<table>
<tr><td>  </td><td>Minutes of Meeting</td><td>Doc.-No:
FRM4SM_QA4SM_Workshop_v1.1</td></tr>
<tr><td>Project:
FRM4SM - Fiducial Reference Measurements for Soil Moisture</td><td></td><td>Doc.-Date:
16.07.2022</td></tr>
<tr><td colspan="3">QA4SM Workshop MOM and Survey Analysis</td></tr>
<tr><td colspan="2">Meeting Place:
6th Satellite Soil Moisture Validation and Application Workshop, Perugia (Italy)</td><td>Date/Time:
07.06.2022, 10.00 – 07.06.2022, 13:00</td></tr>
<tr><td colspan="3">Project Participants (Organisers and Presenters):
ESA: Raffaele Crapolicchio (presenter)
AWST: Monika Tercjak (organizer, techical support on line)
TU Wien: Wolfgang Preimesberger, Irene Himmelbauer, Pietro Stradiotti, Alexander Gruber (organizers, presenters)

Project Participants (Attendees):
AWST: Alexander Boresch
CESBIO: Arnaud Mialon, François Gibon</td></tr>
</table>

Meeting agenda

1. 10:00 - 10:30 CET: QA4SM platform and project overview

10:00 FRM4SM Project Introduction and Overview (Raffaele Crapolicchio)
10:10 QA4SM Introduction and Overview (QA4SM Development Team, Pietro Stradiotti)

- Scientific background of satellite SM validation
- Recent evolutions, current best practices
- Technical background / philosophy of QA4SM
- Status of the platform, goals and future developments

10:25 Q&A session (all)

10:30 BREAK

2. 10:40 - 11:50 CET: Hands-on session

10:40 Guided session on QA4SM use (QA4SM Development Team, Pietro Stradiotti)
10:50 Supervised hands-on session (all)
11:20 Open feedback session (all)

- Introduction Feedback Presentation (QA4SM Development Team, Wolfgang Preimesberger)
- Open discussion (all)

11:50 BREAK

3. 12:00 - 13:00 CET: SAG Session

12:00 Project / QA4SM feedback (SAG Memebers, QA4SM Development Team)

Workshop participants

Table 1: List of workshop attendees (data taken from the workshop registration form). Color code: BLUE = present, online; GREEN = present, in presence; YELLOW = not present. SAG members are highlighted in bold.

Name	***Organisation (optional)***
Raffaele Crapolicchio	ESA-ESRIN
Sly Wongchuig	Université Grenoble Alpes
Kleanthis Karamvasis	National Technical University of Athens
Kehao Yu	China university of geosciences (beijing)
Wouter Dorigo	Tu Wien
Kleanthis Karamvasis	National Technical University of Athens
Berhanu Geremew	BDU
Arnaud MIALON	CESBIO
SHAIK SALMA	NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA, INDIA
Harsh srivastava	Iiit allahabad
KANG CHUEN SIANG	UNIVERSITY TECHNOLOGY MALAYSIA
Amen Al-Yaari	NOT GIVEN
Marian Schönauer	University of Göttingen
Aura Salmivaara	Natural Resources Institute Finland
Abdolnabi Abdeh Kolahchi	Soil Conservation and Watershed Management Research Institute
Philippe Richaume	CESBIO
Maria Piles	**Universitat de València**
François Gibon	CESBIO
Carsten Montzka	**Forschungszentrum Jülich**
VITHLANI NIPA SUNILBHAI	JUNAGADH AGRICULTURE UNIVERSITY
Roberto Sabia	Telespazio-UK for ESA
Stephan Dietrich	**ICWRGC, BfG**
Anneli Ågren	Swedish University of Agricultural Science
Wolfgang Korres	Federal Institute of Hydrology
Matthias Zink	International Centre for Water Resources and Global Change
Robin van der Schalie	Planet
Fay Böhmer	Federal Institute for Hydrology (BfG)
Tunde Olarinoye	ICWRGC

Note: several attendees have joined in presence and not through the official registration form. The total number of participants (including organizers and attendees) was:

- In presence: 10
- Online: 11

Workshop presentations and general discussion

The agenda was maintained albeit with a few minutes of delay for the whole duration of the workshop. The feedback session was adapted due to the absence of most of the SAG members to include all workshop participants. This session extended well beyond the planned workshop conclusion (~20 minutes).

Raffaele Crapolicchio presented the FRM4SM framework, project partners and goals in the opening session of the workshop. Following, Pietro Stradiotti presented the QA4SM platform to the public, including the state-of-the-art in satellite soil moisture validation practices, the concept of QA4SM and the service state. Carsten Montzka (CM) thanked the QA4SM Development Team and the FRM4SM project participants at large for their effort in contributing to create a validation standard for the community.

