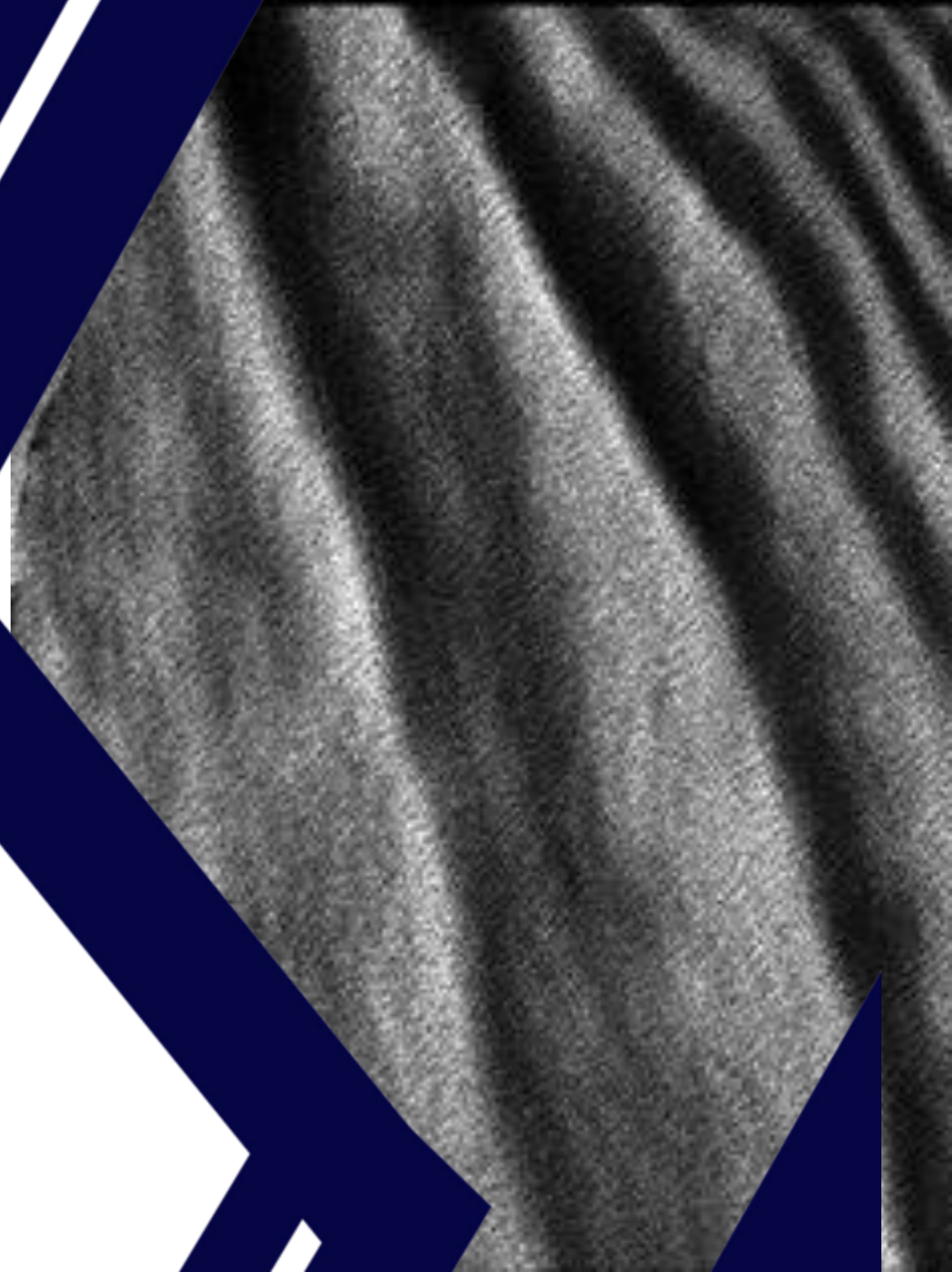
JuliaEO
Global Workshop on Earth Observation with Julia 2025

AIR CENTRE
ATLANTIC INTERNATIONAL RESEARCH CENTRE

Internal Waves Service

Adriana Ferreira
João Gonçalves
Iúri Diogo
João Pinelo



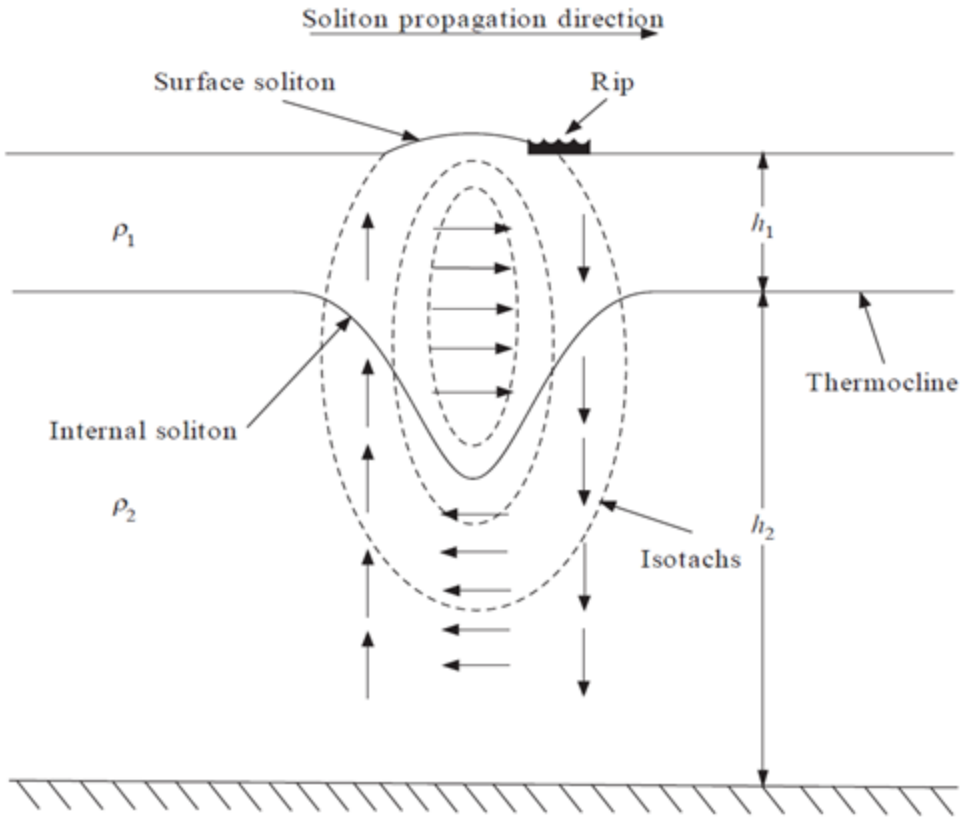
06/01/2025

Purpose and Overview

"At such times, under an abated sun; afloat all day upon smooth, slow heaving swells; seated in his boat, light as a birch canoe; and so sociably mixing with the soft waves themselves, that like hearth-stone cats they purr against the gunwale; these are the times of dreamy quietude, ***when beholding the tranquil beauty and brilliancy of the ocean's skin, one forgets the tiger heart that pants beneath it; and would not willingly remember, that this velvet paw but conceals a remorseless fang.***"

