Environmental Data Science (2026), 5: e16, 1–18
doi:10.1017/eds.2026.10043

CAMBRIDGE UNIVERSITY PRESS

APPLICATION PAPER

One stone three birds: Three-dimensional implicit neural network for compression and continuous representation of multi-altitude climate data

Alif Bin Abdul Qayyum[1], Xihaier Luo[2], Nathan M. Urban[2], Xiaoning Qian[1,2,3] and Byung-Jun Yoon[1,2]

[1]Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA
[2]Computing and Data Sciences, Brookhaven National Laboratory, Upton, NY, USA
[3]Department of Computer Science and Engineering, Texas A&M University, College Station, TX, USA
Corresponding author: Byung-Jun Yoon; Email: bjyoon@tamu.edu

Received: 05 November 2025 UTC; **Revised:** 21 April 2026 UTC; **Accepted:** 29 May 2026 UTC

Keywords: continuous super-resolution; cross-modal prediction; dimensionality reduction; multi-modal representation learning; scientific data compression

Abstract

Wind energy stands out as a promising clean and renewable energy alternative, not only for its potential to combat global warming but also for its capacity to meet the ever-growing demand for energy. However, analysis of wind data to fully harness the benefits of wind energy demands tackling several related challenges: (1) Current data resolution is inadequate for capturing the detailed information needed across diverse climatic conditions; (2) Efficient management and storage of real-time measurements are currently lacking; (3) Extrapolating wind data across spatial specifications enables analysis at costly-to-measure, unobserved points is necessary. In response to these challenges, we introduce the One Stone Three Bird model, a modality-agnostic learning framework utilizing Implicit Neural Network. Our model effectively compresses a large volume of climate data into a manageable latent codec. It also learns underlying continuous climate patterns, enabling reconstruction at any scale and supporting modality transfer and fusion. Extensive experimental results show consistent performance improvements over existing baselines in both (1) continuous super-resolution reconstruction and (2) data compression tasks for different cross-altitude prediction scenarios. Through systematic ablation studies, we demonstrate the effectiveness of each core component, quantifying its individual contribution to the overall performance of the proposed design.

Impact Statement

For climate scientists, the twin goals of shrinking large-volume wind data for efficient storage and providing the full, high-resolution detail required for precise calculation are difficult to reconcile. In this work, we propose a deep learning method-based multi-altitude wind data processing approach that compresses the high-resolution, sensor-acquired discrete data and reconstructs the compressed data into a continuous domain, both simultaneously and in a cross-altitude manner. Extensive empirical evaluations validate the superiority of our proposed approach over existing deep learning and physics-based approaches.

This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/eds.2026.10043 Published online by Cambridge University Press

1. Introduction

The escalating challenge of climate change necessitates immediate and strategic action to mitigate its impacts and steer toward sustainable development (Kaack et al., 2022; Rolnick et al., 2022). As the Earth's climate experiences higher temperatures, rising sea levels, and more extreme weather events, the transition to renewable energy sources becomes inevitable. Wind energy, in particular, offers a notable solution with its ability to deliver a clean power supply while greatly cutting down greenhouse gas emissions (Stengel et al., 2020; Ramesh et al., 2022a; Ramesh et al., 2022b).

However, the deployment and optimization of wind energy face several challenges:

1. *Resolution*: Identifying the most suitable sites for wind turbines necessitates data with a resolution as detailed as 1 square kilometer or finer (Irrgang et al., 2021; Kashinath et al., 2021). However, most wind farm simulations do not achieve this resolution, limiting our capacity to enhance the efficiency of wind energy farms.
2. *Data Storage*: As the granularity of simulated data and the accumulation of field measurements increase (Klöwer et al., 2021; Huang and Hoefler, 2023), the resulting growth in data size requires advanced storage solutions.
3. *Generalization*: Setting up wind measurement stations in specific locations can be challenging due to the high costs of transportation and maintenance. This necessitates cross-altitude inference, such as estimating high-altitude wind speeds from ground-level measurements.

Advancements in deep learning present promising solutions to these challenges. Techniques such as deep learning-based super-resolution can enhance low-resolution data, providing the detailed representations needed for precise analysis (Vandal et al., 2017; Diaconu et al., 2022; Gao et al., 2022). Additionally, deep learning-based data reduction can compress extensive datasets into latent formats, reducing memory and hardware demands. However, most existing deep learning approaches, like convolutional neural networks and autoencoders, are grid-based and fail to offer a continuous representation of wind fields (Requena-Mesa et al., 2021; Nguyen et al., 2023). Due to the inherently continuous nature of wind fields, there is a crucial need for methodologies that can generate and work with continuous data representations (Reichstein et al., 2019; Luo et al., 2023; Luo et al., 2024). Additionally, recent advancements in multi-modal deep learning necessitate its use in the efficient and comprehensive analysis of wind data, including the examination of wind patterns at multiple altitudes (Summaira et al., 2021; Xu et al., 2023).

In this paper, we present OSTB (One Stone Three Bird), a novel deep learning model designed for the efficient dimension reduction and continuous reconstruction of multi-altitude climate data. Our approach utilizes modality-agnostic implicit neural networks within an *encoder-decoder (transfer)-decoder (continuous reconstructor)* framework to process multi-altitude climate data. The *encoder* segment of the three-dimensional implicit neural network functions as a nonlinear data compressor, and the *decoder (transfer)* segment functions as a modality transfer network that exploits inter-altitude data relationships to improve feature extraction. The *decoder (continuous reconstructor)* employs implicit neural representations to reconstruct continuous fields. Overall, our contributions are as follows:

- We design a novel, parameter-efficient deep *encoder-decoder (transfer)-decoder (continuous reconstructor)* framework for simultaneous dimension reduction and continuous reconstruction via super-resolution of multi-altitude climate data.
- We design a three-dimensional implicit neural network for transforming data from one modality/altitude to another modality/altitude. Due to this structure of our designed implicit neural network, our proposed model is scalable to datasets consisting of a large number of modalities, unlike traditional multi-modal deep learning models which face severe scalability issues.
- We employ GAAM (Gaussian Adaptive Attention Mechanism) (Ioannides et al., 2024) in the *decoder (transfer)* segment of our proposed three-dimensional implicit neural network as a parameter-efficient alternative to regular query-key-value based attention mechanisms (Vaswani et al., 2017), which to the best of our knowledge has not been tested in a super-resolution task.

https://doi.org/10.1017/eds.2026.10043 Published online by Cambridge University Press

- We employ a KAN (Kolmogorov-Arnold Network) (Liu et al., 2025) as a superior alternative to the traditional multi-layer perceptron neural network (Hornik et al., 1989) within the state-of-the-art LIIF (Local Implicit Image Function) (Chen et al., 2021) based *decoder (continuous reconstructor)*, which also has not been employed in data reduction or super-resolution tasks.

2. Related work

Climate Downscaling is a critical process in climate science, allowing for the translation of global climate model outputs into finer, local-scale projections. The main methods for downscaling can be categorized into dynamical downscaling and statistical downscaling (Keller et al., 2022; Harder et al., 2023). Dynamical downscaling, while comprehensive, demands significant computational power and depends on the accuracy of the global climate model data (Chau et al., 2021; Chen et al., 2022). Statistical downscaling, in contrast, is more computationally efficient and faster but assumes historical relationships will persist, potentially missing changes in climate variability and extremes (Groenke et al., 2020; Liu et al., 2020).

Alternatively, **Deep Learning-based Super-resolution** techniques are revolutionizing the enhancement of climate data, delivering unparalleled detail and precision in climate models and remote sensing imagery. Different deep neural networks specialize in analyzing both spatial and temporal climate data's complex patterns, enhancing model accuracy and details beyond what traditional downscaling achieves (Vandal et al., 2017; Stengel et al., 2020; Requena-Mesa et al., 2021; Diaconu et al., 2022; Gao et al., 2022). However, these methods often rely on fixed resolutions, highlighting the need for models that offer resolution-independent, continuous climate pattern representations (Luo et al., 2023).

Implicit Neural Representation (INR) uses neural networks to model continuous signals, transcending traditional discrete methods like pixel and voxel grids (Xie et al., 2022; Huang and Hoefler, 2023). This approach has recently made significant strides in climate data analysis, enabling high-resolution reconstructions beyond fixed enhancement scales. In Luo et al., 2024, a context-aware indexing mechanism was introduced to enhance the efficiency of INR in reconstructing fields from sparse observations. In Schwarz et al., 2023, a novel compression algorithm is introduced, utilizing INR within a universal approach to data handling, that effectively generates compact yet comprehensive latent depictions of ERA5 climate data. These advancements highlight INR's capability in precise, scalable data representation. In the meantime, Neural Operator and INR have been applied for solving PDEs and enhancing spatio-temporal resolution (Kovachki et al., 2021; Li et al., 2021). INR-based PDE solvers like MeshfreeFlowNet (Jiang et al., 2020) and MAgnet (Boussif et al., 2022) reconstruct continuous spatio-temporal data from sparse or discrete low-resolution inputs but lack focus on cross-spatial or cross-temporal reconstruction.

Deep Multi-modal Learning transforms machine perception by concurrently integrating diverse data sources such as text, images, audio, and video, unlike traditional single-modal approaches. Multi-modal deep learning has shown impressive performance across different domains (Ngiam et al., 2011; Jing et al., 2021; Boussioux et al., 2022; Cao and Gao, 2022; Tu et al., 2022; Ramesh et al., 2022a, 2022b; Ruan et al., 2023). In Qayyum et al., 2024, a novel multi-modal deep learning method is introduced for simultaneous dimension reduction and continuous cross-altitude reconstruction. However, as the number of modalities increases, the aforementioned multi-modal deep learning models encounter scalability issues, highlighting the need for a parameter-efficient multi-modal model capable of facilitating cross-altitude inference.

3. Preliminaries

3.1. Multi-altitude as multi-modality

Traditionally, multi-modal data refers to the scenario of different data types from different modalities, e.g., image/text, video/text. Similar to the definition used by Qayyum et al., 2024, we follow a more flexible definition of multi-modality where data from each modality is acquired through different sensors, and we consider multi-altitude climate data as multi-modal data with each altitude as a separate modality.

https://doi.org/10.1017/eds.2026.10043 Published online by Cambridge University Press

3.2. Gaussian adaptive attention mechanism

Conventional attention mechanism calculates weights based on the dot-product between different weight matrices (Vaswani et al., 2017), whereas the Gaussian adaptive attention mechanism employs a Gaussian-based modulation of input features, enabling improvement of the standard self-attention mechanism in Transformers along with a reduced number of trainable parameters and lower computational cost (Ioannides et al., 2024).

