EMBED 2SCALE

Ref. Ares(2025)5148782 - 27/06/2025

Earth Observation & Weather Data Federation with AI Embeddings

D4.2
Use Case Implementation Report

Revision: v 2.0

D4.2: Use Case Implementation Report

Work package	WP 4
Task	Task 2
Due date	30/04/2025
Submission date	27/06/2025
Deliverable lead	SatCen
Version	1
Authors	Miguel A. Belenguer-Plomer (SATCEN), Michele Lazzarini (SATCEN), Damien Roberts (U Zürich), Kennedy Adriko (Jülich), Philip Stier (UOXF), Tim Reichelt (UOXF)
Reviewers	Conrad Albrecht (DLR), Damien Robert (U Zürich)
Abstract	This deliverable describes the implementation features of the Use Cases. namely the *Maritime Awareness*, the *Above Ground Biomass Estimation*, *Climate and air pollution prediction from spatio-temporal observational constraints* and the *Crop Stress and Yield Early Detection*. The deliverable resumes the work performed during the first development phase, considering their general description, the data used, the methodology developed (including the use of embeddings) and the different implementation steps. It concludes describing the preliminary results and the way forward for the second implementation cycle, when a new version of this deliverable will be released.
Keywords	Earth Observation, Artificial Intelligence, Embeddings, Maritime Awareness, Climate, Cloud, Air Pollution, Vegetation, Biomass, Crop

Document Revision History

Version	Date	Description of change	List of contributor(s)
1.1	18/06/2024	Improvement of the document based on reviewer's feedback (e.g. introduction inserted, implementation steps in tabular form)	Miguel A. Belenguer-Plomer (SATCEN), Michele Lazzarini (SATCEN), Damien Roberts (U Zürich), Kennedy Adriko (Jülich), Philip Stier (UOXF), Tim Reichelt (UOXF)

Funded by the European Union

Grant Agreement No: 101131841	**Topic:** HORIZON-EUSPA-2022-SPACE-02-55
Call: HORIZON-EUSPA-2022-SPACE	**Type of action:** HORIZON-RIA

DISCLAIMER

Funded by the European Union

EUSPA European Union Agency for the Space Programme

Project funded by

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs, Education and Research EAER
State Secretariat for Education, Research and Innovation SERI

UK Research and Innovation

Embed2Scale (Earth Observation & Weather Data Federation With AI Embeddings) project is funded by the EU's Horizon Europe program under Grant Agreement number 101131841. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them. This work has also received funding from the Swiss State Secretariat for Education, Research and Innovation (SERI) and the UK Research and Innovation (UKRI).

COPYRIGHT NOTICE

Project funded by the European Commission in the Horizon Europe Programme		
Nature of the deliverable:	to specify R	
Dissemination Level		
PU	Public, fully open, e.g. web (Deliverables flagged as public will be automatically published in CORDIS project's page)	
SEN	Sensitive, limited under the conditions of the Grant Agreement	X
Classified R-UE/ EU-R	*EU RESTRICTED under the Commission Decision No2015/ 444*	
Classified C-UE/ EU-C	*EU CONFIDENTIAL under the Commission Decision No2015/ 444*	
Classified S-UE/ EU-S	*EU SECRET under the Commission Decision No2015/ 444*	

* *R: Document, report (excluding the periodic and final reports)*
DEM: Demonstrator, pilot, prototype, plan designs
DEC: Websites, patents filing, press & media actions, videos, etc.
DATA: Data sets, microdata, etc.
DMP: Data management plan
ETHICS: Deliverables related to ethics issues.
SECURITY: Deliverables related to security issues
OTHER: Software, technical diagram, algorithms, models, etc.

TABLE OF CONTENTS

Table of Contents

ABBREVIATIONS

AGB	Above Ground Biomass
AGBD	A Global-scale Biomass Dataset
ALS	Aerial Laser Scanning
AI	Artificial Intelligence
AIS	Automatic Identification System
ARD	Analysis Ready Data
CCI	Climate Change Initiative
CH	Canopy Height
CD	Change Detection
CSCDA	Copernicus Space Components Data Access
GRD	Ground Range Detected
HR	High resolution
EO	Earth Observation
ML	Machine Learning
MODIS	Moderate Resolution Imaging Spectroradiometer
MSE	Mean Squared Error
OLR	Outgoing Longwave Radiation
PAZ	Spanish for "peace": satellite
SAR	Synthetic Aperture Radar
SLC	Single Look Complex
TLS	Terrestrial Laser Scanning
VHR	Very High Resolution

Funded by the European Union

1 INTRODUCTION

This deliverable describes the implementation features of the Use Cases, namely the Maritime Awareness, the Above Ground Biomass Estimation, Climate and air pollution prediction from spatio-temporal observational constraints and the Crop Stress and Yield Early Detection.

The document follows the design of the use cases, described in D4.1, resuming the work performed during the first development phase: it includes the use cases general description, the data used, the methodology developed (including the use of embeddings) and the different implementation steps. It concludes describing the preliminary results. The implementation features so far developed are currently receiving feedback from the respective communities and they will define the way forward for the second implementation cycle, when a new version of this deliverable will be released.

2 MARITIME AWARENESS

2.1 Use case description

Ship detection is an important aspect of maritime awareness, as ships often carry valuable cargo and pose a potential threat to populations and infrastructure. Methods for detecting ships include Synthetic Aperture Radar (SAR) [1, 2], optical imagery [3, 4], and SAR and optical imagery combined with data from the Automatic Identification System (AIS) [5, 6]. Earth Observation (EO) data allow ship traffic monitoring and the identification of potential security threats on large areas and the support of AIS data provides a technology used for maritime safety and security in near real time to identify and track vessels. Receiving timely, reliable, and meaningful information is therefore crucial.

In the last few years, Artificial Intelligence (AI) and Machine Learning (ML) have been used to detect, identify, and classify vessels in an automatic way using EO images [7]. This use case aims to develop tools to identify vessels with the support of AI and ML tools, to:

- Compress the images to improve data transfer latency and facilitate access to relevant sources and collateral data (e.g., AIS) (relying on Embed2Transfer);
- Support the creation tools for ship and port monitoring with minimal data labelling (relying on Embed2Train and Embed2Infer);
- Support the fusion of GeoData with AIS data for anomaly detection of ship movements (Embed2Find).

The application to be developed within E2S will help to define how long it takes to identify the vessels and to transfer the valuable information with the support of compression algorithms, starting from a wide area analysis and focusing on specific target areas (e.g., ports) with the support of High Resolution (HR) and Very High Resolution (VHR) images.

For the use case definition, a dialogue has been initiated with identified stakeholders (SatCen internal teams as Operations, Copernicus Border Surveillance) to provide feedback on the methodology developed and a possible user scenario.

2.2 Data

The datasets employed in this use case are classified into two categories: Earth Observation (EO) data and Automatic Identification System (AIS) data. Currently, these datasets are being utilised as training data to develop vessel detection algorithms and generate embeddings.

- Earth Observation (EO) Data:

Current efforts have been focused on utilising Sentinel-1 data, as SAR imagery is particularly advantageous for maritime applications due to the unique characteristics to detect metal surfaces within the sea. Furthermore, the open-access policy of Sentinel-1 and its high temporal resolution, which offers a six-day revisit period when combining ascending and descending orbits from the two satellites (S1A and S1C, operational since 2014 and 2024, respectively) ensure extensive spatial and temporal coverage. This results in a large dataset suitable for training embedding models based on SAR imagery.

Accordingly, dual-polarized Ground Range Detected (GRD) SAR images, featuring vertical-vertical (VV) and vertical-horizontal (VH) polarizations, acquired by Sentinel-1 in interferometric wide (IW) swath mode, were selected. These images were sourced from the [Copernicus Data Space Ecosystem] and radiometrically normalized to gamma nought ($\gamma 0$) for backscatter coefficient calibration (Figure 2.1).

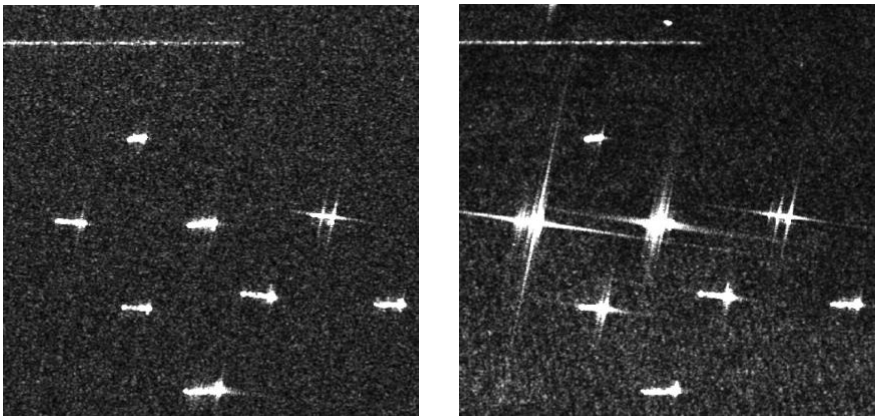


Figure 2.1: Example of SAR backscatter coefficient images for VH (left) and VV (right) polarizations, depicting vessels observed on 10 January 2020 near the Port of Los Angeles.

- Automatic Identification System (AIS) Data:

Additionally, Automatic Identification System (AIS) data has been incorporated to extract vessel locations corresponding to Sentinel-1 acquisitions. The AIS data, obtained from the National Oceanic and Atmospheric Administration (NOAA), provides daily AIS records for the United States coastline, offering sufficient coverage for training dataset extraction (Figure 2.2).

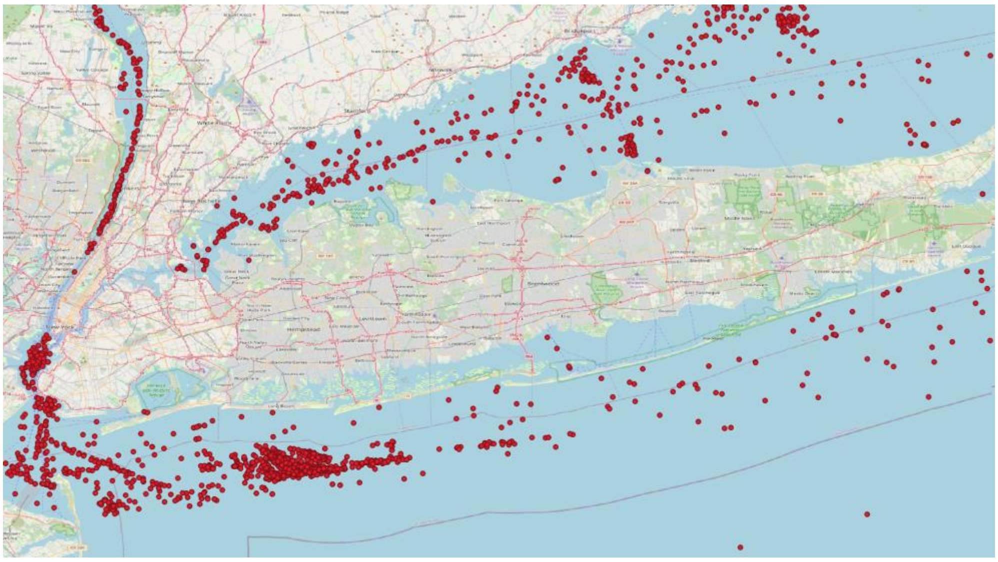


Figure 2.2: Example of AIS data corresponding to vessels detected in Sentinel-1 acquisitions between 2020 and 2024 near the Port of New York.

