PEMANFAATAN DATA KEHATI SKALA BESAR

Safran Yusri

Yayasan TERANGI

PENDAHULUAN

- Dampak kerusakan lingkungan
- Menjaga pemanasan global < 2°C
- Skala pengelolaan besar
- Pengelolaan multidisiplin
- Ekolog perlu menjawab tantangan tersebut dengan cepat dan akurat

TERANGI

BIODIVERSITY ♥ BIG DATA ♥ CLOUD COMPUTING ♥ MACHINE LEARNING

- Data KEHATI banyak dihasilkan berbagai pihak
- Teknologi penyimpanan meningkat pesat
- Variasi format dan inkonsistensi sudah bisa diatasi
- Kemampuan komputasi memungkinkan analisis data skala besar
- Sumberdaya komputasi dapat dikelola, dibagi, serta ditingkatkan sesuai skala yang dibutuhkan
- Pengembangan algoritma pembelajaran mesin untuk mengubah data skala besar menjadi informasi

TERANGI

WOODGROVE BANK

PRIORITAS KERUANGAN

- Bagaimana cara menentukan prioritas ruang untuk konservasi dan pemanfaatan lainnya?
- Tekanan tinggi dari berbagai pihak yang berkepentingan
- Berpacu dengan waktu dengan kerusakan alam dan kepunahan spesies
- Butuh data banyak dan sebarannya secara spasial dimengerti
- Survei lapangan mahal dan Indonesia sangat luas

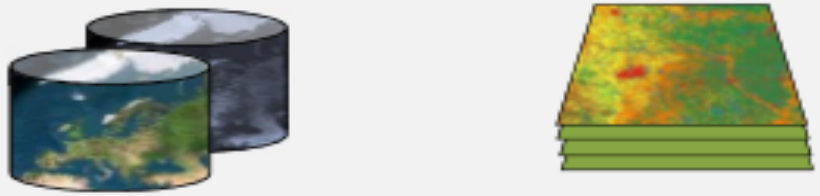
TERANGI

The workflow of spatial prioritization in Zonation

TERANGI

Environmental data sets

Habitats / environmental data



Remote sensing & GIS analyses

Species or BD features

Species distribution modeling

Ecosystem services

Modeling and expert elicitation

Setting the objectives

↓

Ecological model

Preprocessing of data

↓

Spatial prioritization

Result visualization

Verification

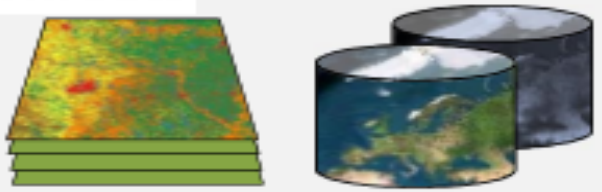
Interpretation

↓

Recommendations for action

Societal data sets

Threats



GIS, registry data, statistics, simulations

Costs

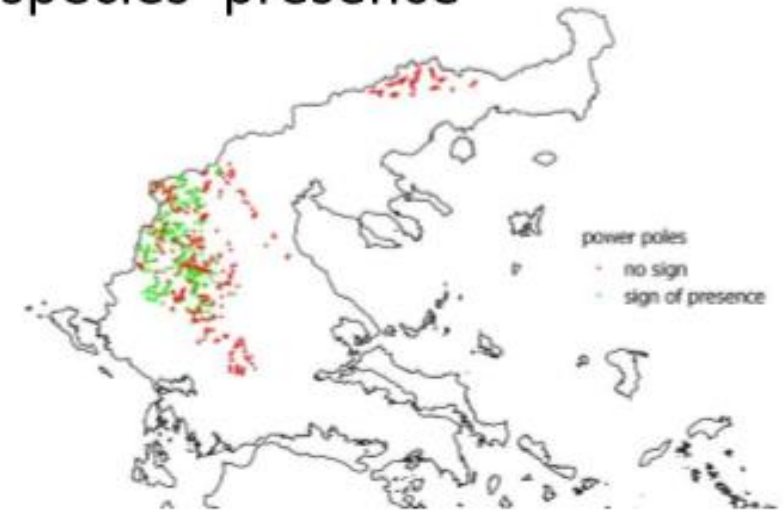
GIS, registry data, statistics, simulations

Current plans

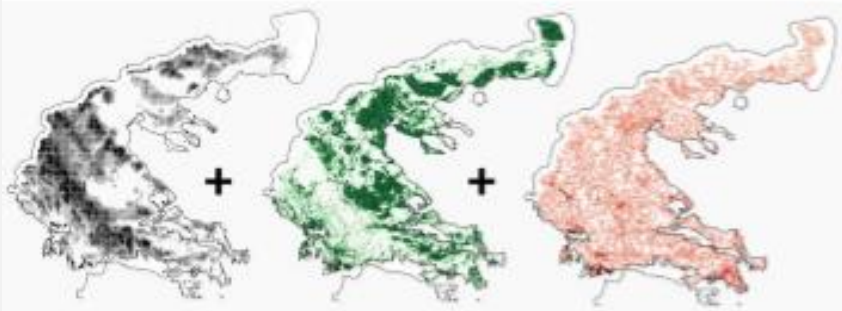
Planning documents

SPECIES DISTRIBUTION MODELLING

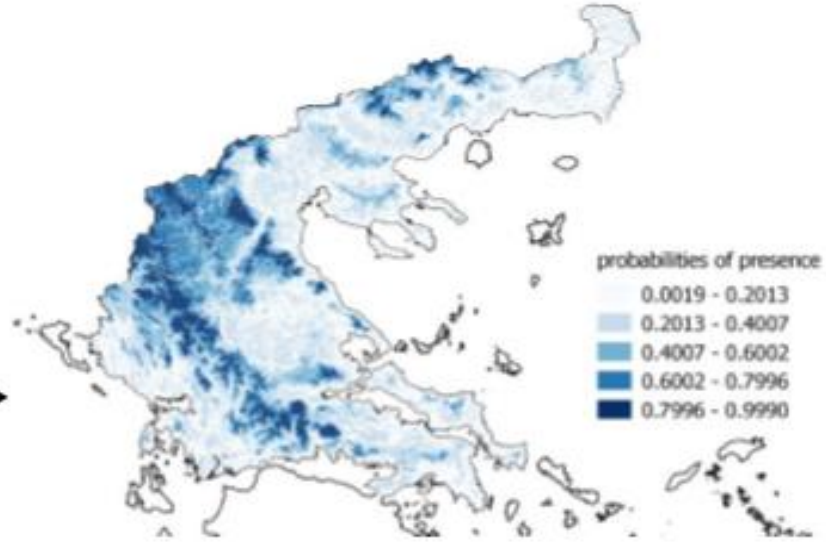
Partial information on the species' presence



