1.5 Connection between In-Situ Data and Algorithm Development

Robert Frouin
Scripps Institution of Oceanography
La Jolla, California

Contents

1. Role of In Situ Data in Algorithm Development:

•Importance in calibrating, validating, and refining algorithms in ocean color remote sensing.

•Ground truth references for satellite and airborne observations, enhancing precision and credibility.

2.Types of In Situ Data:

•Exploration of various types, including water quality measurements and bio-optical properties.

•Significance of diverse datasets for developing tailored algorithms.

3. Calibration and Validation of Ocean Color Remote Sensing Data:

•Definition and process of calibration to harmonize remote sensing measurements with ground truth.

•Validation to evaluate accuracy and reliability using in situ measurements.

•Utilization of in situ data to calibrate sensors, mitigate atmospheric interferences, and validate algorithms.

Contents (contd.)

4. Examples of In Situ Data Applications:

•Presentation of case studies demonstrating the use of in situ data in sensor calibration and algorithm development.

•Radiometric measurements validating satellite-derived products.

5. Challenges and Limitations in In-Situ Data Integration:

•Discussion on challenges including data quality assurance, spatiotemporal representativeness, and scalability.

•Emphasis on addressing challenges for robust and applicable algorithms across diverse marine environments.

6. Future Perspectives and Opportunities in Ocean Color Algorithm Development:

•Exploration of emerging technologies and methodologies like machine learning and data assimilation.

•Opportunities for interdisciplinary collaborations to enhance synergy between in situ measurements and algorithm development.

Importance of In Situ Data in Remote Sensing

1. Introduction to in situ data

•In situ data refers to direct measurements collected within the environment of interest, obtained through field surveys, instrument deployments, or sensor installations.

•Examples include water samples from lakes, rivers, or oceans, buoy observations of oceanographic parameters, and sensor readings from environmental monitoring stations.

2. Validation and Calibration

•In situ data serve as ground truth measurements for validating and calibrating remote sensing observations.

•Comparison of remote sensing data with in situ measurements allows assessment of accuracy, identification of errors or biases, and improvement of algorithm performance.

Importance of In Situ Data in Remote Sensing (contd.)

3. Enhancement of remote sensing products:

• In situ data provide detailed information about local environmental conditions and processes not captured by remote sensing alone.

• Integration of in situ data enhances the interpretation and utility of remote sensing products for scientific research, environmental monitoring, and decision-making applications.

Remote sensing and in situ data are complementary approaches for studying Earth's systems and monitoring environmental changes. Remote sensing provides synoptic views of Earth's surface, while in situ data offer ground truth observations that validate and enhance the accuracy of remote sensing products. Understanding the connection between remote sensing and in situ data is crucial for leveraging both approaches and advancing our knowledge of Earth's complex systems.

Role of In Situ Data in Algorithm Development

1. Fundamental Role of In Situ Data:

•In situ data are essential for calibrating, validating, and refining algorithms in remote sensing applications.

•Calibration adjusts remote sensing measurements to account for sensor characteristics and atmospheric effects, ensuring accuracy and consistency.

•Validation assesses the accuracy and reliability of remote sensing products using ground truth measurements.

•Refinement of algorithms involves adjusting parameters to improve accuracy, precision, and robustness.

Role of In-Situ Data in Algorithm Development (contd.)

2. Importance as Ground Truth:

•In-situ data provide direct, accurate, and reliable observations at specific locations and times.

•In-situ measurements validate satellite-derived products such as chlorophyll-a concentration and water clarity.

•Comparison of remote sensing observations with in-situ measurements validates accuracy, precision, and consistency.

•Identifies errors or biases, guiding algorithm improvements for different environmental conditions and applications.

In situ data are critical for algorithm development. They improve the accuracy, reliability, and utility of remote sensing products for scientific research, environmental monitoring, and decision-making. Understanding the connection between in situ data and algorithm development is essential for advancing our understanding of Earth's systems.

In Situ Data for Ocean Color Remote Sensing

1. Water Quality/Composition Measurements:

•Fundamental for validating satellite-derived ocean color products.

•Parameters include chlorophyll-a concentration, suspended sediment concentration, and CDOM absorption.

•Collected using fluorometers, spectrophotometers, and radiometers on research vessels, buoys, and autonomous platforms.

2. Radiometric Measurements:

•Provide information on water-leaving radiance and inherent optical properties.

•Essential for validating satellite-derived remote sensing reflectance and atmospheric correction algorithms.

•Collected using radiometers, spectroradiometers, and hyperspectral sensors

In Situ Data for Ocean Color Remote Sensing (contd.)

3. Buoy Observations:

•Offer continuous time-series data on oceanographic parameters.

•Include sea surface temperature, chlorophyll-a concentration, and optical properties.

•Collected using moored buoys equipped with sensors and instruments.

4 Validation Matchup Data:

•Involve collocating satellite overpasses with in situ measurements.

•Enable direct comparisons between satellite-derived data and in situ observations.

•Obtained through field campaigns, research cruises, and oceanographic surveys.

Calibration and Validation of Ocean Color Remote Sensing Data

1. Definition of Calibration in Ocean Color Remote Sensing:

•Process of adjusting satellite-derived radiance or reflectance measurements to match ground truth data from in situ measurements.

•Corrects systematic errors, biases, and inconsistencies in satellite observations.

•Ensures accuracy and reliability in quantifying oceanic parameters like chlorophyll-a concentration and water clarity.

2. Definition of Validation in Ocean Color Remote Sensing:

•Process of assessing accuracy and reliability of satellite-derived ocean color products using in situ data.

•Compares satellite-derived observations with independent ground truth measurements collected in situ.

•Provides confidence in the quality and credibility of satellite-derived ocean color products.

Calibration and Validation of Ocean Color Remote Sensing Data (contd.)

3. Use of In Situ Measurements for Calibration and Validation:

•Crucial for calibrating satellite sensors, correcting atmospheric effects, and validating algorithms.

•Provide reference data to refine and improve accuracy of satellite-derived ocean color products.

•Used to calibrate satellite sensors by comparing in situ and satellite radiance or reflectance measurements.

•Used to validate satellite-derived ocean color products by comparing satellite observations with in situ measurements of water properties.

•Matchup analyses enable direct comparisons between satellite and in situ observations, assessing data accuracy and reliability.

Calibration and Validation of Ocean Color Remote Sensing Data (contd.)

4. Illustrative Examples in Ocean Color Remote Sensing:

Example 1: In situ radiometric measurements used to calibrate satellite sensors by comparing in situ and satellite radiance values.

Example 2: In situ measurements of chlorophyll-a concentration, K_d, and POC used to validate satellite-derived products.