In GAAM (Gaussian Adaptive Attention Mechanism), the multi-channel spatial feature $F = \{\mathbf{f}_1, \mathbf{f}_2, \cdots, \mathbf{f}_c\}$ (c as the number of channels of the multi-channel spatial feature) goes through the process of computation of mean $\widehat{\mu} = \frac{1}{c}\sum_{s=1}^{c} \mathbf{f}_s$ and variance $\widehat{\sigma}^2 = \frac{1}{c}\sum_{s=1}^{c} (\mathbf{f}_s - \widehat{\mu})^2$. The channel mean $\widehat{\mu}$ is then adjusted by a learnable offset δ to learn $\psi = \delta + \widehat{\mu}$. Then, the attention $g_a(\mathbf{f}_s)$ is computed through normalized $\mathbf{f}_s$, $\mathbf{f}_{norm,s}$, where $\mathbf{f}_{norm,s} = \frac{\mathbf{f}_s - \psi}{\sqrt{\widehat{\sigma}^2 + \epsilon}}$: $g_a(\mathbf{f}_s) = e^{-\frac{\mathbf{f}_{norm,s}}{2\zeta^2}}$, with a small $\epsilon > 0$.

The output of the Gaussian adaptive attention block $\mathbf{G}$ is described as:

$$\mathbf{G}(F) = \{\mathbf{f}_1 \odot g_a(\mathbf{f}_1), \mathbf{f}_2 \odot g_a(\mathbf{f}_2), \cdots, \mathbf{f}_s \odot g_a(\mathbf{f}_s), \cdots, \mathbf{f}_c \odot g_a(\mathbf{f}_c)\} \tag{3.1}$$

3.3. Three-dimensional positional encoder

Three-dimensional positional encoder $\mathbf{P}$ employs Fourier based positional encoding on 3D coordinate points (Tancik et al., 2020). Considering $\mathbf{v} = (x, y, h)$ to denote a 3D coordinate point, with $x^{(c)} = (x, y)$ being a point in the 2D coordinate space $\mathbf{X}_c$, and h the altitude of the corresponding modality, the output of the 3D positional encoder is: $\mathbf{P}(\mathbf{v}) = \left[\cdots, \cos\left(2\pi\sigma^{j/m}\mathbf{v}\right), \sin\left(2\pi\sigma^{j/m}\mathbf{v}\right), \cdots\right]$ for $j \in \{0, 1, \cdots, m-1\}$, where m denotes the number of frequencies of the learned Fourier features, and σ denotes the frequency constant.

3.4. Kolmogorov-Arnold network

Traditional multi-layer perceptrons (MLPs) model non-linearity through fixed activation functions on nodes (Hornik et al., 1989), whereas KANs employ learnable activation functions on edges (Z Liu et al., 2025). There is no linear weight whatsoever in KANs, each weight parameter is substituted with a univariate function that is parameterized as a spline function instead (Boor, 1978). For a traditional MLP, the output at node j of layer $l+1$, can be defined as

$$x_{l+1,j} = \sigma\left(\sum_{i=1}^{n_l} w_{l+1,i,j} x_{l,i}\right) \tag{3.2}$$

with σ as a non-linear activation function. For a KAN, the output at node j of layer $l+1$, can be defined as

$$x_{l+1,j} = \sum_{i=1}^{n_l} \phi_{l+1,i,j}(x_{l,i}), \tag{3.3}$$

where the non-linear function ϕ can be defined as $\phi(x) = w(b(x) + \text{spline}(x))$, with $b(x) = \text{silu}(x) = x/(1 + e^{-x})$ and $\text{spline}(x)$ a linear combination of B-splines, $\sum_i c_i B_i(x)$ with trainable c_i's.

3.5. Local implicit image function

Local implicit image function (LIIF) (Chen et al., 2021) is a deep learning-based continuous representation method from a discrete image or a multi-channel two-dimensional feature representation. A decoding function $\mathbf{D}_{\theta,liif}$ is typically parameterized as an MLP and takes the form $s = \mathbf{D}_{\theta,liif}(z, x)$, where z is the observation of image pixel values or features at a 2D coordinate point x. Consider, our objective is to predict the output at point $x^{(c)}$, with observed features z_t at neighboring coordinate points $x^{(t)}, t = \{00, 01, 10, 11\}$. S_t is the area of the rectangle between $x^{(c)}$ and $x^{(t')}$ where t' is diagonal to t (i.e. 00 to 11, 10 to 01). Then, LIIF predicts the output at $x^{(c)}$ as:

https://doi.org/10.1017/eds.2026.10043 Published online by Cambridge University Press

$$s\left(x^{(c)}\right)=\sum_{t=\{00,01,10,11\}}\frac{S_t}{S}\cdot\mathbf{D}_{\theta,liif}\left(z_t,x^{(c)}-x^{(t)}\right) \tag{3.4}$$

$S=\sum_t S_t$ is the total area of the four rectangles. Figure 1 illustrates this area-based interpolation through $\mathbf{D}_{\theta,liif}$. Instead of applying the area-based interpolation through $\mathbf{D}_{\theta,liif}$ directly on the input image pixels, multi-channel features z_t are first extracted using an EDSR (Lim et al., 2017) model and then decoded through $\mathbf{D}_{\theta,liif}$.

4. Problem statement

Let $\mathbf{x}_k^H\in\mathcal{M}_k^H\subset\mathbb{R}^{h\times w}$ denote the discrete high resolution data representation of modality k with $\mathcal{M}_k^H$ being the discrete high resolution data space of modality k. We aim to achieve simultaneous data dimension reduction and cross-modal/altitude continuous reconstruction of this data instance. Our goal is to design a model capable of pertaining the following tasks:

1. **Data Dimension Reduction**: Execute $\mathbf{x}_k^H\rightarrow\mathbf{x}_k^L\in\mathcal{M}_k^L\subset\mathbb{R}^{\frac{h}{d}\times\frac{w}{d}}$, where $\mathbf{x}_k^L$ represents the discrete low resolution representation of $\mathbf{x}_k^H$ with $\mathcal{M}_k^L$ being the discrete low resolution data space of modality k and d being the dimension reduction factor.
2. **Cross-Altitude/Modality Continuous Reconstruction**: Execute $\mathbf{x}_k^L\rightarrow\mathbf{x}_{l\neq k}^C\left(x^{(c)}\right)\in\mathcal{M}_l^C\left(x^{(c)}\right)\subset\mathbb{R}$. Let, $x^{(c)}\in\mathbf{X}_c$ be any coordinate point in the continuous 2D coordinate space $\mathbf{X}_c$. Here $\mathbf{x}_l^C\left(x^{(c)}\right)$ represents the target value at coordinate point $x^{(c)}$ for modality l and $\mathcal{M}_l^C\left(x^{(c)}\right)$ represents the data space of target values. The continuous nature of $\mathbf{X}_c$ makes this task a continuous reconstruction from low dimensional representation $\mathbf{x}_k^L$.

We will use cross/multi-altitude and cross/multi-modality interchangably throughout the rest of the paper.

5. Method

Overview. Our proposed *encoder-decoder (transfer)-decoder (continuous reconstructor)* framework, $\mathbf{T}_\theta$, comprises three primary components:

- ***encoder:*** The three-dimensional implicit neural network, $\mathbf{E}_\phi$, reduces dimension of the data from input modality/altitude $\mathbf{x}_{in}^H$ to $\mathbf{x}_{in}^L$. This compression is optimized to retain sufficient information for downstream tasks. h_{in} being the altitude of the input modality, we can define the encoder mapping: $\mathbf{x}_{in}^L=\mathbf{E}_\phi\left(\mathbf{x}_{in}^H,h_{in}\right),\mathbf{E}_\phi:\mathcal{M}_{in}^H\rightarrow\mathcal{M}_{in}^L$.

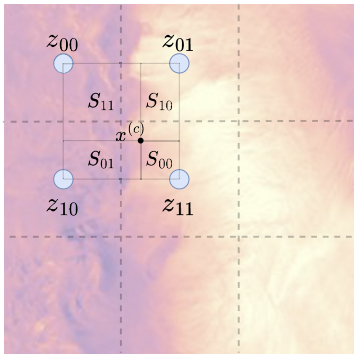


Figure 1. *LIIF method.*

https://doi.org/10.1017/eds.2026.10043 Published online by Cambridge University Press

- ***decoder (transfer):*** The three-dimensional implicit neural network, $\mathbf{D}_{tr}$, transforms the discrete low resolution representation of the input modality/altitude $\mathbf{x}_{in}^{L}$ to the discrete high resolution representation of the target modality/altitude $\mathbf{x}_{out}^{H}$. h_{out} being the altitude of the target modality, we can define the mapping: $\mathbf{x}_{out}^{H} = \mathbf{D}_{tr}\left(\mathbf{x}_{in}^{L}, h_{out}\right), \mathbf{D}_{tr} : \mathcal{M}_{in}^{L} \rightarrow \mathcal{M}_{out}^{H}$.
- ***decoder (continuous reconstructor):*** The implicit continuous decoder, $\mathbf{D}_{cr}$, that utilizes the modality/altitude-transferred discrete high resolution representation $\mathbf{x}_{out}^{H}$ to predict wind data at specific coordinate $\mathbf{x}_{out}^{C}\left(x^{(c)}\right)$.We can define this mapping: $\mathbf{x}_{out}^{C}\left(x^{(c)}\right) = \mathbf{D}_{cr}\left(\mathbf{x}_{out}^{H}, x^{(c)}\right), \mathbf{D}_{cr} : \mathcal{M}_{out}^{H} \times \mathbf{X}_{c} \rightarrow \mathcal{M}_{out}^{C}$.

Figure 2 summarizes the proposed model, $\mathbf{T}_{\theta} \coloneqq \mathbf{D}_{cr} \circ \mathbf{D}_{tr} \circ \mathbf{E}_{\phi}$. We describe the *encoder-decoder (transfer)* segment, $\mathbf{D}_{tr} \circ \mathbf{E}_{\phi}$ in Section 5.1 and *decoder (continuous reconstructor)* segment $\mathbf{D}_{cr}$ in Section 5.2.

