Cloud-based quantification of Spatial Uncertainty of Remote Sensing-based Benthic Habitat Classification and its utilization in the context of Active Learning

S. Christofilakos*[1], A.P. Pertiwi[1], C.B. Lee[1], D. Traganos[1]

[1] DLR - German Aerospace Center, Photogrammetry and image Analysis, Germany (*Spyridon.christofilakos@dlr.de)

This presentation participates in OSPP

Outstanding Student & PhD candidate Presentation contest

Photography encouraged

1. INTRODUCTION

THE SITUATION

With the latest advances in cloud image processing, scientists and policy makers have found an effective and robust platform to process vast satellite data in order to be able to map the extent, monitor the condition and create effective protection policies for different ecosystems across the globe.

THE KNOWLEDGE GAP

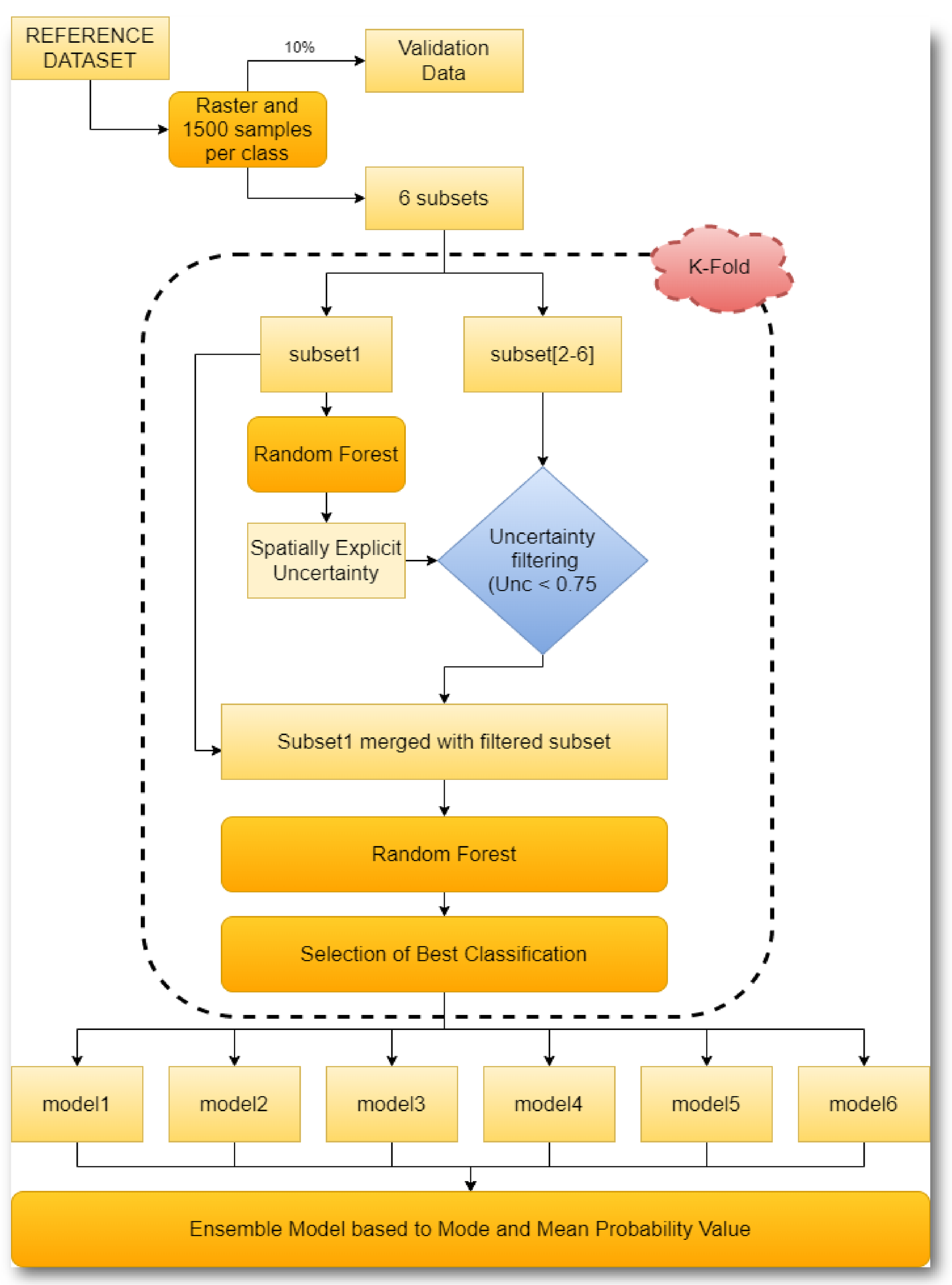
Cloud-based techniques lack information on the spatial explicit uncertainty of Machine Learning related mapping algorithms. Despite letting the researcher know which areas might be misclassified, spatial explicit uncertainty can also orient the researchers to the source of bias.

