Water - ForCE

# Outcomes of the Expert Workshop on *in situ* calibration and validation of satellite products of water quality and hydrology

# Water-ForCE

| Project Identification | |
|---|---|
| **Project Full Title** | Water scenarios for Copernicus Exploitation |
| **Project Acronym** | Water-ForCE |
| **Grant Agreement** | 101004186 |
| **Starting date** | 01.01.2021 |
| **Duration** | 36 months |

| Document Identification | |
|---|---|
| **Dissemination Level** | Public (PU) |
| **Work Package** | WP4 |
| **Lead Partner** | PML, 3edata, IGB, EMU, TU |
| **Cite as** | Simis, S., Walker, P., Ogashawara, I., Cillero, C. Laas, A., Reusen, I. (2021). Outcomes of the Expert Workshop on in situ calibration and validation of satellite products of water quality and hydrology. V1.0.[report]. doi: 10.5281/zenodo.5789232 |
| **License** | Creative Commons Attributions 4.0 International |

| History of Changes | | |
|---|---|---|
| **Date** | **Version** | **Comments** |
| 16.11.2021 | 0.1 | Internal review |
| 17.11.2021 | 1.0 | First public release |

# Table of Contents



Water-ForCE is a CSA that has received funding form the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 101004186.

# 1 Scope

## 1.1 Water-ForCE

Water-ForCE is a H2020 Coordination and Support Action focussed on analysing the strengths and weaknesses, opportunities and bottlenecks of the current provision of water-related services in the Copernicus Earth Observation (EO) programme. The project will result (by 2023) in a roadmap document based on wide community consultation.

## 1.2 Expert workshop

Part of the Water-ForCE effort is the analysis of synergies between research and monitoring communities operating *in situ* instruments and platforms to collect water quality and hydrological information, and those working on (satellite) EO. A working group of experts was formed in the first year of the project and invited to take part in a workshop to discuss, over the course of three sessions:

- Data availability, accessibility, and quality gaps
- Emerging technologies to address current gaps
- Data harmonization and sharing

All registered experts were also encouraged to respond to a pre-workshop survey, which has been separately published, with some results highlighted also in this document. The survey result can be found at doi.org/10.5281/zenodo.5119010[1]

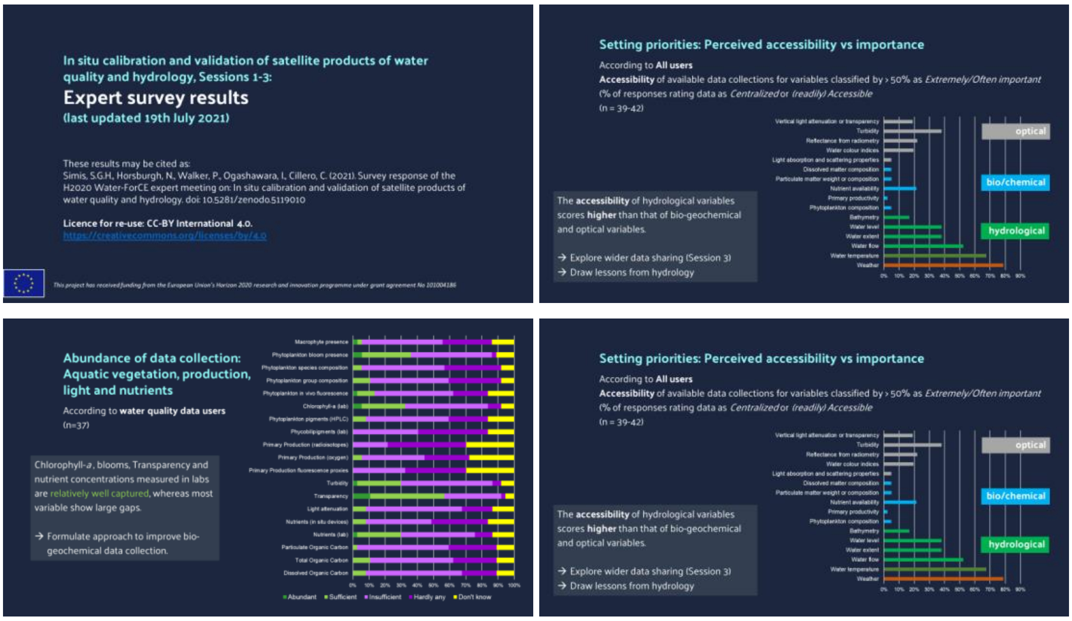


*Figure 1 Survey responses are available under open license as a slide set*

[1] Simis, S., Horsburgh, N., Walker, P., Ogashawara, I., Cillero, C. and Laas, A. (2021). Survey response of the H2020 Water-ForCE expert meeting on In situ calibration and validation of satellite products of water quality and hydrology (1.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.5119010

A record of the announcement, programme and outputs of the workshop, including contributions from invited speakers, is kept on the Water-ForCE project website: waterforce.eu/workshops/in-situ-calibration-and-validation-of-satellite-products-of-water-quality-and-hydrology.