5.1. Three-dimensional encoder-transfer implicit neural network

3D ETINN (Three-Dimensional Encoder-Transfer Implicit Neural Network) consists of the first two segments of the overall framework: (i) *encoder:* a 3D *encoder* INN, $\mathbf{E}_{\phi}$, and (ii) *decoder (transfer)*: a 3D *decoder (transfer)* INN, $\mathbf{D}_{tr}$. Both the *encoder* and *decoder (transfer)* consist of (i) a convolution block, **Conv**, and several (ii) residual implicit neural blocks, $\mathbf{R}_{inn}$. 3D *encoder* INN, $\mathbf{E}_{\phi}$, consists of a convolutional block, $\mathbf{Conv}_{in}$, followed by several residual implicit neural blocks, $\mathbf{R}_{inn,j=\{1,2,\cdots,4\}}^{in}$, and can be described as

$$\mathbf{E}_{\phi} \coloneqq \mathbf{R}_{inn,4}^{in} \circ \cdots \circ \mathbf{R}_{inn,2}^{in} \circ \mathbf{R}_{inn,1}^{in} \circ \mathbf{Conv}_{in} \tag{5.1}$$

3D *decoder (transfer)* INN, $\mathbf{D}_{tr}$, consists of several residual implicit neural blocks, $\mathbf{R}_{inn,j=\{1,2,\cdots,4\}}^{out}$, followed by a convolutional block, $\mathbf{Conv}_{out}$, and can be described as

$$\mathbf{D}_{tr} \coloneqq \mathbf{Conv}_{out} \circ \mathbf{R}_{inn,4}^{out} \circ \cdots \circ \mathbf{R}_{inn,2}^{out} \circ \mathbf{R}_{inn,1}^{out} \tag{5.2}$$

We employed four residual implicit neural blocks in our proposed method. We varied the number of residual implicit neural blocks in our method and report the results in Supplementary Materials.

5.1.1. Residual implicit neural block

Residual implicit neural block $\mathbf{R}_{inn}$ consists of a Gaussian adaptive attention block $\mathbf{G}$, a three-dimensional positional encoder $\mathbf{P}$, and a 2-layer MLP network $\mathbf{D}_{inn}$. Conventional implicit neural networks use a two-

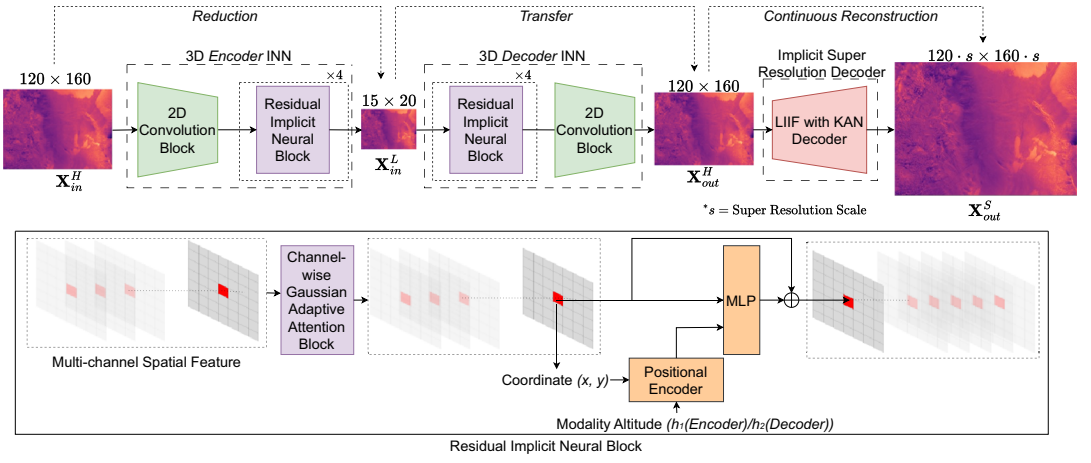


Figure 2. *Overview of the proposed method, which jointly enables data reduction, transfer across modalities, and continuous representation and arbitrary-scale super resolution.*

https://doi.org/10.1017/eds.2026.10043 Published online by Cambridge University Press

dimensional positional encoder to process grid-like 2D data, such as images. In contrast, the residual implicit neural block operates on two-dimensional data while utilizing a positional encoder, **P**, that accepts three-dimensional coordinate points as input. This approach allows the flexibility in transforming 2D data across modalities.

Let the input to $\mathbf{R}_{inn}$ be the multi-channel two-dimensional spatial feature, $F \in \mathbb{R}^{c \times \frac{h}{d} \times \frac{w}{d}}$, and $\mathbf{G}(F) \in \mathbb{R}^{c \times \frac{h}{d} \times \frac{w}{d}}$ be the output of the Gaussian adaptive attention block, **G**. $\mathbf{G}(F)$ can be considered as a stack of $\frac{hw}{d^2}$ numbers of c-dimensional features, $[gf_{x^{(c)}}]_{\frac{hw}{d^2}}$ (visually illustrated in the lower block of Figure 2). Let $gf_{x^{(c)}} \in \mathbb{R}^c$ be the local feature at 2D coordinate point $x^{(c)}$ of $\mathbf{G}(F)$, where c is the number of channels for the spatial feature. $gf_{x^{(c)}}$ is then transformed into $\mathbf{D}_{inn}(\{gf_{x^{(c)}}, \mathbf{P}(x,y,h)\}) + gf_{x^{(c)}}$, where $\mathbf{D}_{inn}$ is a 2-layer MLP network, h is the altitude for the corresponding modality. Similar transformation is done for all those $\frac{hw}{d^2}$ numbers of gf's. So, the feature transformation through $\mathbf{R}_{inn}$ can be described as a two-step transformation:

1. Feature transformation through Gaussian Adaptive Attention, $F \rightarrow \mathbf{G}(F)$.
2. Transformation of $\mathbf{G}(F)$ through residual neural network:

$$\mathbf{G}(F) = [gf_{x^{(c)}}]_{\frac{hw}{d^2}} \rightarrow \mathbf{R}_{inn}(F) = [\mathbf{D}_{inn}(\{gf_{x^{(c)}}, \mathbf{P}(x,y,h)\}) + gf_{x^{(c)}}]_{\frac{hw}{d^2}} \tag{5.3}$$

5.2. Implicit super-resolution decoder

Implicit super-resolution decoder, $\mathbf{D}_{cr} : \mathcal{M}_k^H \times \mathbf{X}_c \rightarrow \mathcal{M}_k^C$, takes the transformed discrete high-resolution representation $\mathbf{X}_k^H$ and any 2D coordinate point $x^{(c)} \in \mathbf{X}_c$ and predicts the output at coordinate point $x^{(c)}$, $\mathbf{x}_k^C\left(x^{(c)}\right)$. We use a modified version of local implicit image function (LIIF) based decoder, which is a coordinate based decoding approach (Chen et al., 2021). The original LIIF decoder primarily consists of two components: (i) EDSR-based feature encoder (Lim et al., 2017), and a (ii) MLP network. In our modified LIIF-KAN decoder, we replace the MLP network with a KAN (Kolmogorov-Arnold Network)-based decoder (Liu et al., 2025).

Decoder (continuous reconstructor): EDSR-based (Lim et al., 2017) feature encoder, **FE**, encodes $\mathcal{M}_{out}^H$ into the encoded feature space $\mathcal{M}_{out}^F$,$\mathbf{FE} : \mathcal{M}_{out}^H \rightarrow \mathcal{M}_{out}^F$. The KAN-based decoder, $\mathbf{D}_{kan}$, then predicts the output at coordinate point $x^{(c)}$ following the similar methodology followed in LIIF (Chen et al., 2021). The overall continuous reconstruction can be described as: $\mathbf{x}_{out}^C\left(x^{(c)}\right) = (\mathbf{D}_{kan} \circ \mathbf{FE})\left(\mathbf{x}_{out}^H, x^{(c)}\right)$.

5.3. Cross-altitude prediction

Let, $\mathbf{x}_{in}^H$ be a discrete high-dimensional datapoint at altitude h_{in}. For simultaneous dimension reduction and continuous cross-altitude reconstruction through super-resolution, we need to infer discrete low-resolution representation $\mathbf{x}_{in}^L$, and $\mathbf{x}_{out}^C\left(x^{(c)}\right)$, where $x^{(c)}$ can be any 2D coordinate point in $\mathbf{X}_c$ at altitude h_{out}. Then,

$$\mathbf{x}_{in}^L = \mathbf{E}_\phi\left(\mathbf{x}_{in}^H, h_{in}\right) \tag{5.4}$$

$$\mathbf{x}_{out}^C\left(x^{(c)}\right) = \mathbf{D}_{kan}\left(\mathbf{FE}\left(\mathbf{D}_{tr}\left(\mathbf{x}_{in}^L, h_{out}\right)\right), x^{(c)}\right) \tag{5.5}$$

$h_{in} \neq h_{out}$ refers to the cross-altitude prediction scenario. At a super-resolution scale s, we evaluate $\mathbf{D}_{KAN}$ on s^2hw number of different $x^{(c)} \in \mathbf{X}_c$, resulting in a super-resolved output with dimension $\mathbb{R}^{sh \times sw}$.

6. Experiments

We first introduce the wind dataset that we used to evaluate our proposed model. Then we continue to elaborate the optimization procedure and obtained results.

https://doi.org/10.1017/eds.2026.10043 Published online by Cambridge University Press

6.1. Wind data

National Renewable Energy Laboratory's Wind Integration National Database (WIND) Toolkit provides high spatial and temporal resolution wind power, wind power forecasting, and meteorological data for over 126,000 locations across the continental United States during a 7-year span (Draxl et al., 2015). The simulated forecasts were developed using the Weather Research and Forecasting Model, which operates on a 2-kilometer (km) by 2-kilometer (km) grid with a 10-meter(m) resolution from the ground to 200 m above ground with several temporal resolutions available at 1-hour, 4-hour, 6-hour, and day-ahead forecast horizons. The spatial resolution of the WIND Toolkit is 2 km × 1 hr in spatio-temporal resolution. As a result, the wind dataset size is 1602(latitude) × 2976(longitude) × 61368(number of instances), or almost 1.2 TB per wind component (wind data at different heights). We randomly cropped data to reduce the resolution to 1500(latitude) × 2000(longitude) for each time instance. Wind velocity components at specific direction were determined using the wind speed and direction at a specific height. For example, if wind speed at height h is V at an angle $\theta°$ with northern direction, the northern and eastern components are $u = V\cos\theta°, v = V\sin\theta°$ accordingly. In this paper, we only report the results with the northern projection of wind speed data.

