FRM4SM – Fiducial Reference Measurements for Soil Moisture

QA4SM Verification Report
DT4-3
Version 2.1
13 March 2023

Prepared by

Angewandte Wissenschaft, Software und Technologie GmbH

in cooperation with

Vienna University of Technology   **CESBIO**

1

Document Revision History:

Issue	Date	Details	Editor
1.0	2022-04-29	Verification report for release 1	M. Tercjak, P. Stradiotti, W. Preimesberger, A. Boresch
1.1	2022-06-01	Updated report addressing MTR comments	M. Tercjak, P. Stradiotti, W. Preimesberger, A. Boresch
2.0	2023-03-03	Verification report for release 2	M. Tercjak, P. Stradiotti, W. Preimesberger, S. Scherrer, A. Boresch
2.1	2023-03-13	Update addressing ESA comments for release 2 acceptance	M. Tercjak, P. Stradiotti, W. Preimesberger, A. Boresch

For any clarifications, please contact support@qa4sm.eu.

Table of Contents

Acronyms

AWST	Angewandte Wissenschaft Software und Technologie GmbH
CCI	Climate Change Initiative
CDRs	Climate Data Records
CEOS	Committee on Earth Observation Satellites
CESBIO	Centre d'Etudes Spatiales de la Biosphère
CMUG	Climate Modelling User Group
DOI	Digital Object Identifier
ECMWF	European Centre for Medium Range Weather Forecasting
ECV	Essential Climate Variable
EODC	Earth Observation Data Centre
ESA	European Space Agency
FAQ	Frequently Asked Questions
FRM	Fiducial Reference Measurements
FRM4SM	Fiducial Reference Measurements for Soil Moisture
GCOS	Global Climate Observing System
GEO	Group on Earth Observation
GEWEX	Global Energy and Water Cycle Experiment
GUI	Graphical user interface
GTOS	Global Terrestrial Observing System
ISMN	International Soil Moisture Network
NASA	National Aeronautics and Space Administration
NRT	Near Real Time
QA	Quality Assurance
QA4SM	Quality Assurance for Soil Moisture
RFI	Radio Frequency Interference
SAR	Synthetic Aperture Radar
SM	Soil Moisture
SMAP	Soil Moisture Active Passive
SMOS	Soil Moisture and Ocean Salinity
TC	Triple Collocation
TOPC	Terrestrial Observation Panel on Climate
TUW	TU Wien

1   Introduction

1.1   Purpose and scope of document

The QA4SM Evolution Verification Report documents the technical software testing and scientific validation activities, which have been performed to verify QA4SM release 2. Verification results and specific observations made during verification are reported for each pertinent verification activity. The test and validation activities have been performed following the associated QA4SM Evolution Test Plan for release 2. The report also serves as input for the Acceptance Review to approve the subsequent deployment of release 2 in the publicly available QA4SM service.

The previous Versions 1 of the Test Plan and Verification Report pertain to QA4SM release 1.

Software issues determined during verification have been fixed and re-tested. Relevant observations in this context are documented in this report for each affected test case.

1.2   Document overview

Sections 1 through 4 are introductory sections to introduce the document, provide an overview of QA4SM release 2, describe the test approach and to identify required test data.

Sections 5 and 6 report on the technical software testing which has been performed. This includes both automatic unit testing, which was continuously performed during development, and manual test cases, which were designed to cover all relevant user features of QA4SM.

Section 7 reports on the scientific evaluation activities of QA4SM results, which were performed for all integrated datasets with respect to independently obtained validation results which were generated by other groups. In particular, CESBIO contributed independent SMOS Level 2 validation results for comparison with validation results generated with QA4SM.

Section 8 concludes the document with references.

1.3   Target audience

This document is primarily addressed to ESA stakeholders who participate in the acceptance review of QA4SM release 2. It also serves for all interested QA4SM users to get informed about the testing and verification which was performed for release 2 and to compare the verification results with their own observations. Relevant observations and questions can be sent to the QA4SM help desk at support@qa4sm.eu to obtain clarifications and dedicated support.

2 Overview of QA4SM release 2

QA4SM evolution within the FRM4SM project started from a publicly available baseline version of the QA4SM service, which was developed in predecessor projects funded by the Austrian research promotion agency FFG (www.ffg.at/en, project numbers 878929, 866004).
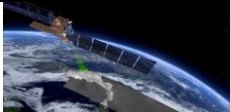
This baseline version is comprehensively described in the QA4SM Service User Manual, v1.1 (FRM4SM deliverable D3-1), which is available at https://qa4sm.eu. Enhancements, which were added in release 1, are described in Version 1 of this Verification Report.

The following major technical and scientific enhancements have been implemented in the evolution towards release 2 within the FRM4SM project:

- SMOS and SMAP Level 2 data were integrated.
- A new feature for users to upload, manage and validate their own soil moisture data has been implemented.
- The user options to define reference datasets were enhanced, so that users can independently select spatial, temporal and scaling references.
- An FRM indicator was introduced to identify representative ISMN sensors. The FRM flag can be used to restrict validation results to representative sensors.
- Handling of validation errors per grid point has been improved to provide better information about the underlying problems which have led to the errors.

The release 2 enhancements are covered by the tests described in sections 6 as summarised in the following tracking table:

Test coverage of enhancements in release 2		
Enhancement	Test Cases	Comments
SMOS and SMAP Level 2 data integration	Section 7.1	QA4SM validation results are compared with independent validations of these datasets.
Upload, management and validation of user data	QA4SM_GUI_UDUF_001, QA4SM_GUI_UDUF_002, QA4SM_GUI_UDUF_003	Manual test cases to verify completeness and proper behavior of the data upload module.
Separation of spatial, temporal and scaling references	QA4SM_GUI_VP_002, QA4SM_GUI_VP_003, QA4SM_GUI_VP_009	These test cases were updated to verify the reference enhancements.
FRM indicator for sensor representativeness	Section 7.4	The quantitative impact of the selected FRM setting on the validation results is evaluated.

Validation error handling per grid point	Section 7.3	See also automated unit test in section 5.

3 Test approach

The following general types of testing and validation activities have been performed:

- Automated tests were continuously performed during incremental development. Software issues caught by these tests were fixed early in the ongoing development process. The automated tests also assured that the software was always kept in a deployable state in the sense that it builds and executes correctly. The performed automatic test activities are described in in section 5.

- Interactive manual tests have been performed after deployment of each new QA4SM version in a dedicated test service instance with limited accessibility for the development team. The manual test cases were performed to systematically verify the correct execution of all QA4SM features from a user perspective. Dedicated test cases have been included for newly implemented enhancements for release 2. The individual test cases, along with the test results and relevant observations made during test execution, are described in section 6. Additional testing was performed to verify the proper execution of validations for all integrated datasets.

- Additional random testing has been performed by internal non-developer users within the project team to catch usability issues, which may otherwise be overlooked by testers who already are fully familiar with the QA4SM service.

- Verification of validation results beyond technical testing has been performed by evaluating QA4SM validation results in comparison with independently produced validation results. The following activities specifically performed within the scope of release 2 are described in section 7: QA4SM validation results for the newly integrated SMOS and SMAP Level 2 datasets were compared against independently produced validation results for these datasets. The impact of the selected FRM indicator on the QA4SM validation results was quantitatively evaluated. New unit tests were implemented to cover the improved validation error handling per grid point.

4 Test data sets

All required test data for automatic and manual software tests are available within the QA4SM services and its source code repositories.

Automated tests use test data which are stored in the relevant test directories in the QA4SM and Pytesmo GitHub repositories. Newly required test data were integrated together with the implementation of the pertinent test cases, which require these data. Hence, all automatic tests can be repeated directly in the repositories, with no need of additional resources.

All manual test cases exclusively use datasets which are integrated into QA4SM for validation by users. Each test case refers to the specific datasets which are used in the test. No additional data are required to independently repeat the described manual tests.

The activities to compare quantitative validation results were based on

- validation results that were generated with QA4SM for each integrated dataset, and
- externally available independent comparison results. For the integrated SMOS Level 2 dataset, CESBIO contributed independently generated validation results which were used for comparison with QA4SM.

The input data to generate QA4SM results are the same integrated datasets, which are available to all QA4SM users in release 2. The selected literature results and validation studies for external comparison are referenced in section 7.

5 Automated software tests and user interface tests

Unit tests are directly integrated into the QA4SM and Pytesmo software repositories on GitHub. They test both individual functions or software features in the back-end and functions related to the data flow between the front-end and back-end parts of the QA4SM software. New unit tests have been added together with the development of new software features during release 2 evolution. In total, more than 140 QA4SM unit tests and 230 Pytesmo unit tests have been set up and executed for release 2.

These automated tests were run by developers each time when new functionality was added to the service, before committing code changes to the master branch of repository. Depending on the individual situation, the tests were executed either collectively or individually to test selected features. All tests have also been executed automatically once per day and additionally upon every pull request to integrate software changes. This test automation was done using Github Actions. The results for executed unit tests are publicly available in the GitHub repository. Issues detected via failed unit tests were resolved before continuing further development. New internal test releases of Pytesmo and QA4SM and the eventual release 2 were created only after all tests on the master branch had passed successfully.

The graphical user interface (GUI) of the QA4SM application was primarily tested manually as described in section 6. Dedicated automated unit tests verify, though, that the individual components of the GUI are properly created. These GUI tests were run in the local development environment during the software development process.

6 Manual test cases

The test cases described in the test plan have been designed to systematically verify the correct performance of all components of the QA4SM user interface and of the interactions between interdependent components. They also verify that the navigation to the individual service pages components works correctly and that the same results are consistently produced for fixed predefined settings.

All test cases have been performed and verified in the release 2 validation phase, after completion of the evolution activities. The results per test case are reported below, along with relevant observations made during testing. Issues, which were identified during testing, have been fixed and the affected test cases have been repeated to verify the successful solution. All tests were performed on the internal test instance of the QA4SM service: test.qa4sm.eu. Upon successful acceptance review, the verified QA4SM release 2 will be deployed in the public QA4SM service instance.

The specification of each test case includes the following points:

- Description: The purpose of the test case indicating which functionality is verified.
- Test data: The input or auxiliary data which is required to execute the test case. Typically, all test data is already available in the tested service. Exceptions are explicitly specified.
- Preconditions: Technical requirements or preparatory actions which are needed to start the test.
- Requirements to be tested: Where explicit software requirements have been specified, which are verified by the test case, they are listed.
- Test Steps and Expected Results: The interactive test procedure is described in the form of a step-by-step execution of the test including detailed user activities in the QA4SM user interface.
  - The expected results describe the expected QA4SM interface responses for each test step.
  - In some cases, failures for specific test steps may imply that the remaining steps cannot be executed and the test case cannot successfully be completed. Each test case is complete when all described test steps have sequentially been executed and the results of each step have been observed.
- Pass/Fail: Each test case passes if the expected results have consistently been observed for all test steps. The test case fails if at least one test step fails in the sense that the actually observed results indicate a failure relative to the specified expected results. In special cases a deviation from the expected result may be acceptable, for example, if the deviation is due to changed circumstances which were not originally foreseen when the test case was developed, the cause of the deviation is understood, a software

failure can be excluded and the functionality of the software is not impaired. In such cases the specific situation and understanding are described in the comments to the test case.

- Comments: Relevant observations during test execution are described in this field. This specifically includes test cases with non-trivial pass/fail results as described above. In such situations the comments describe the observed deviations, the understanding of the situation, the origins of the deviation and the rational for the pass or fail conclusion which was taken.

6.1 Testing of the validation algorithm

Test Case QA4SM_VA_metrics - Test self-validation

Description:	Evaluation of a dataset against itself. This limit case allows only one possible output for each validation metric, with the exception of the significance scores, where the sample size effect could impair the result.
Test Data:	C3S SM combined (v201912)
Preconditions:	-
Requirement(s) to be tested:	-

Test Steps		Expected Results
1.	Go to the "Validate" page and select C3S (v201912) as the spatial reference and candidate dataset, then run the validation. In principle, any other dataset (and version) will yield the same output. Leave all the settings as default and run the validation.	The validation metrics results will need to match the following scores: Bias: ~ 0 Mean square error: ~ 0 Mean square error bias: ~ 0 Mean square error correlation: ~ 0 Mean square error variance: ~ 0 Pearson's r: 1 Sperman's rho: 1 Pearson's/Spearman's p-value : ~ 0 Residual sum of squares: ~ 0 Unbiased root-mean squared deviation: ~ 0

		Note: the results should be as close to the expected values as allowed by the numerical process of the algorithm used. Therefore, a deviation in the order of 8-9 decimal positions can be expected.
Pass/Fail:		Pass
Comments:		

Test Case QA4SM_VA_metrics - Test invalid validation period

Description:	Test the effect of too few points in the sample on the validation output, through the 'Validation Period' settings
Test Data:	C3S (v201912), ISMN (20210131 global)
Preconditions:	-
Requirement(s) to be tested:	-

Test Steps		**Expected Results**
1.	Go to the "Validate" page and run a validation with the default settings	All metrics are calculated normally
2.	Change the fields in 'Validation Period' to compute a validation with the same starting and ending date	The validation metrics are not generated as the calculation fails due to the inadequate sample size
Pass/Fail:		pass
Comments:		

Test Case QA4SM_VA_metrics - Test scaling

Description:	Test the effect of rescaling on the metric output
Test Data:	C3S SM combined (v201912), ISMN (20210131 global)
Preconditions:	-
Requirement(s) to be tested:	-

Test Steps		Expected Results
1.	Go to the "Validate" page and run a validation with the default settings and the option "Mean/standard deviation" in the "Scaling" field	All metrics are calculated normally. The following metrics scores should be equal to 0: • Bias • Mean square error bias • Mean square error variance
2.	Run a validation with the default settings and the option "None" in the "Scaling" field	All metrics are calculated normally. The following metrics scores should also have a non-zero value: • Bias • Mean square error bias • Mean square error variance
3.	Compare the ouput of step 1. and 2.	The correlation scores for the metrics: • Pearson's r • Spearman's rho should be unchanged
Pass/Fail:		pass
Comments:		The scaling option is not called „None" but „No scaling".

Test Case QA4SM_VA_metrics - Test anomalies calculation

Description:	Indirect verification of the anomalies validation routine by testing that the statistical distance between the calculated anomalies of the spatial reference and candidate data sets is smaller than that between the bulk signals.
Test Data:	C3S SM combined (v201912), ISMN (20210131 global)
Preconditions:	-
Requirement(s) to be tested:	-

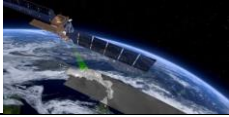| Test Steps		Expected Results
1.	Go to the "Validate" page and run a validation with the default settings (in	All metrics are calculated normally

	principle, any two data sets expressed in the same measuring units can be selected as spatial reference and candidate for this test). Set the "Scaling" field to "None".	
2.	Repeat the validation of step 1. using "Climatology" in the field "Anomalies" and a long term climatology window of at least 20 years.	Verify that the output scores of: • Bias • Mean Square Error are smaller compared to the output of step 1.
Pass/Fail:	Pass	
Comments:	The scaling method name is 'No scaling' and not 'None'	

6.2 Testing of the graphical user interface

6.2.1 Creating and managing user accounts

Test Case QA4SM_GUI_CMUA_001 - Test Sign-up Form

Description:	Testing weather the sign up form behaves as expected and if signing up a new user is possible.
Test Data:	
Preconditions:	
Requirement(s) to be tested:	

Test Steps		**Expected Results**

1.	Go to test.qa4sm.eu/ui//signup	A sign-up form renders, containing:
		
		 • fields:
		  • Username,
		  • Password,
		  • Password confirmation,

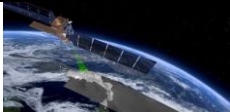|		• Email address,
		• First name,
		• Last name,
		• Organisation,
		• Country (a drop-down list),
		• ORCID,
		• a check-box to accept terms, conditions and privacy policy (unchecked);
		• submit button (disabled).
		There is a question mark icon on the right side of each field, providing additional information on the particular field when hovering over.
2.	Insert values: Username: username, Password: username, Password confirmation: username1, Email address: provide an existing email address, check the terms check-box	'Submit' button enabled.
3.	Click the 'Submit' button	'The two password fields didn't match.' warning shows up under the password confirmation field.
4.	Correct : Password confirmation: username and click 'Submit' again	'The password is too similar to the username. This password is too common.' warning shows up under the password confirmation field.re
5.	Correct: Password: pass, Password confirmation: pass	'This password is too short. It must contain at least 8 characters. This password is too common.' warning shows up under the password confirmation field.

6.	Correct: Username: user1 , Password: *choose a strong password that suits all the requirements,* Password confirmation: *confirm the chosen password* Insert: First name: John, Last name: Smith, Organisation: Some University, Country: choose any country, ORCID: 0000-0000, click the 'Submit' button	'Invalid ORCID identifier' warning shows up under the ORCID field.
7.	Correct: ORCID: 0000-0000-0000-0000 and click 'Submit'	• Form submitted; • redirection to the page test.qa4sm.eu/ui//signup-complete containing: • notification: 'Thank you for signing up with the QA4SM Validation Service. You will be notified via email **as soon as your account is activated** by an admin.' displayed; • 'Back to landing page' button.
8.	Click the 'Back to landing page' button	Redirection to the home page.
9.	Click the 'Sign up' button to go to the sign up page again and insert: Username: user1 , Password: anyPassword, Password confirmation: anyPassword,	'A user with that username already exists' warning shows up under the username field.

10.	Email address: username@something.at, and check the terms check-box, and click 'Submit'	
	Click on the link 'terms, conditions and privacy policy ' next to the terms check-box	Redirection to test.qa4sm.eu/ui//terms page.

Pass/Fail:	Pass
Comments:	

Test Case QA4SM_GUI_CMUA_002 - Test Log in page

Description:	Testing if log in form works as expected
Test Data:	None
Preconditions:	Existing user with data provided in the Test Sign-up Form and activated by an admin
Requirement(s) to be tested:	Test Case QA4SM_GUI_CMUA_001

Test Steps		Expected Results
1.	Go to the test.qa4sm.eu and click the Log in button	A form rendered containing: - a title: 'Please sign in'; 'Username' and 'Password' fields; 'Sign in' button; - link to retrieve a password.
2.	Provide data: Username: test_user, Password: test_user and click the 'Sign in' button	A pop-up window shows up informing about an unsuccessful login attempt.

3.	Click on the 'Forgotten your password?' link	Redirection to the test.qa4sm.eu/ui/password-reset page
4.	Go back to the login page and provide data: Username: user1, Password: *password chosen while registering*,	User logged in, redirection to the home page.

Pass/Fail:	Pass
Comments:	

Test Case QA4SM_GUI_CMUA_003 - Test Password Reset

Description:	Testing if the password reset procedure works properly
Test Data:	None
Preconditions:	Existing user with data provided in the Test Sign-up Form and activated by an admin (remark: the email address must be an existing one); User logged out from the current account - choose 'Profile' => 'Log out' from the navigation bar.
Requirement(s) to be tested:	Test Case QA4SM_GUI_CMUA_001, Test Case QA4SM_GUI_CMUA_002

Test Steps		Expected Results
1.	Go to test.qa4sm.eu/ui/login page and click on the 'Forgotten your password?" link	Redirection to test.qa4sm.eu/ui/password-reset page with an email field and disabled 'Reset my password' button.
2.	Provide a fake email: someUser@some.domain.com and click 'Reset my password' button	Information that there is no user with the given email.

3.	Provide the right email (the one given while registering) and click 'Reset my password' button	Redirection to the page test.qa4sm.eu/ui/password-reset-done, displaying information that password reset instructions have been sent.
4.	Log in to your email address	An emailed entitled '[QA4SM] Password reset for QA4SM webservice' received.
5.	Click on the link in the email	Redirection to the page test.qa4sm.eu/ui/set-password with two input fields (for password and password confirmation) and 'Change my password' button (disabled).
6.	Provide data: Password: 123456 Password confirmation: 123456	'Change my password' button enabled.
7.	Click the 'Change my password' button	Information: "This password is too short. It must contain at least 8 characters. This password is too common. This password is entirely numeric." displayed.
8.	Remove given password and provide a new one: Password: xYbn89Kg Password confirmation: xYbn89	'Change my password' button still disabled.
9.	Choose your own password that fulfill all requirements and confirm it.	'Change my password' button enabled.
10.	Click the 'Change my password' button	'Password change' notification displayed, redirection to the login page.

11.	Log in with the new password	Successful logging in.
12.	Go to your email account and open the link for password resetting again	Information about invalid link displayed with a link to resetting password form provided.
Pass/Fail:	Pass	
Comments:		

Test Case QA4SM_GUI_CMUA_004 - Test User Profile page

Description:	Testing if User Profile renders properly and if it behaves as expected
Test Data:	None
Preconditions:	Existing user account with data provided in the Test Sign-up Form and activated by an admin
Requirement(s) to be tested:	Test Case QA4SM_GUI_CMUA_001, Test Case QA4SM_GUI_CMUA_002

Test Steps		**Expected Results**

1.	Go to test.qa4sm.eu/ui/user-profile	The User Profile renders, containing:
		• fields:
		   • Username (disabled),
		   • Password (empty),
		   • Password confirmation (empty),
		   • Email address,
		   • First name,
		   • Last name,
		   • Organisation,
		   • Country (a drop-down list),
		   • ORCID,
		• 'Save' button (enabled),
		• 'Deactivate my account' button (enabled).
		
		There is a question mark icon on the right side of each field, providing additional

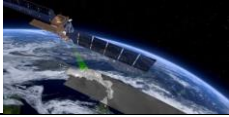|		information on the particular field when hovering over. Apart from the password fields, all the fields are filled with the information given while registering.
2.	Provide a new password, but without confirming it and click the 'Save' button.	'Password do not match' warning shows up under the password field.
3.	Provide the same password confirmation and click the 'Save' button.	A pop-up window with information that the user profile has been updated shows up; Password saved.
4.	Remove email address	Email address field has red border and red label; 'Save' button disabled.
5.	Press ctrl+z to retrieve the email address	'Save' button enabled again.
6.	Remove values from 'First name', 'Last name' and 'Organisation' fields and click the 'Save' button	A pop-up window with information that the user profile has been updated shows up.
7.	Remove four last digits from the ORCID number and click the 'Save' button.	'Invalid ORCID identifier' warning displayed below the ORCID field.
8.	Remove the ORCID number entirely and click the 'Save' button.	A pop-up window with information that the user profile has been updated shows up.
9.	Expand the 'Country' list, choose any country and click the 'Save' button	A pop-up window with information that the user profile has been updated shows up;
10.	Click the 'Deactivate my account' button	An email with information that the account has been deactivated received; Redirection to the page test.qa4sm.eu/ui/deactivate-user-complete, with: • notification that the account will be removed within the next 7 days;

		• a link to the landing page.
11.	Click the link 'Back to landing page'	Redirection to test.qa4sm.eu/ui/home
12.	Click the 'Log in' button on the home page and try to log in with the previous data	A pop-up window shows up, informing that the log in attempt failed.
Pass/Fail:	Pass	
Comments:		
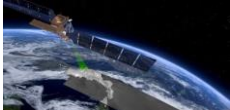
6.2.2 Home page and Navigation Bar

Test Case QA4SM_GUI_HPNB_001 - Test Home Page - as a not logged in user

Description:	Testing whether home page renders properly and if all the buttons and links redirect to appropriate pages
Test Data:	None
Preconditions:	User logged out from the current account - choose 'Profile' => 'Log out' from the navigation bar.
Requirement(s) to be tested:	

Test Steps		Expected Results

1.	Go to test.qa4sm.eu page	Page renders with:
		
		• a navigation bar at the top;
		• a carousel with:
		    • three images and attributions in the lower left corner of each image;
		    • carousel caption containing:
		        • title: 'Quality Assurance for Soil Moisture';
		        • description: 'Validation of satellite soil moisture products

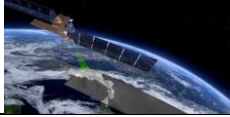|		against in situ and model reference data'; • action buttons: 'See results', 'Sign up' and 'Log in'; • carousel indicators - three white rectangles changing their color to blue when the corresponding picture is displayed; • a row containing partner's logos: FFG, ESA, TU Wien Geo, AWST; • a row containing a short overview of the service functionalities and information on the financial support (left side) and a diagram explaining the platform workflow (right side); • a row containing three screenshots depicting main service functionalities, entitled: 'data set and Settings', 'Results' and 'Download and Visualize'; • a footer containing information about page creators and link to appropriate pages; • a 'go to top' button in the lower right corner.
2.	Click the 'See results' button	Redirection to the 'Published validations' page (test.qa4sm.eu/ui/published-validations).
3.	Go back to the home page and click 'Sign up' button	Redirection to the sign up form (test.qa4sm.eu/ui/signup).
4.	Go back to the home page and click 'Log in' button	Redirection to the log in form (test.qa4sm.eu/ui/login).
5.	Go back to the home page and click one of the carousel indicators	Carousel picture set to the one corresponding to the chosen indicator.
6.	Choose the middle indicator and click the image attribution of the current image	New tab with an ESA website providing current graphics opens.