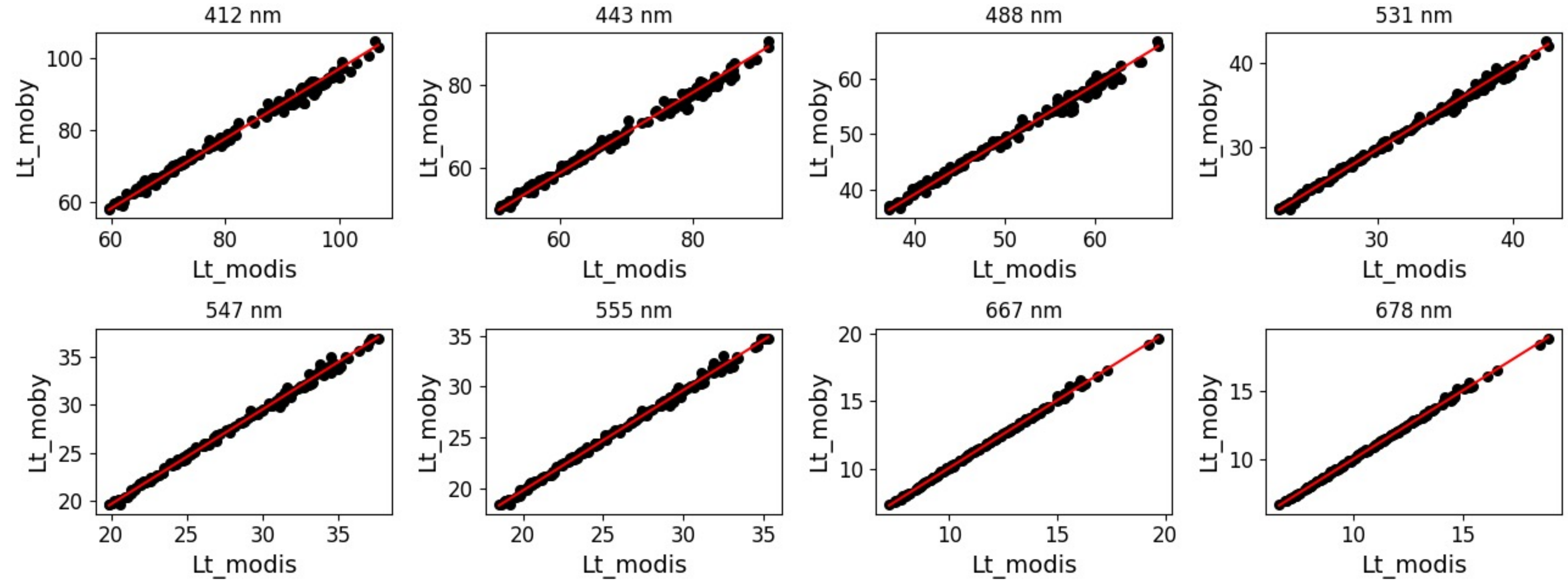
Example 3: In situ radiometric measurements used to evaluate water reflectance (fixed platforms: HYPERNETS)

Example 4: In situ radiometric measurements used to evaluate water reflectance (profiling system: HyperNav)

Example 5: APAR algorithm development

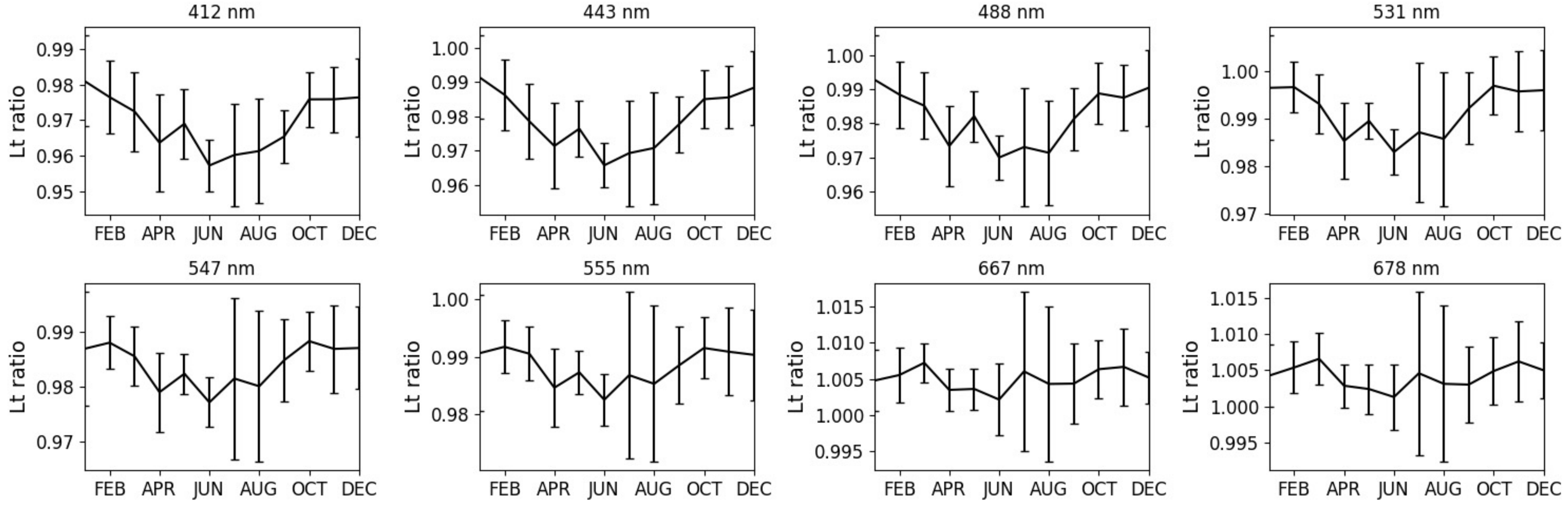
Calibration and validation using in situ measurements are essential for ensuring accuracy and reliability of satellite-derived ocean color products. Comparison with ground truth data enhances algorithm performance and corrects atmospheric effects. Crucial for scientific research, environmental monitoring, and marine ecosystem management.

Example 1: System Vicarious Calibration (SVC) using MOBY data



SVC gains were obtained by propagating Lw_moby to Lt_moby and then dividing Lt_moby over Lt_modis.