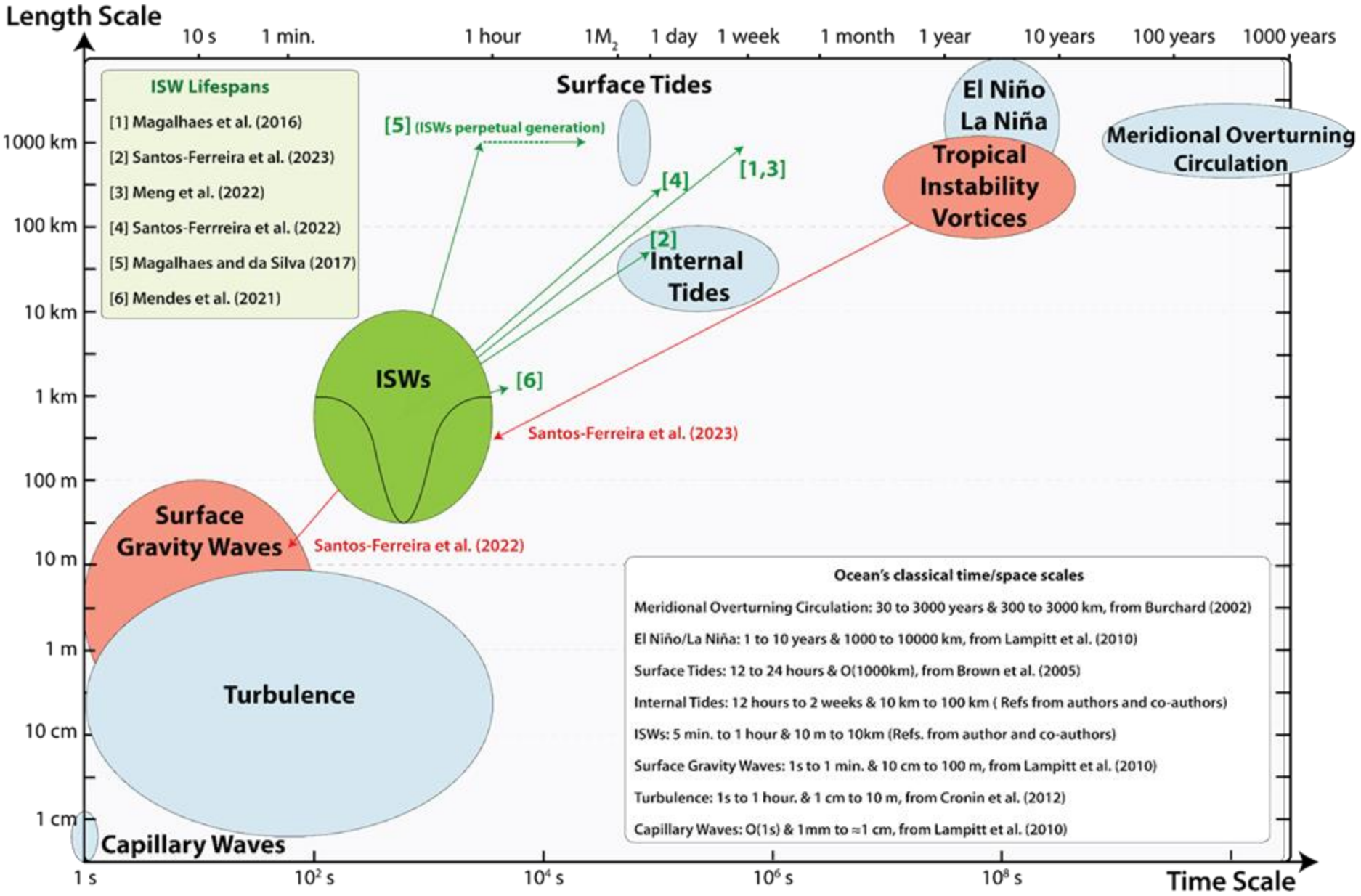
Herman Melville, in Moby Dick

What are Ocean Internal Waves?



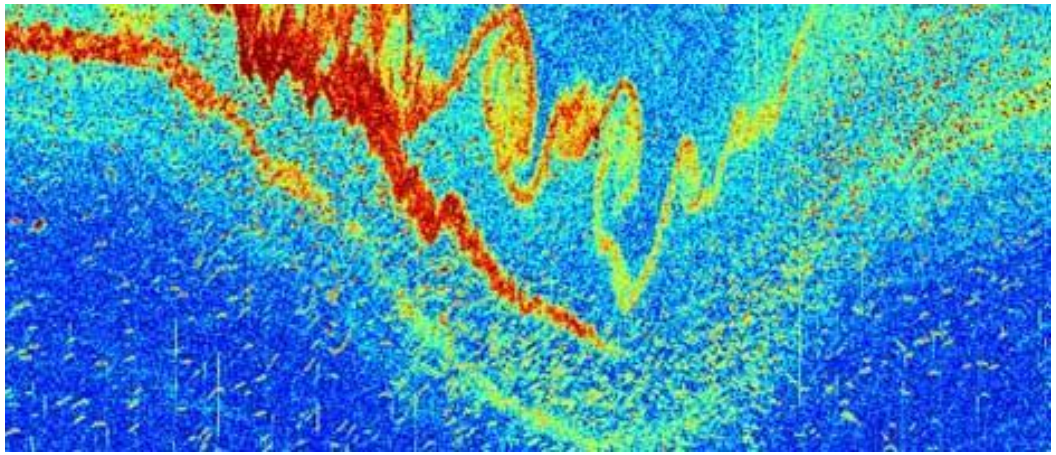
Sea surface roughness pattern generated by a linear internal wave (Osborne & Burch, 1980).

August 30, 2006 (3:30 PM local time), captured by José da Silva in Cape Cod (Massachusetts, USA).

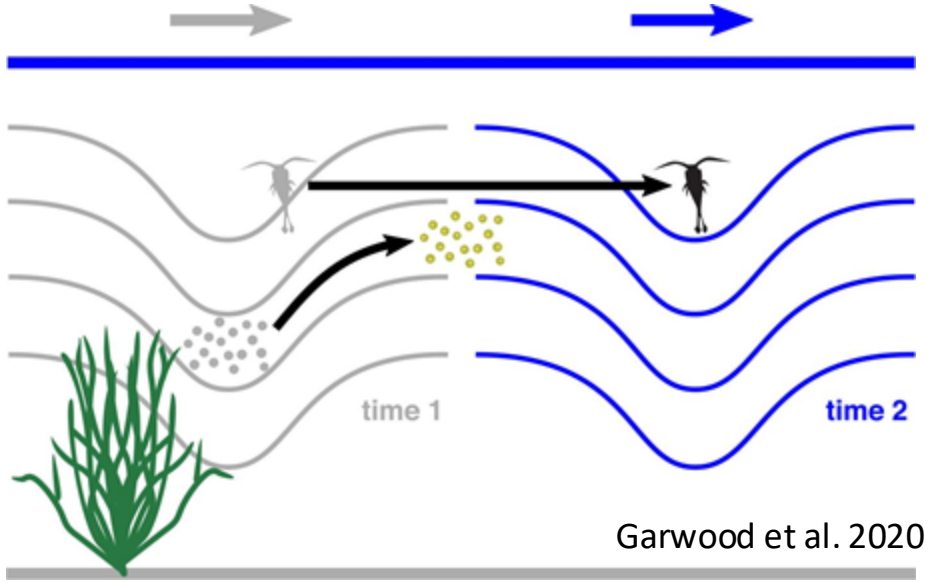

Length Scale
10 s
1 min.
1 hour
1M2
1 day
1 week
1 month
1 year
10 years
100 years
1000 years
1000 km
100 km
10 km
1 km
100 m
10 m
1 m
10 cm
1 cm
ISW Lifespans
[1] Magalhaes et al. (2016)
[2] Santos-Ferreira et al. (2023)
[3] Meng et al. (2022)
[4] Santos-Ferrreira et al. (2022)
[5] Magalhaes and da Silva (2017)
[6] Mendes et al. (2021)
Surface Tides
[5] (ISWs perpetual generation)
[4]
[1,3]
[2]
Internal Tides
El Niño
La Niña
Tropical Instability Vortices
Meridional Overturning Circulation
ISWs
[6]
Santos-Ferreira et al. (2023)
Surface Gravity Waves
Santos-Ferreira et al. (2022)
Turbulence
Capillary Waves
Ocean's classical time/space scales
Meridional Overturning Circulation: 30 to 3000 years & 300 to 3000 km, from Burchard (2002)
El Niño/La Niña: 1 to 10 years & 1000 to 10000 km, from Lampitt et al. (2010)
Surface Tides: 12 to 24 hours & O(1000km), from Brown et al. (2005)
Internal Tides: 12 hours to 2 weeks & 10 km to 100 km (Refs from authors and co-authors)
ISWs: 5 min. to 1 hour & 10 m to 10km (Refs. from author and co-authors)
Surface Gravity Waves: 1s to 1 min. & 10 cm to 100 m, from Lampitt et al. (2010)
Turbulence: 1s to 1 hour. & 1 cm to 10 m, from Cronin et al. (2012)
Capillary Waves: O(1s) & 1mm to ≈1 cm, from Lampitt et al. (2010)
1 s
10^2 s
10^4 s
10^6 s
10^8 s
Time Scale

Why are they important?