7.	Inspect partner's logos - click on each logo	Respective websites open in new tabs: • FFG logo - www.ffg.at; • ESA logo - www.esa.int; • TU Wien Geo - www.geo.tuwien.ac.at; • AWST logo - www.awst.at.
8.	Inspect service overview row - click the link to the international soil moisture network (first bullet point)	The international soil moisture network website (https://ismn.earth) opens in a new tab.
9.	Inspect service overview row - inspect information on the financial support, click both available links	Information lists Austrian Space Application Programme and European Space Agency as two sources of financial support. Clicking on provided links opens new tabs with FFG (www.ffg.at) and ESA (www.esa.int) websites respectively.
9.	Click on the workflow diagram	The diagram opens in full-screen.
10.	Close the diagram and inspect the row containing screenshots depicting main service functionalities by clicking on a picture, closing it and clicking on the next one	Pictures opens in full-size. They depict 'Validate' page (data set and Settings), 'My validations' page (Results) and single result page (Download and Visualize).
11.	Inspect the footer	The footer contains: 'Created by TU Wien GEO and AWST' information, with links opening TU Wien Geo and AWST websites in new tabs.
Pass/Fail:	Pass	
Comments:		

Test Case QA4SM_GUI_HPNB_002 - Test Navigation Bar

Description:	Testing whether navigation bar buttons behave in the expected way and if they redirect to proper urls
Test Data:	None

Preconditions:	An existing account and a logged in user.
Requirement(s) to be tested:	Test Case QA4SM_GUI_HPNB_001

Test Steps		**Expected Results**
1.	Inspect navigation bar buttons	Five buttons: • 'Home' with a home icon, • 'Validate' with a check-box icon, • 'My validations' with a folder icon, • 'Published validations' with a globe icon, • 'Compare validations' with an icon consisting of four squares, • 'My datasets' with an upload icon, two drop-down lists: • 'Info' with an i-in-circle icon, • 'Profile' with a user icon, (all buttons and drop-down list have yellow border, yellow font color and blue background color); and 'QA4SM Validation Service' link title in the right side of the menu bar.
2.	Hover over each single button and drop-down list	Button's/drop-down list's background color changes to yellow and font color changes to white.
3.	Click the 'Validate' button	Button's background color changes to yellow and font color changes to white. Redirection to page: test.qa4sm.eu/ui/validate.
4.	Click the 'My validations' button	Button's background color changes to yellow and font color changes to white. Redirection to page: test.qa4sm.eu/ui/my-validations.

5.	Click the 'Published validations' button	Button's background color changes to yellow and font color changes to white. Redirection to page: test.qa4sm.eu/ui/published-validations.
6.	Click the 'Compare validations' button	Button's background color changes to yellow and font color changes to white. Redirection to page: test.qa4sm.eu/ui/comparison.
7.	Click the 'My datasets' button	Button's background color changes to yellow and font color changes to white. Redirection to page: test.qa4sm.eu/ui/my-datasets.
8.	Click on the 'Info' drop-down list	A drop-down list expanded with following options: • 'About' with an i-letter icon, • 'Help' with a question mark icon, • 'User Manual' with a book icon, • 'Datasets' with a floppy disc icon, • 'Terms' with a briefcase icon. The list has a white background, the option's background changes to yellow when hovering over it.
9.	Expand the 'Info' drop-down list and choose 'About'	Redirection to page test.qa4sm.eu/ui/about.
10.	Expand the 'Info' drop-down list and choose 'Help'	Redirection to page test.qa4sm.eu/ui/help.
11.	Expand the 'Info' drop-down list and choose 'User Manual'	A new tab with the QA4SM user manual opens.
12.	Expand the 'Info' drop-down list and choose 'data sets '	Redirection to page test.qa4sm.eu/ui/datasets.
13.	Expand the 'Info' drop-down list and choose 'Terms'	Redirection to page test.qa4sm.eu/ui/terms.

14.	Click on the 'Profile' drop-down list	A drop-down list expanded with following options: • 'User profile ' with a user icon, • 'Log out' with a log-out icon, • 'Log in' with a log-in icon (disabled as the user is logged in) The list has a white background, the option's background changes to yellow when hovering over it.
15.	Expand the 'Profile' drop-down list and choose 'User profile '	Redirection to page test.qa4sm.eu/ui/user-profile.
16.	Expand the 'Profile ' drop-down list and choose 'Logout '	Redirection to home page.
17.	Expand the 'Profile' drop-down list	'User profile' and 'Logout' options disabled, 'Login' option enabled.
18.	Expand the 'Profile ' drop-down list and choose 'Log in'	Redirection to test.qa4sm.eu/ui/login page.
19.	Click the 'QA4SM Validation Service' title	Redirection to the home page.
Pass/Fail:	Pass	
Comments:	There is additional option to choose on the 'Info' drop down list, called 'Upload Data Help' which redirects to the test.qa4sm.eu/ui/user-data-guidelines page	

Test GUI - Test Log in page - redirection after logging in

Description:	Testing if after logging in the user is redirected to the appropriate page
Test Data:	None
Preconditions:	An existing account, with a logged out user

Requirement(s) to be tested:	Test Case QA4SM_GUI_CMUA_001, Test Case QA4SM_GUI_CMUA_002, Test Case QA4SM_GUI_HPNB_002

	Test Steps	Expected Results
1.	Go to the test.qa4sm.eu and click the 'Log in' button on the main page; provide your data after redirection to the log in form	User logged in, redirection to the home page.
2.	Choose 'Log out' from the 'Profile' drop-down list and go to the 'Validate' page	Redirection to the login page.
3.	Provide your data	User logged in, redirection to the 'Validate' page.
4.	Choose 'Log out' from the 'Profile' drop-down list and go to the 'My validations' page	Redirection to the login page.
5.	Provide your data	User logged in, redirection to the 'My validations' page.
6.	Choose 'Log out' from the 'Profile' drop-down list and go to the 'Compare validations' page	Redirection to the login page.
7.	Provide your data	User logged in, redirection to the 'Compare validations' page.
8.	Choose 'Log out' from the 'Profile' drop-down list and go to the 'My datasets' page	Redirection to the login page.
9.	Provide your data	User logged in, redirection to the 'My datsets' page.

10.	Choose 'Log out' from the 'Profile' drop-down list and then choose 'Log in' from the same list	Redirection to the log in page.
11.	Provide your data	Redirection to the 'User profile' page.
Pass/Fail:	Pass	
Comments:		

6.2.3 Validate Page

Test Case QA4SM_GUI_VP_001 - Test Validate Page - Render Validate Page

Description:	Testing whether the 'Validate' page renders properly
Test Data:	None
Preconditions:	An existing account and a logged in user.
Requirement(s) to be tested:	

Test Steps		**Expected Results**
1.	Go to the 'Validate' page (test.qa4sm.eu/ui/validate/)	Rendered components and buttons: • Data, • Reference, • Map, • Spatial Subsetting, • Temporal Subsetting, • Metrics, • Anomalies, • Scaling, • Name your validation field, • Validate button.
		All the components (apart from the Map and Validation name field) have a question

| mark icon and a collapse button in the right corner of their headers. |
| The Map component has a collapse button only. |
| The 'Name your validation' field is followed by a question mark icon providing additional help when hovering over. |

Pass/Fail:	Pass
Comments:	The scaling component is now placed next to the Reference component, not between Anomalies and Name field, what is suggested by the order in point 1.

Test Case QA4SM_GUI_VP_002 - Test Validate Page - Render Validate Page: Inspect Data Component

Description:	Testing whether the 'Data' component on the 'Validate' page behaves as expected
Test Data:	None
Preconditions:	An existing account and a logged in user.
Requirement(s) to be tested:	Test Case QA4SM_GUI_VP_001

Test Steps		Expected Results
1.	Go to the 'Validate' page (test.qa4sm.eu/ui/validate/) - render default settings	'Data' component renders with two datasets: 1. ISMN data set: • version 20210131 global, • variable 'soil_moisture', • List of available filters with a 'Variable in valid geophysical range' and 'Quality flag is "good" (G)' filters switched on, and two parameterized

		filters ('Use ISMN networks' and 'Use measurements at given depth') switched on and disabled. 2. C3S SM combined data set, • version v202012 • variable 'sm', • List of available filters with a 'Variable in valid geophysical range' filter switched on, • 'Remove data set' red button (disabled), • 'Add data set' green button (enabled). ISMN dataset is marked as spatial and temporal reference.
2.	Expand the variable drop-down list for the C3S data set	Only variable 'sm' available.
3.	Expand the version drop-down list for the C3S data set	Versions: v201706, v201812, v201912 and v202012 available.
4.	Choose version v201706 from the version drop-down list	Date in the field 'To' of the 'Validation Period' component changes to '2017-06-30'.
5.	Check the 'Ascending mode only' filter and then check the 'Descending mode only' filter	Checking the 'Descending mode only' filter unchecks the 'Ascending mode only' automatically.
6.	Expand the C3S data set drop-down list and inspect the list of available data sets	There is no ISMN dataset available.
7.	Expand the ISMN data set drop-down list and choose 'GLDAS Noah'	• The list of available versions changes - NOAH025 3H.2.1 set; • the list of available variables changes - SoilMoi100_200cm_inst set; • the list of available filters changes;

8.	Expand the C3S data set drop-down list and inspect the list one more time	The ISMN data set available again.
9.	Click on the 'Add data set' button	• A new data set tab with default C3S data set settings added below; • 'Remove data set' button enabled; • 'Include triple collocation metrics' checkbox in the 'Metrics' component enabled.
10.	Inspect the 'Dataset' dropdown list of the newly added data set	ISMN data set available.
11.	Click on the 'Add data set' button three more times	• three additional dataset tabs added with default C3S data set settings; • 'Add dataset' button disabled.
12.	Click the 'Remove data set' button on the GLDAS Noah data set tab	• Tab removed; • 'Add data set button' enabled again; • validation period changes to 1978-11-01 (From) and 2020-12-31 (To); • First data set on the list marked as spatial and temporal reference.
13.	Remove three other data set tabs	• 'Remove data set' button disabled; • 'Include Triple Collocation Metrics' checkbox disabled.
14.	Change one data set to ISMN	ISMN set as spatial reference.
15.	Inspect 'Dataset' list of the C3S tab	ISMN unavailable.
15.	Click 'select...' next to the 'Use ISMN Networks' filter	A dialog window opens, with a list of available ISMN networks, ordered by continents.

16.	Click on 'Europe' checkbox and expand the list using the arrow on the left side	The checkbox next to 'Europe' is marked as checked, the same as all the networks assigned to Europe.
17.	Close the dialog window and open it one more time	The dialog window looks the same as before closing - the same networks chosen, the same list folded and expanded.
18.	Close the 'Select networks' dialog window and click 'select...' next to the 'Use measurements at given depth' filter	A dialog window opens, with two fields for introducing depth from and depth to and an 'OK' button. Fields 'Depth from' and 'Depth to' are set by default to 0.0 and 0.1 m.
19.	Change 'Depth to' to 0.2 m and click the 'OK' button. Open dialog window again.	After clicking the 'OK' button the dialog window closes. When opened again the field 'Depth to' is set to 0.2 m.

Pass/Fail:	Step 12 failed
Comments:	The current default ISMN version is 20230110 and the default variable for GLDAS is now SoilMoi0_10cm_inst, i.e. newer than in the test plan.
	ISMN dataset has an additional filter assigned: 'Include only representative sensors (0-10 cm)'.
	In step 12, the date labeled "To" won't change, as C3S v2017 is still chosen. The problem does not come from a software problem, but from a missing step in this test case.
	The test will be updated in a future version of the test plan to address these described issues.

Test Case QA4SM_GUI_VP_003 - Test Validate Page - Render Validate Page: Inspect Reference Component

Description:	Testing whether the 'Reference ' component on the 'Validate' page behaves as expected
Test Data:	None
Preconditions:	An existing account and a logged in user
Requirement(s) to be tested:	Test Case QA4SM_GUI_VP_001

Test Steps		Expected Results
1.	Go to the 'Validate' page (test.qa4sm.eu/ui/validate/)	Component renders with two dropdown lists labeled 'Spatial reference' and 'Temporal reference' and ISMN data set chosen by default in both lists. In the 'Data' component, the ISMN data set is marked as spatial and temporal reference.
2.	Expand the spatial reference dropdown list	There is only ISMN dataset available.
3.	Expand the temporal reference dropdown list	There are both ISMN and C3S data sets available.
4.	Choose C3S as the temporal reference data set	C3S data set in the component 'Data' marked as temporal reference.
5.	Change ISMN data set in the 'Data' component to 'GLDAS Noah'	GLDAS Noah set as the spatial reference.
6.	Expand the spatial reference dropdown list	There are both GLDAS and C3S data sets available.
7.	Add another data set in the 'Data' component and inspect both spatial	There are all three data sets available on both lists;

	and temporal reference dropdown lists – expand and hover over each data set on the list	When hovering over a data set, the proper tab in the 'Data' component gets highlighted.
8.	Choose one of C3S data sets as the spatial reference	Proper C3S data set in the component 'Data' marked as spatial reference.
9.	Change version of the C3S data set that is chosen as the spatial reference	Proper version set in the 'Reference component'.
10.	Change GLDAS data set to ISMN	ISMN set as the spatial reference, C3S left as the temporal reference, information in data sets tab headers updated.
Pass/Fail:	Pass	
Comments:		

Test Case QA4SM_GUI_VP_004 - Test Validate Page - Render Validate Page: Inspect Map Component

Description:	Testing whether the 'Map' component on the 'Validate' page behaves as expected
Test Data:	None
Preconditions:	An existing account and a logged in user
Requirement(s) to be tested:	Test Case QA4SM_GUI_VP_001

Test Steps		**Expected Results**
1.	Go to the 'Validate' page (test.qa4sm.eu/ui/validate/) and expand the 'Map' component	Component renders collapsed and contains: • a map with: a pencil icon in the upper left corner, • attribution icon in the lower right corner,

		• bounding box covering Europe, defined by the coordinates given in the 'Spatial Subsetting' component.
2.	Click on the pencil icon in the upper left corner	• Bounding box cleared; • a blue circle with white border shows up next to the cursor.
3.	Click in any place on the map and move the cursor	Drawing a bounding box starts.
4.	Draw a bounding box of any size covering North America and Europe, and click on the map again	• A rectangle bounding box added on the map; • coordinates in the Spatial Subsetting' component updated with values corresponding to the lower left and upper right corners of the drawn rectangle.
5.	Change data set in the 'Data' component to CGLS SSM 1km	• An alert shows up informing that the chosen spatial subsetting is bigger than the one covered by chosen data sets and that the bounds are corrected to fit available subsetting; • bounding box adjusted to the region covered by the chosen dataset.
6.	Draw a new bounding box, bigger than the current one	• The alert shows up again; • bounding box is adjusted to the available spatial subsetting.
7.	Change data set back to C3S SM combined and draw any bounding box	• Chosen bounding box drawn, • updated coordinates in the 'Spatial Subsetting' component.
8.	Click on the button in the lower right corner of the map	An attributions field expands with a link to the Open Street Maps.
Pass/Fail:	Pass	
Comments:		

Test Case QA4SM_GUI_VP_005 - Test Validate Page - Render Validate Page: Inspect Spatial Subsetting Component

Description:	Testing whether the 'Spatial Subsetting ' component on the 'Validate' page behaves as expected
Test Data:	None
Preconditions:	An existing account and a logged in user
Requirement(s) to be tested:	Test Case QA4SM_GUI_VP_001

Test Steps		Expected Results
1.	Go to the 'Validate' page (test.qa4sm.eu/ui/validate/)	Component renders with: • four input fields - latitude and longitude of the lower left and upper right corners, labeled respectively; • a trash bin button between fields; • default values of the latitude and longitude of the lower left corner set to 34 and -11.2 degrees respectively; • default values of the latitude and longitude of the upper right corner set to 71.6 and 48.3 degrees respectively.
2.	Click on the trash bin button and expand the 'Map' component	• Coordinates field cleared; • bounding box removed from the map.
3.	Input following coordinates: Lat = 40, Lon = -10 for the lower left corner and Lat = 60, Lon = 40 for the upper right corner	Bounding box on the map drawn, corresponding to the given coordinates.
4.	Change the latitude of the lower left corner to 70 degrees	The latitude set to 60 degrees, as it can not exceed the latitude of the upper right corner.

5.	Change data set to CGLS SSM 1km	• An alert shows up informing that the chosen spatial subsetting is bigger than the one covered by chosen data sets and that the bounds are corrected to fit available subsetting; • spatial subsetting values change to 35.004 and -10.996 degrees for the latitude and longitude of the lower left corner and to 71.996 and 49.996 degrees for the latitude and longitude of the upper right corner.
6.	Input following coordinates: Lat = 40, Lon = -10 for the lower left corner and Lat = 60, Lon = 40 for the upper right corner	The bounding box updated.
7.	Click on the trash bin button	Coordinates and the bounding box adjusted to the default spatial subsetting assigned to the chosen data set,
8.	Change the latitude of the lower left corner to 30 degrees	Value changed back to the default one as the introduced value is smaller than the minimum possible.

Pass/Fail:	Pass
Comments:	Setting the minimum and maximum latitude or longitude to the same values will lead to no results if computations are started. In terms of consistency of the user interface, however, the maximum value can not be smaller than the minimum, so there is some space for the user awareness when choosing settings. In steps 3 and 4 instead of using values 60 and 70 degrees values 72 and 75 should be used respectively, because the HR data have the latitude limit 71.996, so to see the described effect the value should be changed. The test case will be updated.

Test Case QA4SM_GUI_VP_006 - Test Validate Page - Render Validate Page: Inspect Validation Period Component

Description:	Testing whether the 'Validation period ' component on the 'Validate' page behaves as expected
Test Data:	None
Preconditions:	An existing account and a logged in user
Requirement(s) to be tested:	Test Validate Page

Test Steps		Expected Results
1.	Go to the 'Validate' page (test.qa4sm.eu/ui/validate/)	Component renders with two sub-fields named Validation Period and Temporal matching. The Validation Period part consists of two date fields described with From and To labels, with default dates set to 1978-11-01 (From) and 2020-12-31 (To). The Temporal matching field contains an input field labeled 'Window size' and two arrows enabling changing values in the field. By default, 12 hours is set.
2.	Click first on the 'From' field and then on the 'To' field	In both cases a calendar widget shows up.
3.	Expand the calendar on the 'From' field and choose January from the month drop-down list and 2021 from the year drop-down list	No date to choose available, as the date 'To' is set to be 2020-12-31
4.	Change date 'From' to 2020-11-01, click on the calendar on the field 'To' and choose October from the month drop-down list	No date available as the date 'From' is set to 2020-11-01;

5.	Press upper arrow next to the 'Window size' field and hold it	Value increases to 24 hours and stops.
6.	Press lower arrow next to the 'Window size' field and hold it	Value decreases to 1 hour and stops.
7.	Insert to the 'Window size' field value bigger than 24 hours	Value 24 hours set.
8.	Insert to the 'Window size' field value smaller than 1 hour	Value 1 hour set.
9.	Try to insert value 1.5 hours	Value 15 hours set, as there is no possibility to introduce a decimal separator.
Pass/Fail:	Pass	
Comments:		

Test Case QA4SM_GUI_VP_007 - Test Validate Page - Render Validate Page: Inspect Metrics Component

Description:	Testing whether the 'Metrics ' component on the 'Validate' page behaves as expected
Test Data:	None
Preconditions:	An existing account and a logged in user
Requirement(s) to be tested:	Test Case QA4SM_GUI_VP_001

Test Steps		**Expected Results**
1.	Go to the 'Validate' page (test.qa4sm.eu/ui/validate/)	Component renders with two checkboxes, described as 'Include Triple Collocation Metrics' and 'Bootstrap Triple Collocation metric confidence intervals (Warning: very slow)', both descriptions are followed by question mark icons providing additional

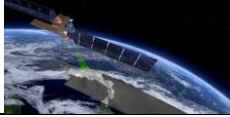|		information when hovering over, both checkboxes are disabled.
2.	Click on the 'Add data set' button in the 'Data' component	• 'Include Triple Collocation Metrics' enabled, • 'Bootstrap Triple Collocation metric confidence intervals' disabled;
3.	Check 'Include Triple Collocation Metrics' box	'Bootstrap Triple Collocation metric confidence intervals' enabled;
4.	Check 'Bootstrap Triple Collocation metric confidence intervals' box and uncheck 'Include Triple Collocation Metrics' box	Both boxes get unchecked;
5.	Check both checkboxes and remove a data set, so only one data set is left	Both boxes get unchecked;
Pass/Fail:	Pass	
Comments:		

Test Case QA4SM_GUI_VP_008 - Test Validate Page - Render Validate Page: Inspect Anomalies Component

Description:	Testing whether the 'Anomalies' component on the 'Validate' page behaves as expected
Test Data:	None
Preconditions:	An existing account and a logged in user
Requirement(s) to be tested:	Test Case QA4SM_GUI_VP_001
Test Steps	**Expected Results**

1.	Go to the 'Validate' page (test.qa4sm.eu/ui/validate/)	Component renders with one drop-down list labeled Method, with the 'Do not calculate' set by default.
2.	Expand the Method drop-down list and choose 'Climatology'	Two empty input date fields show up, labeled 'From' and 'To'.
3.	Expand the calendar widget on the 'From' field, choose January 10, 2021, expand the calendar widget on the 'To' field and choose January 2021	The earliest possible date to choose is January 10, 2021.
4.	Choose January 20, 2021 on the calendar widget on the 'To' field and expand the calendar widget on the 'From' field	The latest possible date to choose is January 20, 2021.
5.	Expand the Method drop-down list and choose '35 day moving average'	Both date selection fields disappear again.

Pass/Fail:	Pass
Comments:	The Anomalies component has been updated, as it is enough to provide dates with accuracy of a year and not day or month, therefore steps 2 to 4 have no application anymore.