As previously explained, these datasets were used for training purposes. The datasets employed in the implementation of the solution specifically cover the Mediterranean Sea, which is the primary area of interest for this use case.

2.3 Methodology

The vessel detection process leverages deep learning techniques, specifically the YOLO (You Only Look Once) architecture. This approach enables efficient and accurate identification of vessels in satellite imagery. The training process relies on a comprehensive dataset constructed using Sentinel-1 backscatter coefficient images, combined with AIS data to provide precise vessel locations at the time of image acquisition. To ensure the accuracy and relevance of the training data, key pre-processing steps are required:

1. <u>Ground truth data extraction (AIS data pre-processing)</u>
   The ground truth, representing vessel locations at the time of Sentinel-1 image acquisition, is derived from AIS data. However, to ensure accuracy and reliability, the data must be refined through the following steps:

- Removing Duplicate Entries: Eliminating redundant AIS records by retaining only a single entry per vessel, corresponding to the timestamp closest to the Sentinel-1 acquisition time.
- Repositioning AIS Points to Align with Vessels: Identifying peaks in backscatter coefficients in both VV and VH polarisation within a 500 m buffer around the AIS-reported locations.
- Extracting Bounding Boxes Around Vessels: Determining vessel dimensions by extracting length information from the AIS dataset (Figure 2.3).

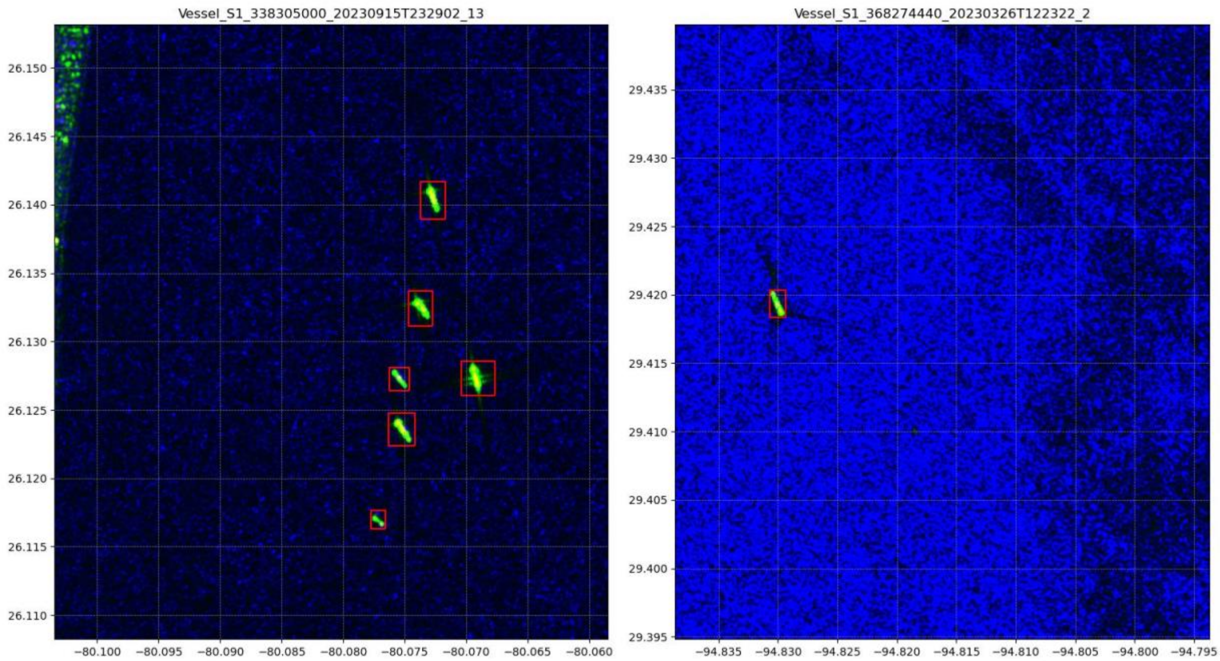


Figure 2.3: RGB composite of VH, VV, and the VH/VV ratio, highlighting areas with vessels. Bounding boxes, derived from the AIS pre-processing, are shown in red. The left image corresponds to the area near Miami Port on 15 September 2023, while the right image depicts the area near Houston Port on 26 March 2023.

2. <u>Input data preparation (EO imagery)</u>
   The Sentinel-1 VV and VH backscatter coefficient images were converted to 8-bit format and transformed into RGB, incorporating the VH/VV ratio as the third band to meet YOLO's input requirements. Additionally, the images were converted to JPG format to optimize storage and processing. This preprocessing step ensures compatibility with YOLO, a widely used algorithm for image and target detection in datasets with these characteristics.

3. <u>Vessel detection algorithm training</u>
   The training process for vessel detection utilises the YOLO architecture [48], a deep learning framework designed for real-time object detection. To enhance efficiency and accuracy, the model is initially configured with pre-trained weights, enabling it to leverage feature representations learned from broader object detection tasks. This transfer learning approach

provides a strong foundation, allowing the model to focus on the unique characteristics of vessels in Sentinel-1 imagery.

To refine its capability, the model undergoes fine-tuning using the Sentinel-1 dataset, optimising it for the detection of vessels based on radar backscatter properties. Training is conducted in batches of 500x500 pixels, with each iteration processing a set of images and their corresponding bounding boxes. This batch-based approach ensures that the model incrementally learns to minimise detection loss, improving its ability to distinguish vessels from surrounding maritime features.

4. Validation and Optimisation
The model's performance will be evaluated using a validation set of unseen Sentinel-1 imagery with annotated vessel locations (i..e, 20% of total labelled vessels reserved for validation). Key metrics such as precision, recall, and mean Average Precision (mAP) will assess detection accuracy, while hyperparameter tuning enhances generalisation and reduces false positives.

- **Use of embeddings**

When it comes to integration of lossy neural compression methodologies, approaches that focus on single timestamp embeddings were targeted. That is: the neural compressor takes as input a single image patch associated with a given scan date when the SAR satellite scans of a given geospatial area. For example, the neural compressor may take a single Sentinel-1 image of the GRD product. In fact, the efforts in generating a benchmark dataset faced the challenge to match ship geolocations to SAR satellite data acquisitions. Time series of Sentinel-1 imagery of a given ship track are far from reach for data assembly when fusing ship track records with SAR satellite data.

In collaboration with Advisory Board (AB) member ICEYE operating a large fleet of SAR satellites, the potential of SAR imagery compression was discussed and a joint team to investigate the use of neural embeddings has been initiated. From the AB perspective, the use of foundation models to generate embeddings implies expectations of resilience to spatial resolution, image distribution shifts, and other aspects that vary from one satellite generation to the next. Along the lines, it is key to develop standards for embedding formats (cf. WP3). Once embeddings are available, Embed2Find may serve to assemble diverse training datasets to fine-tune ship detection models as discussed above.

For the novel E2S data challenge framework developed in WP2, a benchmark dataset of Sentinel-1 data labeled by the AIS coordinates of ships in the scene was assembled. These will be utilized to train linear models that estimate the number and location (in pixel coordinates) of ships in a scene from the embeddings as a regression task. Note that the task of ship type detection is “dense” similar to pixel-wise semantic segmentation. Embeddings that summarize image-level features are insufficient.

2.4 Implementation steps

The implementation of this use case follows a structured workflow which integrates EO and AIS data, deep learning techniques, and embedding methodologies to enhance ship detection capabilities. The key implementation steps are summarised in the table below (Table 2.1).

Table 2.1: Implementation steps and current status for maritime awareness use case.

Step	Description	Implementation Status
Data Collection	Gather Sentinel-1 SAR imagery and AIS data for training	Done
Data Pre-Processing	Convert Sentinel-1 images to 8-bit RGB format, align AIS data with SAR acquisitions, and extract bounding boxes	Done
Ground Truth Generation	Refine AIS data by removing duplicates, repositioning points based on backscatter peaks, and extracting vessel dimensions	Ongoing
Model Training	Train the YOLO model using pre-trained weights and fine-tune on Sentinel-1 imagery with batch-based learning	Ongoing
Validation and Optimisation	Evaluate performance using a validation dataset, measuring precision, recall, and mAP, followed by hyperparameter tuning	Pending
Embedding Integration	Develop embeddings for ship detection and integrate lossy neural compression methods for efficient data handling	Pending
Benchmarking and Refinement	Assemble a benchmark dataset of Sentinel-1 and AIS-labeled data to improve model generalisation and standardise embedding formats	Pending
Deployment and Application	Implement trained models for ship detection and large-scale maritime surveillance	Pending

2.5 Current Results

The vessel detection system has reached key development milestones, with data collection and pre-processing completed, while ground truth generation and model training are ongoing. Preliminary tests using the YOLO object detection model on Sentinel-1 SAR imagery have been conducted, demonstrating the potential of the approach for ship detection.

a) Preliminary vessel detection insights

Accurate vessel detection has been achieved, with most vessels correctly identified (Figure 2.4). However, challenges remain, including small vessel detection, clutter, and bounding box precision, which require further refinement.

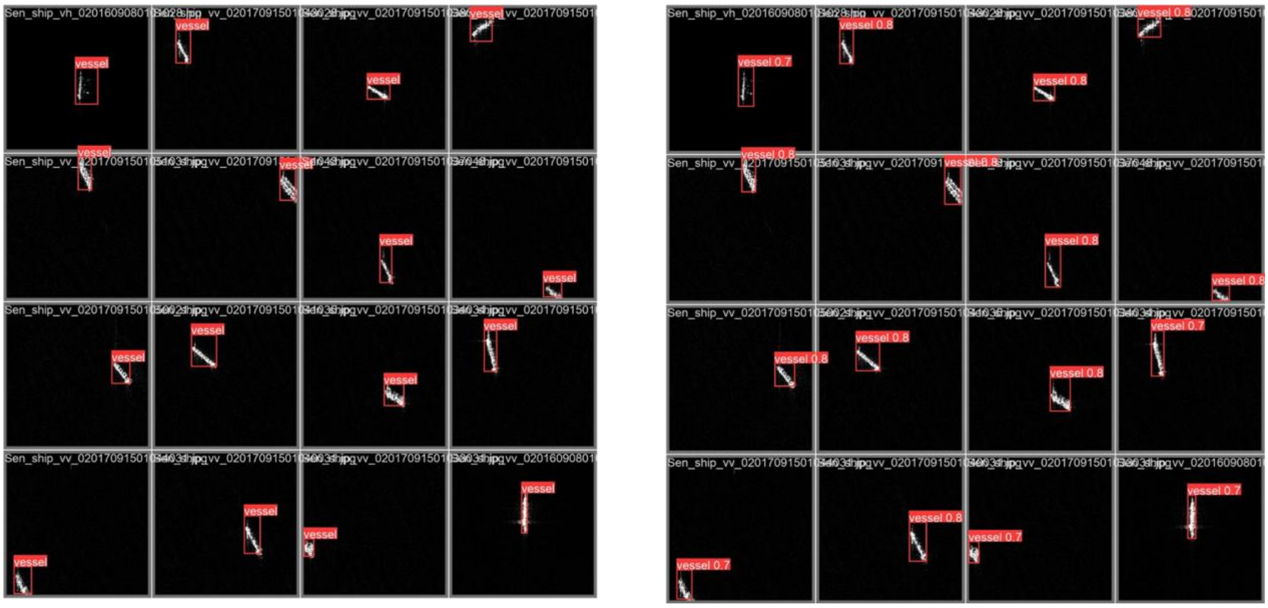


Figure 2.4: Preliminary vessel detections in Sentinel-1 imagery (detections shown in red). The left image represents ground truth, while the right image displays algorithm-detected vessels.

