# Multi-temporal and multi-modal Earth Observation latent space decoding using physically aware Deep Learning

December 19, 2023

## 1 Introduction

This offer is part of the RELEO (REpresentation Learning for Earth Observation) of ANITI-2, the follow-on of the Interdisciplinary Artificial Intelligence Institute in the frame of the French ANR "AI Clusters". The PhD will be funded by CNES (the French Space Agency) and Thales Alenia Space.

RELEO aims at building an AI foundation model for the exploitation of Earth Observation Satellite Image Time Series (EO SITS). This model will fuse multi-modal data (optical, SAR, thermal) into AI-Ready chunks of latent features (also known as embeddings), where traditional spatial, temporal and spectral dimensions of Earth Observation data have been collapsed. These latent chunks will benefit from the complementarity and the correlations of these different data sources and will provide essentialized (fully encoding the useful information) and strongly compressed information. The fusion will be done with deep neural networks whose training will be guided by physical models of the observed processes (bio/geo-physical models) and image formation models (radiative transfer and sensor models).

## 2 Focus of the PhD

The proposed PhD subject will contribute to RELEO's WP3 whose goal is decoding the dimensionless compressed embeddings to produce time series of input variables for the physical models used to constrain the training problem through self-supervision. Some of these variables will be Essential Climate and Biodiversity Variables [1], as for instance soil moisture, land surface temperature or leaf area index. The work consists in building a neural decoder for the data cube unfolding: going from the compressed multi-modal data fusion back to the temporal and spatial resolutions needed by downstream applications. The decoder training will be done through the optimization of the above-mentioned physical models' output. Reference bio/geo-physical products from operational mono-sensor processing chains will be used together with field surveys to validate the decoder outputs.

To give a concrete example, one can start with 2 sources of data, high-resolution optical Sentinel-2 (10 m. resolution every 5 days) and radar Sentinel-1 (10 m. resolution every 6 days) time series covering a full country like France. These are encoded as spatio-temporal chunks of 1km × 1km × 10 days as vectors of 256 features. The goal is to decode these vectors in order to produce weekly Leaf Area Index maps at 10 m. resolution.

The originality of the proposed approach is two-fold.

- First of all, the use of physical models to constrain the learning (PINN, Physically Informed Neural Networks [2]) is very recent in EO [3, 4] and its extension to the multi-modal case is a challenging problem, but it will allow the development of downstream application processors which are independent of the availability of a particular sensor.

- Second, the generation of data at tailored temporal and spatial resolutions (including spatial resolutions not acquired at the chosen time stamps) allows to foresee an accuracy improvement of the downstream products by avoiding repeated resampling steps which are known to degrade the input data. This will entail the development of neural architectures which are able to produce image time series from embeddings which don't have explicit spatial nor temporal dimensions. Generative models either based on diffusion approaches [5], [6], normalizing flows [7] or similar approaches will be investigated.

**Work plan:**

1. State of the art on foundation models and PINN starting from what is already available at CESBIO.
2. Method development: latent space decoding for on-demand resolution using physical model guided training.
3. Validation and assessment using reference products and field surveys.

## 3 Work environment

The PhD will take place at Cesbio[1] in Toulouse. The PhD candidate will be integrated into the *Observation Systems* team and more precisely, within the AI unit.

The team works on CNES' (the French Space Agency) high performance computing (HPC) infrastructure (250 nodes with 8000 CPU, 53 GPU) which also hosts a full mirror of all Sentinel-1 and Sentinel-2 data.

## 4 Applications

**Candidate profile:** Masters in at least one of the following areas: applied mathematics, physics of measure, optimization, machine learning. Skills in and eagerness for computer programming in the areas of scientific computing or machine learning.

Send Curriculum Vitae, motivation letter and recommendation letters to jordi.inglada@cesbio.eu before February 28 2024.

## References

[1] Stephan Bojinski, Michel Verstraete, Thomas C. Peterson, Carolin Richter, Adrian Simmons, and Michael Zemp. The concept of essential climate variables in support of climate research, applications, and policy. *Bull. Amer. Meteor. Soc.*, 95(9):1431–1443, 2014.

[2] M. Raissi, P. Perdikaris, and G.E. Karniadakis. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, Feb 2019.

[1] `https://www.cesbio.cnrs.fr/homepage/`

[3] Yoël Zérah, Silvia Valero, and Jordi Inglada. Physics-driven probabilistic deep learning for the inversion of physical models with application to phenological parameter retrieval from satellite times series. *IEEE Transactions on Geoscience and Remote Sensing*, 61:1–23, 2023.

[4] Yoël Zérah, Silvia Valero, and Jordi Inglada. Physics-constrained deep learning for biophysical parameter retrieval from Sentinel-2 images: inversion of the PROSAIL model. working paper or preprint, December 2023.

[5] Xizewen Han, Huangjie Zheng, and Mingyuan Zhou. Card: Classification and regression diffusion models. *CoRR*, 2022.

[6] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.

[7] Danilo Rezende and Shakir Mohamed. Variational inference with normalizing flows. In *International conference on machine learning (ICML)*, pages 1530–1538, 2015.