Test Case QA4SM_GUI_VP_009 - Test Validate Page - Render Validate Page: Inspect Scaling Component

Description:	Testing whether the 'Scaling ' component on the 'Validate' page behaves as expected
Test Data:	None
Preconditions:	
Requirement(s) to be tested:	Test Case QA4SM_GUI_VP_001

	Test Steps	Expected Results
1.	Go to the 'Validate' page (test.qa4sm.eu/ui/validate/)	Component renders with one drop-down list, labeled Method, with the 'No scaling' set by default
2.	Expand the dropdown list and choose any other method	A 'Scale to' dropdown list shows up, with ISMN data set chosen by default. The ISMN data set marked as scaling reference in the 'Data' component.
3.	Expand the 'Scale to' dropdown list and hover over each option	Both ISMN and C3S datasets (those available in the 'Data' component) shown on the list. When hovering over an option the respective tab in the 'Data' component gets highlighted.
4.	Choose C3S as the scaling reference	C3S data set marked as scaling reference in the 'Data' component
5.	Add another data set in the 'Data' component and remove the C3S component set as scaling reference	First data set on the list set as the scaling reference; Proper information added to the data set tab header in the 'Data' component.
6.	Expand the 'Method' drop-down list and choose method 'No Scaling'	'Scale to' drop-down list vanishes. None of data set is marked as scaling reference;
Pass/Fail:	Pass	
Comments:		

Test Case QA4SM_GUI_VP_010 - Test Validate Page - Start a default validation

Description:	Testing service behavior after starting a validation
Test Data:	None
Preconditions:	
Requirement(s) to be tested:	Test Case QA4SM_GUI_VP_001 to 009

Test Steps		Expected Results
1.	Go to the 'Validate' page (test.qa4sm.eu/ui/validate/), do not change any settings, enter name 'test validation' in the 'Name your validation' field, click the 'Validate' button.	Depending on validations existing in our database it is possible that a modal window opens informing that there exists already a validation with the given settings. It may refer to a validation of the current user, to a validation of another user or the a published validation. There are two buttons on the modal window: 'See the existing validation' and 'Run your own validation'. If the window showed up, go to point 2, if not to point 3.
2.	If the modal window shows up: Click the 'See the existing validation' button	Redirection to the page with results of the existing validation.
3.	Get back to the 'Validate' page, enter name 'test validation' in the 'Name your validation' field and click the 'Validate' button. If the modal window shows up, choose 'Run your own validation'	Redirection to the page test.qa4sm.eu/ui/validation-result/validation_id (where validation_id is a random universal unique identifier) with information about running validation.
4.	Wait until receiving an email about finished validation and refresh the	Validation results displayed in three separate components entitled 'Summary:

page (or go to it using the link given in the email)	test validation', 'Summary statistics' and 'Result files'.
Pass/Fail:	Pass
Comments:	If the dialog window about an existing validation does not show up, just run the same validation one more time. For now, our current state of the validations stored in the database depends on the user activity.

6.2.4 My Validations Page

Test Case   QA4SM_GUI_MVP_001 - Test My Validations Page - Render My Validations Page

Description:	Testing whether the 'My Validations ' page renders properly
Test Data:	None
Preconditions:	Start 3 different validations with the default settings, but a different spatial reference dataset (to be able to choose a different spatial reference dataset than ISMN, there can not be ISMN in dataset pool), give each validation a different name (e.g. test_val_1, test_val_2, test_val_3);
Requirement(s) to be tested:	Test Case QA4SM_GUI_VP_010

Test Steps		**Expected Results**
1.	Go to the 'My validations' page (test.qa4sm.eu/ui/my-validations)	Page renders with: • a title - 'My validations'; • a sorting form, labeled 'Sort by', is rendered with two drop-down lists - one for the sorting feature (by default set to 'Date') and the second one for the order (by default set to 'descending');

		• validation results rows are rendered for each validation that has been run; each row contains: • a header with icon (depending on the validation status it can be a calendar, a spinner, a triangle with an exclamation mark, a book, a box or a ban symbol), date and time of the validation start (in case the validation is a copy of the original, the copy date), validation name and button to change it (or a question mark icon, if the validation has been published); • a body containing information on data sets used for validations, validation status and action buttons; Remark: If more than 10 validations have been run, pagination is added and displayed below validation rows.
2.	Change sorting feature to 'Name'	Validation rows ordered by validation names in the descending order;
3.	Change order to 'ascending'	Validation rows ordered by validation name in the ascending order;
Pass/Fail:	Pass	
Comments:		

Test Case QA4SM_GUI_MVP_002 - Test My Validations Page - Inspect Validation Results Row

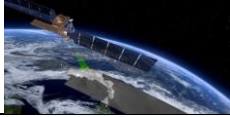| **Description:**	Testing whether Validation Results Row renders properly and if the action buttons works as they should
Test Data:	None

Preconditions:	Run your own validation with default settings, GLDAS Noah data set as the spatial and temporal reference and name 'default_validation', wait until it finishes;
Requirement(s) to be tested:	Test Case QA4SM_GUI_VP_010, Test Case QA4SM_GUI_MVP_001

Test Steps		**Expected Results**

1.	Go to the 'My validations' page (test.qa4sm.eu/ui/my-validations)	Validation row with the newest validation rendered at the top. The row's header contains:
		
		• a calendar icon,
		• date and time of the validation start,
		• name: default_validation followed by a pencil button;
		
		The row's body contains:
		
		• Data:
		
		 1. 'C3S, v202012, sm',
		
		 2. GLDAS Noah, NOAH025 3H.2.1, SoilMoi100_200cm_inst (spatial, temporal reference)
		
		• Status: 'Done';
		• action buttons:
		 • 'View results' folder button,
		 • 'Archive' box button,
		 • double arrow drop-down list with options:
		     ○ Download graphs,
		     ○ Download NetCDF File,
		     ○ Load Validation Settings,
		     ○ Delete Validation Run,
		     ○ Extend Life Span;

2.	Hover over the header calendar icon	Information about the expiration date shows up.

3.	Click on the pencil button next to the validation name	• An input field shows up, filled with the current validation name (empty if there was no name provided); • floppy disc and ban icon buttons displayed.
4.	Click on the ban button	• Input field closed; • name not changed; • pencil button displayed.
5.	Click on the pencil button again and change the validation name to 'test_validation' and press the floppy disc button	• Name changed; • input field closed; • floppy disc and ban icons hidden; • pencil button displayed.
6.	Click on the folder button	Redirection to the page test.qa4sm.eu/ui/validation-result/validation_id (where validation_id is a random universal unique identifier).
7.	Go back to the 'My validations' page and click on the 'Archive' box button and click 'ok' on the pop-up window	• Results archived; • the box button changed into an 'Un-archive' calendar button; • the header calendar icon changed into a box icon.
8.	Hover over the box icon in the header	Information about the validation being archived displayed.
9.	Expand double arrow drop-down list	no 'Extend Life Span' option (not needed for archived results).
10.	Click the 'Un-archive' calendar button and confirm it on the pop-up window	• Results unarchived; • the box icon changed back to the calendar icon; • the calendar button changed back to the box button; • 'Extend Life span' option available again.
11.	Expand the double arrow drop-down list and click on the 'Download Graphs' option	A dialog window for downloading a .zip file opened.

12.	Cancel downloading, expand the double arrow drop-down list and click on the 'Download NetCDF File' option	A dialog window for downloading a result file opened.
13.	Cancel downloading, expand the double arrow drop-down list and click on the 'Load Validation Settings' option	Redirection to the 'Validation' page' with settings of the current validation set.
14.	Go back to the 'My validations' page, hover over the calendar icon and note the date given in the information, expand the double arrow drop-down list and click on the 'Extend Life Span' and confirm it on the pop-up window	New validation expiry date set to be 60 days (including time not only date) from now, the expiry date can be checked when hovering over the calendar icon in the header;
15.	Expand the double arrow drop-down list and click on the 'Delete Validation Run' and confirm it on the pop-up window	• Validation deleted; • appropriate validation row removed; • no access to the results - the link from the email informing about finished validation does not work anymore.

Pass/Fail:	Pass
Comments:	If there are more than 10 validations run, changing pages in pagination bar does not impact the way they are sorted.

Test Case QA4SM_GUI_MVP_003 - Test My Validations Page - Cancel running validation

Description:	Testing if 'Cancel validation' button works as expected
Test Data:	None
Preconditions:	
Requirement(s) to be tested:	Test Case QA4SM_GUI_VP_010

Test Steps		Expected Results

1.	Go to 'Validate' page, choose ERA5-Land as the reference data set, start a new validation and go to 'My validations' page and inspect the running validation row	• Started validation is listed as the first one;
		• there is a spinner icon in the upper left corner of the header;
		• validation status is 'Scheduled';
		• the available action buttons are:
		  • a square 'Cancel validation' button available,
		  • 'View results' folder button,
		  • double arrow drop-down list with options:
		    ○ Delete Validation Run,
		    ○ Extend Life Span;
		
		The rest of the row looks like in case of finished validations.

2.	Wait around one minute and inspect the Status field	Validation status changed to 'Running x%' where x is a integer number between 1 and 100.
3.	Click the 'Cancel validation' button	• Validation stops; • icon changes to a ban symbol; • status changes to 'Canceled'; • 'Cancel validation' button vanishes.
4.	Click the folder button and inspect the validation results page	Only validation summary available, with the same information as for a finished validation, but instead of the number of errors there is information that the validation was canceled. No 'Publish' button available.
Pass/Fail:	Pass	
Comments:		

6.2.5 Managing validation results from Validation Results Page
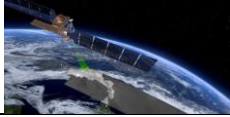
Test Case QA4SM_GUI_VRP_001 - Test Validation Results Page - Render Validation Result Page

Description:	Testing whether the 'Validation Result ' page renders properly
Test Data:	None
Preconditions:	Run a validation with default settings, GLDAS Noah data set as the spatial and temporal reference, and name 'default_validation', wait until it finishes
Requirement(s) to be tested:	Test Case QA4SM_GUI_VP_010

Test Steps		Expected Results
1.	Go to the link from the email informing about the validation has been completed	Validation results displayed in three separate components: • 'Summary: default_validation', followed by a pencil button - containing all the validation settings, used data sets and action buttons (Delete, Renew, Archive, Publish, Load settings), • 'Summary statistics' (collapsed by default) - containing mean, median and standard deviation calculated for all available metrics and a button for downloading a .csv file with the summary; • 'Result files' - containing two plots, two buttons for downloading results (graphs and netCDF file), a drop-down list for choosing a metric to display (by default set to '# observations');
Pass/Fail:	Pass	

Comments:	

Test Case   QA4SM_GUI_VRP_002 - Test Validation Results Page - Inspect Summary component

Description:	Testing whether the 'Summary' component contains all the expected information and if the action buttons behave properly
Test Data:	None
Preconditions:	Run a validation with default settings, GLDAS Noah data set as the reference, and name 'default_validation', wait until it finishes
Requirement(s) to be tested:	Test Case QA4SM_GUI_VP_010, Test Case QA4SM_GUI_VRP_001

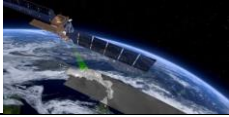| Test Steps		Expected Results
1.	Go to the link from the email informing about the validation has been completed and inspect the 'Summary' component	Information in the summary component body: • validation date and time (start and finish), • list of compared datasets: **1-C3S SM combined (v202012, sm)** Filters: • Variable in valid geophysical range; **0-GLDAS Noah (NOAH025 3H.2.1, SoilMoi0_10cm_inst)** Filters: • Variable in valid geophysical range; • Soil not frozen and no snow-cover; • Spatial filter bounding box: [34.0, -11.2, 71.6, 48.3] • Validation period / temporal filter: Jan. 1, 2000, midnight UTC to Dec. 31, 2017, 11:59 p.m. UTC.

		• Temporal matching window size set to 12 hours. • Validation metrics calculated from absolute values. • Triple collocation analysis was deactivated. • Bootstrapping of confidence intervals for Triple Collocation Analysis was deactivated. • Scaling method: No scaling. • Information on processing time in minutes • for 0% (0 of 21145) of the processed locations (grid points) the validation metrics could not be calculated. • calendar icon and information on the cleanup date.
2.	Click on the pencil button next to the validation name	• An input field shows up, filled with the current validation name (empty if there was no name provided); • floppy disc and ban icon buttons displayed.
3.	Click on the ban button	• Input field closed; • name not changed; • pencil button displayed.
4.	Click on the pencil button again and change the validation name to 'test_validation' and press the floppy disc button	• Name changed, • input field closed, • floppy disc and ban icons hidden, • pencil button displayed;
5.	Click the 'Load Validation Settings' button	Redirection to the 'Validation' page' with settings of the current validation set;
6.	Click the 'Publish' button	A 'Publish results' dialog window shows up
7.	Close the 'Publish' window, click 'Archive' button and confirm it on the pop-up window	• Results archived, • the 'Archive' button changed into an 'Un-archive' calendar button,

		• information on the validation expiry changed into information about results being archived;
8.	Click the 'Un-archive' button and confirm it on the pop-up window	• Results unarchived, • the 'Un-archive' button changed back to the 'Archive' one, • information about expiry date shown again;
9.	Click the 'Renew' button and confirm it on the pop-up window	New validation expiry date set to be 60 days from now (if this option is chosen the same day as the validation was run, there will be no visible effect);
10.	Click the 'Delete' button and confirm it on the pop-up window	• Validation deleted, • user redirected to the 'My validations' page • no access to the results - the link from the email informing about finished validation not working anymore;
Pass/Fail:	Pass	
Comments:		

Test Case QA4SM_GUI_VRP_003 - Test Validation Results Page - Inspect Summary Statistics component for validation with only one data set validated

Description:	Testing whether the 'Summary Statistics' component renders properly and provide a .csv file
Test Data:	None
Preconditions:	Run a validation with default settings, GLDAS Noah data set as the spatial and temporal reference, and name 'default_validation', wait until it finishes;
Requirement(s) to be tested:	Test Case QA4SM_GUI_VP_010, Test Case QA4SM_GUI_VRP_001

Test Steps		Expected Results

1.	Expand the 'Summary Statistic' component	• Short introduction,
		 • table with columns:
		  • 'Metric',
		  • 'Mean',
		  • 'Median',
		  • 'IQ range'
		  • 'data set'
		 and rows referring to the number of observations and following metrics:
		 • Bias,
		 • Mean square error,
		 • Mean square error bias,
		 • Mean square error correlation,
		 • Mean square error variance,
		 • Pearson's r,
		 • Pearson's r p-value,
		 • Residual sum of squares,
		 • Root-mean-square deviation,
		 • Spearman's ρ,
		 • Spearman's ρ value,
		 • Unbiased root-mean-square deviation
		 • 'Download .csv table' button;

2.	Click the 'Download .csv table' button	Dialog window for downloading .csv files opens;
3.	Save the .csv file and open it	File contains exactly the same table as the one rendered on the website.
Pass/Fail:	Pass	
Comments:		

Test Case QA4SM_GUI_VRP_004 - Test Validation Results Page - Inspect Summary Statistics component for validation with multiple data sets validated with triple collocation checked

Description:	Testing whether the 'Summary Statistics' component renders properly and provide a .csv file
Test Data:	None
Preconditions:	Go to the validate page, add another data set and change it to GLDAS Noah, leave ISMN data set as the spatial and temporal reference, check the triple collocation checkbox in Metrics component, leave other settings as set by default, start a validation and wait until it finishes;
Requirement(s) to be tested:	Test Case QA4SM_GUI_VP_010, Test Case QA4SM_GUI_VRP_001

Test Steps		Expected Results
1.	Expand the 'Summary Statistic' component	• Short introduction, • table with columns: • 'Metric', • 'Mean', • 'Median', • 'IQ range' • 'data set' and rows referring to the number of observations and following metrics (separately for each validated data set): • Bias, • Mean square error, • Mean square error bias, • Mean square error correlation, • Mean square error variance, • Pearson's r, • Pearson's r p-value, • Residual sum of squares, • Root-mean-square deviation,

		• Spearman's ρ,
		• Spearman's ρ value,
		• TC scaling coefficient,
		• Unbiased root-mean-square deviation;
		• 'Download .csv table' button;
2.	Click the 'Download .csv table' button	Dialog window for downloading .csv files opens.
3.	Save the .csv file and open it	File contains exactly the same table as the one rendered on the website.
Pass/Fail:	Pass	
Comments:	New metrics added: #observations, Validation Errors, Signal-to-noise ratio in dB, Error standard deviation in m³/m³, TC scaling coefficient in m³/m³, Metrics that are not shown anymore: Pearson's r p-value, Spearman's p value, as there is no point of showing mean and median values of those. The metrics are still calculated and shown in plots.	

Test Case QA4SM_GUI_VRP_005 - Test Validation Results Page - Inspect Result files component for validation with only one data set validated

Description:	Testing whether the 'Result files' component contains all the produced plots and results
Test Data:	None
Preconditions:	Run a validation with default settings, GLDAS Noah data set as the spatial and temporal reference, and name 'default_validation', wait until it finishes;
Requirement(s) to be tested:	Test Case QA4SM_GUI_VP_010, Test Case QA4SM_GUI_VRP_001
Test Steps	**Expected Results**

1.	Inspect 'Result files' component	Component renders with: • two plots - a boxplot on the left side and a map plot on the right side, both referring to the number of observations; • a drop-down list containing names of metrics to be shown on the plots; • two buttons for downloading results - 'Download graphs' on the left side of the drop-down list and 'Dowload NetCDF' on the right side of the drop-down list;
2.	Expand the drop-down list and choose a metric after metric until the end of the list	With every single metric new two plots show up - a boxplot on the left and a map plot on the right - with respective metric name in their titles;
3.	Click on the boxplot	A full-size plot opens as a gallery window
4.	Close the boxplot preview and click on the map plot	A full-size plot opens as a gallery window
5.	Close the map plot window and click the 'Download graphs button'	A dialog window for saving a graphs zip file opens; Default name to save is 'validationId_graphs.zip
6.	Save the file and open it	The zip file contains all the files shown on the website in two formats .png and .svg; Boxplot's file name follows the pattern 'boxplot_metricName'; Map plot's file name follows the pattern: 'overview_referenceName_and_data setName_metricName'.
7.	Click the 'Download NetCDF' button	A dialog window for saving the result netCDF file opens; Default file name to save is '0-GLDAS.SoilMoi0_10cm_inst_with_1-C3S_combined.sm';

8.	Save the netCDF file and check it's size	The saved file should have size around 12 Mb;

Pass/Fail:	Pass

Comments:	There has been a new metric added called '# status', for which there are no box plots available, so if it's chosen, only a map shows up; The file size might be different, as there have been a few more plots added depicting FRM classification and showing source of errors in the validation process.

Test Case QA4SM_GUI_VRP_006 - Test Validation Results Page - Inspect Result files component for validation with multiple data sets validated with triple collocation checked

Description:	Testing whether the 'Result files' component contains all the produced plots and results
Test Data:	None
Preconditions:	Go to the validate page, add another data set and change it to GLDAS Noah, leave ISMN data set as the spatial and temporal reference, check the triple collocation checkbox in Metrics component, leave other settings as set by default, start a validation and wait until it finishes;
Requirement(s) to be tested:	Test Case QA4SM_GUI_VP_010, Test Case QA4SM_GUI_VRP_001

Test Steps		Expected Results
1.	Inspect 'Result files' component	Component renders with: • two plots - a boxplot on the left side and a map plot on the right side, both referring to the number of observations; • a drop-down list containing boxplot classification (note that this is related to

		ISMN dataset used as the spatial reference); • a drop-down list containing names of metrics to be shown on the plots; • two buttons for downloading results - 'Download graphs' on the left side of the drop-down list and 'Dowload NetCDF' on the right side of the drop-down list;
2.	Expand the boxplot classification drop-down list and sequentially select option after option until the end of the list	With every single classification option a boxplot changes.
3.	Expand the drop-down list and sequentially select metric after metric until the end of the list	With every single metric a new boxplot on the left and two map plots on the right side - with respective metric name and data set pairs in their titles; For triple collocation metrics (TC: metricName for data setName) only one map plot is generated on the right side; For TC, Pearson's r p-value and Spearman's rho p-value there is no boxplot classification drop-down list.
4.	Click on the boxplot	A full-size plot opens as a gallery window, navigation arrows display to navigate between different boxplots produced for the chosen metric.
5.	Close the boxplot preview and click on one of the map plots	A full-size plot opens as a gallery window, navigation arrows display to navigate between plots.
6.	Close the map plot window and click the 'Download graphs button'	A dialog window for saving a graphs zip file opens; Default name to save is 'validationId_graphs.zip.
7.	Save the file and open it	The zip file contains all the files shown on the website in two formats .png and .svg;

		Boxplot's file name follows the pattern 'boxplot_metricName' for unclassified boxplots and boxplot_metricName_classificationShortName;
		Map plot's file name follows the pattern: 'overview_referenceName_and_data setName_metricName'.
8.	Click the 'Download NetCDF' button	A dialog window for saving the result netCDF file opens;
		Default file name to save is '0-ISMN.soil_moisture_with_1-C3S_combined.sm_with_2-GLDAS.SoilMoi0_10cm_inst.nc';
9.	Save the netCDF file and check it's size	The saved file should have size around 0.5 Mb;
Pass/Fail:	Pass	
Comments:	The file size might be different, as there have been a few more plots added depicting FRM classification and showing source of errors in the validation process. In step 3 for metric '#status' there is no boxplot assigned, but there is a bar plot.	

Test Case QA4SM_GUI_VRP_007 - Test Validation Results Page - Copying validation belonging to another user

Description:	Testing whether copying validation belonging to another user works as expected
Test Data:	None
Preconditions:	Logged in user

Requirement(s) to be tested:	Test Case QA4SM_GUI_VRP_001

Test Steps		**Expected Results**
1.	Go to the link https://test.qa4sm.eu/ui/validation-result/4625c806-35f8-40b0-85ee-c633492e1ec2	Validation results displayed in three separate components: • 'Summary: validation_for_testing_copying, followed by a question mark icon (informing that there is no possibility of changing name because the validation does not belong to the current user, when hovering over) - containing all the validation settings, used data sets and action buttons (Copy validation and Load settings), • 'Summary statistics' (collapsed by default) - containing mean, median and standard deviation calculated for all available metrics and a button for downloading a .csv file with the summary; • 'Result files' - containing two plots, two buttons for downloading results (graphs and netCDF file), two drop-down lists for choosing a metric to display (by default set to '# observations') and boxplot classification (by default set to 'unclassified').
2.	Click the 'Copy validation' button	Pop-up window shows up informing that the validation will be copied and added to 'my validations' list.
3.	Click 'Ok' on the pop-up window	Redirection to the result page of the copied validation
4.	Inspect new result page	Validation results displayed in three separate components:

| | | • 'Summary: Copy_of_validation_for_testing_copying followed by a pencil button - containing all the validation settings, used data sets and action buttons (Delete, Renew, Archive, Publish, Load settings)
• 'Summary statistics' (collapsed by default) - containing mean, median and standard deviation calculated for all available metrics and a button for downloading a .csv file with the summary;
• 'Result files' - containing two plots, two buttons for downloading results (graphs and netCDF file), two drop-down lists for choosing a metric to display (by default set to '# observations') and boxplot classification (by default set to 'unclassified'). |
|---|---|---|
| **Pass/Fail:** | Pass | |
| **Comments:** | | |

6.2.6 *Publishing validations and managing published results*

Test Case   QA4SM_GUI_PVMRP_001 - Test Validation Results Page - Test Validation Publishing

Description:	Testing whether publishing window renders properly and if publishing works as expected
Test Data:	None
Preconditions:	Go to the 'Validate' page and run a validation with the default settings and name 'test_validation_to_publish', wait until it finishes and open

	it. Go to the user's profile and fill the profile with data: First name: Sam, Last name: Smith, Organization: University, ORCID: 0000-0000-0000-0000, save data;
Requirement(s) to be tested:	Test Case QA4SM_GUI_VP_010, Test Case QA4SM_GUI_VRP_001

Test Steps		**Expected Results**
1.	Click on the 'Publish' button	A modal window for publishing opens with: • a short introduction saying 'Please check the metadata your results will be published with. If you want to permanently set your author details, you can do so on your profile page.' - with 'profile page' redirecting to user's profile; • a title - Validation of C3S SM combined v202012 vs ISMN 20210131 global • description - 'QA4SM validation of soil moisture data: C3S SM combined v202012 vs ISMN 20210131 global. URL: https://test.qa4sm.eu/ui/validation-result/{validationID}/. Produced on QA4SM (https://test.qa4sm.eu)' - where validationId is the id of the validation being published; • keywords: soil moisture, validation, qa4sm, C3S, ISMN; • name: Smith, Sam; • affiliation: University; • orcid: 0000-0000-0000-0000; • information that the publishing result can not be undone; • 'Cancel' and 'Publish now!' buttons, both enabled;

		Each of the fields is followed by a question mark icon than provides additional information when hovering over it.
2.	Remove the title	Publish button disabled, field's frame changes color to red.
3.	Press ctrl+z to retrieve the title and remove description	Publish button disabled, field's frame changes color to red.
4.	Press ctrl+z to retrieve the description and remove all the keywords	Publish button disabled, field's frame changes color to red.
5.	Press ctrl+z to retrieve the keywords, remove the name	Publish button disabled, field's frame changes color to red.
6.	Press ctrl+z to retrieve the name and remove affiliation and orcid	Nothing happens. Publish button still active.
7.	Press ctrl+z to retrieve the affiliation and ORCID, remove qa4sm from the keywords and press 'Publish now!' button	'Missing required keyword' warning displayed below the keywords field.
8.	Press ctrl+z to retrieve the removed key word and remove four last digits of the orcid number and press publish button	'Invalid ORCID identifier' warning displayed below the ORCID field.
9.	Press 'Cancel' button	Publishing windows closes
10.	Open the publish window again and press 'Publish now!' button	Publishing window closes and in place of action buttons there is an information about publishing being in progress.

