
Latent Field Reduction of Earth Observation Foundation Model

Lars Uebbing¹, Harald L. Joakimsen¹, Theodor J. L. Forgaard², Michael C. Kampffmeyer^{1,2}, Kristoffer K. Wickstrøm¹, Sébastien Lefèvre^{1,3}, Arnt-Børre Salberg², Robert Jenssen^{1,2,4}

¹UiT The Arctic University of Norway, ²Norwegian Computing Center

³University of South Brittany, ⁴University of Copenhagen

Abstract

Foundation Models for Earth Observation encode multimodal inputs into high-dimensional latent representations that can be viewed as discrete samples of a latent vector field over the Earth’s surface. Traditional latent space analyses using methods like PCA or t-SNE overlook this spatial continuity. We propose a neural-operator-based framework to model these latent representations as continuous fields, preserving spatial correlations and enabling more coherent interpretation. This approach links discrete latent encodings and the continuous geophysical processes they represent, advancing the understanding of latent structures in Earth observation foundation models.

1 Introduction

Foundation Models (FMs) have demonstrated remarkable success in encoding multimodal inputs into high-dimensional latent vectors, capturing rich information useful for a wide range of downstream tasks [Bommasani et al., 2022, Smith et al., 2024]. In the context of Earth Observation (EO), these latent vectors v_i correspond to discrete points x_i on the continuous domain Ω representing the Earth, allowing us to interpret them as discrete samples of an underlying latent vector field $v : \Omega \rightarrow \mathbb{R}^{d_v}$ [Bodnar et al., 2025, Danish et al., 2025].

Understanding the structure of this latent space is crucial: it provides insight into how FMs represent and organize information, and can reveal relationships between inputs, disentangle semantic factors, and guide improvements in model design and interpretability. Traditionally, analysis of latent spaces relies on dimensionality reduction techniques, such as Principal Component Analysis (PCA), t-distributed stochastic neighbor embedding (t-SNE) [Pareek and Jacob, 2021], or Uniform Manifold Approximation and Projection (UMAP) [McInnes et al., 2020], which project high-dimensional vectors into a lower-dimensional space for visualization or clustering. While these approaches are effective for exploring global or local structure, they treat latent vectors as isolated points and typically ignore the continuous nature of the underlying domain. Further they struggle with over-clustering, which yields misleading representations of data [Liu et al., 2025].

To address this limitation, we propose leveraging Neural Operators (NOs) [Li et al., 2020, Lu et al., 2021, Li et al., 2021, Azizzadenesheli et al., 2024, Wang and Wang, 2024] to learn a continuous, lower-dimensional representation of the latent vector field. By explicitly modeling the latent space as a field over Ω , neural operators allow us to capture spatial continuity and correlations, enabling more faithful interpolations and analysis of the structure of the latent space across the entire Earth surface. This approach bridges the gap between discrete latent representations and the continuous phenomena they encode, offering a principled framework for interpreting and manipulating latent structures in Earth Observation Foundation Models (EOFMs).

2 Theoretical Background and Methodology

2.1 EOFMs – Latent Vector Field Encoding

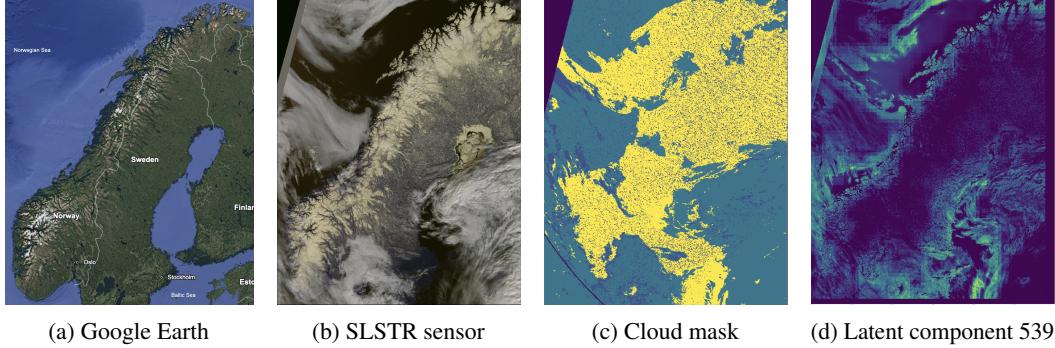


Figure 1: Input data for latent vector field reduction. For reference, (a) shows a Google Earth screenshot from the region of which data was sampled. (b)–(d) represent data from an exemplary day of 2018. (b) shows the RGB band of Sentinel 3’s SLSTR sensor. (c) provides a cloud mask for the given recording. (d) shows a plot of a randomly chosen component of FM4CS’ latent vector field.

