living planet symposium BONN 2022

A CLOUD-BASED APPROACH ON REMOTE SENSING BASED UNCERTAINTY MAPS, IN MARINE HABITAT MAPPING

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

[1]DLR-Geman Aerospace Center, Remote Sensing Technology Institute (IMF), Department of Photogrammetry and Image Analysis, 12489, Berlin, Germany. * spyridon.christofilakos@dlr.de

[2]German Aerospace Center (DLR), Remote Sensing Technology Institute (IMF), 82234 Wessling, Germany

Introduction

THE INFORMATION GAP

The necessity of monitoring and expanding the existing Marine Protected Areas has led to vast and high-resolution map products which, even if they feature high accuracy, they lack information on the spatially explicit uncertainty of the habitat maps, a structural element in the agendas of policy makers and conservation managers for designation and field efforts.

THE SUGGESTION

The target of this study is to fill the gaps in the visualization and quantification of the uncertainty of benthic habitat mapping by producing an end-to-end continuous layer using relevant training datasets.

Methods & Data

By applying a semi-automated function in Google Earth Engine's cloud environment we are able to estimate the spatially explicit uncertainty of a supervised benthic habitat classification product. In this study we explore and map the aleatoric uncertainty of multi-temporal data driven, per-pixel classification . Aleatoric uncertainty, also known as data uncertainty, is part of the information theory that seeks for the data driven random and inevitable noise under the spectrum of Bayesian statistics.

Study area: Bahamas' coastal extent (114,059.25 km^2)

Satellite Imagery: 2 year timeseries of Sentinel-2 lvl2a data

Classification's training and validation data (figure 1) :

	Seagrass	Sand	Coral	Rock	Source
Training points	500	500	500	500	Annotation
Validation points	400	400	400	400	Allen Coral Atlas [1] https://allencoralatlas.org/atlas/#5.45/21.7490/-75.2633