b) Final Training Dataset

The final training dataset comprises **25,901 vessels**, identified by their MMSI codes across multiple acquisition times. Consequently, the same vessel may appear more than once at different acquisition dates. The dataset includes 2,008 processed Sentinel-1 images collected from 12 highly trafficked ports in the United States between 2020 and 2024. These images have been segmented into **90,348 batches of 500×500 pixels, with 53,046 containing vessels and 37,302 without.**

2.6 Way Forward

To enhance the vessel detection accuracy based on the use of embeddings as well as the operational efficiency, the following steps will be pursued:

a) Model optimisation and validation

- Complete ground truth refinement for better training labels
- Finalise YOLO training and tune hyperparameters to improve detection accuracy

b) Exploring detection accuracies of different input

- Raw Sentinel-1 SAR Data (VV/VH dual-polarization)
- Neural compression of SAR images to feature vectors
- SAR imagery decompressed from feature vectors

c) Incorporation of PAZ data from Hisdesat

- Enhanced spatial resolution, enabling both vessel detection and classification
- Inclusion of multiple polarizations, requiring embedding training specific to this dataset

d) Deployment and application

- Test the operability of the algorithm when monitoring large areas (i.e., Mediterranean sea)

Funded by the European Union

EUSPA

3 GLOBAL ABOVE-GROUND BIOMASS ESTIMATION

3.1 Use case description

Worldwide mapping of vegetation properties is of prime importance for understanding the global carbon cycle [9], the impact of human activities on carbon emissions [11], and the study of ecosystem services [15]. The accurate and frequent mapping of a small set of vegetation structure indicators such as canopy height (CH) and aboveground biomass (AGB) is key to the study of terrestrial ecosystem functions [12, 16].

The traditional protocol for estimating such indicators requires in-situ (sometimes destructive) manual measurement surveys. Due to the poor spatial-temporal scalability of this approach (Figure 3.1), much research effort has been invested in characterising vegetation structure from remote sensing data with terrestrial laser scanning (TLS) and aerial laser scanning (ALS) [19]. While ALS provides accurate, dense, very high-resolution data, acquisition campaigns remain costly, limited to regional scales, with revisit rates of several years. The ultimate need to scale vegetation mapping to global scale with revisit rates below one year and low-cost data hence calls for space-borne acquisition. Ideal satellite observations for global forest analysis need to capture vegetation properties at high spatial resolution with high revisit rate and be freely available. Several works have proposed to map forest structure from time-series of NASA/USGS Landsat or ESA Sentinel-1/2 acquisitions [17]. Recently, combining spaceborne LiDAR measurements from the NASA GEDI mission [10] with Sentinel imagery has shown great potential for regressing forest biophysical variables like AGB or CH at a global scale and 10 m resolution [14].

Still, Lang et al. [14] find that the prediction of a single, global map for the year of 2020 requires extensive computational power. To cover the entire landmass of the Earth (excluding Antarctica), a total of ~160 terabytes of Sentinel-2 image data need to be downloaded and processed. Running the model on these images takes ~27,000 GPU-hours (~3 GPU years) of computation time, parallelized on a high-performance commercial cluster to obtain the global map in ten days real time. Yet, the breakdown of the entire process reveals that more than half of the time is spent downloading and moving data between devices.

Therefore, a pipeline capable of efficiently and accurately transmitting neurally-compressed sensor data (Embed2Transfer) or pretrained feature representations (Embed2Infer, Embed2Train) would allow producing and frequently updating vegetation structure maps. By lowering the compute, bandwidth, and AI expertise cost for using deep learning models to regress vegetation structure variables from remote sensing data, more stakeholders may take part in the production and analysis of such products. This would in turn benefit crucial applications such as ecosystem protection and global carbon cycle monitoring.

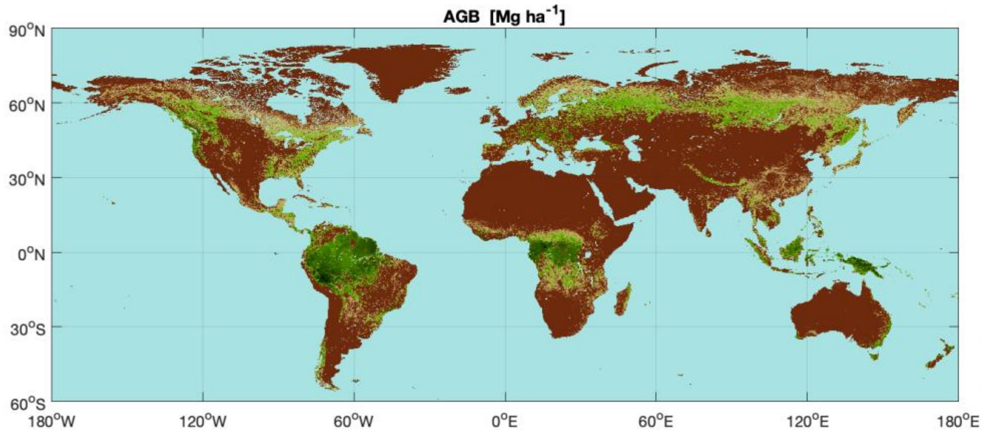


Figure 3.1: Aboveground biomass map from the Climate Change Initiative (CCI) biomass project. Despite its global coverage, this product has a 100 m ground sampling resolution and is only available for 2017, 2018, and 2020. This is limiting for applications needing to monitor the evolution of vegetation structure at higher spatial-temporal resolution.

This use case has emerged from discussions between the UZH EcoVision lab, IBM Research Zurich, and Google Research Zurich which have a shared interest for researching and producing high-resolution global maps of vegetation structure. Other stakeholders include companies from the chocolate industry such as Barry Callebaut, Nestlé, which all need tools ensuring their supply chains are deforestation-free [14, 9]. Finally, insurance companies such as SwissRe and Axa have also shown interest in vegetation maps for diverse risk assessments.

3.2 Data

The focus of this use case is to assess the quality of Embed2Scale embeddings for addressing the task of dense aboveground biomass (AGB) estimation. To this end, we will use the NASA GEDI L4A product as target AGB data. We choose this specific data source because it is, at the moment, the best available global dataset of biomass estimates at high resolution. Alternatives such as the ESA CCI map do cover the globe but at a 100 m resolution, much lower than the 10 m resolution we can hope to achieve using Embed2Train representations extracted from Sentinel imagery. Meanwhile, other datasets do provide biomass estimates derived from high-resolution aerial or terrestrial LiDAR acquisitions, but these are always limited to small areas and cannot be combined into a larger global dataset due to their heterogeneous acquisition and modelling protocols.

Figure 3.2: Geographic distribution of the AGBD dataset. The spatial sampling ensures a complete and balanced representation of the diversity of vegetation ecosystems and landscapes across the globe. At each geolocation, biomass estimates and rasters from various remote sensing modalities are provided.

In order to assess the potential of Embed2Scale embeddings for biomass estimation, we will compare the performance of models trained to regress the GEDI L4A AGB estimates either from Embed2Scale embeddings or directly from satellite imagery. This hence requires a dataset of satellite imagery along with co-registered AGB GEDI estimates. To this end, we will use A Global-scale Biomass Dataset (AGBD) dataset [20], which defines a stratified sampling of locations around the globe ensuring a complete and balanced representation of the diversity of vegetation ecosystems and landscapes. This dataset provides diverse modalities at each location Figure 3.2): Sentinel-2 imagery, ALOS-2 PALSAR-2 imagery, land cover rasters, digital elevation maps, canopy height maps. In addition, the geolocation of each raster image patch is provided, which allows for aligning Embed2Scale embeddings with the dataset for benchmarking purposes.

3.3 Methodology

We plan to assess the Embed2Scale embeddings against a baseline model inspired from existing works [14, 20]. The Lang et al. [14] backbone will be trained to regress the sparse AGB estimates provided by the NASA GEDI L4A product. The AGBD dataset allows training on a diverse set of inputs: single-date Sentinel-2 imagery, ALOS-2 PALSAR-2 imagery, land cover, digital elevation, canopy height predictions, and geolocation. Sialelli et al. [20] provide metric performance for several combinations of those, demonstrating that the more input modalities, the more accurate the AGB predictions. However, it is worth noting that comparison between such a multimodal model and Embed2Scale embeddings solely based on Sentinel-2 imagery would not be entirely fair. For this reason, we will train a baseline solely based on Sentinel-2 for comparison purposes. Furthermore, we will compute fine-grained performance metrics characterizing the distribution of prediction errors. This is of particular interest for this task, as all models in the literature tend to overestimate low biomass values and underestimate high values.

■ Use of embeddings

To leverage the Embed2Scale embeddings, we will first gather a set of embeddings for Sentinel-2 patches centered at the same locations as the AGBD dataset. These embeddings may optionally be computed from different patch sizes, thus providing features capturing diverse levels of context.

Then, a small model will be trained to regress the AGBD targets based on the embeddings. Since our task requires dense predictions (i.e. a value per pixel in the image), we identify two possible directions, depending on the nature of the embeddings. If the Embed2Train embeddings are one-dimensional vectors summarizing the content of a patch, we will use an MLP and deconvolution operators to decode the embedding into a feature map from which the biomass will be predicted at each pixel. If the Embed2Train embeddings already provide a feature map we will directly regress biomass from each pixel feature (Figure 3.3).

Additionally, as observed in [20], we expect additional modalities to improve performance. For this reason we may explore model architectures combining the pretrained Embed2Scale Sentinel-2 embeddings with auxiliary information, such as the geolocation.