New gains (136 SVC match-ups):
0.97253 0.98251 0.98532 0.99317
0.98491 0.98904 1.00432 1.00367

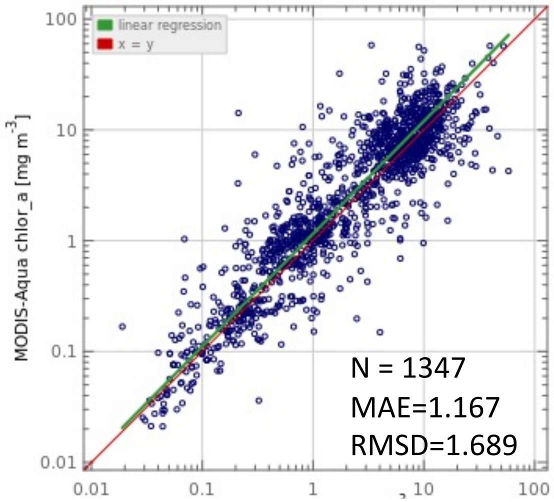


SVC gains exhibit seasonal changes, which yields seasonal bias in MODIS-derived water-leaving radiance.

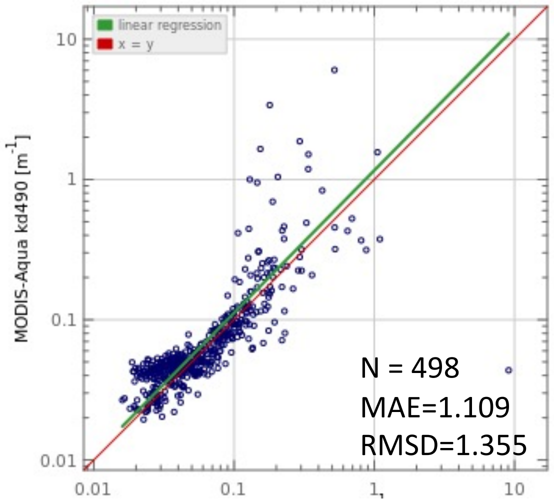
Origin of the seasonal dependence in SVC gains is still unknown.

Example 2: In-situ validation of NASA OBPG MODIS-A [Chl-a], K_d, and POC products

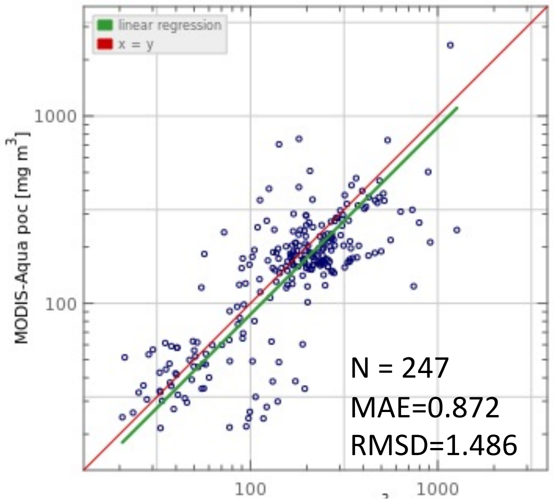
[Chl-a]

