MMOcean: Multi-Pretext Self-Supervised Learning for Ocean and Coastal Earth Observation

Stefan Oehmcke, Ayush Prasad
Institute for Visual & Analytic Computing, University of Rostock, Germany
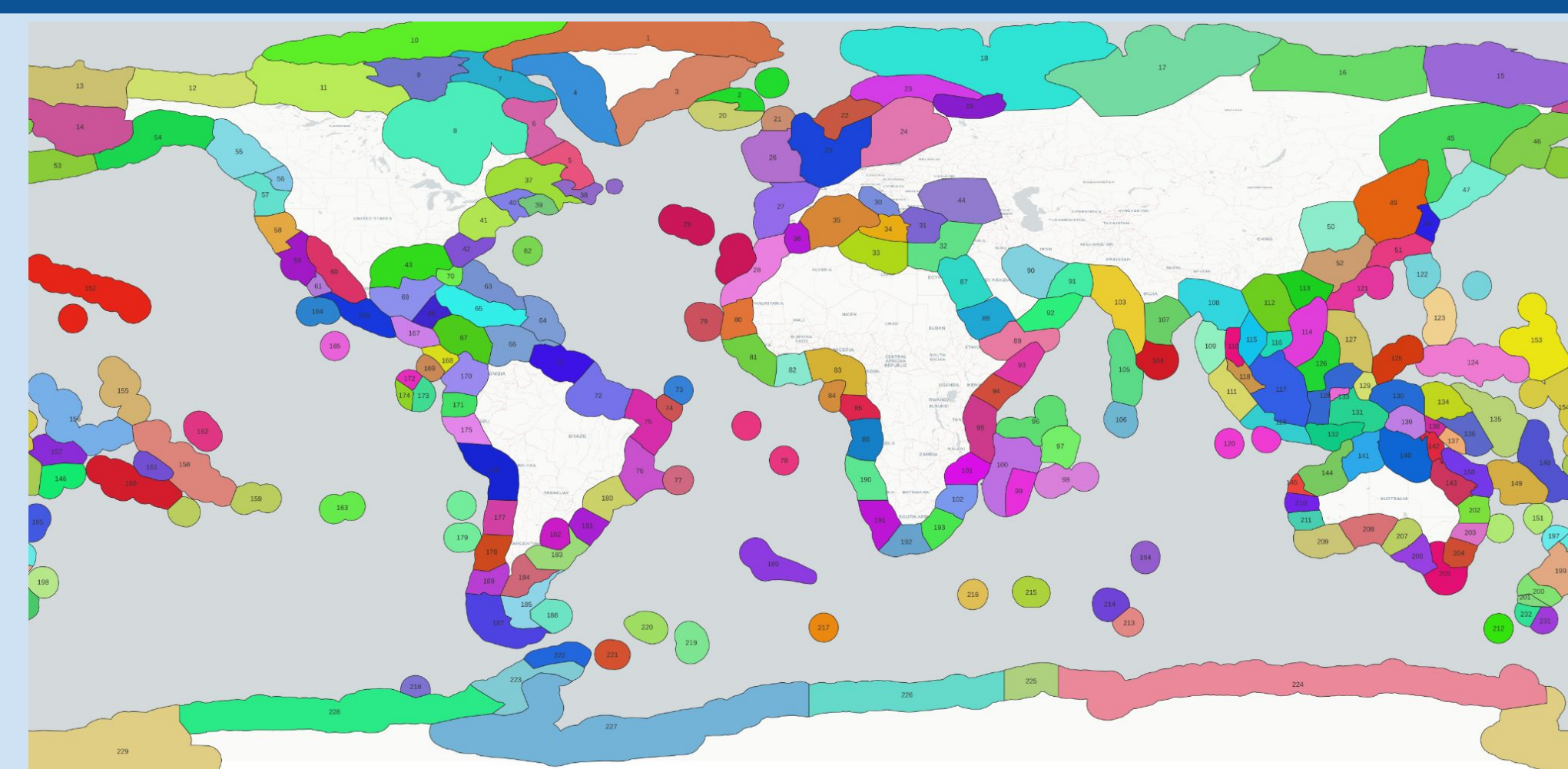Email: stefan.oehmcke@uni-rostock.de