The hands-on session was started with a live demonstration of the QA4SM platform by Pietro Stradiotti, with the online technical support from Monika Tercjak. Here, an overview of the main service components was made, including how to run a validation, access and share the results, analyze the outputs and compare different validations. Several questions were raised by the participants who showed good engagement during this part of the program. After the demonstration, the attendees had time to use the platform and perform validations of their interest, with the help of the QA4SM Development Team. While there was some concern on the performance of QA4SM due to the number of parallel usages, the service worked as expected and no abnormal behavior or service crash was observed. All together there were 26 validations run, which took between 1 and 7 minutes, with the average validation time of 2.5 minutes. The development team addressed questions and comments from the participants; many questions were related to misunderstanding of the information provided by the Graphical User Interface. This provided a few learning points for the QA4SM development team, including the necessity of more detailed information provided with the data sets.

Wolfgang Preimesberger presented the future evolution of the platform to open the feedback session. The presentation raised a few questions that guided the feedback from the attendees. The following discussion involved all the workshop participants and continued until the end of the program. At the break (11:50), few of the participants online left the meeting, while the SAG members where asked to participate to the last session.

Figure 1: The QA4SM Development Team presents the QA4SM Platform to the Workshop participants.

(SAG) Feedback session (11:20-13:20)

The feedback session was mostly centered around the User Upload feature that is planned for the upcoming QA4SM release, at the conclusion of the second year of the FRM4SM project. This was recognized as the most beneficial (technical) improvement for the service and one which most of the participants would find useful in their own application. Several technical and scientific aspects where discussed:

Philippe Richaume argues that the netCDF file format is acceptable from a user perspective; the Climate and Forecast (CDF, https://cfconventions.org/conventions.html) metadata convention is a good starting point. However, care should be taken in the formatting of irregular gridded data sets and accounting for the error that is introduced in the resampling. Especially with Level 2 data this could pose a threat that should be dealt with. As a starting point 3-dimensional netcdf files should be the baseline for the service, however, in future versions there should be more flexibility to also allow the validation of L2 data (2 netcdf dimensions).

Alexander Gruber agrees that the re-gridding should be performed in the fairest way possible, avoiding to benefit a specific data set in a comparison. However, the way forward to introducing this in the 'best practices' documentation is not clear.

Philippe Richaume argues that to deal with the issue of re-projection and possible distortions of the validation scores that may originate from it, the upload of data should not need resampling. This will facilitate the users by making the service capable of dealing with various data formats in a flexible way. To help the developers, several L2 grids that are more commonly used should be considered for a start. In particular the EASE grid (SMAP, SMOS) and CCI grids should be considered.

A general discussion on preprocessing finds that certain preprocessing steps are necessary in order to handle differences in the structures of data sets (e.g. spatial resampling). However, the service should be more transparent about these steps that could have a potential impact on the validation results.

Wolfgang Preimesberger raises a point on the acceptable size for users to find a compromise between finding the tool useful and facilitating the handling of the data on the server side. It is the general opinion in the discussion group that a few Gigabytes (up to ~10) should be enough for the various purposes. Another point raised by Wolfgang is how to find a trade-off between the traceability of the data sets uploaded by the users and the hardware limitations related to permanent data storage for potentially tens of users. Through discussion with the participants, it seems that a possible way forward is to have a 'temporary' storage of data sets in an allocated space to each user, where test data and non final versions can be uploaded and validated; here, the cleanup will be performed with a higher frequency. Along with this, a more permanent space could be made available, where users can upload data that can be validated by other users and used to publish validation results. In this case, the cleanup will happen less often.

The point of filtering user uploaded data is brought up. An option to consider flags in the uploaded data should be implemented to avoid unnecessary pre-processing on the user side before using the service.

Raffaele Crapolicchio makes a point of including 'temporal' metrics in the service, resp. time series showing the performance of the various data sets through time. This could be helpful especially in the (stability) evaluation of operational data streams, such as "drifts" in sensor measurements. The group discusses whether a framework exists for this. Alexander Gruber points out that there is currently no indication of best practices in 'spatial' validation metrics, i.e. metrics that could be calculated on a spatial sample and evolve at the time resolution of the satellite data set. "Autocorrelation metrics" could also be considered in the service.

Irene Himmelbauer brings up the point of uploading in situ (reference data) into the service. In general the service should focus on evaluating satellite data. However, the origin of the measurements is not relevant for the service, and any data can be uploaded as long as the format requirements are fulfilled. All user uploaded data will be treated as "candidates" for validation. Identifying "reference data" should be handled on the service developer side.

