DLR Institute of Remote Sensing Technology
Photogrammetry and Image Analysis Department (PBA)
Team: Optical Remote Sensing of Water Bodies (BA)

First Steps in Estimating the Spatial Uncertainty of Maximum Likelihood Tasks in a Cloud-based Environment in Context of Marine Remote Sensing

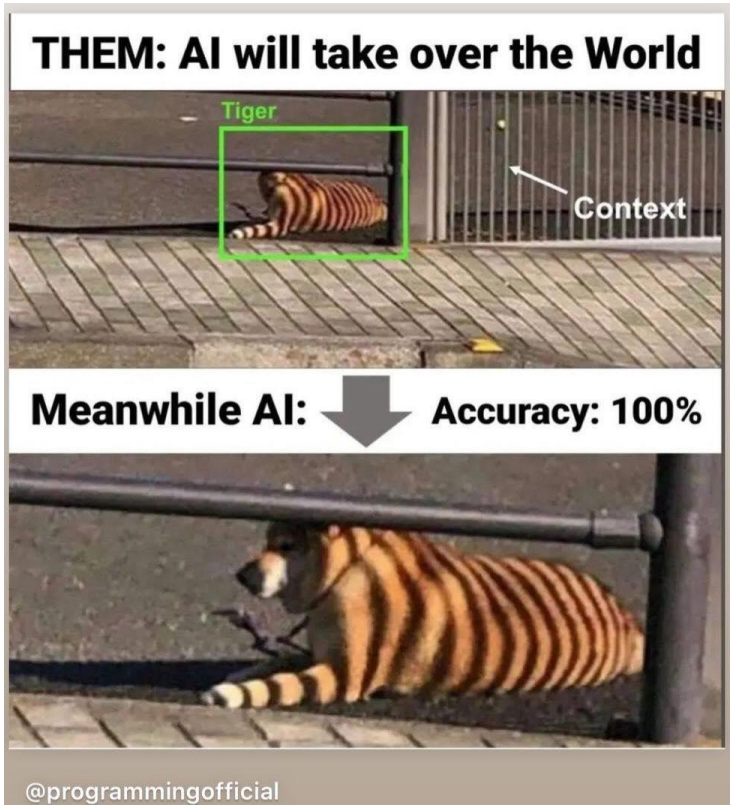
Spyros Christofilakos

GLOBAL SEAGRASS WATCH
serverless is more

DLR

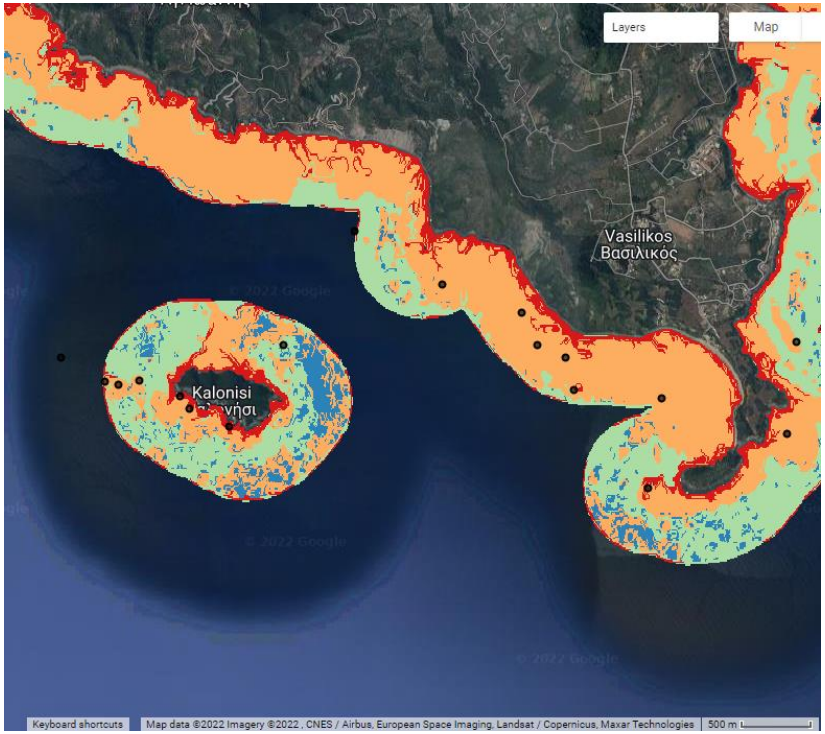
DAAD
Deutscher Akademischer Austausch Dienst
German Academic Exchange Service

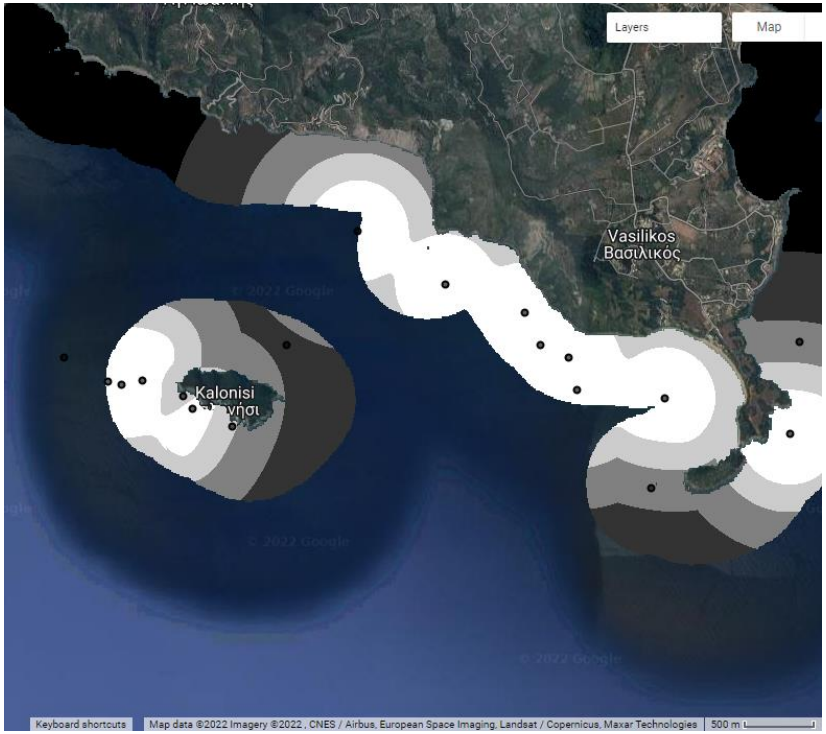
How accurate is a classification, spatially?