## 1.3 Structure of this report

This report includes all recommendations collected during the workshop from invited speakers and breakout discussion groups in Appendix A. They have been filtered to combine closely related statements. All statements relevant to the discussion themes were kept for reference and their inclusion here does not imply community consensus. These recommendations were grouped by their relevance to common themes:

- Data gaps
- Data management
- Resourcing/Funding
- Coordination
- Solutions

The project team has since worked to analyse the community feedback and provides, in the next section, a summary view of recommendations which are **supported** but **not yet adopted**. This overview will be input to further drafting of the Roadmap, which will again be open to community feedback. In the meantime, we welcome any comments on this overview from the expert group through the project and respective work package leads.

# 2 Summary recommendations

This section lists the summary recommendations from the expert workgroup, organized per overarching objective of the Water-ForCE effort to coordinate *in situ* and remote sensing communities.

## 2.1 Towards a Standard of Practise for in situ observation networks, including high-frequency monitoring

### Observations from survey response

- 50% of the surveyed data users consider that precipitation, air temperature, wind speed and river flow/discharge have good availability. Other hydrological variables are perceived as gaps in data collection, including soil properties.
- Data collection of key variables, as water temperature, is considered insufficient by > 40% of the respondents.
- Phytoplankton bloom presence and Chl-a, as primary production descriptors, are considered to be well captured by >30% of the respondents. The rest of related variables showed even larger gaps.
- Light and nutrients data collection is considered insufficient by most of water quality data users, with only transparency perceived as well captured by >50% of respondents.

- The survey results highlight data gaps in particulate/dissolved matter, optical properties, colour and reflectance data collection.
- There was a strong agreement on the need of adding quality data descriptors in the metadata.

## Key variables (and gaps) to include in monitoring networks

Colour/bio-optical/physical variables

- Turbidity
- Secchi depth
- Chlorophyll-a
- Coloured Dissolved Organic Matter
- Total Suspended Solids / Total Suspended Matter
- Inherent Optical Properties (linking bio-optical models and water colour), including particulate backscatter
- Phycocyanin and Phycoerythrin (diagnostic pigments of cyanobacteria)

Biological/chemical indicators

- Aquatic Vegetation
- Primary Productivity
- Nutrients

Hydrology variables

- Flow data
- Soil moisture

### *Summary recommendations on key variables (and gaps)*

- Standardisation and cross-calibration activities (protocols) for *in situ* data collection are needed. Sensor calibration, sensor characterization (e.g. temperature dependence) need to be specified. Quantification of uncertainties, observational gaps, precision is needed.
- Add redundant sensors for cross-validation.
- Skilled people are needed to handle data collection.
- Establish more autonomous data collection.
- Coordinate time-series of hydrology data collection.
- Focus on global flow data (EU well covered, limited elsewhere)
- Capture soil moisture continuously *in situ*, including 3D nature

## Observation principles and techniques

Data collection principles to express key and desirable variables include:

- Hyperspectral data
- Vertical profiles
- Representation of different water types, locations
- Higher frequency and real-time measurements
- Phytoplankton biomass proxies include fluorescence / biovolume / pigment concentration / absorption, but they are not alike.
- Registration of sensor drift
- Pairing hyperspectral radiometers with IOP and atmospheric sensors, cameras, weather station.
- Sun glint corrections, bottom reflectance avoidance to be considered.
- Documentation on collection principles and algorithms in fluoroprobes.

- Aerosol characterization for atmospheric correction under different atmospheric conditions
- Work towards providing *in situ* data in near real time
- Transnational monitoring of (large) rivers
- Artificial intelligence approaches or methodologies (deep learning) could help to study shallow areas within lakes

Specific instrumentation needs:

- Underwater spectrophotometers which could measure in UV-Visible range
- CDT and oxygen measurements
- An adequate CDOM (absorption) or fDOM (fluorescence) sensor

### *Summary recommendations on observation principles and techniques*

- Prioritise collection and accessibility of data used for atmospheric correction validation (including nearshore), inherent optical properties, and vertical light profiles.
- Determine comparability of protocols/methods/sensors targeting the same substance through different (physical/chemical) means.

### Coordination and capacity building

- Promote lab/field campaigns with intercomparison opportunities
- Site selection:
    - Improve selection of water level monitoring sites to benefit EO cal/val
    - Develop strategic site selection criteria for hyperspectral devices.
    - Improve coordination of *in situ* measurements to fill observation gaps.
    - Inform site selection from current gaps (e.g. where ACIX-II shows gaps)
    - Exploit complementarity with existing platforms
- Establish wider contribution to cal/val:
    - Establish cal/val contributing activities within (national) monitoring context
    - Develop protocols for site operators to do local cal/val
    - train *in situ* monitoring community to facilitate data collection for use in satellite cal/val (i.e sample on clear days during satellite overpass)
    - Community-driven improvement of commercial sensor offering, e.g. with radiometers and  drones.
    - Provide QGIS tools that allow for cal/val using local data
- Training:
    - Training of students, globally
    - Seek funding for training programmes
- Community building and coordination:
    - Coordinate efforts with existing communities such as GEO Aquawatch
    - Holistic future observation for energy and water to be supported by WMO

### Cost considerations for R&D

- Low-cost instrumentation can have significant operational cost.
- A range of radiometry solutions is needed, from fiducial reference sensors to economic (incl. self-build) options.
- Consider low-cost versions of both platforms and sensors
- Low(er)-cost *in situ* absorption spectrometers need improvement to meet the usability of fluorescence probes, which would give way to broader use for algae pigment (chlorophyll-a) monitoring and absorption-based cal/val.

