System
PS	Processing System
SSD	System Specification Document
SVR	System Verification Report

1. Introduction

This document is the System Verification Report (SVR) of the biomass processing system of the ESA CCI Biomass project. A core task of the CCI Biomass project is the development and prototyping of a Processing System (PS) for global biomass products. This report describes the processing system qualification review protocol and results.

The report (Deliverable 3.3, SVR) documents that the system as outlined in the System Specification Document (SSD) is properly working when executed in the operational software environment. The SVR gives a complete report of all activities executed and the results achieved from a technical assessment of the end-to-end prototype system with all its subcomponents to verify that the implementation is consistent and that it fulfils its intended purpose and replicates the results of the prototype algorithms.

The SVR includes for each of the tested modules:

- a description of the objectives and scope of the module
- a list of all elements and components of the module that have been tested including a description of the platform, the network, and the interfaces with other systems
- a description of all test activities carried out and of the criteria on how the module was tested to ensure that the requirements are fulfilled, and that the system performs as specified
- a summary description of all test cases, test procedures, and test data used
- a record of all test results
- a description of all acceptable and stated limitations in the module and the steps taken to work around anomalous, inappropriate, or undesired operating conditions.

The tests are performed at the level of the modules (BIOMASAR-C, BIOMASAR-L) by the Systems Engineers on May 29th, 2019. For all modules the objectives of the processor and the components tested, the performed test activities and the achieved results are documented in this report. The tests are completed by conclusions relevant for the system engineers towards system development and sustainability.

2. Verification Methodology

The system needs to be verified to ensure system integrity after software updates or installation on a new platform. It is not the purpose of the system verification process to validate the product scientifically but to ensure reproducibility of a defined process, and processing system.

The verification processes should distinguish between testing after a system upgrade and the installation on a new platform. While after a software upgrade the results may differ if e.g. a classification algorithm was changed, no or minor differences are to be expected if the PS is installed on a new system. In any case deviations must be understood and rectified.

In general, a benchmark test scenario is defined for the system verification. Such a test scenario covers:

- Hardware requirements
- Software (availability) requirements
- Input Data
- Benchmark Data
- Scenario Process

- Other Resources

The Scenario Process description describes the processing steps and pass/fail tests to be conducted. The processing steps, if based on multiple executables, are best complemented by a script that conducts the different processing steps in an automated fashion to minimize operator errors.

All data (input, intermediate, output, benchmark) must be checked for integrity and consistency. Tests to be done are:

- Availability
- Integrity
- Format
- Content

The use of hash values derived from a hash function is the preferred method to check the data. If the data are consistent the hash value is identical. The preferred hash function is based on the Message-Digest Algorithm 5 (MD5) a widely used and implemented cryptographic algorithm. It computes a 128-bit hash value of any dataset. For some products where deviations on the byte level must be expected (e.g. meta data holding processing dates), special tools may be necessary to only compare or hash the data part that is not affected by dynamic content. Fall-back strategies are value thresholds and visual inspection. The latter might be the only useful method after software upgrades affecting the product algorithm.

Other tests in the scenario address the processing environment. Tests need to cover the expected processing/production time, disk space as well as memory space usage. The results are OS and hardware dependent and will usually be checked against a threshold.

3. Qualification

3.1. Description of the test environment

Hardware: Xeon(R) CPU E5-2630 v2 @ 2.60GHz (6 cores, 12 threads), RAM: 64GB
Operating System: Ubuntu Linux 18.04.1 LTS (x86_64), Kernel: 4.15.0-47-generic
Software: Matlab version R2018a
Network, Interfaces: Input data are received from GAMMA network storage.

3.1.1. Operator

Dr. Oliver Cartus, GAMMA Remote Sensing AG

4. Test Protocol

4.1. Objectives and Scope

The BIOMASAR-C and BIOMASAR-L modules compute Growing Stock Volume (GSV) on a 100m grid based on Sentinel-1 and ALOS-2 data tiles.