How accurate is a classification, spatially?

- Accuracy assessment is spatially bound





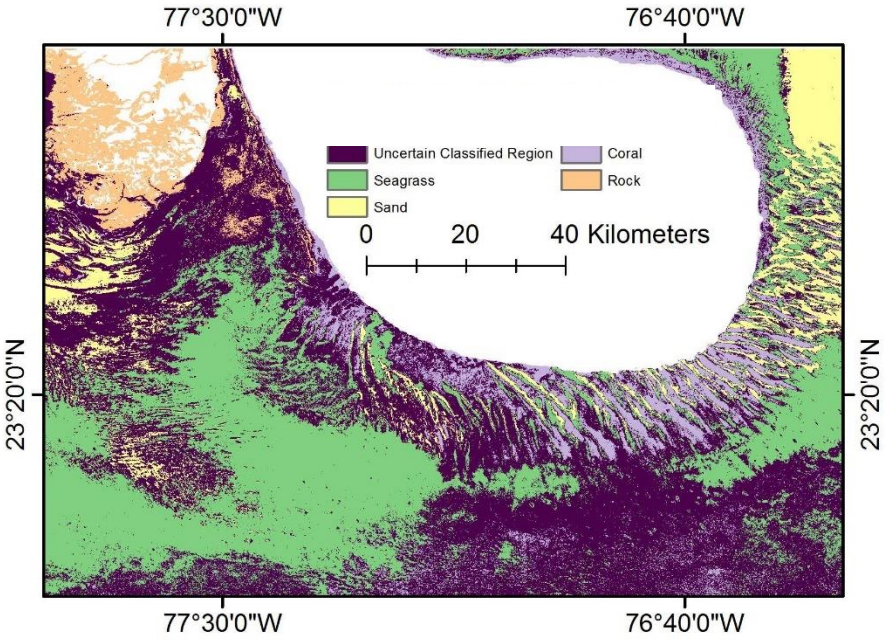
Goal of the research

- Develop a semi-automated workflow to estimate the spatial explicit uncertainty of classification and regression procedures that take place in coastal ecosystems

Goal of the research

- Develop a semi-automated workflow to estimate the spatial explicit uncertainty of classification and regression procedures that take place in coastal ecosystems

1) Highlight the uncertain areas



Goal of the research

- Develop a semi-automated workflow to estimate the spatial explicit uncertainty of classification and regression procedures that take place in coastal ecosystems

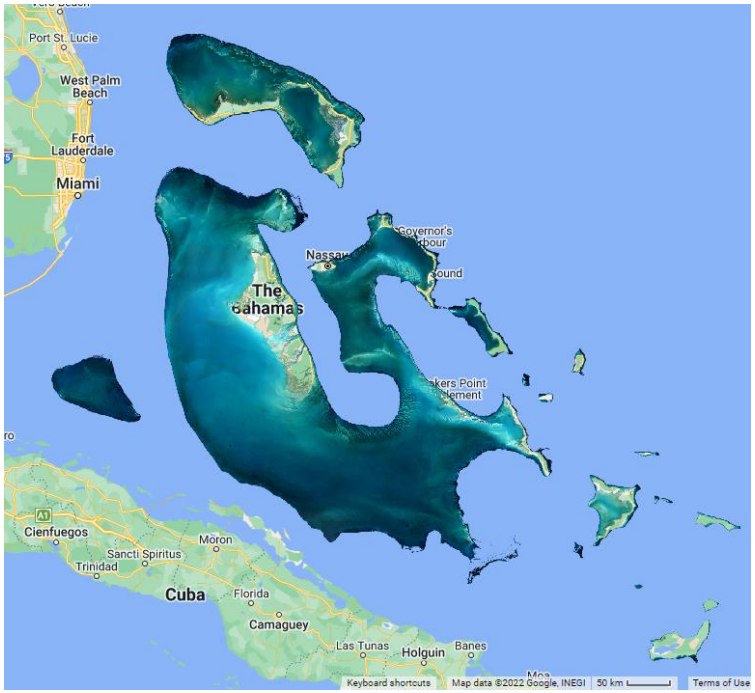
1) Highlight the uncertain areas
2) Acquire training data from the uncertain areas and re-train the model

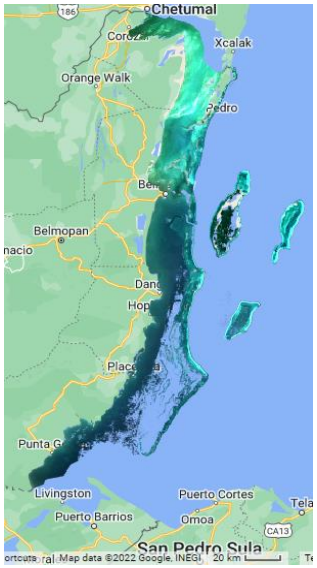
Goal of the research

- Develop a semi-automated workflow to estimate the spatial explicit uncertainty of classification and regression procedures that take place in marine ecosystems

1) Highlight the uncertain areas
2) Acquire training data from the uncertain areas and re-train the model
3) Be able to tell how accurate is the classification/regression spatially (EU Habitats Directive)

Study Areas





CLASSIFICATION

- Task: Benthic Habitat classification
- Case study: Bahamas
- Satellite Data: Four years timeseries of Sentinel-2 system, lvl-2A data
- Validation Points: 300 per class
- Training Points: 1000 per class

(Allen Coral Atlas)