Raffaele Crapolicchio brought up the issue that it is not clear that the ERA5 data in the service has a 6-hourly temporal sampling, while most people familiar with the data set would expect the original, hourly sampling.

FRM/QA4SM feedback analysis

The feedback form was distributed in online format to the members of the QA4SM Users Workshop as well as to the QA4SM users mailing list and (separately) to the Scientific Advisory Group (SAG), collecting a total of 7 responses. Of the SAG team, only Carsten Montzka signed their name in the survey. The questions asked in the form are reported in Appendix A.

Note: providing responses to the questions in the feedback form is optional for the participants. Therefore, the number of responses varies from question to question and from the total number of participants.

Participants background

The participants to this survey label themselves as researchers or scientists and are part of research institutions, universities or governmental agencies. All the participants perform soil moisture validations often (2) or sometimes (5), indicating some familiarity with the concepts inquired in the form.

The focus of the different participants is well distributed among the given options, where 'application' (comparison of different products, own product development) and validation 'method development' where selected an equal number of times. The 'application' refers to the development of upcoming versions of own data (2) while under 'method development', 'best practices', 'metrics robustness' and 'merging indicators' were mentioned. Almost all participants (5) know and use the best practice documents (Gruber et al., 2020; Montzka et al, 2021) and/or literature related to specific validation aspects.

The regularly used data sets vary between the participants, although 'satellite', 'model' and 'in situ' were equally selected; SMOS, SMAP, AMSR2/E, H-SAF, ERA5(-Land) and GLDAS were mentioned as satellite and model data sets, while the in situ data used comes generally from ISMN or undistributed sources.

FRM4SM

In general, all participants responding to the question (5) agree with the provided definition of FRMs. One participant remark is that *"[…] I would emphasize the inner trust in the in-situ data to be correct, stable in quality (aging probes, aging plot setup characteristics) and fully characterized in term of errors, calibration. The main goal […] is that any departures to in-situ shall come only (or mostly) from scales representativeness, sampling and/or satellite SM limitation, issues and/or errors."*, and other stress 'reliability', 'traceability' and 'trustworthiness'.

Of the responses (6), 3 participants are familiar with the concept of 'committed area' and think that it should be included in the validation process; 2 of these responses have been provided by participants that have indicated a 'method development' focus.

In terms of the feature importance in the qualification of in situ measurements as FRMs (question 7), the 'time series length and temporal coverage' of the measurements is agreed by the participants as one of the most important qualities in the measurements; the 'sensor calibration' and 'number of stations in a satellite pixel' come immediately after (Figure 2). This indicates that there a high interest in the quality and quantity of the data production. According to the participants, all other features, including factors that are determined in the data analysis and post-production (e.g., 'committed area', 'spatial representativeness', 'scale mismatch') are less relevant.

The tradeoff between spatial distribution and density (question 8) of the measurements was divisive among the participants, with 2 preferring 'High-quality / error characterized stations in a small number of networks' and 3 preferring 'A large number of (rather globally distributed) in situ stations..', where the point is made that: *"the quality assessment of a satellite based on only a specific surface condition gives a very limited idea of the satellite's overall quality […]. Moreover, the goal of validation is not to determine the highest possible accuracy, but to quantify the difference between the satellite measurement and the measurement from a probe under all conditions"*.