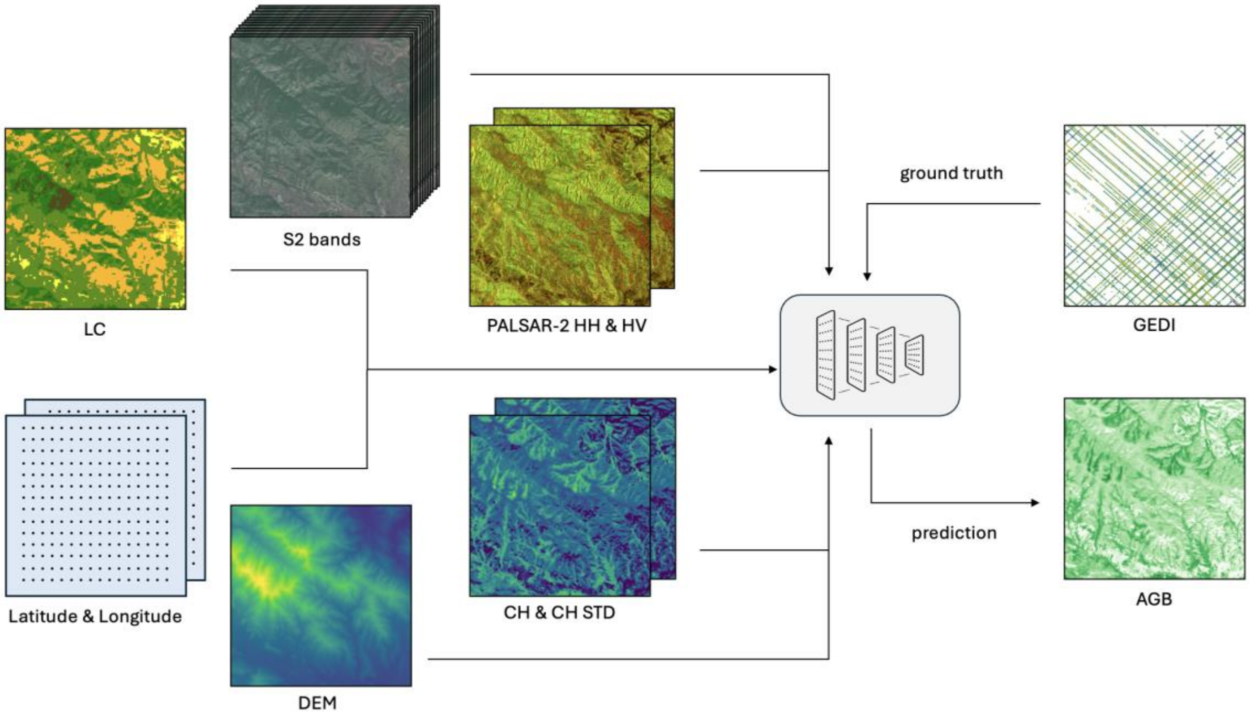


Figure 3.3: Pipeline for our baseline model, based on Sialelli et al. [20]. Given a (flexible) combination of co-registered rasters and geolocation information, the model is trained to predict AGB at each pixel location. The supervision signal comes from the sparsely sampled GEDI L4A estimates.

3.4 Implementation steps

The implementation steps and the status for the above ground biomass estimation use case are displayed in table 3.1.

Table 3.1: Implementation steps and current status for the above ground biomass estimation use case.

Steps	**Implementation Status**
Downloading and preprocessing the AGBD [20] dataset	Done
Implementing metrics for tracking the distribution of errors on AGBD training	Done
Training baseline models on AGBD with Sentinel-2 only as input	Done
Training baseline models on AGBD with other modalities as input	Done
Implementing the pipeline for training a model on E2S embeddings	Ongoing
Training a model on E2S embeddings	Pending
Training a model on E2S embeddings with auxiliary modalities	Pending
Comparing performance of image-based and embedding-based models	Pending

3.5 Current Results

We show in Figure 2.4 the results of the exploration of several baseline model designs on the AGBD dataset [20]. Several model architectures are explored: GBDT, FCN, UNet, Lang et al. [14]. Lang et al. performs best, which motivates our choice of this architecture for our baseline model. Furthermore, the results indicate that the more input modalities are fed to the model, the better the performance. However, for the sake of fair comparison, the Lang et al model operating on Sentinel-2 only (third row) should be our main point of comparison for models trained on Embed2Scale embeddings derived from Sentinel-2 patches.

Features						Methods						
CH	RGBN	S2 (all)	LC	DEM	P2	GBDT (1×1)	FCN (15×15)	FCN (25×25)	UNet (15×15)	UNet (25×25)	Lang et al. (15×15)	Lang et al. (25×25)
×						70.73 ±0.18	69.29 ±0.16	69.33 ±0.07	69.08 ±0.08	68.54 ±0.12	68.54 ±0.27	68.11 ±0.24
	×					72.92 ±0.18	64.97 ±0.25	64.75 ±0.11	63.14 ±0.29	61.97 ±0.09	61.60 ±0.13	60.39 ±0.19
		×				69.41 ±0.11	63.52 ±0.22	63.48 ±0.08	62.00 ±0.09	60.69 ±0.05	60.36 ±0.13	59.15 ±0.22
		×	×			66.48 ±0.05	62.40 ±0.19	62.47 ±0.12	61.09 ±0.18	59.99 ±0.12	59.65 ±0.10	58.82 ±0.14
		×		×		67.57 ±0.06	62.43 ±0.24	61.78 ±0.69	60.47 ±0.67	59.33 ±0.40	58.55 ±0.19	57.90 ±0.21
		×			×	67.27 ±0.08	61.95 ±0.19	61.81 ±0.18	60.59 ±0.14	59.27 ±0.11	59.12 ±0.11	58.22 ±0.23
						65.16	60.97	60.97	59.84	58.50	58.16	57.14

Figure 2.4: Comparison of various baseline model designs from Sialelli et al. [20]. Mean test RMSE (↓) and associated standard deviation per model, with various inputs. Crosses denote the presence of a feature, values in brackets denote patch size. Value

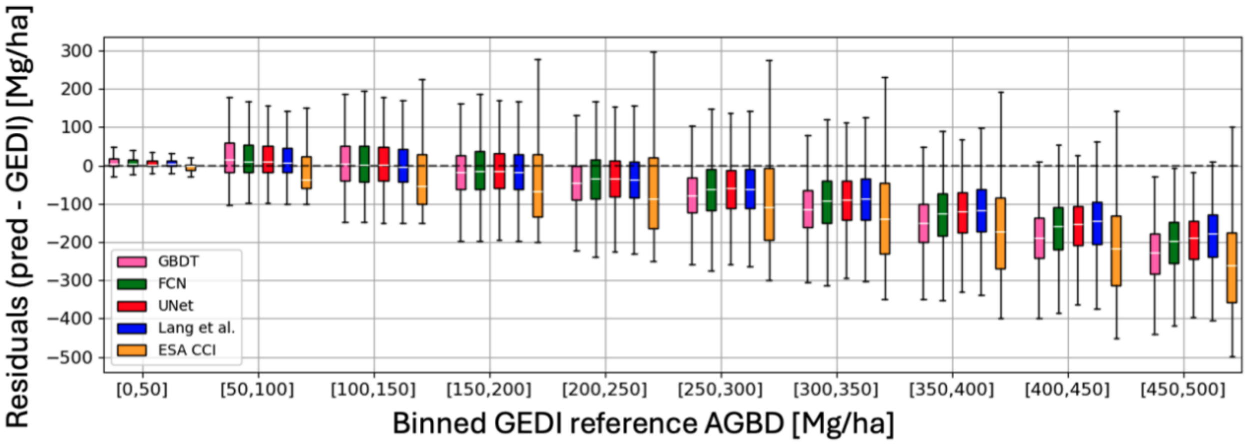


Figure 2.5: Distribution of biomass prediction errors for various baseline architectures and for the ESA CCI biomass map.

We provide in Figure 2.5 an illustration of the distribution of biomass prediction errors of several model designs on the AGBD dataset [20]. Consistent with canopy height and biomass estimation literature, all models tend to overestimate low values and underestimate high values. This phenomenon can be explained by the long-tailed distribution of biomass across the globe. In other words, because high-biomass locations are extremely rare, models shy away from making predictions in this range. We hope that Embed2Scale embeddings will provide rich representations capable of better capturing the high biomass values and avoiding the low overprediction of low values where no AGB should be predicted.

3.6 Way Forward

We are currently preparing the pipeline for training a model on AGBD and Embed2Scale embeddings. If the pretrained model is trained to only accept a certain Sentinel-2 patch size as input, we will have to download Sentinel-2 patches of the required size, centered at the geolocations of the AGBD dataset. Otherwise, we will use the already available 15x15 or 25x25 patches present in the AGBD dataset. An open question at the moment is the nature of the embeddings. As explained previously, biomass estimation is a dense task: it requires per-pixel predictions, and therefore per-pixel features. How the spatial arrangement of compressed features can be recovered from the embeddings has yet to be clarified. In the event where the spatial arrangement could not be recovered from the compressed embeddings, we would need to re-formulate our downstream task as an image-level biomass estimation. To this end, we would simply regress aggregated biomass statistics (mean, variance, or even histogram) at the scale of a patch.

4 CLIMATE AND AIR POLLUTION PREDICTION FROM SPATIO-TEMPORAL OBSERVATIONAL CONSTRAINTS

4.1 Introduction

This use case is developing downstream applications of embedding models for neural compression in two main strands:

Strand 1 focuses on understanding the utility of embeddings for specific use cases in climate science. The model we are developing in this strand directly enable the Embed2{Infer,Train,Find}, with a particular focus on the ability to derive unsupervised cloud classification schemes and constraining of next generation km-scale/storm-resolving climate models, as developed by e.g. the NextGEMS project (grant number 101003470, coordinated by the Max Planck Institute for Meteorology and the European Center for Medium-Range Weather Forecasts).

In **Strand 2** we are developing a new benchmark for the compression of climate data. This is to ensure that the development of novel neural compression schemes is targeted towards the particular requirements of climate scientists as key user community. This will address a gap in the existing literature that we identified in the review of existing neural compressors we conducted as part of deliverable D2.1. Neural compressors are mainly developed by the machine learning community whose members are generally not experts in climate science and might not be aware of the specific requirements for compressors in the climate domain. With the development of our benchmark, which specifies not just the data that should be compressed but also the evaluation mechanisms that should be used to rank compressors, we give the neural compression community a useful guide and evaluation environment to develop novel neural compression schemes that are targeted directly towards climate science applications. To ensure the benchmark will be relevant to the stakeholders and has community support, we are collaborating with multiple stakeholders that include: members of the ESiWACE3 EU project (grant number 101093054) at the European Centre for Medium-Range Weather Forecasts (ECMWF) and the University of Helsinki, and researchers at ETH Zurich, the UK National Centre for Atmospheric Science, and the US NSF National Center for Atmospheric Research.

4.2 Data

Strand 1 at the moment mainly leverages outgoing longwave radiation (OLR) data derived from the GOES-16 geostationary satellite. However, the methodology that we are developing should be agnostic to the specific choice of satellite. We chose the GOES-16 data source for practical reasons because we already had access to it on our computing infrastructure, so by utilising GOES-16 data we were able to significantly accelerate the implementation process.