6.2. Experimental setup

6.2.1. Dataset

To evaluate the continuous reconstruction capability of our proposed method, we created a dataset for multi-modal super-resolution tasks using simulated wind data discussed in Section 6.1.

Dataset creation. From the total 61368 data instances at different timepoints, we randomly sampled 4000 data points from various timestamps at heights of 10 m, 60 m, 160 m and 200 m. We used 3000 data points for training, 200 for validation and 800 for testing.

Data sample processing. We used bicubic interpolation to generate a pair of discrete and continuous high-resolution samples for each instance. For example, if the discrete input dimension is 120 × 160 and the super-resolution scale is 2.5 ×, the continuous output dimension is 300 × 400. Data sample creation for training and validation/test has been conducted differently for better learning of the model.

- **Train**: During training, the continuous high-resolution sample is created by randomly cropping from the actual 1500 × 2000 resolution data sample and then generating the discrete high-resolution sample via bicubic interpolation. The discrete high-resolution dimension was set to 120 × 160, with super-resolution scales ranging from 1 × to 4 ×.
- **Validation/Test**: At testing, we split the original 1500 × 2000 resolution data into 4 data samples of resolution 750 × 1000 by splitting them evenly into four regions. As a result, the test dataset consists of 3200 data samples, each with a dimension of 750 × 1000. To avoid randomness during testing, both continuous and discrete high-resolution samples were generated by bicubic interpolation from this data sample with dimension 750 × 1000. For example, if the task is to super-resolve the discrete super-resolution sample of dimension 120 × 160 to a super-resolution scale 2.5 × (the continuous high-resolution sample dimension is 300 × 400), both the discrete and continuous high-resolution samples are generated through bicubic interpolation from the original cropped sample of dimension 750 × 1000.

6.2.2. Training details

Adam was adopted as the optimizer (Kingma and Ba, 2015), to train the model for 600 epochs. The learning rate during the optimization was set following the cyclical learning rate technique with a minimum learning rate of 10^{-5} and a maximum learning rate of 10^{-4} (Smith, 2017). At super-resolution scale s, the super-resolution ground truth $\mathbf{x}^S_{out}$ has a dimension of $sh \times sw$. Coordinate-based implicit neural networks, e.g., LIIF (Y Chen et al., 2021), predict outputs pixel-by-pixel instead of predicting the whole super-resolved image at a single forward pass. We randomly selected 2048 coordinate points, at each

https://doi.org/10.1017/eds.2026.10043 Published online by Cambridge University Press

optimization step, among the s^2hw coordinate points of the target super-resolved continuous representation, and optimized the parameters based on the predictions $\widehat{\mathbf{x}}^C_{out}\left(x^{(c)}\right)$ and the ground truth $\mathbf{x}^C_{out}\left(x^{(c)}\right)$ on those coordinate points to extradite the optimization process. We used the L_1 loss function, $l_1\left(\widehat{\mathbf{x}}^C_{out}, \mathbf{x}^C_{out}, x^{(c)}\right) = |\mathbf{x}^C_{out}\left(x^{(c)}\right) - \widehat{\mathbf{x}}^C_{out}\left(x^{(c)}\right)|$, to optimize the model. It took about 12 hours to train the model on a workstation with a single NVIDIA A100 GPU.

6.2.3. Evaluation metrics

We employed two metrics to evaluate the continuous reconstruction through super resolution. Peak signal-to-noise ratio (PSNR) is the ratio of a signal's maximum possible value (power) to the power of distorting noise that affects the quality of its representation. Structural similarity index (SSIM) is a perceptual metric that evaluates the degradation of image quality, which compares the spatial structures between the target image and reproduced image. For evaluation of the compression performance, we employed the Compression Ratio (CR) metric, which measures the ratio between the required memory for storing the compressed versus the uncompressed data. These evaluation metrics have been discussed in detail in the Supplementary Materials.

6.3. Results

6.3.1. Task

We evaluated the model's continuous reconstruction capability by performing super-resolution at different scales on a test dataset containing 800 data points. As we split each data point into 4 regions, the total number of effective test data samples is 3200.

6.3.2. Setup

High-resolution input dimension was set to 120×160, with a dimension reduction factor $d = 8$, yielding a low-resolution representation of 15×20. We focus on cross-altitude predictions where the input modality height h_{in} is closer to the ground (10 m, 60 m) and the output modality height h_{out} is significantly higher (160 m, 200 m). Evaluation results for the opposite cross-altitude prediction scenarios have been discussed in the Supplementary Materials. Due to the lack of existing methodologies for simultaneous dimension reduction and reconstruction through super-resolution, we use the following three methods as baselines:

- **MAINN (Multi-Altitude Implicit Neural Network):** Multi-altitude simultaneous reduction and continuous reconstruction through a multi-modal implicit neural network (Qayyum et al., 2024). We trained 4 different models with combinations of data from (10 m, 160 m), (60 m, 160 m), (10 m, 200 m), and (60 m, 200 m) heights.
- **GINO:** Bicubic downscaling followed by Geometry-Informed Neural Operator (GINO) (Li et al., 2024).
- **WindLaw:** We used the wind power law $v_1/v_2 = (h_1/h_2)^{\alpha}$ with our proposed One Stone Three Bird (OSTB) model. We eliminated the *3D Decoder INN* and trained four different models for four altitudes, and then used the wind power law for cross-altitude prediction.

6.3.3. Result summary

As shown in Figure 3, our model consistently outperforms all the baselines across all cross-altitude scenarios and super-resolution scales for PSNR and SSIM. Our model obtained low LPIPS for the super resolution scale $2\times$, but obtained higher LPIPS for higher super-resolution scales. The GINO model obtained superior results in terms of LPIPS. Results for super-resolution scales $s \in [2,4]$ and $s \in [5,6]$ respectively illustrate the results for in-distribution scales and out-of-distribution scales.

https://doi.org/10.1017/eds.2026.10043 Published online by Cambridge University Press

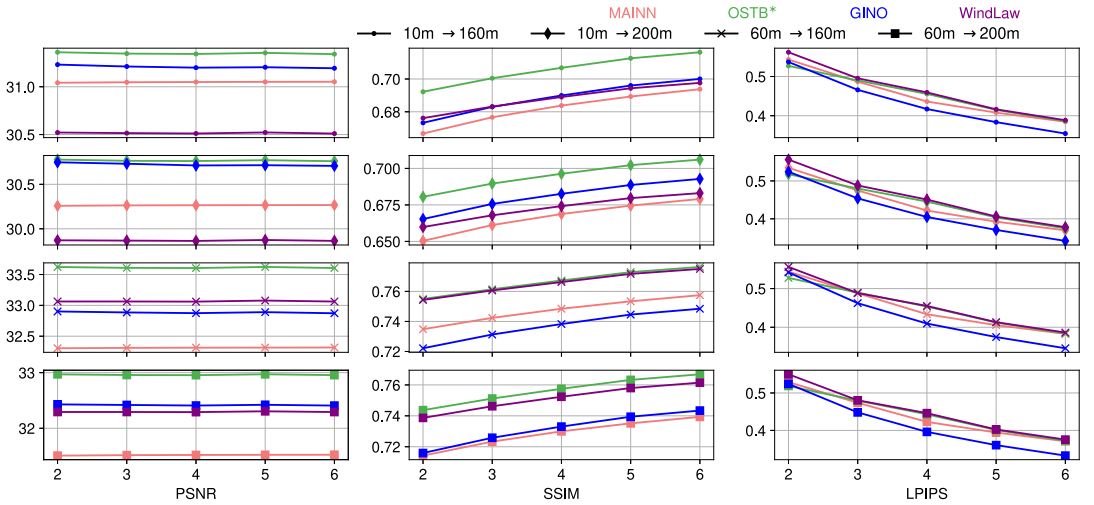


Figure 3. *Comparative Results with Baselines. x−axis shows different super resolution scales for all three subplots, and y−axis shows the metric values (PSNR, SSIM and LPIPS metric values are shown in the three columns from left to right column accordingly). The 4 rows show the evaluation results in 4 different cross-altitude prediction scenarios:* $10\ m \rightarrow 160\ m$, $10\ m \rightarrow 200\ m$, $60\ m \rightarrow 160\ m$, *and* $60\ m \rightarrow 200\ m$ *from top to bottom row accordingly. For the WindLaw baseline, the displayed results are observed with* $\alpha = 0.16$.

6.4. Additional Experiment 1: data compression

6.4.1. Task

We evaluate the data compression performance of our model and compare it against existing compression methods, focusing on cross-altitude prediction and reconstruction accuracy.

6.4.2. Setup

We employed the Prediction by Partial Matching (PPM) data compression algorithm with μ-law encoding at various quantization levels (Q) for data compression and reconstruction, as described in Moffat (1990). Additionally, we tested bicubic interpolation for compressing and decompressing the data. For cross-altitude predictions, we used the wind power law $v_1/v_2 = (h_1/h_2)^{\alpha}$ to transform the reconstructed data from one height to another for both PPM and bicubic methods. To assess performance, we measured the average compression ratio, PSNR, and SSIM over the test set. Since traditional compression methods lack super-resolution capabilities, we limit super-resolution comparison to the scale $s = 1$ for our model.