After some time (usually a few seconds):
- page refreshes,
- two buttons show up - 'Load settings' and 'Pin validation',
- information about validation being published added along with its DOI |

		number and link to the ZENODO website,
		• pencil button in the header is replaced with a question mark icon saying (when hovering over) that changing name of a published validation is not possible.
11.	Go to 'My validations' page	Validation row with the published validation changed:
		• there is a book icon in the header saying that the validation has been published (when hovering over),
		• pencil button in the header is replaced with a question mark icon saying (when hovering over) that changing name of a published validation is not possible,
		• no 'Archive' button,
		• no 'Extend Life Span' and 'Delete Validation Run' options when expanding double-arrow drop-down list.
12.	Go to 'Published validations' page	The validation added to the published validations list.

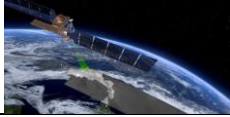| **Pass/Fail:**	Pass
Comments:	This test depends on an external service, so it may happen that it is impossible to conduct it if the ZENODO Sandbox (or in case of the main instance - ZENODO) is unavailable.
	The default ISMN version is now 20230110, and the full name C3S might be use which is C3S SM Combined.

Test Case QA4SM_GUI_PVMRP_002 - Test Published Validations Page - Render Published Validations Page

Description:	Testing whether the 'Published validations ' page renders properly
Test Data:	None
Preconditions:	A validation created and published during Validation Publishing test
Requirement(s) to be tested:	Test Case QA4SM_GUI_PVMRP_001

Test Steps		Expected Results

1.	Go to 'Published validations' page (test.qa4sm.eu/ui/published-validations)	Page renders with:
		
		 • a title - 'Published validations',
		 • short introduction containing links Zenodo (https://zenodo.org/) and DOI (https://www.doi.org/) services,
		 • sorting form, labeled 'Sort by' is rendered with two drop-down lists - one for the sorting feature (by default set to 'Date') and the second one for the order (by default set to 'descending'),
		 • validation results rows rendered for each validation that has been published; each row contains:
		   • a header with a name on the left side and DOI number (being a link to the results published on ZENODO) on the right side,
		   • a body containing information on data sets used for validations, creation date and action buttons;
		
		 Remark: If more than 10 validations have been run, pagination is added and displayed below validation rows;

Pass/Fail:		Pass

Comments:	Links to the results are prepared to be working with the main ZENODO instance and not with ZENODO Sandbox, therefore, they may not work properly for the validation published using the test instance of QA4SM.

Test Case QA4SM_GUI_PVMRP_003 - Test Published Validations Page - Inspect Validation Results Row

Description:	Testing whether Validation Results Row renders properly and if the action buttons works as they expected
Test Data:	None
Preconditions:	A validation created and published during Validation Publishing test
Requirement(s) to be tested:	Test Case QA4SM_GUI_PVMRP_001, Test Case QA4SM_GUI_PVMRP_002

Test Steps		**Expected Results**
1.	Go to 'Published validations' page (test.qa4sm.eu/ui/published-validations)	Validation row with the newest published validation rendered at the top; The row's header contains: • validation name, • DOI assigned during publishing, being a link to the results published on Zenodo (on the test instance the link might not work properly, as publishing here is done trough Zenodo Sandbox) The row's body contains: • Data: 'C3S, v202012, sm', ISMN, 20210131 global, soil_moisture (spatial, temporal reference)

		• Created : date of creation;
		• action buttons:
		    • 'View results' folder button,
		    • 'Load Validation Settings' round arrow button,
		    • 'Pin validation' plus button.

2.	Click on the folder button	Redirection to the page test.qa4sm.eu/ui/validation-result/validation_id (where validation_id is a random universal unique identifier);
3.	Go back to the published validations page and click the Load validation settings button	Redirection to the 'Validation' page' with settings of the current validation set;
4.	Click the 'Pin validation' button and choose 'ok' on the pop-up window	'Pin validation' plus button changes into 'Un-pin' X button
5.	Go to 'My validations' page	Pinned validations list added, with rows referring to each pinned validation and containing a validation name and four action buttons: • folder button for opening results, • download button for downloading graphs, • download button for downloading netCDF result file, • X button for unpinning a validation;
6.	Get back to the 'Published validations' page and click the 'Un-pin' button and choose 'ok' on the pop-up window	'Un-pin validation' X button changes into 'Pin' plus button, a pop-up window shows up informing that the validation has been removed from the list of pinned validations.
7.	Go to 'My validations' page	No 'Pinned validations' list.

Pass/Fail:	Pass
Comments:	

Test Case QA4SM_GUI_PVMRP_004 - Test Pinned Validations List

Description:	Testing whether the 'Pinned validations' list renders properly and if the action buttons behave as expected
Test Data:	None
Preconditions:	Go to 'Published validations' page (test.qa4sm.eu/ui/published-validations) and click the 'Pin validation' button.
Requirement(s) to be tested:	Test Case QA4SM_GUI_PVMRP_001, Test Case QA4SM_GUI_PVMRP_002

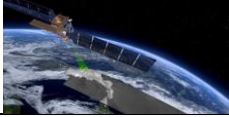|	Test Steps	Expected Results
1.	Go to the 'My validations' page and click the folder 'View results' button	Redirection to the page test.qa4sm.eu/ui/validation-result/validation_id (where validation_id is a random universal unique identifier);
2.	Go back to the 'My validations' page and click the 'Download graphs .zip' button	A dialog window for downloading a .zip file opened;
3.	Cancel downloading, click the 'Download NetCDF File' button	A dialog window for downloading a result file opened;
4.	Cancel downloading and click the 'Un-pin validation' button and confirm on the pop-up window.	The 'Pinned validations' list vanishes.
Pass/Fail:	Pass	
Comments:	Depending on the browser settings, the dialog window may not open, and the downloading process may start right away after clicking the proper download button.	

6.2.7 Validation comparison module

Test Case QA4SM_GUI_VCM_001 - Test Render Validation Comparison Module

Description:	Testing if comparison module renders properly
Test Data:	None
Preconditions:	Logged in user
Requirement(s) to be tested:	

	Test Steps	Expected Results
1.	Go to test.qa4sm.eu/ui/comparison	A Validation Comparison module rendered with: • Dataset configuration tab; • Validations selection tab; • Spatial extent tab; • 'Compare' button. All tabs have their title on the left side of the header and a question mark icon on the right side, which displays help text when hovering over. The Spatial extent tab has a collapse button.
2.	Inspect 'Dataset configuration' tab	The tab consists of two components: • Spatial Reference selection with 'Dataset', 'Version' and 'Variable' drop-down lists. By default fields are set to 'ISMN', '20210131 global' and 'soil_moisture' respectively. • Non-reference datasets section with a checkbox labeled 'Multiple non-reference datasets' and a question mark icon, providing help text when

		hovering over. The checkbox is unchekced by default.
3.	Inspect 'Validations selection' tab	The tab consist of one component labeled 'Validations available for comparison', containing a drop-down list and an 'Add validation' button.
		If there are no validations available with the chosen spatial reference dataset, then the 'Add validation' button is disabled.
4.	Inspect 'Spatial extent' tab	The tab contains a disabled check box labeled 'Include all points from the selected validations' and a question mark icon providing help text when hovering over.

Pass/Fail:	Pass
Comments:	

Test Case QA4SM_GUI_VCM_002 - Test Validation Comparison Module

Description:	Testing if comparison module behaves as expected
Test Data:	None
Preconditions:	Logged in user,
	Existing at least two validations with the same spatial reference dataset, but with common spatial subsetting and one validation with two non-reference datasets.
Requirement(s) to be tested:	QA4SM_GUI_VCM_001

Test Steps		Expected Results
1.	Render the 'Compare validations' module and press 'Compare' button	A warning pops up informing that there were no validations chosen for comparison.

2.	In the 'Dataset configuration' section choose the dataset and its version which you used as the spatial reference in at least two validations	The list of validations available for comparison updates.
3.	In the 'Validations selection' section select two validations with common spatial range you want to compare from the drop-down list, every time confirm your choice with the 'Add validation' button	Two red buttons added below the drop-down list. Each button contains a trash bin icon and a label: 'Remove: Validation date: …, Non-reference-dataset: …'; 'Add validation' button disabled; The checkbox in the 'Spatial extent' section enabled.
3.	Remove one of the added validations	'Add validation' button enabled again.
4.	Retrieve the removed validation, check the 'Spatial extent' checkbox and click the 'Compare' button and wait until results are generated	'Validation comparison results' section shows up, consisting of: • Comparison summary tab, • Selected comparison extent tab, • Comparison statistics tab, • Comparison plots tab. Comparison plots tab renders unfolded. Other tabs render collapsed. All tabs have: • their titles in the middle of their headers, • a question mark icon that provides help text when hovering over, • collapse buttons.
5.	Remove one validation from the list	'Validation comparison results' section vanishes. The 'Spatial extent' checkbox gets unchecked and disabled.
6.	Check the 'Multiple non-reference datasets'	All the chosen validations vanish and the list of validations available for comparison updates.

| | | If there is no validation with the chosen spatial reference dataset that contains two non-reference datasets, the list is empty. Go to point 7.

If there exists, for the chosen spatial reference dataset, a validation with two non-reference datasets, go to point 8. |
|---|---|---|
| 7. | Choose the spatial reference dataset for which there exists a validation with two non-reference datasets | The list of validations available for comparison updates. |
| 8. | Choose on validation from the list and confirm it with the 'Add validation' button. | A red button added below the drop-down list. 'Add validation' button disabled; |
| 9. | Click the 'Compare' button | 'Validation comparison results' section shows up, consisting of the same sections as in point 4. |
| 10. | Remove the chosen validation | The 'Validation comparison results' section vanishes. |

Pass/Fail:	Pass
Comments:	

Test Case QA4SM_GUI_VCM_003 - Test Validation Comparison Results Section

Description:	Testing if comparison module results part renders and behaves as expected
Test Data:	None
Preconditions:	Logged in user,

Existing at least two validations with the same spatial reference dataset, but with common spatial subsetting (not necessarily the same). |

Requirement(s) to be tested:	QA4SM_GUI_VCM_001, QA4SM_GUI_VCM_002

	Test Steps	Expected Results
1.	Choose the spatial reference dataset, for which there exist at least two validations with one non-reference data set, from the 'Spatial reference selection' drop down window; Choose two validations for comparison from the 'Validations available for comparison' list; Click the 'Compare' button	'Validation comparison results' section shows up, consisting of: • Comparison summary tab, • Selected comparison extent tab, • Comparison statistics tab, • Comparison plots tab.
2.	Unfold and inspect 'Comparison summary' tab	The tab contains two columns with summarized information about each validation used for comparison. The summary consists of: • validation date and time (start and finish), • list of compared data sets, • spatial filter bounding box, • validation period / temporal filter, • information about validation metrics, • information about triple collocation analysis, • information about bootstrapping of confidence intervals for Triple Collocation Analysis, • information about scaling reference, • information about Scaling method.
3.	Unfold and inspect 'Selected comparison extent' tab	The tab contains: • information that all the points were taken into account during the comparison process,

		• a plot showing spatial extent of the comparison with bounding box referring to each validation (if chosen validations cover exactly the same region, there will be only one bounding box visible), • a button labeled 'Download image'.
4.	Click the spatial extent plot	A full-size plot opens as a gallery window.
5.	Close the gallery and click the 'Download image' button	A dialog window for saving the picture opens, with default name 'spatial_extent' and .png format.
6.	Close the dialog window and unfold and inspect 'Comparison statistics' tab	The tab contains • a table with four columns containing: metric name, median values for each metric for each validation, differences between median values coming from two validations; • a button labeled 'Download .csv table'.
7.	Click the 'Download .csv table' button	A dialog window for saving the table in .csv format opens, with default name 'Comparison_summary'.
8.	Save the file and open it	The file contains exactly the same table as the one displayed on the page.
9.	Inspect 'Comparison plots' tab	The tab contains: • a boxplot depicting a chosen metric for each validation and the difference between them, • a map of the common spatial subsetting, depicting difference of the chosen metric, • two buttons labeled 'Download image' located under each plot, • a drop-down list with available metrics.
10.	Click on the boxplot	A full-size plot opens as a gallery window

11.	Close the boxplot preview and click on the map	A full-size plot opens as a gallery window
12.	Close the map plot window and click any of the 'Download image' button, close the dialog window and click the second one	Every time a dialog window for saving a plot opens, with a default name 'plot0_metric_name' for the boxplot and 'plot1_metric_name' for the map.
13.	Choose another metric from the metric drop-down list	Plots update according to the chosen metric.

Pass/Fail:	Pass
Comments:	

6.2.8 User datasets uploading form and list of uploaded files

Test Case QA4SM_GUI_UDUF_001 - Test Uploading User Data - Uploading File

Description:	Testing if the file is uploaded along with its metadata.
Test Data:	Test file
Preconditions:	Logged in user, Existing netCDF file prepared according to our guidelines
Requirement(s) to be tested:	
Test Steps		**Expected Results**
1.	Go to 'My datasets' page https://test.qa4sm.eu/ui/my-datasets	The page contains: • A title: 'My datasets'; • Uploading section with two buttons 'Select file' and 'Upload file' and information that no file has been chosen yet;

		• List of uploaded files or information that no files have been uploaded yet, if that is the case.
2.	Click the 'Select file' button	'Upload your own data' window opens, with a link to the file standard guidelines on the help page and a button labeled 'Add file'.
3.	Click the 'Add file' button	Window for choosing a file opens.

4.	Choose a proper netCDF file to upload and close the window	• File name displayed;
		• 'Add file' button label changed into 'Change file';
		• Metadata form displayed, with four fields labeled:
		  o Dataset name,
		  o Dataset display name (optional),
		  o Version name,
		  o Version display name (optional),
		
		and question mark icons on the right side of each field, showing guidelines for each field, when hovering over them;
		
		• 'Save' button displayed, but deactivated

5.	Provide: • 'test_dataset' as dataset name, • 'Test dataset' as dataset display name, • 'test_version' as version name, • 'Test version' as version display name	'Save button' gets activated.
6.	Click the 'Save button'	Uploading file window closes.
7.	Click the 'Select file' button again	Uploading file window opens, filled with the data provided in the previous steps.

8.	Close the window and click the 'Upload file' button	Page gets blocked, a spinner shows up, information about uploading file and its progress is displayed.
9.	Wait until the file is uploaded and the list of uploaded file gets refreshed.	A new tab with information on the uploaded file added.
Pass/Fail:	Pass	
Comments:		

Test Case QA4SM_GUI_UDUF_002 Test Uploading User Data - Uploaded File Panel

Description:	Verifying if a panel showing information of an uploaded file displays properly and if all the anticipated actions can be performed on the file
Test Data:	Test file
Preconditions:	Logged in user, Uploaded netCDF file
Requirement(s) to be tested:	QA4SM_GUI_UDUF_001

Test Steps		**Expected Results**
1.	Go to 'My datasets' page https://test.qa4sm.eu/ui/my-datasets and inspect the list of uploaded files	Above the first panel there is folded information about user file management, with 'Read more' link. There are as many panels as many files have been uploaded so far.
2.	Click 'Read more' link	Full information on user file management displayed, with 'Read less' link displayed.
3.	Click 'Read less' link	Information folded again.
4.	Inspect user data panel	The user data panel contains:

		• A header with the file size on the left side and the upload date on the right side;
		
		• A body with information on:
		   ○ dataset – its name, version (from the metadata form) and variable (as derived from the file),
		   ○ coordinates names derived from the file,
		   ○ action buttons.

5.	Hover over dataset name, version name, variable and coordinates names	Cursor appearance changes and tooltips show up explaining possible name changes.
6.	Click on the dataset name	Text field opens with the current name in it, save and cancel buttons show up.
7.	Click the cancel (ban icon) button	Text field closes, name does not change, buttons vanish.
8.	Click on the dataset name one more time, change the name to 'My dataset' and click the save (floppy disc) button	Text field closes, name changes, buttons vanish, a green notification shows up in the upper right corner saying that the metadata has been updated.
9.	Click on the version name	Text field opens with the current name in it, save and cancel buttons show up.
10.	Click the cancel (ban icon) button	Text field closes, name does not change, buttons vanish.
11.	Click on the version name one more time, change the name to 'My version' and click the save (floppy disc) button	Text field closes, name changes, buttons vanish, a green notification shows up in the upper right corner saying that the metadata has been updated.
12.	Click on the variable, latitude, longitude and time names	Dropdown lists and cancel buttons show up.
13.	Unfold and inspect each dropdown list (do not click on any names)	All dropdown lists contain the same list of variable names, those are variable names retrieved from the uploaded file.

14.	Click the cancel buttons next to the latitude, longitude and time dropdown lists	Latitude, longitude and time dropdown lists and cancel buttons vanish
15.	Unfold the variable dropdown list and choose a different name	Name changes, the list and cancel button vanish, a green notification shows up in the upper right corner saying that the metadata has been updated.
16.	Click the 'Remove dataset' button	A pop-up window shows up requesting about removal confirmation
17.	Click yes on the pop-up window	The dataset panel is removed from the list.

Pass/Fail:	Pass
Comments:	There has been change in the row layout. As it has been decided not to present dimension names on the data row, points 5, 12, 13 and 14 should refer only to the variable name drop-down list.

Test Case QA4SM_GUI_UDUF_003 - Test Uploading User Data - Running A Validation With User Data

Description:	Verifying if the validation can be run with the uploaded data
Test Data:	Test file
Preconditions:	Logged in user
Requirement(s) to be tested:	QA4SM_GUI_UDUF_001, QA4SM_GUI_VRP_001, QA4SM_GUI_MVP_002

Test Steps		Expected Results
1.	Go to 'My datasets' page https://test.qa4sm.eu/ui/my-datasets and upload a netCDF file, providing: • 'test_dataset' as dataset name,	A new tab with information on the uploaded file added.

	• 'Test dataset' as dataset display name, • 'test_version' as version name, • 'Test version' as version display name	
2.	Verify if the names of variable and coordinates are the one used in the file	If the file indicated as the test data in this test case is used, the names should be: • soil_moisture (soil_moisture) for Variable, • lat for Latitude, • lon for Longitude, • Time for Time
3.	Go to 'validate' page https://test.qa4sm.eu/ui/validate, unfold dataset dropdown list and scroll to the top of the list	'Test dataset' name is on the top of the list
4.	Choose 'Test dataset' from the dropdown list	'Test version' set as version and 'soil_moisture' set as variable.
5.	Leave other settings as they are and start a validation by clicking the 'Validate' button. Go to 'My validations' page https://test.qa4sm.eu/ui/my-validations, wait until the validation finishes and open it.	Validation results displayed, as described in the test case QA4SM_GUI_VRP_001.
6.	Go to 'My datasets' page and find the 'Test dataset' panel	The 'Remove dataset' button belonging to the dataset is deactivated.
7.	Go to 'My validations' page, remove the validation with the uploaded dataset and go back to 'My datasets' page (Removing validation is described in e.g. test case QA4SM_GUI_MVP_002)	The 'Remove dataset' button belonging to the dataset is active again.

8.	Remove the dataset, go to the 'Validate' page and unfold the dropdown list with datasets	'Test dataset' is no longer available on the list.
Pass/Fail:	Pass	
Comments:	There has been change in the row layout. As it has been decided not to present dimension names on the data row, point 2 should refer only to the variable name drop-down list.	

6.3 Summary and evaluation of test results

The test plan and the individual test cases were developed to comprehensively cover all functions and user interactions of the QA4SM service. The test cases enable systematic verification of the correct behaviour of each service component and can easily be repeated in the development process to ensure service stability and re-test service components when they are changed.

All test cases have systematically been performed and evaluated twice by two independent testers. Due to platform updates after development of the test cases, a number of small inconsistencies were observed between the current version of the software and the test cases. For example, the default version of the ISMN dataset has been updated from '20210131 global' to '20230110 global', which was not yet reflected in the test case description. These observed inconsistencies, are explained in the comment section of the affected test cases. A few remaining software issues have been identified in a number of tests. All such issues have been fixed and successfully re-tested before completion of this report. The tested QA4SM release 2 version is thus technically ready for acceptance and deployment in the public service.

7 Verification of outputs and results

Having provided a first verification of the QA4SM service in the first version of the QA4SM Evolution Verification Report (DT4-3, Version 1.1), the verification activities performed for Release 2 of QA4SM focused on the main back-end[1] enhancements and improvements that have been implemented on top of release 1. These consist of the following:

- Integration of SMOS Level 2 (hereafter SMOSL2) in the service
- Integration of SMAP Level 2 (hereafter SMAPL2) in the service
- Handling and communication of (known) processing errors
- Integration of Fiducial Reference Measurements flag (hereafter FRM-flag) from ISMN in QA4SM

7.1 Verification of SMOS Level 2 integration

7.1.1 Test Cases

7.1.1.1 Test SMOSL2-1: comparison with independent validation run

For these tests, the ISMN data set used by CESBIO was shared with the QA4SM team and integrated in the platform, to ensure that differences in ISMN versions or reprocessing interfere with the verification output.

Comparison through User Upload

For this exercise, a subset of the SMOS L2 data set was extracted by CESBIO and shared with the QA4SM team, corresponding to the DGGs listed in Table 1. Such points are matched to the closest ISMN station (and sensor depth), which is also provided in the table. The reference scores have been computed by CESBIO and are relative to the SMOS DGGs indicated in the table. The QA4SM validation was run using the settings detailed in: https://test.qa4sm.eu/ui/validation-result/f188955b-f1bf-4bda-afe6-246757e94ae1.