- Energy Transport
- Mixing
- Nutrient Transport

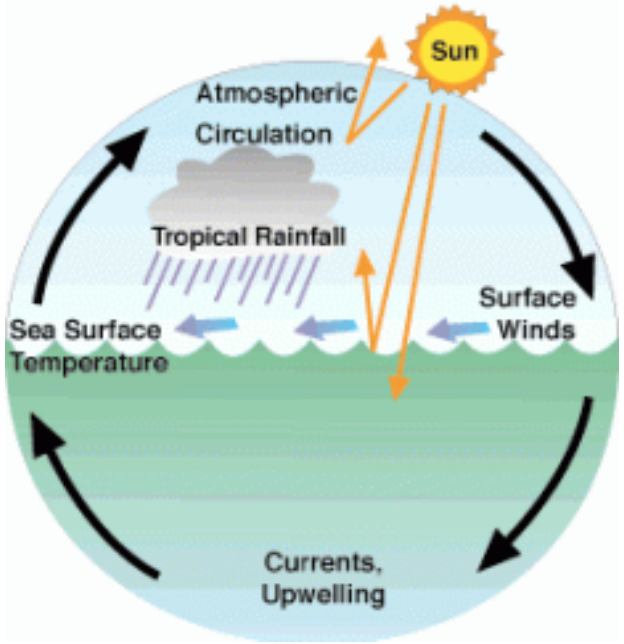


Robert Hallberg courtesy



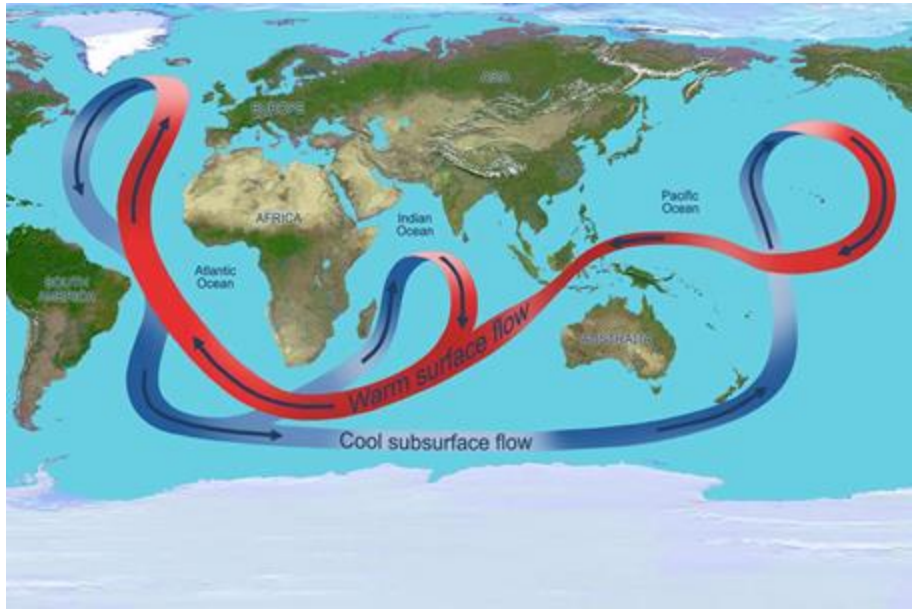
Garwood et al. 2020

Climate Regulation



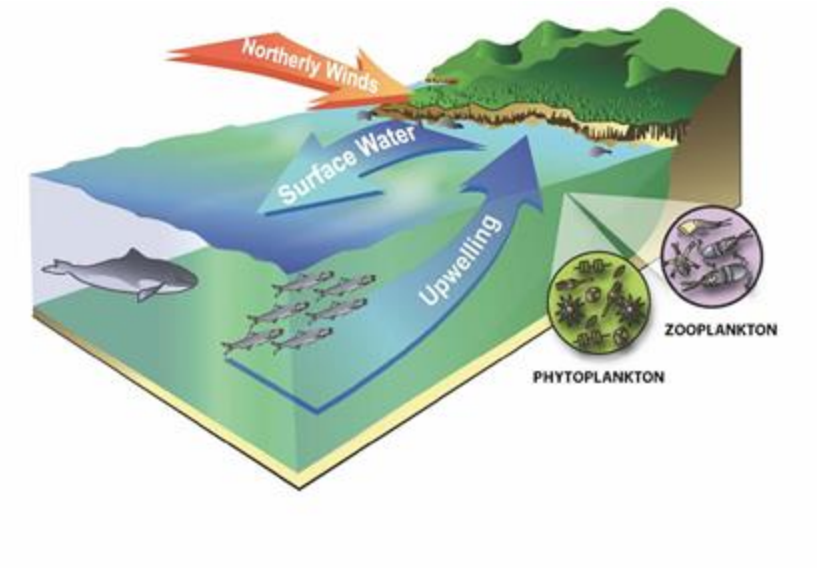
https://isabelnew.wordpress.com/2011/12/07/one-of-the-reasons-why-the-earth-is-warm/

Connection to Global Ocean Circulation



https://www.1ocean.org/ocean-tales/the-thermohaline-circulation

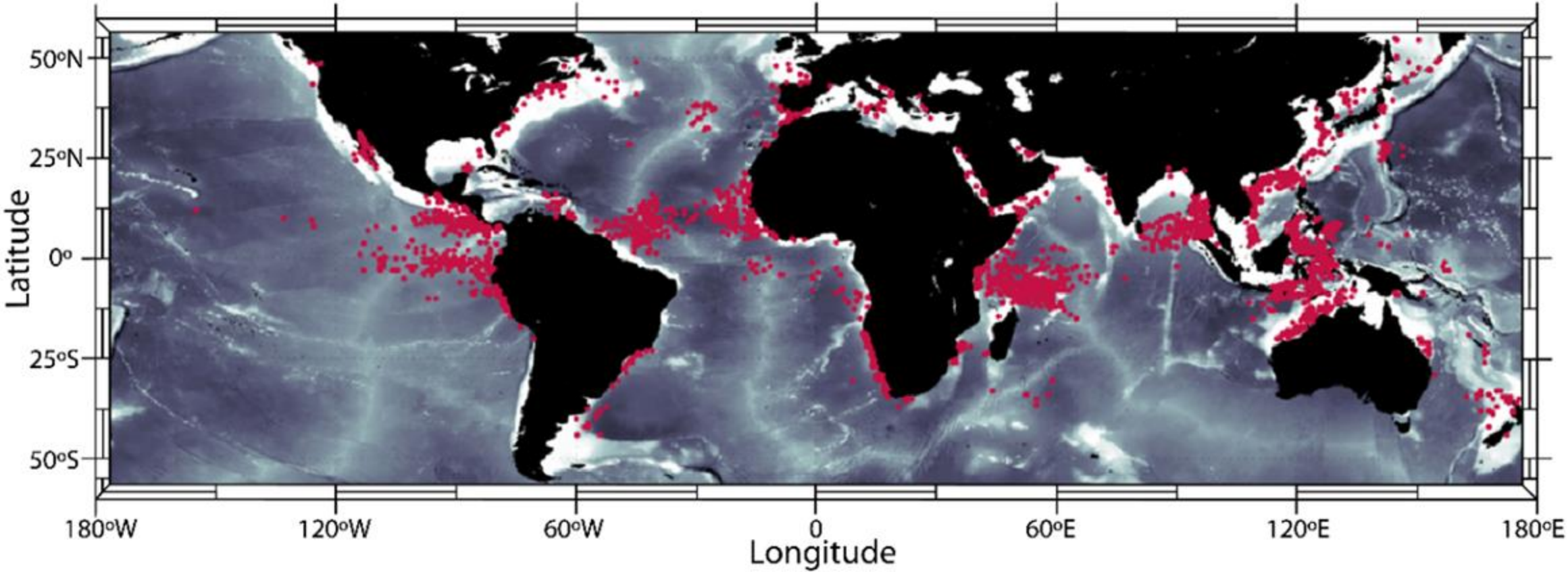
Coastal Impacts



Influence Engineering and Submarine Operations

Internal waves are essential for understanding the physical, chemical, and biological dynamics of the ocean.

Where can we find internal waves in the ocean?



Jackson et al., 2012

Despite significant research advances, Internal Waves remain a complex phenomenon with several aspects still not fully understood.

Our Goal:

To support researchers by delivering detailed analyses of internal waves hotspots, thereby improving our understanding of the Earth's oceans.