6.4.3. Result summary

Figure 4 illustrates the cross-altitude prediction methodology of these methods, while Table 1 presents a detailed comparison. For cross-altitude predictions where the input modality height is 10 m, our model outperforms the baseline methods across three metrics—compression ratio, PSNR, and SSIM. In cases where the input modality height is 60 m, while our model does not achieve the best PSNR and SSIM, it significantly surpasses others in terms of compression ratio. For bicubic models with a reduction factor $d = 4$, the low-resolution dimension is 30×40. With a higher reduction factor ($d = 8$), bicubic models achieve a compression ratio similar to ours but at the cost of lower PSNR and SSIM. Likewise, PPM models attain higher compression ratios at lower Q values, but at the expense of PSNR and SSIM. Only our model consistently delivers a high compression ratio alongside either the best or near-best PSNR and SSIM. Regarding the LPIPS metric, our model is outperformed by other models across all cross-altitude prediction scenarios, however their compression ratio is significantly higher than our model.

https://doi.org/10.1017/eds.2026.10043 Published online by Cambridge University Press

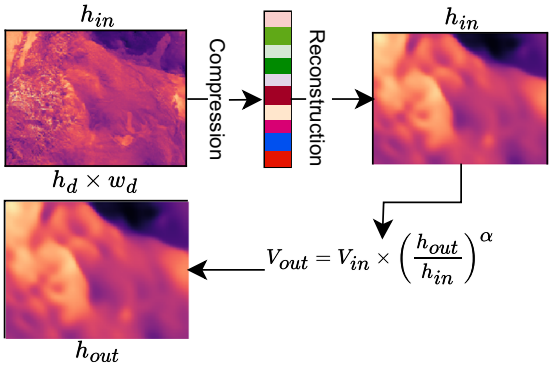


Figure 4. *Data compression followed by cross-altitude prediction using wind power law.*

6.5. Additional experiment 2: modality transfer

6.5.1. Task

We evaluate the performance of our model without the 3D ETINN segment by designing and training separate models for different altitude levels, then comparing them in cross-altitude prediction tasks.

6.5.2. Setup

We removed the 3D ETINN segment and developed four distinct models, each trained on data from a specific height, following consistent optimization procedures. For cross-altitude predictions, we used the wind power law (Touma, 1977) to transform the reconstructed data from one altitude to another. For instance, for an input height $h_{in} = 10$ m and output height $h_{out} = 200$ m, we applied the model trained at 10 m to predict super-resolved wind data and then used the wind power law to transform it to 200 m. Figure 5 summarizes this cross-altitude prediction methodology without the 3D ETINN. The downsampling architecture for dimension reduction followed the design of the invertible UNet (Etmann et al., 2020), as depicted in Figure 5. We tested different values of α for cross-altitude predictions at various super-resolution scales.

6.5.3. Result summary

Figure 6 presents the average PSNR, SSIM, and LPIPS across the test set at different super-resolution scales. In all cross-altitude prediction scenarios, the model incorporating the 3D ETINN consistently outperforms the models without it, demonstrating the critical role of the 3D ETINN in enhancing performance.

6.6. Additional experiment 3: ablation studies

6.6.1. Task 1

Ablation of decoder. We evaluate the performance of the proposed decoder by comparing it against other implicit neural network-based super-resolution models.

Setup. To benchmark the performance of our model, we replaced the LIIF-KAN decoder segment with the LIIF (Y Chen et al., 2021), ITNSR (Yang et al., 2021), HiNOTE (Luo et al., 2024), SRNO (Wei and Zhang, 2023), DIINN (QH Nguyen and Beksi, 2023), and LTE (Lee et al., 2022) models. We also tried LTE-KAN, replacing MLP with KAN in the LTE model, similar to the modification done in LIIF-KAN. All models were trained using the same optimization procedures. Figure 7 illustrates the average PSNR and SSIM across the test set at varying super-resolution scales, with the x-axis of each plot showing the respective scales.

https://doi.org/10.1017/eds.2026.10043 Published online by Cambridge University Press

Table 1. *Comparative compression performance at different cross-altitude prediction scenarios. Best performing methods are boldfaced according to respective metrics.*

	$h_{in} = 10$ m $\rightarrow h_{out} = 160$ m				$h_{in} = 10$ m $\rightarrow h_{out} = 200$ m			
Model	PSNR(↑)	SSIM(↑)	LPIPS(↓)	CR(↑)	PSNR(↑)	SSIM(↑)	LPIPS(↓)	CR(↑)
$\text{PPM}_{Q=4,\alpha=0.16}$	16.1586	0.2220	0.4652	97.0277	16.1174	0.2179	0.4753	97.0277
$\text{PPM}_{Q=4,\alpha=0.24}$	16.1619	0.2218	0.4695	97.0277	16.1139	0.2169	0.4807	97.0277
$\text{PPM}_{Q=4,\alpha=0.32}$	16.1120	0.2160	0.4754	97.0277	16.0408	0.2094	0.4878	97.0277
$\text{PPM}_{Q=8,\alpha=0.16}$	24.6997	0.4343	0.2864	95.6397	24.2223	0.4127	0.3034	95.6397
$\text{PPM}_{Q=8,\alpha=0.24}$	22.9060	0.4043	0.3188	95.6397	22.3011	0.3813	0.3400	95.6397
$\text{PPM}_{Q=8,\alpha=0.32}$	20.8745	0.3696	0.3597	95.6397	20.1082	0.3442	0.3864	95.6397
$\text{PPM}_{Q=12,\alpha=0.16}$	27.4357	0.5559	0.2270	94.2337	26.7911	0.5228	0.2446	94.2337
$\text{PPM}_{Q=12,\alpha=0.24}$	25.7281	0.5135	0.2576	94.2337	24.9666	0.4764	0.2792	94.2337
$\text{PPM}_{Q=12,\alpha=0.32}$	23.5273	0.4572	0.2949	94.2337	22.7020	0.4168	0.3193	94.2337
$\text{PPM}_{Q=16,\alpha=0.16}$	28.9872	0.6221	**0.1893**	93.3278	28.2311	0.5839	**0.2070**	93.3278
$\text{PPM}_{Q=16,\alpha=0.24}$	27.3643	0.5773	0.2168	93.3278	26.5088	0.5334	0.2383	93.3278
$\text{PPM}_{Q=16,\alpha=0.32}$	25.1355	0.5125	0.2502	93.3278	24.1513	0.4629	0.2764	93.3278
$\text{Bicubic}_{d=4,\alpha=0.16}$	30.3240	0.6559	0.3387	93.7500	29.6082	0.6306	0.3380	93.7500
$\text{Bicubic}_{d=4,\alpha=0.24}$	29.1067	0.6318	0.3468	93.7500	28.2963	0.6007	0.3477	93.7500
$\text{Bicubic}_{d=4,\alpha=0.32}$	26.9890	0.5805	0.3587	93.7500	26.0394	0.5417	0.3618	93.7500
$\text{Bicubic}_{d=8,\alpha=0.16}$	29.5809	0.6024	0.6427	**98.4375**	28.9719	0.5850	0.6411	**98.4375**
$\text{Bicubic}_{d=8,\alpha=0.24}$	28.4870	0.5840	0.6745	**98.4375**	27.7818	0.5626	0.6748	**98.4375**
$\text{Bicubic}_{d=8,\alpha=0.32}$	26.5472	0.5438	0.7036	**98.4375**	25.6907	0.5162	0.7048	**98.4375**
OSTB*	**31.3332**	**0.6801**	0.4174	**98.4375**	**30.7483**	**0.6657**	0.4099	**98.4375**
	$h_{in} = 60$ m $\rightarrow h_{out} = 160$ m				$h_{in} = 60$ m $\rightarrow h_{out} = 200$ m			
Model	PSNR(↑)	SSIM(↑)	LPIPS(↓)	CR(↑)	PSNR(↑)	SSIM(↑)	LPIPS(↓)	CR(↑)
$\text{PPM}_{Q=4,\alpha=0.16}$	14.1466	0.2907	0.4921	97.3825	14.1382	0.2881	0.4992	97.3825
$\text{PPM}_{Q=4,\alpha=0.24}$	14.1615	0.2963	0.4920	97.3825	14.1566	0.2944	0.4994	97.3825
$\text{PPM}_{Q=4,\alpha=0.32}$	14.1761	0.3014	0.4921	97.3825	14.1741	0.2998	0.4998	97.3825
$\text{PPM}_{Q=8,\alpha=0.16}$	25.3451	0.5294	0.2627	96.2921	24.9154	0.5081	0.2758	96.2921
$\text{PPM}_{Q=8,\alpha=0.24}$	24.8199	0.5217	0.2703	96.2921	24.2698	0.4988	0.2857	96.2921
$\text{PPM}_{Q=8,\alpha=0.32}$	24.2433	0.5132	0.2785	96.2921	23.5485	0.4884	0.2965	96.2921
$\text{PPM}_{Q=12,\alpha=0.16}$	28.3028	0.6492	0.2038	94.9586	27.6685	0.6137	0.2208	94.9586
$\text{PPM}_{Q=12,\alpha=0.24}$	27.6473	0.6360	0.2133	94.9586	26.8731	0.5970	0.2332	94.9586
$\text{PPM}_{Q=12,\alpha=0.32}$	26.9218	0.6203	0.2238	94.9586	25.9739	0.5769	0.2472	94.9586
$\text{PPM}_{Q=16,\alpha=0.16}$	30.7661	0.7342	**0.1583**	94.1765	29.9271	0.6925	**0.1761**	94.1765
$\text{PPM}_{Q=16,\alpha=0.24}$	30.1380	0.7210	0.1668	94.1765	29.1738	0.6750	0.1871	94.1765
$\text{PPM}_{Q=16,\alpha=0.32}$	29.3750	0.7039	0.1761	94.1765	28.2249	0.6522	0.1994	94.1765
$\text{Bicubic}_{d=4,\alpha=0.16}$	33.4129	**0.7627**	0.3196	93.7500	32.4829	**0.7356**	0.3183	93.7500
$\text{Bicubic}_{d=4,\alpha=0.24}$	33.1206	0.7572	0.3214	93.7500	32.0726	0.7270	0.3207	93.7500
$\text{Bicubic}_{d=4,\alpha=0.32}$	32.4632	0.7470	0.3236	93.7500	31.2240	0.7115	0.3240	93.7500
$\text{Bicubic}_{d=8,\alpha=0.16}$	32.0933	0.6916	0.6241	**98.4375**	31.3809	0.6743	0.6220	**98.4375**
$\text{Bicubic}_{d=8,\alpha=0.24}$	31.8357	0.6870	0.6352	**98.4375**	31.0246	0.6673	0.6355	**98.4375**
$\text{Bicubic}_{d=8,\alpha=0.32}$	31.2836	0.6785	0.6462	**98.4375**	30.2944	0.6549	0.6488	**98.4375**
OSTB*	**33.5972**	0.7500	0.4138	**98.4375**	**32.9439**	0.7354	0.4062	**98.4375**

https://doi.org/10.1017/eds.2026.10043 Published online by Cambridge University Press