***Summary recommendations on Coordination, Capacity Building efforts and cost considerations***

- Promote sustained long-term funding to support data providers for network, infrastructure, data collection, Q&A and training.
- Focus on sustainability of funding for high quality measurements - e.g., use of Research Infrastructures.
- Build support for long-term continuous measurements in countries that may not be able to afford to do so.
- Forge direct cooperation between *in situ* and RS communities through common funding support.

## 2.2 Best practices for integration of *in situ* and satellite earth observation data

### Observations from survey response

- The current use of harmonized metadata is inadequate, there is a general lack of metadata and shared terminology
    - 78% of *in situ* data providers were interested in harmonising variable names with other providers. 43% of providers did not know if their variable names were defined in a published vocabulary.
    - 64% did not use (or were unaware they used) a common vocabulary service
- Improve data formatting: only 27% of Water Quality providers thought that the *in situ* data which they produced and/or used were generally formatted, structured and shared in line with the state of the art.
- Where data are being collected, sharing needs to improve
- Metadata should be searchable in geospatial context

### Interoperability requirements

- For any (incl. non-centralised) data repository, these must be interoperable
- Define common metadata requirements and promote machine-readable (self-describing) formats where feasible
    - All datasets should provide at least a minimum set of metadata (position, universal time, methods, parameter / variable / observed property).
    - Adopt common vocabularies, units
    - Define suitability or purpose of observation data to contribute to calibration versus validation, as these may have different requirements.
    - Include traceability of observations, methodologies, calibrations.

### Accessibility requirements

- Improve timeliness, harmonization of formats and openness of data collected by environment agencies
- Ensure the continuity of databases and associated metadata
- Improve cataloguing and centralized access of data stores between multiple organisations
- A central body to coordinate EU-based data repositories.
- Common infrastructures for data sharing, possibly by repurposing global data centres for satellite EO
- Provide data producers common interfaces to provide and update data contributions

***Summary recommendations on network interoperability***

- Create a Standard of Practise for metadata collection and data storage
    - Develop and adopt a controlled vocabulary, informed by existing services (NERC vocabulary services, GLEON/NETLAKE common vocabulary, CF conventions, EDMO codes, and ORCID for persons).
    - Promote FAIR principles.
- Train individuals to reach the required data standard, particularly small research groups with limited expertise.
- Develop tools to help centralize: 'upload API' & forms/templates, QA/QC
- Provide platforms to explore and query *in situ* and satellite observation data together.
- Pursue the same level of data integration of inland waters as achieved for marine (e.g. SeaDataNet with 100 National Oceanographic Data Centres, 34 European coastal states; physical, chemical, biological, geology, to geophysics and bathymetry)
- Clarify the respective roles of EEA, Copernicus *In Situ*, Copernicus 'Water' as coordinating bodies for i*n situ* observation data.

## 2.3 Advanced data sharing for wider use and impact

### Promote transparent licensing options

- Only 37% of Water Quality data providers included the licence or a reference to a licence in the metadata.
- Provide data purpose (license, sharing) at time of collection (or before). For *in situ* sensor data providers: build the metadata addition into the data flow.

### FAIR principles

- Data / repositories should be machine searchable (OGC CSW)
- Discourage wild-growth of data hubs, and evidence their value with usage reports.
- Include data purpose (license, sharing) as key search criterion
- Make observation uncertainty traceable through calibration records, protocols.

### Attribution, incentives and rewards:

- Explore data commons, non-monetary compensation to incentivize sharing.
- Acknowledge effort and time to deal with data requests, comments, queries.
- Incentivise countries to provide data / what is the value proposition?
- Linking to existing large scale observation networks (flux towers, meteo).

***Summary recommendations on FAIR , attribution, transparency***

- Develop community principles / guidance on data sharing, promoting FAIR principles. Work with existing communities such as GEO Aquawatch.
- Identify organisations who should/could be responsible for coordinating and funding for data sharing/storing/support.

## 2.4 Emerging monitoring methods

### Observations from survey response

- 67% of respondents to the survey were 'definitely willing' to use emerging technologies to address current data gaps.
- The main barriers were related to quality control, accuracy and cost.

- Reflectance sensors, autonomous monitoring stations, UAVs (drones) with commercial sensors and Citizen Science official programs were perceived by >50% of the respondents as technologies adding value to satellite products cal/val.
- The use of UAVs (drones) remains under-explored for cal/val activities.

### Barriers to uptake

- Complementary use of remote sensing is not always clear due to lacking core monitoring variables (e.g. nutrients/pollutants)
- Acceptance of alternative non-statutory methodologies by monitoring agencies is difficult

### Opportunities

- Assess the potential of using Citizen Science
- Integrate citizen science data into common networks
- Remotely piloted surface vessels – R&D needed for satellite cal/val, opportunity to assess bathymetry
- Fixed instrumentation on beaches for proximal sensing - not limited by clouds.
- Assess opportunities for Unpiloted Aerial Vehicles to carry payload (optical, thermal to bridge observations scales incl. for soil moisture spatial variability.
- Low-cost cameras for level / flow measurement, e.g. *openrivercam*
- Integration of fiducial reference sensors and low-cost alternatives – R&D need.