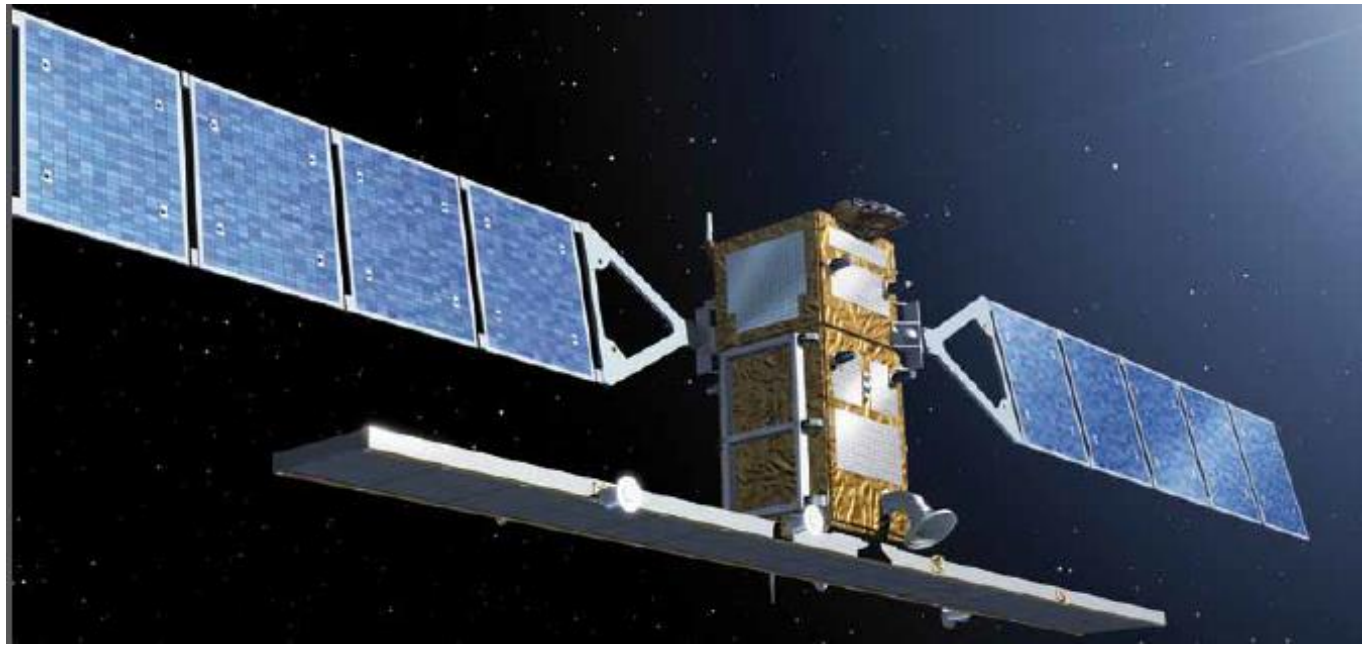
How?

We have created the "Internal Waves Service" (IWS)

https://www.aircentre.org/internal-waves-service/

Satellite Data: Sentinel-1

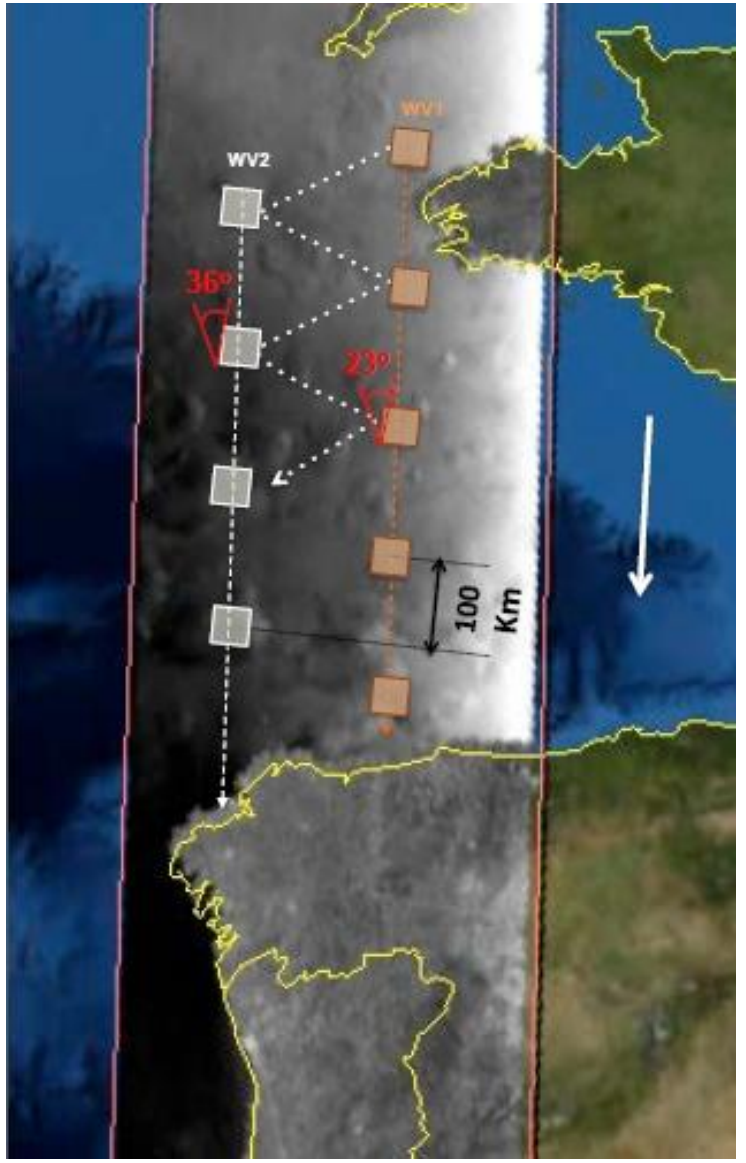
- It operates in SAR (Synthetic Aperture Radar) mode and, therefore, is not affected by clouds or atmospheric conditions.



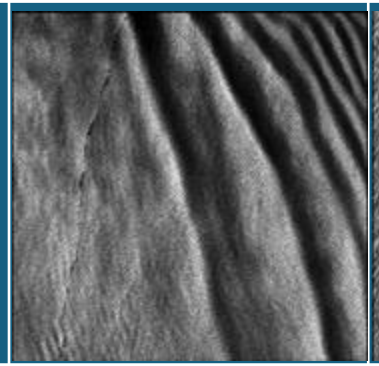
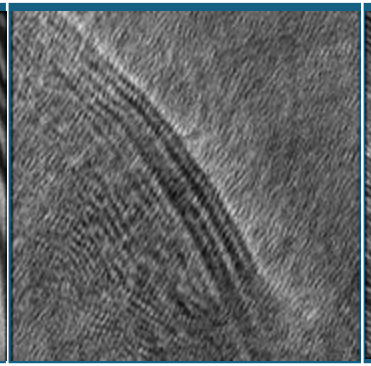
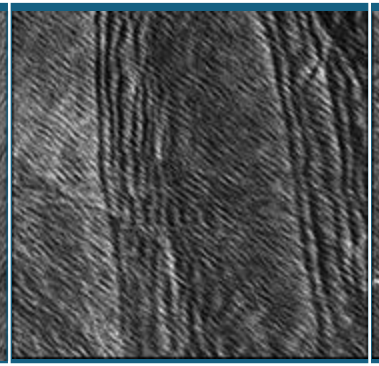
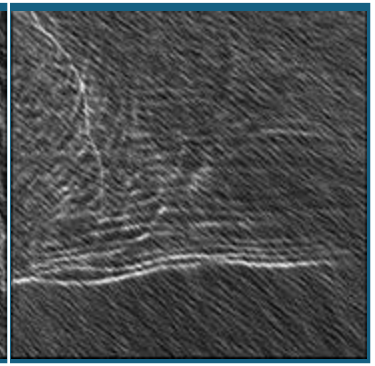
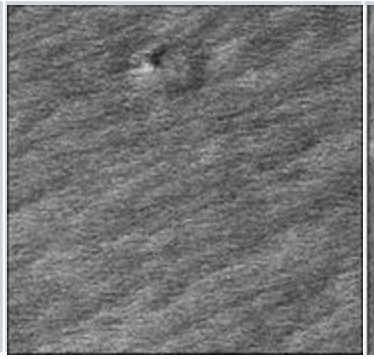
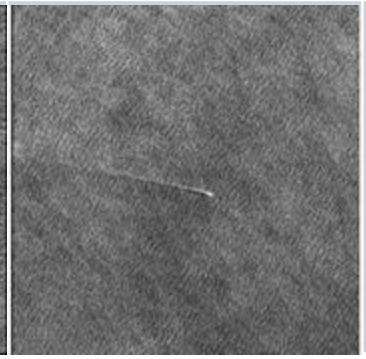
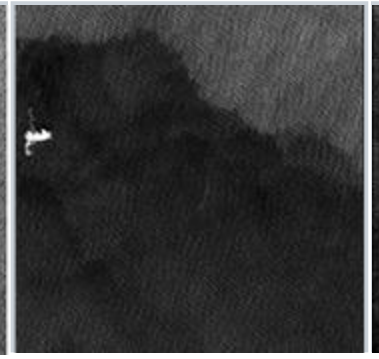
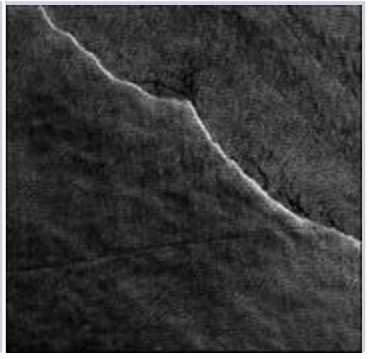
https://sentiwiki.copernicus.eu/web/s1-mission