4.2. Components Tested

This is an end-to-end test of the BIOMASAR-L module. It is tested by running 3 steps:

1. Retrieve Sigma
2. Spatial interpolation of model parameter estimates
3. Multi-temporal retrieval

4.3. Input Data

BIOMASAR-L :

`Directory with ALOS-2 imagery: /raid/SEN3APP/data/jaxa_palsar_mosaics/`

`GLAS data: /home/oliver/data/GLAS/gla14_all_reduced_dec2015_v4.mat`

`Landsat Canopy Density: /raid/SEN3APP/data/global_forest_change/treecover2010.vrt`

`Land Cover Map: /raid/SEN3APP/data/cci_lc2010/ESACCI-LC-L4-LCCS-Map-300m-P5Y-2010-v1.6.1.tif`

`Maximum GSV: /home/oliver/data/globbiomass/max_gsv/gsvdf_glaspred_interp_gapfil_1km_20180528.tif`

`Beta: /raid/SEN3APP/data/globbiomass/beta_map/beta_map.tif`

4.4. Reference Output Data

Uncertainty Map for 1 degree Tile with upper left coordinates 50°N and 11°E:

`/raid/SEN3APP/data/globbiomass/maps_new/fbd/N50E011/N50E011_mt16_88_gsv_err.tif`

Hash: 5a5ff0b493d8b2e5e748ba771fc257ab

GSV Map:

`/raid/SEN3APP/data/globbiomass/maps_new/fbd/N50E011/N50E011_mt16_88_gsv.tif`

Hash: 7f3f1b65d6347155c614d03eebb3baa3

Model Parameter estimates stored in MATLAB file format:

`/raid/SEN3APP/data/globbiomass/maps_new/fbd/N50E011/N50E011_Y16_sl_HV.mat`

Hash: 27d286727c84b188de4a2e4d616b9dd5

`/raid/SEN3APP/data/globbiomass/maps_new/fbd/N50E011/N50E011_Y17_sl_HV.mat`

Hash: a39fb80296c0a2c804b7555ee9b8e166

Interpolated Model parameter estimates in MATLAB file format:

`/raid/SEN3APP/data/globbiomass/maps_new/fbd/parameters.mat`

Hash: 0702154b7dac8b9cd1f6bae8a2695254

4.5. Requirements for successful testing

Compliance/Integrity of output products:
Integrity by md5sum comparison of main output (gsv map in geotiff)
Alternatively if md5sum fails the datasets are compared visually and by differencing with the reference datasets. If compliance test fails the same procedure applies to all steps of biomasar_call.sh to find the source of the failure

Processing time:
All processing steps are done below the time thresholds:

- biomasar_call.sh 1 < 100 seconds
- biomasar_call.sh 2 < 10000 seconds
- biomasar_call.sh 3 < 300 seconds

Test condition:
Correct reproduction of sample output data set using the test installation of the prototype (input and auxiliary data). Test installation is defined by date/size and/or md5sum of datasets.

4.6. Test Procedure

Run test script run_test.sh (See Appendix A) and log to a report file. The script includes the following steps:

1. Write out Start Date and Time
2. Write out version of code used
3. ./biomasarc_call.sh 1
4. ./biomasarc_call.sh 2
5. ./biomasarc_call.sh 3
6. Report output file dates and time and if applicable md5sum

7. Write out End Date and Time

The test is run without interruption. Results are checked against the benchmark. If the end-to-end test is successful the intermediate results need not to be checked.

4.7. Test results

Test	Benchmark	Result	Pass
Start Date/Time		Mi Mai 29 10:48:35 CEST 2019	
Version		Revision 30	
./biomasar_call.sh 1	<100s	Files produced in 15s	PASS
./biomasar_call.sh 2	<10000s	Files produced in 6034s	PASS
./biomasar_call.sh 3	<300s	Files produced in 49s	PASS
Final Result		MD5SUM match	PASS
Start Date/Time		Mi Mai 29 12:30:13 CEST 2019	
BIOMASAR-L Overall Test condition			**PASS**

MD5 hash of output products:

```
/raid/SEN3APP/data/globbiomass/maps_new/fbd_op/N50E011/N50E011_mt16_88_gsv_
    err.tif
Hash: 5a5ff0b493d8b2e5e748ba771fc257ab


/raid/SEN3APP/data/globbiomass/maps_new/fbd_op/N50E011/N50E011_mt16_88_gsv.
    tif
Hash: 7f3f1b65d6347155c614d03eebb3baa3


/raid/SEN3APP/data/globbiomass/maps_new/fbd_op/N50E011/N50E011_Y16_sl_HV.ma
    t
Hash: 27d286727c84b188de4a2e4d616b9dd5
/raid/SEN3APP/data/globbiomass/maps_new/fbd_op/N50E011/N50E011_Y17_sl_HV.ma
    t
Hash: a39fb80296c0a2c804b7555ee9b8e166
```

```
/raid/SEN3APP/data/globbiomass/maps_new/fbd_op/parameters.mat
Hash: 0702154b7dac8b9cd1f6bae8a2695254
```

4.8. Relevant Issues for System Engineering

The processing line is based on Matlab, GDAL, and shell scripts.

5. Conclusions

On 29 May 2019 the verification of the BIOMASAR-L module was successful. It can now be used for production.