Table 1: overview of the ISMN points used in the comparison exercise

Network	Station	SMOS DGG	Lat	Lon	Sensor type	Sensor depth
USCRN	Wolf_Point_34_NE	160036	48.536	-105.271	Stevens-Hydraprobe-II-Sdi-12	0.05
USCRN	Wolf_Point_34_NE	160036	48.536	-105.271	Stevens-Hydraprobe-II-Sdi-12	0.1
USCRN	Stillwater_2_W	222578	36.129	-97.133	Stevens-Hydraprobe-II-Sdi-12	0.05
USCRN	Stillwater_2_W	222578	36.129	-97.133	Stevens-Hydraprobe-II-Sdi-12	0.1
USCRN	Stillwater_5_WNW	222578	36.129	-97.133	Stevens-Hydraprobe-II-Sdi-12	0.05
USCRN	Stillwater_5_WNW	222578	36.129	-97.133	Stevens-Hydraprobe-II-Sdi-12	0.1
USCRN	Oakley_19_SSW	205135	38.887	-101.043	Stevens-Hydraprobe-II-Sdi-12	0.05
USCRN	Oakley_19_SSW	205135	38.887	-101.043	Stevens-Hydraprobe-II-Sdi-12	0.1
USCRN	Northgate_5_ESE	162101	49.028	-102.208	Stevens-Hydraprobe-II-Sdi-12	0.05
USCRN	Manhattan_6_SSW	210797	39.036	-96.627	Stevens-Hydraprobe-II-Sdi-12	0.05
USCRN	Manhattan_6_SSW	210797	39.036	-96.627	Stevens-Hydraprobe-II-Sdi-12	0.1
USCRN	Jamestown_38_WSW	174921	46.734	-99.507	Stevens-Hydraprobe-II-Sdi-12	0.05
USCRN	Jamestown_38_WSW	174921	46.734	-99.507	Stevens-Hydraprobe-II-Sdi-12	0.1
USCRN	Goodridge_12_NNW	172379	48.338	-95.781	Stevens-Hydraprobe-II-Sdi-12	0.05
USCRN	Goodridge_12_NNW	172379	48.338	-95.781	Stevens-Hydraprobe-II-Sdi-12	0.1
USCRN	Gadsden_19_N	244166	34.342	-85.972	Stevens-Hydraprobe-II-Sdi-12	0.05
USCRN	Gadsden_19_N	244166	34.342	-85.972	Stevens-Hydraprobe-II-Sdi-12	0.1

---

[1] i.e., related to the validation algorithm or the input data, and not to the GUI

USCRN	Bronte_11_NNE	232283	32.02	-100.21	Stevens-Hydraprobe-II-Sdi-12	0.05
USCRN	Bronte_11_NNE	232283	32.02	-100.21	Stevens-Hydraprobe-II-Sdi-12	0.1
USCRN	Avondale_2_N	222699	39.916	-75.797	Stevens-Hydraprobe-II-Sdi-12	0.05
USCRN	Avondale_2_N	222699	39.916	-75.797	Stevens-Hydraprobe-II-Sdi-12	0.1
USCRN	Austin_33_NW	240491	30.673	-98.109	Stevens-Hydraprobe-II-Sdi-12	0.05
USCRN	Austin_33_NW	240491	30.673	-98.109	Stevens-Hydraprobe-II-Sdi-12	0.1
USCRN	Aberdeen_35_WNW	179533	45.719	-99.181	Stevens-Hydraprobe-II-Sdi-12	0.05
USCRN	Aberdeen_35_WNW	179533	45.719	-99.181	Stevens-Hydraprobe-II-Sdi-12	0.1

Note that only the pre-filtered SMOS L2 time series and the respective coordinates were shared, therefore the spatial matching was also tested in the process. Figure 1 shows the location of the SMOS L2 DGGs shared by CESBIO (blue) and the matched ISMN points (orange). While we have a perfect match of the ISMN stations, it can be noticed how in QA4SM each *spatial reference* validation point is matched with the uploaded SMOS L2 data set, leading to validation of points that can be illogically far apart. Note that this issue is only occurring for the case where a non-gridded data set (like the uploaded SMOS L2, consisting only of individual time series) is validated against ISMN (which must be selected as spatial reference). Before QA4SM Release 2, this aspect will be corrected by setting a distance threshold equal to the resolution of the spatial reference data set (and a default of 30km for ISMN).

Figure 1: Overview of the ISMN validation points used for the comparison. Blue and orange circles indicate the CESBIO and QA4SM points considered, respectively, while green crosses over validation points indicate that (at any depth) the scores could not be calculated in QA4SM. Note that only points where CESBIO scores could be calculated have been provided, and are plotted in the blue circles.

The SMOS L2 data extracted correspond to the June-August 2016 period. For reference, the time series received for the DGG *179533* is shown in Figure 2, together with the reference

ISMN data at 0.05 m depth and the complete SMOS L2 data for the time series, from the integrated QA4SM data set.
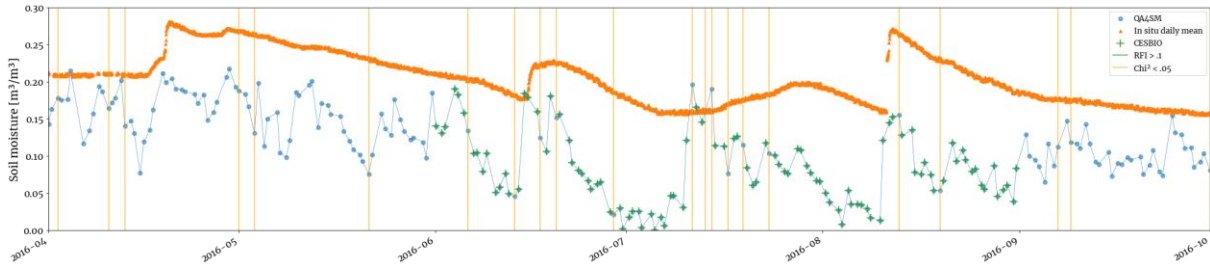
Figure 2: Time series showing the shared SMOSL2 data set ('CESBIO'), the SMOSL2 integrated data set ('QA4SM') and respective flags for RFI and CHI² probability, and the daily averaged ISMN data from the 0.05 m Aberdeen-35-WNW sensor.

Figure 4 shows the comparison between the reference and the QA4SM scores with respect to the 1:1 line. The agreement between the two sets of scores is perfect, with all points centered on the 1:1 line, which also lies well within the confidence intervals of the scores. No difference in the number of samples of the validation time series is evidenced for any of the validation points.
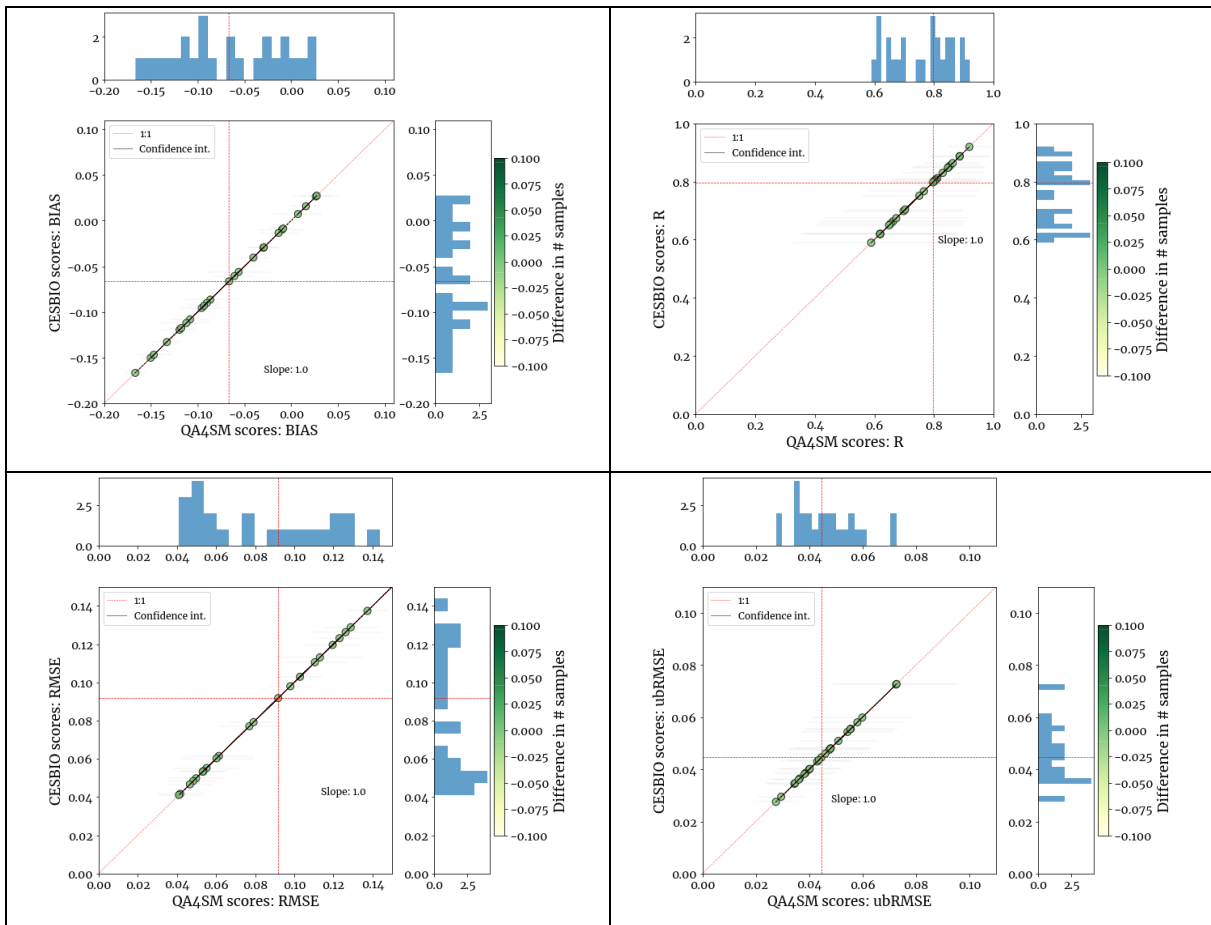
Figure 3: Comparison of BIAS, Root Mean Squared Error, unbiased-RMSE and correlation between the QA4SM-run validation (through the user upload feature) and the reference scores. The color bar indicates the difference in the number of samples in the two time series.

Figure 4 shows the distribution of the absolute and relative differences between the two scores sets. In general, the differences are several orders of magnitude smaller than the actual scores and their confidence intervals, with median relative differences always lower than 0.01 %, and overall differences never exceeding 1%.
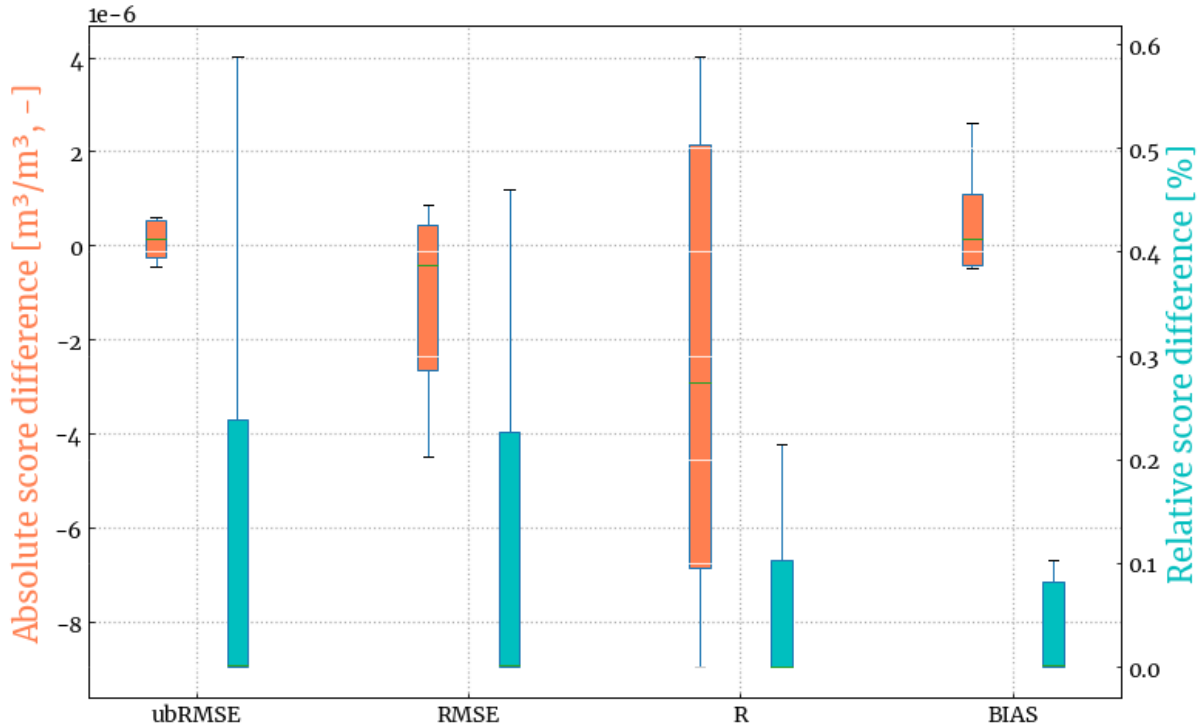
Figure 4: Comparison of the scores between the QA4SM-run validation (through the user upload feature) and the reference scores through absolute difference (left vertical axis) and difference relative to the reference scores (right vertical axis).

The test has demonstrated that there is no substantial difference between the methodology of QA4SM and of the independent validation, including all the validation (temporal/spatial matching, scores computation) steps that occur after the filtering of the data set. Eventual differences might remain in the selection (spatial/temporal sub-setting) and masking of the data, and are therefore assessed in the following exercise.

Comparison through integrated data set

For this exercise, the full 2010.01.01 (technically, SMOS L2 begins on 2010.06.01) - 2020.12.31 temporal sub-set of the SMOS L2 data set was validated against ISMN by CESBIO and with the QA4SM platform (https://test.qa4sm.eu/ui/validation-result/ccfe5968-a656-4a29-8982-6dae5b3914c5). The full extent of the validation can be appreciated in Figure 5. The global ISMN validation points in the 0.0-0.10 m depth from all networks are 3102; of these, around 41% led to validation errors (e.g., due to lack of temporal overlap), leaving 1780 validation

points for which the scores could be cross-compared between the reference and the QA4SM validation. Note that the reference scores include 118 more points with valid results, where however QA4SM did not yield an output, due to the stricter threshold on the number of observations in QA4SM (n=10) than in the reference validation (n=3).
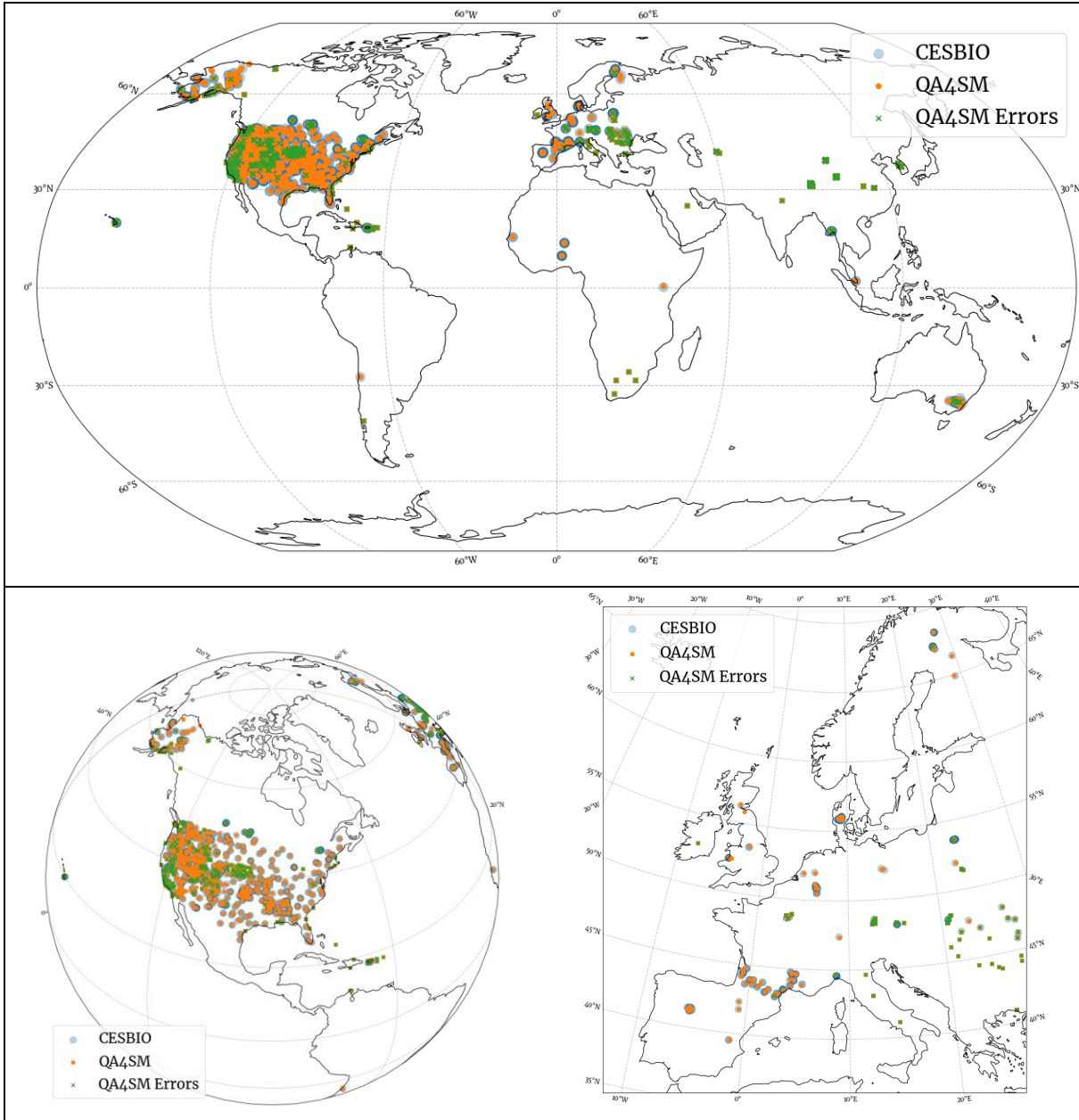

Figure 5: As for Figure 1, but for the integrated data set verification.

Figure 6: As for Figure 3, but for the integrated data set verification.Figure 6 shows the results of the cross-comparison for the correlation and difference metrics. Similar to the user upload test, the results show a perfect 1:1 agreement well within the confidence intervals of the individual data points. A slight difference in the number of observations still persists: 423 of 1780 points have a difference in the number of observations, with 273 having a difference of

1 point, and the rest having a difference between 1 and 20 (not shown). Differences that were initially found in the cross-comparison were explained by:

- An error in the temporal window implementation in QA4SM, which affected the number of samples but not the output scores (corrected).
- The different use of ">" (reference validation) or ">=" (QA4SM validation) operator in the value range filtering of the in situ reference for 0.0 m³/m³ values. Many malfunctions in the probes are not captured by the ISMN "G" flag, which leads to 0.0 m³/m³ values in the reference data. The lower range threshold in QA4SM was moved to 0.001 m³/m³ to compensate for this.

The persisting difference is possibly explained by the different time stamp implementation in the reference and QA4SM validation codes. Overall, it is considered minimal (a maximum of 2 observations per validation year, on average) and has a negligible impact on the validation scores.
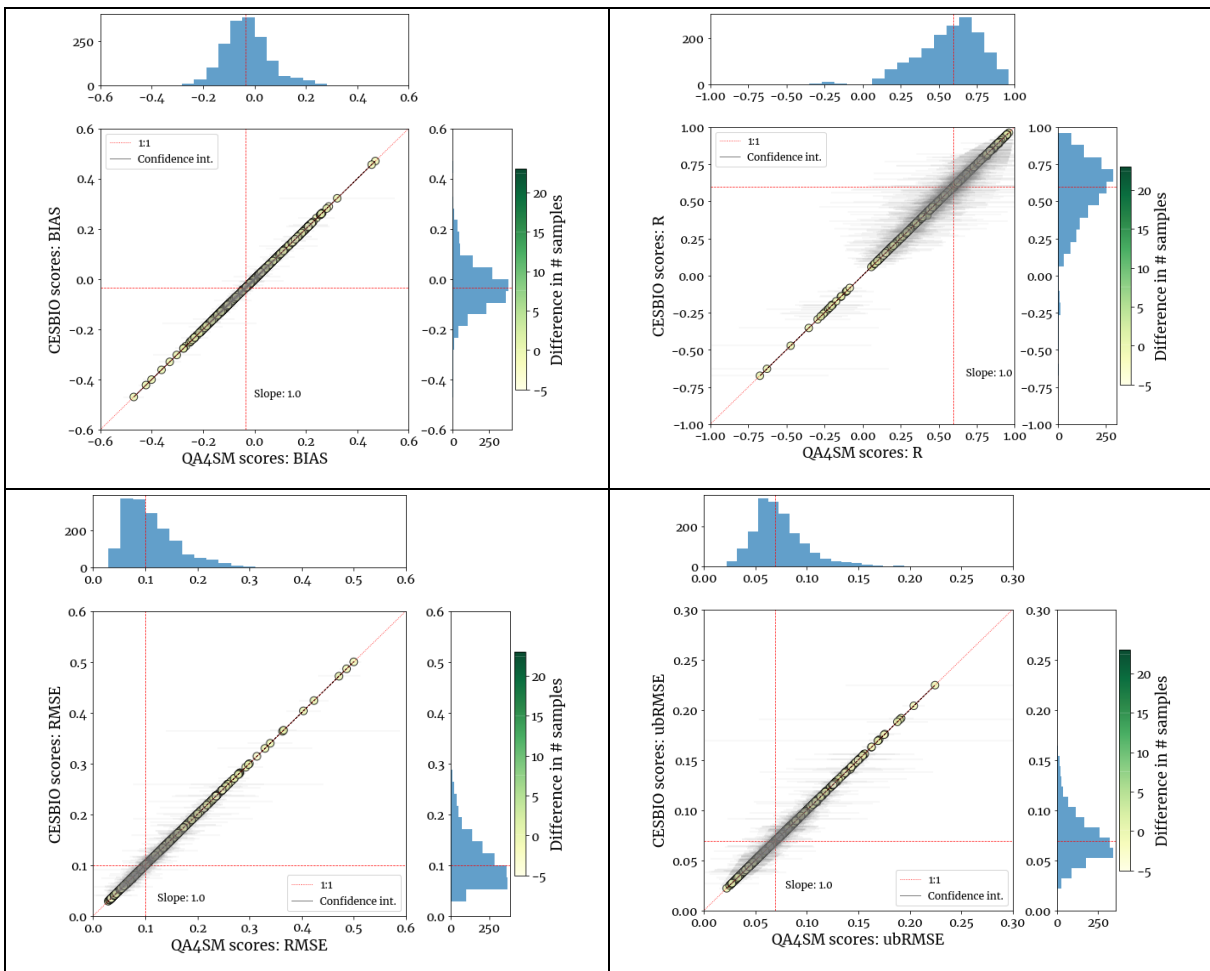
Figure 6: As for Figure 3, but for the integrated data set verification.