For the benchmark in **Strand 2**, we are including a wide variety of data source with heterogeneous spatial correlations and varying grid resolutions; see below Figures 4.1-4.5 for individual visualisations of the fields. The current list of data sources is:

- *GOES-16* OLR data as used in **Strand 1**
- Sea-level pressure and 10m wind vector data from *ERA5 (Figure 4.1)* as published by ECMWF
- Temperature *(Figure 4.2)* and sea-surface temperature data from the *Coupled Model Intercomparison Project Phase 6 (CMIP6)*
- Above-ground biomass (AGB) estimates *(Figure 4.3)* as captured by the *European Space Agency's (ESA) Climate Change Initiative* (https://climate.esa.int/en/)

- OLR *(Figure 4.4)* and precipitation data from the ICON km-scale model developed as part of *NextGEMS project*
- Nitrogen Dioxide reanalysis data *(Figure 4.5)* from the *Copernicus Atmospheric Monitoring Service (CAMS)*

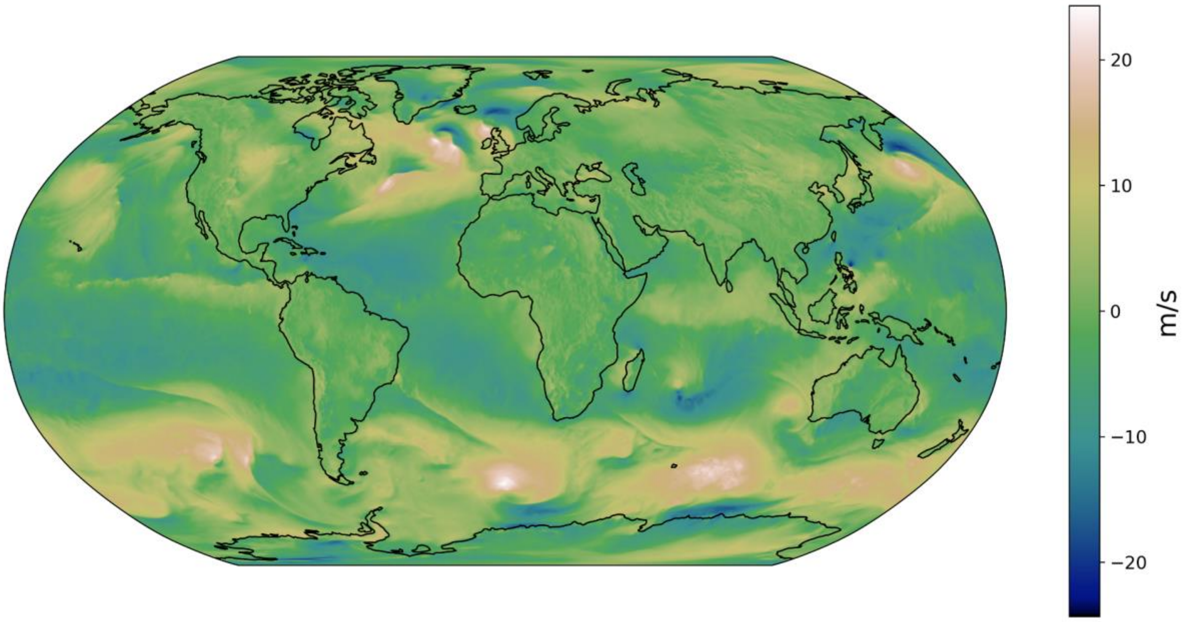


Figure 4.1: Horizontal wind component of the ERA5 dataset.

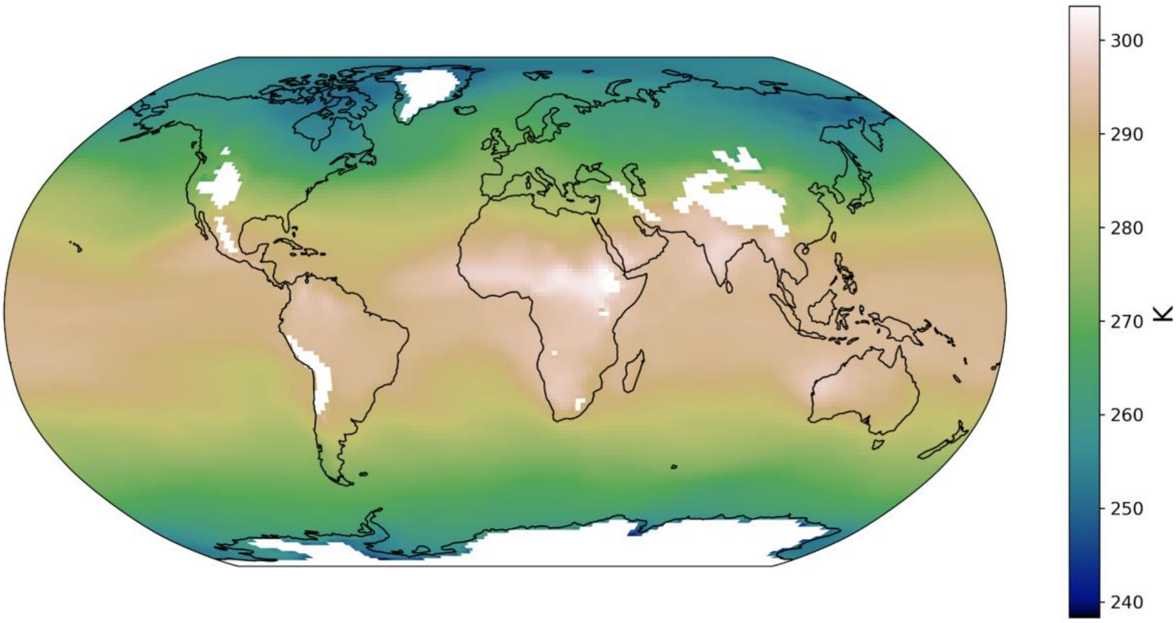
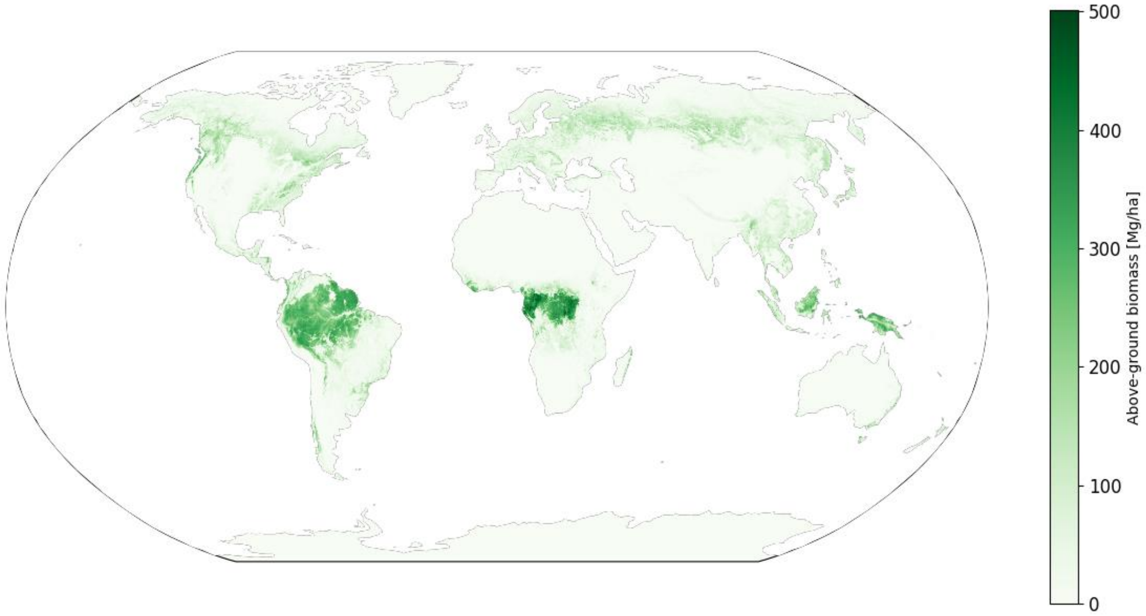


Figure 4.2: Air temperature field at 850hPa as generated by the UKESM1 climate model as part of CMIP6.



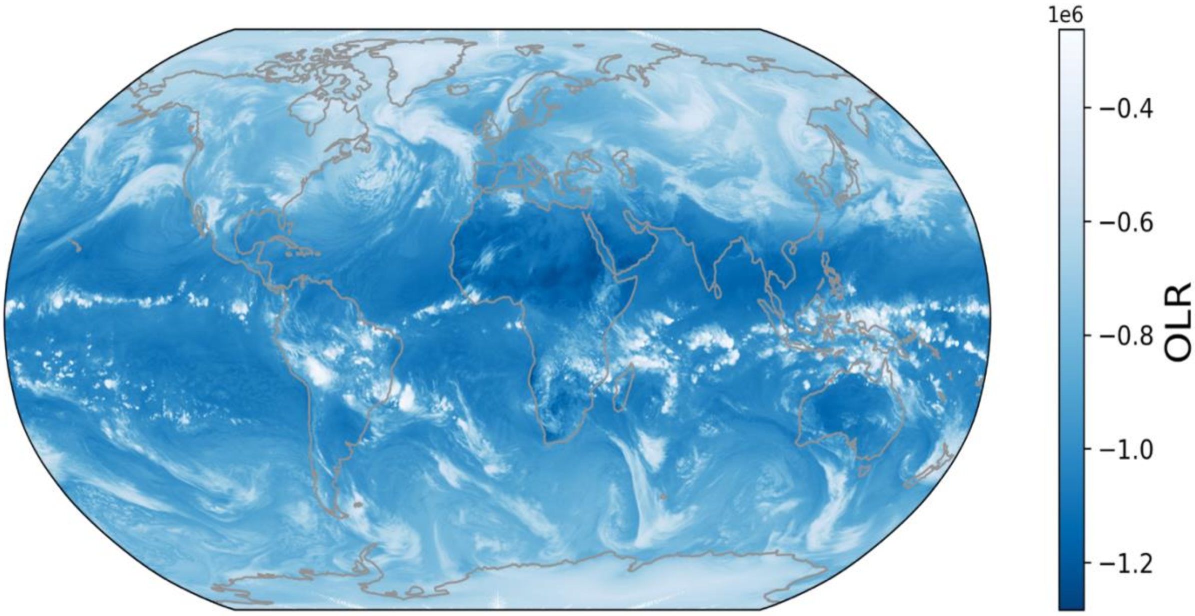


Figure 4.3: Global snapshot of the above ground biomass estimates of the ESA Climate Change Initiative.

Figure 4.4: Outgoing longwave radiation (OLR) field as generated by the ICON model as part of the NextGEMS project.

Funded by the European Union

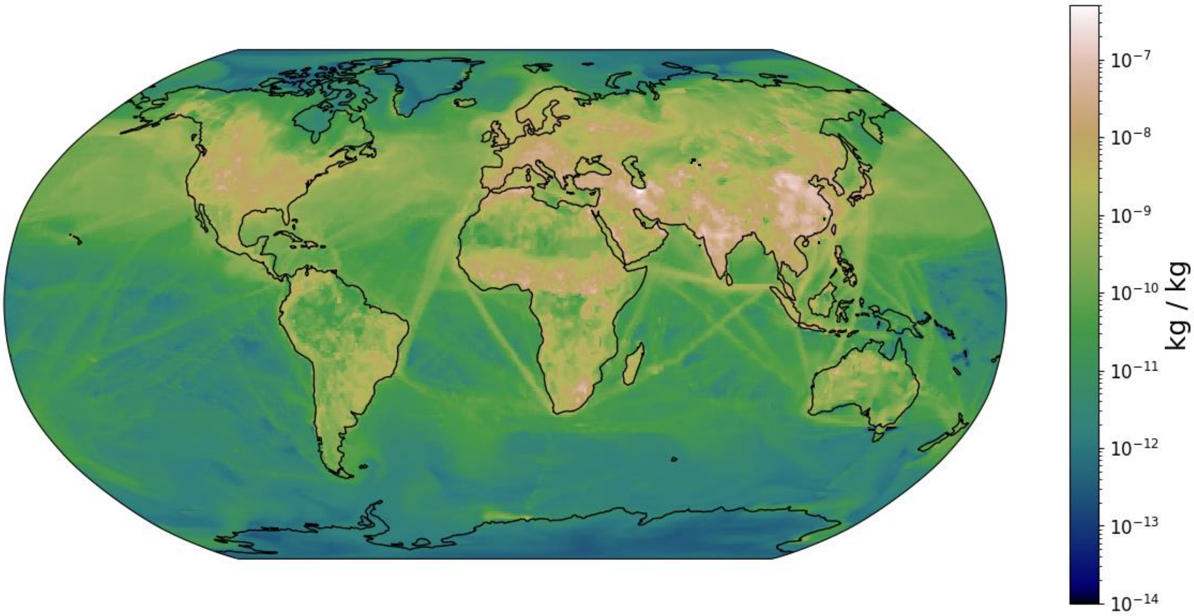


Figure 4.5: Nitrogen Dioxide reanalysis field at 1000hPa pressure level as produced by the Copernicus Atmospheric Monitoring Service (CAMS).