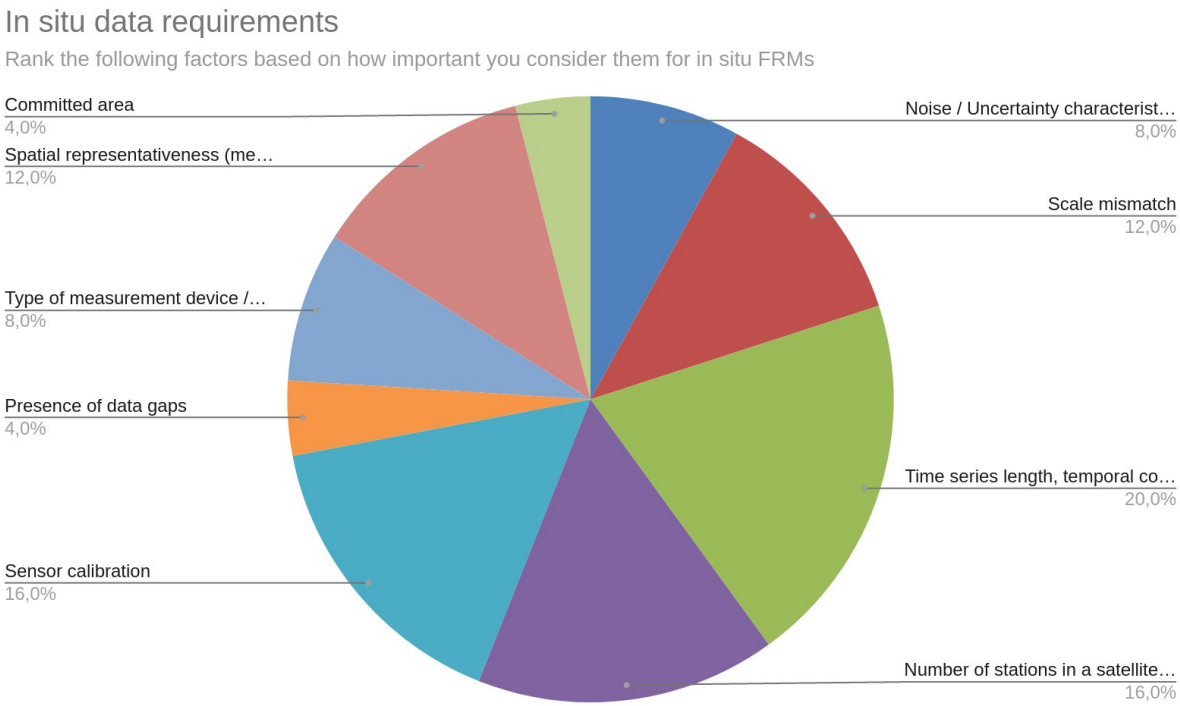

Figure 2: Feedback responses to question 7 (Appendix A). The hit counts per feature in the first 5 position in the ranking (by importance, according to the participants) is shown.

When asked about the measurement characteristics that FRMs should necessarily provide (question 9, Figure 3), 'Flagging information for time series quality' is the one on which most users agree (6 of 7). All others are subject to disagreement. In general, distributing accurate flagging information on time series quality and environmental conditions with the in situ data seems to be high in the priority of the participants.

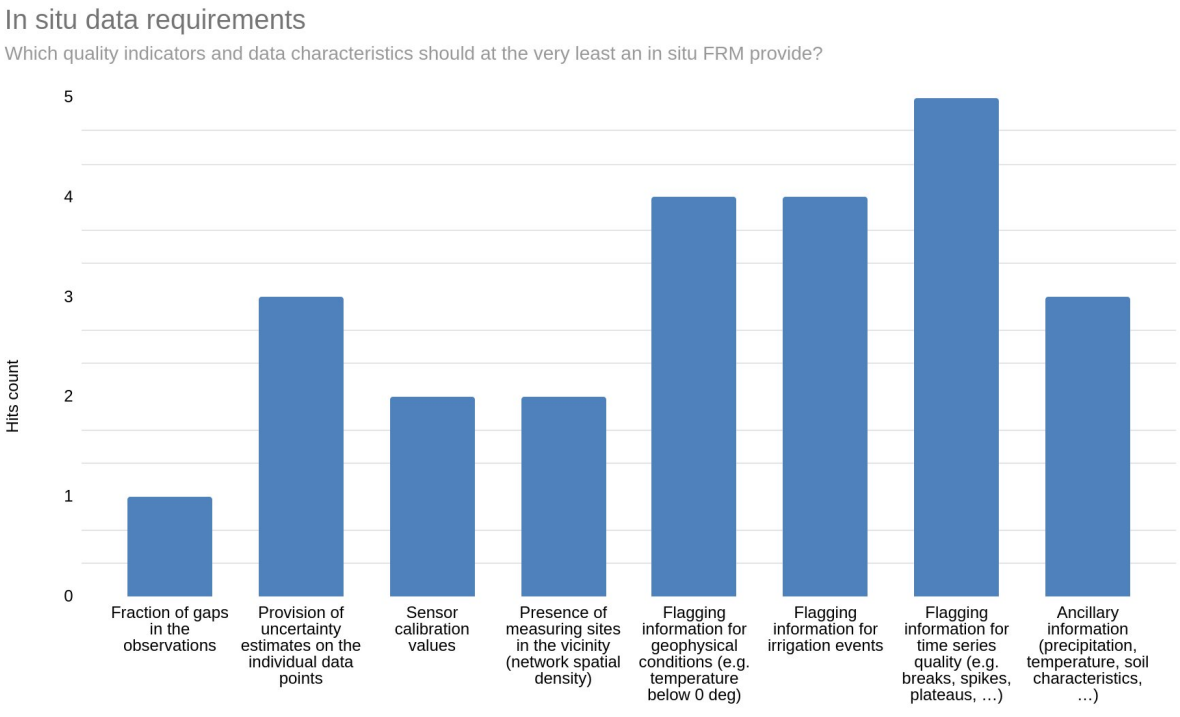

Figure 3: Feedback responses to question 9 (Appendix A).