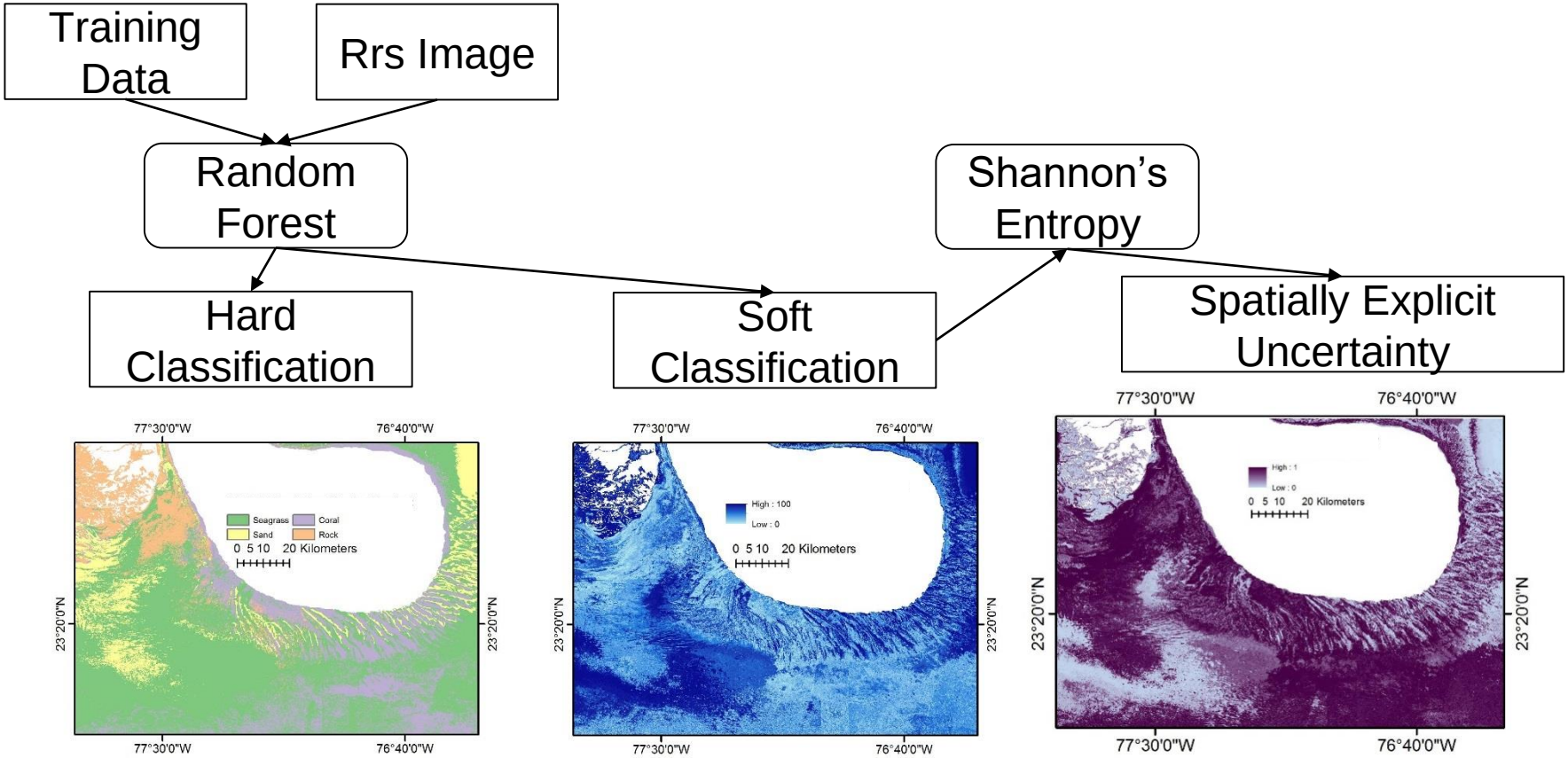
REGRESSION

- Task: Satellite Derived Bathymetry
- Case study: Belize, Quirimbas (Mozambique)
- Satellite Data: Two years timeseries of Sentinel-2 system, lvl-2A data
- Validation Points: 800 (777 after rescaling)
- Training Points: 3200 (3110 after rescaling)

Blume, Alina (2021) *Development of cloud-native and scalable algorithms to estimate seagrass composition and related carbon stocks in support of the Nationally Determined Contributions of the Paris Agreement.* Master's, University of Aachen. (https://elib.dlr.de/148787/)

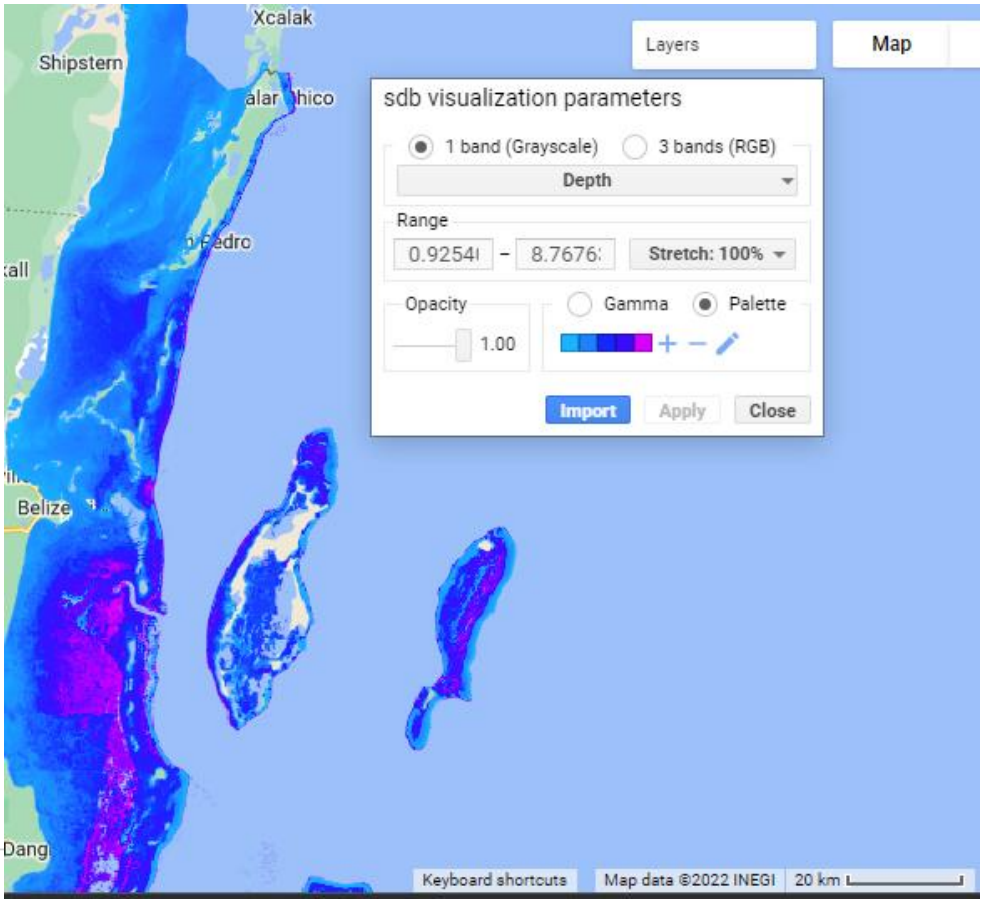
N. Marc Thomas et all., (2020).**SPACE-BORNE CLOUD-NATIVE SATELLITE-DERIVED BATHYMETRY (SDB) MODELS USING ICESat-2 and SENTINEL-2**