4.3 Methodology

■ Use of embeddings

.

For strand 1, we are pre-processing the GOES-16 OLR data into a dataset of 256 by 256 pixel images. On these images we are training a finite scalar quantized variational auto-encoder (FSQ-VAE) [36] to compress GOES-16 OLR data into a lower-dimensional embedding. Notably, the values FSQ-VAE architecture discretizes the embedding for each image into a collection of discrete tokens; see Figure 4.6 for a visualization of the discretization process. This collection of tokens is then the *latent representation* of the input image. For our use case, we can now discover different cloud regimes through clustering in latent space.

For the compression benchmark, the task is to reconstruct the original input. Standard autoencoders tend to measure reconstruction quality using simple point-wise error metrics such as root mean-squared error (RMSE). While these error metrics are also important for climate applications, they do not constitute a completely sufficient evaluation metric by themselves. For example, autoencoder models often produce reconstructions which look “blurry” to the human eye (see Figure 4.7 for an example) but that lead to low RMSE. However, for high-resolution climate models like the ones developed during the NextGEMS project, it is especially important to capture the “sharp” features from the original inputs, as a lot of computational resources have been used to create the high-resolution features in the first place.

How to ideally choose error metrics to evaluate climate compression algorithms is still a topic of ongoing research. For now, we evaluate the reconstruction quality according to a set of metrics that we chose based on discussion with the stakeholders:

- **d-SSIM:** An adaption of the widely used SSIM measure to scientific floating point data [37]. d-SSIM aims to measure the visual similarity between the compressed and uncompressed data as it would be perceived by the human eye.

- **Peak Signal-to-Noise (PSNR) ratio** is a standard measure in compression research that takes into account both the average pointwise error in the compressed data, as well as the overall dynamic range of the input data.
- **Spatial relative error (SRE).** The SRE measures the proportion of reconstructed data points that exceed a given relative error threshold [38].
- **Spectral Error** calculates the squared error between the power spectrum of the input and output**.** This error metric ensures that *high-frequency features* are preserved.

While the primary goal of the benchmark is to assess neural compressors based on their ability to reconstruct a given input, the benchmark will also allow us to study what computations and analysis can be run in the embedding space of neural compression models. If a model is able to accurately reconstruct the output image based on a given embedding, the embedding should contain enough information about the output to do useful scientific analysis like detecting distribution shifts.

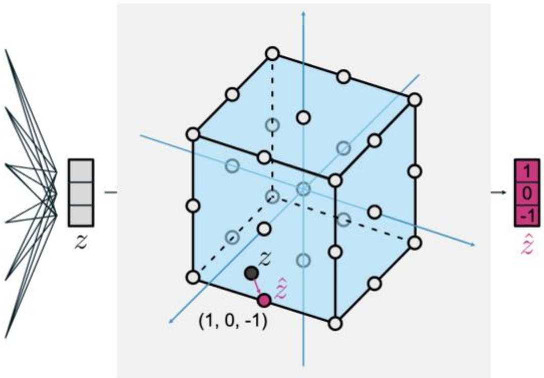


Figure 4.6: Illustration of finite-scalar quantization (FSQ) mechanism. Each latent z gets mapped to one of the equally spaced cluster centres. Crucially, the cluster centres are fixed and not learnable which drastically simplifies the training process. Figure taken from https://arxiv.org/pdf/2309.15505.

4.4 Implementation steps

We have successfully implemented the major steps of our implementation plan outlined in D4.1 (see Table 4.1). Specifically, we have completed the following objectives:

- Implemented the pre-processing pipeline for tiling GOES-16 data into a data format that can be processed by an embedding model
- We have implemented the FSQ-VAE model
- We successfully trained the FSQ-VAE on the GOES-16 data (see discussion of results below) on the JASMIN compute server (https://jasmin.ac.uk/) that allows us to leverage NVIDIA A100 GPUs for accelerated model training. JASMIN also fulfils all the other necessary software requirements we have for our training process.
- We have implemented evaluation procedures for comparing the reconstruction quality of different models

This constitutes the main chunk of engineering work for Strand 1. The next phase of development will now involve refining our training procedure, architecture, and evaluation process to improve the performance of the FSQ-VAE model.

Table 4.1: Implementation steps and current status for Strand 1 of the climate and air pollution prediction use case.

Steps	**Implementation Status**
Pre-processing pipeline for tiling GOES-16 data into data format that can be processed by an embedding model	Done
Implementation of FSQ-VAE model	Done
Train FSQ-VAE model on GOES-16 data	Done
Implement evaluation procedures to assess reconstruction quality	Done
Tune the performance of the FSQ-VAE to improve reconstruction quality	Ongoing
Investigate the utility of using compression metrics to rank unsupervised cloud classification schemes	Pending

For the compression benchmark we have also completed major milestones in our implementation. Notably, we have a data pipeline in place that can handle all the data sources described in Section 4.2. Additionally, we have a set of baseline compressors that we can evaluate on the data sources to get initial results for what compression performance is achievable with current compressors. Currently, we can evaluate the following compressors (the choice of baselines was informed from our review as part of D2.1): ZFP [39], Bit Rounding [40], SZ3 [41], Stochastic Rounding [42], TTHRESH [43]. Initial results are shown in Section 3.5.

Table 4.2: Implementation steps and current status for the compression benchmark of the climate and air pollution prediction use case.

Steps	**Implementation Status**
Collecting feedback on benchmark design from existing compression developers and potential stakeholders	Done
Assemble and curate input datasets for compression benchmark	Done
Evaluate baseline compressors on the input dataset	Done
Implement evaluation procedures to assess reconstruction quality	Done
Prepare benchmark for publication	Ongoing

Funded by the European Union

4.5 Current Results

Figure 4.7 shows initial qualitative results for comparing the compression performance of the FSQ-VAE against the established JPEG2000 compressor. We observe that the FSQ-VAE gives visually quite similar reconstructions. By investigating the power spectrum of the original images and the reconstructed images we observe that JPEG2000 can preserve detail at the smaller scales slightly better. One interesting qualitative difference between the compressor is that the FSQ-VAE seems to "blur" the image quite uniformly while the JPEG2000 compressor has some regions which are more blurred than others. Given that the trained FSQ-VAE should be considered only an initial prototype, the results are generally encouraging. We believe that there is ample room for improvement of the neural compressor by further optimizing the training process and the model architecture.

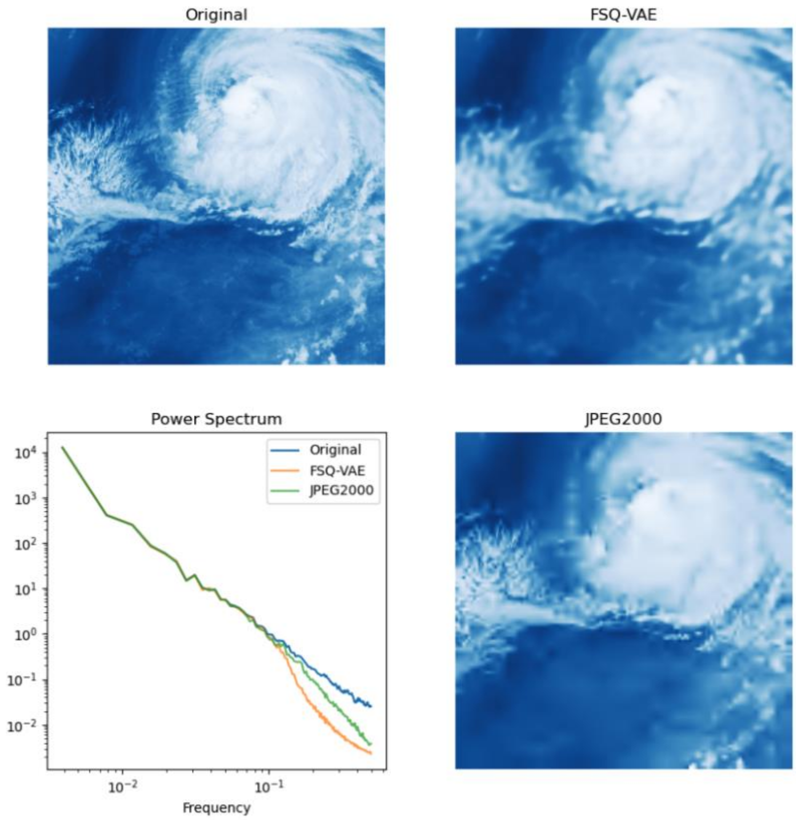
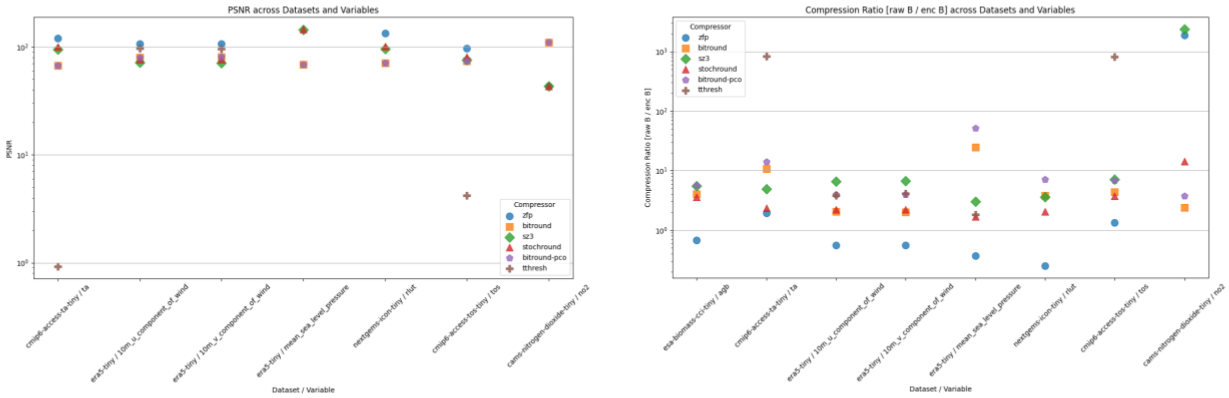


Figure 4.7: Comparison of the FSQ-VAE against JPEG2000 for the GOES-OLR data. Both compressors lead to compression ratios of roughly ~100x (i.e. the compressed representation is 100 times smaller than the original file).



(a) Peak Signal-to-Noise Ratio *(b) Compression Ratio*

Figure 4.8: (a) Peak signal-to-noise ratio (PSNR; higher is better) and (b) compression ratio (higher is better) for each of the climate compression benchmark data sources described in Section 3.2. Different markers represent the different baseline compressors outlined in Section 3.4. Notably, the performance of TTHRESH is significantly lower compared to the other baselines on the CMIP6 datasets (additionally, execution of TTHRESH timed out on the ESA biomass, CAMS nitrogen dioxide, and NextGEMS Icon datasets). While ZFP is able to achieve the highest PSNR values it suffers from low compression ratios, with most compressors achieving a compression ratio below 10.