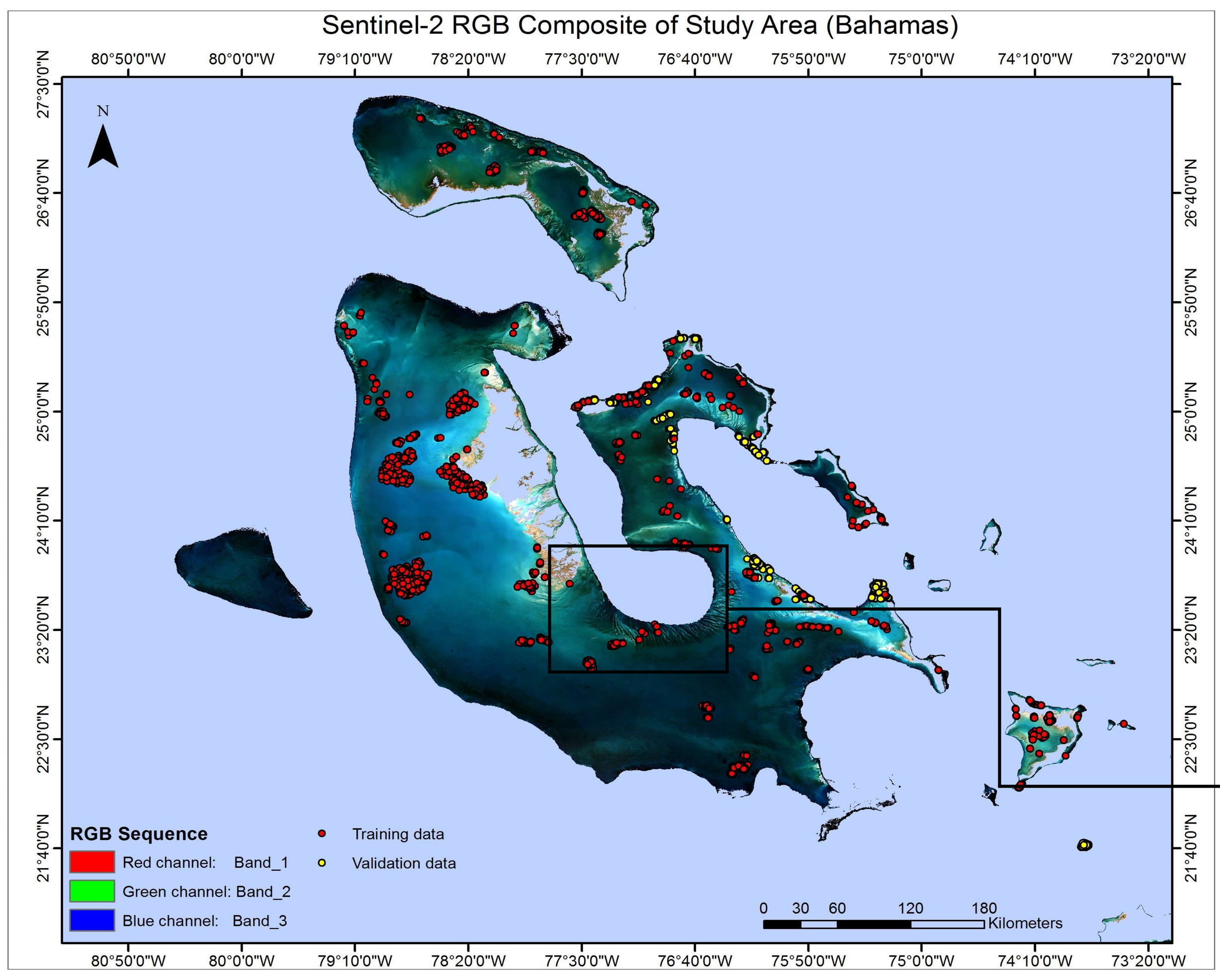
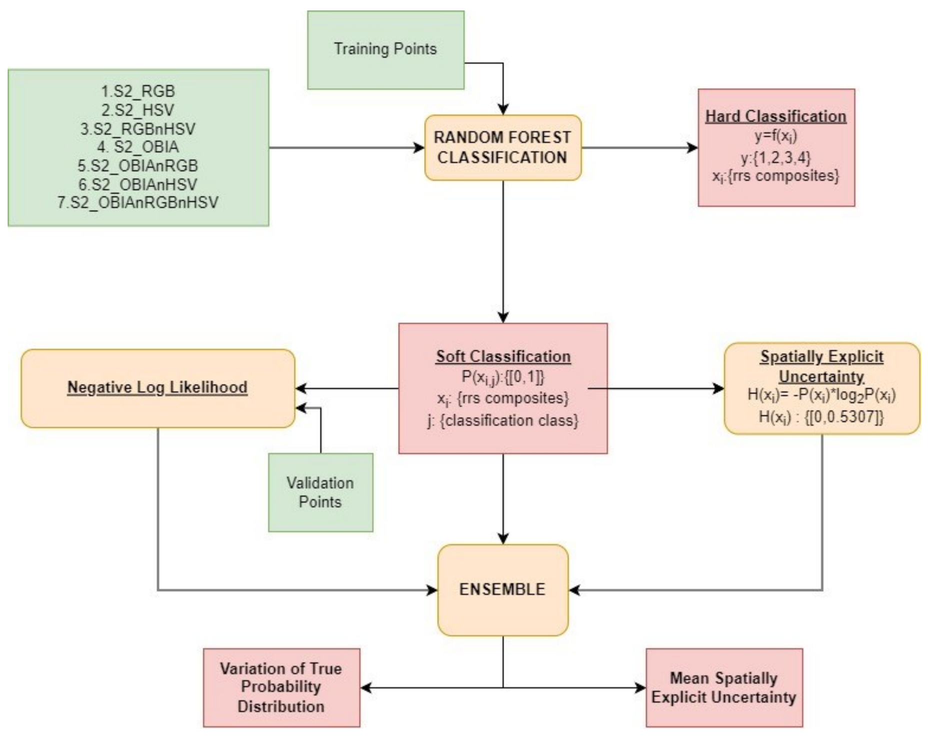


Figure 1. RGB composite of the study area and reference data for the classification. Coordinate system: GCS_WGS_1984

The aim by implementing the presented workflow is to quantitatively identify and minimize the spatial residuals in large-scale coastal eco-system accounting.