### *Summary recommendations on emerging monitoring methods*

- Increase acceptance of alternative technologies and data sources for monitoring by
    - Raising awareness
    - Capacity building (National/European Agencies/DGs)
    - Expanding the science base (publications)
- Support citizen science products; narrow down on uncertainties and build on best practises (e.g. Eye on Water Australia)
- Establish target and threshold requirements similar to observation requirements for Essential Climate Variables (GCOS, CEOS)

# Appendix 1: all workshop outcomes

This appendix includes all categorized output recorded during the expert workshop, including recommendations from invited speakers who have agreed for their contributions to be identified. All other inputs are anonymous.

Discussion outcomes are broadly categorised into four categories:

**Data gaps** - Including lack of observation or sharing thereof.

**Data management** - Issues leading to a lack of value-adding activities, such as data dissemination, harmonization, and adequate description (metadata)

**Resourcing/Funding** - Issues and solutions related to funding, or lack of clarity on funding responsibilities, between actors with the potential to contribute to data collection and sharing. Further, lack of trained staff, equipment, or other resources.

**Coordination** - Issues, barriers and potential solutions related to funding, sharing and knowledge exchange, or clarity over responsibilities, between actors with the potential to contribute to data collection and sharing. Further, lack of trained staff, equipment, facilities, or research/resource frameworks e.g. for cross-validation of methodologies, setting a common agenda, a coordinating body.

**Solutions** – any validated solutions ready to be implemented (but may need funding).

These have been further divided into identification of improvements needed, current barriers, and recommendations for improvement to specific stakeholders or communities.

## Session 1: Data availability, accessibility, and quality gaps

### Session overview

This session took place on 17 May 2021 13:00-16:00 CET and was attended by 58 invited experts (counted at 15 minutes past start), with 23 F / 35 M.

The session opened with a Water-ForCE project overview (Tiit Kutser) followed by two invited contributions from Copernicus *In Situ*, on current practises and needs in hydrology (Matthew Fry, CEH) and water quality (Kerstin Stelzer, Ana Ruescas, Brockmann Consult). The session continued with an overview and initial analysis of the pre-workshop expert survey, a networking break and a discussion session organized in break-out groups focussed on either hydrology, water quality, or data quality, and where each group was asked to consider (1) improvement needed (based on available options), (2) known barriers, and (3) recommendations to specific stakeholders or communities.

## Contributions from invited speakers

The presentations in PDF format are available from the workshop page on the Water-ForCE website.

**Copernicus *in situ* – hydrology** (Matthew Fry) observed:

- River flows – there are still data gaps e.g. in Poland but EFAS is a large data resource. Data sharing happening through JRC, still some issues to work out.
- Global data are established in the WMO GRDC database but not focussed on EO requirements. Also looking at National Hydro services.
- River & (some) lake levels. Note that data collection is mostly manual activity. Recommend making use of HDCC for large European rivers. There are some national resources (US, Aus, Argentina) online.

Future activities include: requirements analysis continues, CIS2 database under development (Water-ForCE could feed into it or benefit?), work with services notably CEMS, looking into harmonized data licensing -> recommend an open license?

**Copernicus *in situ* – water quality** (Kerstin Stelzer, Ana Ruescas) observed:

- Looked at CLMS requirements, MONOCLE survey, existing data sources, to inform gap analysis.
- Multiple angles for data need: algo. dev, calibration, product validation
- Requirements include: good match up window, either by synchronizing or continuous measurements; location: shallow water issues, away from land but also characterize this effect; different water types, particularly off-shore. Data descriptions: methodology used, multiple pigments requested.
- Recommend close organisation into and between data centres, room for improvement.
- Quality control; harmonization (e.g. units); common metadata (incl. contact details)
- Clear access policy needed
- Availability for both algorithm developers and calibration/validation

## Recommendations from breakout discussions: Hydrology

### Improvements needed

**Data gaps**

- Flow data typically used for modelling – EU well covered, limited elsewhere.
- Water levels – products in existence but data lacking for validation (satellite altimetry) - for large rivers globally.

- Measured locations are often outside the altimeters footprint leading to fewer match-ups.

Data management

- More organised/available metadata needed.
- Some hydro data only exist at unsuitable temporal and/or spatial resolution (often aggregated).

### Known Barriers

Resourcing/Funding

- Keeping time series of hydro data going, sensor and data maintenance, calibration, making data available. Sustainable funding is a global issue.

Coordination

- Security issues associated with some of the data (is this an increasing issue?)
- Rivers are often transnational, monitoring is not.
- Data spread across multiple organisations, even in single countries – difficult to collect and share.

### Recommendations

Coordination

- Can WMO help feed (historical and real-time) data into the right places?
- Incentivise countries to provide data / what is the value proposition?
- Target big rivers globally
- Holistic future observation for energy and water supported by WMO
- Linking to existing large scale observation networks (flux towers, meteo)
- Do we know what the Earth Observation community really need? Consider energy budgets, climate.