THE SUGGESTION

A novel approach on the estimation of uncertainty in a benthic habitat classification context. Furthermore, we explore the benefits of such information in the context of better classification results through an ensemble classifier and the visualization of the uncertain areas in an attempt to provide better maps to the policy makers.

2. MATERIALS AND METHODS

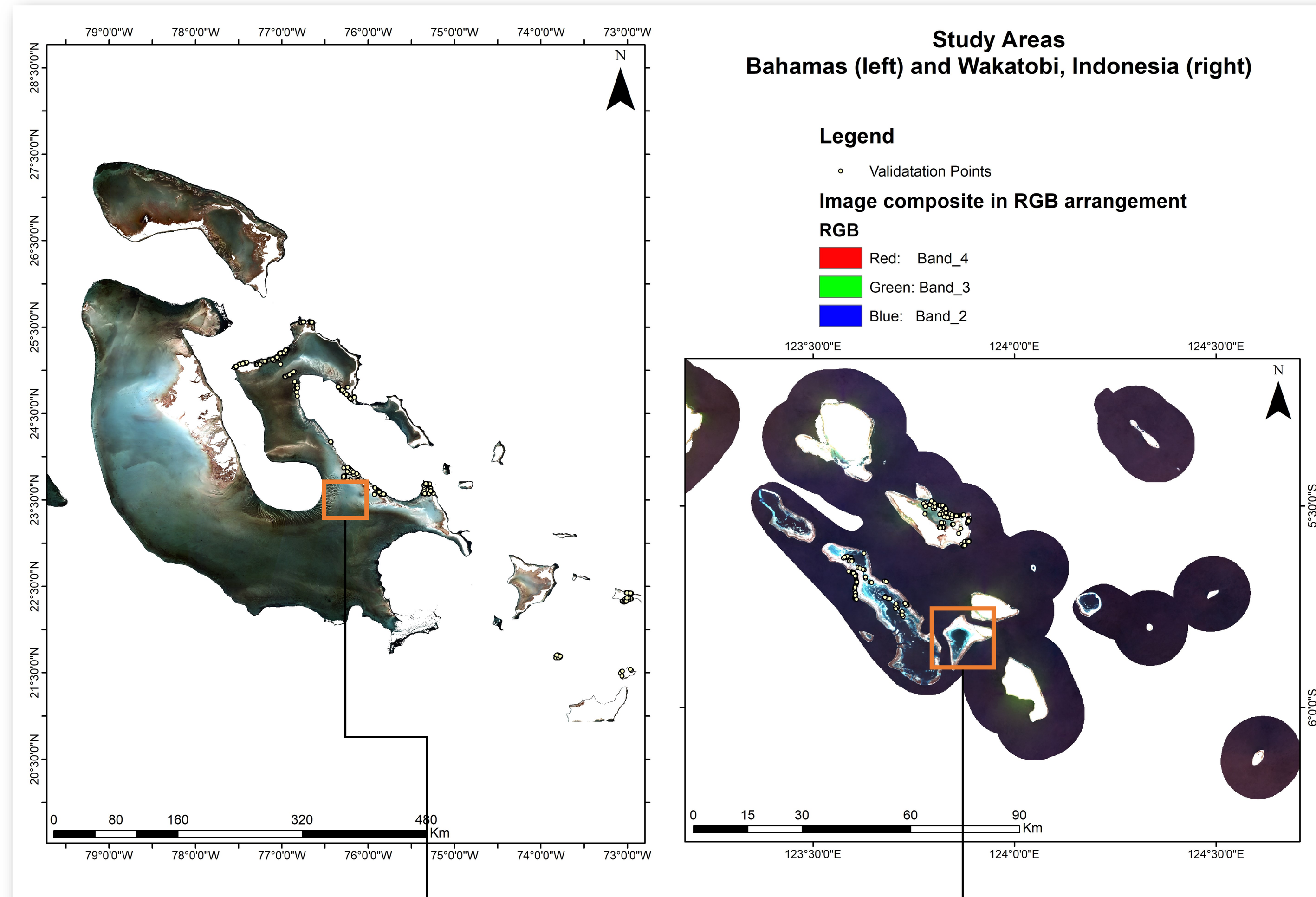
STUDY AREA	SIZE OF STUDY AREAS (Km^2)	SATELLITE TYPE	TIME PERIOD	BANDS THAT WERE USED
Wakatobi isle (Indonesia)	85255.81	PlanetScope, NICFI archive (5m resolution)	Whole dataset until 2022	['B', 'G', 'R', 'N', 'B_mean', 'B_stdDev', 'B_median', 'G_mean', 'G_stdDev', 'G_median', 'R_mean', 'R_stdDev', 'R_median', 'N_mean', 'N_stdDev', 'N_median'][1,2]
Bahamas[3]	114059.25	Sentinel 2– lvl 2A (10m resolution)	2 years	['B1', 'B2', 'B3', 'B4', 'B2B3', ' hue', 'value', 'B1_stdDev', 'B2_stdDev','B3_stdDev', 'B4_stdDev', 'B2B3_stdDev', 'hue_stdDev', 'value_stdDev', 'B1_median','B2_median', 'B3_medi-anv','B4_median','B2B3_median', 'hue_medi-an','value_median']