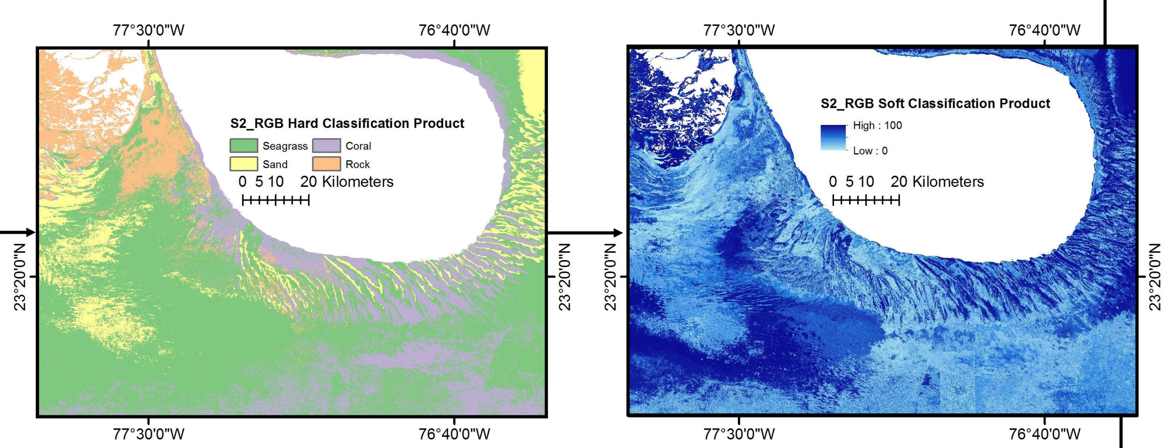


Figure 2. Hard classification (left) and Soft classification (right) with S2_RGB as input of the classification

1.By introducing different bands in each classification we develop 7 classifications products. These products carry the categorical (Hard Classification) and continuous (Soft Classification) values of the classification (figure 2).

Inputs of the classification	i)Sentinel 2 RGB	ii)Sentinel 2 HSV	iii) Sentinel 2 RGB and HSV	iv) Sentinel 2 OBIA OBIA: Object-Based Image Analysis	v) Sentinel 2 OBIA and RGB	vi) Sentinel 2 OBIA and HSV	vii) Sentinel 2 RGB and HSV
	Bands (4): B1,B2,B3,B4	Bands(3): Hue, Saturation, ,Value	Bands(7): i) + ii)	Bands(21) : mean, stdDev and median values of iii)	Bands(25): iv) and i)	Bands(24): iv) and ii)	Bands(28): iv) and iii)
Outputs of the classifications	S2_RGB	S2_HSV	S2_RGBnHSV	S2_OBIA	S2_OBIAnRGB	S2_OBIAnHSV	S2_OBIAnRGBnHSV

CROSS-REFERENCE WITH CURRENT LITERATURE

To visualize spatially the divergence between the soft classification values per pixel and the soft classification values per class category, we parametrized the NLL as follows :

$$D = \frac{(\hat{\mu}(X_{i,j}) - Y_{i,j})^2}{2\sigma^2(X_{i,j})} + \frac{1}{2}\log\left(\widehat{\sigma^2}(X_{i,j})\right)$$

μ(X) = mean soft probability of each classification class

Y = soft probability value per pixel of each class

i = classification product corresponding to the initial composite that was classified

j = classification category

2. By calculating the marginal and conditional distribution's divisions given the available training data, we can estimate the **Expected Entropy, Mutual Information and Spatially Explicit Uncertainty** of a maximum likelihood model outcome.

3. Moreover, a visual comparison (figure 3) takes place between the Spatially Explicit Uncertainty and the True Conditional Probability distribution. In order to approximate the true conditional distribution, we apply a Gaussian **Negative Log Likelihood (NLL)**[5] estimation.

DID YOU KNOW?

*To download **one** Sentinel 2 image, you would need a few minutes.*

To execute the current study through Google Earth Engine, took only 5 minutes.[6]

Expected Conditional Entropy[7]

Predicts the overall data uncertainty of the distribution P(x,y), with x:training dataset and y:model outcome.

Mutual Information[7]

Estimates in total and per classified class the level of independence and therefore the relation of y and x distributions.