## Recommendations from breakout discussions: Water quality (multiple discussion groups including focus group on data quality)

### Improvements Needed

Data gaps

**Key variables highlighted (in no particular order):**

- Phycocyanin and Phycoerythrin. Data collection not standardized.
- Inherent Optical Properties
- Turbidity
- Coloured Dissolved Organic Matter
- Chlorophyll-*a*

- Total Suspended Solids / Total Suspended Matter / Suspended Particulate Matter
- Secchi depth
- Aquatic Vegetation
- Primary Productivity
- Nutrients
- Optical backscattering

**Observation principles and techniques highlighted:**

- Hyperspectral data
- Vertical profiles
- Representation of different water types, locations
- Higher frequency and  real-time measurements
- Phytoplankton biomass proxies include fluorescence / biovolume / pigment concentration / absorption, but they are not alike.
- Measurement / calibration standards need to be addressed
- Quantification of uncertainties, observation gaps, precision
- Registration of sensor drift

**Data management**

- Atmospheric correction data need to be harmonized (protocols, data formats).
- Transparency of algorithms used in different fluoroprobes.
- Transparency in quality control procedures applied to shared data
- Records of sampling depth and universal time
- Metadata should be searchable. E.g. sampling location is crucial
- Sensor calibration, sensor characterization (e.g. temperature dependence)

**Coordination**

- Where data are being collected, sharing needs to be improved
- Standardisation and cross-calibration activities for *in situ* data collection needed

### Known Barriers

**Data management**

- Need common definitions of data uncertainty; although mathematically simple, it has different components and different interpretation among fields.
- Lack of shared terminology, standardized metadata
- Environment agencies collect and store data but: time delays to availability, no common centralized exchange between agencies, openness not guaranteed

**Resourcing/Funding**

- Lacking capacity building on sampling, protocols for EO cal/val
- Lacking funding for spectroradiometers
- Calibration of sensors to maintain sufficient data quality
- Effort by data providers not always sustainable due to lack of funding for cal/val
- Continuity of databases (Long term data storage) and the associated metadata.

Coordination

**Uptake/acceptance:**

- Remote sensing is sometimes not considered part of a monitoring solution due to lacking core variables (e.g. nutrients/pollutants)
- Acceptance of new methodologies by monitoring agencies is difficult

**Leadership:**

- No single organisation responsible (eg Copernicus *in situ*, EEA, ESA) = no proper funding and data continuity

## Recommendations

### Data gaps

- Prioritise collection and accessibility of data used for atmospheric correction validation (including nearshore), inherent optical properties, and vertical light profiles.
- Assess the potential of using Citizen Science
- Determine comparability of protocols/methods/sensors targeting the same substance through different (physical/chemical) means.
- Promote large field campaigns with intercomparison opportunities

### Data management

- Define the role of EEA, Copernicus *In Situ*, Copernicus 'Water', as centralised body for EU to store and make data available.
- For any (incl. non-centralised) data repository, these must be interoperable
- Define common metadata requirements and promote machine-readable (self-describing) formats
- Recommend metadata on suitability of observation data to contribute to calibration versus validation, as these may have different requirements.
- Merge horizontal and vertical data with machine learning approaches, data assimilation techniques

### Resourcing/Funding

- Establish cal/val activities within monitoring context, not just on research basis
- Promote sustained funding to support data providers.

Coordination

- Develop community principles / guidance on data sharing, promoting FAIR principles
- Recommend *in situ* protocols for satellite cal/val (where already available, disseminate them to the appropriate community). Work with existing communities such as GEO Aquawatch.
- Encourage cooperation to have field campaigns whilst satellite is overpassing.
- Inform / train monitoring community to facilitate data collection for use in satellite cal/val (i.e sample on clear days during satellite overpass)

## Session 2: Emerging technologies to address current gaps

### Session overview

This session took place on 18 May 2021 13:00-16:00 CET. The session included five invited talks on ongoing projects that either make use of, or are developing, new technologies in support of satellite product calibration and validation of water quality or hydrology. The session continued with an overview from the pre-workshop expert survey on technological gaps and priorities followed by break-out group discussion focussed on four topics:

- What technologies should be used in existing moored platforms to support satellite cal/val. Specifically, what type of sensors are needed (incl. cost, performance)?
- Which alternative technologies (citizen science, UAV, smartphone apps) should (or can) be adopted for satellite cal/val?
- What should be the outcome of further R&D in terms of technologies? What should be developed in 5 years' time?
- Recommendations for capacity building and funding (i.e. to install new equipment on existing platforms)

### Contributions from invited speakers

The presentations in PDF format are available from the workshop page on the Water-ForCE website.

**Advancing aquatic science and EO cal/val using optical measurements by an automated profiler** (Daniel Odermatt et al., EAWAG/EPFL).
Four prominent challenges and corresponding best-practise were highlighted:

- Uncertainties in satellite derived reflectance: this was tackled in this study by including an independent modelled reflectance to attribute uncertainty metrics to the EO derived quantity.

- Presence of vertical gradients: an autonomous profiler (Thetis) including inherent optical properties observation was used to characterize the vertical dimension.
- Cal/val activities vs aquatic science interests: the added value of serving both interests by selecting appropriate observation variables was demonstrated.
- Lake morphology – exploit deep locations to minimise fouling, but this is not easily achieved with (above) surface sensors.

Further recommendations include networking between Thetis operators, sharing QC and operational protocols, and securing long-term funding.

**Soil moisture ground-truth: Bridging the scale gap** (Nick van de Giesen, TU Delft)
The TAHMO network linked to the H2020-TWIGA project has set up 600+ soil moisture sensors in Africa, which can be used to validate SMAP and SMOS, AMSR-E and Sentinel-1 operations. The BLOSM neutron detector enables automation at low cost ($100). The funding scope to maintain the network is costed at €20,000 / yr.