On the minimum length of a measurement time series (question 10), the answers vary depending on the aimed validation scope; however, most participants consider one year as the minimum requirement for the representation of at least one seasonal cycle that includes the impact of vegetation. 2 participants indicate 3 years as a sufficiently robust sample for the metric calculation, one indicates 5 years as the minimum plausible period to capture hydrologic extremes.

Similarly to question 8, question 11 inquires the relevance of measurement representativeness in the specific application field of the users. Once again, most answers (4 of 5) consider measurement distribution and representativeness of environmental conditions important. One participant highlights the importance of densely instrumented, controlled in situ networks for the validation of high resolution (Sentinel 1) data.

All users point at hourly sampling intervals in the measuring of in situ soil moisture for FRMs, which is compatible with the resolution provided by ISMN. According to the participants, all stations with a lower temporal resolution or frequent data gaps will compromise the ability to capture diurnal cycles and wetting/drying events, and will introduce temporal biases in the validation of satellite overpasses.

With regards to the environmental factors that are considered important in the use of FRMs (question 13, Figure 4), land cover type and measuring depth are considered equally important. The measuring technique used by the in situ sensor is mostly disregarded; one participant includes 'ground temperature', 'vegetation state' as relevant environmental factors.

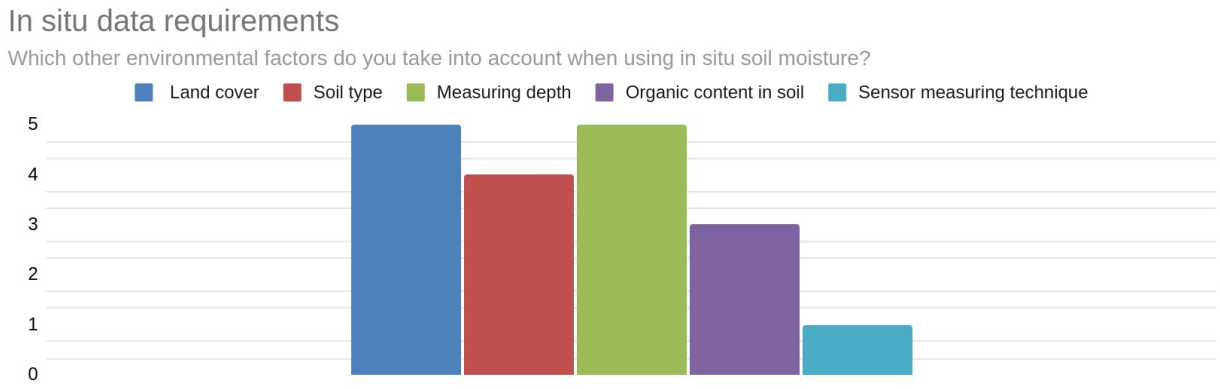

Figure 4: Feedback responses to question 13 (Appendix A).

QA4SM

This section is geared towards understanding the user requirements and expectations of QA4SM in view of the second release foreseen in the FRM4SM project. The answers to the first question (# 15) are equally in favor of 'A scientific platform where users can play around with different validation setting (where "wrong" settings can be chosen)' (2) and 'A "best practice" platform, where options are limited and where the platform maintainers make many decisions in the background.' (2). The answers are unrelated to participants choosing 'method development' and 'application' as validation focus (question 2). Users who opt for a 'best practice' option argue in favor of standards in validation: *"We need a highly standardized validation. If too many options can be chosen, the comparability to other validation exercises is lost. I would prefer a single standard validation setting, with the option to modify it […], and automated information of the changes in the validation report"*. Users who opt for a more flexible service argue that a very transparent documentation combined with freedom for the users is the way forward, and that making all the decisions in the background could be detrimental: *"As there are still unresolved issues, providing a black box to the user could be hazardous"*.

Data sets requirements

All users consider the provision of Level 3 data with the highest priority, while Level 2 and 4 respectively come after (Figure 5).

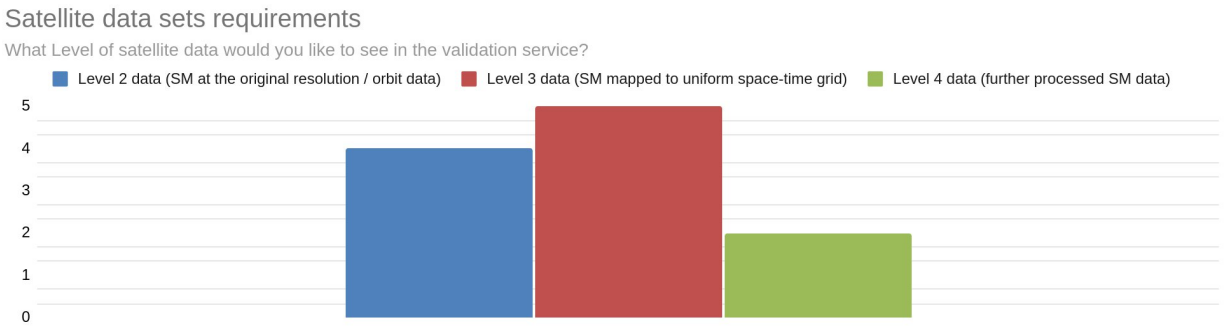

Figure 5: Feedback responses to question 16 (Appendix A).