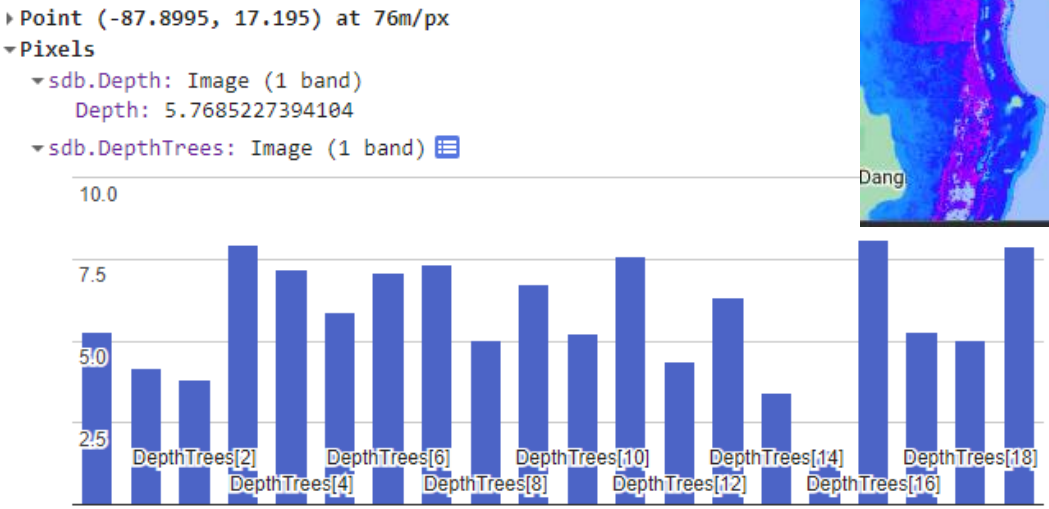
Uncertainty in classification
Benthic habitat Classification



Uncertainty in regression

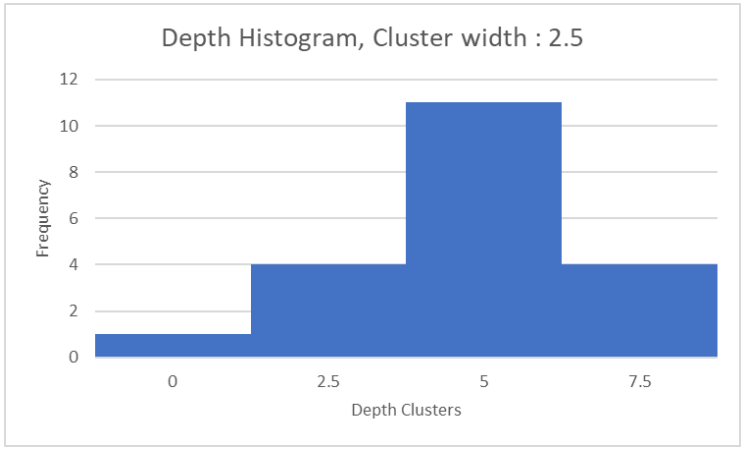
Bathymetry regression with Random Forest classifier of 20 trees