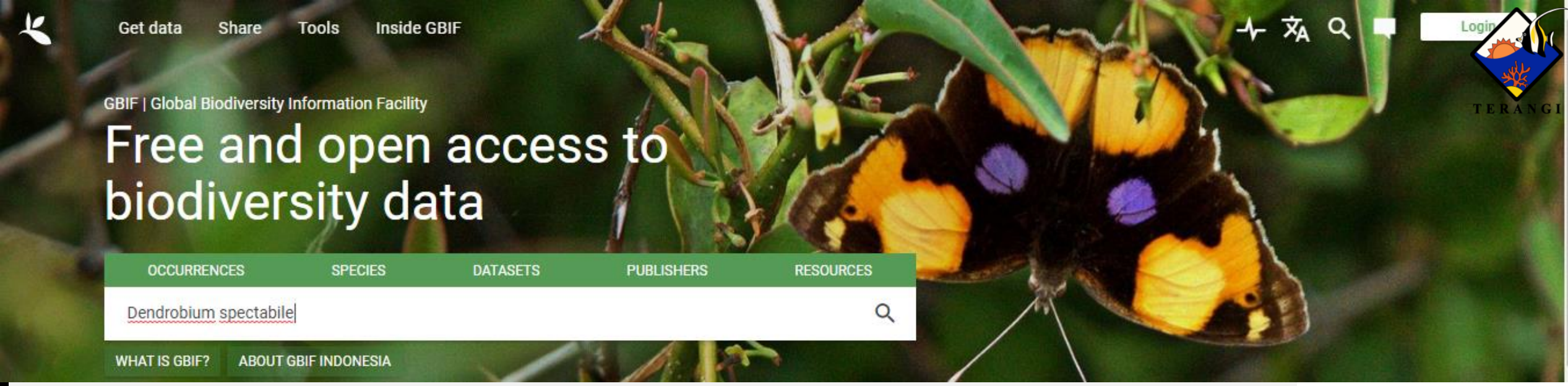
+



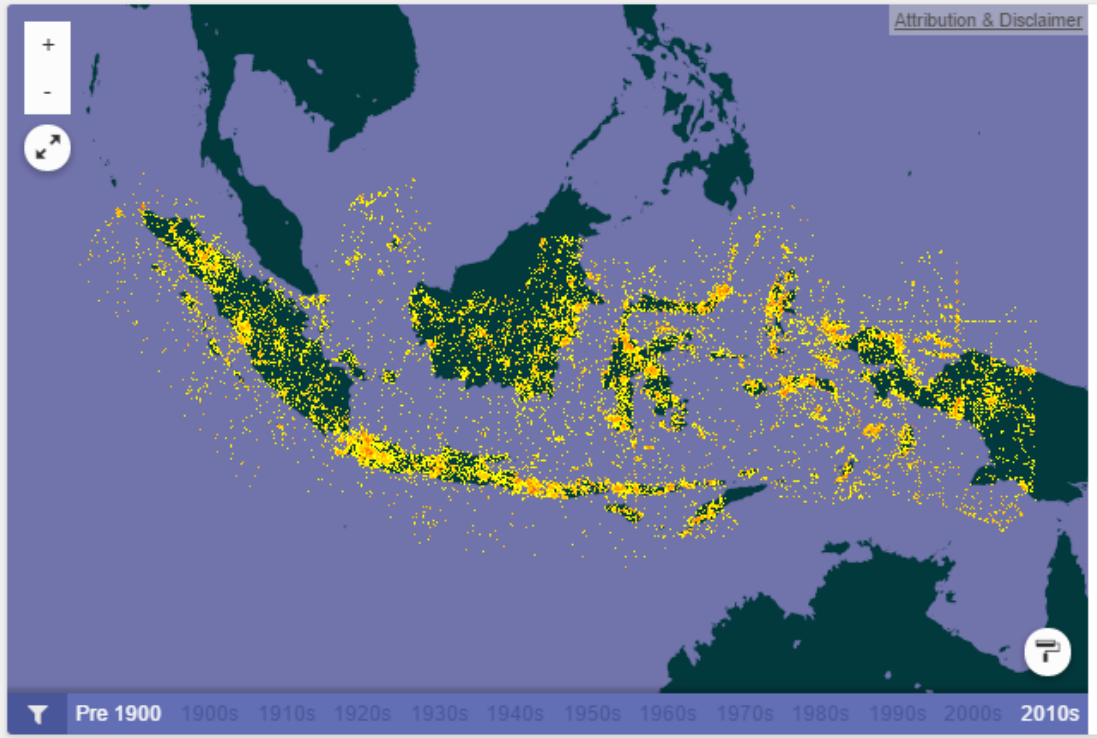
Environmental variables



Probabilities of presence in the whole area of interest



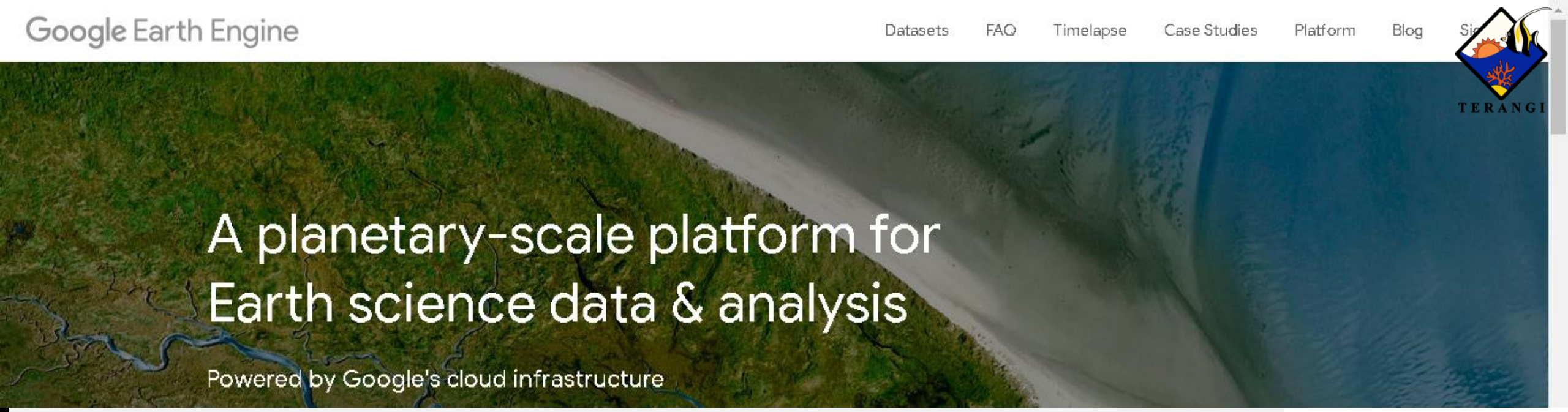
- Global Biodiversity Information Facility (GBIF) merupakan insiatif global pengumpulan data bersama
- Menyediakan 1,5 juta data kemunculan di Indonesia
- https://gbif.org



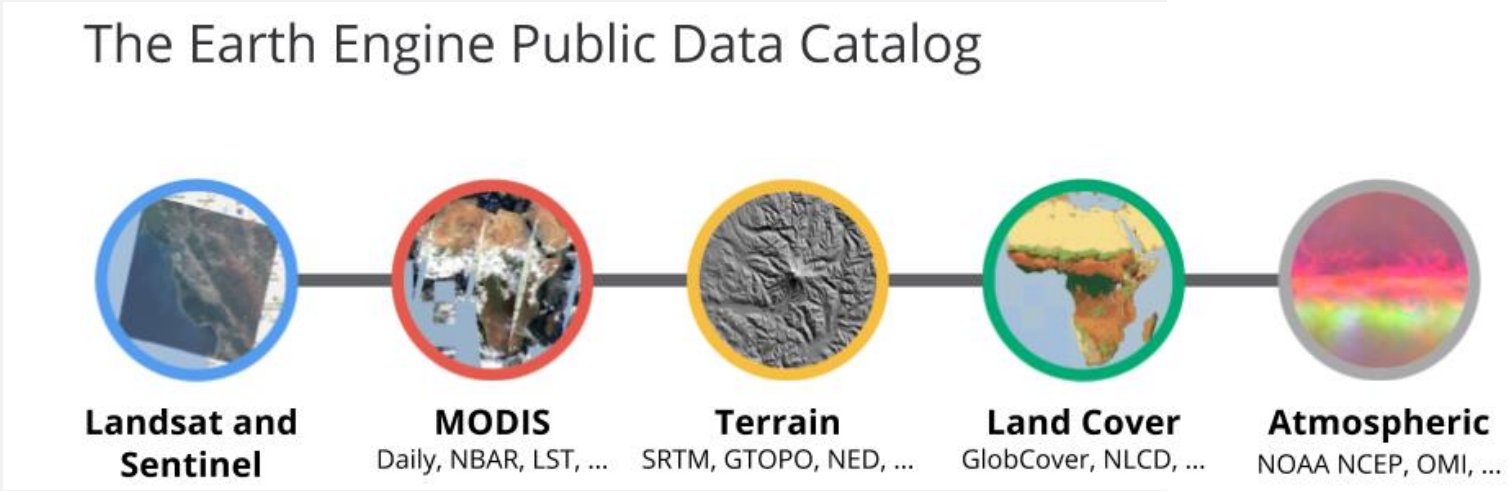
Data about Indonesia

- 814 occurrence datasets with 1,567,566 records.
- No checklist datasets.
- No metadata-only datasets relevant to Indonesia.
- 37 countries contribute data about Indonesia.

View records shown on the map



- Google Earth Engine (GEE) menyediakan beberapa terabit data pengamatan kebumian
- Wahana pemetaan dan analisis citra berbasis komputasi awan
- https://earthengine.google.com