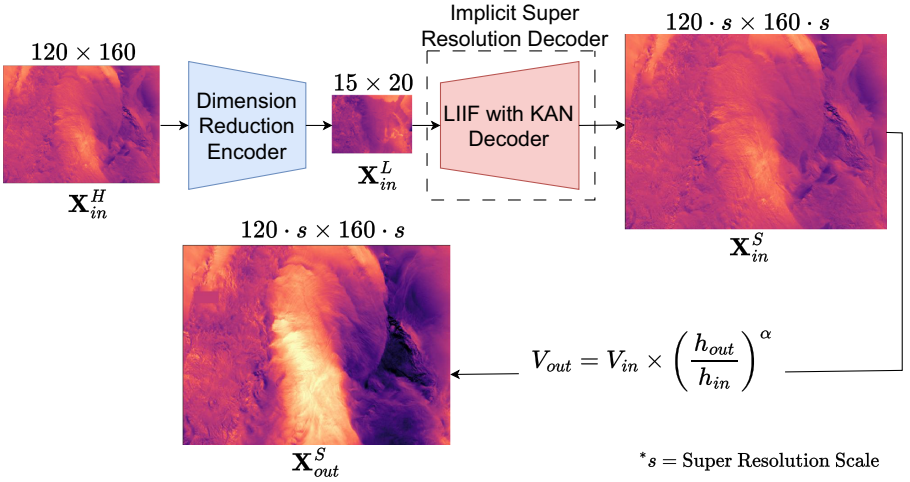


Figure 5. *Methodology for cross-altitude prediction without modality transfer INN.*

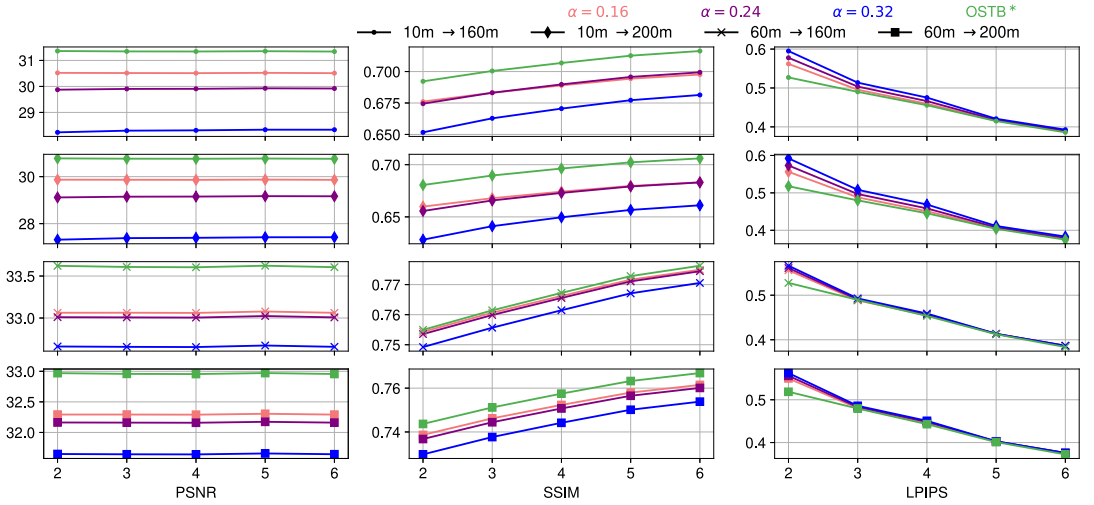


Figure 6. *Results for Modality Transfer Experiment. $x-$axis shows different super-resolution scales for all three subplots, and $y-$axis shows the metric values (PSNR, SSIM and LPIPS metric values are shown in the three columns from left to right column accordingly). The 4 rows show the evaluation results in 4 different cross-altitude prediction scenarios:* $10\ m \rightarrow 160\ m$, $10\ m \rightarrow 200\ m$, $60\ m \rightarrow 160\ m$ *and* $60\ m \rightarrow 200\ m$ *from top to bottom row accordingly.*

Result summary. Our modified LIIF-based decoder, LIIF-KAN, consistently outperforms all other implicit super-resolution models, including the original LIIF decoder. The results show a significant performance improvement in both PSNR and SSIM, demonstrating that LIIF-KAN offers superior super-resolution capabilities compared to existing models. However, the SRNO decoder performs best among all the decoders in terms of the LPIPS metric for most of the super-resolution scales.

6.6.2. Task 2

Ablation of attention mechanisms. We assess the efficacy of the Gaussian adaptive attention in the Residual Implicit Neural blocks by comparing it with other attention mechanisms.

https://doi.org/10.1017/eds.2026.10043 Published online by Cambridge University Press

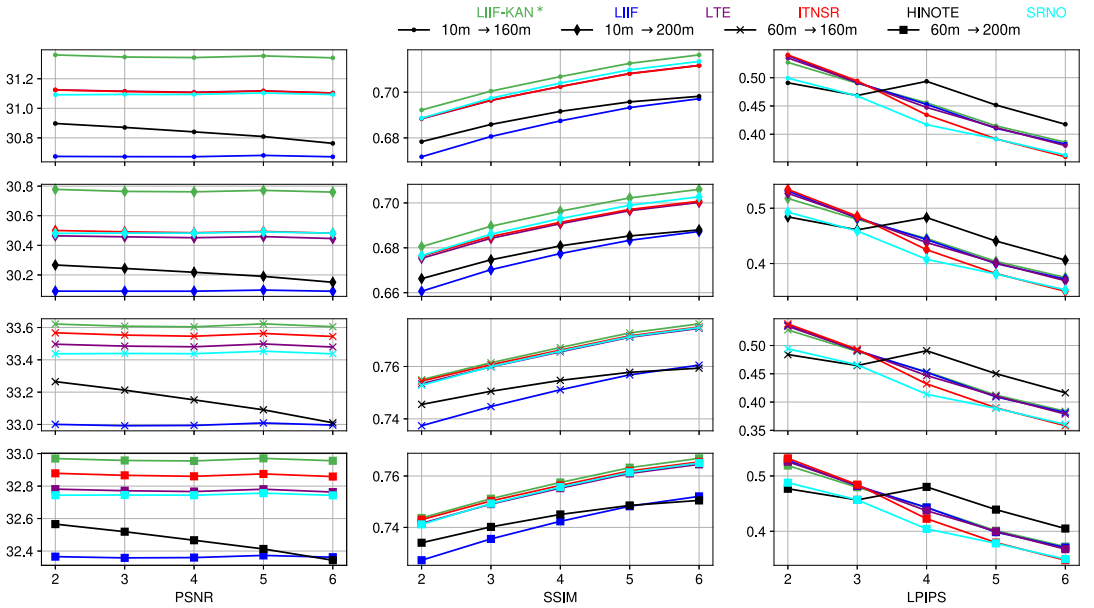


Figure 7. *Ablation of decoders. x−axis shows different super-resolution scales for all three subplots, and y−axis shows the metric values (PSNR, SSIM, and LPIPS metric values are shown in the three columns from left to right column accordingly). The 4 rows show the evaluation results in 4 different cross-altitude prediction scenarios:* $10\ m \rightarrow 160\ m$, $10\ m \rightarrow 200\ m$, $60\ m \rightarrow 160\ m$ *and* $60\ m \rightarrow 200\ m$ *from top to bottom row accordingly.*

Setup. We replaced the Gaussian adaptive attention mechanism in the Residual Implicit Neural blocks with the standard Query, Key, Value (QKV) attention mechanism (Vaswani et al., 2017), and also tested the performance without any attention mechanism. All models were trained following identical procedures. Figure 8 presents the results, showing the performance of each variant.

Result summary. The results demonstrate that our proposed method, incorporating the Gaussian adaptive attention mechanism, is outperformed by both the QKV-based attention and the version without attention in terms of PSNR and SSIM. However, it achieves lower LPIPS than its counterparts. This observation leaves us to balance among the three metrics. We have observed that our proposed model outperforms other models in all previous experiments, but it often failed to obtain low LPIPS. However, we observed that with QKV-based attention and without the attention method obtained even higher PSNR and LPIPS, but at the cost of even more suboptimal LPIPS. Also, the Gaussian adaptive attention mechanism is parameter-efficient compared to the QKV-based attention mechanism.

7. Conclusion

We have developed an innovative deep learning approach for simultaneous continuous super-resolution, data dimensionality reduction, and multi-altitude learning for climatological data. We designed a three-dimensional implicit neural network specifically for learning continuous, rather than discrete, representations of multi-altitude velocity fields used for wind farm power modeling across the continental United States. Unlike traditional multi-modal deep learning models, which handle only a limited number of modalities due to scalability issues, our three-dimensional implicit neural network is scalable to a large number of modalities. We employed two very recently proposed techniques: the Gaussian adaptive attention mechanism and the Kolmogorov–Arnold network to design our model and modify existing

https://doi.org/10.1017/eds.2026.10043 Published online by Cambridge University Press