- At the moment, we are using the WV mode, which operates primarily over open oceans with VV polarization. It captures 20 km x 20 km vignettes at a resolution of 5 m x 5 m, acquired at regular 100 km intervals along the orbit.

- However, Sentinel-1 WV mode does not operate over the entire ocean. For instance, in coastal zones, the satellite typically uses other modes.



What is this service?

Internal Waves				
No Internal Waves				

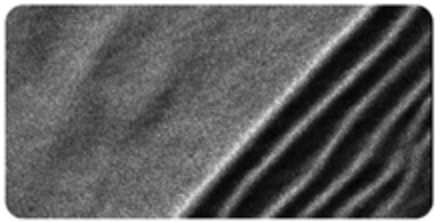
Which ML model do we use?

ATLANTIC INTERNATIONAL RESEARCH CENTRE · COMMUNITY PREDICTION COMPETITION · A MONTH AGO

Late Submission

Automatic Identification of Internal Waves

Probabilistic image classification competition - image classification challenge.



Overview | Data | Code | Models | Discussion | Leaderboard | Rules | Team | Submissions

Overview

Goal

*The **goal** of this competition is to automate the identification of simplified satellite images that have internal waves.*

The exercise consists of an image classification challenge of satellite images from Sentinel 1 (Copernicus Constellation) for the presence of internal waves.

Start
Aug 16, 2024

Close
Oct 31, 2024

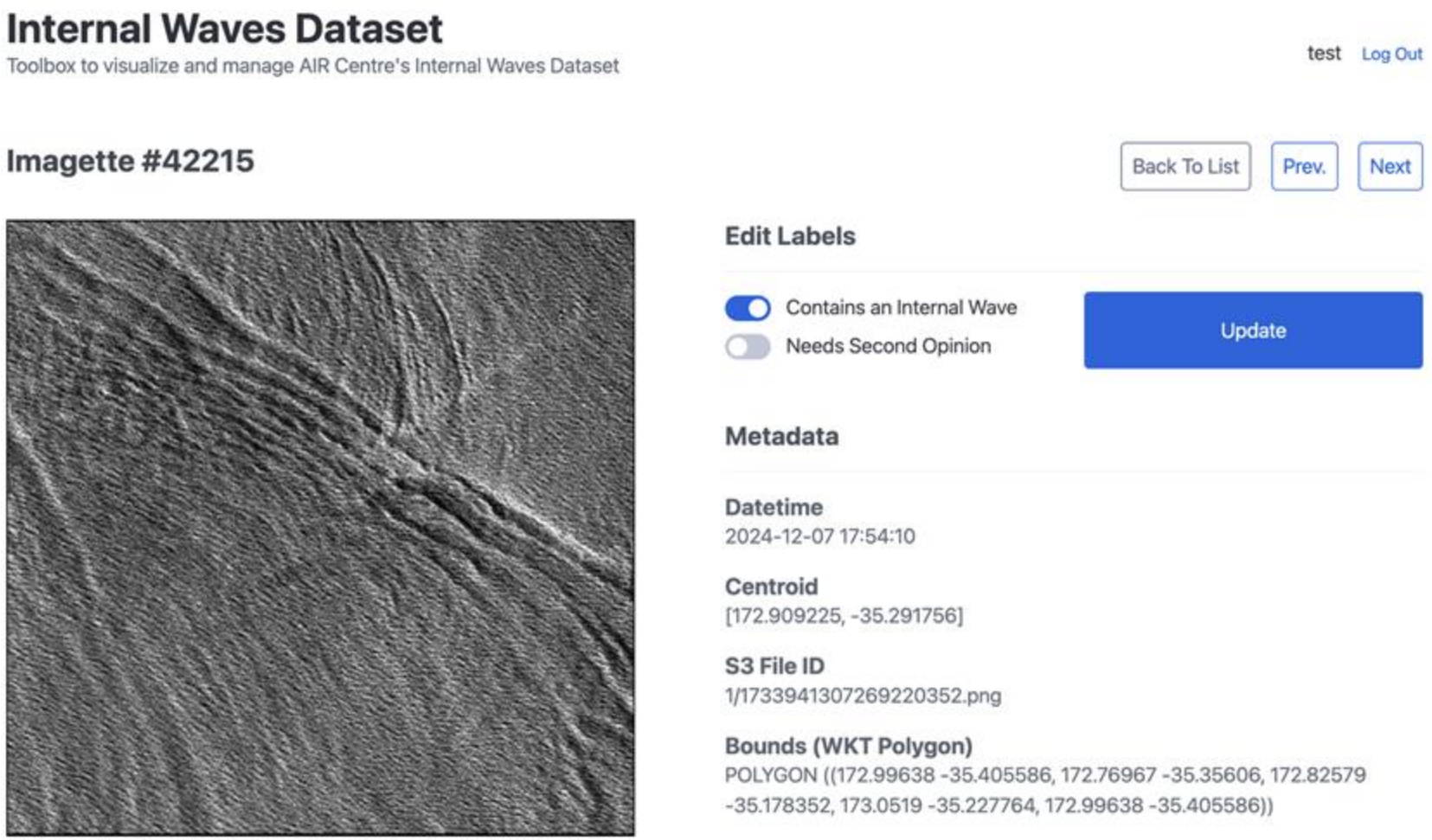
Competition Host
Atlantic International Research Centre

Prizes & Awards
Kudos
Does not award Points or Medals

Participation
67 Entrants
15 Participants
14 Teams
55 Submissions

Tags

Validation Platform



IWS Internal Waves Service WORKSHOP 2025

- Date: 3 - 4 April 2025
- Location: TBC, Angra do Heroísmo, Terceira Island, Azores , Portugal
- Timezone: UTC -1; GMT-1
- Event Type: In person only.
- Registration for attendance: by invitation only.

AIR CENTRE
ATLANTIC INTERNATIONAL RESEARCH CENTRE

U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Confirmed experts from:

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

Ifremer

NERSC
NANSEN ENVIRONMENTAL AND REMOTE SENSING CENTER
THE NANSEN CENTER · BERGEN · NORWAY

THE UNIVERSITY OF SOUTHERN MISSISSIPPI

ciimar
Centro Interdisciplinar de Investigação Marinha e Ambiental

Exploring the Winning Model

Competition Notebook

Automatic Identification of Internal Waves

Private Score	Best Score
0.98551	0.98551 V6

Version 6 of 6

Runtime

11m 6s · GPU T4 ×2

```
train = pd.read_csv('/kaggle/input/internal-waves/train.csv')
test = pd.read_csv('/kaggle/input/internal-waves/test.csv')

train['file'] = train['id'].apply(lambda x: f'/kaggle/input/internal-waves/images_train-20240709T
094004Z-001/images_train/{x}.png')
test['file'] = test['id'].apply(lambda x: f'/kaggle/input/internal-waves/images_test-20240709T093
512Z-001/images_test/{x}.png')

train
```

Exploring the Winning Model

kaggle

In [4]:

```
class Dataset(Dataset):
    def __init__(self, files, transform):
        self.files = files
        self.transform = transform

    def __len__(self):
        return len(self.files)

    def __getitem__(self, idx):
        img = Image.open(self.files[idx]).convert('RGB')
        img = self.transform(img)
        return img

model = timm.create_model('timm/eva02_large_patch14_448.mim_m38m_ft_in1k', pretrained=True)
data_config = timm.data.resolve_model_data_config(model)
transforms = timm.data.create_transform(**data_config, is_training=False)
model = model.to('cuda').eval()
model = torch.nn.DataParallel(model)
```