WOODGROVE BANK

STUDI KASUS

TERANGI

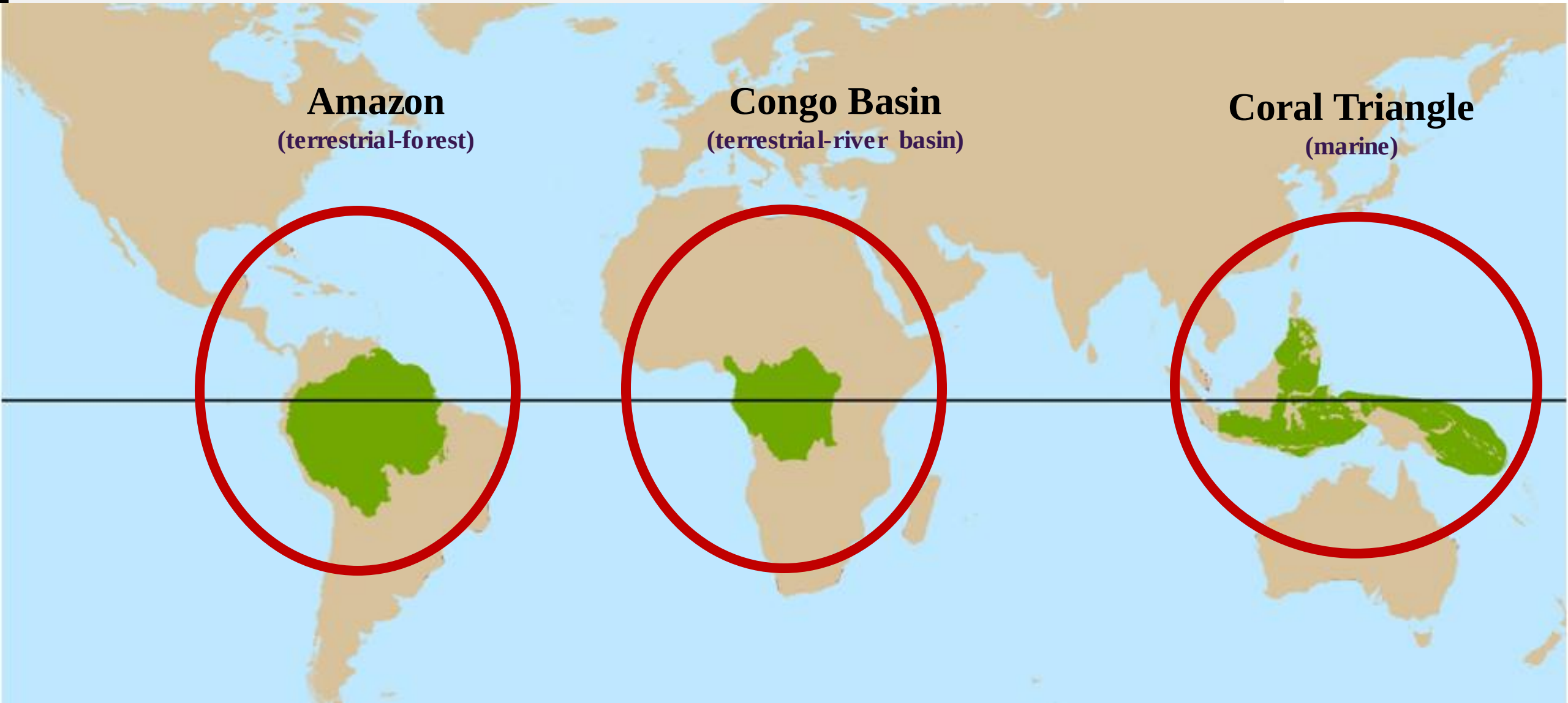
SPATIAL MODELLING OF SCLERACTINIAN CORAL DISTRIBUTION IN INDONESIA

MSc in IT for NRM

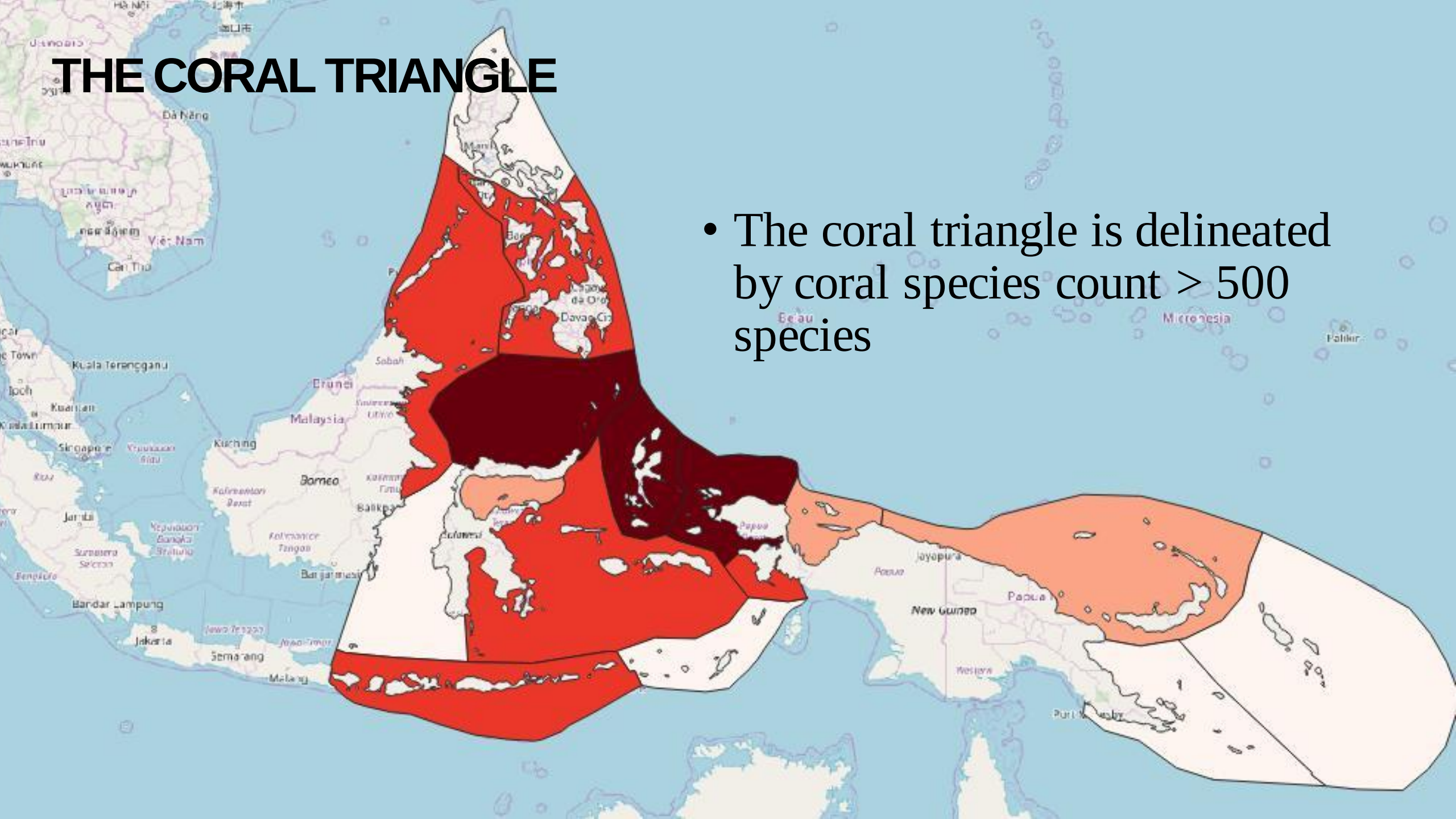
Safran Yusri | G051160021

INTRODUCTION

TERANGI



THE CORAL TRIANGLE



- The coral triangle is delineated by coral species count > 500 species

CORAL CONSERVATION IN INDONESIA

- Coral reefs extent : 2.5 million ha (BIG 2013)
- 569 species of hard corals (Suharsono 2017)
- Corals are traded for ornamentals
- Increasing local and global threats
- Managed through coastal zonation and quota

Euphyllia spp.

RESEARCH OBJECTIVE

- To develop genera distribution model of hard corals (Scleractinia) in Indonesia based on the available presence records and publicly available remote sensing imageries and oceanographic models using Maximum Entropy.

RESEARCH LOCATION

TERANGI

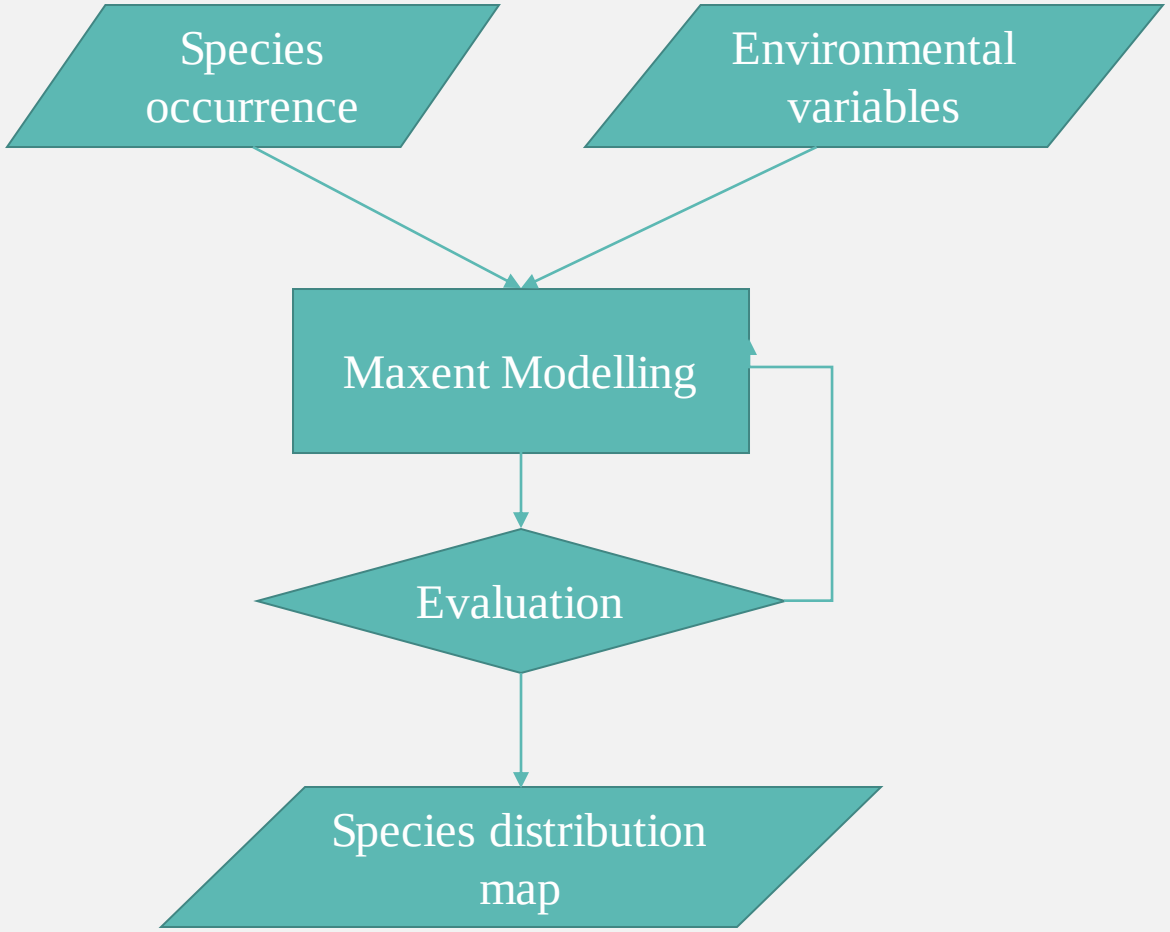
MAXIMUM ENTROPY: THE BASIC IDEA

- Goal: estimate pi
- Choose p with maximum entropy (or "uncertainty") subject to the constraints (or "evidence").
- $H = -\sum_{i=1}^{N} p_i \ln(p_i)$
- and p(y=x|z) = fx(z)p(y=x) / f(z)