**State of the art and future developments on remote sensing for water quantity estimation** (María José Escorihuela, IsardSAT)
The main challenge in validating altimetry for Water Level assessment has always been the narrow satellite track which can easily miss reference targets. Using fully-focussed SAR increases the spatial footprint while maintaining 0.5m resolution. On soil moisture, high resolution is achieved using a multi-sensor approach combined with thermal/optical data. The International Soil Moisture Network contains global, searchable *in situ* measurements.

**Radiometric measurements: requirement, solution offered and recommendations (HYPERNETS)** (Kevin Ruddick, RBINS)
The case is made to support hyperspectral *in situ* reflectance radiometers for fiducial cal/val of optical satellites. The ideal sensor system has a broad spectral range and continuous calibration features using a built-in light source. HYPERNETS develops this sensor and several prototypes are currently in place.

**Alternative technologies for calibration and validation of water quality EO (MONOCLE) –** (Stefan Simis, Plymouth Marine Laboratory)
MONOCLE focusses sensor development for water quality cal/val on lowering the cost of individual sensors and their maintenance. Solutions include automation (stationary (WISP-M) and shipborne (So-Rad) Reflectance sensors/systems), dedicated low-cost atmospheric transmission sensors which characterize the downwelling light field to improve Rrs (HSP-1), obtaining reflectance from UAV (with optional added payload) as well as smartphones equipped with a low-cost spectropolarimeter. Complementary observation technologies that can be operated by non-experts (incl. citizen

scientists) are developed. The solutions offered currently range between TRL4 and TRL9.

## Recommendations from breakout discussions: Hydrology

**What technologies should be used in existing moored platforms to support satellite cal/val. Specifically, what type of sensors are needed (incl. cost, performance)?**

Data gaps

- Soil moisture: need permanent *in situ* measurements, as well as campaigns. Capturing 3D nature of problem.

Data gaps Coordination

- Water level: accuracy of standard methods is suitable, but locations need to benefit EO cal/val.

Resourcing/Funding

- It is still often better to have “in-house” developed sensors than “black box”/commercial sensors due to flexibility / price.

**Which alternative technologies (citizen science, UAV, smartphone apps) should (or can) be adopted for satellite cal/val?**

- R&D need: Can results from lots of cheap sensors be combined with a smaller number of high-quality measurements to get the best of both worlds?

Solutions

- There is scope to use cheaper cameras for level / flow measurements, e.g. openrivercam https://openrivercam.readthedocs.io/en/latest/.
- Similarly, the BLOSM for soil moisture is ready for wider adoption.

**What should be the outcome of further R&D in terms of technologies? What should be developed in 5 years' time?**

- Need analysis of sensitivities - how to define the level of accuracy needed for a given output - could this lead to guidelines, or support from e.g., WMO. Also looking at different scales, for soil moisture, from point to field scale to satellite grid cell.
- Need better understanding of potential for UAV mounted sensors (thermal camera, etc.) to bridge scales for soil moisture spatial variability

Resourcing/Funding Coordination

- Need plans to sustain funding for *in situ* measurements. Also programmes to statistically assess where measurements are needed.

**Recommendations for capacity building and funding (i.e. to install new equipment on existing platforms)**

Data gaps Coordination

- Improve coordination of *in situ* measurements to fill observation gaps.
- Build support for long-term continuous measurements in countries that may not be able to afford to do so.

Data gaps Resourcing/Funding

- Seek support from WMO and others for guidance on use of technologies, and coordination and support to help make data available to international community and preservation.
- Focus on sustainability of funding for high quality measurements - e.g., use of Research Infrastructures.

Data management

- Need common infrastructure for data sharing. Repurpose global data centres for EO, allow data producers to provide and update data, providing benefits to data providers such as QA/QC services.

## Recommendations from breakout discussions: Water Quality (multiple groups)

**What technologies should be used in existing moored platforms to support satellite cal/val. Specifically, what type of sensors are needed (incl. cost, performance considerations)?**

These responses have been filtered to remove overlap with Session 1, Water Quality requirements.

Data gaps **Spectroradiometry**

Most groups focussed on including more (reflectance) spectroradiometry in *in situ* observation networks, to improve complementary value of *in situ* water quality measurements and satellite observations, reducing uncertainties in atmospheric correction. Multiple solutions exist from a range of providers (hyperSAS, PANTHYR, Iexpore, WISPstation, hydraspectra) although some are still prototypes; for non-static platforms the So-Rad and hyperSAS and Dalec can be considered suitable. The sensors should be networked, hyperspectral and continuous.

Further considerations included:

- Pairing the hyperspectral radiometers with IOP and atmospheric sensors, additional cameras, weather station.
- Site selection should be informed by current gaps (e.g. where ACIX-II paper shows gaps) and complementary with existing platforms (additional water quality variables already observed).

- Hyperspectral radiometers may be too expensive, covering large regions we need to have economic (incl. self-build) options too, if maintenance/labour is not a factor.
- Not all sites will directly benefit from adding on hyperspectral devices.
- Sun glint corrections, bottom reflectance avoidance to be considered.

**Other sensors and considerations** (insofar as not already covered in Session 1 – Requirements):

- Underwater spectrophotometers which could measure in UV-Visible range
- CDT and oxygen measurements.
- Address depth distributions using profilers.