Figure 7 shows that the median deviation between the metrics is in all cases very close to 0.0 (m³/m³, -), with inter-quartile ranges always below 0.5 * 10⁻⁵. The median relative difference

is always below 0.002 %, and is several orders of magnitude below the relative confidence intervals on the scores.
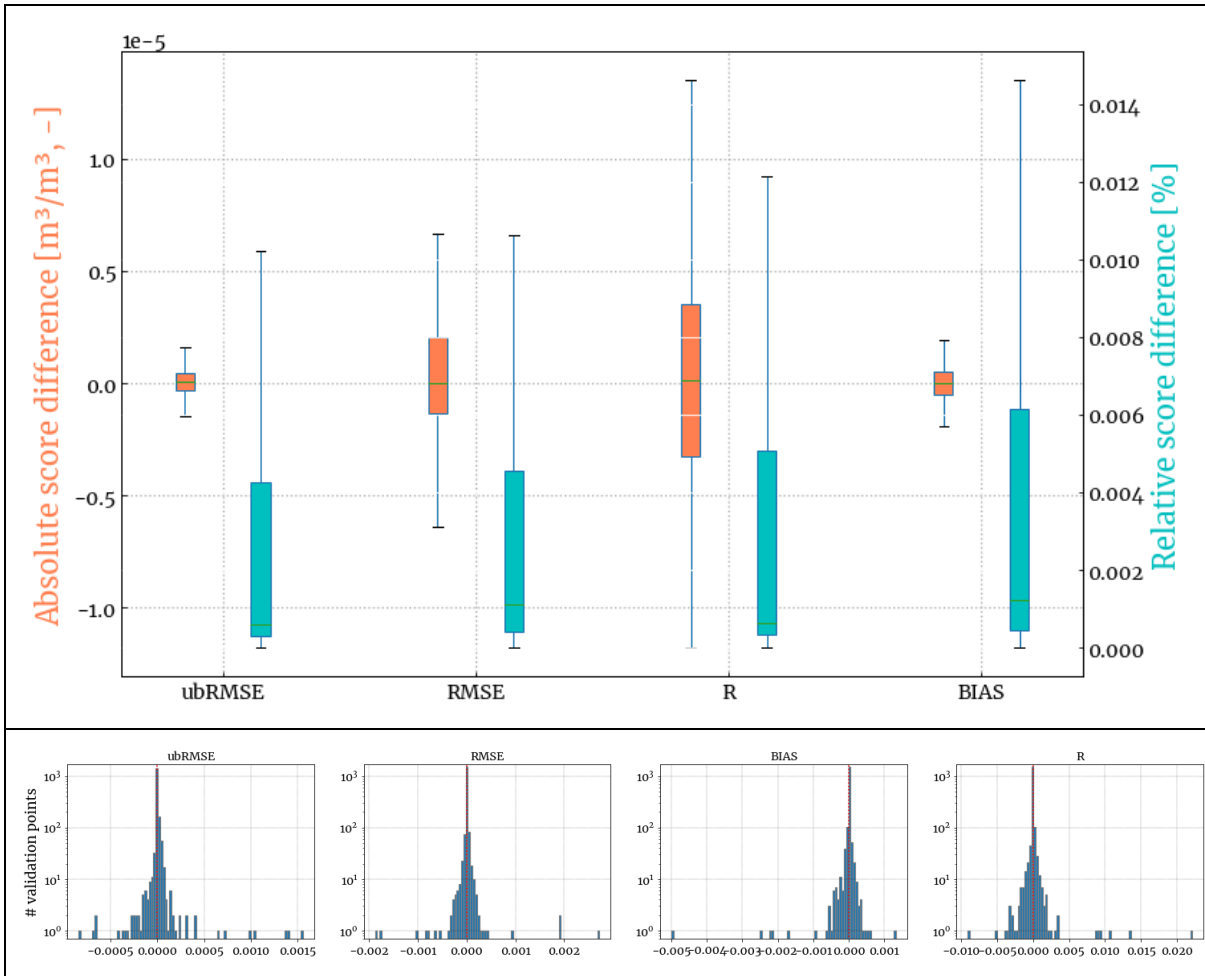
Figure 7: (above) as for Figure 4, but for the integrated data set verification (below) distribution of the scores' differences by metric with logarithmic vertical scale.

Overall, the test has allowed to assess the whole validation processing chain, from data set pre-processing, to reading, filtering, spatial- and temporal-matching and validating. The differences originally found where traced to differences in the validation codes that were homogenized during the exercise. The remaining differences are in practical terms negligible in terms of impact on the validation scores. Therefore, the test has resulted in an increased confidence in the QA4SM platform functionality.

7.1.1.2 Test SMOSL2-2.1: test impact of flags

Note: For this test, with respect to the Test Plan document, the "baseline" scenario was modified to include a 0.05 χ^2 probability threshold, corresponding to the default QA4SM settings. The "low Retrieval Confidence (RC)" scenario was modified to apply a 0.1 χ^2 probability threshold.

For this test, the flagging options provided with QA4SM were assessed by running a validation against ERA 5 over Europe and alternating the settings as detailed in the Test Plan document. Figure 8 shows the impact of the various configurations on the collocated validation sample. Noticeably, the "baseline" scenario (with frozen ground, moderate topography, 0.05 RC probability and low RFI filtering applied) shows the highest results coverage and data density. Gaps in the validation output, as for instance over part of the Balkan peninsula, Italy and Scandinavia, are explained by filtering of moderate topography and low RFI probability. Low-, medium- and high-RC filtering make a notable impact on the available points for validations, and in all cases more than using a high RFI filtering option. In absolute terms, the median difference in sample size from the medium RC, high RC and high RFI filtering are similar and close to on average 2 observations per year (observation difference of 20), while in the high RC scenario, few areas were reduced by as much as on average 80 observations per year or more (difference of > 800). Looking at the coverage maps, it is clear that while RFI determines a localised masking, in correspondence of the RFI sources, the RC threshold has a more homogeneous impact in space.
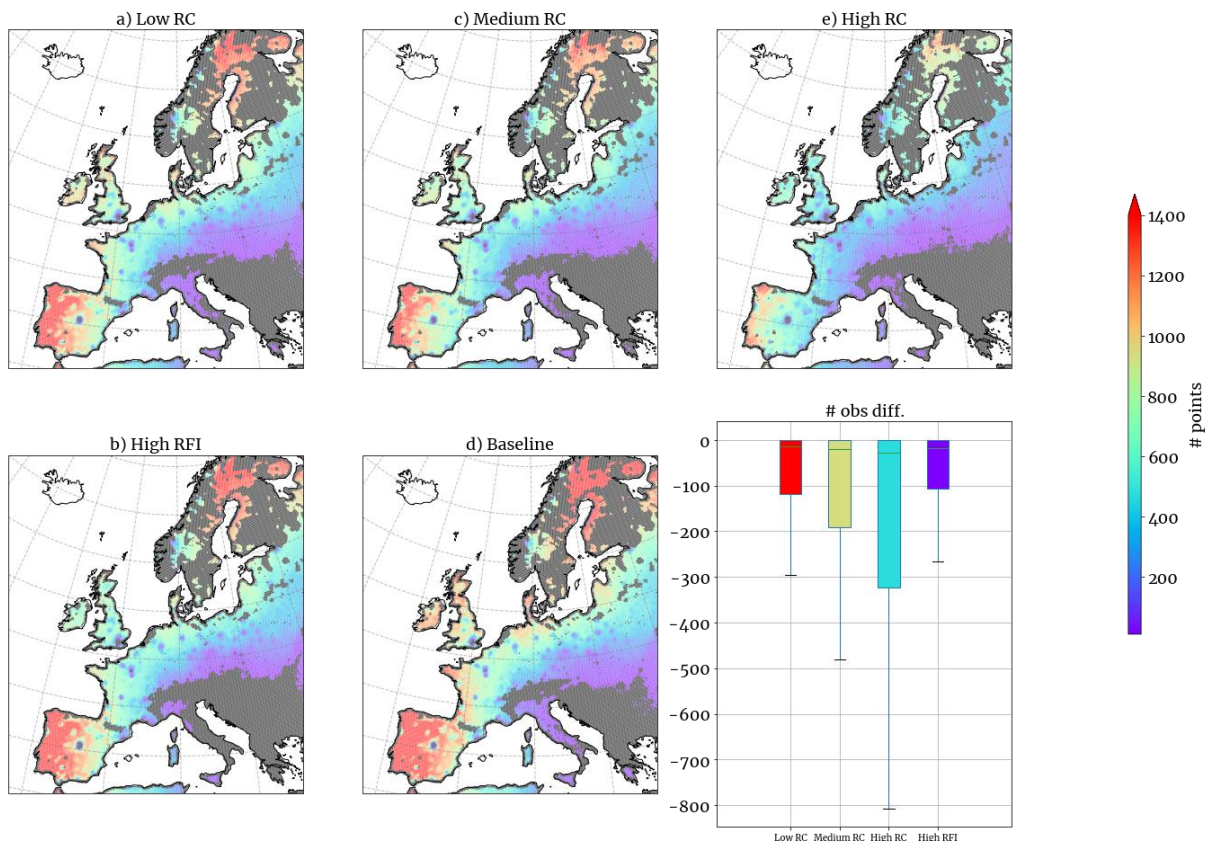
Figure 8: Impact of the various flagging configurations on the number of observations in the validation time series. Grey areas mean that the validation scores could not be computed.

Figure 9 shows how the filtering scenarios play out in terms of validation scores. All flagging scenarios do not provide a univocal impact on the validation scores, but instead introduce degradations together with score improvements. The low Retrieval Confidence (RC) and high

Radio Frequency Interference (RFI) scenarios provide a similar impact on the scores, while the medium and high RC filtering scenarios generate a median increase in correlation and a decrease in unbiased-RMSE that is proportional to the level of filtering applied. Interestingly, the high RC filtering is denoted by a larger bias than the low and medium RC cases, and the RFI filtering introduces the highest median bias. The high RFI filtering pertains an identical median unbiased RMSE to the baseline, while all the RC scenarios are characterised by a smaller random difference compared to the baseline. Independently of the validation reference used, such result indicates that there is a non-homogeneous impact of the flagging on the SMOS L2 data set which can be further investigated on different scales or by location.
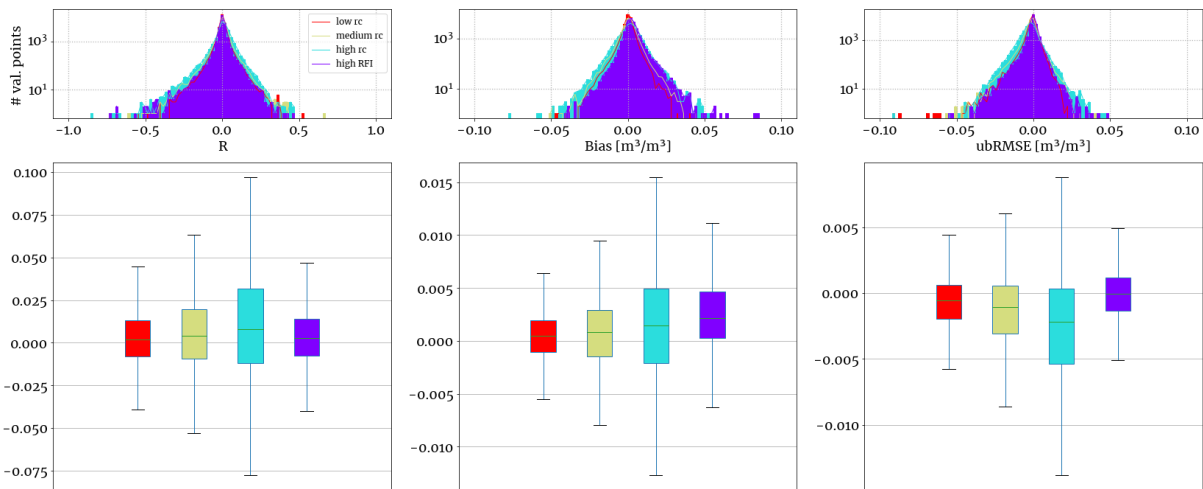
Figure 9: Comparison of the scores distribution as box plots (below) and histograms (above) for the various flagging configurations, in terms of difference from the "baseline" scenario.

Overall, the SMOSL2-2.1 test has produced an expected behavior in terms of sample size, with the validation sample being reduced at each of the filtering scenarios. The spatial patterns in the sample size are also consistent with the sources of degradation that the filtering is applied to. The impact on the validation scores is not only positive, but the filtering options may introduce degradations both in the correlation as well as distance (bias, unbiased RMSE) scores.

7.1.1.3 Test SMOSL2-2.2: comparison with SMOS L3

For this test, the SMOS L2 and L3 data sets were compared against ERA 5 over the US. Since the two data sets were downloaded from different repositories (respectively, earth.esa.int/eogateway/missions/smos/data and www.catds.fr/Products/Available-products-from-CPDC) and processed independently, it is expected that any issue in the data originating from e.g. missing data in the download or malfunctioning in the processing infrastructure would result in obvious anomalies in the comparison of the two. Overall, the test evidenced no such issues (Figure 10, validation reference: https://test.qa4sm.eu/ui/validation-result/92918d5f-2f5c-4bbf-be44-16554392ef6d). In general, SMOS L2 was found to have a higher agreement with the reference than the

simplified L3 product, both in terms of correlation and unbiased RMSD. Such result could be explained by the fact that SMOS L2 was chosen as temporal reference, which might have slightly penalized the L3 observations, because of the temporal distance to the ERA 5 observations, which are provided in QA4SM at 6-hourly intervals (0AM, 6AM, 12PM, 18PM). The scatter plots on the right side of Figure 10 are color coded to reflect the average size of the confidence intervals at each point. Interestingly, many of the outliers showing a large performance difference between SMOS L2 and L3 are also characterized by high uncertainty in the scores, and many of the points close to the 1:1 line present a lower uncertainty in the scores. This implies that the sample sizes might affect the robustness of this comparison.
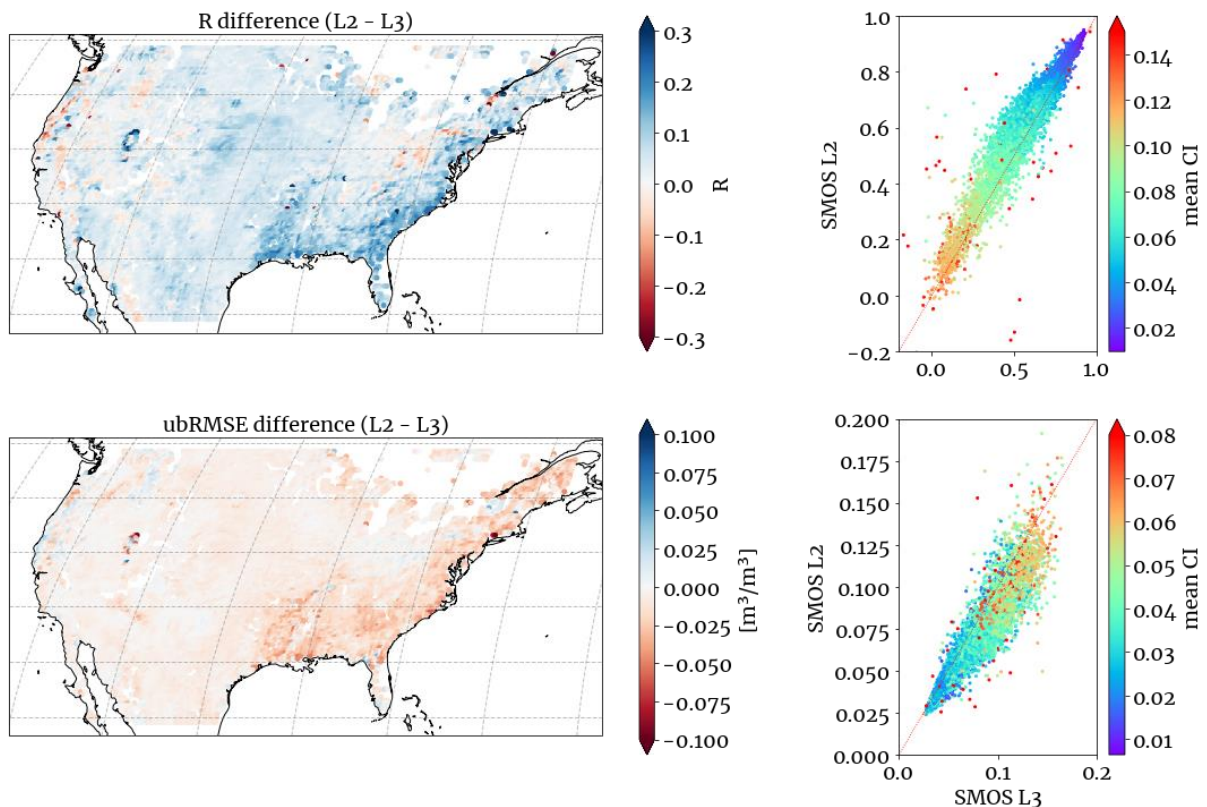
Figure 10: Difference of correlation (above) and unbiased-RMSE (below) between the SMOS Level 2 and Level 3 data sets integrated in QA4SM. The color range of the points in the scatter plot indicate the mean confidence interval characterizing the 2 scores.

7.2 Verification of SMAP L2 data integration

7.2.1 Test SMAPL2-1

7.2.1.1 Comparison with Chen et al., 2018

Chen et al. (2018) performed a global validation of the SMAP-, SMOS- and ASCAT-derived soil moisture data sets using a distributed model reference data set and Triple Collocation Analysis-based skill estimates over the globe. Error cross correlation between the data sets

were verified using in situ reference data from the ISMN data set. The objectives of the analysis were to:

- Perform a cross-assessment of the satellite-based products at the global scale
- Demonstrate the validity of the error orthogonality assumptions for the considered (Extended)TC triplets

For the analysis, short-term anomalies with respect to a 30-days sliding windows were calculated for each data set. The skill of the different products against the unknown truth were computed using Extended TCA theory (McColl et al., 2014), where the squared anomaly correlation can be computed through:

$$\rho_{t,X_i}^2 = \frac{(\beta_i \sigma_t)^2}{(\beta_i \sigma_t)^2 + \sigma_{\varepsilon_i}^2} = \frac{ubSNR}{ubSNR + 1}$$

With $(\beta_i \sigma_t)^2$ representing the unbiased SNR, i.e., the ratio of the signal variance to random error variance, β_i^2 being the scaling coefficient of the time series X_i and σ_t^2 the true signal variance, and with $\sigma_{\varepsilon_i}^2$ representing the error variance of X_i. Confidence intervals on the estimated correlation values were computed by means of moving-block bootstrapping (MBB) (Ólafsdóttir and Mudelsee, 2014), which accounts for the auto-correlated nature of the soil moisture signal. The error orthogonality assumption is validated in the study by comparing the ETC-based correlation results with Quadruple Collocation Analysis results including the in situ data set.

Differences in data and validation methods

- Due to the availability of data sets on the platform, the validation was run with different versions of the data sets:
    - SMAP L3 daily images gridded on the EASEv2 grid was replaced in the QA4SM-based validation with SMAP L2 v6
    - SMOS L3 v300 was replaced in the QA4SM-based validation with SMOS L3 v700
    - ECMWF land surface modelling scheme was replaced in the QA4SM-based validation with ERA5
- ETC-based correlation results were masked for positive cross-correlation estimates between the pairs in the TCA triplet. Such method can be easily replicated in QA4SM using the correlation outputs.
- A minimum of 50 points was used for the TCA results.
- Currently, the bootstrapping method implemented in QA4SM does not account for auto-correlation in soil moisture, and therefore neglects the real sample distribution.

Discussion of results and differences

The current QA4SM validation scores suite does not include the correlation parameter ρ_{t,X_i}^2, which poses the biggest limitation to the reproducibility of this particular literature source.

With respect to this, the differences in the input products used is almost negligible, and especially the difference in the model data set, which is not expected to affect the TC results. Nevertheless, the comparison between the correlation scores from the reference and the SNR from the QA4SM run are shown in Figure 11. Generally, the patterns agree quite well, especially over Australia, Africa and the US. However, it should be considered that the correlation parameter does contain additional information on the sensibility of the measured signal (McColl et al., 2014), and therefore differences in space should be expected. Moreover, the impact of the different threshold on the sample size for TCA in QA4SM (100) and in the reference (50), together with the additional flagging for correlation used in Chen et al. (2018) results in sometimes very different coverage of the validation scores.
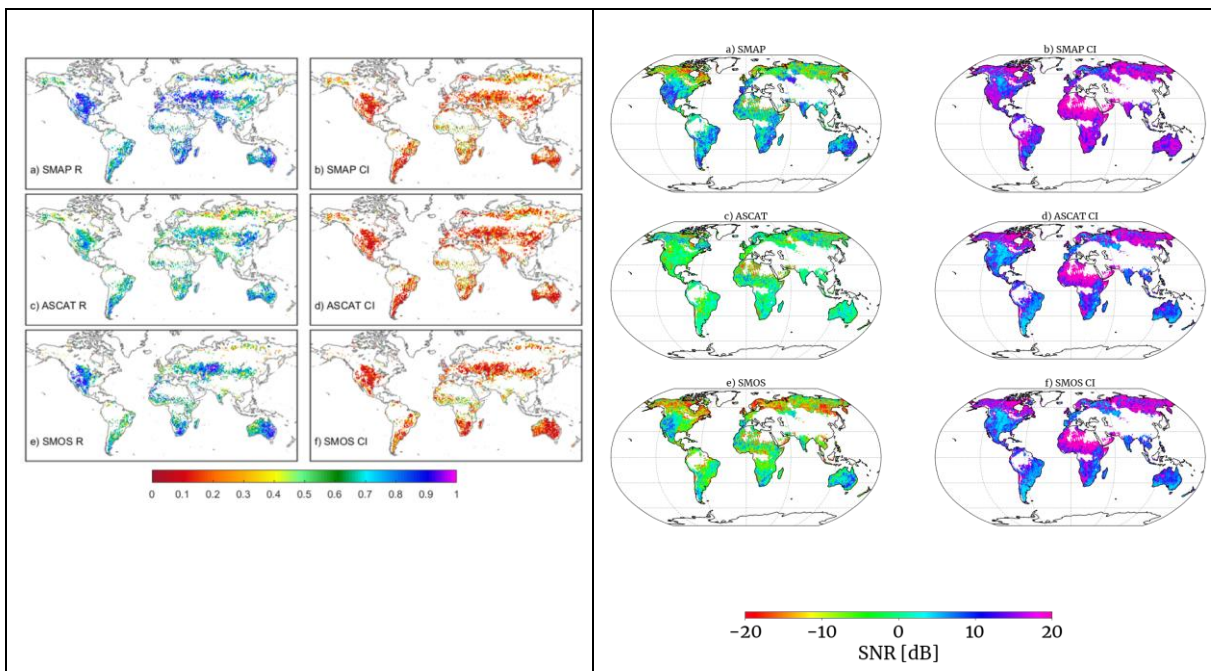
Figure 11: Qualitative comparison between the ETC-based correlation and relative confidence intervals from Chen et al. (2018) (left) and the QA4SM-based SNR estimate and relative confidence intervals (right).

Key learning points

From the point of view of the SMOS L2 integration assessment, the current comparison shows that a large-scale analysis of the product, involving a gridded reference and several other products to cross-compare, is possible. However, several methodological difference between the reference and the QA4SM-based validation pose important limitations in the analysis. Nevertheless, they offer points of potential improvements in the platform:

- Although it is not discernible from the current analysis, neglecting the soil moisture auto-correlation has an impact on the estimated confidence intervals, possibly resulting in an underestimation of their size. This aspect should be quantified in future verification activities and the application of a method to correct for this, such as MBB

or "thinning" (i.e., a reduction of the bootstrapping sample size based on the estimated auto-correlation) should be discussed.