Recently, a variety of EOFMs has emerged, focussing on different types of data and applications. For this case study we employ FM4CS¹ —a Foundational Model for Climate and Society, encoding data from Sentinel 1,2 and 3 satellite images of different resolution. The FM4CS also supports data from the Sea and Land Surface Temperature Radiometer (SLSTR), with 6 optical and 3 thermal bands in 500 m and 1000 m ground sampling distance, respectively. FM4CS uses a ViT architecture to encode combined Sentinel images associated with a given region into a 768 dimensional latent vector and is trained on several tasks including pixel reconstruction and prediction of climate variables provided by the ERA5 dataset².

In this study we will work with latent representations generated by FM4CS for daily Sentinel-3 SLSTR data over Scandinavia from 2018, see Figure 1. We have selected 4 optical (S1 – S3 and S5) and 2 thermal (S7 and S8) bands. The area is covered by 599×791 points, resulting in a resolution of approximately 2–3 km. The 768 dimensional latent vectors are regarded as discrete realizations $v_i = v(x_i)$ of the latent vector field $v(x)$ with $x \in \Omega$ representing the domain (here: Norway).

2.2 Neural Operators

NOs are a type of machine learning frameworks that—in contrast to traditional methods, which learn to approximate functions to map from data space to data space—approximate operators \mathcal{L} mapping from function to function $(\mathcal{L}v)(x) = u(x)$ [Azizzadenesheli et al., 2024]. Therefore, NOs work in a conceptually different setting, as they assume a continuous underlying function to the data provided to them. While this has some practical advantages like resolution independence, it also captures the nature of many problems better. Viewing the latent space of EOFMs as a continuous vector field—essentially a function—NOs are the natural choice to learn its underlying structure.

We choose the NO implementation as Graph Kernel Operator (GKO) [Li et al., 2020], a Graph neural network consisting of three main components:

1. **Lifting:** $h_i = Pv_i$, a single neural layer P , mapping node features (here: latent vectors) v_i into a potentially higher-dimensional latent space H .
2. **Message-passing:**

$$h_i^{(t+1)} = \sigma \left(Wh_i^{(t)} + \sum_{j \in N_i} \frac{K_\phi(e(i, j))h_j^{(t)}}{|N_i|} \right), \quad (1)$$

where N_i denotes neighboring points of x_i and K_ϕ is a learnable kernel.

¹FM4CS presentation: <https://nr.no/en/publication/10255345/>

²<https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5>

3. **Projection:** $\hat{u}(x) = Qh_i^{(T)}$, mapping back to the output space.

This framework aligns naturally with existing dimensionality reduction methods and allows clearer interpretation than more recently developed NO frameworks.

2.3 NO based Vector Field Reduction

In order to use the GKO framework for latent vector field reduction, we need to construct a graph based on a set of realized vectors $\{v_i\}_{i=1}^N$. We construct a graph with one node for each realized location x_i on our domain Ω , see Figure 2. Each node is described by a set of features, which in this case is given by the corresponding latent vector v_i . Each node x_i is connected to its k nearest neighbors $x_j \in N_i$ with respect to the spatial domain Ω . We choose $k = 5$ to keep the model slim and computationally efficient. The edges are themselves equipped with a vector of edge-attributes $e(i, j)$ that encodes the positional encoding of x_i , as well as the difference of the positional encodings $(x_i - x_j)$. This allows the network to learn spatial relations and patterns in order to predict meaningful kernel matrices $K_\phi(e(i, j))$.

For more efficient computation, P first maps the 768 dimensional node-features v_i into a 32 dimensional vector space $h_i \in \mathcal{H}$. After the message-passing steps, these nodes are then projected into the final low-dimensional space. Here we choose to map the data into a 3 dimensional space, which allows us to plot the result as a color map by interpreting each dimension as a RGB color-channel. Taking inspiration from t-SNE, we use the Kullback-Leibler divergence between input and output distributions as a loss function. For training, we generate graphs by sampling 10,000 random locations to realize the vector field. The training dataset is composed of 20 graphs per day from January to March 2018.