Spatially Explicit Uncertainty

A per pixel estimation of the uncertainty of the classification.

Results

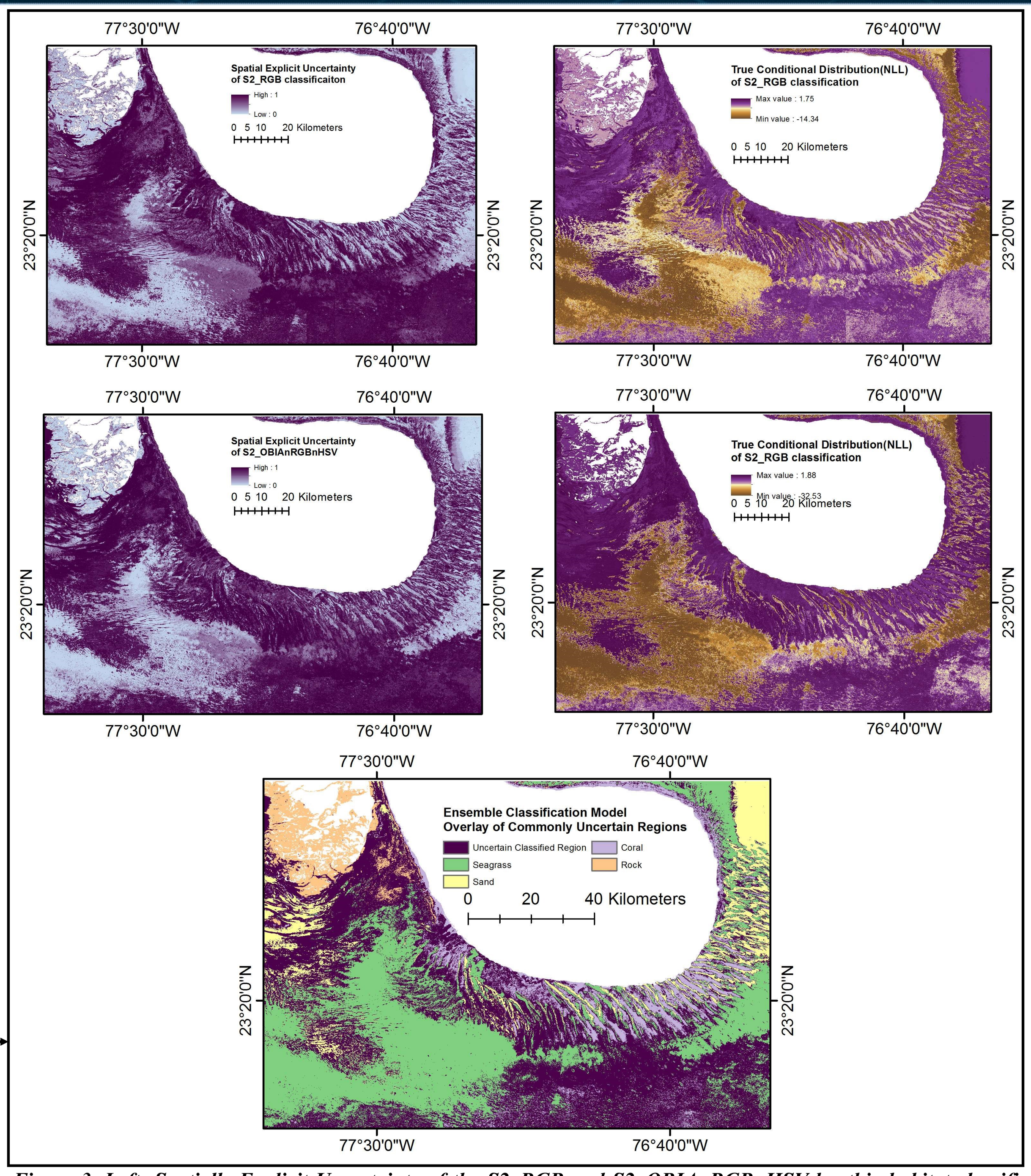
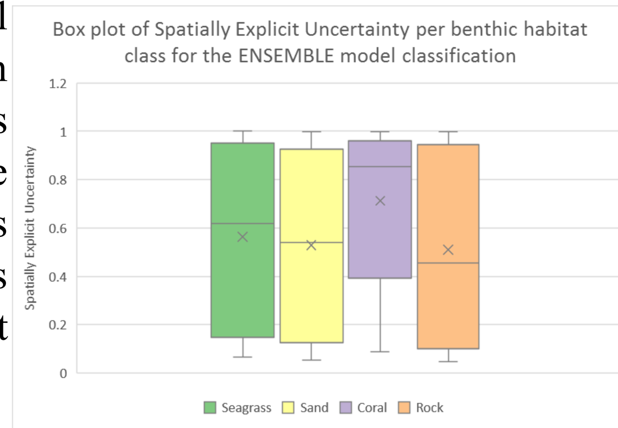