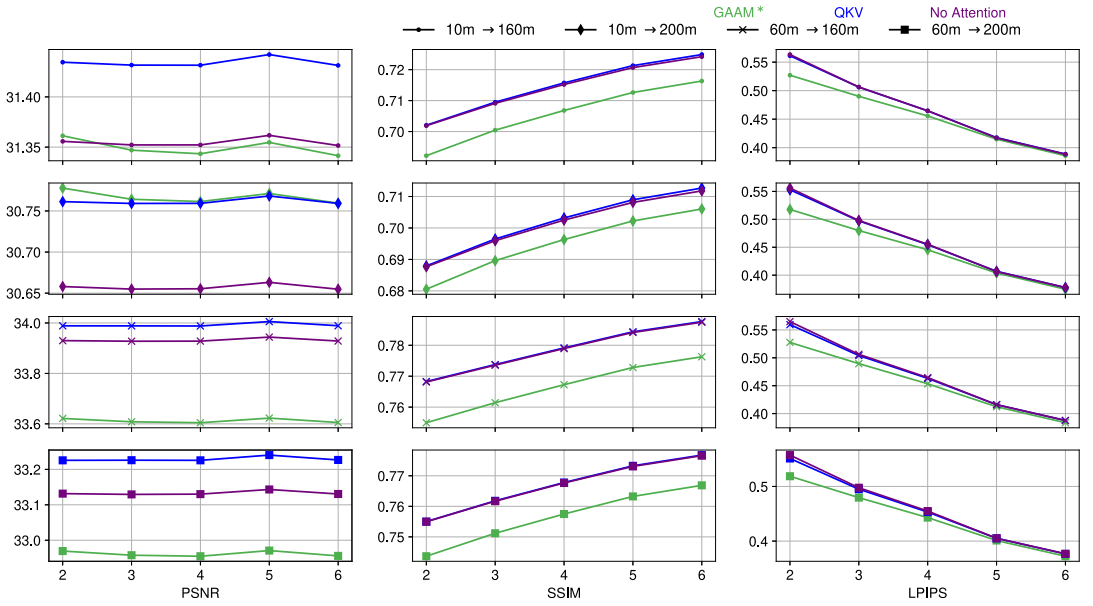


Figure 8. *Ablation of attention mechanism. x−axis shows different super resolution scales for all three subplots, and y−axis shows the metric values (PSNR, SSIM, and LPIPS metric values are shown in the three columns from left to right column accordingly). The 4 rows show the evaluation results in four different cross-altitude prediction scenarios:* 10 $m \rightarrow$ 160 m, 10 $m \rightarrow$ 200 m, 60 $m \rightarrow$ 160 m *and* 60 $m \rightarrow$ 200 m *from top to bottom row accordingly.*

models. Experimental results have shown promising potential in improving wind energy assessment for electricity generation, efficient large data storage via dimensionality reduction, and extrapolation to inaccessible spatial areas. Both Gaussian adaptive attention and Kolmogorov–Arnold networks have the potential to enhance interpretability in machine learning models, making this an area for future research.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/eds.2026.10043.

Acknowledgements. The authors acknowledge the use of (Google Gemini, 2025) for assistance in language editing and improving the clarity of the manuscript's presentation. All conceptual development, technical content, analysis, and conclusions are entirely the work of the authors.

Author contribution. Conceptualization: A.B.A.Q; X.L; X.Q; B.Y. Data curation: A.B.A.Q. Formal analysis: A.B.A.Q; X.L; N.M.U; X.Q; B.Y. Funding acquisition: X.Q; B.Y. Investigation: X.L; X.Q; B.Y. Methodology: A.B.A.Q; X.L; X.Q; B.Y. Project administration: X.Q; B.Y. Resources: X.Q; B.Y. Software: A.B.A.Q; X.Q. Supervision: X.L; X.Q; B.Y. Validation: A.B.A.Q; X.L; X.Q; B.Y. Visualization: A.B.A.Q; X.L. Writing—original draft: A.B.A.Q; X.L; X.Q; B.Y. Writing - review & editing: A.B.A.Q; X.L; N.M.U; X.Q; B.Y.

Competing interests. The authors declare no competing interests.

Data availability statement. Data and code implementations can be found at https://github.com/alifbinabdulqayyum/OSTB.

Funding statement. This work was supported in part by the Department of Energy (DOE) award DE-SC0012704.

References

Boor CD (1978) *A Practical Guide to Splines*, Vol. *27*. New York: Springer New York. https://doi.org/10.1007/978-1-4612-6333-3.

Boussif O, **Bengio Y**, **Benabbou L and Assouline D** (2022) MAgnet: Mesh agnostic neural PDE solver. *Advances in Neural Information Processing Systems 35*, 31972–31985.

Boussioux L, **Zeng C**, **Guénais T and Bertsimas D** (2022) Hurricane forecasting: A novel multimodal machine learning framework. *Weather and Forecasting 37*(6), 817–831. https://doi.org/10.1175/WAF-D-21-0091.1.

Cao ZJ and Gao G (2022) Multi-omics single-cell data integration and regulatory inference with graph-linked embedding. *Nature Biotechnology 40*(10), 1458–1466. https://doi.org/10.1038/s41587-022-01284-4.

Chau SL, **Bouabid S and Sejdinovic D** (2021) Deconditional downscaling with Gaussian processes. *Advances in Neural Information Processing Systems 34*, 17813–17825.

Chen X, **Feng K**, **Liu N**, **Ni B**, **Lu Y**, **Tong Z and Liu Z** (2022) RainNet: A large-scale imagery dataset and benchmark for spatial precipitation downscaling. *Advances in Neural Information Processing Systems 35*, 9797–9812.

Chen Y, **Liu S and Wang X** (2021) Learning continuous image representation with local implicit image function. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8628–8638.

Diaconu CA, **Saha S**, **Günnemann S and Zhu XX** (2022) Understanding the role of weather data for earth surface forecasting using a convlstm-based model. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1362–1371.

Draxl C, **Clifton A**, **Hodge BM and McCaa J** (2015) The WIND integration National Dataset (WIND) toolkit. *Applied Energy 151*, 355–366. https://doi.org/10.1016/j.apenergy.2015.03.121.

Jiang C, **Esmaeilzadeh S**, **Azizzadenesheli K**, **Kashinath K**, **Mustafa M**, **Tchelepi HA**, **Marcus P**, **Prabhat M and Anandkumar A** (2020) Meshfreeflownet: A physics-constrained deep continuous space-time super-resolution framework. In *SC20: International Conference for High Performance Computing, Networking, Storage and Analysis*. Los Alamitos, CA, USA: IEEE, pp. 1–15. doi: 10.1109/SC41405.2020.00013.

Etmann C, **Ke R and Schönlieb CB** (2020) iUNets: Learnable invertible up-and downsampling for large-scale inverse problems. In *2020 IEEE 30th International Workshop on Machine Learning for Signal Processing (MLSP)*. Piscataway, NJ, USA: IEEE, pp. 1–6.

Gao Z, **Shi X**, **Wang H**, **Zhu Y**, **Wang YB**, **Li M and Yeung DY** (2022) Earthformer: Exploring space-time transformers for earth system forecasting. *Advances in Neural Information Processing Systems 35*, 25390–25403.

Google Gemini (2025) *Large Language Model (Version 2.5).* Generative AI tool used for rephrasing on [Insert Date(s) of Use]. Available at https://gemini.google.com/.

Groenke B, **Madaus L and Monteleoni C** (2020) ClimAlign: Unsupervised statistical downscaling of climate variables via normalizing flows. In *Proceedings of the 10th International Conference on Climate Informatics*, New York, NY, USA: Association for Computing Machinery. 60–66. doi: 10.1145/3429309.3429318. Available at https://doi.org/10.1145/3429309.3429318.

Harder P, **Hernandez-Garcia A**, **Ramesh V**, **Yang Q**, **Sattegeri P**, **Szwarcman D**, **Watson C and Rolnick D** (2023) Hard-constrained deep learning for climate downscaling. *Journal of Machine Learning Research 24*(365), 1–40.

Hornik K, **Stinchcombe M and White H** (1989) Multilayer feedforward networks are universal approximators. *Neural Networks 2*(5), 359–366. https://doi.org/10.1016/0893-6080(89)90020-8.

Huang L and Hoefler T (2023) Compressing multidimensional weather and climate data into neural networks. In *The Eleventh International Conference on Learning Representations*. Available at https://openreview.net/forum?id=Y5SEe3dfniJ.

Ioannides G, **Chadha A and Elkins A** (2024) *Density Adaptive Attention is All You Need: Robust Parameter-Efficient Fine-Tuning Across Multiple Modalities.* arXiv: 2401.11143 [cs.LG]. Available at https://arxiv.org/abs/2401.11143.

Irrgang C, **Boers N**, **Sonnewald M**, **Barnes EA**, **Kadow C**, **Staneva J and Saynisch-Wagner J** (2021) Towards neural earth system modelling by integrating artificial intelligence in earth system science. *Nature Machine Intelligence 3*(8), 667–674. https://doi.org/10.1038/s42256-021-00374-3.

Jing L, **Vahdani E**, **Tan J and Tian Y** (2021) Cross-modal center loss for 3D cross-modal retrieval. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Los Alamitos, CA, USA: IEEE. 3142–3151. doi: 10.1109/CVPR46437.2021.00316.

Kaack LH, **Donti PL**, **Strubell E**, **Kamiya G**, **Creutzig F and Rolnick D** (2022) Aligning artificial intelligence with climate change mitigation. *Nature Climate Change 12*(6), 518–527. https://doi.org/10.1038/s41558-022-01377-7.

Kashinath K, **Mustafa M**, **Albert A**, **Wu J**, **Jiang C**, **Esmaeilzadeh S**, **Azizzadenesheli K**, **Wang R**, **Chattopadhyay A**, **Singh A**, **Manepalli A**, **Chirila D**, **Yu R**, **Walters R**, **White B**, **Xiao H**, **Tchelepi H**, **Marcus P**, **Anandkumar A**, **Hassanzadeh P and Prabhat** (2021) Physics-informed machine learning: Case studies for weather and climate modelling. *Philosophical Transactions of the Royal Society A 379*(2194), 20200093. https://doi.org/10.1098/rsta.2020.0093.

Keller AA, **Garner KL**, **Rao N**, **Knipping E and Thomas J** (2022) Downscaling approaches of climate change projections for watershed modeling: Review of theoretical and practical considerations. *PLOS Water 1*(9), e0000046. https://doi.org/10.1371/journal.pwat.0000046.

Kingma DP and Ba J (2015) Adam: A method for stochastic optimization. In Bengio Y and LeCun Y (eds), *3rd International Conference on Learning Representations*, ICLR 2015, San Diego, CA, USA, May 7–9, 2015, Conference Track Proceedings.

Klöwer M, **Razinger M**, **Dominguez JJ**, **Düben PD and Palmer TN** (2021) Compressing atmospheric data into its real information content. *Nature Computational Science 1*(11), 713–724. https://doi.org/10.1038/s43588-021-00156-2.

Kovachki NB, **Li Z**, **Liu B**, **Azizzadenesheli K**, **Bhattacharya K**, **Stuart AM and Anandkumar A** (2021) *Neural Operator: Learning Maps Between Function Spaces. CoRR* abs/2108.08481.

Lee J and Jin K (2022) Local texture estimator for implicit representation function. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1929–1938.

https://doi.org/10.1017/eds.2026.10043 Published online by Cambridge University Press

Li Z, **Kovachki N**, **Choy C**, **Li B**, **Kossaifi J**, **Otta S**, **Nabian MA**, **Stadler M**, **Hundt C**, **Azizzadenesheli K and Anandkumar A** (2024) Geometry-informed neural operator for large-scale 3d pdes. *Advances in Neural Information Processing Systems 36*.