**Data management**

- Add redundant sensors for cross-validation.
- User-friendly (robustness) and scalable solutions preferred.

**Resourcing/Funding**

- Skilled peopled needed to handle data

**Which alternative technologies (citizen science, UAV, smartphone apps) should (or can) be adopted for satellite cal/val?**

**UAV/Drones**

- Drones can be used if they have calibrated sensors -> transect measurements.
- Drones with RGB, multispectral, hyperspectral sensors and atmospheric correction for shallow hard to reach areas; US setting guidelines for drone operations/processing, cfr. MONOCLE; drone data for water hyacinth as a mask; drones for scum; drones for taking samples in remote locations.
- UAV autonomy is of interest but faces regulatory barriers

**Solutions**

- Identified opportunities include spectrometers on UAV for transect studies (adjacency effects), for smaller water bodies and shallow waters, micro-scale horizontal variability

**Other (semi)autonomous platforms**

- Remotely piloted "boats" can be used but development is needed for satellite cal/val.
- Fixed camera on beaches for near remote sensing - not limited by clouds.
- Identified opportunities: USV for water depth

**Citizen science**

- Smartphones/citizen science can be used for validation (algal bloom), also for shallow areas, qualitative validation.
- Motivation of citizens unknown, pre-education needed.

- Thus far citizen science data have not been widely used for cal/val, but inclusion can help public acceptance for sampling infrastructure. Demonstrations needed.
- Water colour (e.g. eye on water) is very promising. How to bring measurements initiatives (and data) together globally?

Solutions

- Identified opportunities: Secchi disk, QA/QC, validation for optically shallow waters (bathymetry algorithms, bottom).

**General considerations:**

Data management

- Emerging alternative technologies may be suitable for validation but less so for calibration, important to separate these.

Resourcing/Funding

- Methodologies needed to allow operators to do their own local cal/val & R&D
- Cheap instrumentation can still have large operational expense.
- Remotely piloted and autonomous instruments might become cheaper.

**What should be the outcome of further R&D in terms of technologies? What should be developed in 5 years' time?**

Data gaps

- Floating platforms (closer to the coast for shallow water)
- An adequate CDOM (absorption) or fDOM (fluorescence) sensor is needed.
- Better suited equipment for IOPs in general and more affordable backscattering measurement devices in particular.
- Aerosol characterization for atmospheric correction under different atmospheric conditions
- Using apps - could potentially expand available information (e.g iSPEX)

Data management

- Integrating data from different sensors (drones, smartphones, …). Metadata, interoperability.

Resourcing/Funding & cost

- Low-cost versions of platforms and sensors
- Low(er)-cost *in situ* absorption spectrometers (like ACS) should be improved to meet the usability of fluorescence probes, which would give way to broader use for algae pigment (chlorophyll-a) monitoring.
- Leverage human effort - citizen science (learn from other industries - medical).

Coordination

- Community-driven improvement of commercial sensor offering, e.g. radiometers, drones.
- Provide QGIS tools that allow for cal/val using local data.

**Other R&D needs:**

- Artificial intelligence approaches or methodologies (deep learning) could help to study shallow areas on lakes
- Simultaneous deployment of *in situ* plankton Imagers that analyse the images in near real-time using AI - this would create a complete system for early warning of e.g. harmful algal blooms - and intercalibration with *in situ* absorption.

**Recommendations for capacity building and funding (i.e. to install new equipment on existing platforms)**

**Resourcing/Funding** including training

- Long-term funding for network, infrastructure, data collection, Q&A, and maintenance needed
- Funding for citizen science programmes
- Funding for training programmes from e.g. Copernicus
- Training of students, globally
- Training for citizen science programmes, including young people

Coordination

- Provide standard procedures to maintain data quality (e.g., instrument care)
- Address complicated permissions around drone flights, e.g. research licensing
- Facilitate loans of expensive equipment
- Incentives for data providers to feed into global data programmes.

## Session 3: Data harmonization and sharing

### Session overview

This session took place on 20 May 2021 13:00-16:00 CET. The session included four invited talks from leading data sharing and harmonization, interoperability initiatives. This was followed by breakout group discussion focussed on four questions:

- Are existing network(s) fit for purpose, what should be improved? What practises should be adopted more widely?
- Which stakeholders can/should be activated to improve data accessibility
- Which actions should be taken to improve (meta)data harmonization
- Which actions should be taken to improve data sharing

## Contributions from invited speakers

The presentations in PDF format are available from the workshop page on the Water-ForCE

**Global Terrestrial Network of Water Resources Observation Infrastructures** (Stephan Dietrich, ICWRGC)

GTN-H represents a global effort to coordinate 12 datacentres covering terrestrial water observations, including water quantity and quality. Parts of it have been in operation since decades. It is a member of the UN Water family (WMO, UNESCO, FAO, UNEP), with Climate observations in support of GCOS and WMO. Sustainable Development Goals are another main driver.

However, observational gaps exist, both spatially and temporally. Timeliness of data availability is an issue which relates to the data approval mechanism. This can take years. Data include *in situ* and satellite derived observations.

There is a pressing need to improve data management and connect stakeholders across the value chain from producers to data centres.

Sustained funding is an issue both at local/regional network scale and global data centres.

Data access policies need to become more open, free and accessibility should be further improved. GCOS recommendations underline these aims.