Figure 3. Left: Spatially Explicit Uncertainty of the S2_RGB and S2_OBIAnRGBnHSV benthic habitat classifications. Right: NLL distribution of the probability distribution of the S2_RGB and S2_OBIAnRGBnHSV classifications. Bottom: Overlay of benthic habitat classification and mutual uncertain regions with value >0.8 of the ensemble models of Spatially Explicit Uncertainty and NLL distribution procedures.

The overlay of the mutual uncertain regions seems to come in an agreement with the low user's and producer's accuracy of the ensemble model which indicates the coral and seagrass classes as the most and second most misclassified benthic habitats.



		S2_RGB	S2_OBIAnRGBnHSV	Ensemble model
User's accuracy	**Coral**	0.316	0.296	0.1
	Rock	0.464	0.480	0.491
	Sand	0.940	0.971	0.962
	Seagrass	0.271	0.287	0.311
Producer's accuracy	**Coral**	0.393	0.373	0.378
	Rock	0.715	0.773	0.768
	Sand	0.67	0.753	0.77
	Seagrass	0.138	0.103	0.118
	Overall accuracy	0.479	0.5	0.508
	Expected Conditional Entropy	0.092	0.049	0.027
	Mutual information	0.919	0.93	0.935

Conclusions

- Our results indicate regions and classes with high and low uncertainty that can either be used for a better selection of the training dataset or to identify, in an automated fashion, areas and habitats that are expected to feature misclassifications not highlighted by existing qualitative accuracy assessments.
- By doing so, we can streamline more confident, cost-effective, and targeted benthic habitat accounting and ecosystem service conservation monitoring , resulting in strengthened research and policies, globally.

References

1) B. Lyons, M., M. Roelfsema, C., V. Kennedy, E., M. Kovacs, E., Borrego-Acevedo, R., Markey, K., Roe, M., M. Yuwono, D., L. Harris, D., R. Phinn, S., Asner, G.P., Li, J., E. Knapp, D., S. Fabina, N., Larsen, K., Traganos, D., J. Murray, N., 2020. Mapping the world's coral reefs using a global multiscale earth observation framework. Remote Sens. Ecol. Conserv. 1–12. https://doi.org/10.1002/rse2.157

2) A. Blume, 2021. composition and related carbon stocks in support of the Nationally Eidesstattliche Versicherung Statutory Declaration in Lieu of an Oath

3) C.B. Lee, D. Traganos, P. Reinartz, 2022. A Simple Cloud-Native Spectral Transformation Method to Disentangle Optically Shallow and Deep Waters in Sentinel-2 Images. https://doi.org/10.3390/rs14030590

4) D. Traganos, B. Aggarwal, D. Poursanidis, K. Topouzelis, N. Chrysoulakis, P. Reinartz, 2018. Towards Global-Scale Seagrass Mapping and Monitoring Using Sentinel-2 on Google Earth Engine: The Case Study of the Aegean and Ionian Seas. Remote Sens. 10, 1227. https://doi.org/10.3390/rs10081227

5) N. Lang, N. Kalischek, J. Armston, K. Schindler, R. Dubayah, J.D. Wegner, 2022. Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles. Remote Sens. Environ. 268, 112760. https://doi.org/10.1016/j.rse.2021.112760

6) N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, R. Moore, 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2017.06.031

GLOBAL SEAGRASS WATCH

serverless is more

#LPS22