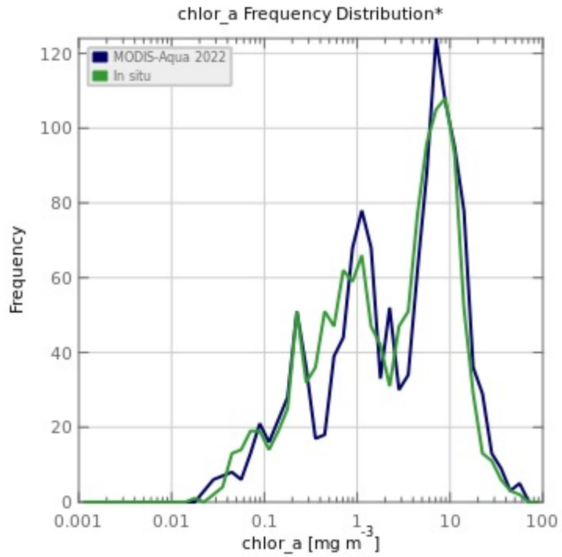


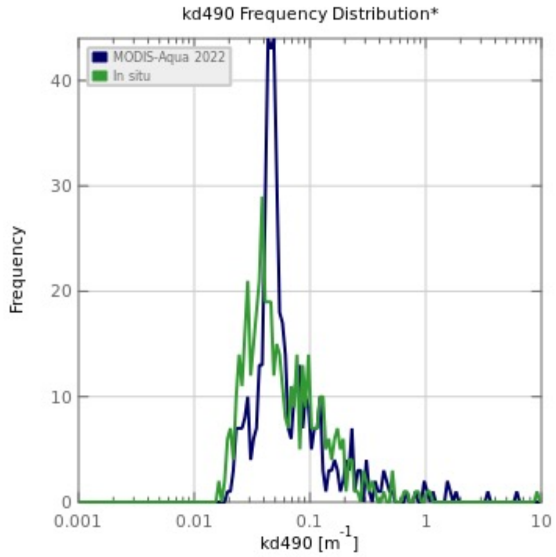
K_d

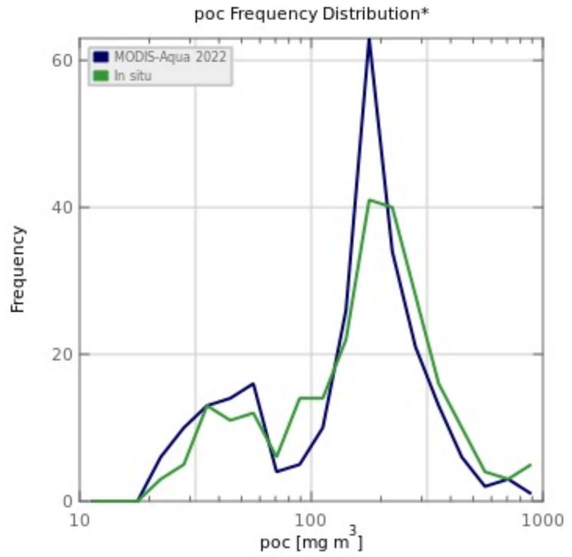


POC

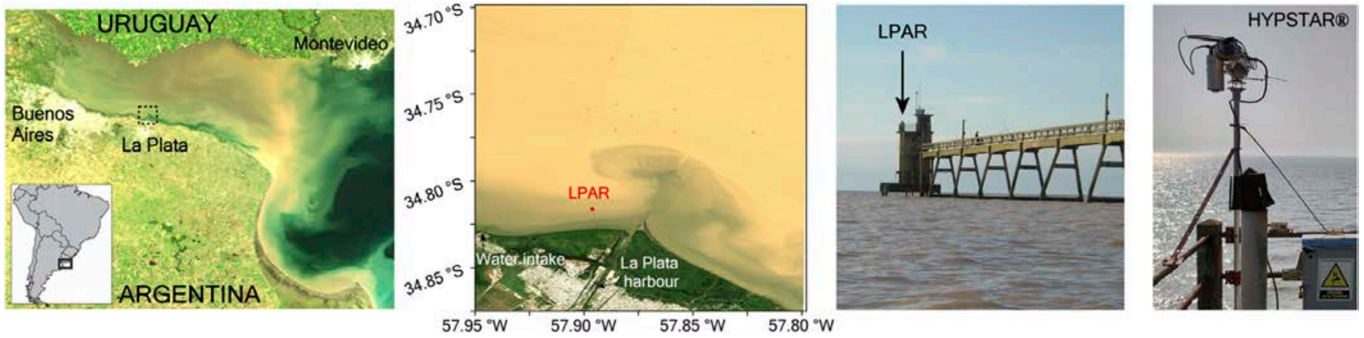


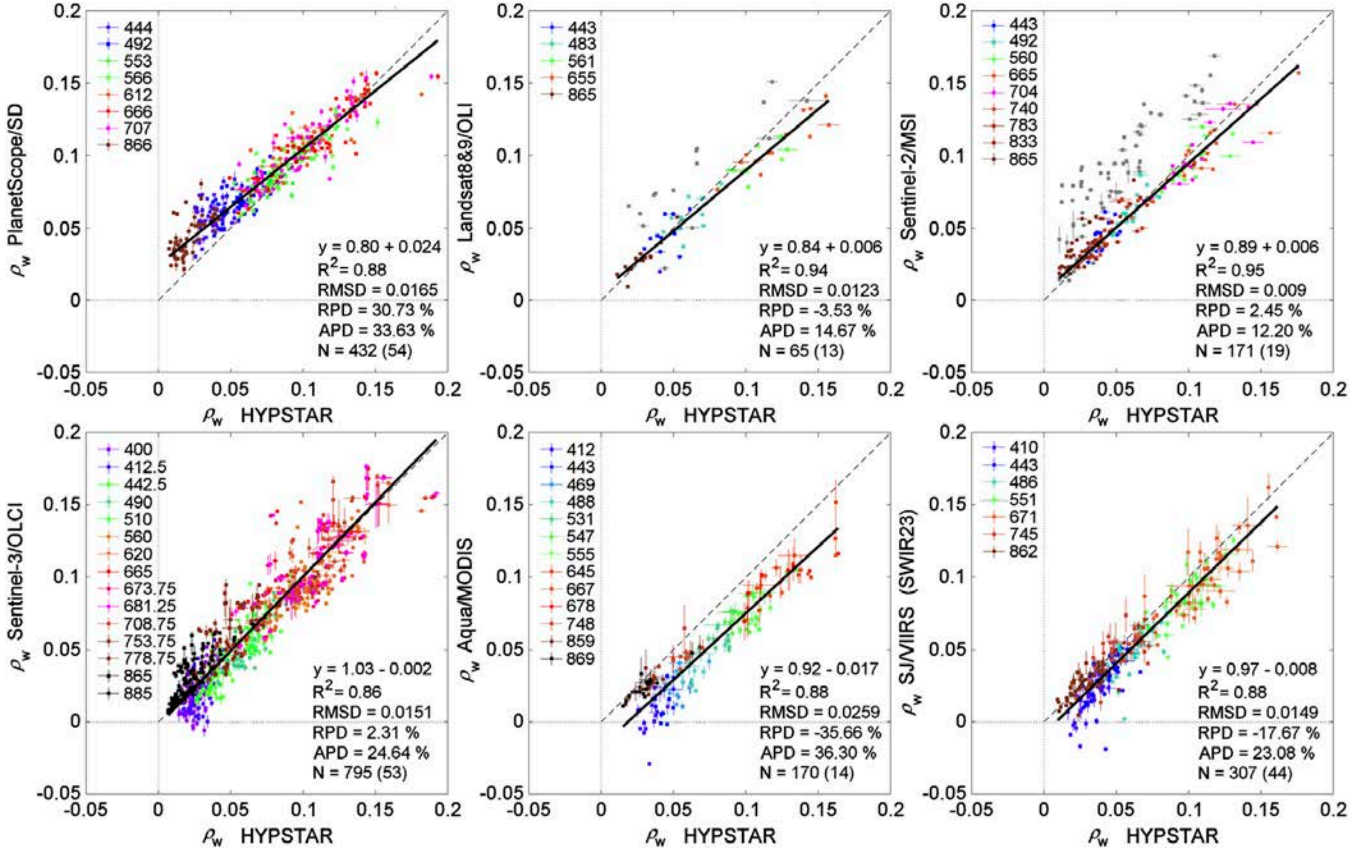


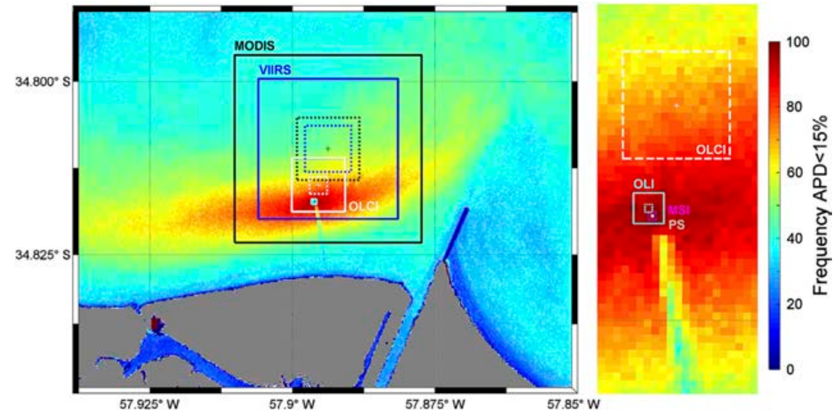


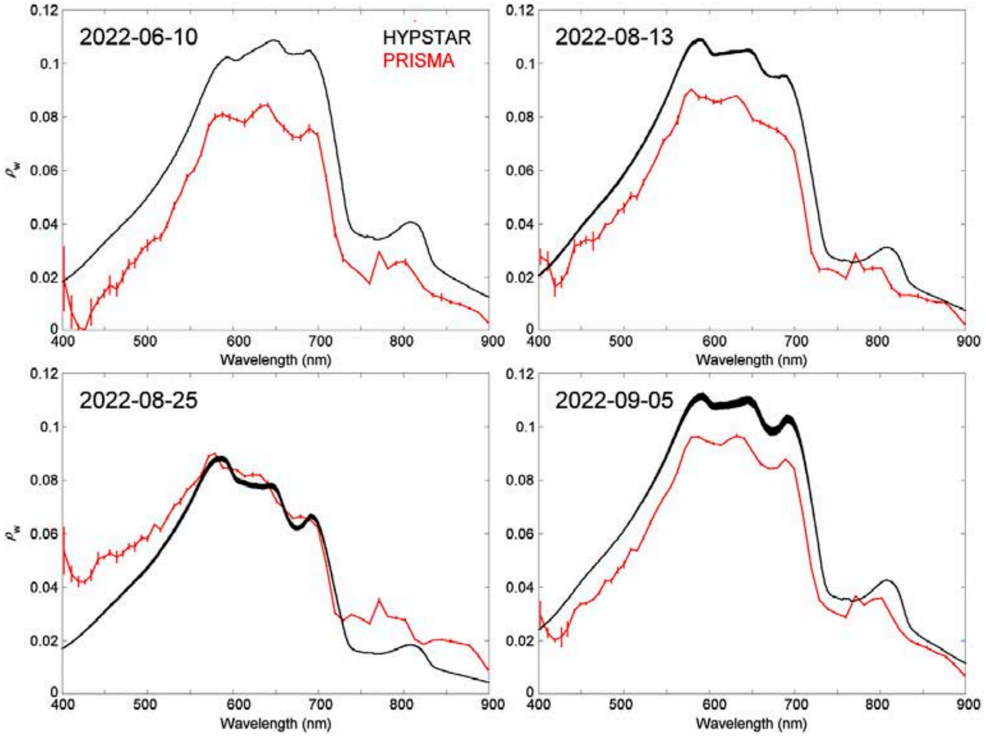


Example 3: Validation of water reflectance using HYPERNETS, Rio de la Plata









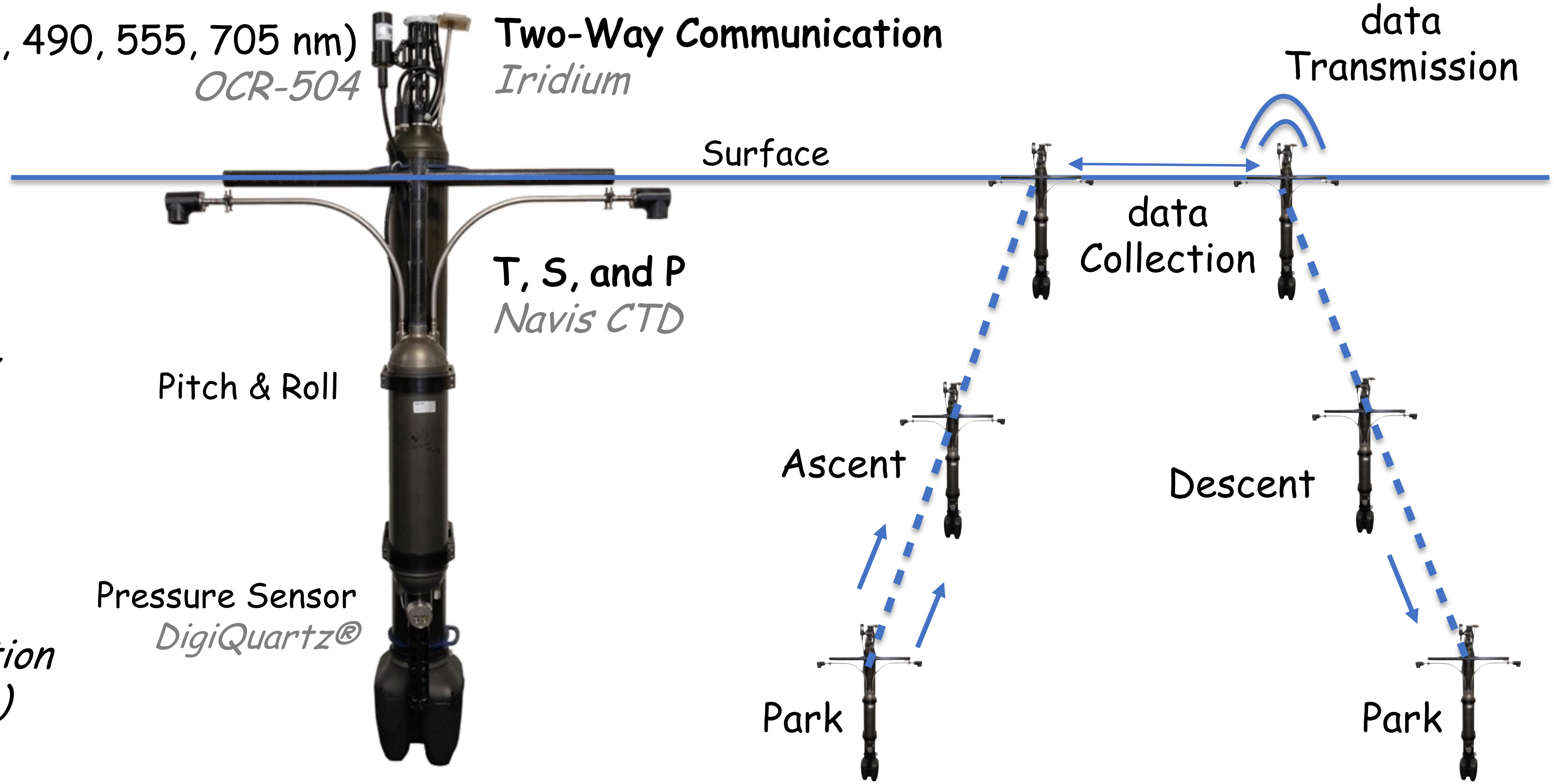
Dogliotti et al., 2024)

Example 4: Evaluation of SGLI water reflectance using HyperNav systems

-HyperNav is a spectroradiometer/float system developed by SeaBird, Inc. to support hyperspectral ocean color satellite missions, such as PACE.

Dual Upwelling Radiance (Lu)
Spectral Range: 350-780 nm
Channel spacing: 0.4 nm
Full Width Half-Max: ~2.2 nm
9° FOV
Integrated shutters for darks
Uncertainty: < 4% UV-G, 6 % R-NIR