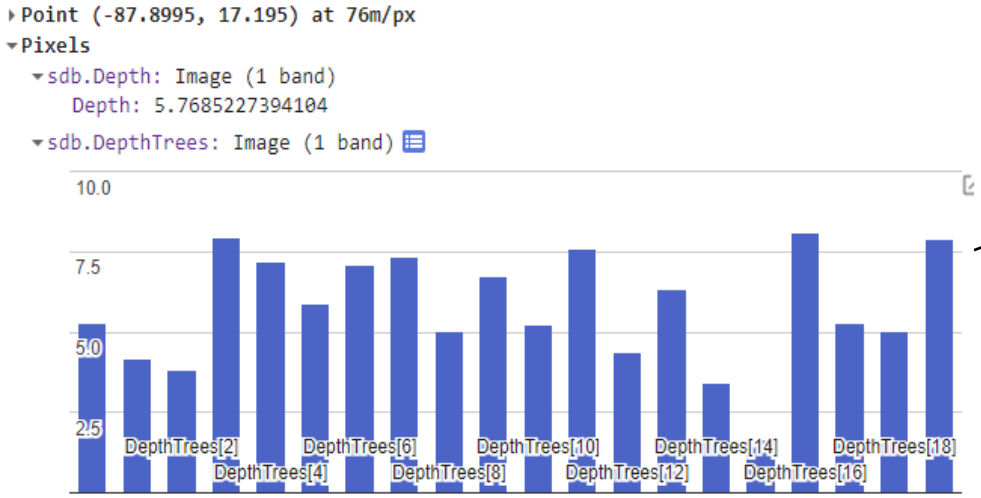


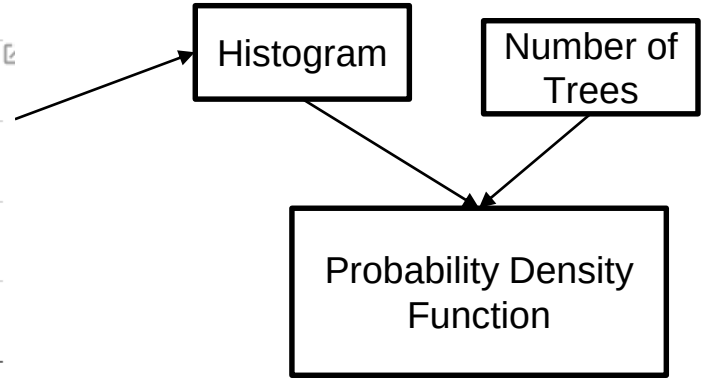


Uncertainty in regression

Bathymetry regression with Random Forest classifier of 20 trees

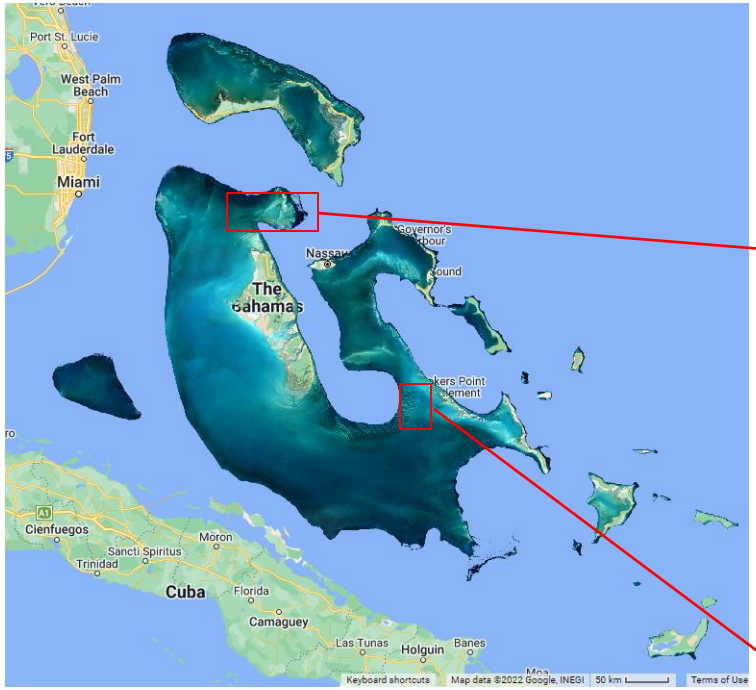


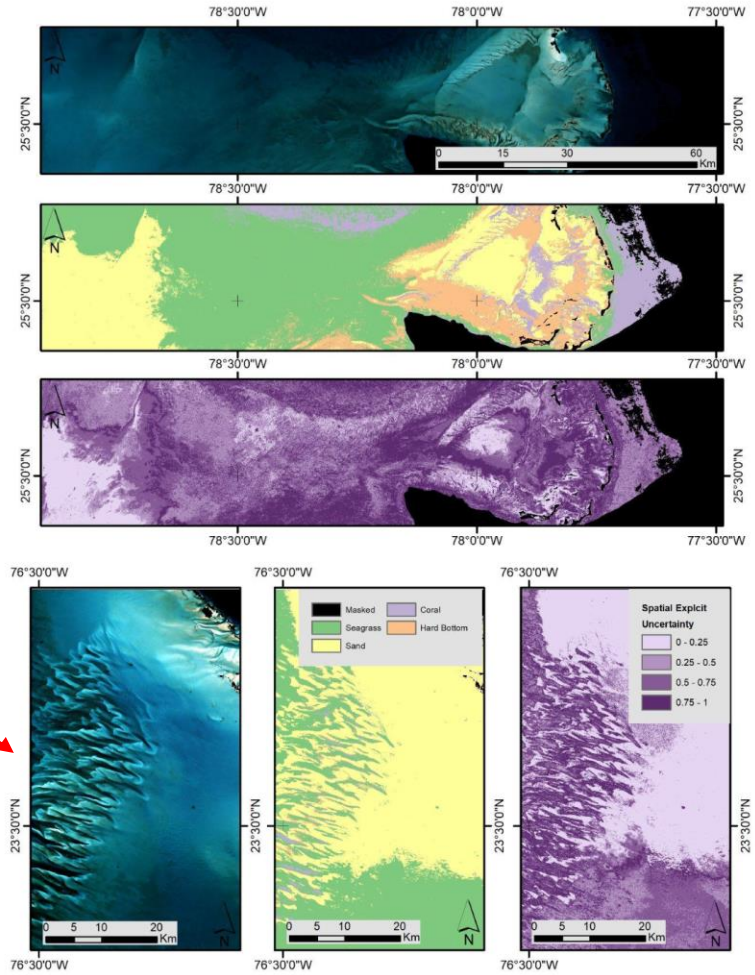




Results: Accuracy Assessment in Classification

OBIA			
	Initial Classification	**Retrained from Uncertain Areas lt(0.25)**	**Accuracy Gain**
Overall Accuracy	57.83%	62.08%	4.25%
User's Accuracy	53.82%	60.30%	6.48%
Producer's Accuracy	54.00%	67.33%	13.33%


Port St. Lucie
West Palm Beach
Fort Lauderdale
Miami
Nassau
The Bahamas
Cienfuegos
Sancti Spiritus
Trinidad
Cuba
Florida
Camaguey
Las Tunas
Holguin
Banes
Moron
Keyboard shortcuts
Map data ©2022 Google, INEGI
50 km
Terms of Use


78°30'0"W
78°0'0"W
77°30'0"W
25°30'0"N
23°30'0"N
76°30'0"W
0 15 30 60 Km
Masked
Coral
Seagrass
Hard Bottom
Sand
Spatial Explcit Uncertainty
0 - 0.25
0.25 - 0.5
0.5 - 0.75
0.75 - 1
0 5 10 20 Km
N