Universität Rostock

Need for Coastal and Ocean Foundational Models

- EO enables large-scale monitoring of critical ocean and coastal environments
- Marine EO lacks labeled data compared to terrestrial applications
- Self-supervised models tailored to marine data can address this gap
- Existing methods (e.g., MMEarth) rely on pretext tasks ill-suited to marine contexts
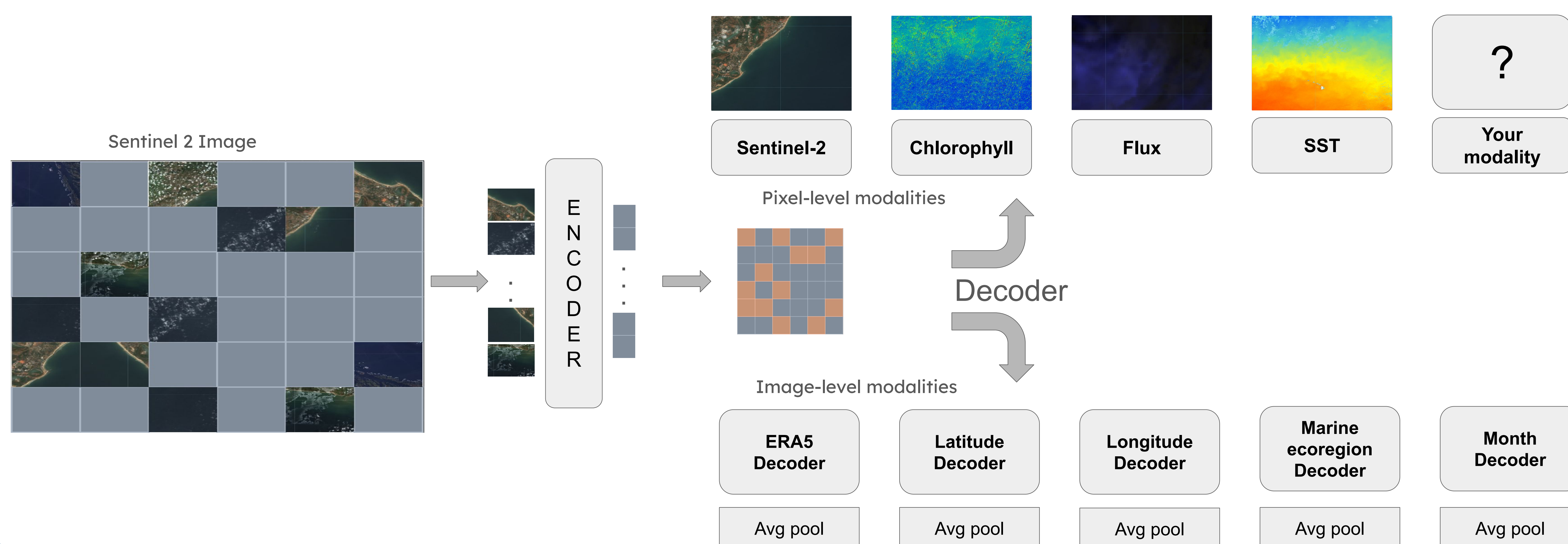
SUSTAINABLE DEVELOPMENT GOALS

14 LIFE BELOW WATER

MMOcean Dataset: At Least 1.1M Samples

Input:	Sentinel-2 (all bands)
Sampling:	WWF Marine Ecoregions
Pretext Tasks:	Pixel- and image-level
Resolution:	10–60m
Focus:	Marine self-supervised learning

WWF Marine Ecoregions

MMOcean Model: Multi-Pretext Masked Autoencoder

Sentinel 2 Image

ENCODER

Sentinel-2 | Chlorophyll | Flux | SST | ? Your modality

Pixel-level modalities

Decoder

Image-level modalities

ERA5 Decoder | Latitude Decoder | Longitude Decoder | Marine ecoregion Decoder | Month Decoder

Avg pool | Avg pool | Avg pool | Avg pool | Avg pool

Marine-Specific Pretext Tasks

Surface Reconstruction:
- SST, salinity, chlorophyll-a (pixel-level)
- ERA5 variables, geolocation, time (image-level)

Surface-to-Subsurface:
- Predict subsurface temperature, salinity, and current profiles from surface imagery

Temporal Forecasting:
- Forecast future SST, chlorophyll-a distribution, surface currents

Planned Downstream Applications

Forecasting:
- Marine heatwave prediction (e.g., SST time series forecasting)

Semantic Segmentation:
- Bathymetry mapping (e.g., MagicBathyNet)
- Coastal land-use/change classification

Regression / Classification:
- Water quality assessment (e.g., chlorophyll-a concentration, turbidity classes)

Open Questions

Task Selection:
Which pretext tasks best capture marine-specific phenomena?

Benchmarking:
How to design a marine and coastal SSL benchmark (analogous to SSL4EO-S12)?

Robustness:
How to ensure stable representations across seasons and climate-driven shifts?

References

MMEarth: Exploring multi-modal pretext tasks for geospatial representation learning. *Nedungadi, V., Kariryaa, A., Oehmcke, S., Belongie, S., Igel, C., & Lang, N. (2024, September). In European Conference on Computer Vision (ECCV) (pp. 164-182).*

Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdaña, Z. A., Finlayson, M. A. X., ... & Robertson, J. (2007). BioScience, 57(7), 573-583.*

open for collaborations