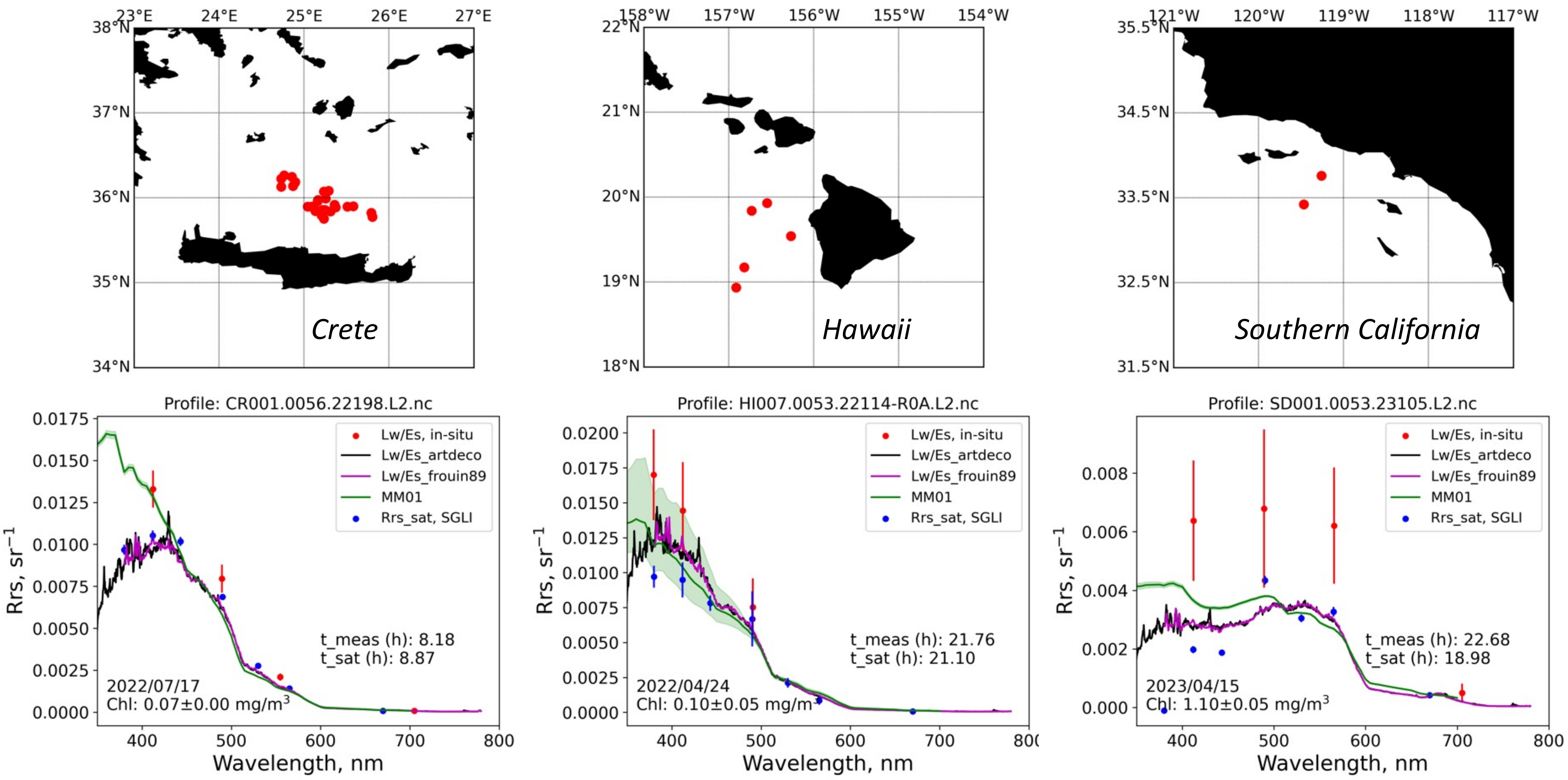
Advantages of using a float
Full profile (0.1 to 500 m)
Reduced errors in $L_U(0^-)$ extrapolation
Reduced bio-fouling (park at 500 m)
Portable (easy to deploy/recover)



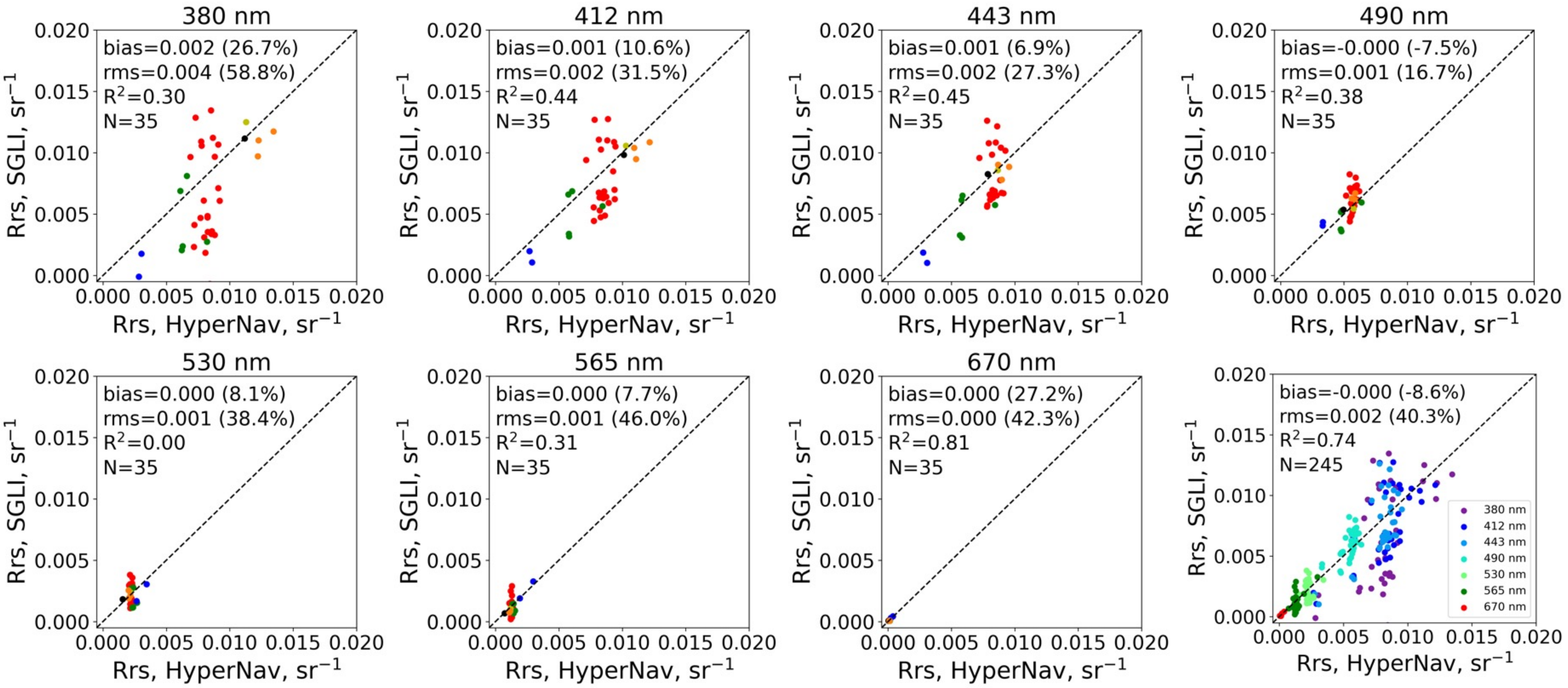
-3 sites tested (Crete, Hawaii, San Diego); 2 long deployments (58 and 69 profiles); 19 deployments to date (5% loss; HI004 - 0052); 166 HyperNav enabled profiles (2-year period).