Results: Accuracy Assessment in Regression

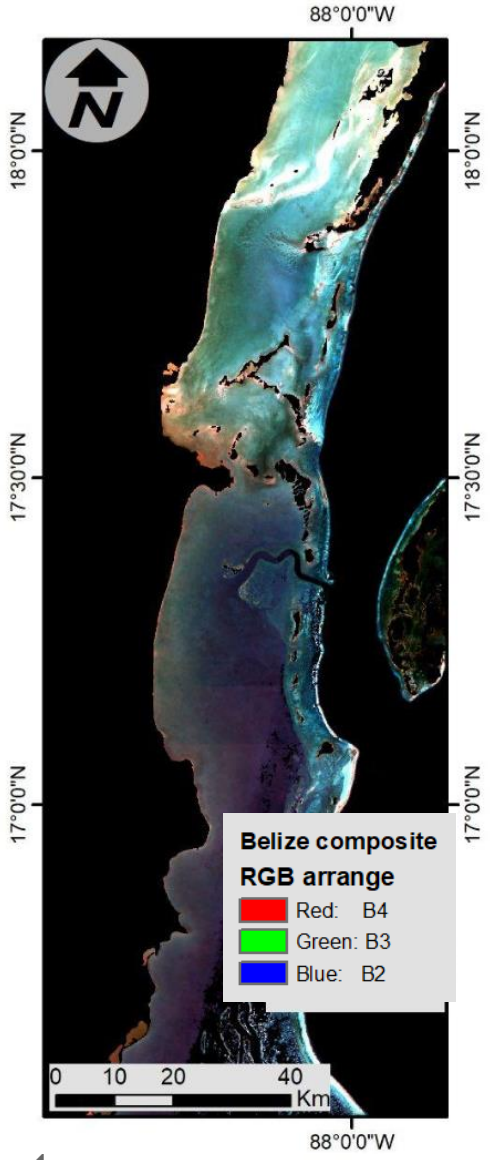
Quirimbas

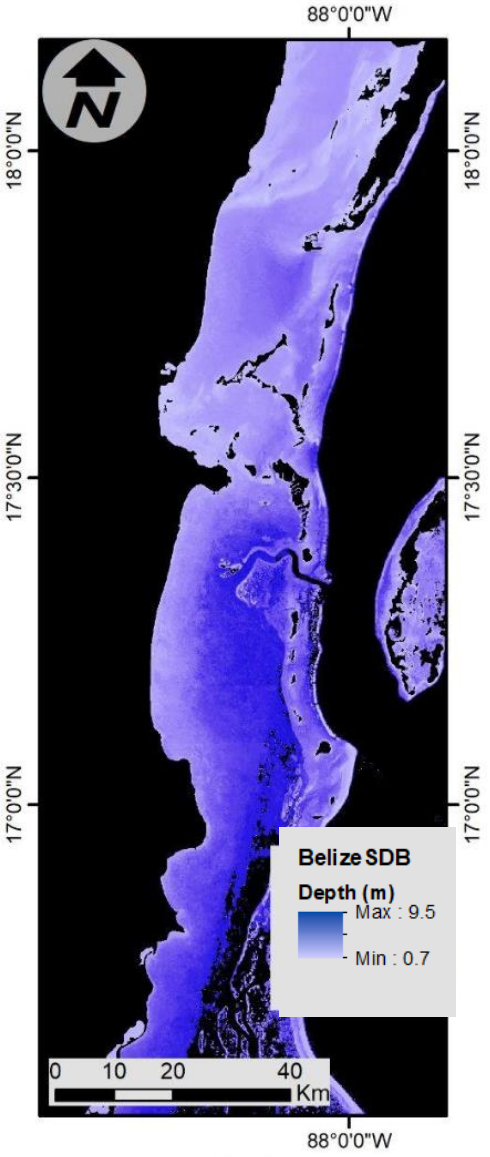
OBIA			
model	Initial Regression	Retrained from Uncertain Areas lt(0.25)	Accuracy Gain
MeanSqr Error	2.6328	2.1955	0.4373
r_sqr	0.6289	0.6162	0.0127

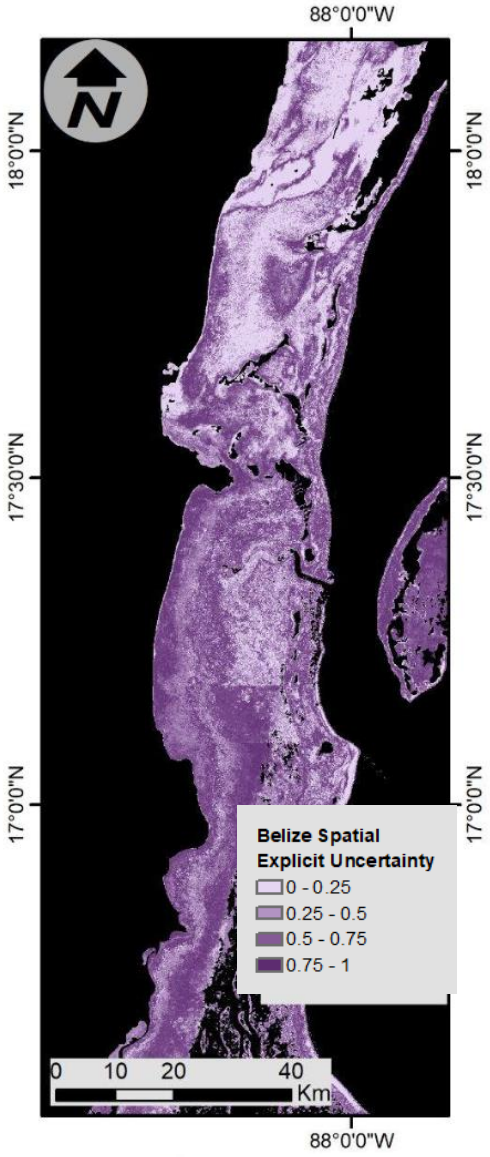
Belize

OBIA			
model	Initial Regression	Retrained from Uncertain Areas lt(0.25)	Accuracy Gain
MeanSqr Error	1.2306	1.1479	0.0827
r_sqr	0.6104	0.6026	0.0078

Uncertainty in Regression (Belize)