Reference Data	Seagrass	Sand	Coral	Rock
TD (both areas)	1350	1350	1350	1350
VD (Wakatobi)	158	158	158	158
VD (Bahamas)	300	300	300	300

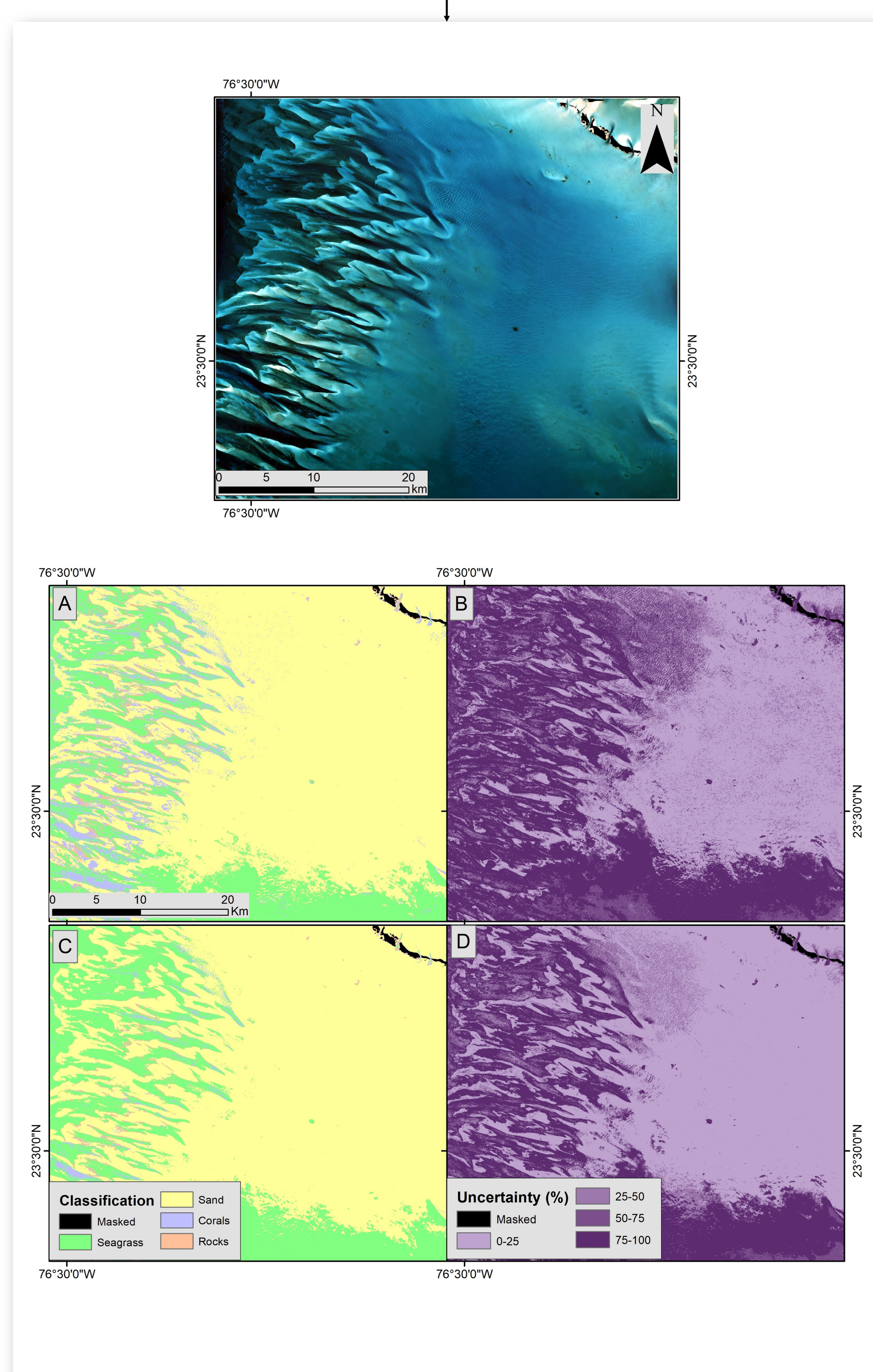
Bahamas' validation dataset comes from an independent source, while Wakatobi's comes from the same source as its training data.