Initial results for the climate compression benchmark are presented in Figure 4.8. The figure shows the measured peak signal-to-noise ratio (PSNR) achieved by the baseline compressor (Section 4.4) on the data sources described in Section 3.2. We are currently in the process of conducting quality checks to ensure that we configured all the baseline compressors correctly. Notably, Figure 4.8 highlights that using the TTHRESH compressor leads to considerably lower reconstruction quality. We will present the current state of the climate compression benchmark as an *oral presentation* at the *EGU General Assembly 2025* at the session titled “*ESSI1.6 - Machine Learning and Digital Twins for Earth System Observation and Prediction*”.

4.6 Way Forward

The FSQ-VAE model and the wider class of auto-encoder/embedding models will allow us to derive unsupervised cloud classification schemes. However, an open unanswered question is how we can meaningfully rank the quality of different cloud classification schemes. The wealth of different auto-encoder architectures leads to an equally large number of possible cloud classifications and assessing which classification scheme a practitioner should use is not immediately clear. In the near future we will investigate whether reconstruction quality can be used as a comparison metric. This is motivated by a fact that embeddings from models that lead to better reconstructions should capture more meaningful semantic concepts of the input (assuming a fixed embedding dimension).

Our immediate goal is to prepare the climate compression benchmark for publication and make it publicly accessible. Naturally, we are planning to use the benchmark in our own work in the development of neural compression schemes for climate data. In effect the benchmark can be used as a stepping stone to help to generalise the work of strand 1 so that we are building cnbmpressors that are performing well on heterogeneous input sources.

5 CROP STRESS AND YIELD EARLY DETECTION

5.1 Introduction

The new European Green Deal [21] has recently emphasised the need to accelerate the digital and green transitions to achieve a climate-neutral and sustainable economy. European agriculture is consistently impacted by weather extremes [22, 23], which are expected to increase in magnitude and frequency in the near future. The effect of adverse weather conditions on crops strongly depends on their development stage. Timely monitoring systems for crop phenology are essential for understanding and assessing the impact of climate change on crop production [24]. Due to their high temporal frequency and spatial resolution, the Sentinel satellite missions [25] have made significant contributions to agricultural monitoring. Despite the development of crop maps and crop yield forecasting activities at the European scale [27], integrating Earth observation with weather and climate data is necessary to capture the effects of increasing weather extremes on agricultural phenology.

This use case aims to monitor crop phenology at a country or continental scale using Sentinel-2 data fused with climate data (e.g., ERA5) and publicly available European crop-type maps from farmers' declarations. It will use Embed2Transfer to access Sentinel-2 data covering vast spatio-temporal areas, which will be used to define reference curves of crop development under no-stress conditions. With Embed2{Infer, Train}, comparing these reference curves to the early ones will enable the detection of crop stress in a timely manner. The results will be cross-validated with spatially explicit evidence of weather extremes—such as heatwaves, droughts, or storms—utilising data from weather radars and stations. These events will be efficiently identified through **Embed2Find**, allowing for the attribution of crop stress signals to specific adverse climatic conditions.

This use case is expected to support a range of actors in the agricultural community: (i) farmers and agricultural organisations (e.g., improved monitoring/forecasting of field damage assessment), (ii) the public sector responsible for governing the transition of agriculture, (iii) the private sector, including the agricultural technology and machinery industry, seed companies and agribusiness retailers, the agrochemical industry, and the insurance sector for risk management, and (iv) environmental agencies conducting crop forecasting activities.

5.2 Data

This use case aims to implement a scalable, embedding-driven framework for crop monitoring, yield forecasting, and stress detection using multi-source Earth observation (EO) and meteorological datasets. It integrates high-resolution satellite imagery from Sentinel-2 and MODIS, climate variables from ERA5 and Daymet-4, and ancillary datasets like SoilGrids (soil properties) and SMAP (surface soil moisture). Central to this are temporal embedding models that condense spectral and climate time series into compact representations of crop development dynamics. The implementation is divided into three key components:

- Crop Yield Prediction
  This module predicts end-of-season yield by using embeddings derived from time series data of vegetation indices (e.g., EVI, LAI) along with weather and soil information. Temporal encoders, including autoencoders, transformers, or Embed2Train, create latent representations that are input into regression models to estimate yield on field, county, or

national levels. For model training and evaluation, public yield statistics from sources such as USDA QuickStats[1] and the European Union Joint Research Centre (JRC) are used.

- Crop Type Mapping
Leveraging Sentinel-2 time series and reference labels, this module utilises embedding models to capture phenological patterns specific to crop types. The embeddings are classified using supervised or semi-supervised classifiers. We will leverage Sen4Map [46], a large-scale benchmark dataset, which provides 335,125 geo-tagged multispectral patches (64 × 64 pixels) across 28 EU countries, annotated with LUCAS 2018 land cover and land use survey data [47]. This will enable the generation of high-accuracy crop maps even in low-label settings.

- Crop Stress Detection
This component detects early crop stress by comparing real-time embeddings to historical reference curves. Deviations in the embedding space are cross-referenced with environmental extremes, such as droughts and heatwaves, detected using ERA5, SMAP, and Embed2Find models. Tools like Embed2Infer provide early alerts and link stress signals to weather-related causes.

This modular, embedding-based architecture, as shown in Figure 5.1, allows for the scalable monitoring of regions and seasons. It supports precision agriculture, risk management, and the design of climate-resilient policies.

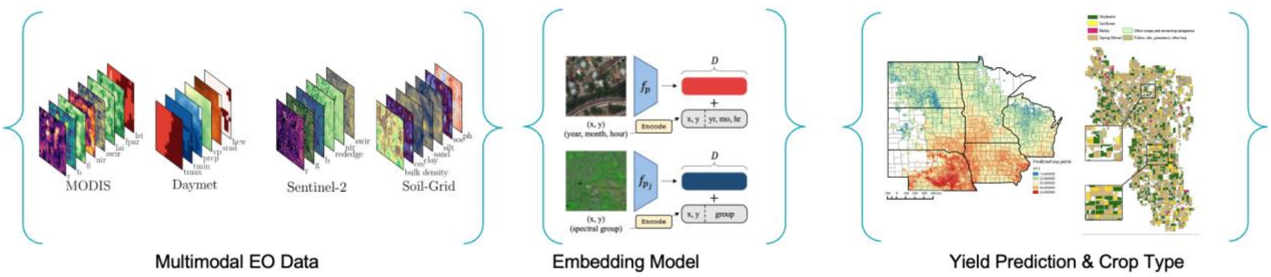


Figure 5.1: Embedding-based crop monitoring framework. Satellite imagery (Sentinel-2, MODIS), weather data (ERA5, Daymet), and ancillary datasets (SoilGrids) are fused and processed through temporal embedding models. Using embeddings will allow the scalability of yield studies by reducing the amount of ingested data as the embeddings retain meaningful features from raw data.

5.3 Methodology

■ Use of embeddings

Embeddings will be generated from the multi-modal datasets used in this task. The goal is to explore deep learning compression algorithms that can embed both the temporal and spatial information of satellite and climate time series datasets. These embeddings would then be evaluated across the three subtasks: crop type mapping, crop yield detection, and stress detection.

While traditional deep-learning compression methods focus on the reconstructed images, this use case will focus on the embeddings and how to reduce data volume while retaining model performance for downstream tasks.

[1] https://quickstats.nass.usda.gov/

5.4 Implementation steps

The implementation steps so far performed are displayed in Table 5.1. The initial steps have involved an exhaustive literature review, encompassing contemporary state-of-the-art machine learning methodologies for various subtasks and compression methods applied to Earth observation (EO) and climate data. We have retrieved and pre-processed key datasets (Sentinel-2, MODIS, Daymet-4, SoilGrids) for the baseline machine-learning pipeline. Baseline experiments for the three subtasks have been established using pre-processed datasets and both classical and deep machine learning. The baseline experiments also included comparing various EO datasets (MODIS and Sentinel-2) for crop yield prediction using explainable AI, to understand the strengths and limitations of each sensor. By identifying optimal data preprocessing techniques and uncovering sensor-specific performance metrics, we can establish a robust preprocessing pipeline for compression methods. The current step in implementing this task involves exploring deep learning compression methods for the subtasks of crop type mapping and crop yield prediction. These include VQGANs, Masked Autoencoders (MAE), Temporal Attention Encoders (TAE), and Variational Autoencoders (VAEs), among others. Once a robust method has been established, the next step will involve scaling the use case to continental-scale crop stress and early yield detection by deploying the pipeline on high-performance computing (HPC) systems.

Table 5.1: Implementation steps and current status for the crop stress and yield early detection use case

Steps	**Implementation Status**
Data acquisition and preprocessing	Done
Training of baseline models on multimodal data for yield prediction and crop type mapping	Done
Experimentation with existing compression and foundation models architectures	Done
Implementing pipeline for training a model on E2S embeddings	Ongoing
Training model on E2S embeddings using satellite, climate and auxiliary data	Pending
Comparison of baseline models and E2S embedding-based models	Pending

5.5 Current results

Figures 5.2 and 5.3 below present the results from the baseline experiments for crop yield prediction and crop type mapping, respectively. They consist of tables detailing model performance for both subtasks across various machine learning algorithms, as measured by accuracy and additional performance metrics, including mean absolute error (MAE), R-squared (R2), F1-score, and root mean square error (RMSE).

Dataset	Model	Cotton (lb/acre)				Corn (Bu/acre)			
		MAE	RMSE	MAPE	R^2	MAE	RMSE	MAPE	R^2
MODIS	EBM	96.75	129.20	0.14	0.74	6.62	8.69	0.04	0.78
	Linear Regression	98.32	135.69	0.14	0.71	7.58	9.65	0.04	0.72
	KNN	115.83	160.23	0.16	0.60	9.46	12.27	0.05	0.56
	Random Forest	97.79	136.36	0.14	0.71	7.86	9.78	0.04	0.72
	XGBoost	105.87	154.30	0.15	0.63	8.07	10.06	0.05	0.70
	LightGBM	97.32	135.00	0.13	0.72	7.38	9.35	0.04	0.74
Sentinel-2	EBM	94.83	123.52	0.13	0.76	**6.26**	**8.40**	**0.03**	**0.79**
	Linear Regression	**92.72**	**121.57**	0.13	**0.77**	6.77	8.64	0.04	0.78
	KNN	115.72	160.11	0.16	0.60	9.44	12.24	0.05	0.56
	Random Forest	90.92	123.72	0.13	0.76	7.10	9.36	0.04	0.74
	XGBoost	101.28	137.93	0.14	0.70	7.86	10.08	0.04	0.70
	LightGBM	93.13	128.50	0.13	0.75	6.84	8.89	0.04	0.76

Figure 5.2: Crop yield prediction baseline experiment comparing the performance of MODIS and Sentinel-2 datasets.