- The ρ^2 metric provides a complementary way to assess soil moisture products with respect to the metrics listed in the good validation practices reference of Gruber et al. (2020), and does not pose particular challenges with respect to its implementation in QA4SM. Therefore, it should also be considered for a follow-up QA4SM release.

7.2.1.2 Comparison with Zhang et al., 2019

Zhang et al. (2019) perform a global validation of the SMAP L3 product (v2) against in situ measurements from the ISMN network. The aim of the study is to determine the product performance against static (organic carbon content, land cover, ...) and dynamic (vegetation, etc.) parameters, for ascending and descending overpass times separately. The validation period used is April 2015 to March 2018, and a thorough ISMN site selection is applied:

- A coverage of at least 100 data points is required
- Only the most representative station per SMAP grid point is selected, based on the correlation with SMAP itself and model data
- Stations in areas at high topographic complexity or with wetland fraction are excluded

Standard difference (BIAS, ubRMSE, RMSE) and correlation (Pearson's R) metrics are used to assess the data set, and stratified based on a range of environmental parameters.

Differences in data and validation methods

In QA4SM, the validation was repeated using the newly integrated SMAP L2 (mixed ascending/descending, v8), and the ascending/descending orbit observations of SMAP L3 (v6), https://test.qa4sm.eu/ui/validation-result/a814eee9-e625-49bd-82f6-1fd6852d96f3.

Besides the difference in the input data sets, the main deviations in the methods come from the selection criteria applied to the reference data. However, all the pre-selection criteria (which are based on the station location and sample size) can be applied to the QA4SM results as post-processing steps, using the coordinates of the validation points and the # observations output (not shown here). Besides this, all validation parameters could be replicated in QA4SM, although many were not explicitly mentioned in the reference (i.e., temporal collocation methods). Finally, the anomalies computation could be performed in the same manner as suggested in Zhang et al., but with a 35 days window (instead of 30). Given the selection differences, the number of stations in QA4SM are increased by a 10-fold factor (Figure 12).
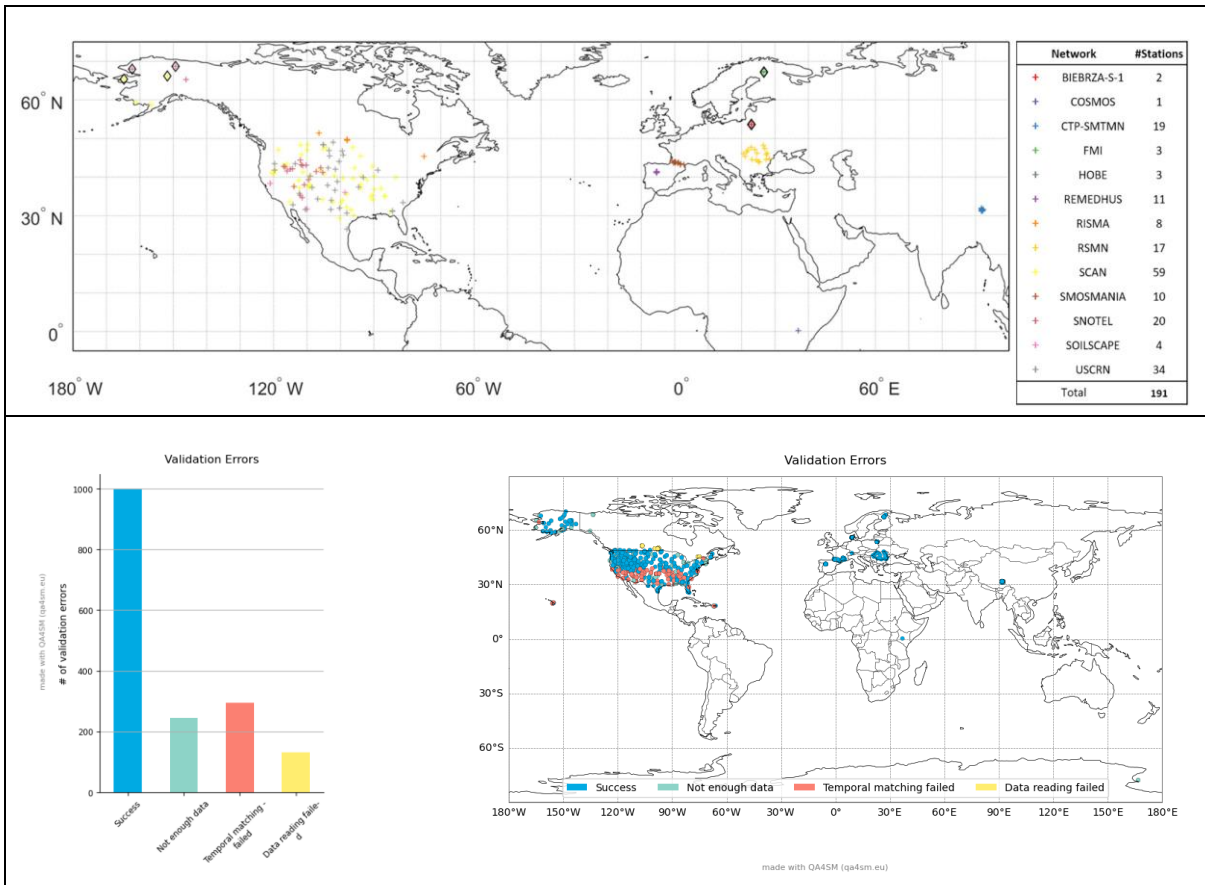
Figure 12: comparison of the ISMN validation points from Zhang et al. (2018) (above) and valid QA4SM outputs (below)

All of the environmental parameters used in the reference are also present in the QA4SM metadata-stratified plots (i.e., climate type, land cover for ESA CCI data set, soil property), or can in general be derived from the station locations. However, the "dynamic" parameters (i.e., vegetation conditions, land surface temperature) are not considered in the QA4SM validation routine.

Discussion of results and differences

Figure 13 compares the main, unclassified results obtained by Zhang et al., with those obtained in the QA4SM run. Despite the SMAP data sets versions and the ISMN reference points selection being overall different, the results appear close both in terms of absolute values and relative to one another. The analysis of Zhang et al., just like the one reproduced with QA4SM, finds no substantial difference between the AM and PM SMAP L3 product, with the AM product generally performing better. This can be similarly learned from the QA4SM scores, the only difference being the AM data set correlation being slightly higher compared to the PM. Hence, the main conclusion of the study can be qualitatively replicated in QA4SM.
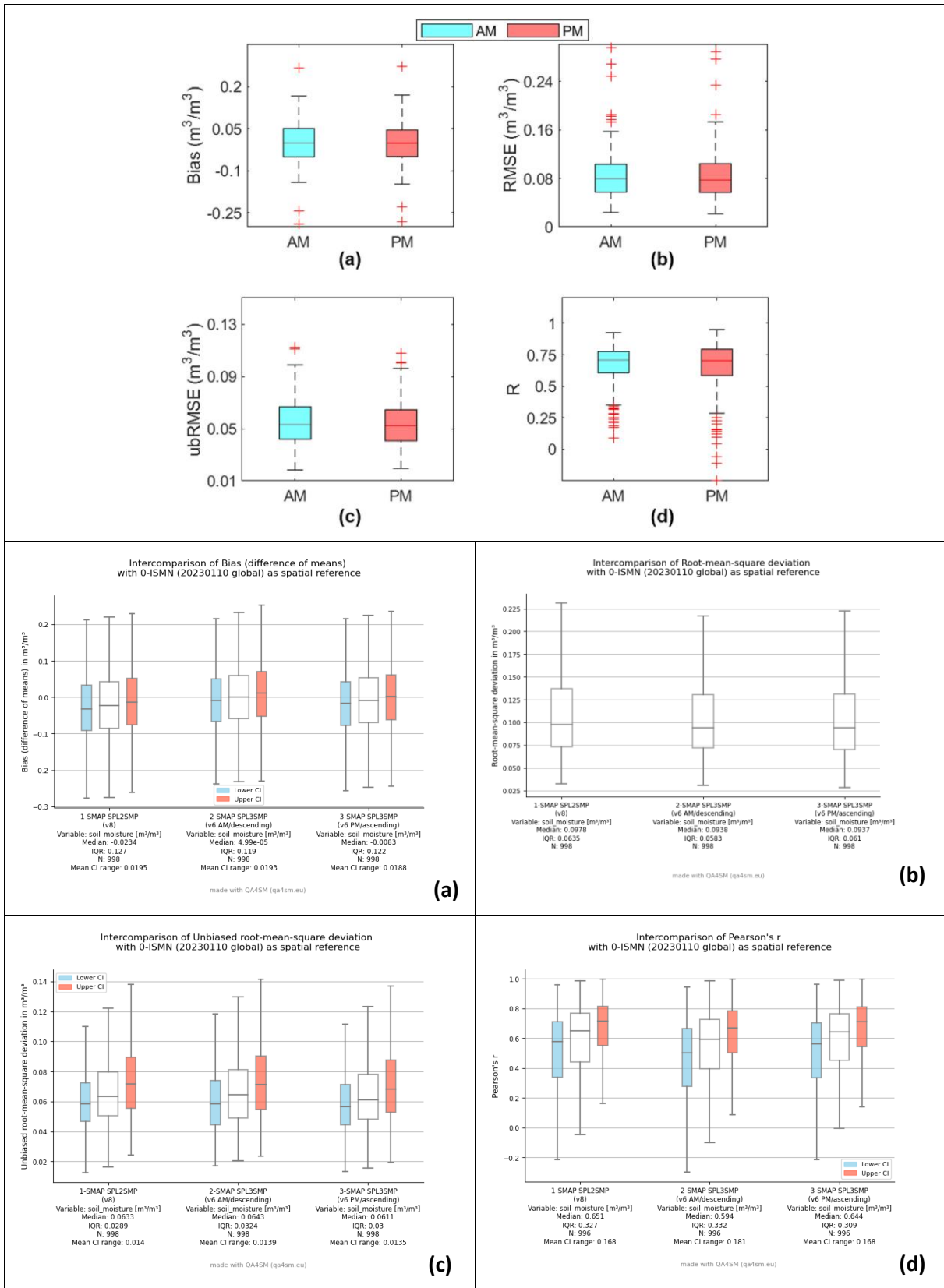
Figure 13: Comparison of the unclassified results from Zhang et al. (2018) (above) and QA4SM (below)

The comparison of the SMAP L2 and L3 (AM/PM) products shows that there is no substantial difference over the validation sites. This increases the confidence in the overall data set integration in the QA4SM platform.

Key learning points

- The study by Zhang et al. (2018) could be almost entirely replicated in QA4SM, the only difference being the in situ points selection and the use of "dynamic" conditions as stratifiers for the results. The former difference can be easily levelled by performing a post-validation results selection, while the latter is at the current moment outside the scope of the QA4SM platform.
- The main conclusion of the paper was replicated in QA4SM in qualitative terms.
- Overall, the validation of the newly integrated SMAP L2 product did not raise particular issues and yielded comparable results to those found in literature.

7.3 Verification of error handling and logging in QA4SM

Verification of proper error handling and reporting consists of 2 steps: (1) Verification of proper error handling and passing of status codes in the backend package pytesmo (automated unit tests), and (2) verification of proper displaying of error information to the user (manual tests on QA4SM platform).

7.3.1 Pytesmo unit tests for error handling

Unit tests for pytesmo error handling have been set up and can be accessed publicly on Github:

https://github.com/TUW-GEO/pytesmo/blob/master/tests/test_validation_framework/test_error_handling.py

These tests are run as part of the continuous integration (CI) pipeline after any changes made to the master branch, and as part of the automated release CI pipeline, so that releases can only be published if all tests run successfully.

The current tests verify proper raising and handling of all foreseen errors that can arise from malformed data, as well as proper handling in case no error occurs, i.e., all of the following status codes:

- OK: Validation run successfully and returned valid results
- METRICS_CALCULATION_FAILED: Data for all data sets is available, but due to unforeseen reasons the calculation failed.
- INSUFFICIENT_DATA: Data for all data sets exists, but not enough to calculate metrics (by default >10 data points are required).
- NO_TEMP_MATCHED_DATA: At least one of the datasets doesn't have data, or no temporal overlap with other datasets.

- TEMPORAL_MATCHING_FAILED: Temporal matching failed due to other reasons
- SCALING_FAILED: The scaling procedure failed, e.g., due to insufficient data
- NO_GPI_DATA: One dataset does not have any data at the given grid point index.
- DATA_MANAGER_FAILED: For some unforeseen reason, the calls to get data from the data manager fail

The error codes UNCAUGHT and VALIDATION_FAILED cannot be tested, because they refer to behaviour that is not foreseen in any way and can therefore not be tested.

7.3.2 Verification of QA4SM error handling

For the verification of the proper error handling and displaying, 3 manual test cases have been designed and performed:

7.3.2.1 Complete success

This test case has been set up to show 100% success by choosing a domain (Southern France, 40.7, -4.52, 46.3, 5.93) and time period (2010-2019) with good data coverage of both involved data sets (ISMN, C3S). As expected, the test summary indicates that "for 0% (0 of 72) of the processed locations (grid points) the validation metrics could not be calculated", and the map of status codes indicates only successful validations.

Validation errors for 1-C3S SM combined (v202012) with 0-ISMN (20230110 global) as spatial reference

7.3.2.2 Mixed success and failure

By choosing an earlier time period (2000-2009), at some of the locations not enough data is left for calculating metrics. Therefore, in this case "for 41.667% (30 of 72) of the processed locations (grid points) the validation metrics could not be calculated". The mix of successful and failed validations is also visible from the map, which gives further information on which locations did not provide enough data.

Validation errors for 1-C3S SM combined (v202012) with 0-ISMN (20230110 global) as spatial reference

7.3.2.3 Complete failure

A failure at all validation points can be artificially triggered by choosing a very short period of only a few days, here we chose the same domain as above but with a validation period from 2009-31-12 to 2010-01-04. In this case, the results overview indicates that "for 100% (72 of 72) of the processed locations (grid points) the validation metrics could not be calculated". The status map is produced nevertheless and shows that all locations failed due to insufficient data.

Validation errors for 1-C3S SM combined (v202012) with 0-ISMN (20230110 global) as spatial reference

7.4 Verification of Fiducial Reference Measurements flag integration

7.4.1 Overview

A filter option was added to all available ISMN versions in the service, to only consider sensors from the selected ISMN stations classified as "representative" or "very representative" by the Triple Collocation Analysis (TCA) based FRM Quality Identifier (QI; described in more detail in the FPP_SM document; definition also given in Figure 14) in a validation run. In total 1880 ISMN time series have a valid QI assigned.

For 2946 time series TCA was possible (all sensors below 10 cm depth and sensors for which TCA was not possible are therefore excluded in the following plots). 24 % of the tested sensors are classified as "not representative", 19 % as "representative" and 21% as "very representative". For the remaining 36 % TCA was possible, but the sample size was too small (<100 data points) and/or the SNR 80%-confidence interval was too large, to assign a QI, so that these sensors are marked as "undeducible".

Estimated SNR[(1)]	QI class	# sensors
NaN	undeducible[(2)]	1066
< 0 dB	not representative	720
0 – 3 dB	representative	556
> 3 dB	very representative	604
[(1)] based on TCA with ERA5-Land and CCI Passive (0.25 deg)		
[(2)] when # samples < 100 or CI (80%) < 3 dB		

Figure 14 - QI classification based on TCA between ISMN, ERA5-Land and CCI Passive (25 km)

Figure 15 shows the distribution of FRM QIs by TCA based SNR (a) and Scaled Error Standard Deviation (b), and TCA sample size (c). As expected, SNR and sample size increases with "representativeness", while random error variance decreases.
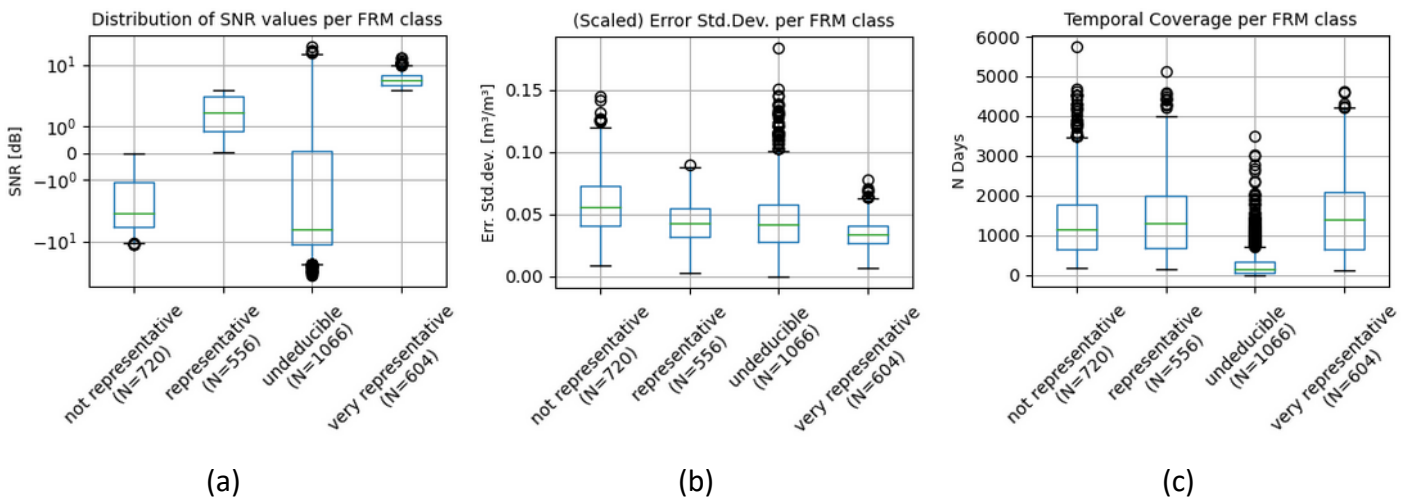

(a)   (b)   (c)

Figure 15 - Distribution ISMN QIs by TCA metrics: SNR (a), (scaled) random error standard deviation (b) and TCA sample size (c). "undeducible" here contains only sensors for which TCA was performed, but the number of observations was found too low, resp. the CI was found to be too large to perform a classification based on the estimated SNR.

### 7.4.2 FRM sensor coverage and FRM filter verification

The following test case, assesses the availability of FRM data. The same validation between ERA5-Land (data in the service available between 2001-01-01 and 2019-04-30; filtering for frozen soils is applied) and ISMN ("G"-flagged measurements only) is computed twice:

(a) once with all available ISMN stations (FRM QI filter **off)**

(b) once with only the (very) representative sensors.

Figure 16 shows the difference in in the number of available (in situ) time series by comparing the newly added "status" plots (note that dots overlap when multiple sensors in depth 0-10 cm are available). The lower number of available sensors when using only FRMs is clearly visible in Figure 16 (b).
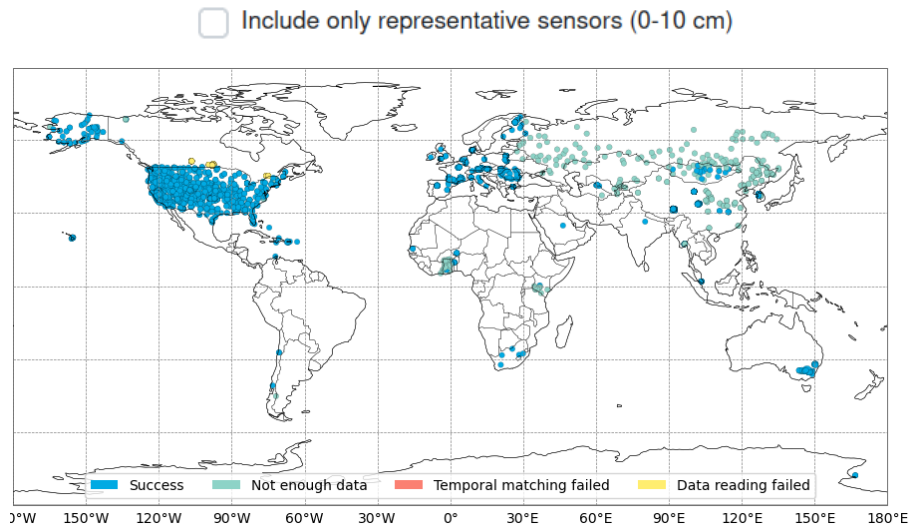
(a) ISMN vs. ERA5-Land, using ALL available ISMN sensors in 0-10 cm depth. N=3329
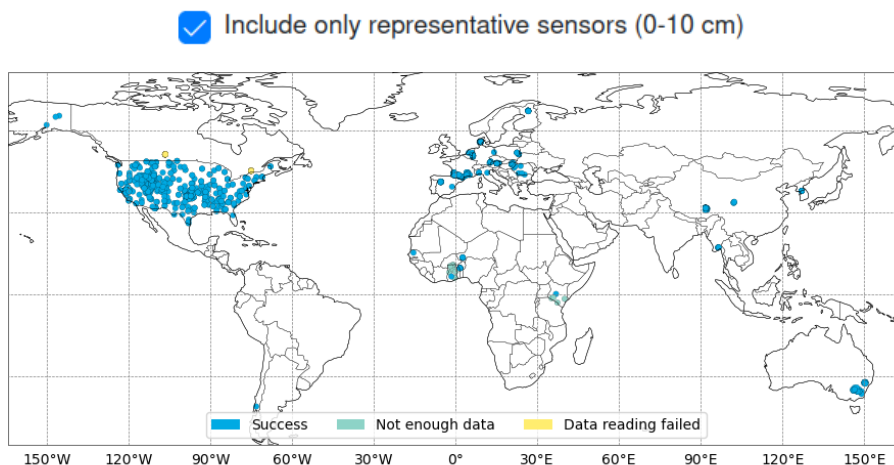
(b) Validation errors: ISMN vs. ERA5-Land, using only FRMs in 0-10 cm depth. N=1016

Figure 16 - Location of available ISMN time series in a validation run with ERA5-Land when using all available ISMN stations in QA4SM (a) vs. when using only FRMs (b).

In case (a) for 2834 of 3329 ISMN time series, validation metrics were computed. In case (b) this was the case for 982 of 1016 attempted ISMN time series. The 34 non-successful instances are due to the limited time period of ERA5-Land data in the service, the fact that ERA5-Land soil moisture is masked out in the validation runs when the soil temperature is negative and the potential inaccuracy of the ERA5-Land land mask.

The discrepancy between the total number of sensors used in validation runs (a) resp. (b) and the number of sensors with a QI assigned in Figure 15 (3329 vs 2946 in (a), resp. 1160 vs. 1016 in (b)) has two reasons:

1. The number of sensors described in section 7.4.1 only refers to those where TCA was possible. However, validation run (a), contains all available ISMN sensors, even those for which no TCA was possible (QI "undeducible" is still assigned during preprocessing)

2. The difference in number of "representative" plus "very representative" sensors (i.e. total number of points) in validation run (b) and the numbers in section 7.4.1 is due to vertically installed ISMN sensors that range across the defined max. depth threshold of 10 cm set in the validation run. While some of these sensors have a valid FRM QI assigned, they are not used in QA4SM when a max. validation depth of 10 cm is selected. Assigning a single depth value to each ISMN sensor instead of a depth range could resolve such ambiguities.

Figure 17 shows the impact this selection has on the computed correlation scores. R values in (b) are overall higher than in (a), especially outliers close to or below 0 are removed, resulting in an overall median increase of ~0.1.

The FRM classification is treated the same ways as any other metadata provided with ISMN data (e.g., landcover type, soil composition). Therefore, QA4SM creates box plots for each metric, where results are split based on the newly introduced metadata class. It is therefore possible to verify that only the "representative" and "very representative" sensors were used in validation run (b).