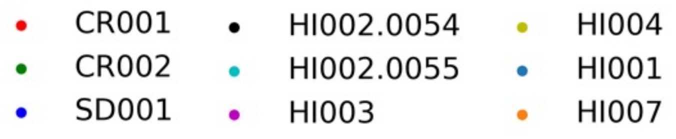
-Data were used to evaluate SGLI-derived R_{rs} (35 quality match-ups).

Example 4: Evaluation of SGLI water reflectance using HyperNav systems (contd.)



Example 4: Evaluation of SGLI water reflectance using HyperNav systems (contd.)





-Large variability in SLI R_{rs} in UV and blue during CR001 deployment, despite stable water conditions, with underestimation in many cases (possibly due to the presence of dust from North Africa).

Example 5: APAR modeling using environmental variables

-We used an ensemble of regular MLPs to model APAR from in-situ R_{rs} in all 6 SGLI spectral bands in the PAR spectral range (1266 situations) with and without environmental variables, i.e., Sun zenith angle, SST, daylength, and latitude. No noise was included in the data.

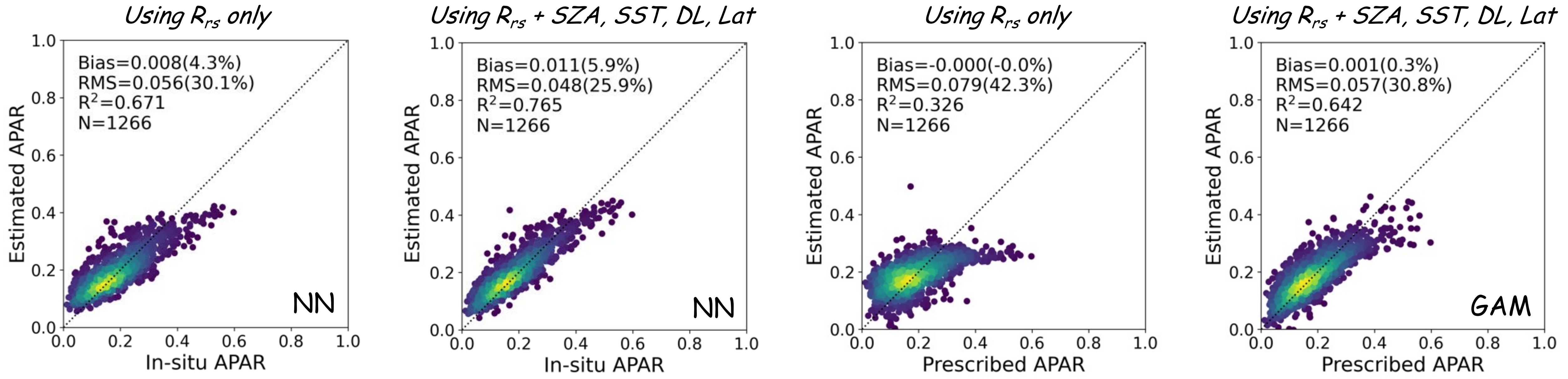
-We divided the training data into 5 partitions, use 4 partitions for training and 1 partition for evaluation, and rotate until all 5 partitions have been used as evaluation partition once. Thus, we can have 5 models. We do this 5 rounds to have a total of 25 models in the ensemble.

-We also used an ensemble of GAMs with coefficients dependent on the environmental variables. In this case, we did 10 random realizations and in each realization the data were split randomly (80% for training and 20% for validation).

-In the end, we used the ensemble of models together with their respective standard scalers to do inference on the full dataset. For each model we have a set of predictions, and the final prediction is the mean across all the models, associated with a standard deviation.

-The APAR uncertainty consists of two parts: (1) the model uncertainty e1, i.e., the standard deviation obtained using the model ensembles, and (2) the uncertainty between modeled and actual APAR. Final uncertainty was then calculated as the square root of e1^2+e2^2.

Example 5: APAR modeling using environmental variables (contd.)