Inquired about the integration of NRT products (question 17), none of the participants chose: 'Available NRT updates of data sets should be integrated directly and automatically in QA4SM to keep data sets up-to-date'. Instead, all think that either 'Delayed data updates are fine' (3) or that QA4SM should be made available to NRT producers through a programmatic interface (2). One user adds that: *"A lot of effort to make NRT for something that is done over a long temporal time span [...] does not make much sense"*.

Asked if they would like to see more sophisticated spatial re-sampling options (question 18), all answers consider the currently employed Nearest Neighbor (NN) technique sufficient. However, one participant points out the importance of providing the QA4SM users with information on the native resolution of the products.

QA4SM platform development

No participant sees the need for additional features (question 20) or user-defined metrics (question 21) in the service.

The 'stratification' of the validation data and results (question 22, Figure 6) is considered relevant for Leaf Area Index by most participants (4), followed by the topography data (3).

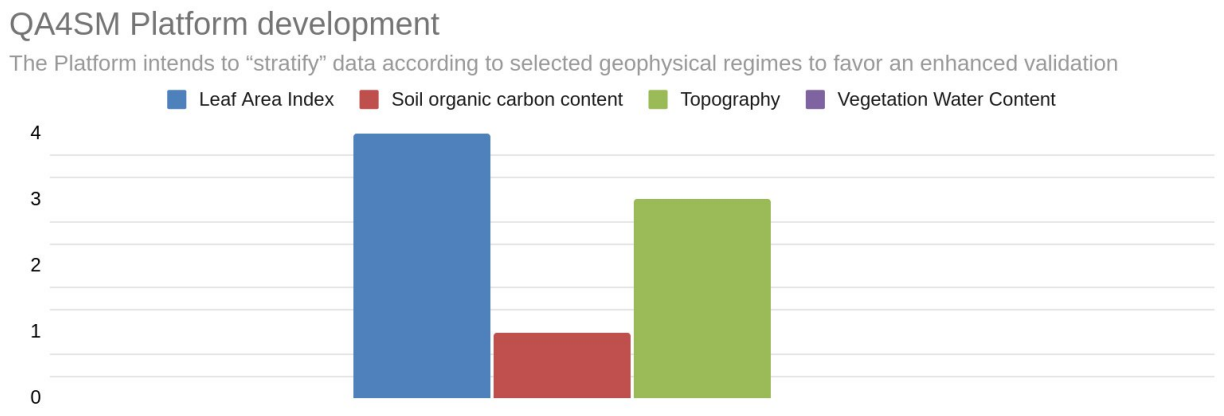

Figure 6: Feedback responses to question 22 (Appendix A).

The benefits of QA4SM with respect to off line validations (question 23) as mentioned by the participants are generally: 1) the availability of computational resources 2) the easy-to-use interface and availability of reports. Moreover, all participants think that the processing options offered by the service are complete (question 24) and the only suggestion made with respect to the graphical interface (question 25) is: *"A good tutorial and some extra explanation why certain choices are made so users become much more aware of the good practices underneath the hood of the validation tool"*.

3 out 4 of the participants replying to question 26 would find a programmatic interface preferable, indicating that this could be a more convenient way of integrating the platform with the participants' own research necessities.

2 (out of 4) users indicate that 'some data are missing' from the service (question 27). At question 29, the indicated missing data sets are: HSAF SM DAS, ESA CCI SM v7.1, MERRA2.

2 (out of 4) users indicate that it would be crucial to update their own data to the service (question 28). At question 30, the data sets that the users would like to upload to the service are: *'satellite data*, *downscaled*, *own retrievals*, *model results'* and *'test data on Sentinel-1*, *time series'*. The data size seen fitting for this purpose is 1-10 Gigabytes for all respondents (question 31); furthermore, these users are willing to reformat their data to netCDF (with CF formatting convention), csv or similar to use the service, and indicate that supporting a 3-dimensional (time, lat, lon) netcdf image stack will be sufficient for the purpose (questions 32, 33, 34). Finally, one of the respondents declares a strong aversion to permanent storage of proprietary data (question 35): *"option 1 will work*, *option 2 not. I would not use the service is this is the case with test data*". The other indicates: *"no license limitations. It is important to have a private version and the possibility to publish it*".

