Advancing Global Carbon Stock Mapping with the 3DABC Foundation Model

Gunjan Joshi*†, **Weikang Yu***†, Aldino Rizaldy†, Peter Steinbach†, Pedram Ghamisi†

†Helmholtz-Zentrum Dresden-Rossendorf, Dresden, Germany

*Main authors; G. Joshi is the presenting author

Foundation models (FMs) are large-scale deep learning architectures that typically follow a two-phase training process: pre-training on vast domain-agnostic datasets, followed by fine-tuning for specific tasks. This pretrain-then-finetune paradigm enhances performance and robustness by leveraging self-supervised learning techniques, allowing models to extract meaningful representations from unlabeled data and reducing the dependence on large labeled datasets, which had been a major hurdle in training past models. While some recent foundation models emphasize zero-shot learning for a variety of tasks, fine-tuning remains essential in many cases to adapt the models for specific, task-oriented applications. In Earth observation, FMs have been successful in tasks such as global land cover mapping, fire scar detection, and flood mapping, though multimodal data integration remains a challenge in this domain.

The global carbon budget plays a crucial role in understanding carbon sources and sinks, providing essential insights into the global carbon cycle. Accurate carbon stock mapping is vital for both scientific and economic purposes, informing emission pathways and climate policies. A foundation model for mapping global soil and vegetation carbon stocks using remote sensing and field data would mark a significant step forward. Building on this need, the Helmholtz Foundation Model Initiative is developing the 3D Above and Below Ground Carbon Stocks FM (3D-ABC FM) to enable precise, high-resolution mapping of terrestrial carbon stocks. By integrating multimodal datasets, the model enhances understanding of carbon distribution across various land types thereby supporting climate strategies.

The integration of multiple remote sensing data modalities necessitates the use of a framework that can effectively manage multimodality. 3D-ABC FM builds upon the Massively Multimodal Masked Modeling (4M) framework, which has recently demonstrated state-of-the-art performance. The 4M architecture employs a unified Transformer encoder-decoder with a masked modeling objective, processing diverse input and output modalities, including text, images, and geometric data. A multi-stage training process further optimizes its representations across different modalities. By adapting the 4M framework, 3D-ABC FM ensures seamless integration of multimodal remote sensing datasets, advancing carbon stock mapping capabilities.

On behalf of the HFMI - 3DABC Project. The 3D-ABC project is funded by the Helmholtz Foundation Model Initiative supported by the Helmholtz Association.