Exploring the Winning Model

kaggle

In [4]:

```
dataset = Dataset(train['file'].tolist(), transforms)
data_loader = DataLoader(dataset, batch_size=64, shuffle=False, num_workers=4)
print( len(dataset), len(data_loader) )
ypred = []
for batch in tqdm(data_loader):
    with torch.no_grad():
        ypred.append( model(batch.to('cuda')).softmax(dim=1).cpu().numpy() )
ypred = np.vstack(ypred)
print(ypred.shape)


dataset = Dataset(test['file'].tolist(), transforms)
data_loader = DataLoader(dataset, batch_size=64, shuffle=False, num_workers=4)
print( len(dataset), len(data_loader) )
ypred_test = []
for batch in tqdm(data_loader):
    with torch.no_grad():
        ypred_test.append( model(batch.to('cuda')).softmax(dim=1).cpu().numpy() )
ypred_test = np.vstack(ypred_test)
print(ypred_test.shape)
```

Exploring the Winning Model

kaggle

In [4]:
```
xgb_params = {
    'tree_method':        'hist',
    'random_state':       1,
    'learning_rate':      0.05,
    'max_depth':          7,
    'subsample':          0.50,
    'colsample_bytree':   0.75,
    'device':             'cuda',
}

val_score = cross_val_score(
    estimator=XGBClassifier(**xgb_params, n_estimators=550),
    X=ypred,
    y=train['ground_truth'],
    cv=StratifiedKFold(5, shuffle=True, random_state=2024),
    scoring = make_scorer(roc_auc_score),
)
np.mean(val_score), val_score
```

Out[4]:
```
(0.9380407923548514,
 array([0.9500647 , 0.939801  , 0.93703223, 0.92943976, 0.93386628]))
```

Exploring the Winning Model

kaggle

In [5]:
```
fullmodel = XGBClassifier(**xgb_params, n_estimators=550)
fullmodel.fit(ypred, train['ground_truth'])

test['predicted'] = fullmodel.predict_proba(ypred_test)[:, 1]
test[['id', 'predicted']].to_csv('submission.csv', index=False)
```

In [ ]:
```

```

In [ ]:
```

```

Advantages of Using Julia for Training and Inference

- Ideal for intensive workloads such as training and inference of machine learning models.
- Enhances inference due to optimized hardware usage and lower execution overhead.

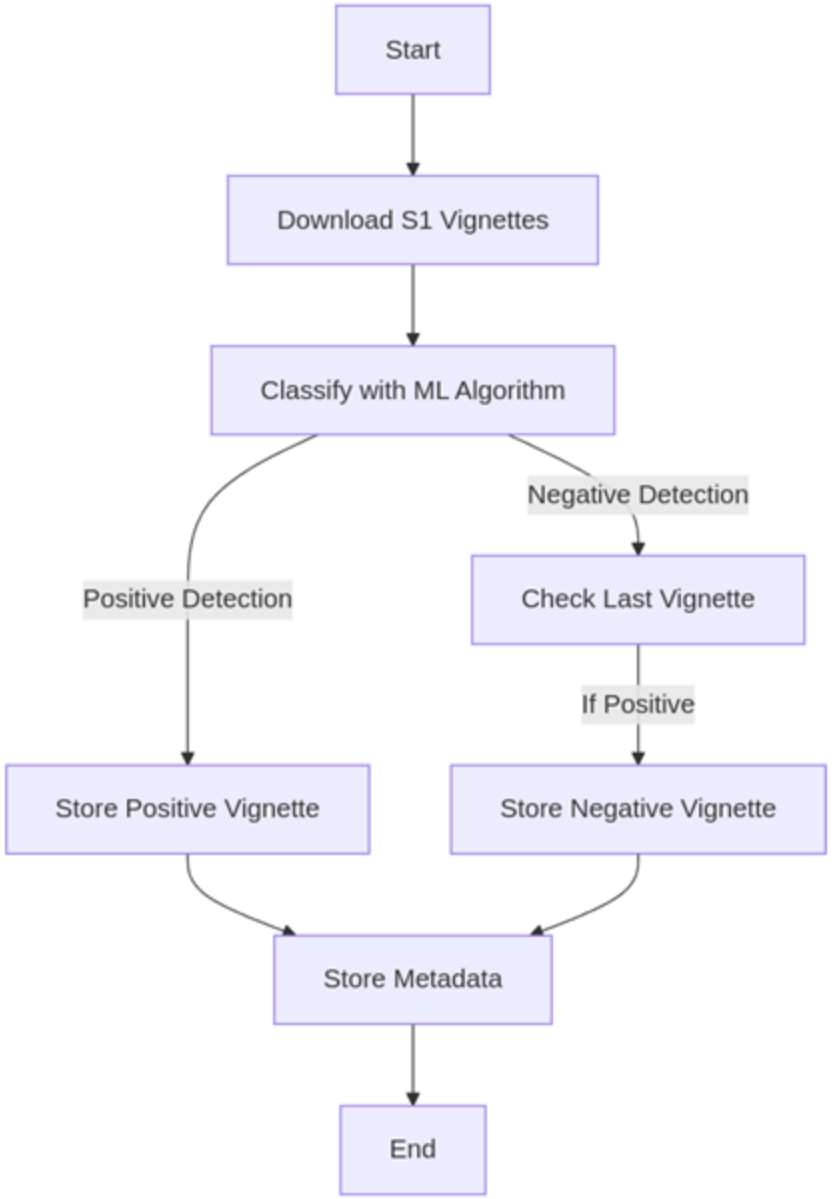
  Current code with warnings about mismatched devices (CPU and GPU) when running XGBoost. → Julia has packages (e.g. Flux.jl) optimized for GPUs and matrix operations and supports direct inference on the GPU.

- Suited for projects that require scalability, such as GPU clusters.
- Potential reduction in time per iteration by optimizing computational resource utilization.

Internal Waves Dataset Background Service

Key Features

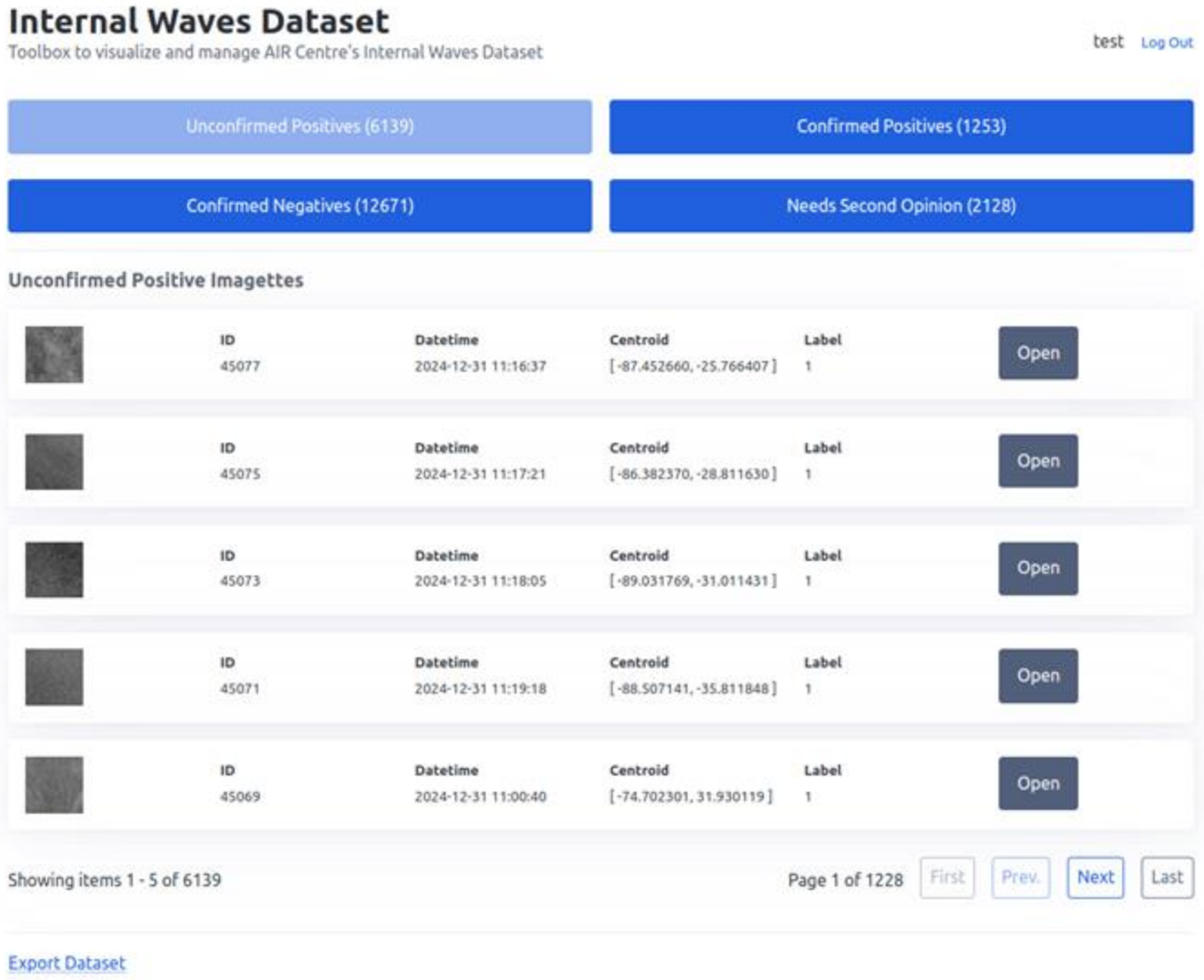
- Periodically fetches S1 vignettes from defined sources.
- Identifies IWs using a trained ML algorithm.
- Extracts and stores location, geo-bounds, timestamps, and other metadata.
- Saves all positives and one negative per positive for dataset balance.