Li Z, **Kovachki NB**, **Azizzadenesheli K**, **B l**, **Bhattacharya K**, **Stuart A and Anandkumar A** (2021) Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations* Available at https://openreview.net/forum?id=c8P9NQVtmnO.

Lim B, **Son S**, **Kim H**, **Nah S and Lee KM** (2017) Enhanced deep residual networks for single image super-resolution. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.

Liu Y, **Ganguly AR and Dy J** (2020) Climate downscaling using YNet: A deep convolutional network with skip connections and fusion. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 3145–3153.

Liu Z, **Wang Y**, **Vaidya S**, **Ruehle F**, **Halverson J**, **Soljacic M**, **Hou TY and Tegmark M** (2025) KAN: Kolmogorov–Arnold networks. In *The Thirteenth International Conference on Learning Representations* Available at https://openreview.net/forum?id=Ozo7qJ5vZi.

Luo X, **Qian X**, **Urban N and Yoon BJ** (2023) Reinstating continuous climate patterns from small and discretized data. In *1st Workshop on the Synergy of Scientific and Machine Learning Modeling@ ICML2023*.

Luo X, **Qian X and Yoon BJ** (2024) Hierarchical neural operator transformer with learnable frequency-aware loss prior for arbitrary-scale super-resolution. In Salakhutdinov R, Kolter Z, Heller K, Weller A, Oliver N, Scarlett J and Berkenkamp F (eds), *Forty-First International Conference on Machine Learning*, Vol *235*. Proceedings of Machine Learning Research. Honolulu, Hawaii, USA: PMLR, 33466–33485. Available at https://openreview.net/forum?id=LhAuVPWq6q.

Luo X, **Xu W**, **Nadiga B**, **Ren Y and Yoo S** (2024) Continuous field reconstruction from sparse observations with implicit neural networks. In *The Twelfth International Conference on Learning Representations* Available at https://openreview.net/forum?id=kuTZMZdCPZ.

Moffat A (1990) Implementing the PPM data compression scheme. *IEEE Transactions on Communications 38*(11), 1917–1921. http://doi.org/10.1109/26.61469.

Ngiam J, **Khosla A**, **Kim M**, **Nam J**, **Lee H and Ng AY** (2011) Multimodal deep learning. In *International Conference on Machine Learning (ICML)*. ICML'11. Madison, WI, USA: Omnipress, 689–696. isbn: 9781450306195.

Nguyen QH and Beksi WJ (2023) Single image super-resolution via a dual interactive implicit neural network. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, Los Alamitos, CA, USA: IEEE, 4936–4945. doi: 10.1109/WACV56688.2023.00491.

Nguyen T, **Brandstetter J**, **Kapoor A**, **Gupta JK and Grover A** (2023) ClimaX: A foundation model for weather and climate. In Krause A, Brunskill E, Cho K, Engelhardt B, Sabato S and Scarlett J (eds), *Proceedings of the 40th International Conference on Machine Learning*, *Proceedings of Machine Learning Research.* Honolulu, Hawaii, USA: PMLR, Vol. *202*, pp. 25904–25938. Available at https://proceedings.mlr.press/v202/nguyen23a.html

Qayyum ABA, **Luo X**, **Urban NM**, **Qian X and Yoon BJ** (2024) Implicit neural representations for simultaneous reduction and continuous reconstruction of multi-altitude climate data. In *2024 IEEE 34th International Workshop on Machine Learning for Signal Processing (MLSP)*, Piscataway, NJ, USA: IEEE, pp. 1–6. https://doi.org/10.1109/MLSP58920.2024.10734742.

Ramesh A, **Dhariwal P**, **Nichol A**, **Chu C and Chen M** (2022a) *Hierarchical Text-Conditional Image Generation with CLIP Latents.* arXiv: 2204.06125 [cs.CV].

Ramesh SS, **Sessa PG**, **Krause A and Bogunovic I** (2022b) Movement penalized Bayesian optimization with application to wind energy systems. *Advances in Neural Information Processing Systems 35*, 27036–27048.

Reichstein M, **Camps-Valls G**, **Stevens B**, **Jung M**, **Denzler J**, **Carvalhais N and f P** (2019) Deep learning and process understanding for data-driven earth system science. *Nature 566*(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1.

Requena-Mesa C, **Benson V**, **Reichstein M**, **Runge J and Denzler J** (2021) EarthNet2021: A large-scale dataset and challenge for earth surface forecasting as a guided video prediction task. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Los Alamitos, CA, USA: IEEE pp. 1132–1142. doi: 10.1109/CVPRW53098.2021.00124

Rolnick D, **Donti PL**, **Kaack LH**, **Kochanski K**, **Lacoste A**, **Sankaran K**, **Ross AS**, **Milojevic-Dupont N**, **Jaques N**, **Waldman-Brown A and Luccioni AS** (2022) Tackling climate change with machine learning. *ACM Computing Surveys (CSUR) 55*(2), 1–96.

Ruan L, **Ma Y**, **Yang H**, **He H**, **Liu B**, **Fu J**, **Yuan NJ and JinQand Guo B** (2023) MM-diffusion: Learning multi-modal diffusion models for joint audio and video generation. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Los Alamitos, CA, USA: IEEE Computer Society, 10219–10228. doi: 10.1109/CVPR52729.2023.00985. Available at https://doi.ieeecomputersociety.org/10.1109/CVPR52729.2023.00985.

Schwarz JR, **Tack J**, **Teh YW**, **Lee J and Shin J** (2023) Modality-agnostic variational compression of implicit neural representations. In Krause A, Brunskill E, Cho K, Engelhardt B, Sabato S and Scarlett J (eds), *Proceedings of the 40th International Conference on Machine Learning*, *Proceedings of Machine Learning Research*, Vol. *202*. Honolulu, Hawaii, USA: PMLR, pp. 30342–30364. Available at https://proceedings.mlr.press/v202/schwarz23a.html.

Smith LN (2017) Cyclical learning rates for training neural networks. In *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*. Los Alamitos, CA, USA: IEEE, 464–472. doi: 10.1109/WACV.2017.58.

Stengel K, **Glaws A**, **Hettinger D and King RN** (2020) Adversarial super-resolution of climatological wind and solar data. *Proceedings of the National Academy of Sciences 117*(29), 16805–16815. https://doi.org/10.1073/pnas.1918964117.

https://doi.org/10.1017/eds.2026.10043 Published online by Cambridge University Press

Summaira J, **Li X**, **Shoib AM**, **Li S and Abdul J** (2021) Recent advances and trends in multimodal deep learning: A review. arXiv preprint arXiv:2105.11087.

Tancik M, **Srinivasan P**, **Mildenhall B**, **Fridovich-Keil S**, **Raghavan N**, **Singhal U**, **Ramamoorthi R**, **Barron J and Ng R** (2020) Fourier features let networks learn high frequency functions in low dimensional domains. *Advances in Neural Information Processing Systems 33*, 7537–7547.

Touma JS (1977) Dependence of the wind profile power law on stability for various locations. *Journal of the Air Pollution Control Association 27*(9), 863–866. https://doi.org/10.1080/00022470.1977.10470503.

Tu X, **Cao ZJ**, **Xia CR**, **Mostafavi S and Gao G** (2022) Cross-linked unified embedding for cross-modality representation learning. In Oh AH, Agarwal A, Belgrave D and Cho K (eds), *Advances in Neural Information Processing Systems*. vol *35*. Red Hook, NY, USA: Curran Associates, Inc., 15942–15955. Available at https://proceedings.neurips.cc/paper_files/paper/2022/file/662b1774ba8845fc1fa3d1fc0177ceeb-Paper-Conference.pdf.

Vandal T, **Kodra E**, **Ganguly S**, **Michaelis A**, **Nemani R and Ganguly AR** (2017) Deepsd: Generating high resolution climate change projections through single image super-resolution. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '17. New York, NY, USA: Association for Computing Machinery, 1663–1672. isbn: 9781450348874. doi: 10.1145/3097983.3098004. Available at https://doi.org/10.1145/3097983.3098004.

Vaswani A, **Shazeer N**, **Parmar N**, **Uszkoreit J**, **Jones L**, **Gomez AN**, **Kaiser L and Polosukhin I** (2017) Attention is all you need. In Guyon, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S and Garnett R (eds), *Advances in Neural Information Processing Systems*, Vol. *30*. Red Hook, NY, USA: Curran Associates, Inc Available at https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

Wei M and Zhang X (2023) Super-resolution neural operator. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Los Alamitos, CA, USA: IEEE Computer Society, pp. 18247–18256. doi: 10.1109/CVPR52729.2023.01750. Available at https://doi.ieeecomputersociety.org/10.1109/CVPR52729.2023.01750.

Xie Y, **Takikawa T**, **Saito S**, **Litany O**, **Yan S**, **Khan N**, **Tombari F**, **Tompkin J**, **Sitzmann V and Sridhar S** (2022) Neural fields in visual computing and beyond. *Computer Graphics Forum 41*(2), 641–676. https://doi.org/10.1111/cgf.14505.

Xu P, **Zhu X and Clifton DA** (2023) Multimodal learning with transformers: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence 45*(10), 12113–12132. https://doi.org/10.1109/TPAMI.2023.3275156.

Yang J, **Shen S**, **Yue H and Li K** (2021) Implicit transformer network for screen content image continuous super-resolution. In Beygelzimer A, Dauphin Y, Liang P and Vaughan JW (eds), *Advances in Neural Information Processing Systems*, Vol *34*. Red Hook, NY, USA: Curran Associates, Inc., 13304–13315. Available at https://proceedings.neurips.cc/paper_files/paper/2021/file/6e7d5d259be7bf56ed79029c4e621f44-Paper.pdf.

Cite this article: Qayyum ABA, Luo X, Urban NM, Qian X and Yoon B.-J. (2026). *One stone three birds*: Three-dimensional implicit neural network for compression and continuous representation of multi-altitude climate data. *Environmental Data Science*, 5: e16. doi:10.1017/eds.2026.10043

https://doi.org/10.1017/eds.2026.10043 Published online by Cambridge University Press