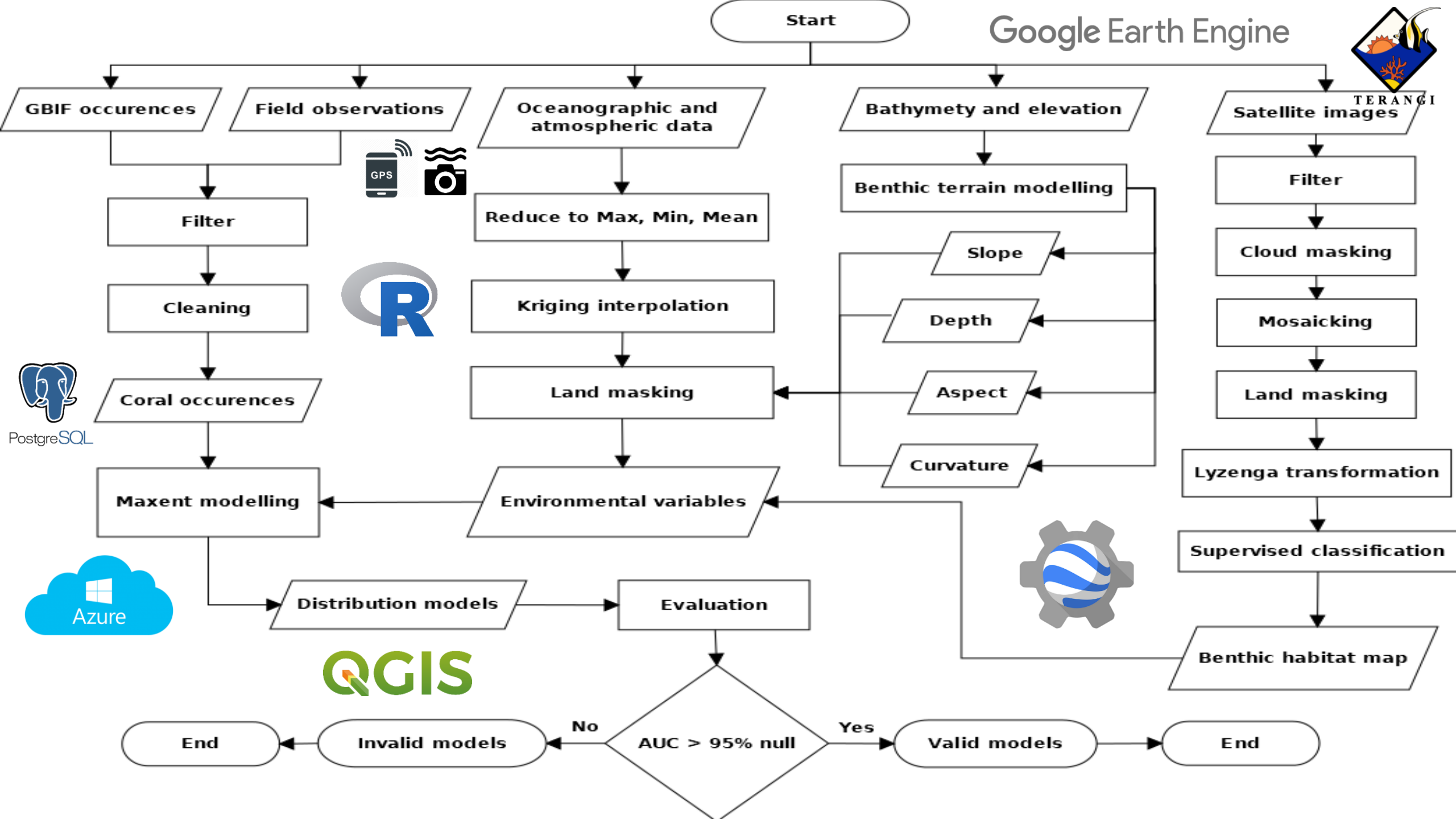
TERANGI

WOODGROVE BANK

BASIC FLOWCHART

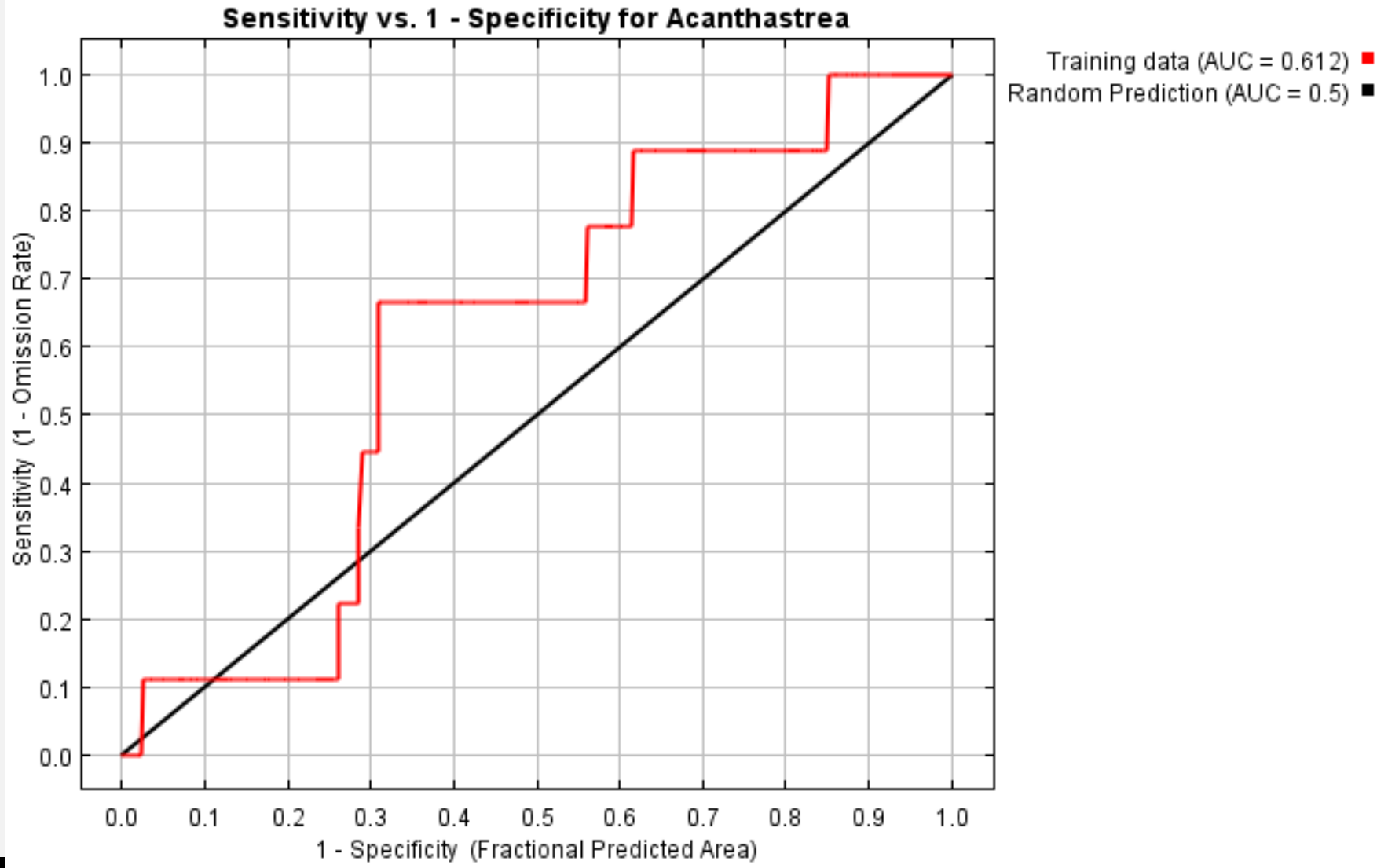


- Target: 73 coral genera (Suharsono 2017)
- Environmental variables
  - Substrate type
  - Chlorophyll-A
  - Bathymetry
  - Sea surface temperatures
  - Particulate organic carbon
- Species occurrence
  - Specimens from GBIF
  - Observations


Start
Google Earth Engine
TERANGI
GBIF occurences
Field observations
Oceanographic and atmospheric data
Bathymety and elevation
Satellite images
GPS
Filter
Reduce to Max, Min, Mean
Benthic terrain modelling
Filter
Cleaning
Kriging interpolation
Slope
Depth
Aspect
Curvature
Cloud masking
Mosaicking
Land masking
Coral occurences
Land masking
Lyzenga transformation
PostgreSQL
Maxent modelling
Environmental variables
Supervised classification
Azure
Distribution models
Evaluation
QGIS
Benthic habitat map
End
Invalid models
No
AUC > 95% null
Yes
Valid models
End

MODEL EVALUATION AND VALIDATION WITH AUC

- Area under the curve (AUC)
- AUC value close to 1 = Good prediction
- Compared to 99 null model's AUC
- AUC > top 5% of null model's AUC = Significant and valid model

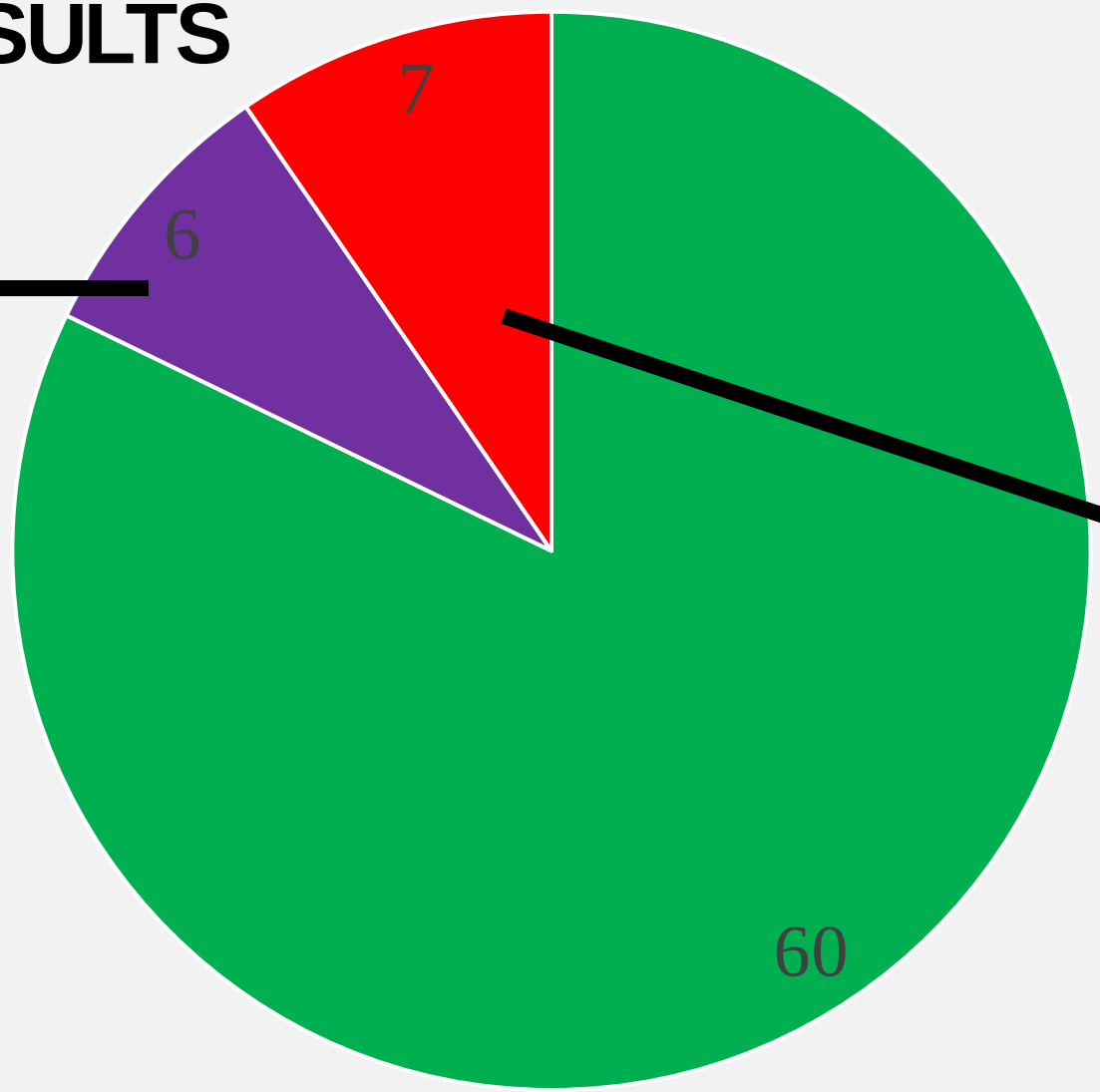


Results and Discussions

ENVIRONMENTAL PARAMETERS

- 32 rasters generated
- Benthic geomorphology (2014)
  - Depth: 0 – 6827 m
  - Slope: 0 – 34.33°
  - Aspect: 0 – 359.99°
  - Curvature: 0 – 0.94.
- Water energy (2002 – 2017)
  - Mean sea surface elevation: 0 – 0.85 m
  - Mean water velocity: 0 – 4 m/s
- Water quality (2002 – 2017)
  - Mean particulate organic carbon: 25.31 – 953.47‰
  - Mean clorophyll-A concentration: 0.05 – 13.63‰.
  - Mean surface salinity: 20.09 – 35.32‰.

MODELLING RESULTS



1. *Balanophyllia*
2. *Blastomussa*
3. *Madracis*
4. *Pleisiastrea*
5. *Pseudosiderastrea*
6. *Stylocoeniella*.