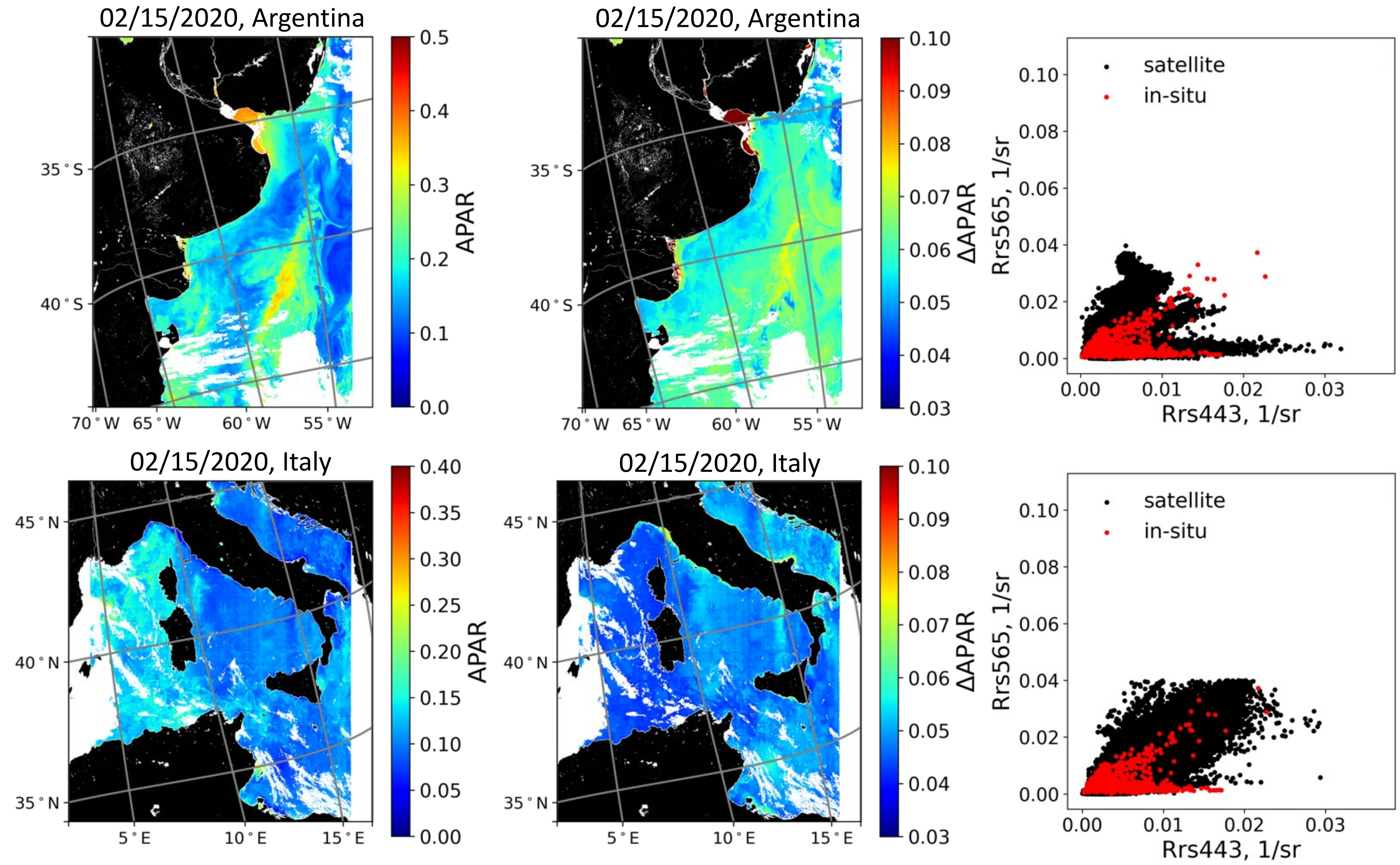


-Performance is significantly increased by adding environmental variables, with R^2 of 0.765 instead of 0.671 using NN ensemble and 0.642 instead 0.326 using GAM ensemble.

-Results are also much better than classifying R_{rs} spectra in multi-linear combinations (leading to R^2 of only 0.535), i.e., environmental variables might be better predictors than R_{rs} classes.

-Generalization may be difficult because theoretical data sets do not include environmental variables; Need for more (representative) in situ data.

Application of NN ensemble using Rrs + SZA, SST, DL, Lat to SGLI imagery



-APAR imagery does not exhibit obvious artifacts, despite different domains of R_{rs} variability for in-situ and satellite data, i.e., NN has some generalizing capability.

-Need to identify observations that do not fit model, i.e., to associate to each APAR estimate, not only uncertainty, but also information on whether data and model are compatible.

Challenges and Limitations in Ocean Color Remote Sensing

1. Spatial and Temporal Representativeness:

•In situ data may not adequately represent spatial and temporal variability of oceanic processes.

•Limited spatial coverage and temporal gaps in observations can lead to uncertainties in algorithm performance.

•Comprehensive coverage and continuous monitoring are essential for accurate algorithm development.

2. Data Quality Assurance:

•In situ data quality assurance is crucial for accurate algorithm development.

•Errors, biases, and inconsistencies in in situ measurements must be identified and mitigated.

•Quality control procedures such as sensor calibration and data validation are necessary.

Challenges and Limitations in Ocean Color Remote Sensing (contd.)

3. Scalability:

•Scaling up in situ data collection presents challenges due to limited resources and logistical constraints.

•Collaborative efforts and innovative technologies are needed to enhance scalability.

•Integration of diverse data sources and advances in sensor technology can improve scalability.

4. Importance of Addressing Challenges:

•Overcoming challenges associated with in situ data utilization is crucial for advancing remote sensing algorithms.

•Comprehensive spatial and temporal coverage, coupled with high-quality in situ data, enhances the accuracy and reliability of remote sensing products.

•Collaborative efforts and interdisciplinary approaches are essential for addressing these challenges and advancing ocean color remote sensing.

Future Perspectives and Opportunities

1. Emerging Technologies and Methodologies:

•Machine Learning:

-Neural networks and deep learning models can enhance satellite-derived ocean color products by learning from in situ measurements.

-Data-driven approaches improve algorithm development, calibration, and validation.

•Data Assimilation Techniques:

-Ensemble Kalman filters and particle filters integrate in situ data into remote sensing algorithms.

-Improves spatial and temporal resolution, enhances predictive capabilities, and reduces uncertainties.

•Crowdsourcing Initiatives:

-Citizen science projects augment observational networks and expand spatial coverage.

-Engages volunteers in data collection, enhancing diversity of observations.

Future Perspectives and Opportunities (contd.)

2. Interdisciplinary Collaborations

•Opportunities:

-Remote sensing scientists, field researchers, and data providers collaborate to address challenges.

-Capitalizes on complementary expertise and resources.

•Collaborative Initiatives:

-Co-design of observational campaigns and sensor deployments tailored to research objectives.

-Integration of in situ measurements into remote sensing algorithms.

•Engaging Data Providers:

-Government agencies, research institutions, and non-profit organizations promote data sharing and interoperability.

-Open data policies enhance accessibility and quality assurance.

Conclusions and Summary

We have delved into the intricate realm of ocean color remote sensing, exploring the vital role of in situ data in improving the accuracy and reliability of remote sensing algorithms.

1. Recap of key concepts:

- We explored the critical role of in situ data, encompassing measurements collected directly from the ocean's surface through buoys, research vessels, and autonomous platforms, in calibrating, validating, and refining remote sensing algorithms.
- We discussed the applications of in situ data in algorithm development and showcased examples of using buoy measurements for calibrating satellite sensors and validating ocean color products.

Conclusions and Summary (contd.)

2. Encouragement for Continued Exploration:

•I encourage each of you to continue your exploration and engagement in ocean color remote sensing research and applications in collaborative ways.

•By harnessing the power of in situ data, we have the potential to make significant strides in understanding and monitoring our oceans, thereby contributing to the conservation and sustainable management of marine ecosystems.

3. Leveraging In Situ Data for Environmental Challenges:

•Incorporating in situ data into remote sensing algorithms enables us to address pressing environmental challenges, including harmful algal blooms, coastal pollution, and climate change impacts on marine ecosystems.

•By leveraging in situ measurements, we can enhance the accuracy and reliability of satellite-derived ocean color products, providing valuable insights for policymakers, resource managers, and stakeholders.

Conclusions and Summary (contd.)

4. Advancing Scientific Knowledge:

•The integration of in situ data into remote sensing algorithms facilitates advancements in scientific knowledge, allowing us to unravel the complexities of ocean dynamics, biogeochemical cycles, and ecosystem functioning.

•Through interdisciplinary collaborations and innovative methodologies, we can expand our understanding of marine processes and phenomena, paving the way for informed decision-making and sustainable ocean stewardship.

In conclusion, the synergy between in situ data and remote sensing technologies offers unprecedented opportunities for unlocking the mysteries of the ocean and addressing global environmental challenges.