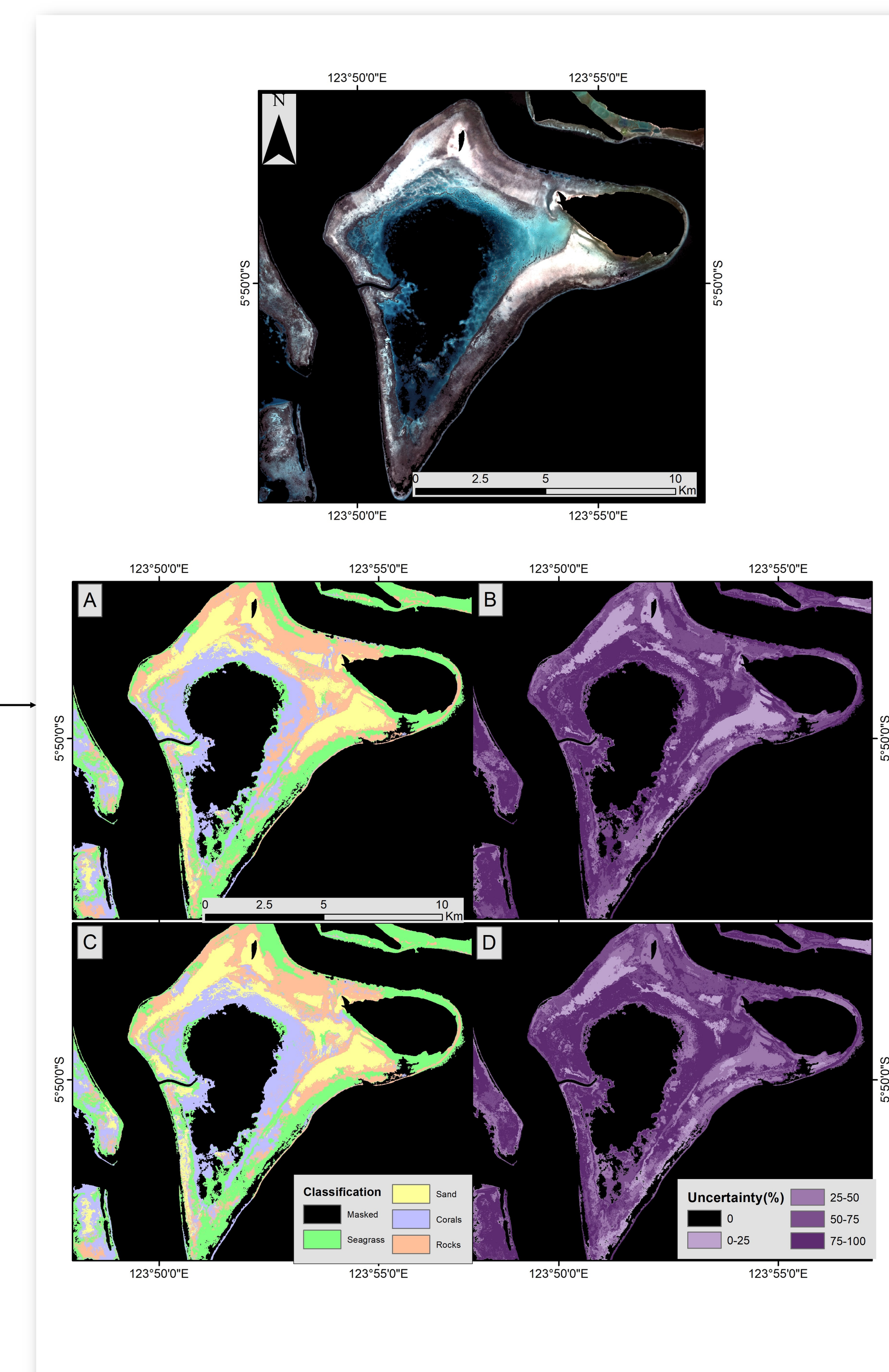
3. RESULTS



RF trees :15 Unc trhsld: 0.75	Ensemble Model		Classification w whole dataset	
	Overall Accuracy	F1 Score	Overall Accuracy	F1 Score
Wakatobi	0.993	0.994	0.995	0.995
Bahamas	0.603	0.603	0.598	0.597

RF trees :15 Unc trhsld: 0.75	Ensemble Model		Classification w whole dataset	
Class: Seagrass	Producer's Accuracy	User's Accuracy	Producer's Accuracy	User's Accuracy
Wakatobi	100%	98.6%	99.3%	100%
Bahamas	54.6%	58.9%	52.3%	60.3%





The above maps show the resulted classification and spatially explicit uncertainty layer with the presented workflow for both study areas.

A—*Initial Classification (with one training subset)
B—Spatially Explicit Uncertainty of initial classification
C—Ensemble model classification
D—Spatially Explicit Uncertainty of ensemble model

4. DISCUSSION

- While the ensemble model's accuracy was not statistically significantly bigger than the accuracy of the classification with the whole training data, the reduction of the noise into the classifier is certain due to the uncertainty threshold during the k-fold.

.

- The estimation of the uncertainty in the ensemble model is based on the average uncertainty values of the 6 best models. However, uncertainty values are expected to be relatively high due to the size of the training data (two subsets were used out of 6) at the end of k–fold process. With that said, future studies will also observe the uncertainty flex based on training data size.

5. CONCLUSIONS

- According to our results, the uncertainty workflow might not achieve a significant better overall accuracy score, but it improves producer accuracy , minimizes the difference between user and producer accuracy and smoothens down the classification and uncertainty surface information.