Ongoing and planned activities are identified to address the major issues listed above.

***In situ* water quality data harmonization and sharing, challenges and opportunities** (Philipp Saile, GEMStat)

88 countries have now provided data to GEMStat, however there is lack of standardized reporting. Linked water quality vocabularies provide opportunity to improve, with some good examples (e.g. UK water quality archive).

Openness of the data is currently limited for some data sets. Expected to improve, albeit under multiple licenses.

Recommendations:

- Further develop and publish upper-level ontologies/vocabularies for water quality parameters, analytical methods, … Governance unclear

- Encourage data providers to share data as open as possible referencing international licenses (compatible), such as Creative Commons or Open Data Commons
- Further develop and agree upon standard protocols and formats for making data accessible. Use e.g. OGC data format standards (WaterML 2 WQ), OGC APIs (SOS, EDR).
- Develop tools to deal with complex standards

**Increasing FAIRness of marine data within ENVRI-FAIR** (Peter Thijsse, MARIS)
Focussing on the marine data landscape, it is evident that the data represent large value (1.4B€ per year on acquisition) and there are diverse data providers and aggregators active.
The FAIR principles do not require data alterations, just additional specifications at the interface level, to improve discoverability and access between machine interfaces. Nevertheless, data management (providing the appropriate metadata) at the source (data acquisition) is crucial.

ENVRI-FAIR (2019-2022) aims to connect ENVRIs to European Open Science Cloud (EOSC) by Upgrading search/discovery options for data and metadata, implementing APIs, Upgrading metadata using common vocabularies (CF conventions, EDMO codes, and ORCID for persons), upgrading provenance information (unique IDs for data and methods).

**Recommendations from OGC / E-shape *in situ* first findings** (Marie-Françoise Voidrot, OGCE)
*In situ* observations are crucial to the success of any global observing system, and crucial to the success of any postprocessing. There is greater value in integrated remotely sensed and *in situ* than in its individual parts. Historical datasets were often linked to a single issue, project or network. There is an ongoing rise in demand for *in situ* data.

Tools are being made increasingly available (pilots in e-shape) to create and manage FAIR *in situ* data. Essential to adopt the GEO Data Management Principles, use Open Standards whenever possible and implement good quality Metadata (including Credit, Provenance, licences…), and finally to test the compliance of the implementation.

## Recommendations from breakout discussions

**Which practises should be more widely adopted (and by whom)? Which are priority actions?**

Forming network(s) is key to interact with all data operators. The networks should:
- Create common capacity building programmes
- Consider ‘data stewards’ to push interoperability in their organisations

- Development of APIs
- Work to include statutory/national monitoring data
- Promote FAIR principles
- Promote GEO data management principles
- Promote transparent licensing options
- Promote timeliness of data sharing
- Work with sensor industry to adopt same standards
- Promote appropriate centralisation
- Offer a catalogue of repositories
- Provide tools to help data sharing, validation.
- Seek funding to improve metadata, also for historical data

Data centres include (this is not a comprehensive list):

- LIMNADES for water quality validation data
- National data centres (often without access, or only through portal)
- EEA WaterBase (limited metrics available)
- ILEC world lake database
- GEMS-Water - looks great, but we are unsure what type of QA/QC rules have been taken into account
- USGS and LAGOS aggregate water quality across the United States
- NETLAKE metadata databases

Network members should include:

- GEO Aquawatch

Common vocabularies to be informed by:

- NERC vocabulary services
- GLEON/NETLAKE common vocabulary

**Community data sharing principles to address current shortcomings**

Towards FAIR data:

- Data / repositories should be machine searchable (OGC CSW)
- Provide a minimum set of metadata: position, time, methods, parameter / variable / observed property.
- Work towards a complete set of metadata which allows traceability of observations and methodologies, calibrations.
- Provide data purpose (license, sharing) at time of collection (or before). For *in situ* sensor data providers: build metadata addition into the data flow.
- Adopt common vocabularies, units

Attribution, incentives and rewards:

- Explore data commons, non-monetary compensation principles to incentivize sharing. Need to give something back to the data providers (publications?)
- Acknowledgement to effort and time to deal with data requests, comments, queries.

Data quality management:

- The next big issue for data re-use is quality control, confidence, uncertainties. Needs to be built-in.
- Dedicated project funding to address final data QA/QC and harmonization.

**Recommendations to remote sensing and *in situ* data services / networks on how synergies can be improved**

- Train individuals to reach the required data standard, particularly small research groups with limited expertise.
- Develop tools to help centralize: ‘upload API’ & forms/templates, QA/QC, e.g. towards LIMNADES for optical/water quality data.
- Evidence where efforts are most clearly needed to get data moving, curated, structured, and FAIR.
- Provide common platforms to explore and query *in situ* and satellite observation data.
- Develop and adopt common vocabularies, address any conflicts.
- Work towards providing *in situ* data in near real time.
- Seek support from large actors (EEA, GEO, ESA, Copernicus programme)
- Pursue the same level of data integration of inland waters as achieved for marine (e.g. SeaDataNet with 100 National Oceanographic Data Centres, 34 European coastal states; physical, chemical, biological, geology, to geophysics and bathymetry)
- Integrate citizen science data into common networks.
- Discourage wild-growth of data hubs, and evidence their value with usage reports. This does not prevent organisations from using their own systems, too.