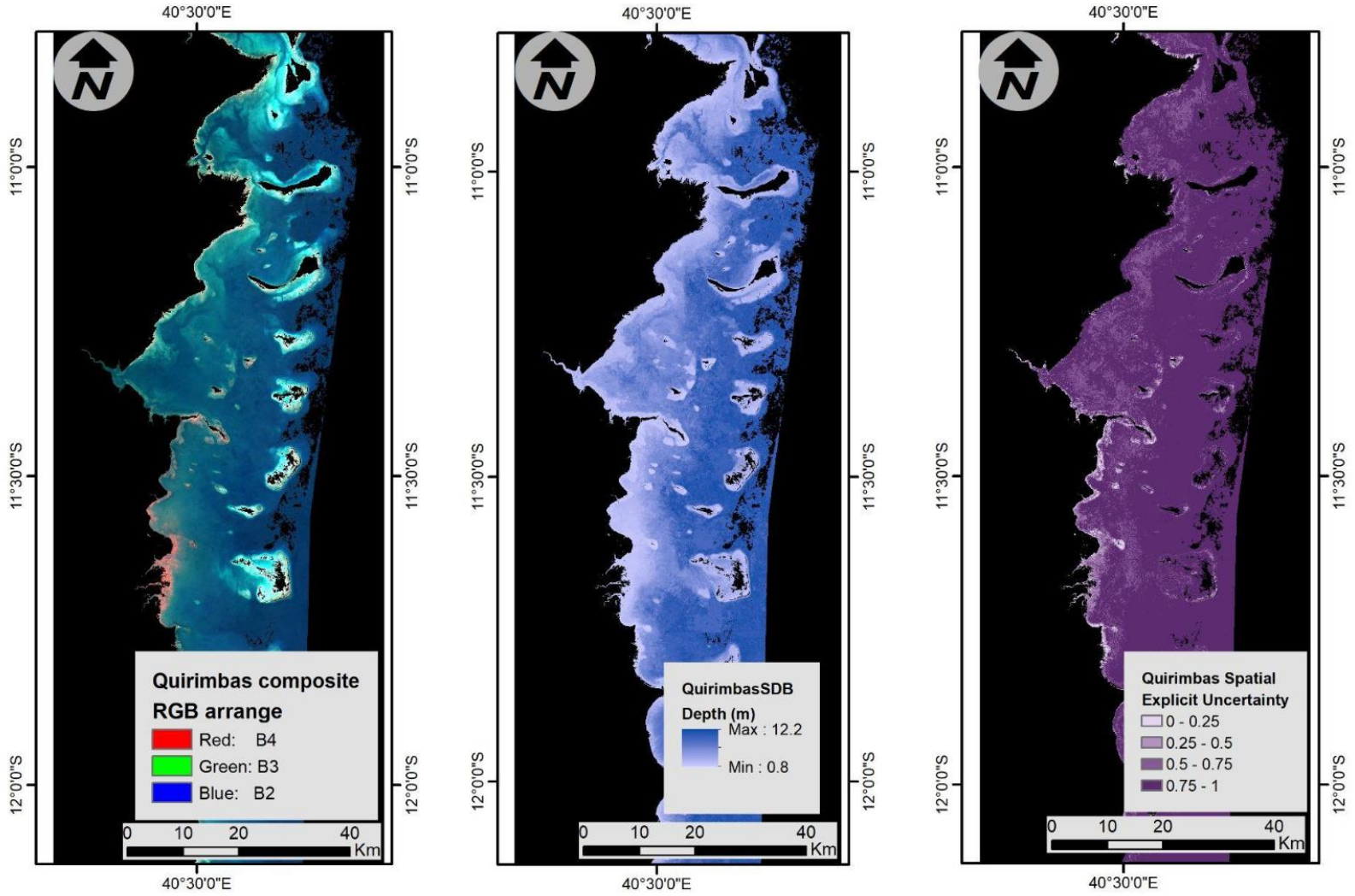






DLR

Uncertainty in Regression (Quirimbas)



Conclusions and Future Steps

- Spatial Explicit Uncertainty seems as a promising variable to improve the understanding of remote sensing data, models, applications.

i) Experimentation with larger training dataset

ii) Experimentation with PlanetScope data (3m)

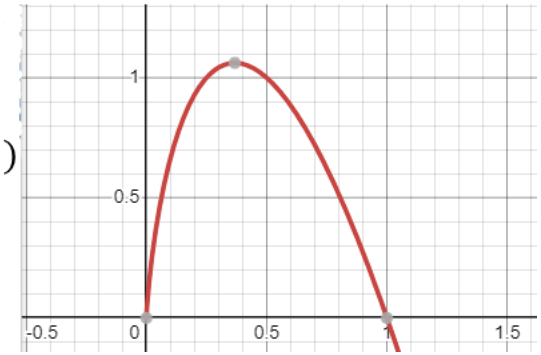
Thank you!

Uncertainty in classification
Head or Tails?

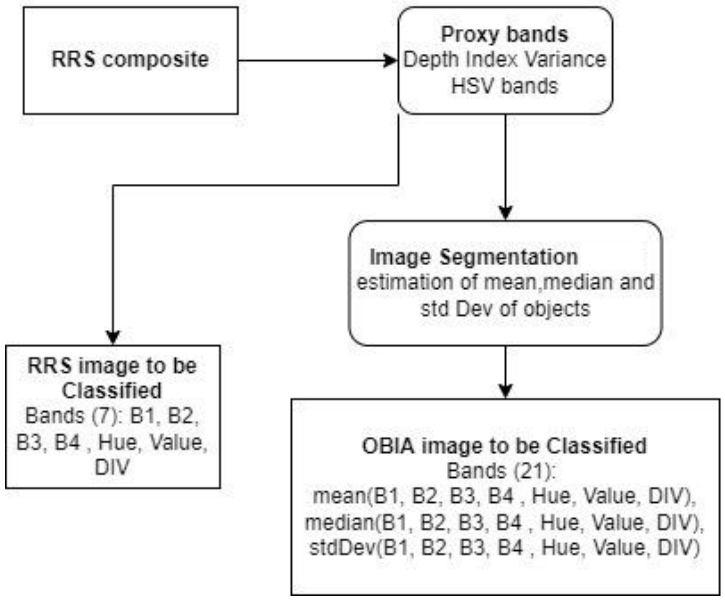
- 1) Possible outcome: Head , Tails
- 2) Probabilities of the outcome: P(H)= 50%
  P(T)=50%
- 3) Shannon's Entropy