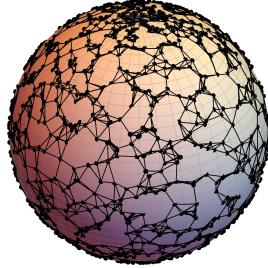


Figure 2: Graph structure on domain (Earth).

3 Results

To evaluate the NO reduction, we sample graphs from dates outside the training set (April–May 2018) and compare results against the standard methods: PCA, t-SNE and UMAP. For a first qualitative comparison we sample data from 01.04.2018 as for this day most of the region was recorded by the satellite and cloud coverage is relatively low. We sample 100,000 random latent field vectors and reduce them to three dimensions. The computation times show that PCA is most efficient (~ 3 sec.), followed by the NO method (~ 11 sec.) and UMAP (~ 1 min.), t-SNE being by far the slowest method (~ 15 min.).

Results are plotted by interpreting the dimensions as RGB color channels as shown in Figure 3, together with the SLSTR sensor image for comparison. As UMAP completely failed to create any meaningful results (see Figure 3e), new sets of data points were sampled until UMAP yielded something meaningful (see Figure 3f). Visual inspection of the results shows PCA and our NO method yield comparable results, that structurally resemble the SLSTR image. In contrast, t-SNE

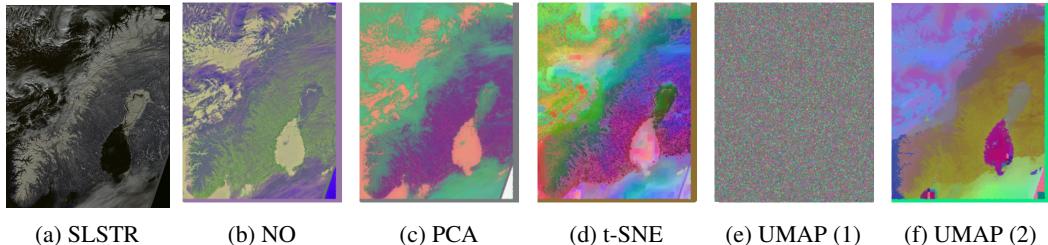


Figure 3: FM4CS latent space representation for 01.04.2018 for 100,000 randomly sampled latent vectors. (a) shows the original SLSTR image. (b)–(f) represent the results of a reduction into 3 dimensions (interpreted as RGB channels) using different methods. A second sample was created for UMAP (f) as this method failed to find a meaningful embedding for the original sample (e).

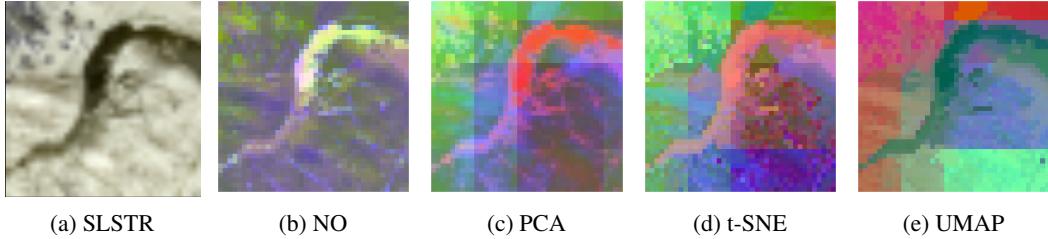


Figure 4: Dimensionality reduction for densely sampled 1600 latent vectors for a small region, representing a “zoom” into a small region of interest. Plots compare (4a) the original SLSTR image and (4b–4e) reduction for the same samples across methods.

and UMAP yield strong clustering, including very strictly rectangular shaped clusters. These are likely triggered by underlying patches going into the ViT of FM4CS. This also leads to artifacts in the resulting figures that can be seen in particular in the Gulf of Bothnia. Another obvious difference is that PCA and the NO method seem to encode similar regions in similar vectors, i.e., clouds, water, snow, etc. always yield shades of the same color, while t-SNE and UMAP provide different encoding for the same type of region within the same image. This makes PCA and NO results much more interpretable and consistent.