1. *Catalaphyllia*
2. *Dendrophyllia*
3. *Heterocyathus*
4. *Oulastrea*
5. *Palauastrea*
6. *Scapophyllia*
7. *Zoopilus*

Valid Insignificant Can not be modeled

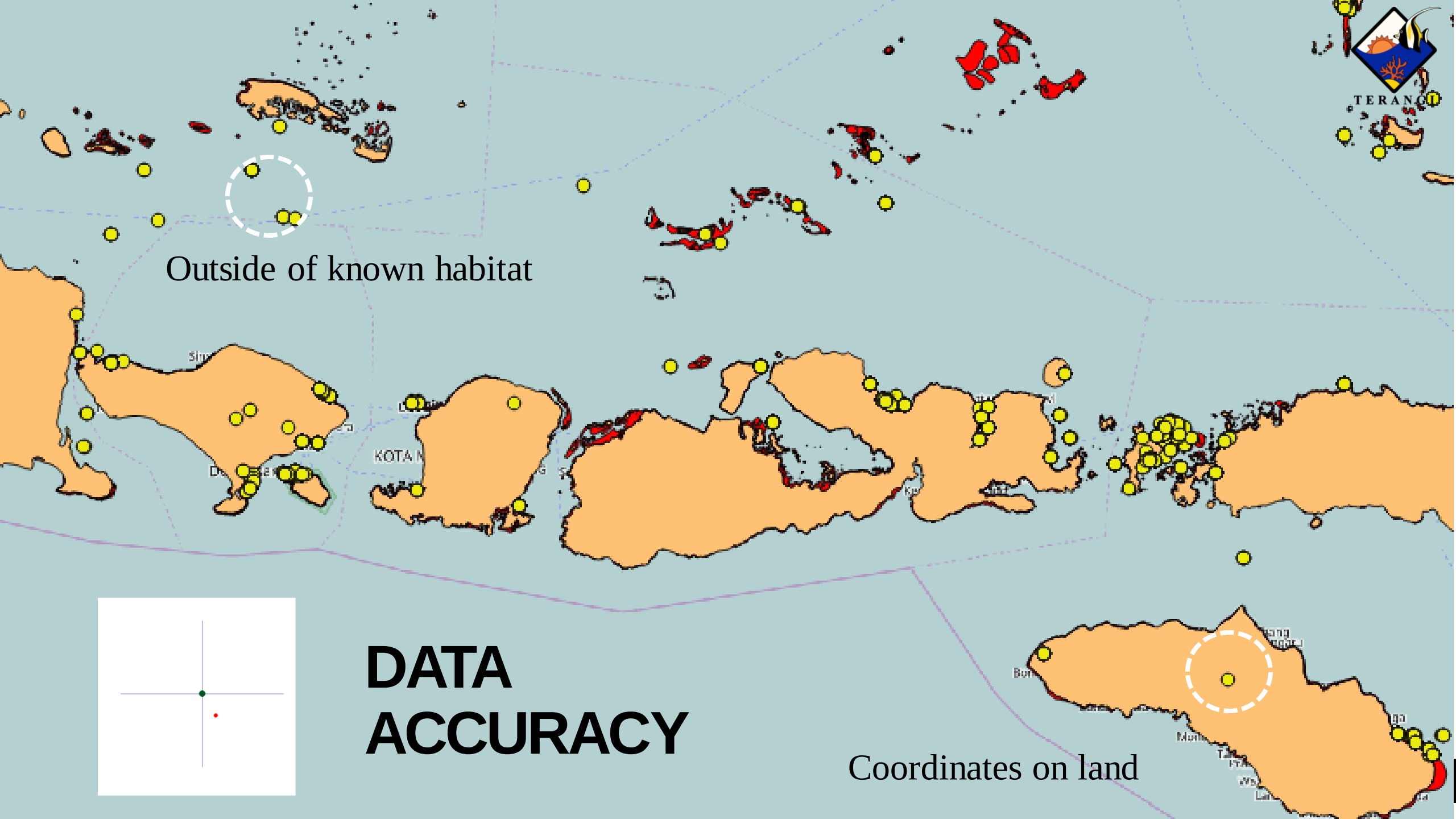
TERANGI

WOODGROVE BANK

UNSUCCESSFUL MODELS

- Even though the insignificant models have high AUCs, cross validation and low sample size can make even random null models' AUC values of 0.85 to be unusual (Merckx et al. 2011).
- All unsuccessful models have low sample size
- Due to low records and low data accuracy

Genus	AUC		Sample
	Test	95% null	
Balanophyllia	0.9452	0.9792	6
Blastomussa	0.8406	0.9999	3
Madracis	0.9287	0.9989	4
Plesiastrea	0.9731	0.9952	5
Pseudosiderastrea	0.8751	0.9974	5
Stylocoeniella	0.9233	0.9708	7
Cataphyllia	0.5	N/A	1
Dendrophyllia	0.9999	N/A	2
Heterocyathus	0.9743	N/A	2
Oulastrea	0.9695	N/A	5
Palauastrea	0.5	N/A	1
Scapophyllia	N/A	N/A	1
Zoopilus	N/A	N/A	0


Outside of known habitat
TERANGI
KOTA
DATA
ACCURACY
Coordinates on land

UNSUCCESSFUL MODELS

- These genera except for *Oulastrea*, are generally rare in nature
- *Oulastrea* are usually found in turbid water
- *Dendrophyllia*, *Madracis*, and *Stylocoeniella* are cryptic and usually found in the shade of rocks and caves
- *Catalaphyllia*, *Heterocyathus*, *Balanophyllia*, and *Blastomussa* are mostly found in deeper water (more than 20 m)
- These areas are where coral reef monitoring programs seldom reach
- These genera require targeted surveys and the use of another modelling method

TERANGI

WOODGROVE BANK

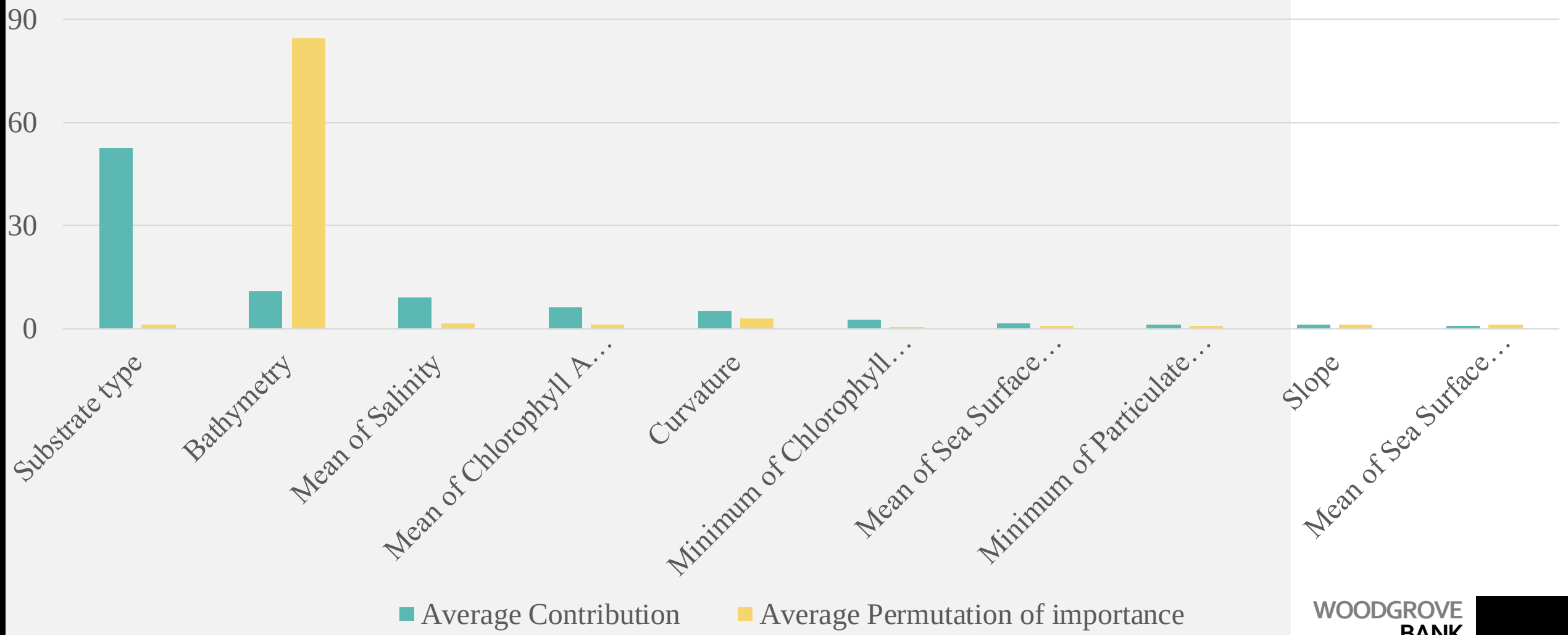
SUCCESSFUL MODELS

- 60 genera can be modelled successfully and tested significant against the upper 5% of the null distribution's AUC.
- Successful model can be built using as low as 7 + 1 training samples with cross validation (*Gardineroseris* distribution model).
- Model with lowest training sample and null AUC below 0.9 is *Alveopora* distribution model with 33 samples with cross validation