Additional remarks

"More important than NRT implementation is the automated (monthly?) update of validation reports or when new product versions are delivered. This is important to reach validation stage 4. Maybe in line with a dashboard where the current main metrics of all available products are presented."

"Not allowing users to upload their own data, but yes to creating a programmatic interface which people can apply to their local self stored data sets"

"Focus on quality, not on near-real-time. NRT has little added value for validation over time series."

Appendix A. FRM4SM Feedback form

1. User background and validation practices

This part of the survey helps us to understand how diverse the field of soil moisture validation really is.

At the bottom of this page you find a free text field, which can be used for any additional comments (on one of the previous question and in general).

1. Do you routinely perform validation of soil moisture data? Choose one of the following answers:
    - Yes, often
    - Yes, sometimes,
    - No, never
2. What is your focus in terms of Soil Moisture validation? Can you briefly describe your application in the comment field(s) on the right?
    - I mostly use it to compare the performance of different products to be used in my own studies (application)
    - I mostly use it in the development of my own product, to verify its strength and weaknesses and steer my research (development)
    - I am developing validation methods
3. Are you following any best practice document for Soil Moisture validation?
4. What kind of Soil Moisture data are you working with?
    - Satellite soil moisture data. Which?
    - Model / reanalysis soil moisture data. Which?
    - In situ soil moisture data. Which?

2. FRM4SM

This part of the survey deals with characterising in situ measurements as "fiducial". <u>You can skip this part of the survey if you have no experience with in situ data (for validation purpose).</u>

At the bottom of this page you find a free text field, which can be used for any additional comments (on one of the previous question and in general).

Fiducial Reference Measurements (FRMs):

The suite of independent ground measurements that provide the maximum return on investment for a satellite mission by delivering, to users, the required confidence in data products, in the form of independent validation results and satellite measurement uncertainty estimation, over the entire end-to-end duration of a satellite mission (FRM4SM, FRM4VEG)

5. Do you agree with the above definition of soil moisture FRMs? Do you have an alternative definition?
6. Are you familiar with the concept of ‘committed area’ in the validation of satellite data for soil moisture observations?
    - Yes, and I think committed areas should be considered in soil moisture validation.
    - Yes, but I think committed areas are not relevant for soil moisture validation.
    - I am not famliliar with with the concept of ‘committed area’.
7. Rank the following factors based on how important you consider them for in situ FRMs (drag and drop them from the left to the right side). Frommost important (top) to least important (bottom).
    - Noise / Uncertainty characteristics
    - Spatial representativeness (measuring technology)
    - Time series length, temporal coverage, temporal sampling
    - Presence of data gaps

- Scale mismatch
- Type of measurement device / sensor type, measuring technique
- Number of stations in a satellite pixel
- Sensor calibration
- Committed area

8. What is more important?
   - A large number of (rather globally distributed) in situ stations, with long time series but high uncertainties / low
   - data quality, respectively missing error information.
   - High-quality / error characterized stations in a small number of networks (but only regional).
9. Which quality indicators and data characteristics should at the very least an in situ FRM provide?
   - Fraction of gaps in the observations
   - Provision of uncertainty estimates on the individual data points
   - Sensor calibration values
   - Presence of measuring sites in the vicinity (network spatial density)
   - Flagging information for geophysical conditions (e.g. temperature below 0 deg)
   - Flagging information for irrigation events
   - Flagging information for time series quality (e.g. breaks, spikes, plateaus, …)
   - Ancillary information (precipitation, temperature, soil characteristics, …)
10. Should a FRM cover a min. number of years? How many and why?
11. What in situ data would you prefer for your application?
    - Many in situ stations within different classes (climate, soil, land-cover global distribution)
    - A large number of stations but in a small / homogeneous area
12. Should a FRM provide a specific temporal sampling frequency? Which and why?
13. Which other environmental factors do you take into account when using in situ soil moisture?
    - Land cover
    - Climate class
    - Soil type
    - Measuring depth
    - Organic content in soil
    - Sensor measuring technique
14. If you have any additional comments on any of the questions from this page (or in general), you can use this field:

3. QA4SM

*The following section focuses on **QA4SM**. This is an online validation service for soil moisture data products, which is publicly accessible for registered users and addresses the scientific soil moisture community. As such, it needs to be kept up to date with the latest scientific and technical standards which continuously evolve in the target community, both in general terms and in the specific FRM context (https://qa4sm.eu (http://qa4sm.eu)).*

At the bottom of this page you find a free text field, which can be used for any additional comments (on one of the previous question and in general).