To illustrate the strength of NOs for dimensionality reduction in the context of EO, we create a realistic scenario by “zooming” into a potential region of interest (here: a random region of approx. $100 \text{ km} \times 100 \text{ km}$) and sample a high resolution grid of latent vectors. We produce 3 dim. embeddings and plot the results as a RGB map like before and compare across methods together with the original SLSTR data, shown in Figure 4. While the NO map shows smooth, continuous results that are consistent with previous findings from Figure 3, all traditional methods show artifacts like rectangular clusters with sharp border and strong fluctuations between individual neighboring pixels, making the resulting low-dim. vector fields discontinuous and disrupting smoothness. In addition to this, the maps show that the color scheme of the NO map is consistent with results from Figure 3 (water \rightarrow yellow, land \rightarrow green, ice \rightarrow purple, etc.), making results directly comparable, which is not given for the traditional methods. This widens the range of applications and increases practicality of the NO method.

The results found by the above experiments align with our theoretical arguments for using NOs for the vector space reduction. The continuous structure of NOs prevent the method from over-clustering and produces smooth results, much closer to what would be expected for the representation of a physical system. By using a neural network to predict the kernel matrix for the message-passing steps (Eq. 1), the model effectively learns the structure of the underlying domain, making this the only one of the above discussed methods that takes spatial relations into account. In future work we will asses the advantages of the NO method more quantitatively and apply it to explicit downstream tasks like generating predictors for variable forecasting. Taking into account the flexibility of NOs in terms of resolution as well as the fact that they can be fine-tuned for specific tasks, we expect this method to perform strong in a wide range of EO applications.

4 Conclusion

In this work we proposed NOs as a method for dimensionality reduction to analyze latent spaces of EOFMs. We argued that due to their continuous nature, NOs are more suitable to retain the structure of the latent space than traditional methods. We provided a concrete way of implementing NOs for domain-aware, continuous dimensionality reduction using a GKO approach and showcased experimental results for the latent space of FM4CS. Results show that NOs are more efficient than t-SNE or UMAP and yield more interpretable and consistent results. Furthermore, the NO method proved to be the only amongst the tested methods that did not yield strong artifacts (pixel and cluster wise) due to over-clustering and discontinuities. Future work will provide a more quantitative analysis of these advantages. Due to their flexibility and continuous, domain-aware nature, we find NOs to be an important tool to find low-dimensional representations in the field of EO. In future work we will make use of this directly by using NO outputs for downstream tasks.

Acknowledgments and Disclosure of Funding

This work was partially funded by the Research Council of Norway (Visual Intelligence, grant no. 309439 and KnowEarth, grant no. 337481).

