D6 – Technical Documentation for Trained Regional Model v1

Deliverable Overview

Lead(s)	HZDR, FZJ
Contributor(s)	
Work Package	WP5.1 (Model implementation, technical side)
Stage	Foundation Stage (I)
Duration	PM 1 (05/2024) – PM 12 (04/2025; planned), PM 17 (11/2025; revised)
Dependencies	WP5.1 (D6) depends on this deliverable
Outcome (Type)	Report
Link	
Status	☐ To be done ☐ In progress ☒ Completed

Executive Summary

This deliverable provides the complete technical documentation for the 3D-ABC regional foundation model released under Deliverable D5. It describes the environment setup, data preprocessing workflows, model configuration, tokenization pipeline, and inference procedures used in pretraining the first regional version of the 3D-ABC foundation model. The documentation builds on the ml-4m framework developed by Apple and EPFL and its adaptation within the 3DABC-MODEL environment for multimodal Earth Observation data.

Results

The report corresponding to D6 details the system configuration, training environment, and data workflow used to develop the regional 3D-ABC foundation model. The model was trained in a Python 3.11 environment with an editable ml-4m installation and dependencies such as torch, torchgeo, mpi4py, and cmake, optimized for a CUDA-enabled HPC cluster.

The pretraining used TanDEM-X (20 m) and Harmonized Landsat and Sentinel (HLS, 30 m) datasets, converted into WebDataset format for efficient distributed processing. A Diffusion VAE tokenizer was trained with a 16 × 16 patch size on 128 × 128 tiles, generating 64 tokens per tile for consistent multimodal representation.

The public GitHub repository provides model weights, tokenizers, configuration files, and setup documentation, ensuring transparency and full reproducibility. This setup will be reused for regional and later global-scale training within the 3D-ABC framework.