TERANGI

WOODGROVE BANK

ENVIRONMENTAL VARIABLES CONTRIBUTION TOWARDS MODEL GAIN

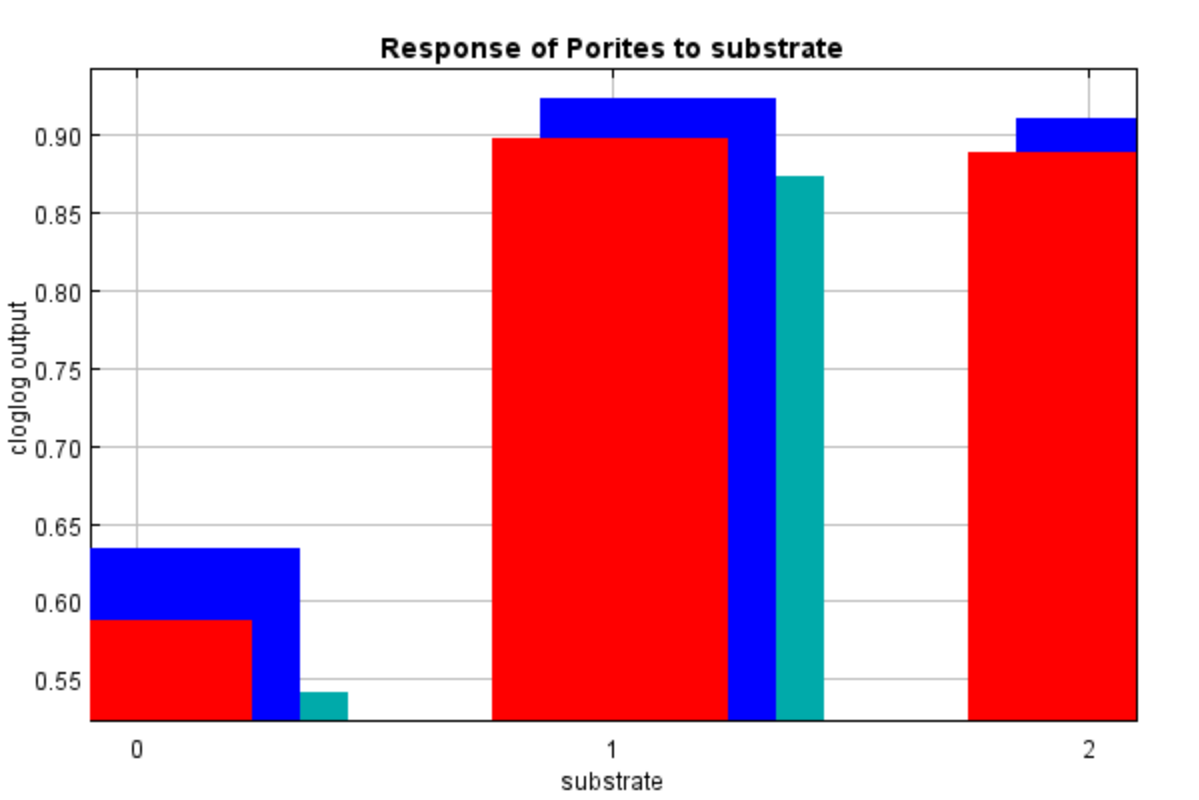


TERANGI

WOODGROVE BANK

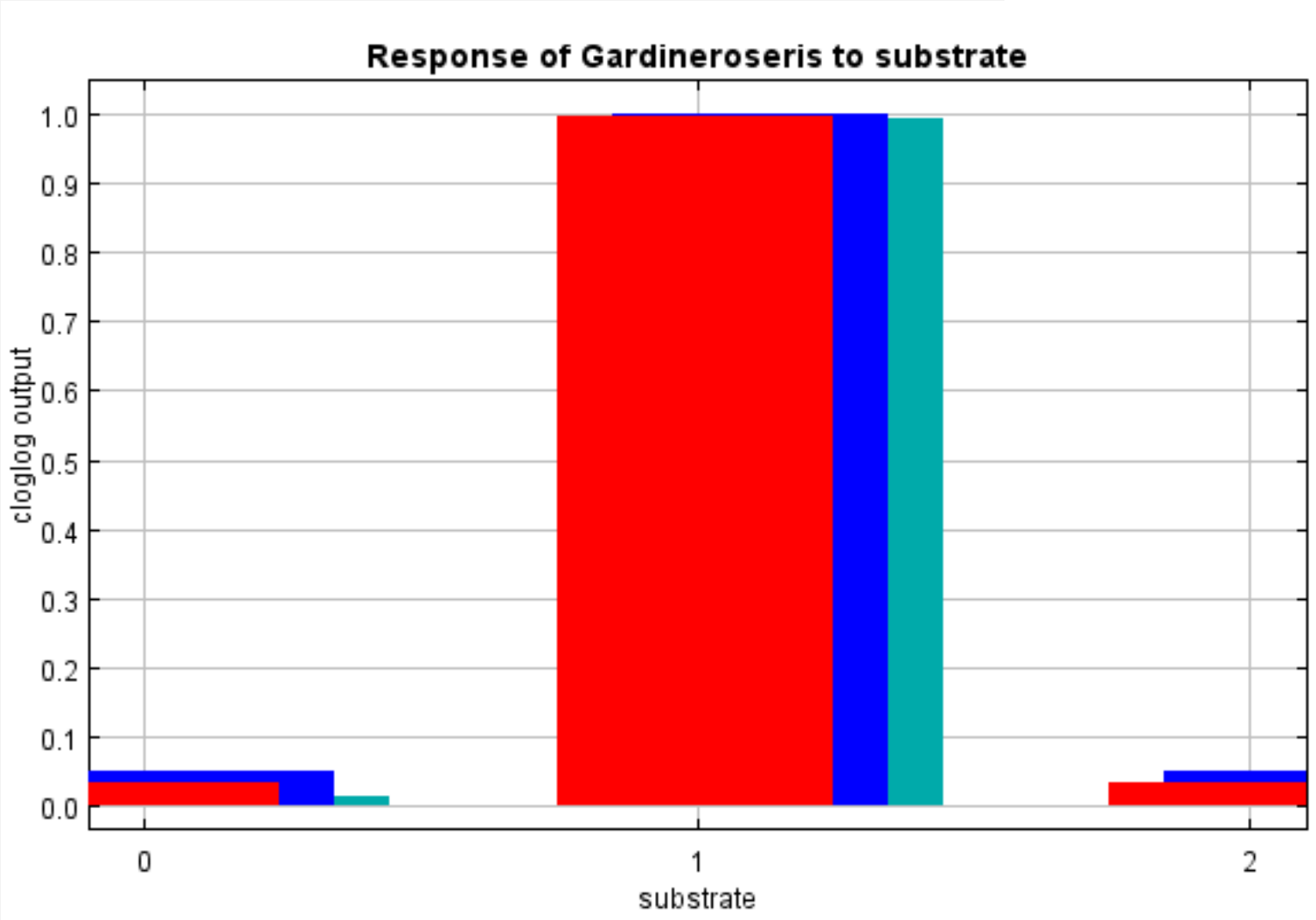
SUBSTRATE TYPE

- *Porites* preferred coral dominated reefs (1) than sand dominated reefs (2) but can live in both



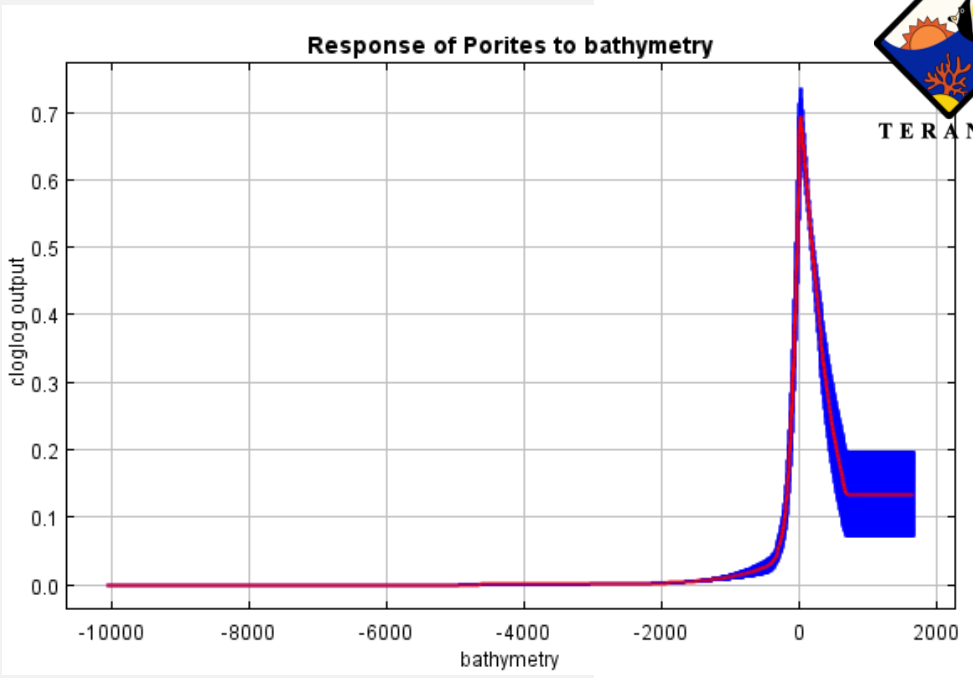
GARDINEROSERIS

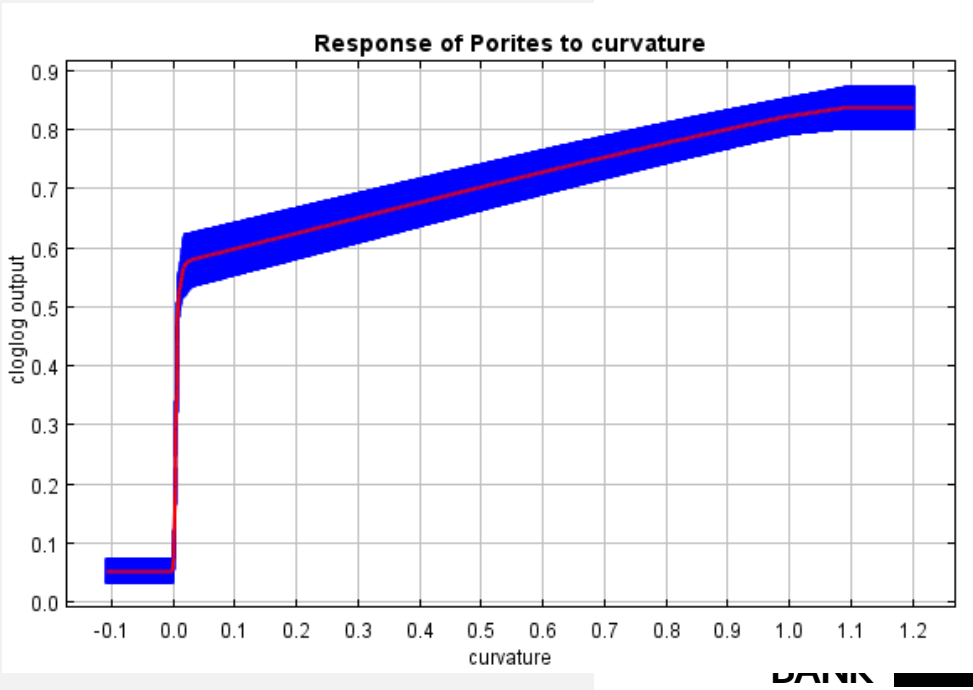
- Almost entirely dependent on coral dominated reefs