References

- Kamyar Azizzadenesheli, Nikola Kovachki, Zongyi Li, Miguel Liu-Schiaffini, Jean Kossaifi, and Anima Anandkumar. Neural operators for accelerating scientific simulations and design. *Nature Reviews Physics*, 6(5):320–328, 2024. doi: 10.1038/s42254-024-00712-5. URL <https://doi.org/10.1038/s42254-024-00712-5>.
- Cristian Bodnar, Wessel P. Bruinsma, Ana Lucic, Megan Stanley, Anna Allen, Johannes Brandstetter, Patrick Garvan, Maik Riechert, Jonathan A. Weyn, Haiyu Dong, Jayesh K. Gupta, Kit Thambiranam, Alexander T. Archibald, Chun-Chieh Wu, Elizabeth Heider, Max Welling, Richard E. Turner, and Paris Perdikaris. A foundation model for the earth system. *Nature*, 641(8065):1180–1187, 2025. doi: 10.1038/s41586-025-09005-y. URL <https://doi.org/10.1038/s41586-025-09005-y>.
- Rishi Bommasani, Drew A. Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S. Bernstein, Jeannette Bohg, Antoine Bosselut, Emma Brunskill, Erik Brynjolfsson, Shyamal Buch, Dallas Card, Rodrigo Castellon, Niladri Chatterji, Annie Chen, Kathleen Creel, Jared Quincy Davis, Dora Demszky, Chris Donahue, Moussa Doumbouya, Esin Durmus, Stefano Ermon, John Etchemendy, Kawin Ethayarajh, Li Fei-Fei, Chelsea Finn, Trevor Gale, Lauren Gillespie, Karan Goel, Noah Goodman, Shelby Grossman, Neel Guha, Tatsunori Hashimoto, Peter Henderson, John Hewitt, Daniel E. Ho, Jenny Hong, Kyle Hsu, Jing Huang, Thomas Icard, Saahil Jain, Dan Jurafsky, Pratyusha Kalluri, Siddharth Karamcheti, Geoff Keeling, Fereshte Khani, Omar Khattab, Pang Wei Koh, Mark Krass, Ranjay Krishna, Rohith Kuditipudi, Ananya Kumar, Faisal Ladhak, Mina Lee, Tony Lee, Jure Leskovec, Isabelle Levent, Xiang Lisa Li, Xuechen Li, Tengyu Ma, Ali Malik, Christopher D. Manning, Suvir Mirchandani, Eric Mitchell, Zanele Munyikwa, Suraj Nair, Avanika Narayan, Deepak Narayanan, Ben Newman, Allen Nie, Juan Carlos Niebles, Hamed Nilforoshan, Julian Nyarko, Giray Ogut, Laurel Orr, Isabel Papadimitriou, Joon Sung Park, Chris Piech, Eva Portelance, Christopher Potts, Aditi Raghunathan, Rob Reich, Hongyu Ren, Frieda Rong, Yusuf Roohani, Camilo Ruiz, Jack Ryan, Christopher Ré, Dorsa Sadigh, Shiori Sagawa, Keshav Santhanam, Andy Shih, Krishnan Srinivasan, Alex Tamkin, Rohan Taori, Armin W. Thomas, Florian Tramèr, Rose E. Wang, William Wang, Bohan Wu, Jiajun Wu, Yuhuai Wu, Sang Michael Xie, Michihiro Yasunaga, Jiaxuan You, Matei Zaharia, Michael Zhang, Tianyi Zhang, Xikun Zhang, Yuhui Zhang, Lucia Zheng, Kaitlyn Zhou, and Percy Liang. On the opportunities and risks of foundation models, 2022. URL <https://arxiv.org/abs/2108.07258>.
- Muhammad Sohail Danish, Muhammad Akhtar Munir, Syed Roshaan Ali Shah, Muhammad Haris Khan, Rao Muhammad Anwer, Jorma Laaksonen, Fahad Shahbaz Khan, and Salman Khan. Terrafm: A scalable foundation model for unified multisensor earth observation, 2025. URL <https://arxiv.org/abs/2506.06281>.
- Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: Graph kernel network for partial differential equations, 2020. URL <https://arxiv.org/abs/2003.03485>.
- Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations, 2021. URL <https://arxiv.org/abs/2010.08895>.
- Zhexuan Liu, Rong Ma, and Yiqiao Zhong. Assessing and improving reliability of neighbor embedding methods: a map-continuity perspective. *Nature Communications*, 16(1):5037, 2025. doi: 10.1038/s41467-025-60434-9. URL <https://doi.org/10.1038/s41467-025-60434-9>.
- Lu Lu, Pengzhan Jin, Guofei Pang, Zhongqiang Zhang, and George Em Karniadakis. Learning nonlinear operators via deeponet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3(3):218–229, March 2021. ISSN 2522-5839. doi: 10.1038/s42256-021-00302-5. URL <http://dx.doi.org/10.1038/s42256-021-00302-5>.

Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction, 2020. URL <https://arxiv.org/abs/1802.03426>.

Jyoti Pareek and Joel Jacob. Data compression and visualization using pca and t-sne. In Vishal Goar, Manoj Kuri, Rajesh Kumar, and Tomonobu Senju, editors, *Advances in Information Communication Technology and Computing*, pages 327–337, Singapore, 2021. Springer Singapore. ISBN 978-981-15-5421-6.

Michael J. Smith, Luke Fleming, and James E. Geach. Earthpt: a time series foundation model for earth observation, 2024. URL <https://arxiv.org/abs/2309.07207>.

Tian Wang and Chuang Wang. Latent neural operator for solving forward and inverse pde problems, 2024. URL <https://arxiv.org/abs/2406.03923>.