(a) R between ISMN and ERA5-Land, using ALL available ISMN sensors in 0-10 cm depth

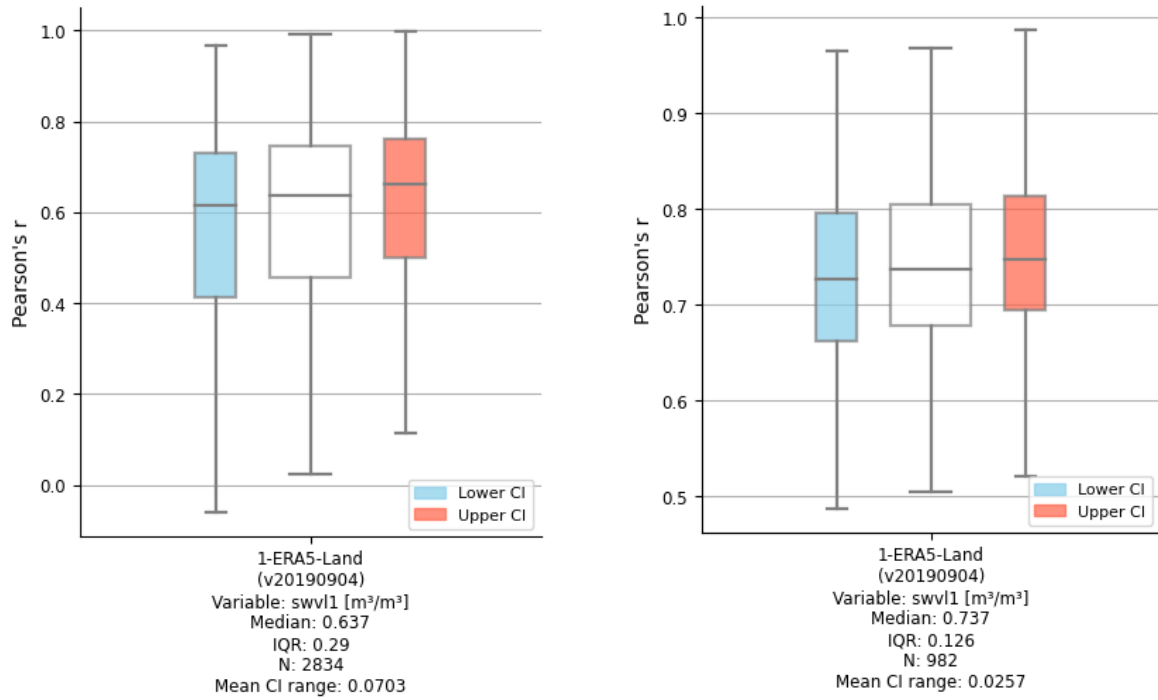(b) R between ISMN and ERA5-Land, using only FRMs in 0-10 cm depth

Figure 17 - Comparison of R between ISMN and ERA5-Land using all available ISMN measurements in 0-10 cm depth, vs. only using "representative" and "very representative" FRMs.
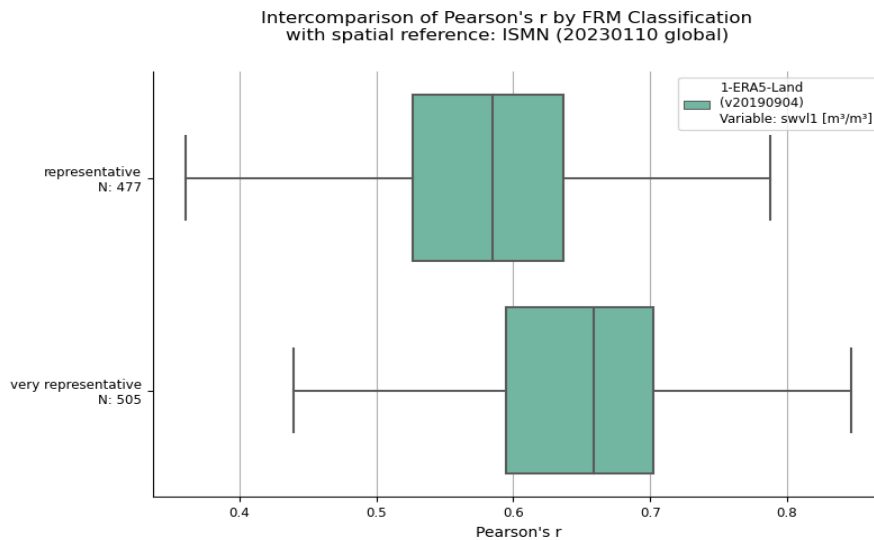

Figure 18 - R split by FRM class in validation run (b) where only FRMs are used

This test case confirms that ISMN data is properly filtered when then option to use only FRMs is activated in the QA4SM graphical interface.

A more detailed analysis on the impact the new filter has on the computed validation scores is given in the next chapter for various validation runs with satellite and model data sets.

7.4.3 Impact of using FRMs on validation results

Validation runs in Table 2 were preformed to test the impact on validation scores and number of available time series, when using only FRMs vs. when using all available ISMN data. The two active products (CGLS SSM and SWI) are given in units of [% sat.], GLDAS Noah Soil Moisture is given in [kg/m²]. They are therefore scaled to ISMN using linear regression scaling (affects the computed ubRMSD but not the R score). In all cases, the validation runs were successful, FRM classifications were transferred into the results file and all metadata-based box plots (including the new plots by FRM class) were created.

Table 2 - List of validation runs to verify impact of FRM flags option in QA4SM (*i.e., validation run was carried out successfully, validation metrics were computed, output netcdf file and plots generated.)

ID	ISMN data	Satellite Data	Comment	Success*
VR1.1	ALL of ISMN v202312	ESA CCI SM Passive v6.1 (0.25 deg)	This satellite data set was also used to perform TCA that the FRM QIs are based on.	Yes
VR1.2	FRMs of ISMN v202312	ESA CCI SM Passive v6.1 (0.25 deg)	Same as above	Yes
VR2.1	ALL of ISMN v202312	CGLS SSM (~1 km)	Higher resolution than expected representative area of FRMs. Different units, satellite data was scaled to ISMN via linear regression matching.	Yes
VR2.2	FRMs of ISMN v202312	CGLS SSM (~1 km)	Same as above	Yes
VR3.1	ALL of ISMN v202312	CGLS SWI T=5 (~1 km)	Same as above	Yes
VR3.2	FRMs of ISMN v202312	CGLS SSM T=5 (~1 km)	Same as above	Yes
VR4.1	ALL of ISMN v202312	ERA5-Land (~9 km)	This satellite data set was also used to perform TCA that the FRM QIs are based on.	Yes
VR4.2	FRMs of ISMN v202312	ERA5-Land (~9 km)	Same as above	Yes
VR5.1	ALL of ISMN v202312	GLDAS Noah (0.25 deg)	Different units, satellite data was scaled to ISMN via linear regression matching.	Yes
VR5.2	FRMs of ISMN v202312	GLDAS Noah (0.25 deg)	Different units, satellite data was scaled to ISMN via linear regression matching.	Yes
VR6.1	ALL of ISMN v202312	SMAP L2 (~36 km)		Yes
VR6.2	FRMs of ISMN v202312	SMAP L2 (~36 km)		Yes

| VR7.1	ALL of ISMN v202312	SMOS L2 (~15 km)		Yes
VR7.2	FRMs of ISMN v202312	SMOS L2 (~15 km)		Yes
VR7.3	FRMs of ISMN v202312	SMOS L2 (~15 km)	Anomaly values (wrt. 1991-2020 reference)	Yes

The box plots in Figure 19 show changes in aggregated Pearson's R estimates for all available ISMN stations by FRM classification (VR7.1, VR5.1, and VR 2.1). In all three cases, an increase in R is found for the "(very) representative" sensors over the "not representative" ones. This applies to the radiometer scale, the model data set with similar spatial resolution and the high-resolution radar-based data. Note that the 3 data sets shown here are all independent of the data used to perform TCA to compute the FRM QIs.
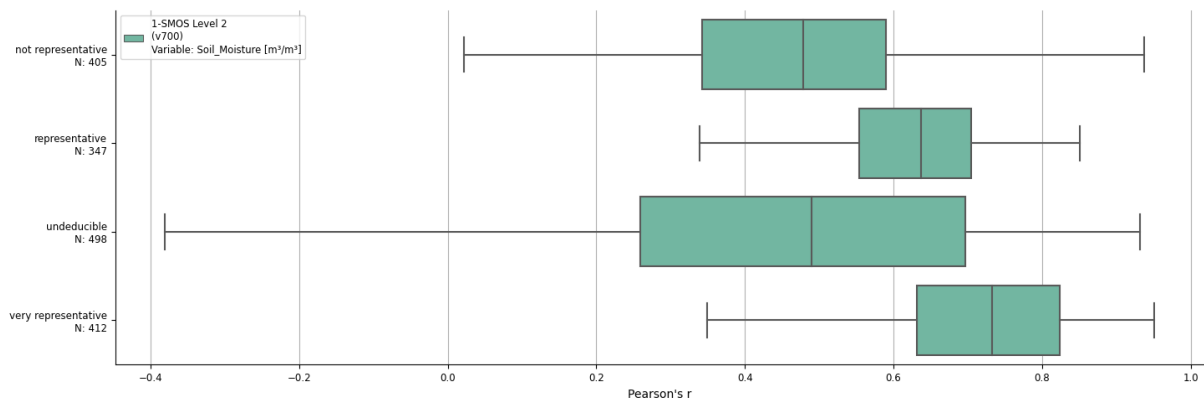
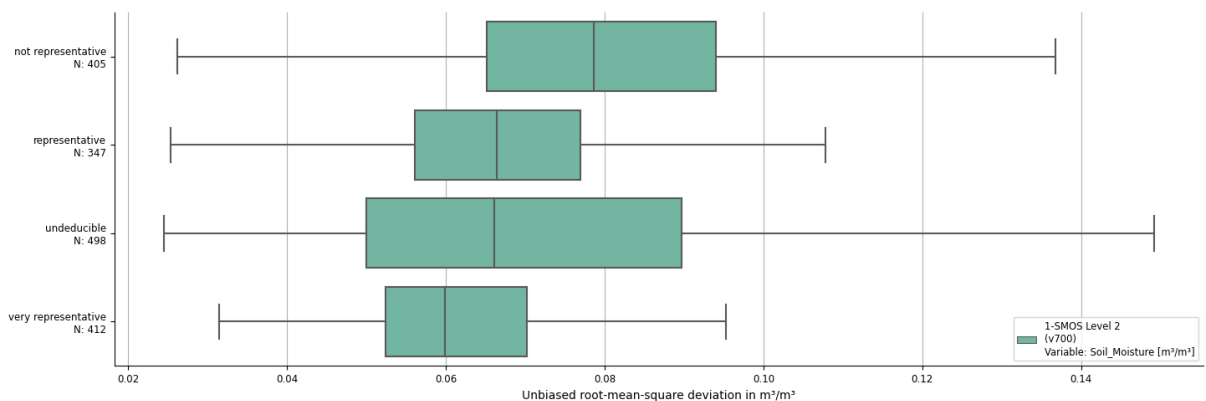(a) Change in R between ISMN and SMOS L2 over different FRM QIs (VR 7.1)
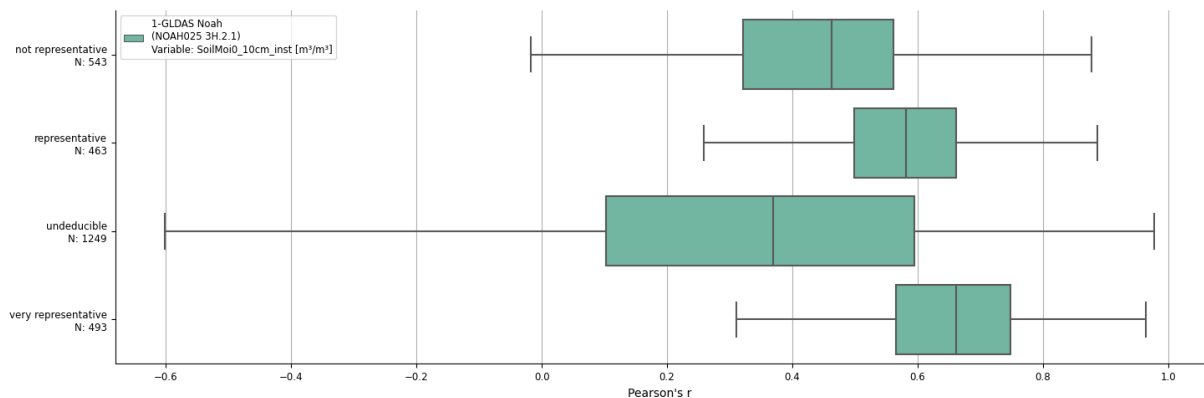
(b) Change in R between ISMN and GLDAS Noah over different FRM QIs (VR 5.1)
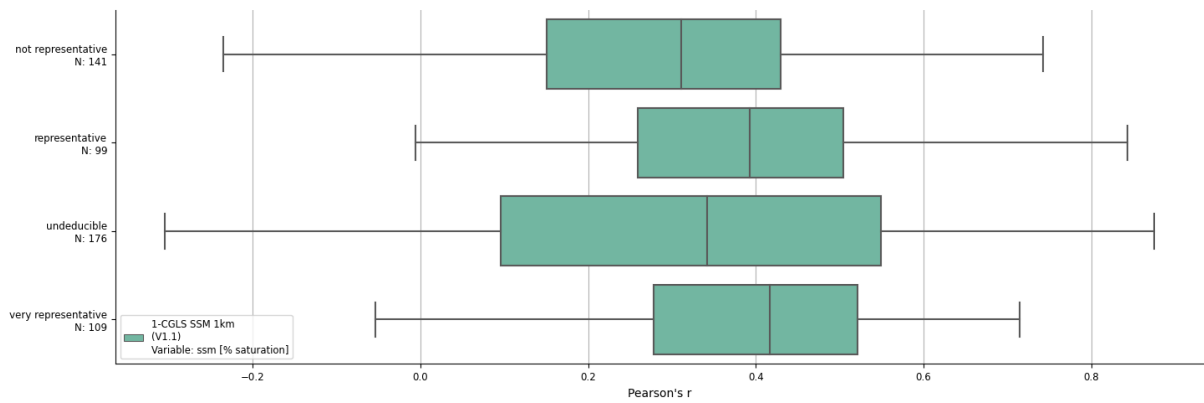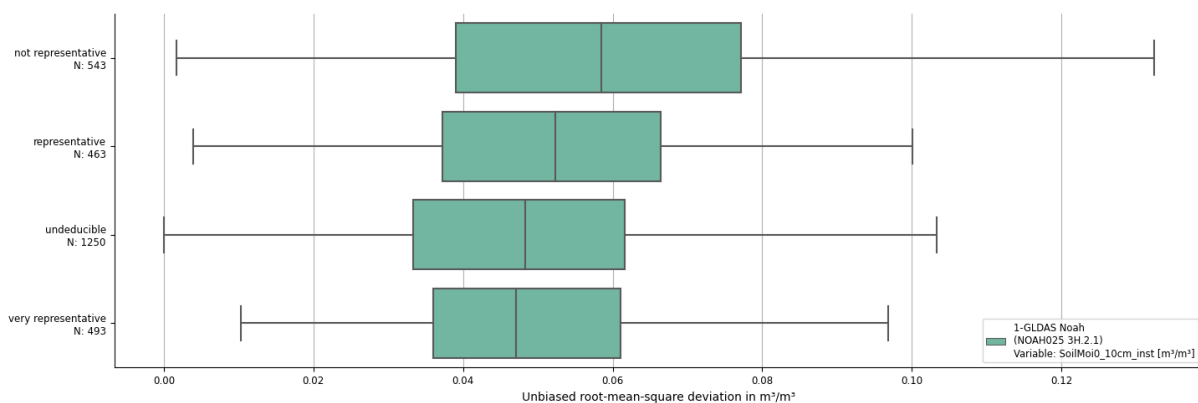
(c) Change in R between ISMN and CGLS SSM over different FRM QIs (VR 2.1)

Figure 19 - Change in R across different FRM classes between ISMN and SMOS L2 (a), GLDAS Noah (b) and CGLS SSM (c) Soil Moisture (taken from VR 7.1, VR 5.1, VR 2.1 respectively)

Similar changed across different FRM QIs are found when looking at ubRMSD (Figure 20). For all coarse resolution data sets, ubRMSD decreases with increased representativeness of ISMN sensors. However, in the case of the 1 km SSM data - even after removing the additive and multiplicative biases – the median of the "non representative" cases is below the median of the "representative" ones. The "very representative" sensors still show the best agreement with the satellite data. However, the spatial coverage of the CGLS data is limited to Europe only, leading to much smaller sample sizes in metrics compared to the coarse resolution data. The shorter temporal coverage of Sentinel 1 (data is only available after 2014) should also be considered here.

(a) Change in ubRMSD between ISMN and SMOS L2 over different FRM QIs (VR 7.1)

(b) Change in ubRMSD between ISMN and GLDAS Noah over different FRM QIs (VR 5.1)
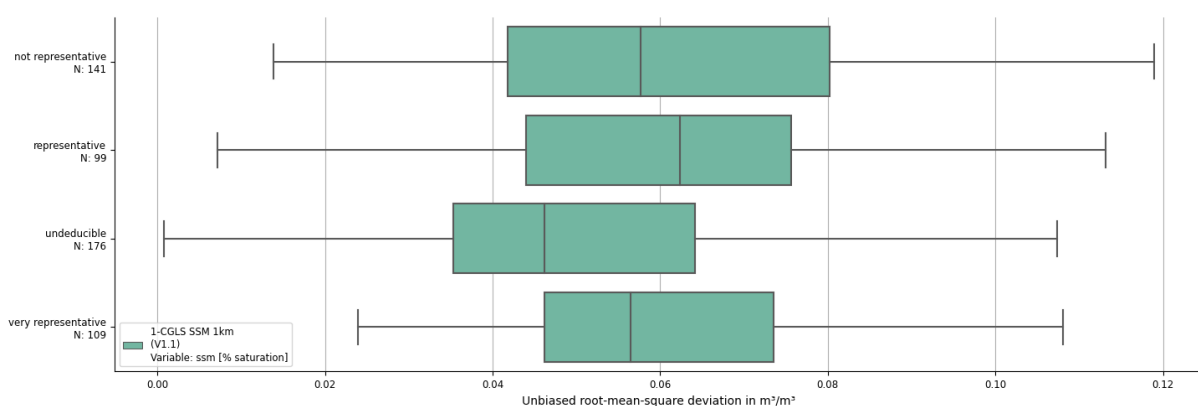
(c) Change in ubRMSD between ISMN and CGLS SSM over different FRM QIs (VR 2.1)

Figure 20 - Change in ubRMSD across different FRM classes between ISMN and SMOS L2 (a), GLDAS Noah (b) and CGLS SSM (c) Soil Moisture (taken from VR 7.1, VR 5.1, VR 2.1 respectively)

For the remaining validation runs, Table 3 compares the difference in median Pearson's R and ubRMSD when using all available ISMN sensors vs. using only the "very representative" and "representative" ones. In all cases, R increases when the FRM QI filtering option in QA4SM is activated, while ubRMSD decreases in all cases except for the two 1 km products. In addition, the number of time series decreases as expected.

Table 3 - Comparison of R and ubRMSD between validation runs (from previous table) that use all available ISMN time series between 0 and 10 cm depth (left half) and only FRM qualified time series (right half). Bold numbers indicate an improvement in the median of the computed metrics.

Validation Runs	Using all available ISMN stations ☐ Include only representative sensors (0-10 cm)			Only "very representative" or "representative" ☑ Include only representative sensors (0-10 cm)		
	R (median)	ubRMSD (median)	N (time series)	R (median)	ubRMSD (median)	N (time series)
VR1.1 &. VR1.2	0.496	0.0815	2577	**0.649**	**0.0732**	996
VR2.1 & VR2.2	0.364	0.0537	525	**0.398**	0.0586	212
VR3.1 & VR3.2	0.527	0.0482	545	**0.614**	0.0499	209
VR4.1 & VR4.2	0.637	0.0589	2834	**0.737**	**0.0554**	982
VR5.1 & VR5.2	0.515	0.0501	2748	**0.625**	**0.0493**	956
VR6.1 & VR6.2	0.629	0.0624	2007	**0.717**	**0.0574**	845
VR7.1 & VR7.2	0.603	0.0667	1662	**0.677**	**0.0624**	759

This first version of a FRM subset for ISMN data therefore identifies in situ sensors that generally correspond better with all tested satellite data sets than the global set. The generally low noise level and long time series available through SM FRMs currently comprise the best available reference data for validating satellite soil moisture products on a scale of 1 - 25 km. It should be noted that QIs are only assigned to in situ sensors in the top 10 cm soil layer. FRMs therefore only represent surface soil moisture.

8 References

Chen, F., Crow, W. T., Bindlish, R., Colliander, A., Burgin, M. S., Asanuma, J., & Aida, K. (2018). Global-scale evaluation of SMAP, SMOS and ASCAT soil moisture products using triple collocation. *Remote Sensing of Environment*, *214*, 1-13.

Cosh, Michael, C. Montzka, A. Al Bitar, A. Berg, R. Bindlish, H. Bogena, F. Cabot, T. Caldwell, S. Chan, A. Colliander, W. Crow, N. Das, G. de Lannoy, E. Lopez-Baeza, W. Dorigo, A. Gruber, S. Hahn, T. Jagdhuber, S. Jones, Y. Kerr, S. Kim, C. Koyama, M. Kurum, F. Mattia, K. McColl, S. Mecklenburg, B. Mohanty, P. O´Neill, T. Pellarin, G. Petropoulos, M. Piles, R. Reichle, N. Rodriguez-Fernandez, C. Rüdiger, T. Scanlon, D. Spengler, P. Srivastava, S. Suman, R. van der Schalie, W. Wagner, U. Wegmüller, F. Camacho, and J. Nickeson (2020): Soil Moisture Product Validation Best Practice Protocol. Version 1.0.

Dorigo, W., Himmelbauer, I., Aberer, D., Schremmer, L., Petrakovic, I., Zappa, L., Preimesberger, W. and others (2021). The International Soil Moisture Network: serving Earth system science for over a decade. Hydrol. Earth Syst. Sci., 25, 5749–5804. https://doi.org/10.5194/hess-25-5749-2021.

Gruber, A., De Lannoy, G., Albergel, C., Al-Yaari, A., Brocca, L., Calvet, J.C., Colliander, A., Cosh, M., Crow, W., Dorigo, W., and others 2020. Validation practices for satellite soil moisture retrievals: What are (the) errors?. Remote Sensing of Environment, 244, p.111806.

McColl, K. A., Vogelzang, J., Konings, A. G., Entekhabi, D., Piles, M., & Stoffelen, A. (2014). Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target. Geophysical research letters, 41(17), 6229-6236.

Ólafsdóttir, K. B., & Mudelsee, M. (2014). More accurate, calibrated bootstrap confidence intervals for estimating the correlation between two time series. Mathematical Geosciences, 46(4), 411-427.

Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, IJsbrand Jan; Appleton, Gabrielle; et al. (15 March 2016). *"The FAIR Guiding Principles for scientific data management and stewardship"*. Scientific Data. 3: 160018. https://doi.org/10.1038/sdata.2016.18.

Zhang, R., Kim, S., & Sharma, A. (2019). A comprehensive validation of the SMAP Enhanced Level-3 Soil Moisture product using ground measurements over varied climates and landscapes. *Remote sensing of environment*, *223*, 82-94.