BATHYMETRY AND CURVATURE

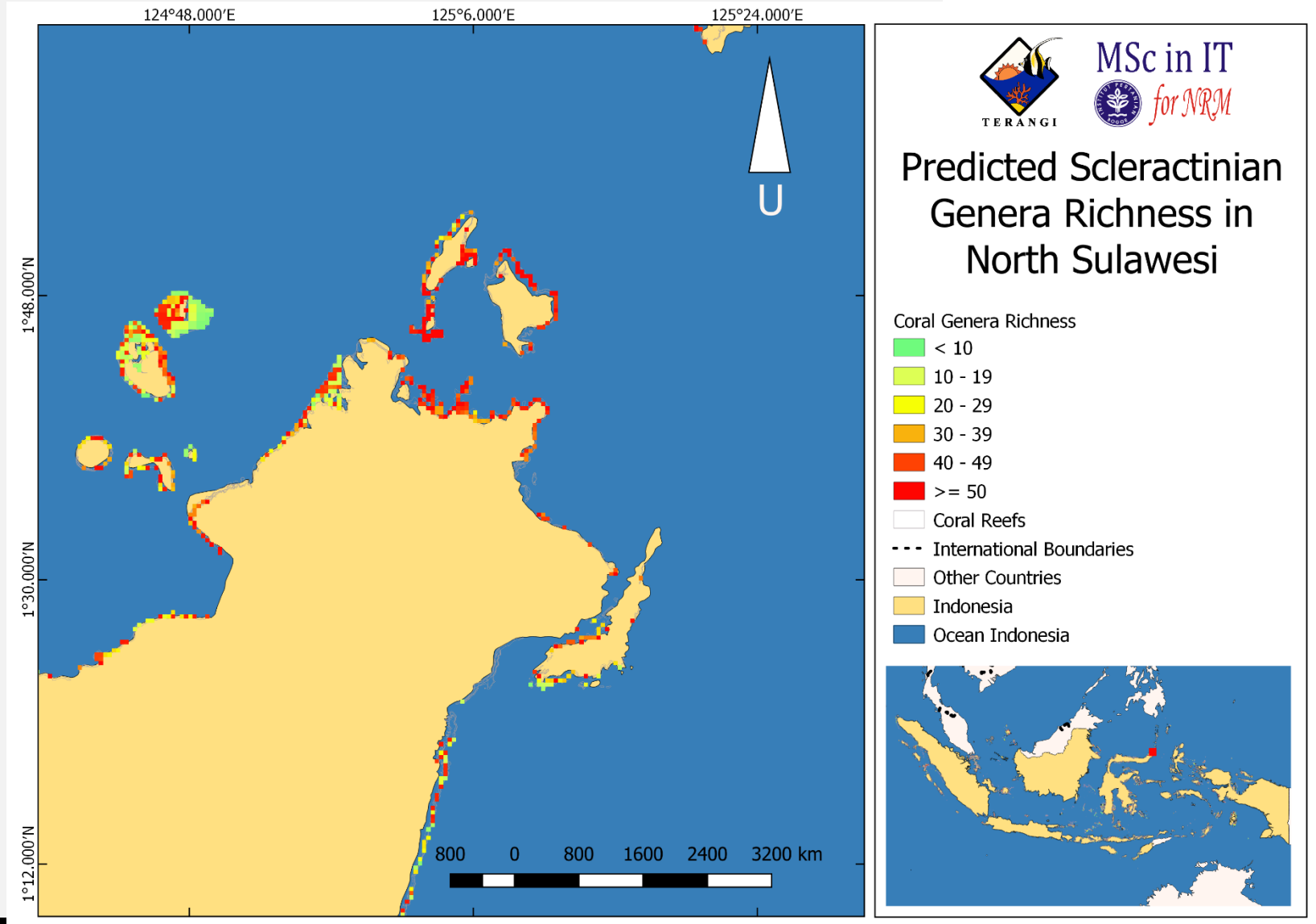
- Most habitable is shallow water
- Most preferred is reefs with curvature from 0 - 1°



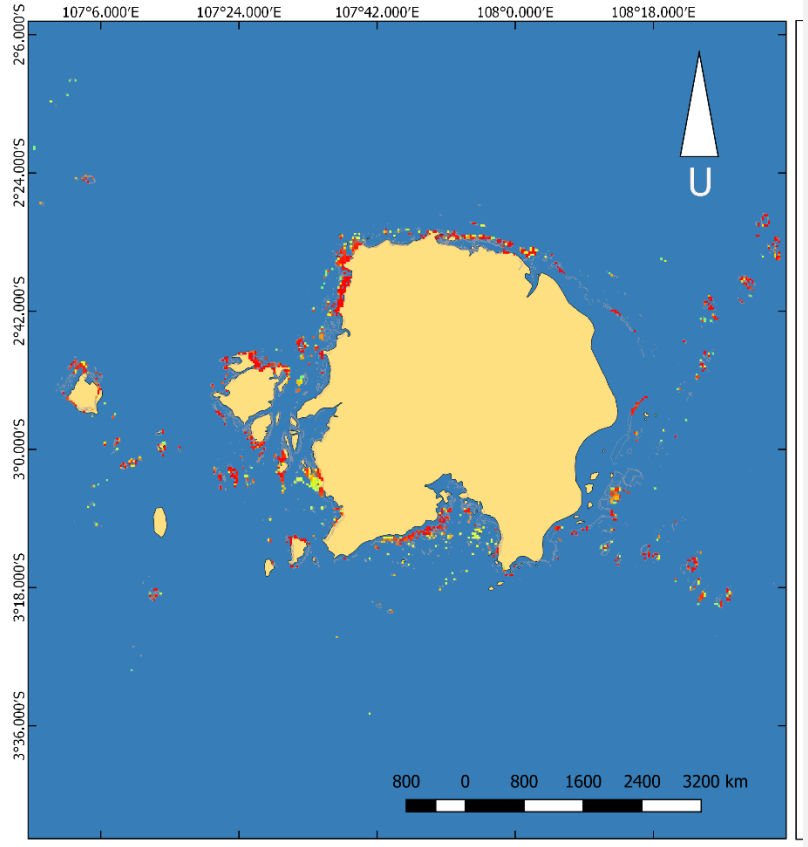


GENERA RICHNESS

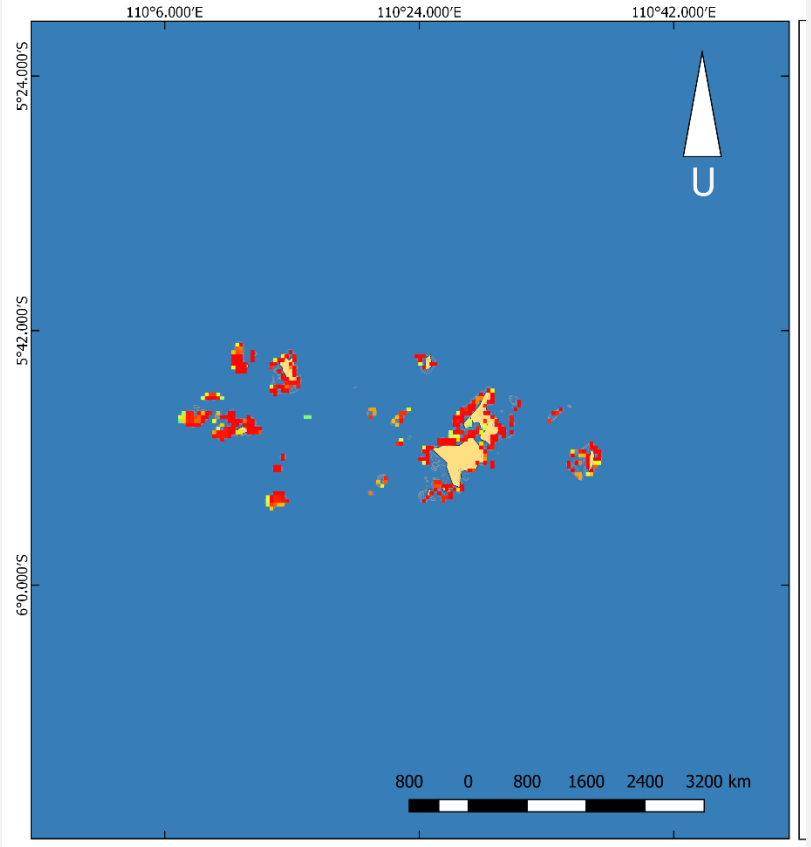
- The maximum genera richness in a location is 58 genera
- Can be found in several places, such as the middle part of Seribu Islands chain, Bangka, Gangga, and Lembeh Island in North Sulawesi, and Raja Ampat.