15. At the moment QA4SM tries to implement validation best practices. This reduces the flexibility in terms of validation options in the service. What do you prefer?
    - A scientific platform where users can play around with different validation setting (where “wrong” settings can be chosen)
    - A “best practice” platform, where options are limited and where the platform maintainers make many decisions in the background.
16. What Level of satellite data would you like to see in the validation service?
    - Level 2 data (SM at the original resolution / orbit data)
    - Level 3 data (SM mapped to uniform space-time grid)

- Level 4 data (further processed SM data)

17. At the moment all datasets in QA4SM are sporadically updated by the platform maintainers. However, some datasets are produced in Near Real Time (NRT) and could be integrated/updated automatically in the service with a short latency. Which option(s) do you think make sense for an online validation platform (consider traceability of validation results)?
    - Available NRT updates of datasets should be integrated directly and automatically in QA4SM to keep datasets up-to-date.
    - The QA4SM functionality should be made available to the data producers (e.g. through a programmatic interface) to validate their NRT products by themselves.
    - Delayed data updates are fine for a validation service, the integration of NRT data is secondary.
18. Currently, all data sets in QA4SM are uploaded with their native resolution and collocated using a nearest-neighbor appraoch. Do you think this is sufficient, or would you like to see more sophisticated resampling options (e.g., bilinear interpolation). If yes, which ones?
19. How important is it to update operational data sets in the service on a regular basis (i.e. add new data from operational models / satellites)? From (1) not important to(5) very important.
20. Any other statistics/metrics/plots to be included/visualized in the service?
21. Would you like to define your own (validation) metrics in the service? Which metrics? How would you like to define them in the service?
22. The Platform intends to "stratify" data according to selected geophysical regimes to favor an enhanced validation. Any suggestions for additional regimes to be considered? Example regimes are listed below:
    - Leaf Area Index
    - Soil organic carbon content
    - Topography
    - Vegetation Water Content
23. What benefit do you see in an online validation platform compared to offline processing or the other way around?
24. Data pre-processing: Do you think the current options for data-preprocessing in the service are sufficient (flagging/filtering, scaling, anomaly computation)? Do you think these options are too limited? Any suggestions for additional preprocessing tools?
25. Do you think the graphical interface is intuitive enough? Do you miss any options? Can you suggest changes?
26. Currently validation settings, results management, etc. are all handled directly on the QA4SM website through a graphical user interface. Would a (programmatic) interface (e.g. via command line tools, python package etc.) as an alternative way to control the service make sense in your opinion?
    - Yes, a programmatic interface would be preferred.
    - No, a graphical interface is fine.
27. How would you rate the data selection in the service at the moment?
    - All necessary data are in the service
    - Some data are missing
    - Many data are missing
    - No relevant data are in the service
28. Would it be crucial to upload your own data to the service?
    - Yes, I would like to evaluate my own data.
    - No, I would leave integrating data to the service developers.

This part of the survey refers to answers that were given by you on the previous page.

At the bottom of this page you find a free text field, which can be used for any additional comments (on one of the previous question and in general).

29. You said on the previous page that not all relevant data are in the service. Can you name which ones you are missing?
30. You said on the previous page that you would like to upload your own data. What kind of data would you upload (satellite, model, time series, images, …)? What is the size of the data you would want to upload to validate? Please choose all that apply:
    - A few Megabytes (MB)
    - 1-10 Gigabytes (GB)
    - 10-100 Gigabytes
    - 100-1024 Gigabytes
    - >1 Terabyte (TB)
    - Other:
31. Would you be willing to bring your data into a format that the validation service can handle before uploading it? Please choose only one of the following:
    - Yes, I would reformat my data in order to validate it through QA4SM
    - No, I would only use QA4SM if it can handle my data directly
32. What format (csv, netcdf, …) is the data you would upload to the service in?
33. Are you following any data standards at the moment (e.g. CF standard for netcdf files) when producing data you would potentially upload to the validation service? Which ones?
34. Would supporting a 3-dimensional (time, lat, lon) netCDF image stack be sufficient for uploading your SM data to the serice? Considering e.g. irregular time stamps or irregularly gridded data. If not, how else do you structure your data?
35. Would data license constraints prevent you from uploading your data to the service? Consider two options to handle your data: 1) private upload with limited traceability/reproducibility to the service (other users don't have access to your data) 2) permanent storage of your data in the service to make validations traceable/reproducible (by you and other users)
36. You considered updating operational data sets in the service on a regular basis rather important on the previous page. What delay is acceptable from the moment the data is distributed by the producers until when it is available in QA4SM?
    - A few days
    - A few months
37. Do you know any existing (open) platforms/services that could be connected to QA4SM to manage (standardized) user uploaded data (instead of uploading your data directly into QA4SM)?