Internal Waves Dataset Validation Dashboard

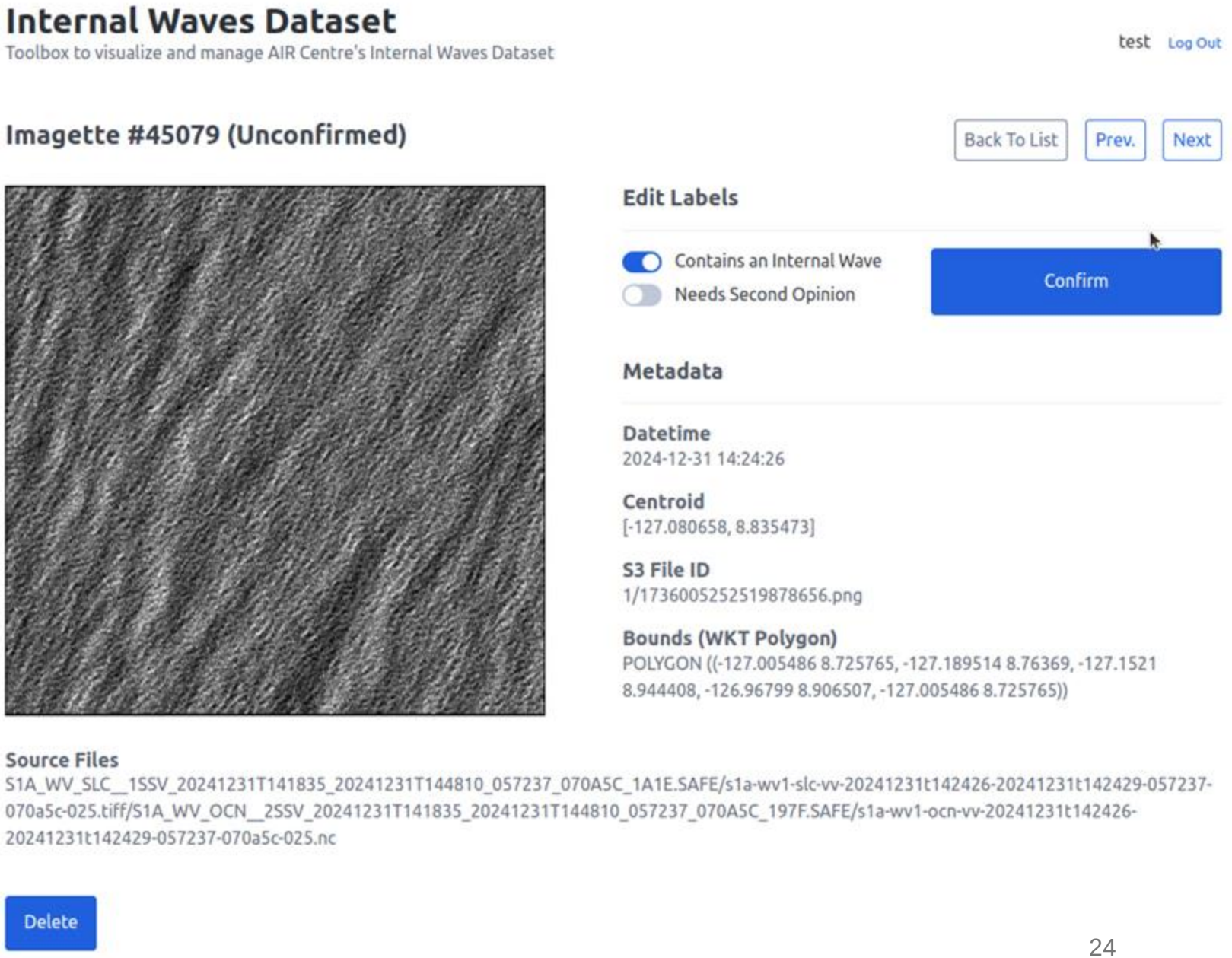
Key Features:

- Web application with user-friendly interface
- Secure access via email/password
- Inspect IW vignettes and labels
- Update labels and organize images into unvalidated and validated lists
- Dataset Export in standardized format for ML training
- Map interface showing the locations of IW events around the world



Internal Waves Dataset Validation

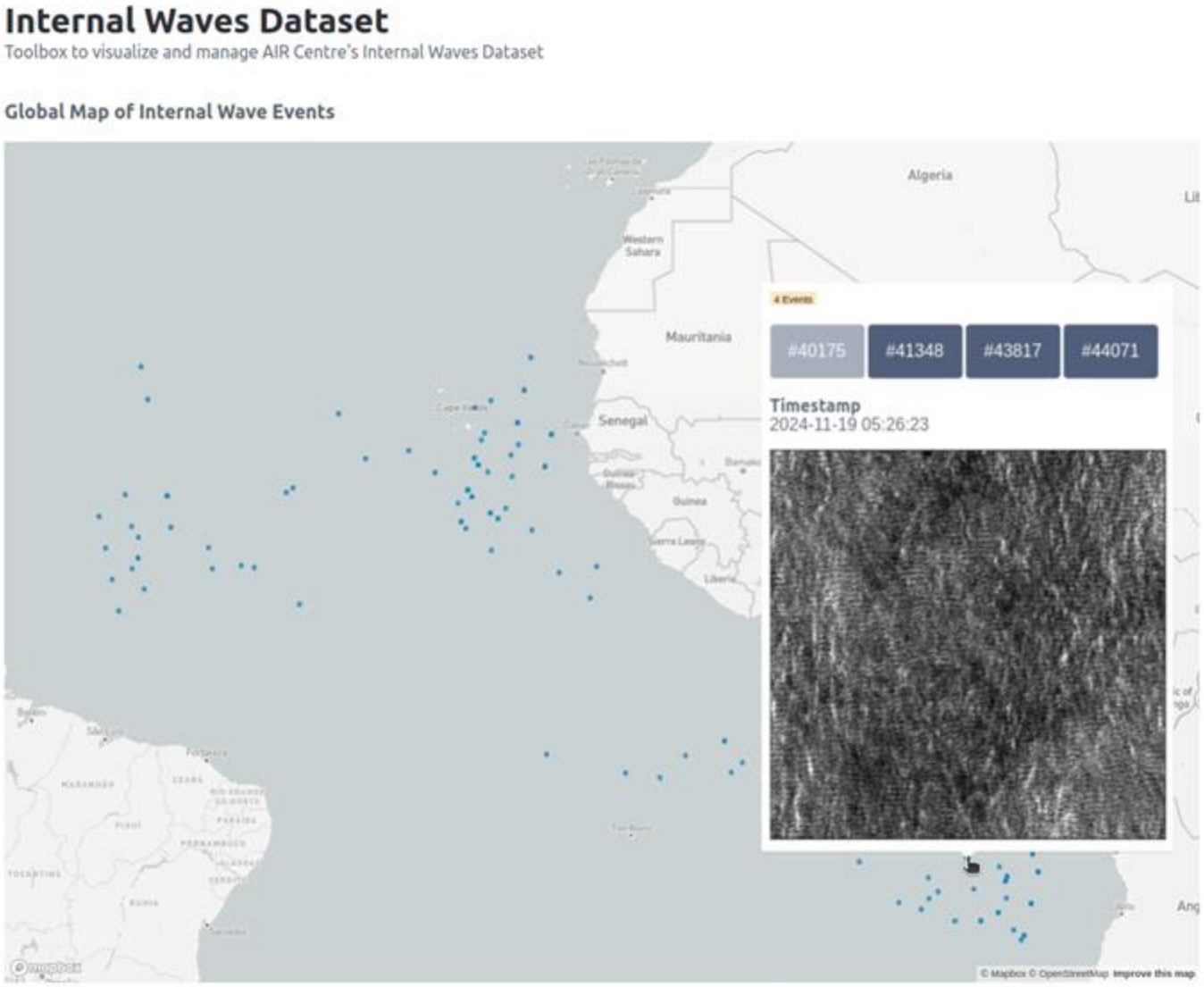
Vignette Classification Validation Interface