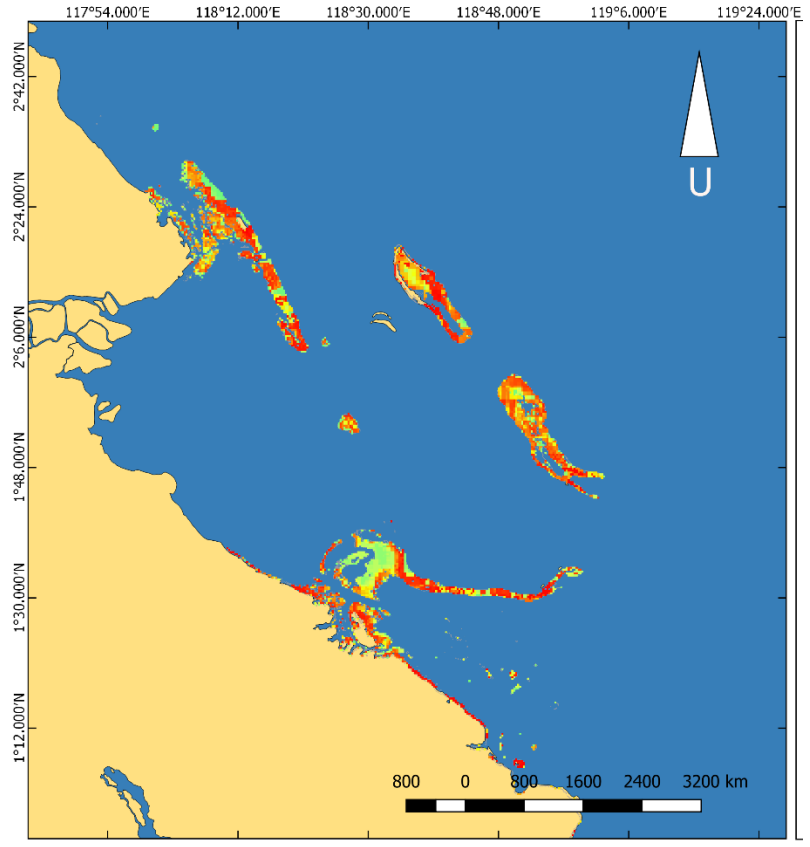
GENERA RICHNESS



Belitung Island



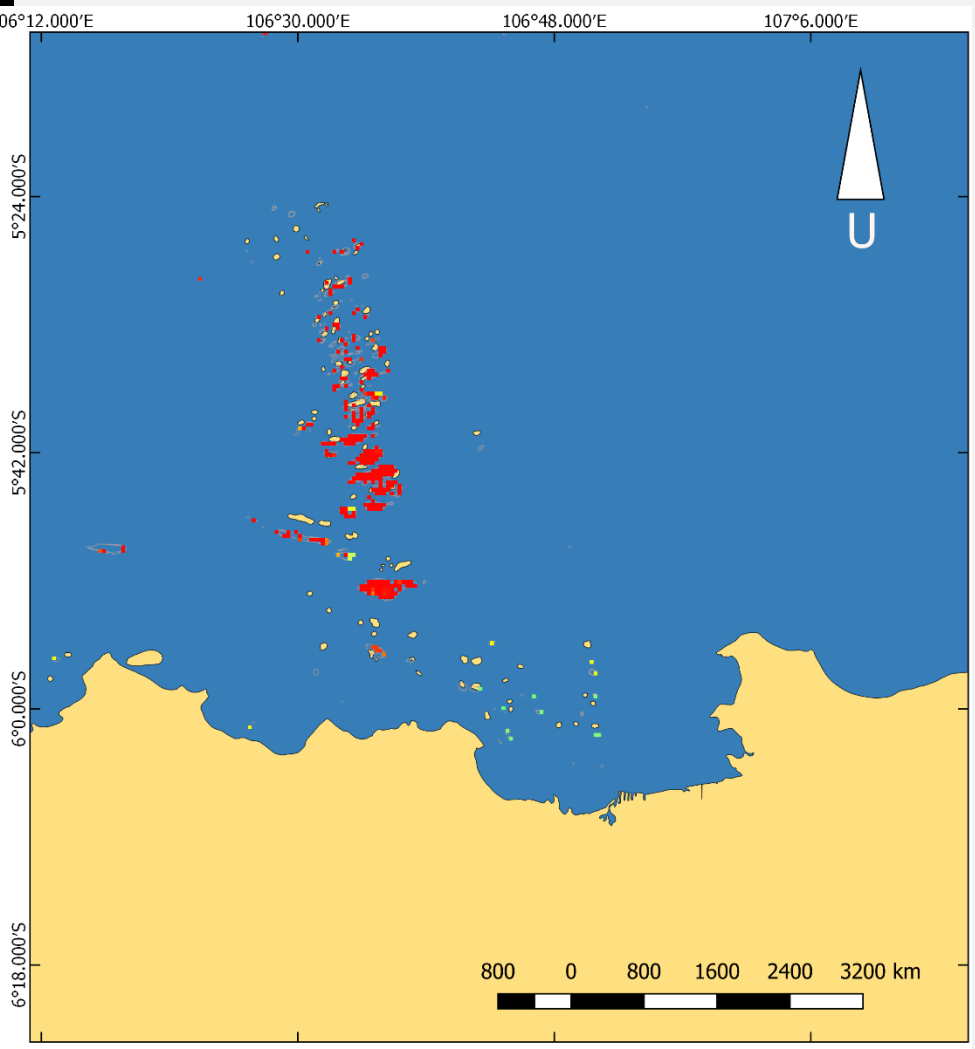
Karimun Jawa Islands

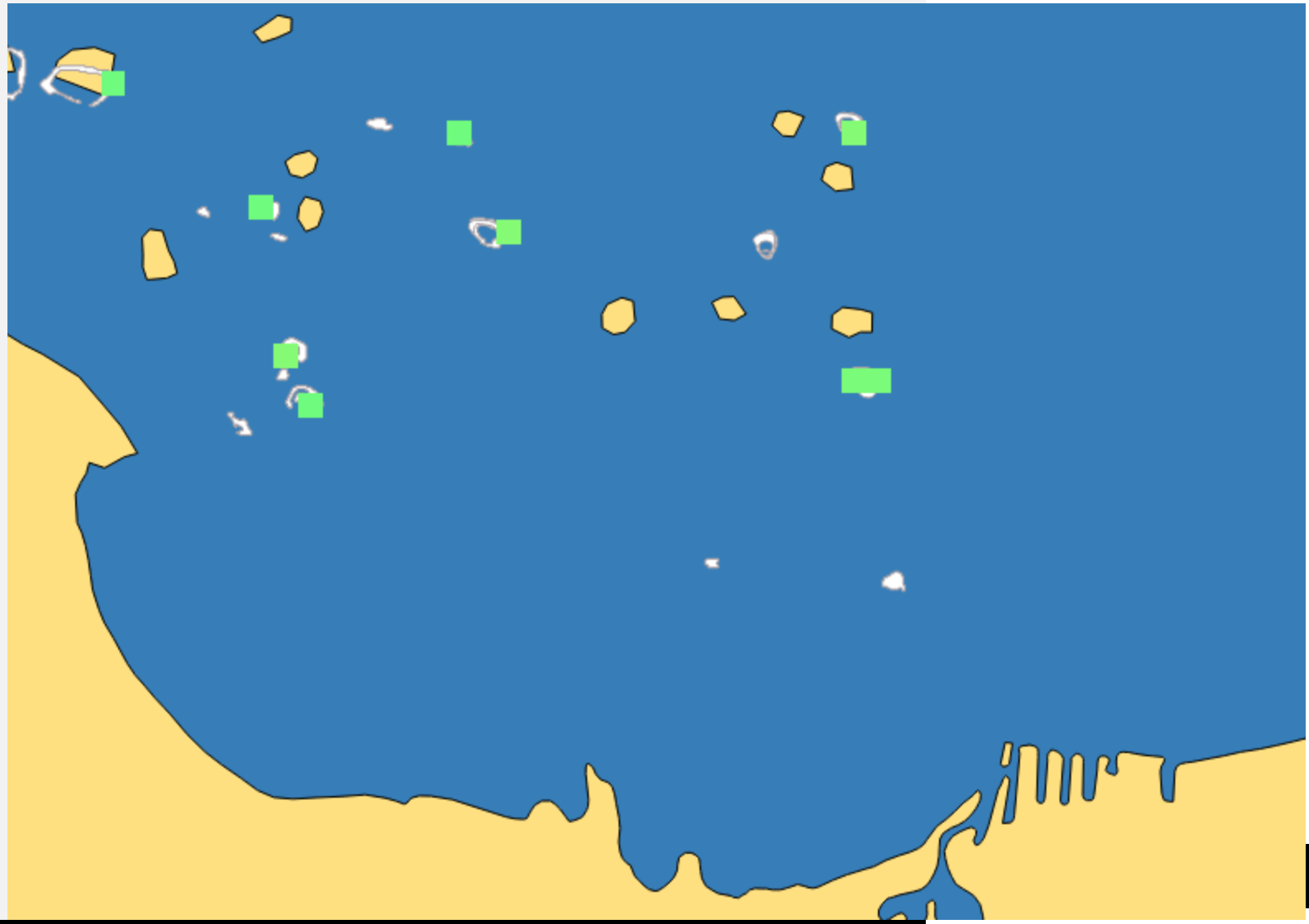


Derawan Islands

GENERA RICHNESS IN SERIBU ISLAND

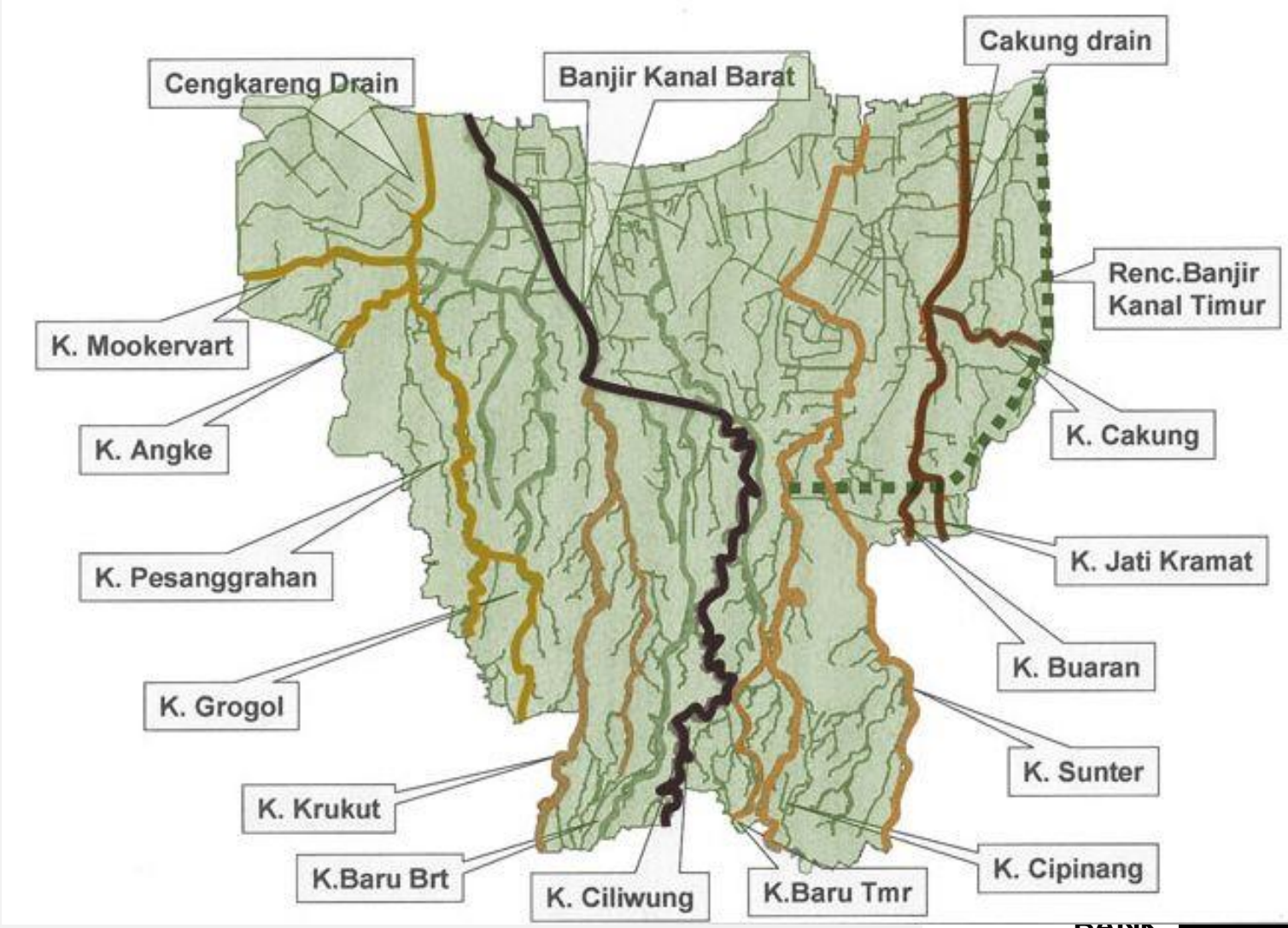
TERANGI





CORAL ABSENCE IN JAKARTA BAY

- No preferable substrate type
- Both coral reefs and coral-sand reefs are smothered with sediment
- 13 polluted rivers dumping their sediment, heavy metal, liquid and solid waste altogether, this area has high sedimentation that made life hard for corals and they are unable to photosynthate
- Sediment accumulation rate is up to 0.852 cm/year
- Jakarta will experience coastal reclamation, that will increase the benthic sediment thickness to 2.49 m compared to 0.84 m today



CONCLUSIONS

- Scleractinian distribution models can be built using openly available environmental data and occurrence records.
- Successful model can be built using as low as 7 + 1 training samples with cross validation, but recommended to be at least 33.
- Rare and cryptic species requires targeted surveys.
- Improvement in spatial accuracy can be done with higher resolution input

TERANGI

WOODGROVE BANK

THANK YOU VERY MUCH

TERANGI

MSc in IT for NRM

- Questions
- Critics
- Suggestions