- Spatially Explicit Uncertainty is able to provide policy makers with crucial information regarding the extent, the condition and the conditions that compile the parameter under investigation. On top of that, its usage as a tool to distinguish certain from uncertain classified areas can be mitigate to continuous distributions like satellite derived bathymetry, sea surface temperature or even blue carbon accounting.

6. ACKNOWLEDGEMENTS

i) The Bahamian image composite was created by Alina Blume during her master thesis in DLR.
ii) The reference points for both areas were sampled by the same source, Allen Coral Atlas (https://allencoralatlas.org/) . The validation points though for Bahamas were collected in situ with scientific dives and were used by Alan Coral Atlas team as their validation dataset.

7. BIBLIOGRAPHY

1.Dimosthenis Traganos & Peter Reinartz (2018) Machine learning-based retrieval of benthic reflectance and *Posidonia oceanica* seagrass extent using a semi-analytical inversion of Sentinel-2 satellite data, International Journal of Remote Sensing, 39:24, 9428-9452, DOI: 10.1080/01431161.2018.1519289
2.. Lee, C.B.; Traganos, D.; Reinartz, P. A Simple Cloud-Native Spectral Transformation Method to Disentangle Optically Shallow and Deep Waters in Sentinel-2 Images. *Remote Sens.* 2022, *14*, 590. https://doi.org/10.3390/rs14030590
3. Blume, A.; Pertiwi, A.P.; Lee, C.B.; Traganos, D. (2023). Bahamian seagrass extent and blue carbon accounting using Earth Observation, Frontier in Marine Science, 2023, 10. 2296-7745, DOI: 10.3389/fmars.2023.1058460

GLOBAL SEAGRASS WATCH
serverless is more

DAAD
Deutscher Akademischer Austausch Dienst
German Academic Exchange Service

DLR Deutsches Zentrum für Luft- und Raumfahrt
German Aerospace Center