Internal Waves Dataset Web Map

Key Features:

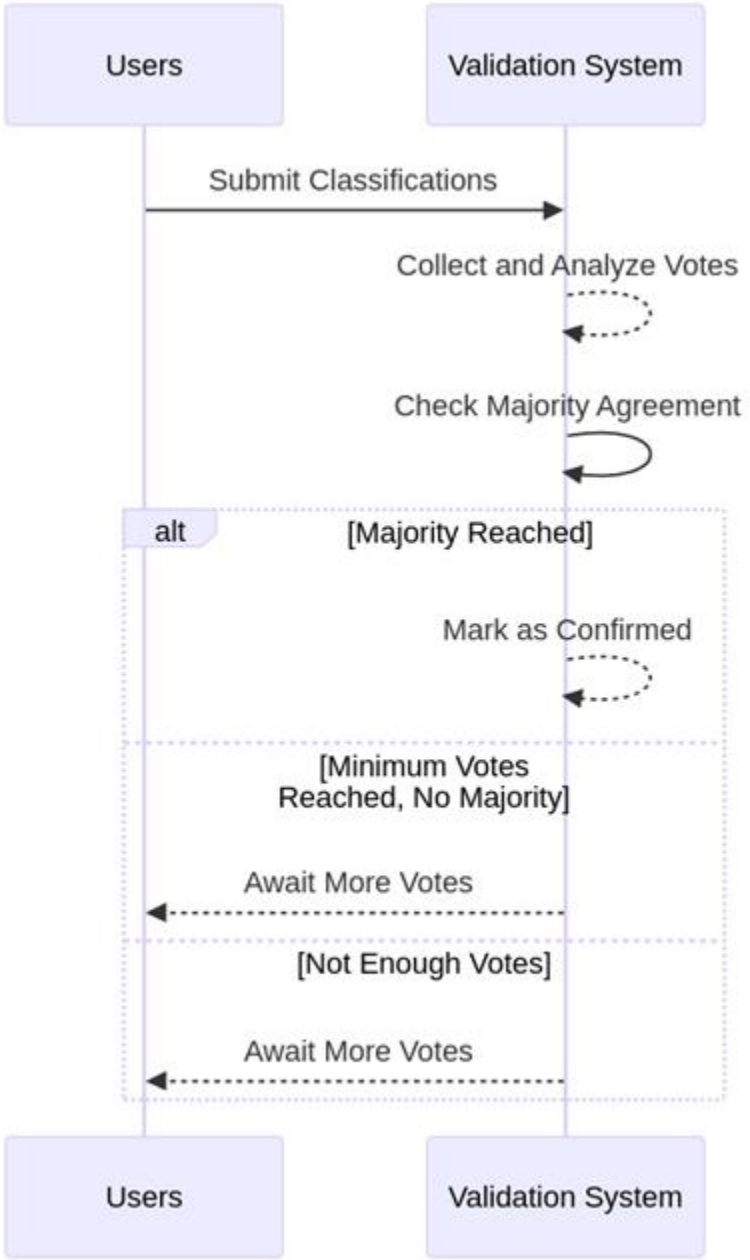
- Map visualization of validated IW events
- Click to view IW Vignette(s)
- Frequently updated



Internal Waves Dataset Validation Dashboard

Features currently in development:

- Multi-User Validation - Consensus algorithm for higher accuracy
- Text Comments - Users can add collaborative comments per vignette



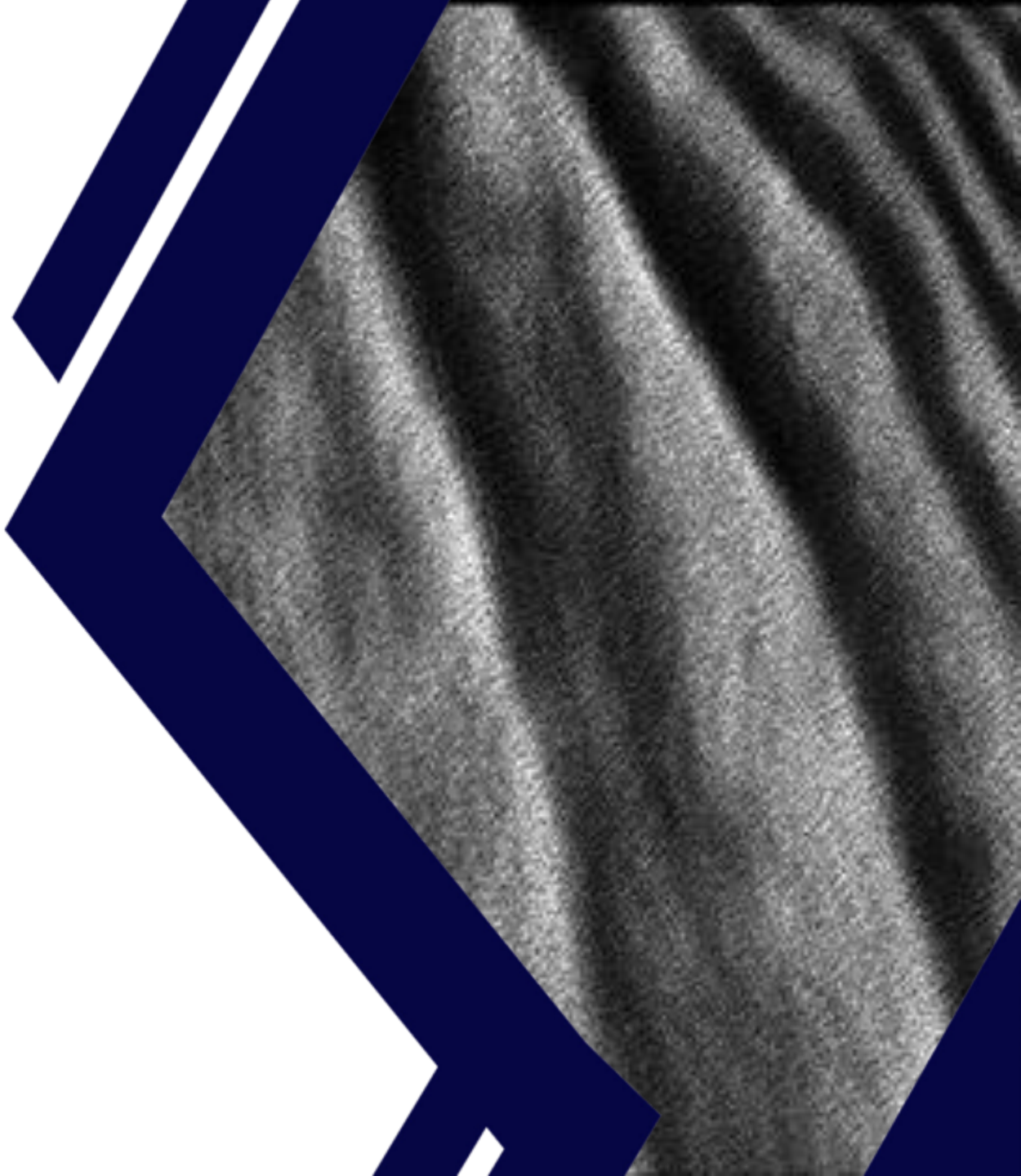
JuliaEO
Global Workshop on Earth Observation with Julia 2025

AIR CENTRE

Thank you for your attention!

IWS

Internal Waves Service



06/01/2025