Classes	Random Forest		Transformer (pixel-based)		ViT		VViT	
	Center pixel	3x3 window	Center pixel	3x3 window	Center pixel	3x3 window	Center pixel	3x3 window
Cereals	0.78	0.88	0.83	0.93	0.66	0.78	0.79	0.86
Root crops	0.56	0.69	0.69	0.88	0.24	0.44	0.64	0.75
Industrial crops	0.64	0.76	0.72	0.88	0.42	0.58	0.67	0.78
Dry pulses,flowers	0.07	0.1	0.31	0.68	0.04	0.08	0.18	0.35
Fodder crops	0.01	0.01	0.31	0.64	0.19	0.34	0.37	0.54
Bareland	0.33	0.47	0.40	0.64	0.21	0.37	0.32	0.47
Wood/Shrubland	0.88	0.96	0.90	0.97	0.85	0.93	0.88	0.95
Grassland	0.71	0.88	0.76	0.93	0.66	0.84	0.71	0.87
W.A. F-score	0.77	0.88	0.82	0.93	0.69	0.82	0.76	0.86
Overall Accuracy	0.79	0.89	0.82	0.93	0.70	0.83	0.77	0.87

Figure 5.3: F-score and Overall Accuracy obtained on the crop-type mapping baseline experiment [Sharma et al., 2024].

5.6 Way forward

Subsequent actions include finalising experiments with different deep-learning-based compression algorithms and testing these for the different subtasks. Crop yield forecasting experiments have, up until this point, focused on the continental United States; the next goal is to explore domain adaptation by shifting focus to European continent yield studies. While we are internally exploring different compression methods for this use case, we are also awaiting the embeddings from the upstream activities.

6 CONCLUSIONS

This deliverable described the current implementation status of the different use cases within E2S. While each use case has its own goals, development steps and level of maturity, it can be observed that most of the efforts of the first development cycle were devoted to collect, pre-process and analyze the input data. The embeddings selection is an ongoing process for most of the use cases, which will be applied in the second phase of the project. It is expected to have also a significant contribution on the embeddings selection from the **CVPR EarthVision 2025 Data Challenge** currently ongoing.

Funded by the European Union

7 REFERENCES

[1] Gerard Margarit et al., "Operational ship monitoring system based on synthetic aperture radar processing", Remote Sensing, 2009

[2] Li, J., Xu, C., Su, H., Gao, L., & Wang, T. (2022). Deep learning for SAR ship detection: Past, present and future. *Remote Sensing*, *14*(11), 2712.

[3] Changren Zhu et al., "A Novel Hierarchical Method of Ship Detection from Spaceborne Optical Image Based on Shape and Texture Features," IEEE Transactions on Geoscience and Remote Sensing, 2010

[4] Ciocarlan, A., & Stoian, A. (2021). Ship detection in sentinel 2 multi-spectral images with self-supervised learning. Remote Sensing, 13(21), 4255.

[5] Ramona-Maria Pelich et al., "AIS based Evaluation of Target Detectors and SAR Sensors Characteristics for Maritime Surveillance", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015

[6] Ai, J., Pei, Z., Yao, B., Wang, Z., & Xing, M. (2021). AIS data aided Rayleigh CFAR ship detection algorithm of multiple-target environment in SAR images. IEEE Transactions on Aerospace and Electronic Systems, 58(2), 1266-1282.

[7] Jiexiong Tang et al., "Compressed-domain ship detection on spaceborne optical image using deep neural network and extreme learning machine", IEEE Transactions on Geoscience and Remote Sensing, 2015.

[8] Jiang, P., Ergu, D., Liu, F., Cai, Y., & Ma, B. (2022). A Review of Yolo algorithm developments. Procedia computer science, 199, 1066-1073.

[9] Pieter De Frenne et al., "Global buffering of temperatures under forest canopies", Nature Ecology & Evolution, 2019

[10] Ralph Dubayah et al., "The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography", Science of remote sensing, 2020

[11] Nguyen Tien Hoang et al., "Mapping the deforestation footprint of nations reveals growing threat to tropical forests", Nature Ecology & Evolution, 2021

[12] Walter Jetz et al., "Essential biodiversity variables for mapping and monitoring species populations", Nature ecology & evolution, 2019

[13] Nikolai Kalischek et al., "Cocoa plantations are associated with deforestation in Cˆote d'Ivoire and Ghana", Nature Food, 2023

[14] Nico Lang et al., "A high-resolution canopy height model of the Earth", Nature Ecology & Evolution, 2023

[15] Peter Manning et al., "Redefining ecosystem multifunctionality", Nature ecology & evolution, 2018

[16] Mirco Migliavacca et al., "The three major axes of terrestrial ecosystem function", Nature, 2021

[17] Peter Potapov et al., "Mapping global forest canopy height through integration of GEDI and Landsat data", Remote Sensing of Environment, 2021

Funded by the European Union

[18] Cecile Renier et al., “Transparency, traceability and deforestation in the Ivorian cocoa supply chain”, Environmental Research Letters, 2023

[19] Benjamin Brede et al., “Non-destructive estimation of individual tree biomass: Allometric models, terrestrial and UAV laser scanning”, Remote Sensing of Environment, 2022

[20] Ghjulia Sialelli et al., "AGBD: A Global-scale Biomass Dataset", arXiv preprint arXiv:2406.04928, 2024

[21] Fetting, Constanze. "The European green deal." ESDN Report, December 2, no. 9 (2020).

[22] Beillouin, Damien, Bernhard Schauberger, Ana Bastos, Phillipe Ciais, and David Makowski. "Impact of extreme weather conditions on European crop production in 2018." Philosophical Transactions of the Royal Society B 375, no. 1810 (2020): 20190510.

[23] Hristov, Jordan, Andrea Toreti, I. Pérez Domínguez, Franciscus Dentener, Thomas Fellmann, Christian Elleby, Andrej Ceglar et al. "Analysis of climate change impacts on EU agriculture by 2050." Publications Office of the European Union, Luxembourg, Luxembourg (2020).

[24] White, Michael A., Forrest Hoffman, William W. Hargrove, and Ramakrishna R. Nemani. "A global framework for monitoring phenological responses to climate change." Geophysical Research Letters 32, no. 4 (2005).

[25] Berger, Michael, Jose Moreno, Johnny A. Johannessen, Pieternel F. Levelt, and Ramon F. Hanssen. "ESA's sentinel missions in support of Earth system science." Remote sensing of environment 120 (2012): 84-90.

[26] Van der Velde, M., and L. Nisini. "Performance of the MARS-crop yield forecasting system for the European Union: Assessing accuracy, in-season, and year-to-year improvements from 1993 to 2015." Agricultural Systems 168 (2019): 203-212.

[27] Justice, C. O., Louis Giglio, S. Korontzi, J. Owens, J. T. Morisette, D. Roy, J. Descloitres, S. Alleaume, F. Petitcolin, and Y. Kaufman. "The MODIS fire products." Remote sensing of Environment 83, no. 1-2 (2002): 244-262.

[28] Bolton, Douglas K., and Mark A. Friedl. "Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics." Agricultural and forest meteorology 173 (2013): 74-84.

[29] Khabbazan, Saeed, Paul Vermunt, Susan Steele-Dunne, Lexy Ratering Arntz, Caterina Marinetti, Dirk van der Valk, Lorenzo Iannini, Ramses Molijn, Kees Westerdijk, and Corné van der Sande. "Crop monitoring using Sentinel-1 data: A case study from The Netherlands." Remote Sensing 11, no. 16 (2019): 1887.

[30] Celik, Mehmet Furkan, Mustafa Serkan Isik, Gulsen Taskin, Esra Erten, and Gustau Camps-Valls. "Explainable artificial intelligence for cotton yield prediction with multisource data." IEEE Geoscience and Remote Sensing Letters (2023).

[31] Bruhwiler, Kevin, Paahuni Khandelwal, Daniel Rammer, Samuel Armstrong, Sangmi Lee Pallickara, and Shrideep Pallickara. "Lightweight, embeddings based storage and model construction over satellite data collections." In 2020 IEEE International Conference on Big Data (Big Data), pp. 246-255. IEEE, 2020.

[32] Schneider, Maja, Amelie Broszeit, and Marco Körner. "Eurocrops: A pan-european dataset for time series crop type classification." arXiv preprint arXiv:2106.08151 (2021).

[33] d'Andrimont, Raphaël, Astrid Verhegghen, Guido Lemoine, Pieter Kempeneers, Michele Meroni, and Marijn Van Der Velde. "From parcel to continental scale–A first European crop type map based on Sentinel-1 and LUCAS Copernicus in-situ observations." Remote sensing of environment 266 (2021): 112708.

[34] Esser, Patrick, Robin Rombach, and Bjorn Ommer. "Taming transformers for high-resolution image synthesis." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 12873-12883. 2021.

[35] Stewart, Adam J et al., "Torchgeo: deep learning with geospatial data", Proceedings of the 30th International Conference On Advances In Geographic Information Systems, 2022

[36] Mentzer, Fabian, David Minnen, Eirikur Agustsson, and Michael Tschannen. "Finite Scalar Quantization: VQ-VAE Made Simple." In The Twelfth International Conference on Learning Representations. 2024.

[37] Baker, Allison H., Alexander Pinard, and Dorit M. Hammerling. "On a structural similarity index approach for floating-point data." *IEEE Transactions on Visualization and Computer Graphics* 30, no. 9 (2023): 6261-6274.

[38] Baker, Allison H., Haiying Xu, Dorit M. Hammerling, Shaomeng Li, and John P. Clyne. "Toward a multi-method approach: Lossy data compression for climate simulation data." In *High Performance Computing: ISC High Performance 2017 International Workshops, DRBSD, ExaComm, HCPM, HPC-IODC, IWOPH, IXPUG, P^ 3MA, VHPC, Visualization at Scale, WOPSSS, Frankfurt, Germany, June 18-22, 2017, Revised Selected Papers 32*, pp. 30-42. Springer International Publishing, 2017.

[39] Lindstrom, Peter. "Fixed-rate compressed floating-point arrays." *IEEE transactions on visualization and computer graphics* 20, no. 12 (2014): 2674-2683.

[40] Zender, Charles S. "Bit Grooming: statistically accurate precision-preserving quantization with compression, evaluated in the netCDF Operators (NCO, v4. 4.8+)." *Geoscientific Model Development* 9, no. 9 (2016): 3199-3211.

[41] Liang, Xin, Kai Zhao, Sheng Di, Sihuan Li, Robert Underwood, Ali M. Gok, Jiannan Tian et al. "Sz3: A modular framework for composing prediction-based error-bounded lossy compressors." *IEEE Transactions on Big Data* 9, no. 2 (2022): 485-498.

[42] Croci, Matteo, Massimiliano Fasi, Nicholas J. Higham, Theo Mary, and Mantas Mikaitis. "Stochastic rounding: implementation, error analysis and applications." *Royal Society Open Science* 9, no. 3 (2022): 211631.

[43] Ballester-Ripoll, Rafael, Peter Lindstrom, and Renato Pajarola. "TTHRESH: Tensor compression for multidimensional visual data." *IEEE transactions on visualization and computer graphics* 26, no. 9 (2019): 2891-2903.

[44] Poggio, Laura, et al. "SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty." Soil 7.1 (2021): 217-240.

[45] Entekhabi, Dara, et al. "The soil moisture active passive (SMAP) mission." Proceedings of the IEEE 98.5 (2010): 704-716.

[46] Sharma, Surbhi, et al. "Sen4map: Advancing mapping with sentinel-2 by providing detailed semantic descriptions and customizable land-use and land-cover data." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (2024).

Funded by the European Union

[47] D'Andrimont, Raphaël, et al. "Harmonised LUCAS in-situ land cover and use database for field surveys from 2006 to 2018 in the European Union." Scientific data 7.1 (2020): 352.

[48] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Las Vegas, 27-30 June 2016, 779-788.

Funded by the European Union