$$E(x) = -\sum_{i=1}^{N} P(x_i) * \log_2 P(x_i)$$

$$-\sum_{i=1}^{2} P(x_i) * \log_2 P(x_i)$$



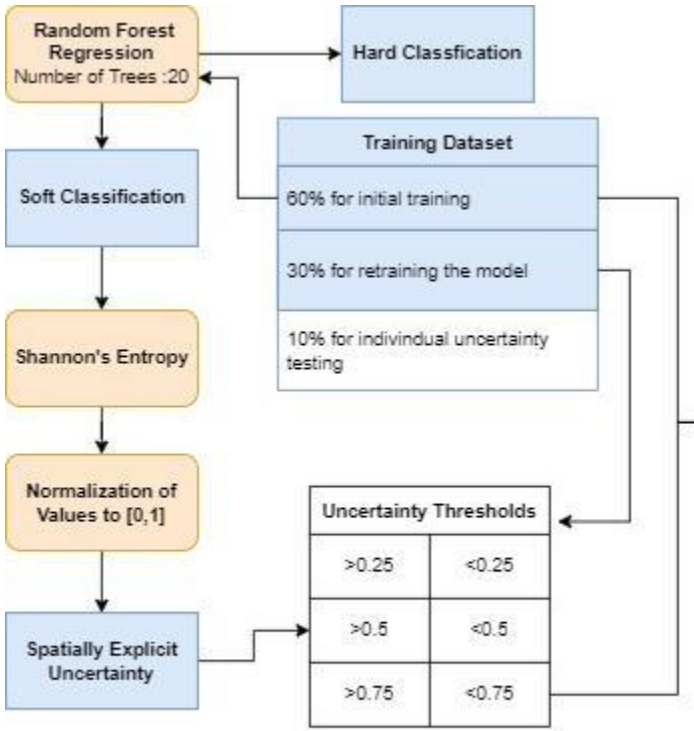
Data Pre-processing



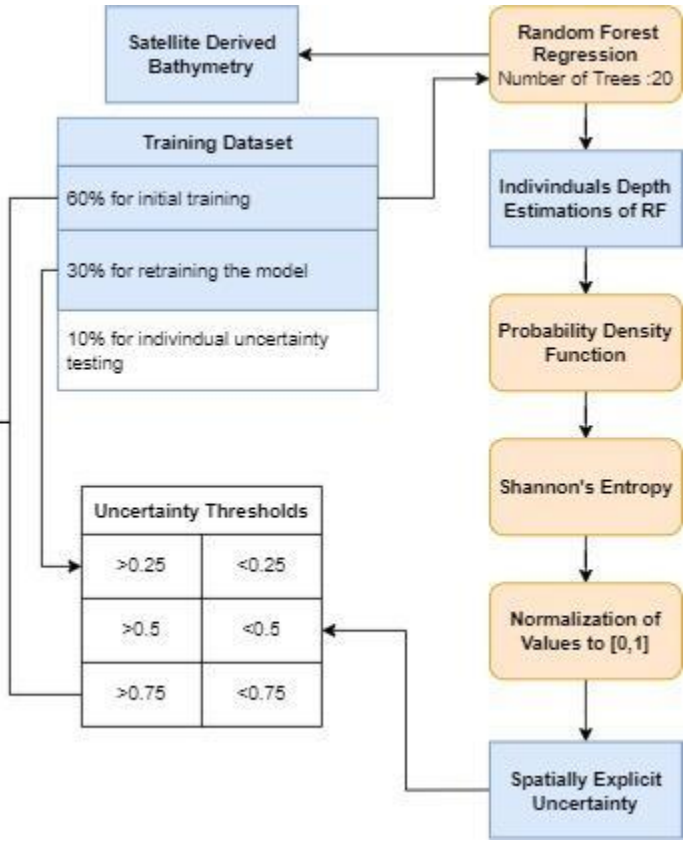
Training Dataset
60% for initial training
30% for retraining the model
10% for indivindual uncertainty

Data Processing

Distinct Distribution (Classification)



Continuous Distribution (Regression)



Retrain Model

Results: Accuracy Assessment in Classification

OBIA	lt: Less than	gt: Greater than						
model	**Retrained from Uncertain Areas lt(0.25)**	**Initial Classification**	**Retrained from Uncertain Areas lt(0.5)**	**Retrained from Uncertain Areas lt(0.75)**	**Retrained from Uncertain Areas gt(0.25)**	**Retrained from Uncertain Areas gt(0.5)**	**Retrained from Uncertain Areas gt(0.75)**	**Classification with 90% of Data**
Overall Accuracy	62.08%	57.83%	60.92%	58.83%	59.58%	60.42%	58.83%	59.17%
	Percentage Gain	4.25%	1.17%	3.25%	2.50%	1.67%	3.25%	2.92%
User's Accuracy	60.30%	53.82%	58.86%	55.56%	53.94%	56.01%	57.19%	56.37%
	Percentage Gain	6.48%	1.44%	4.74%	6.36%	4.29%	3.11%	3.93%
Producer's Accuracy	67.33%	54.00%	62.00%	61.67%	61.67%	59.00%	61.67%	59.00%
	Percentage Gain	13.33%	5.33%	5.67%	5.67%	8.33%	5.67%	8.33%

RGB	lt: Less than	gt: Greater than						
model	**Retrained from Uncertain Areas lt(0.5)**	**Initial Classification**	**Retrained from Uncertain Areas lt(0.25)**	**Retrained from Uncertain Areas lt(0.75)**	**Retrained from Uncertain Areas gt(0.25)**	**Retrained from Uncertain Areas gt(0.5)**	**Retrained from Uncertain Areas gt(0.75)**	**Classification with 90% of Data**
Overall Accuracy	59.33%	56.92%	56.75%	56.83%	57.17%	57.67%	58.25%	57.25%
	Percentage Gain	2.42%	2.58%	2.50%	2.17%	1.67%	1.08%	2.08%
User's Accuracy	48.35%	44.62%	45.08%	44.44%	45.28%	46.73%	47.73%	47.19%
	Percentage Gain	3.74%	3.27%	3.91%	3.07%	1.62%	0.62%	1.16%
Producer's Accuracy	58.67%	48.33%	47.33%	48.00%	46.33%	50.00%	49.00%	47.67%
	Percentage Gain	10.33%	11.33%	10.67%	12.33%	8.67%	9.67%	11.00%